Uniform Distribution Theory **7** (2012), no.1, 147–154

uniform distribution theory

DISCRETE LOGARITHMS AND THEIR EQUIDISTRIBUTION

D. Jason Gibson

ABSTRACT. Let $g, r \geq 2$. Extending a result of Cobeli, we establish a multidimensional equidistribution result for the discrete logarithms $(\log_g x_1, \ldots, \log_g x_r)$ as $(x_1, \ldots, x_r)$ ranges over certain subsets of $[1, p-1]^r$ with $p \to \infty$ along a sequence of primes having g as a primitive root modulo p.

Communicated by Sergei Konyagin

Dedicated to the memory of Gérard Rauzy

1. Introduction

Let p be a prime number. We identify $\mathbb{F}_p$ with the set $\mathbb{Z}/p\mathbb{Z} = \{0, 1, \ldots, p-1\}$. Let g be a primitive root modulo p. For $x \in \mathbb{F}_p \backslash \{0\}$, the discrete logarithm problem requires finding the least nonnegative integer n such that $g^n \equiv x \bmod p$. Because g is a primitive root, such an n exists, and it lies in the interval $[0, p-1)$. In the sequel, we write $n = \log_g x = \log x$, and we call n the *discrete logarithm* of x to the base g.

Discrete logarithms, of considerable importance in cryptography (see Diffie and Hellman [3]), should exhibit various types of random behavior. Here, we answer a question raised by Cobeli [2] concerning the discrete logarithm along arithmetic progressions. To be precise, consider integers $a \geq 0$, $d > 0$, and $N > 0$, and define

$$\mathcal{J} = \{a + d, \ldots, a + Nd\} \subset [1, p-1]. \tag{1}$$

We show that, for fixed $g, r \geq 2$, the tuples $\left(\frac{\log_g x_1}{p-1}, \ldots, \frac{\log_g x_r}{p-1}\right)$ become equidistributed as $(x_1, \ldots, x_r)$ ranges over $\mathcal{J}^r$. (Here, $p \to \infty$ over a sequence of primes

2000 Mathematics Subject Classification: Primary 11A07; Secondary 11B50, 11L07.
Keywords: Discrete logarithms, exponential sums, primitive roots.
I thank the referee for detailed and helpful comments and corrections.

such that g is a primitive root mod p, and N must be sufficiently large in comparison to p, say $N \approx p^{1-1/(2r)+\epsilon}$.) To establish this fact, we connect the distribution of these discrete logarithms to discrepancy, which we then estimate using exponential sums.

2. Equidistribution

When one samples from our ambient space, e.g., $[0, p-1)^r$ or one of its subsets, it might be expected that, given some subset Ω of the unit cube with (normalized Haar) measure $\mu(\Omega)$, the tuples of scaled discrete logarithms $\left(\frac{\log_g x_1}{p-1}, \ldots, \frac{\log_g x_r}{p-1}\right)$ fall within Ω a fraction $\mu(\Omega)$ of the time. Under suitable hypotheses, this turns out to be the case, as will be demonstrated in Theorem 2.1. (This result should be compared to those of Granville, Shparlinski, and Zaharescu [5], which deal with the distribution of points under the action of rational maps on curves defined over $\mathbb{F}_p$.)

To make this precise, let $\mathbb{T}^r = [0,1)^r$ be the unit cube, and, for $\Omega \subset \mathbb{T}^r$, we define

$$\mu_{\log}(\Omega) = \frac{\#\{(x_1, \ldots, x_r) \in \mathcal{J}^r : \left(\frac{\log x_1}{p-1}, \ldots, \frac{\log x_r}{p-1}\right) \in \Omega\}}{\#\mathcal{J}^r}, \tag{2}$$

the number of points $(x_1, \ldots, x_r) \in \mathcal{J}^r$ for which the tuple of (appropriately scaled) discrete logarithms $\left(\frac{\log x_1}{p-1}, \ldots, \frac{\log x_r}{p-1}\right)$ falls within Ω.

Theorem 2.1. *Let $p > 2$ be prime, with g a primitive root mod p. Let $r \geq 2$. Let $\Omega \subset \mathbb{T}^r$ be a domain with piecewise smooth boundary. Then, for any integer $L > 1$, we have*

$$\mu_{\log}(\Omega) = \mu(\Omega) + O_{r,\Omega}\left(L^{-1/r} + \frac{1}{N} p^{1-1/(2r)} (\log p)\,(\log L)\right).$$

(Here, μ denotes normalized Haar measure on $\mathbb{T}^r$.)

3. Discrepancy

In this section, we closely follow the work of Granville, Shparlinski, and Zaharescu [5] in order to connect the distribution of the sequence of discrete logarithms to discrepancy via exponential sums. By the identification of $\mathbb{F}_p$ with $\{0, \ldots, p-1\}$, for $x \in \mathbb{F}_p$, we can consider the rational number $t(x) = x/(p-1)$.

(For this choice of normalization, recall that this discrete logarithm satisfies $\log x \in [0, p-1)$.) This allows us to map tuples of discrete logarithms $\mathbf{x} = (\log x_1, \ldots, \log x_r) \in [0, p-1)^r$ to

$$t(\mathbf{x}) = \left(\frac{\log x_1}{p-1}, \ldots, \frac{\log x_r}{p-1} \right) \in \mathbb{T}^r,$$

the unit cube $\mathbb{T}^r = [0,1)^r$.

For a finite set $A \subseteq \mathbb{T}^r$ and domain $\Omega \subseteq \mathbb{T}^r$, define the *discrepancy* $\Delta(A, \Omega)$ and *box discrepancy* $D(A)$ by

$$\Delta(A, \Omega) = \left| \frac{\#\{a \in A : a \in \Omega\}}{\#\{a \in A\}} - \mu(\Omega) \right|$$

and

$$D(A) = \sup_{\mathbb{B} \subseteq \mathbb{T}^r} \Delta(A, \mathbb{B}),$$

respectively, where in the latter expression, the supremum is taken over all boxes of the form $\mathbb{B} = \prod_{i=1}^{r} [\alpha_i, \beta_i]$.

For a vector $\mathbf{u} \in \mathbb{T}^r$ and a set $\Gamma \subset \mathbb{T}^r$, define the distance $\mathrm{dist}(\mathbf{u}, \Gamma)$ by

$$\mathrm{dist}(\mathbf{u}, \Gamma) = \inf_{\mathbf{w} \in \Gamma} \|\mathbf{u} - \mathbf{w}\|,$$

where $\|\mathbf{v}\|$ denotes the Euclidean norm of $\mathbf{v}$. For $\epsilon > 0$ and a domain $\Omega \subseteq \mathbb{T}^r$, define the sets

$$\Omega_\epsilon^+ = \{\mathbf{u} \in T^r \backslash \Omega \mid \mathrm{dist}(\mathbf{u}, \Omega) < \epsilon\}$$

and

$$\Omega_\epsilon^- = \{\mathbf{u} \in T^r \mid \mathrm{dist}(\mathbf{u}, \mathbb{T}^r \backslash \Omega) < \epsilon\}.$$

Let $b(\epsilon)$ be any increasing function defined for $\epsilon > 0$ with the additional property that $\lim_{\epsilon -> 0^+} b(\epsilon) = 0$. Define the class M_b of domains to be those domains $\Omega \subseteq \mathbb{T}^r$ for which

$$\mu(\Omega_\epsilon^+) \leq b(\epsilon) \qquad \text{and} \qquad \mu(\Omega_\epsilon^-) \leq b(\epsilon).$$

For $\Omega \in M_b$, we have a relation between $D(A)$ and $\Delta(A, \Omega)$, given by the following result (Theorem, page 106) from [8].

Lemma 3.1. *For any domain $\Omega \in M_b$, we have*

$$\Delta(A, \Omega) = O_r(b(r^{1/2} D(A)^{1/r})).$$

Box discrepancy bounds can be reduced to estimating certain exponential sums via the *Koksma-Szüsz inequality* [7, 10] (also see Theorem 1.21 of [4]).

Lemma 3.2. *For a set $A \subseteq \mathbb{T}^r$ of N points and any integer $L > 1$,*

$$D(A) = O\left(\frac{1}{L} + \frac{1}{N} \sum_{\substack{\mathbf{c}=(c_1,\ldots,c_r)\in\mathbb{Z}^r\setminus\{\mathbf{0}\} \\ |c_i|<L \text{for each } i}} \frac{1}{\prod_{i=1}^r (1+|c_i|)} \left|\sum_{\mathbf{a}\in A} \mathbf{e}(\mathbf{c}\cdot\mathbf{a})\right|\right).$$

To apply this result, it remains to estimate the appropriate exponential sum. In the next section, we first analyze the corresponding 1-dimensional sum, and then show how the needed estimate follows.

4. Some Exponential Sums

To establish Theorem 2.1, one requires an estimate for sums of multiplicative characters. We show explicitly the connection between the discrete logarithm and the exponential sums for the sake of directness.

To that end, let ζ be a pth root of unity. We first capture a single discrete logarithm via an exponential sum, by means of the expression

$$\frac{1}{p}\sum_{n=1}^{p} \zeta^{n(g^j - z)} = \begin{cases} 1 & g^j \equiv z \mod p, \\ 0 & g^j \not\equiv z \mod p. \end{cases} \tag{3}$$

This sum acts as an indicator function for detecting whether $j = \log z$, provided that $0 \le j \le p-2$, as will be the case in what follows.

Next, for eventual use in Lemma 3.2, we must now incorporate $(p-1)$st roots of unity. In the sequel, we write $e_{p-1}(z) = \exp\left(\frac{2\pi i z}{p-1}\right)$. Then, for k with $0 \le k \le p-2$, we need to detect $e_{p-1}(k \log z)$ via exponential sums. To that end, write

$$\begin{aligned} e_{p-1}(k\log z) &= \sum_{j=0}^{p-2} e_{p-1}(kj) \frac{1}{p} \sum_{n=1}^{p} \zeta^{n(g^j - z)} \\ &= \frac{1}{p} \sum_{n=1}^{p} \sum_{j=0}^{p-2} e_{p-1}(kj) \zeta^{n(g^j - z)}. \end{aligned}$$

We have turned the type of expression that will occur in our use of Lemma 3.2 into something that can be estimated by standard techniques. Specifically, we

must estimate

$$\sum_{z\in\mathcal{J}} e_{p-1}(k\log z) = \frac{1}{p}\sum_{n=1}^{p}\sum_{j=0}^{p-2} e_{p-1}(kj)\zeta^{ng^j}\sum_{z\in\mathcal{J}} e_p(-nz). \tag{4}$$

The current arrangement of this expression suggests bounding the innermost sum first, removing the dependence on the particular structure of $\mathcal{J}$, and then handling the remaining portion. The following lemmas give both of the required bounds.

Lemma 4.1. *Let $a \geq 0$, $d > 0$, and $N > 0$. Let p be a prime, and let*

$$\mathcal{J} = \{a+d, \ldots, a+Nd\} \subset [1, p-1].$$

Then

$$\left|\sum_{z\in\mathcal{J}} e_p(-nz)\right| \leq \min\left(N, \left(2\left\|\frac{nd}{p}\right\|\right)^{-1}\right). \tag{5}$$

P r o o f. See (2.6) of [2]. □

Lemma 4.2. *Let p be prime, and let g be a primitive root mod p. For $0 \leq k \leq p-2$ and $n \bmod p$ with the additional condition that $1 \leq k \leq p-2$ or $n \not\equiv 0 \bmod p$ (so that $e_{p-1}(k) = \zeta^n = 1$ does not hold), then*

$$\left|\sum_{j=0}^{p-2} e_{p-1}(kj)\zeta^{ng^j}\right| \leq \sqrt{p}. \tag{6}$$

P r o o f. Write $x = g^j$, so that $\log x = j$. Then $e_{p-1}(kj)$ can be viewed as a multiplicative character χ, where $\chi(x) = e_{p-1}(kj)$. Our sum (6) reduces to a Gauss sum,

$$\sum_{j=0}^{p-2} e_{p-1}(kj)\zeta^{ng^j} = \sum_{x=1}^{p-1}\chi(x)\zeta^{nx},$$

and the stated bound follows from [9], Theorem 4.4.19 (also see [6], Chapter 3.4). □

Now, we require a version (see (2.7) of [2]) of the Pólya-Vinogradov inequality (also see Theorem 12.5 in Chapter 12 of [6]) for our sum (4).

Lemma 4.3. *With our previous notation, if $1 \leq k \leq p-2$, we have*

$$\left|\sum_{z\in\mathcal{J}} e_{p-1}(k\log z)\right| \leq \sqrt{p}(2+\log p). \tag{7}$$

P r o o f. Recalling (4) and then using Lemma 4.1 and Lemma 4.2, we have

$$\begin{aligned}\left|\sum_{z\in\mathcal{J}} e_{p-1}(k\log z)\right| &= \left|\frac{1}{p}\sum_{n=1}^{p}\sum_{j=0}^{p-2} e_{p-1}(kj)\zeta^{ng^j}\sum_{z\in\mathcal{J}} e_p(-nz)\right| \\ &\le \frac{1}{p}\sum_{n=1}^{p} p^{1/2}\min\left(N,\left(2\left\|\frac{nd}{p}\right\|\right)^{-1}\right) \\ &\le p^{1/2}+p^{-1/2}\sum_{n=1}^{p}\left(2\left\|\frac{nd}{p}\right\|\right)^{-1} \\ &\le p^{1/2}+p^{-1/2}\sum_{n=1}^{\frac{p-1}{2}}\frac{p}{n} \\ &\le \sqrt{p}(2+\log p),\end{aligned}$$

verifying the claim. □

5. Conclusion

The distribution of our sequence claimed in Theorem 2.1 now follows readily.

P r o o f o f T h e o r e m 2.1. We will use Lemma 4.3 to bound the exponential sum appearing in the Lemma 3.2 statement of the Koksma-Szüsz inequality. For the values c_i in Lemma 3.2 with $c_i \not\equiv 0 \pmod{p-1}$, the bound from Lemma 4.3 applies. For the values c_i in Lemma 3.2 with $c_i \equiv 0 \pmod{p-1}$, we use the trivial bound p. Then, focusing on the contribution from one of the r coordinates in Lemma 3.2, we get

$$\sum_{\substack{|c_i|<L\\ c_i\neq 0}} \frac{1}{1+|c_i|}\left|\sum_{z\in\mathcal{J}} \mathbf{e}(c_i t(z))\right| = O\left((\log L)\sqrt{p}\log p\right) \tag{8}$$

from the $c_i \neq 0$ terms, and a trivial bound of p for the $c_i = 0$ term.

To complete the estimate of the sum appearing in Lemma 3.2, note that any particular r-tuple $(c_1,\dots,c_r)$ contains at most $r-1$ zero entries. From Lemma 3.2, we then obtain

$$D\left(\mathcal{J}^r\right) = O_r\left(\frac{1}{L}+\frac{1}{N^r}p^{r-1/2}(\log p)^r(\log L)^r\right)$$

for the box discrepancy of $\mathcal{J}^r$. This establishes that, for boxes $\mathbb{B}$, we have

$$\mu_{\log}(\mathbb{B}) = \mu(\mathbb{B}) + O_r\left(\frac{1}{L} + \frac{1}{N^r}p^{r-1/2}\left(\log p\right)^r\left(\log L\right)^r\right). \tag{9}$$

The piecewise smooth boundary of Ω allows us to conclude that, for sufficiently small $\epsilon > 0$, we have $\mu\left(\Omega_\epsilon^{\pm}\right) \ll_\Omega \epsilon$.

By Lemma 3.1, we get

$$\begin{aligned}|\mu_{\log}(\Omega) - \mu(\Omega)| &\leq \Delta(\mathcal{J}^r, \Omega) \ll_{r,\Omega} D(\mathcal{J}^r)^{1/r} \\ &\ll_{r,\Omega} L^{-1/r} + \frac{1}{N}p^{1-1/(2r)}(\log p)\left(\log L\right).\end{aligned}$$

□

6. Further Work

This line of inquiry suggests several further questions. The framework described in Cobeli's work [2] asks for analysis of the distribution of the discrete logarithm under (nonlinear) transformations, e.g., under polynomial transformations, which would require bounds for

$$\sum_{z\in\mathcal{J}^r} e_{p-1}\left(P(\log z)\right),$$

where P is some fixed polynomial of degree greater than 2. See the work of Banks and Shparlinski [1] for a treatment of a sum of similar type. Similarly, one can seek to answer questions on the distribution of discrete logarithms to distinct bases $g_1, \ldots, g_r$.

Finally, other nice subsets (not merely subintervals or arithmetic progressions) of $[1, p-1]$ should also exhibit random distributions of the values of the discrete logarithm function.

REFERENCES

[1] William D. Banks and Igor E. Shparlinski. Exponential sums with polynomial values of the discrete logarithm. *Unif. Distrib. Theory*, 2(2):67–72 (electronic), 2007.

[2] Cristian Cobeli. On the discrete logarithm problem, 2008. arxiv:0811.4182v1 [math.NT].

[3] Whitfield Diffie and Martin E. Hellman. New directions in cryptography. *IEEE Trans. Information Theory*, IT-22(6):644–654, 1976.

[4] Michael Drmota and Robert F. Tichy. *Sequences, discrepancies and applications*, volume 1651 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1997.

[5] Andrew Granville, Igor E. Shparlinski, and Alexandru Zaharescu. On the distribution of rational functions along a curve over $\mathbb{F}_p$ and residue races. *J. Number Theory*, 112(2):216–237, 2005.

[6] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[7] J. F. Koksma. *Some theorems on Diophantine inequalities*. Scriptum no. 5. Math. Centrum Amsterdam, 1950.

[8] M. Laczkovich. Discrepancy estimates for sets with small boundary. *Studia Sci. Math. Hungar.*, 30(1-2):105–109, 1995.

[9] Steven J. Miller and Ramin Takloo-Bighash. *An invitation to modern number theory*. Princeton University Press, Princeton, NJ, 2006. With a foreword by Peter Sarnak.

[10] Péter Szüsz. Über ein Problem der Gleichverteilung. In *Comptes Rendus du Premier Congrès des Mathématiciens Hongrois, 27 Août–2 Septembre 1950*, pages 461–472. Akadémiai Kiadó, Budapest, 1952.

Received November 7, 2010
Accepted August, 29, 2011

Department of Mathematics & Statistics
Eastern Kentucky University
521 Lancaster Avenue
Richmond, KY 40475

E-mail: jason.gibson@eku.edu