MATHEMATICS OF COMPUTATION
Volume 85, Number 297, January 2016, Pages 379–399
http://dx.doi.org/10.1090/mcom/2978
Article electronically published on May 12, 2015

THE SELF-POWER MAP AND COLLECTING ALL RESIDUE CLASSES

CATALINA V. ANGHEL

Abstract. The self-power map is the function from the set of natural numbers to itself which sends the number n to n^n. Motivated by applications to cryptography, we consider the image of this map modulo a prime p. We study the question of how large x must be so that $n^n \equiv a \bmod p$ has a solution $1 \leq n \leq x$, for every residue class a modulo p. While $n^n \bmod p$ is not uniformly distributed, it does appear to behave in certain ways as a random function. We give a heuristic argument to show that the expected x is approximately $p^2 \log \phi(p-1)/\phi(p-1)$, using the coupon collector problem as a model. We prove the bound $x < p^{2-\alpha}$ for sufficiently large p and a fixed constant $\alpha > 0$ independent of p, using a counting argument and exponential sum bounds.

1. Introduction

The self-power map is the function from the set of natural numbers to itself which sends the number n to n^n. We consider the image of this map modulo a prime p.

The self-power map is related to the discrete logarithm problem in $(\mathbb{Z}/p\mathbb{Z})^\times$. This problem is as follows: Given g and h in $(\mathbb{Z}/p\mathbb{Z})^\times$, find an integer a such that $h = g^a \bmod p$. While raising a group element to a power is computationally easy, the inverse operation of finding the discrete logarithm is believed to be difficult in general. Thus, the security of many cryptography applications depends on the difficulty of the discrete logarithm problem.

One such application is the El-Gamal signature algorithm. However, two variants of this algorithm also depend on the difficulty of finding n such that $n^n \equiv a \bmod p$ for a given a, and $1 \leq n \leq p-1$. [16, Notes 11.70, 11.71].

Another application of the discrete logarithm is the pseudo-random bit generator of Blum and Micali [5], which iterates the function $f : x \to g^x \mod p$. The residue classes in $(\mathbb{Z}/p\mathbb{Z})^\times$ are represented as integers between 1 and $p-1$. The pseudo-random bit sequence depends on the sequence $\{x, f(x), f(f(x)), \ldots\}$. Having the iteration fall into a fixed point or a short cycle would be problematic. Cycles of length two of the discrete logarithm, of the form

$$g^h = a \bmod p \quad \text{and} \quad g^a = h \bmod p, \tag{1.1}$$

Received by the editor April 21, 2013 and, in revised form, April 22, 2014.
2010 *Mathematics Subject Classification.* Primary 11A07, 11K45, 11T23; Secondary 94A60.

are related to collisions[1] of the self-power map

$$h^h \equiv a^a \mod p. \tag{1.2}$$

For instance, given values g, h, and a satisfying (1.1), then a and h also satisfy (1.2). However, in certain cases solutions to (1.2) produce solutions to (1.1), as described in [13].

The self-power map has not been studied as extensively as the discrete logarithm problem. Past work from Crocker [7] and Balog et al. [4] focused on bounding the number of solutions of $n^n \equiv a \bmod p$ when $1 \leq n \leq p-1$. In this article, we consider the case when n can be greater than p and we show that the sequence $n^n \bmod p$ has random-like properties. The behaviour of the self-power map is an interesting mathematical problem in its own right.

Before starting, we collect here some of the definitions and notations used. We use common notations for the standard arithmetic functions. The Euler totient function $\phi(n)$ denotes the number of positive integers from 1 to n which are relatively prime to n. The function $\omega(n)$ denotes the number of distinct prime divisors of n. The Möbius function $\mu(n)$ takes the value 0 if n is divisible by a square prime factor, otherwise takes the value $(-1)^{\omega(n)}$.

For the sake of brevity, we introduce the notations

$$\begin{aligned} e(z) &= \exp(2\pi i z), \\ e_k(z) &= \exp((2\pi i z)/k), \end{aligned}$$

given a positive integer k. We also often denote the greatest common divisor of integers a and b as (a, b).

Finally, we borrow the notation from Flajolet et al. [10], [11], and use $[z^n]f(z)$ to denote the nth Taylor coefficient of $f(z)$.

2. Preliminaries

The self-power map is not injective when n is restricted to be between 1 and $p-1$, since $1^1 \equiv (p-1)^{p-1} \bmod p$. On the other hand, we can find a solution for $n^n \equiv a \bmod p$ for any a if we let $1 \leq n \leq p(p-1)$ by taking

$$n = a + p(p-a) = 1 + (p-1)(p+1-a). \tag{2.1}$$

Our question is as follows:

Question 2.1. How large must x be so that $n^n \equiv a \bmod p$ has a solution $n \leq x$, for any given residue class a modulo p?

Since the self-power map is not injective when $n < p$, the inequality $p \leq x$ holds, and using equation (2.1), we have that $x \leq p(p-1)$. Thus, the trivial bounds for x are $p \leq x \leq p(p-1)$. We give a heuristic argument to estimate the expectation of x to be $p^2 \log \phi(p-1)/\phi(p-1)$ in Section 3, and we give the deterministic bound $x < p^{2-\alpha}$ for an $\alpha > 0$, in Section 4.

To obtain these results, it is useful to organize the values of $n^n \bmod p$ where $\gcd(n, p) = 1$ as in Table 1. Note that the self-power map is periodic modulo p with period $p(p-1)$.

[1]The term "collision" is used to describe hash functions in cryptography. A collision occurs when a (hash) function maps a larger domain to a smaller range and two arguments result in the same function value. We use the same term here, following [13], as the situation is similar.

TABLE 1. The values of n^n considered modulo p, where the base is reduced modulo p, and the exponent modulo $p-1$. Values of n which are multiples of p are omitted.

	Values of $n^n \bmod p$, where $n = kp + b$					
b	1	2	3	...	$(p-2)$	$(p-1)$
$k=0$	1^1	2^2	3^3	...	$(p-2)^{p-2}$	$(p-1)^{p-1}$
$k=1$	1^2	2^3	3^4	...	$(p-2)^{p-1}$	$(p-1)^1$
$k=2$	1^3	2^4	3^5	...	$(p-2)^1$	$(p-1)^2$
$\vdots$	$\vdots$					
$k=(p-2)$	1^{p-1}	2^1	3^2	...	$(p-2)^{p-3}$	$(p-1)^{p-2}$

The number of times a residue $a \in (\mathbb{Z}/p\mathbb{Z})^\times$ of order t appears in the table was given by Somer as $\sum_{k|\frac{p-1}{t}} \phi(kt)\cdot(p-1)/(kt)$ in [20, Theorem 1]. If $p-1$ is square-free, it can also be estimated as follows, obtained independently from Somer by R. Balasubramanian, [3].

Proposition 2.2. *Let a be an element of order t in $(\mathbb{Z}/p\mathbb{Z})^\times$. Define a multiplicative function on square-free numbers by*

$$g(q) = 2q - 1$$

where q is prime. Assuming $p-1$ is square-free,[2] *the number of times it will appear in Table 1 is*

$$\phi(t) g\left(\frac{p-1}{t}\right).$$

Proof. In Table 1, for $b \in (\mathbb{Z}/p\mathbb{Z})^\times$, each column is a permutation of

$$1, b, b^2, \ldots, b^{p-1}.$$

The residue class a will appear in any column where the order of b is a multiple of the order of a. Letting $\operatorname{ord} a = t$ and $\operatorname{ord} b = kt$, then a will occur $(p-1)/kt$ times in this column. Thus, the total number of times a will appear in Table 1 is

$$\sum_{k|\frac{p-1}{t}} \phi(kt)\cdot\frac{p-1}{kt} = \frac{\phi(t)}{t}(p-1)\sum_{k|\frac{p-1}{t}}\frac{\phi(k)}{k} = \phi(t)g\left(\frac{p-1}{t}\right). \qquad \square$$

Note that since there are $\phi(t)$ elements of order t in $(\mathbb{Z}/p\mathbb{Z})^\times$, the following equality holds for $p-1$ square-free:

$$\sum_{t|(p-1)} \phi(t)^2 g\left(\frac{p-1}{t}\right) = (p-1)^2.$$

[2]The case when $p-1$ is square-free gives a nice closed-form expression which is used in later calculations. The proof does not extend to non-square-free $p-1$ as ϕ is not completely multiplicative.

3. A heuristic estimate based on the coupon collector problem

The motivation for the probabilistic model is simply the ideal scenario. If elements were randomly distributed in Table 1, then the expected x in Question 2.1 would be the same as the expected number of random draws required to obtain all residue classes.

The latter question is the coupon collector problem under a non-uniform random distribution. If m different types of coupons are randomly distributed in cereal boxes, how many boxes should be bought in order to obtain at least one coupon of each type? In our case the cereal boxes correspond to the n's and the coupons to the residue classes $a \bmod p$.

The calculations are inspired by the classical coupon collector problem. In the classical case, each of the m coupons occurs with probability $1/m$ in any cereal box. If T is the number of cereal boxes required until all coupons are collected, the expectation for T is

$$E(T) = m\left(\log m + \gamma + o(1)\right)$$

where γ is the Euler-Mascheroni constant [11, p. 117, equation (34)]. For any $t \in \mathbb{R}$ the cumulative probability function of T satisfies

$$\lim_{m\to\infty} P\left(T \leq m(\log m + t)\right) = e^{-e^{-t}}.$$

The previous result was proven by Erdős and Rényi [8], but can also be proven using the saddle point method [11, p. 118, Note II.10].

In our case, the coupon collector model is modified to give an estimate of the number of elements of the sequence $n^n \bmod p$ needed to collect all the residue classes $a = 1, 2, \ldots, p-1$. The estimate is

$$E(T) = \frac{(p-1)^2}{\phi(p-1)}\left(\log\phi(p-1) + \gamma + o(1)\right).$$

Analogously to the classical case, the saddle-point method is used to show that, for $p-1$ square-free,

$$\lim_{p\to\infty} P\left(T \leq \frac{(p-1)^2}{\phi(p-1)}\left(\log\phi(p-1) + t\right)\right) = e^{-e^{-t}}.$$

We begin by deriving the expectation, and then calculating the cumulative distribution function.

3.1. **An estimate for the expectation.**

The estimate for the expectation of x is motivated by previous work which suggested that the primitive roots are the bottleneck for the collection of all residue classes [1]. Since higher order elements occur least often in Table 1, we assume that it is more difficult to obtain a solution n to the congruence $n^n \equiv a \bmod p$ for a's which are primitive roots.

In the following calculations, we set the primitive roots to be the coupons of value, or the 'useful coupons'. We modify the coupon collector problem by assuming that no residue classes other than the primitive roots need be collected. That is, all other residue classes are 'blank coupons'. This assumption simplifies the calculation for the expectation, and matches computational results.

Consider a set of m coupons under a general probability distribution $\{p_i\}_{i=1}^m$, which are drawn with replacement. Let m_B coupons, which appear with probabilities $p_1, p_2, \ldots, p_{m_B}$, be blank coupons. Let

$$p_1 + p_2 + \ldots + p_{m_B} = \beta.$$

In the coupon collector problem with blank coupons, our goal is to collect only the $m - m_B$ useful coupons, not all m coupons.

Proposition 3.1. *The expectation $E(T_j)$ of the time (or the number of draws) necessary to gather a collection of j useful coupons under a general probability distribution with m_B blank coupons is*

$$E(T_j) = \sum_{q=0}^{j-1} \int_0^\infty [u^q] \left(\prod_{i=m_B+1}^{m} \left(1 + u\left(e^{p_i t} - 1\right)\right) \right) e^{(\beta-1)t} dt, \tag{3.1}$$

where $[u^q]f(u)$ denotes the qth coefficient in the Taylor expansion of $f(u)$. For a full collection of the $m - m_B$ useful coupons the expectation is

$$E(T) = \int_0^\infty \left(1 - \prod_{i=m_B+1}^{m} (1 - e^{-p_i t}) \right) dt. \tag{3.2}$$

We apply Proposition 3.1 to a random sampling of the elements from Table 1, where we assume that only the primitive elements are 'useful' coupons. All smaller order elements are considered to be 'blank'.

Corollary 3.2. *Assume that the elements from Table* 1 *are randomly selected, with replacement. Then the expected time required to collect all primitive roots modulo p is*

$$E(T) = \frac{(p-1)^2}{\phi(p-1)} \left(\log \phi(p-1) + \gamma + o(1) \right).$$

Since the proof of Proposition 3.1 follows the proofs of Theorems 3.1 and 4.1 in [10] very closely, it is included in Appendix A. Corollary 3.2 follows from a calculus calculation, by letting $p_i = \phi(p-1)/(p-1)^2$ for all $1 \le i \le \phi(p-1)$ in Proposition 3.1 and using the substitutions $u = e^{-(\phi(p-1)/(p-1)^2)t}$ followed by $v = 1 - u$ in the integration.

3.2. **An estimate for the cumulative distribution function.** In this section, n and t will be different from the notation above. We hope their use will be clear from the context, which follows Flajolet and Sedgewick [11]. From [11, p. 192], the cumulative distribution function for the coupon collector problem relative to a general probability distribution $\{p_i\}_{i=1}^m$ is given by

$$P(T \le n) = n![z^n] \prod_{j=1}^{m} \left(e^{p_i z} - 1\right), \tag{3.3}$$

where again we use the notation $[z^n]f(z)$ for the nth coefficient in the Taylor expansion of $f(z)$.

We will find the asymptotic behaviour of $P(T \le n)$ in the case when coupons are randomly distributed, but according to the frequencies of the elements in Table 1,

provided $p-1$ is square-free. By Proposition 2.2, when $p-1$ is square-free, if a residue class has order d, the probability of selecting it at random from Table 1 is

$$\frac{\phi(d)g\left(\frac{p-1}{d}\right)}{(p-1)^2}.$$

There are $\phi(d)$ elements of order d, thus by equation (3.3) our goal is to estimate

$$n![z^n]\prod_{d|(p-1)}\left(e^{\frac{\phi(d)g\left(\frac{p-1}{d}\right)}{(p-1)^2}z}-1\right)^{\phi(d)},$$

so we let

$$G(z)=\prod_{d|(p-1)}\left(e^{\frac{\phi(d)g\left(\frac{p-1}{d}\right)}{(p-1)^2}z}-1\right)^{\phi(d)}. \tag{3.4}$$

By the Cauchy integral formula,

$$[z^n]G(z)=\frac{1}{2\pi i}\int_{\mathcal{C}}\frac{G(z)}{z^{n+1}}dz$$

where $\mathcal{C}$ is a contour that encircles the origin. For Hayman-admissible functions, the integral (and thus the Taylor coefficient) can be evaluated using the saddle point. Roughly speaking, we can choose a contour $\mathcal{C}$ such that the function is large only on a small part of the contour and small everywhere else, and the integral can be approximated by a Gaussian integral which captures the area under the peak.

The following definition of Hayman admissibility is from [12, p. 68], with notation from [11, Definition VIII.1, p. 565]. Letting $z=re^{i\theta}$, we have

$$\log G\left(re^{i\theta}\right)=\log G(r)+\sum_{\nu=1}^{\infty}\alpha_\nu(r)\frac{(i\theta)^\nu}{\nu!}$$

and define

$$h(z)=\log G(z)$$

and

$$\begin{aligned} a(r) &= \alpha_1(r)=rh'(r),\\ b(r) &= \alpha_2(r)=r^2h''(r)+rh'(r). \end{aligned}$$

Definition 3.3 (Hayman-admissibility)**.** Let $G(z)$ have a radius of convergence R with $0<R\le\infty$ and be always positive on some subinterval (R_0,R) of $(0,R)$. The function $G(z)$ is said to be H-admissible (or Hayman-admissible) if, with $a(r)$ and $b(r)$ defined as above, it satisfies the following three conditions:

(1) (Capture condition) $\lim_{r\to R}a(r)=\infty$ and $\lim_{r\to R}b(r)=\infty$.
(2) (Locality condition) For some function $\theta_0(r)$, defined over (R_0,R) and satisfying $0<\theta_0<\pi$, one has
$$G\left(re^{i\theta}\right)\sim G(r)e^{i\theta a(r)-\theta^2b(r)/2}$$
as $r\to R$, uniformly in $|\theta|\le\theta_0(r)$.
(3) (Decay condition) Uniformly in $\theta_0(r)\le|\theta|\le\pi$,
$$G\left(re^{i\theta}\right)=o\left(\frac{G(r)}{\sqrt{b(r)}}\right).$$

One of the convenient properties of Hayman-admissible functions is the closure under multiplication [12, Theorem VII]. The function

$$G(z) = \prod_{d|(p-1)} \left(e^{\frac{\phi(d) g\left(\frac{p-1}{d}\right)}{(p-1)^2} z} - 1 \right)^{\phi(d)}$$

can be shown to be Hayman-admissible by showing that each of its factors is Hayman-admissible. Thus we can apply the saddle point method from [11, Section VIII.5.1].

Theorem 3.4. *Let p be a prime such that $p-1$ is square-free. If the elements of Table 1 are sampled at random with replacement, and T is the time (or the number of draws) until obtaining a complete collection of residue classes modulo p, then*

$$\lim_{p\to\infty} P\left(T < \frac{(p-1)^2}{\phi(p-1)} \log \phi(p-1) + t \frac{(p-1)^2}{\phi(p-1)} \right) = e^{-e^{-t}}. \tag{3.5}$$

Proof. Given equation (3.3), we would like to estimate $n![z^n]G(z)$, where $G(z)$ is given by equation (3.4), using [11, Theorem VIII.4, p. 565].

Letting $h(r) = \log G(r)$, we have

$$h(r) = \sum_{d|(p-1)} \phi(d) \log \left(e^{\frac{\phi(d) g\left(\frac{p-1}{d}\right)}{(p-1)^2} r} - 1 \right)$$

and

$$h'(r) = \frac{G'(r)}{G(r)} = \sum_{d|(p-1)} \frac{\phi(d)^2 g\left(\frac{p-1}{d}\right)}{(p-1)^2} \left(e^{\frac{\phi(d) g\left(\frac{p-1}{d}\right)}{(p-1)^2} r} \right) \left(e^{\frac{\phi(d) g\left(\frac{p-1}{d}\right)}{(p-1)^2} r} - 1 \right)^{-1}.$$

We have that $\lim_{r\to\infty} h'(r) = 1$ since $\sum_{d|(p-1)} \phi(d)^2 g\left(\frac{p-1}{d}\right) = (p-1)^2$ from Proposition 2.2.

Since the radius of integration should be chosen to be a good approximation[3] to $\xi G'(\xi)/G(\xi) = n$, we may take

$$\xi = n.$$

Here we must use the estimate of the expectation of T from Corollary 3.2 in our expression for n,

$$n = \frac{(p-1)^2}{\phi(p-1)} \left(\log \phi(p-1) + t \right),$$

to find an approximation of the cumulative distribution function in terms of t (see the discussion at the beginning of Section 3).

From [11, Theorem VIII.4, p. 565],

$$n![z^n]G(z) = n! \left(\frac{G(\xi)}{\xi^n \sqrt{2\pi b(\xi)}} \right),$$

where $b(r) = r^2 h''(r) + r h'(r)$.

[3] In [11, Theorem VIII.4, p. 565] the radius of integration is taken to be the exact solution. In the general discussion of the saddle-point method, it is noted that the solution can be approximate: "On another register, it often proves convenient to adopt integration paths that come close enough to the saddle-point but need not pass exactly through it."[11, p. 554] In our case, it can be shown that $h'(n) = G'(n)/G(n) = 1 + O\left(1/\phi(p-1)\right)$ and $b(n) = n\left(1 + o(1)\right)$, so that the estimation of the Taylor coefficient is valid.

Using the Stirling approximation $n! \sim \sqrt{2\pi n}\,(n/e)^n$ and that $b(\xi) = \xi$ (and thus can be replaced by n), the factor $n!/\left(\xi^n\sqrt{2\pi b(\xi)}\right)$ tends to $1/e^n$. We have

$$\begin{aligned}\frac{1}{e^n}G(n) &= \frac{1}{e^n}\prod_{d|(p-1)}\left(e^{\phi(d)g\left(\frac{p-1}{d}\right)\frac{n}{(p-1)^2}}-1\right)^{\phi(d)}\\ &= e^{-n}\prod_{d|(p-1)} e^{\frac{n\phi(d)^2 g\left(\frac{p-1}{d}\right)}{(p-1)^2}}\left(1-e^{-\frac{n\phi(d)g\left(\frac{p-1}{d}\right)}{(p-1)^2}}\right)^{\phi(d)}\\ &= \prod_{d|(p-1)}\left(1-e^{-n\cdot\frac{\phi(d)g\left(\frac{p-1}{d}\right)}{(p-1)^2}}\right)^{\phi(d)}.\end{aligned}$$

Substituting $n = (p-1)^2\left(\log\phi(p-1)+t\right)/\phi(p-1)$ and taking out the term $d = p-1$ in front of the product, we obtain

$$\frac{1}{e^n}G(n) = \left(1-\frac{1}{\phi(p-1)}e^{-t}\right)^{\phi(p-1)}\prod_{d|\frac{(p-1)}{2}}\left(1-e^{-(\log\phi(p-1)+t)g\left(\frac{p-1}{d}\right)\phi\left(\frac{p-1}{d}\right)^{-1}}\right)^{\phi(d)}.$$

We have

$$\lim_{p\to\infty}\left(1-\frac{1}{\phi(p-1)}e^{-t}\right)^{\phi(p-1)} = e^{-e^{-t}},$$

as desired. We can show that the product over divisors of $(p-1)/2$ tends to 1 by showing that its logarithm tends to zero, using the Taylor expansion for $\log(1-x)$ when x is small and the bound $g\left(\frac{p-1}{d}\right)\phi\left(\frac{p-1}{d}\right)^{-1} \geq 2$ for $d \neq 1$. □

3.3. **Computational results.** We may rewrite equation (3.5) as

$$P\left(\left(T-\frac{(p-1)^2\log(p-1)}{\phi(p-1)}\right)\left(\frac{\phi(p-1)}{(p-1)^2}\right) < t\right) \to e^{-e^{-t}}.$$

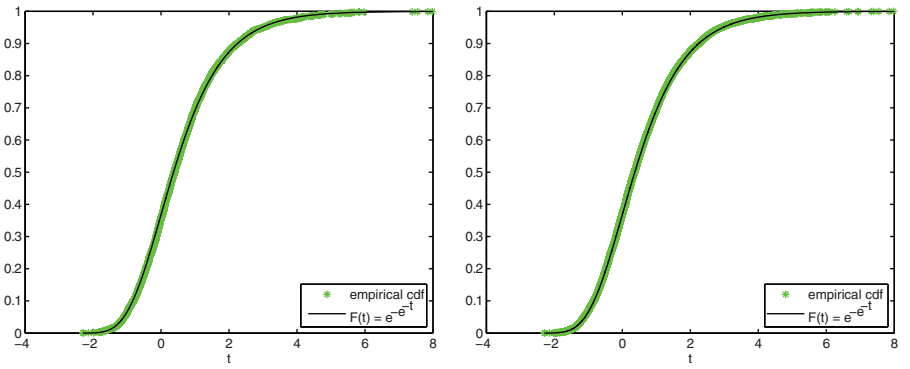
Since T should be a good approximation of x, let

$$X_p = \left(x-\frac{p(p-1)\log(p-1)}{\phi(p-1)}\right)\left(\frac{\phi(p-1)}{p(p-1)}\right), \tag{3.6}$$

where x depends on p in each case, and is the smallest natural number so that $n^n \equiv a \bmod p$ and $n \leq x$ has a solution for every residue class a modulo p. (The replacement of $(p-1)^2$ by $p(p-1)$ in the expression for X_p corrects for the omission of multiples of p in Table 1. However, this change is not important as both could be replaced by p^2.)

For primes p less than 80000, such that $p-1$ is square-free, the graph comparing the empirical and theoretical distributions are similar but not identical; the Kolmogorov-Smirnov test accepts the null hypothesis of the double exponential distribution $e^{-e^{-t}}$, at a significance value 0.05, with p-value of 0.1647. For all odd primes less than 80000, the test accepts the null hypothesis at a significant value 0.05, with p-value of 0.7928.

The number of primes tested was limited by the computational time. The calculation was performed using Matlab®, Student Version 7.12.0.635 (R2011a) on a personal laptop and took an average of 2.3 minutes per prime for the last 100 primes calculated.



(a) Square-free primes. (b) Odd primes.

FIGURE 1. Comparison between the cumulative distribution function derived from the coupon collector problem and the empirical distribution function for the variables X_p from equation (3.6), for primes less than 80000.

4. A deterministic bound

We return to the central question, restated below, from a different angle.

Question 2.1. How large must x be so that $n^n \equiv a \bmod p$ has a solution $n \leq x$, for any given residue class a modulo p?

In the previous section we conjectured that the value of x should be approximately $p^2 \log \phi(p-1)/\phi(p-1)$, and in this section we will give the deterministic bound of $p^{2-\alpha}$ for an $\alpha > 0$. Residue classes of higher order again play an important role in the proof.

Theorem 4.1. *For p a large prime, and any fixed residue class $a \bmod p$, there exists an n such that*

$$n^n \equiv a \bmod p$$

and $n < p^{2-\alpha}$, for some $\alpha > 0$ independent of p and of the residue class $a \bmod p$.

Any value $\alpha \leq 1 \times 10^{-3}$ is sufficient for large enough p.

The theorem is equivalent to the statement that every residue class $a \bmod p$ appears in the first $p^{1-\alpha}$ rows of Table 1. Thus, we can rephrase the question as:

Question 4.2. Given a, an element of order t in $(\mathbb{Z}/p\mathbb{Z})^\times$, is there an $n < p^{2-\alpha}$ such that $n^n = a \mod p$?

If this is true for all a of order t, and true for all orders t, then $x < p^{2-\alpha}$.

The question is interesting for residue classes of large order. Notice that an element $a \bmod p$ of (small) order less than $p^{1/2+\epsilon}$ will appear in the first $p^{1/2+\epsilon}$ rows of the table. Thus, it is sufficient to consider the case when $n^n \equiv g \bmod p$, for an element g of large multiplicative order. The change in notation from a to g is motivated by now considering the residue class as a group element of $(\mathbb{Z}/p\mathbb{Z})^\times$. In particular, the proof is easiest to follow in the case when g is a primitive root.

The proof for elements of large order follows the strategy of Theorem 8 in [17]. The main tools are the Erdős-Turán inequality and a bound on exponential sums. Both are given below.

Let B be a box

$$B = [\alpha_1, \beta_1) \times [\alpha_2, \beta_2) \times \cdots \times [\alpha_k, \beta_k) \subseteq [0,1)^k$$

of volume

$$|B| = \prod_{i=1}^{k} (\beta_i - \alpha_i).$$

Definition 4.3. The discrepancy of the sequence $\{\mathbf{x}_n\}_{n=1}^N$ in $[0,1)^k$ is defined as

$$\Delta = \sup_{B \subseteq [0,1)^k} \left| \frac{\mathcal{N}(B)}{N} - |B| \right|$$

where $\mathcal{N}(B)$ is the number of points of the sequence which fall in the box B, and the supremum runs over all boxes B inside the k-dimensional unit cube.

Theorem 4.4 (Erdős-Turán Inequality [4])**.** *Let $\{\mathbf{x}_n\}_{n=1}^N$ be points in the k-dimensional unit cube and M an arbitrary positive integer. Then*

$$\Delta \leq \left(\frac{3}{2}\right)^k \left(\frac{2}{M+1} + \sum_{0 < \|\mathbf{h}\| \leq M} \frac{1}{r(\mathbf{h})} \left| \frac{1}{N} \sum_{n=1}^{N} e(\mathbf{h} \cdot \mathbf{x}_n) \right| \right)$$

where $\|\mathbf{h}\| = \max_{1 \leq j \leq k} |h_j|$ and $r(\mathbf{h}) = \prod_{j=1}^{k} \max(1, |h_j|)$ for $\mathbf{h} = (h_1, \ldots, h_k) \in \mathbb{Z}^k$. Recall that $e(z) = \exp(2\pi i z)$.

The exponential sum bound required comes from Theorem 4 in [6]. Let g be a fixed element in $(\mathbb{Z}/p\mathbb{Z})^\times$ of order t and let

$$S_{m,n} = \sum_{(u,t)=1} e_p\left(m g^{u^{-1}}\right) e_t(nu).$$

Recall that we are using the notation $e_k(z) = \exp((2\pi i z)/k)$.

Theorem 4.5. *For any $\epsilon > 0$ there exists $\delta > 0$ such that for $t \geq p^\epsilon$, uniformly over $m \in (\mathbb{Z}/p\mathbb{Z})^\times$ and $n \in \mathbb{Z}/t\mathbb{Z}$, we have the bound*

$$S_{m,n} \ll t^{1-\delta}.$$

For our purposes, let g have order $t > p^{1/2+\epsilon}$ for any $\epsilon > 0$. Denote the representatives mod p of all the elements of order t as

$$1 < g_1 < g_2 < \ldots < g_N < p$$

where $N = \phi(t)$. Note that g is equal to one of the g_i's. Let $1 \leq a_i < t$ be the numbers relatively prime to t such that

$$g_i^{a_i} \equiv g \bmod p.$$

Then,

$$g_i \equiv g^{a_i^{-1}} \bmod p,$$

[4]The one-dimensional version of this inequality is due to Erdős and Turán [9]. The generalization to k dimensions is by Koksma [14] and Szűsz [21], and the notation used here is from [15] p. 112–116.

where the inverse of a_i is taken modulo t.

To show that $n^n \equiv g \bmod p$ has a solution in n with $n < p^{2-\alpha}$, it will be enough to show that the sequence $\{(a_i, g_i)\}_{i=1}^N$ is sufficiently scattered in the $t \times p$ rectangle, even though it is not uniformly distributed. As an illustration, Figure 2 shows the position of one primitive root for $p = 337$ in the $(p-1)\times p$ rectangle. First normalize to the unit square $[0,1)^2$ to obtain the sequence

$$\left\{\left(\frac{a_i}{t}, \frac{g_i}{p}\right)\right\}_{i=1}^N = \left\{\left(\frac{a_i}{t}, \frac{g^{a_i^{-1}}}{p}\right)\right\}_{i=1}^N.$$

We begin by showing that the discrepancy of this sequence is small.

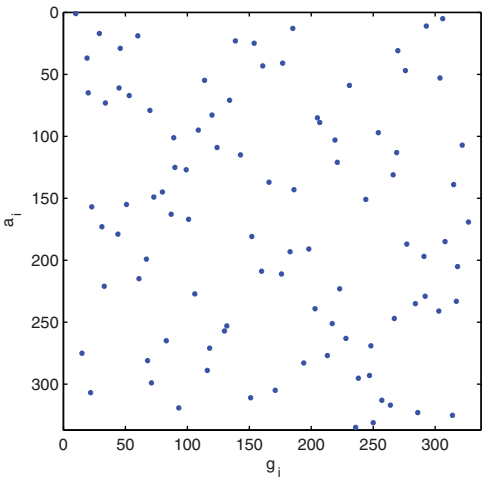


FIGURE 2. If g is primitive, the rectangle of interest has dimensions $(p-1) \times p$. The figure shows the positions of $\{(a_i, g_i)\}_{i=1}^N$ corresponding to the smallest primitive root $g = 10$ for the prime 337.

Proposition 4.6. *Let p be a prime. The discrepancy of the sequence*

$$\left\{\left(\frac{a_i}{t}, \frac{g_i}{p}\right)\right\}_{i=1}^N$$

is bounded by $O\left((\log^2 t)N^{-1}t^{1-\delta}\right)$, for some $\delta > 0$, where $N = \phi(t)$.

Proof. Consider the sum

$$\begin{aligned} S_{m,n} &= \sum_{i=1}^N e\left(m\frac{g^{a_i^{-1}}}{p} + n\frac{a_i}{t}\right) \\ &= \sum_{i=1}^N e_p\left(mg^{a_i^{-1}}\right) e_t(na_i) \\ &= \sum_{\substack{u=1 \\ (u,t)=1}}^t e_p\left(mg^{u^{-1}}\right) e_t(nu). \end{aligned}$$

We will restrict m, n to be integers less than t in magnitude. Theorem 4.5 gives the bound $S_{m,n} \ll t^{1-\delta}$ for the case $m \neq 0$ and $-t < m, n < t$.

If $n \neq 0$ and $m = 0$ the sum reduces to the Ramanujan sum:

$$S_{0,n} = \sum_{\substack{u=1 \\ (u,t)=1}}^{t} e_t(nu).$$

Since $S_{0,n} = \overline{S_{0,-n}}$, we need only consider the cases when m, n are non-negative. The Ramanujan sum can be evaluated using von Sterneck's formula, where $k = \gcd(n, t)$:

$$S_{0,n} = \mu\left(\frac{t}{k}\right) \frac{\phi(t)}{\phi\left(\frac{t}{k}\right)}. \tag{4.1}$$

Given that the Möbius function only takes values 0, 1, and -1, we have that $S_{0,n} = 0$ or

$$|S_{0,n}| = \frac{\phi(t)}{\phi\left(\frac{t}{k}\right)}.$$

Thus $|S_{0,n}|$ may be larger than $t^{1-\delta}$ in cases when k is large.

By Theorem 4.4 (Erdős-Turán inequality) for two dimensions, we have

$$\Delta \leq C\left(\frac{2}{M+1} + \frac{1}{N} \sum_{\substack{(m,n)\neq(0,0) \\ 0\leq m,n\leq M}} \left(\frac{1}{1+m}\right)\left(\frac{1}{1+n}\right)|S_{m,n}|\right).$$

Taking m and n non-negative above only changes the constant in Erdős-Turán inequality by essentially a multiple of 4.

We can split the sum above into two parts, depending on whether $|S_{m,n}| \leq t^{1-\delta}$ or $|S_{m,n}| > t^{1-\delta}$. In the latter case, $m = 0$. Let A be the set

$$A = \left\{1 \leq n < t \text{ such that } |S_{0,n}| > t^{1-\delta}\right\}.$$

Taking $M = t - 1$ in the bound for discrepancy, we have

$$\begin{aligned}\Delta &\leq C\left(\frac{2}{t} + \frac{1}{N} \sum_{\substack{(m,n)\neq(0,0), n\notin A \\ 0\leq m,n<t}} \left(\frac{1}{1+m}\right)\left(\frac{1}{1+n}\right) t^{1-\delta} + \frac{1}{N} \sum_{\substack{n\in A \\ 1\leq n<t}} \left(\frac{1}{1+n}\right)|S_{0,n}|\right) \\ &\leq C\left(\frac{2}{t} + (\log^2 t)\frac{t^{1-\delta}}{N} + \frac{1}{N} \sum_{\substack{n\in A \\ 1\leq n<t}} \left(\frac{1}{1+n}\right)|S_{0,n}|\right).\end{aligned}$$

It remains to show that

$$\frac{1}{N} \sum_{\substack{n\in A \\ 1\leq n<t}} \left(\frac{1}{1+n}\right)|S_{0,n}|$$

is small. This follows from Lemma 4.7 and Lemma 4.8.

Lemma 4.7. *If $n \in A$, then $n > t^{1-2\delta}$.*

Proof. Using von Sterneck's formula (4.1), recall that $S_{0,n} = 0$ or

$$|S_{0,n}| = \frac{\phi(t)}{\phi\left(\frac{t}{k}\right)},$$

where $k = \gcd(n, t)$. For $n \in A$, it follows that

$$\frac{\phi(t)}{\phi\left(\frac{t}{k}\right)} > t^{1-\delta}.$$

Thus $\phi\left(\frac{t}{k}\right) < t^{\delta}$. Since for all $x > 2$,

$$\phi(x) > \frac{x}{e^{\gamma}\log\log x + \frac{3}{\log\log x}}$$

and $\log\log(t/k) > 1$ for $(t/k) > e^e$, we have

$$t^{\delta} > \phi\left(\frac{t}{k}\right) > \frac{\frac{t}{k}}{e^{\gamma}\log\log\frac{t}{k} + 3}$$

provided $(t/k) > e^e$. Thus

$$t^{\delta} > \frac{\frac{t}{k}}{e^{\gamma}\log\log t + 3},$$
$$k > \frac{t^{1-\delta}}{e^{\gamma}\log\log t + 3}.$$

We remark that in the case $\log\log(t/k) < 1$, we have $k > (t/e^e)$. Thus in either case, $n > k > t^{1-2\delta}$ for t large enough. □

Lemma 4.8. *For p large enough, $|A| \le t^{2\delta}$.*

Proof. We have from [2] the identity

$$\sum_{n=1}^{t} |S_{0,n}| = \phi(t)2^{\omega(t)}$$

where $\omega(t)$ is the number of distinct prime divisors of t. From the inequality [18, p. 394]

$$2^{\omega(t)} < t^{\frac{(1+\epsilon)\log 2}{\log\log t}}$$

we obtain $2^{\omega(t)} < t^{\delta}$ if t is large enough. Thus, there are at most

$$\frac{2^{\omega(t)}N}{t^{1-\delta}} \le 2^{\omega(t)}t^{\delta} \le t^{2\delta}$$

elements of A, where $N = \phi(t)$. □

We can now return to the proof of Proposition 4.6. Using the trivial estimate $|S_{0,n}| < N$, we have

$$\frac{1}{N}\sum_{\substack{n\in A\\ 1\le n<t}} \left(\frac{1}{1+n}\right)|S_{0,n}| \le t^{2\delta}t^{-1+2\delta} = \frac{1}{t^{1-4\delta}}.$$

Substituting this into the expression for the discrepancy, we have

$$\Delta \le C\left(\frac{2}{t} + (\log^2 t)\frac{t^{1-\delta}}{N} + \frac{1}{t^{1-4\delta}}\right) < C'(\log^2 t)\frac{t^{1-\delta}}{N}.$$ □

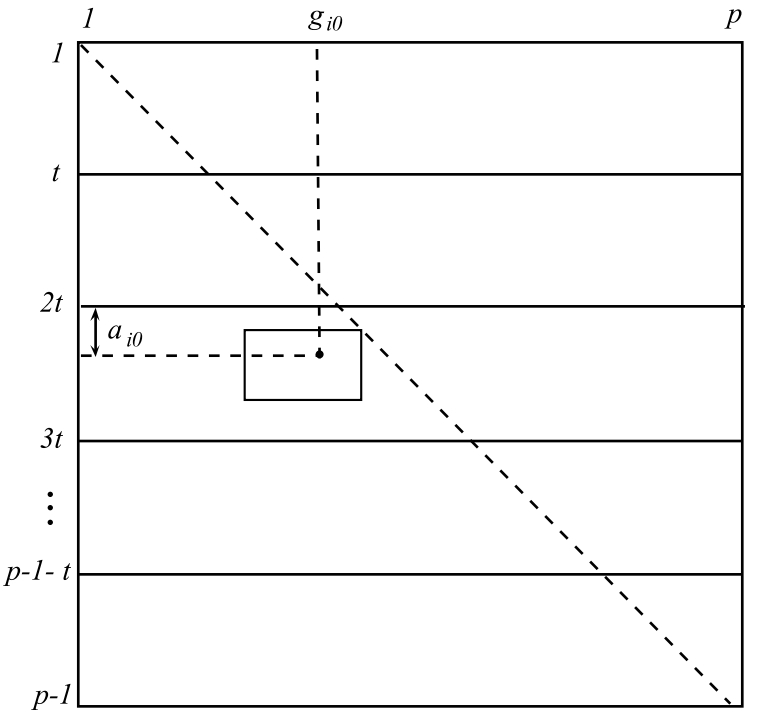


FIGURE 3. The small box has length $pt^{-\delta''/2}$ and height $t^{1-\delta''/2}$ and contains one point of the sequence.

Proposition 4.9. *For large enough prime p, every element $g \in (\mathbb{Z}/p\mathbb{Z})^{\times}$ of order $t > p^{1/2+\epsilon}$, can be represented as $n^n \equiv g \bmod p$ for some natural number $n < p^{2-\alpha}$ where $\alpha > 0$ is independent of p.*

Proof. For the equivalence $n^n \equiv g \mod p$ to have a solution with $n < p^{2-\alpha}$, we must have that

$$g_i^{g_i+k_i} \equiv g \mod p$$

for $k_i < p^{1-\alpha}$, for at least one value $i = i_0$. Then the required n is equal to $g_{i_0} + pk_{i_0}$. We now find such a solution using a counting argument.

As in [17], using the bound for the discrepancy in Proposition 4.6, the number of elements in the sequence that fall into the box B is

$$\mathcal{N}(B) > N(|B| - \Delta) > \phi(t)|B| - C'(\log^2 t)t^{1-\delta}$$

since $N = \phi(t)$. Therefore, if

$$|B| = \phi(t)^{-\delta'}, \text{ where } \delta' < \delta$$

we have $\mathcal{N}(B) > 0$ for p large enough. In other words, B contains at least one point,

$$\left\{\left(\frac{a_i}{t}, \frac{g_i}{p}\right)\right\}_{i=1}^{N}$$

from the sequence.

To renormalize we must multiply the area by pt since the length and width of the box were stretched by different factors. Thus, in the $t \times p$ rectangle every box of volume

$$|B| = pt\phi(t)^{-\delta'} = pt^{1-\delta''}, \text{ where } \delta'' < \delta' < \delta \tag{4.2}$$

contains a point of the sequence $\{(a_i, g_i)\}_{i=1}^{N}$.

The rectangles of size $t \times p$ will repeat within the larger $(p-1) \times p$ rectangle. Consider a box of height $t^{1-\delta''/2}$ and length $pt^{-\delta''/2}$. Position it so that its upper right corner is along the diagonal $(c, c), 1 \le c \le p-1$ and so that it falls within one of the $t \times p$ rectangles as in Figure 3. Then we have a point (a_{i_0}, g_{i_0}) contained in that box, so that

$$g_{i_0}^{a_{i_0}} \equiv g_{i_0}^{g_{i_0}+k_{i_0}} \equiv g \bmod p$$

where

$$k_{i_0} < pt^{-\delta''/2} + t^{1-\delta''/2} < p^{1-\alpha}$$ □

as required.

An explicit value of α which holds for all sufficiently large p can be calculated with the deterministic bound for $S_{m,n}$ from [19]. The best value found using this method is $\alpha < 0.00125$.

5. Conclusion

Given the difference between the expected bound $p^2 \log \phi(p-1)/\phi(p-1)$ and the deterministic bound $p^{2-\alpha}$, the natural question to ask is whether or not the deterministic bound can be improved. A stronger bound on exponential sums of the form $S_{m,n}$ would help, but a different approach might be needed. (A limitation of Proposition 4.9 is that the volume of $|B_{\text{renormalized}}|$ will be at least p, even if the bounds on $S_{m,n}$ are small; thus one cannot reach a bound lower than $p^{3/2}$ when t is close to $p^{1/2}$.) However, the counting argument presented here may be useful in finding non-trivial small solutions to other functions modulo p, as it was in [17].

Statistical methods were useful in predicting the behaviour of the self-power map. We anticipate modifying the coupon collector model to improve and obtain other estimates. A first improvement would be to try to calculate the cumulative distribution function for all primes. Also, it would be useful to find the conditions under which the least frequent coupons are the bottleneck in determining expectation. Then the coupon collector problem with blank coupons may be used to obtain estimates for a full collection of residue classes with a specified property.

Appendix A. The coupon collector problem with blank coupons

In this appendix we describe the modification of Flajolet et al.'s argument in [10] to the case of a coupon collector problem under non-uniform distribution with blank coupons.

A.1. **Formal Languages and Probabilities.** This introduction on formal languages, generating functions, and probabilities is taken from Section 2 of [10]. The aim is to relate formal descriptions of regular languages to generating functions and probability estimates.

Definition A.1. Let $\mathcal{A} = \{a_1, a_2, \ldots, a_m\}$ be a fixed set called the alphabet whose elements are letters. Then $\mathcal{A}^*$ denotes the set of all finite sequences, called the *words* or strings of $\mathcal{A}$. A *language* is any subset of $\mathcal{A}^*$.

Given languages L_1, L_2, and L, we define the following operations:

- *union* of L_1 and L_2 is $L_1 + L_2 = \{w | w \in L_1 \text{ or } w \in L_2\}$,
- *product* of L_1 and L_2 is $L_1 L_2 = \{w_1 w_2 | w_1 \in L_1 \text{ and } w \in L_2\}$,
- *star* L^* is the language formed from all possible sequences from elements of L,
$$L^* = \{\varepsilon\} + L + (L \cdot L) + +(L \cdot L \cdot L) + \ldots,$$
where ε denotes the empty word.

Definition A.2. The class of *regular languages* is the smallest class of languages containing the finite sets and closed under the three operations of union, product and star.

We extend the class of regular languages by the addition of another operation, the *shuffle product*. Given two words w_1 and w_2, their shuffle, denoted $(w_1 \sqcup\!\sqcup w_2)$, is the set of all words obtained by interlacing the letters of w_1 and w_2 in all possible ways, while maintaining their order inside w_1 and w_2. For example, if $w_1 = ab$ and $w_2 = xy$, then

$$(ab \sqcup\!\sqcup xy) = \{abxy, axby, axyb, xaby, xayb, xyab\}.$$

The shuffle of two languages L_1 and L_2 is

$$L_1 \sqcup\!\sqcup L_2 = \bigcup_{\substack{w_1 \in L_1 \\ w_2 \in L_2}} (w_1 \sqcup\!\sqcup w_2)\,.$$

These operations translate into generating functions which are used to estimate probabilities. Given a language L, let $l_{n_1,\ldots,n_m}$ be the number of words in L that have n_1 occurrences of letter $a_1, \ldots, n_m$ occurrences of a_m. The *multivariate function* of L is

$$l(z_1, \ldots, z_m) = \sum_{n_1,\ldots,n_m \geq 0} l_{n_1,\ldots,n_m} z_1^{n_1} z_2^{n_2} \cdots z_m^{n_m}.$$

We use the notation $[x^m y^n z^k] f(x, y, z)$ to denote the coefficient of the term $x^m y^n z^k$ in $f(x, y, z)$.

Associate to the alphabet $\mathcal{A} = \{a_1, a_2, \ldots, a_m\}$ a weight distribution where letter a_i has weight p_i. The weight extends multiplicatively to words so that $w = a_{j_1} a_{j_2} \ldots a_{j_n}$ has weight

$$\pi[w] = p_{j_1} p_{j_2} \cdots p_{j_n}.$$

In our case, the weight distribution is a probability distribution and $\sum_{i=1}^m p_i = 1$. Furthermore, $\pi[w]$ is the probability of a word w being in $\mathcal{A}^n$, the set of words with n letters.

Definition A.3. The function

$$l(z) = l(p_1 z, p_2 z, \ldots, p_m z) = \sum_{n_1,\ldots,n_m} l_{n_1,\ldots n_m} p_1^{n_1} \cdots p_m^{n_m} z^{n_1+\ldots+n_m} = \sum_{w \in \mathcal{A}^*} \pi[w] z^{|w|}$$

is called the *ordinary generating function* of the language L with respect to the weight distribution $\vec{p} = (p_1, \ldots, p_m)$ and is denoted by $l(z)$. The *exponential generating function* is defined with $z^n/n!$ replacing z^n,

$$\hat{l}(z) = \sum_{n_1,\ldots,n_m} l_{n_1,\ldots n_m} p_1^{n_1} \cdots p_m^{n_m} \frac{z^{n_1+\ldots+n_m}}{(n_1+\ldots+n_m)!} = \sum_{w\in\mathcal{A}^*} \pi[w] \frac{z^{|w|}}{|w|!}.$$

The value $[z^n]l(z)$ is the probability that a random word in $\mathcal{A}^n$ is in L.

The ordinary and exponential generating functions are related by the Laplace-Borel transform

$$l(z) = \int_0^\infty \hat{l}(zt)e^{-t}dt$$

following from the classical relation

$$\int_0^\infty t^n e^{-t}dt = n!$$

An operation on languages is unambiguous if every word of the resulting language is obtained only once. The following is Theorem 2.1 in [10]. Along with the Laplace transform integral, it is the key to determining the generating function of a language defined by a combination of the four basic operations.

Theorem A.4. *When they operate unambiguously on their arguments, the operations of union, product, star, and shuffle product translate into generating functions:*

$$\begin{array}{lllll} (a) & L = L_1 + L_2 & \Rightarrow & l(z) = l_1(z) + l_2(z), \\ (b) & L = L_1 \cdot L_2 & \Rightarrow & l(z) = l_1(z) l_2(z), \\ (c) & L = L_1^* & \Rightarrow & l(z) = (1 - l_1(z))^{-1}, \\ (d) & L = L_1 \sqcup\!\sqcup L_2 & \Rightarrow & \hat{l}(z) = \hat{l}_1(z)\hat{l}_2(z). \end{array}$$

A.2. **Proof of Proposition 3.1.** The proof of Proposition 3.1 parallels the proofs of Theorems 3.1 and 4.1 in [10]. We begin by rewriting Theorem 3.1 of [10] to obtain Proposition A.6 below, of which Proposition 3.1 is a corollary.

The alphabet $\mathcal{A}$ consists of m different coupons, and p_i is the probability of coupon $a_i \in \mathcal{A}$. Flajolet et al. considered the following problem.

Question A.5. Determine the expectation of the number B_j of elements that need to be drawn from $\mathcal{A}$ (with replacement) until we first obtain j distinct coupons that are each repeated at least k times.

The case $k = 2$ and $j = 1$ is the classical birthday problem of determining the expected number of random people needed such that two will share a birthday. The case $k = 1$ and $j = m$ is the coupon collector problem where a full collection is desired. Below we determine $E(B_j)$ for the case when the alphabet $\mathcal{A}$ contains blank coupons.

Proposition A.6. *Consider a set of m coupons under a general probability distribution $\{p_i\}_{i=1}^m$, which are drawn with replacement. Let m_B coupons, which appear with probabilities $p_1, p_2, \ldots, p_{m_B}$, be blanks. Let*

$$p_1 + p_2 + \ldots + p_{m_B} = \beta.$$

The expectation $E(B_j)$ of the time for obtaining $j \leq m - m_B$ distinct useful (i.e. non-blank) coupons, each appearing k times, is given by

(A.1)

$$E(B_j) = \sum_{q=0}^{j-1} \int_0^\infty [u^q] \left(\prod_{i=m_B+1}^{m} \left(\tilde{e}_{k-1}(p_i t) + u \left(e^{p_i t} - \tilde{e}_{k-1}(p_i t) \right) \right) \right) e^{(\beta-1)t} dt$$

where $\tilde{e}_{k-1}(t)$ represents the truncated exponential

$$\tilde{e}_{k-1}(t) = 1 + \frac{t}{1!} + \frac{t^2}{2!} + \ldots + \frac{t^k}{k!}.$$

Proof. We will use the terms 'useful' and 'non-blank' interchangeably. Let the random variable B_j be the number of draws for first obtaining at least k copies of j useful coupons in an infinite sequence of trials. Thus, it is a random variable defined on $\mathcal{A}^\infty$ with the product measure.

Let Y_n be the random variable defined on $\mathcal{A}^n$ representing the number of k-occurrences of different useful coupons in a sequence of n trials. As in [10], we have

$$\Pr\{Y_n \geq j\} = \Pr\{B_j \leq n\}$$

and

$$E(B_j) = \sum_{q=0}^{j-1} \left(\sum_{n \geq 0} \Pr\{Y_n = q\} \right).$$

It remains to estimate the sum $\sum_{n\geq 0} \Pr\{Y_n = q\}$.

Let H_q be the language consisting of words with exactly q useful letters that occur at least k times. Label the omitted letters (corresponding to the blank coupons) as $\alpha_1, \alpha_2, \ldots, \alpha_{m_B}$ and these may occur any number of times. All other letters occur at most $k-1$ times. With

$$\begin{aligned} \alpha^{<k} &= \epsilon + \alpha + \alpha^2 + \ldots + \alpha^{k-1}, \\ \alpha^{\geq k} &= \alpha^k \cdot \alpha^* \end{aligned}$$

the language H_q can be expressed as

$$\begin{aligned} H_q = &\left(\alpha_1^* \sqcup\!\sqcup \alpha_2^* \sqcup\!\sqcup \ldots \sqcup\!\sqcup \alpha_{m_B}^* \right) \\ &\sqcup\!\sqcup \bigcup_{I,J} \left(\alpha_{i_1}^{\geq k} \sqcup\!\sqcup \alpha_{i_2}^{\geq k} \sqcup\!\sqcup \ldots \sqcup\!\sqcup \alpha_{i_q}^{\geq k} \right) \sqcup\!\sqcup \left(\alpha_{j_1}^{<k} \sqcup\!\sqcup \alpha_{j_2}^{<k} \sqcup\!\sqcup \ldots \sqcup\!\sqcup \alpha_{j_r}^{<k} \right) \end{aligned}$$

where the union is over all sets I and J of cardinality q and $r = m - m_B - q$ such that

$$\begin{aligned} I &= \{i_1, i_2, \ldots, i_q\}, \\ J &= \{j_1, j_2, \ldots, j_r\}, \\ I \cap J &= \emptyset, \\ I \cup J &= \{m_B + 1, m_B + 2, \ldots, m\}. \end{aligned}$$

If α is a letter with probability σ, the exponential generating functions of α^*, $\alpha^{<k}$ and $\alpha^{\geq k}$ are $e^{\sigma z}$, $\tilde{e}_{k-1}(\sigma z)$ and $e^{\sigma z} - \tilde{e}_{k-1}(\sigma z)$, respectively. Thus, the exponential

generating function of H_q is

$$\hat{h}(z) = e^{\beta z} \sum_{I,J} \left(e^{p_{i_1} z} - \tilde{e}_{k-1}(p_{i_1} z)\right) \cdots \left(e^{p_{i_q} z} - \tilde{e}_{k-1}(p_{i_q} z)\right) \tilde{e}_{k-1}(p_{j_1} z) \cdots \tilde{e}_{k-1}(p_{j_r} z).$$

Therefore,

$$\hat{h}(z) = e^{\beta z} [u^q] \Phi(z, u)$$

where $[u^q]$ denotes the coefficient of the qth power of u of

$$\Phi(z, u) = \prod_{i=m_B+1}^{m} \left(\tilde{e}_{k-1}(p_i t) + u \left(e^{p_i t} - \tilde{e}_{k-1}(p_i t)\right)\right). \tag{A.2}$$

The ordinary generating function of H_q is given by the Laplace-Borel transform,

$$h_q(z) = \int_0^\infty e^{\beta z t} [u^q] \Phi(zt, u) e^{-t} dt.$$

Therefore,

$$\sum_{n \geq 0} \Pr\{Y_n = q\} = h_q(1) = \int_0^\infty [u^q] \Phi(t, u) e^{(\beta-1)t} dt,$$

and summing over q ranging from 0 to $j-1$, we have the statement of the proposition. □

We can now prove Proposition 3.1.

Proof. The equation (3.1) is a specialization of equation (A.1) in Proposition A.6 to the case $k = 1$.

To obtain equation (3.2) from (3.1), introduce the function

$$\Phi(t, u) = \prod_{i=m_B+1}^{m} \left(1 + u \left(e^{p_i t} - 1\right)\right) \tag{A.3}$$

which is the function Φ in equation (A.2) in the case $k = 1$. Expand $\Phi(t, u)$ as follows:

$$\Phi(t, u) = \sum_{q=0}^{m-m_B} \phi_q(t) u^q.$$

We have

$$\begin{aligned} E(T) = E(T_{m-m_B}) &= \int_0^\infty \sum_{q=0}^{m-m_B-1} [u^q] \Phi(t, u)|_{u=1} \, e^{(\beta-1)t} dt \\ &= \int_0^\infty \left(\phi_0(t) + \phi_1(t) + \ldots + \phi_{m-m_B-1}(t)\right) e^{(\beta-1)t} dt \\ &= \int_0^\infty \left(\Phi(t, 1) - \phi_{m-m_B}(t)\right) e^{(\beta-1)t} dt. \end{aligned}$$

We have that $\phi_{m-m_B}(t) = \prod_{i=m_B+1}^{m} (e^{p_i t} - 1)$ and, using equation (A.3),

$$\begin{aligned}\Phi(t,1) &= \prod_{i=m_B+1}^{m} e^{p_i t} = \prod_{i=1}^{m} e^{p_i t} \prod_{i=1}^{m_B} e^{-p_i t} \\ &= e^{(p_1+p_2+\ldots+p_m)t} e^{-(p_1+p_2+\ldots+p_{m_B})t} \\ &= e^{(1-\beta)t}.\end{aligned}$$

Once these expressions for $\phi_{m-m_B}(t)$ and $\Phi(t,1)$ are substituted in the integral, we obtain the required form for $E(T)$. □

Acknowledgments

The author is grateful to Professor Kumar Murty for suggesting the question and for his comments and constant encouragement, to Professor Balasubramanian for two very useful discussions, and to Professor Ram Murty for suggesting his article to the author. Thank you also to the two referees of this article for their detailed comments. In particular, the discussion of the saddle-point which method corrected an error in the proof. The results of this paper are contained in the author's doctoral thesis [1].

References

[1] C. V. Anghel, *The Self-power Map and Its Image Modulo a Prime*, ProQuest LLC, Ann Arbor, MI, 2013. Thesis (Ph.D.)–University of Toronto (Canada). MR3211738

[2] G. Bachman, *On an optimality property of Ramanujan sums*, Proc. Amer. Math. Soc. **125** (1997), no. 4, 1001–1003, DOI 10.1090/S0002-9939-97-03650-2. MR1363445 (97f:11067)

[3] R. Balasubramanian, private communication, 2011.

[4] A. Balog, K. A. Broughan, and I. E. Shparlinski, *On the number of solutions of exponential congruences*, Acta Arith. **148** (2011), no. 1, 93–103, DOI 10.4064/aa148-1-7. MR2784012 (2012f:11071)

[5] M. Blum and S. Micali, *How to generate cryptographically strong sequences of pseudorandom bits*, SIAM J. Comput. **13** (1984), no. 4, 850–864, DOI 10.1137/0213053. MR764183 (86a:68021)

[6] J. Bourgain and I. E. Shparlinski, *Distribution of consecutive modular roots of an integer*, Acta Arith. **134** (2008), no. 1, 83–91, DOI 10.4064/aa134-1-6. MR2429637 (2009e:11156)

[7] R. Crocker, *On residues of n^n*, Amer. Math. Monthly **76** (1969), 1028–1029. MR0248072 (40 #1326)

[8] P. Erdős and A. Rényi, *On a classical problem of probability theory* (English, with Russian summary), Magyar Tud. Akad. Mat. Kutató Int. Közl. **6** (1961), 215–220. MR0150807 (27 #794)

[9] P. Erdős and P. Turán, *On a problem in the theory of uniform distribution, I, II*, Indag. Math. **10** (1948), 370–378; 406–413.

[10] P. Flajolet, D. Gardy, and L. Thimonier, *Birthday paradox, coupon collectors, caching algorithms and self-organizing search*, Discrete Appl. Math. **39** (1992), no. 3, 207–229, DOI 10.1016/0166-218X(92)90177-C. MR1189469 (93i:68107)

[11] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2009. MR2483235 (2010h:05005)

[12] W. K. Hayman, *A generalisation of Stirling's formula*, J. Reine Angew. Math. **196** (1956), 67–95. MR0080749 (18,293f)

[13] J. Holden, *Fixed points and two-cycles of the discrete logarithm*, Algorithmic number theory (Sydney, 2002), Lecture Notes in Comput. Sci., vol. 2369, Springer, Berlin, 2002, pp. 405–415, DOI 10.1007/3-540-45455-1_32. MR2041100 (2005h:11277)

[14] J. F. Koksma, *Some Theorems on Diophantine Inequalities*, Scriptum no. 5, Math. Centrum Amsterdam, 1950. MR0038379 (12,394c)

[15] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Dover Publications, Inc., 2006.
[16] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, *Handbook of Applied Cryptography*, CRC Press Series on Discrete Mathematics and its Applications, CRC Press, Boca Raton, FL, 1997. With a foreword by Ronald L. Rivest. MR1412797 (99g:94015)
[17] M. R. Murty, *Small solutions of polynomial congruences*, Indian J. Pure Appl. Math. **41** (2010), no. 1, 15–23, DOI 10.1007/s13226-010-0015-z. MR2650097 (2011d:11288)
[18] I. Niven, H. S. Zuckerman, and H. L. Montogomery, *An Introduction to the Theory of Numbers*, 5th edition, John Wiley & Sons, Inc., 1991.
[19] I. E. Shparlinski, *Exponential sums with consecutive modular roots of an integer*, Q. J. Math. **62** (2011), no. 1, 207–213, DOI 10.1093/qmath/hap023. MR2774362 (2012c:11166)
[20] L. Somer, *The residues of n^n modulo p*, Fibonacci Quart. **19** (1981), no. 2, 110–117. MR614045 (82g:10007)
[21] P. Szűsz, *On a problem in the theory of uniform distribution (Hungarian)*, Compt. Rend. Premier Congrès Hongrois (1952), 461–472.

Department of Mathematics, University of Toronto, 40 St. George St., Toronto, ON M5S 2E4, Canada

E-mail address: catalina.anghel@alum.utoronto.ca