PRIMITIVE ROOTS: A SURVEY

Shuguang Li
Department of Mathematics, Natural Sciences Division, University of Hawaii–Hilo, 200 W. Kawili Street, Hilo, HI 96720-4091
shuguang@hawaii.edu

Carl Pomerance
Fundamental Mathematics Research, Mathematics Center, Lucent Technologies, Bell Laboratories, 600 Mountain Ave., Murray Hill, NJ 07974-0636
carlp@lucent.com

Abstract For primes p, the multiplicative group of reduced residues modulo p is cyclic, with cyclic generators being referred to as primitive roots. Here we survey a few results and conjectures on this subject, and we discuss generalizations to arbitrary moduli. A primitive root to a modulus n is a residue coprime to n which generates a cyclic subgroup of maximal order in the group of reduced residues modulo n. [1]

Keywords: primitive root, Artin's conjecture, Generalized Riemann Hypothesis

Introduction

For a prime p, the multiplicative group $(\mathbf{Z}/p\mathbf{Z})^*$ is cyclic. Number theorists refer to any cyclic generator of this group as a primitive root modulo p. There are many attractive theorems and conjectures concerning primitive roots, and we shall survey some of them here. But it is also our intention to broaden the playing field, so to speak, and introduce the concept of a primitive root for a composite modulus n. It is well-known that for most numbers n, the multiplicative group $(\mathbf{Z}/n\mathbf{Z})^*$ is *not* cyclic (namely, $(\mathbf{Z}/n\mathbf{Z})^*$ is not cyclic for any number $n > 4$ that is not of the form $p^a, 2p^a$ for p an odd prime). So what then do we mean by a primitive root for n? In any finite group G one may look at elements whose

[1]This article also appeared in "New Aspects of Analytic Number Theory" (RIMS Kokyuroku No. 1274), 2002.

order is the maximum order over all elements in G. We do precisely this, and say that such elements for the group $G = (\mathbf{Z}/n\mathbf{Z})^*$ are primitive roots modulo n. That is, a primitive root modulo n is an integer coprime to n such that the multiplicative order of this integer modulo n is the maximum over all integers coprime to n. This concept reduces to the usual notion in the case that G is cyclic, so there should be no confusion. We shall see that there are a few surprises in store when we consider primitive roots for composite moduli. For a more traditional survey on primitive roots, see Murty [12].

The number of primitive roots for a given modulus

A basic question that one might ask is a formula for $R(n)$, the number of primitive roots for a given modulus n, and beyond that, a study of the order of magnitude of $R(n)$ as a function of n. For primes, the situation is straightforward. If g is a primitive root modulo p then all of the primitive roots for p are of the form g^a where a is coprime to $p-1$. Thus $R(p) = \varphi(p-1)$ where φ is Euler's function. This fact is well-known, but less well-known is that $\varphi(p-1)/(p-1)$ has a continuous distribution function. That is, let $D(u)$ denote the relative asymptotic density in the set of all primes of the set $\{p \text{ prime} : R(p)/(p-1) \le u\}$. Then $D(u)$ exists for every real number u, $D(u)$ is a continuous function of u, and $D(u)$ is strictly increasing on $[0, 1/2]$, with $D(0) = 0$, $D(1/2) = 1$. This beautiful result, which echoes Schoenberg's theorem on $\varphi(n)/n$, is due to Kátai [5].

It is not so easy to get a formula for $R(n)$ in general. It may be instructive to first consider the case of a general finite abelian group G. Write G as a product of cyclic groups of prime power order. For each prime p dividing the order of G, let p^{λ_p} be the highest power of p that appears as an order of one of these cyclic factors, and let ν_p be the number of times that this cyclic factor appears. Then the maximal order of an element in G is

$$\prod_{p||G|} p^{\lambda_p},$$

and the number of elements of G with this order is

$$|G| \prod_{p||G|} \left(1 - p^{-\nu_p}\right).$$

To see the latter assertion, note that an element g will have p^{λ_p} dividing its order if and only if at least one of its projections in the ν_p cyclic factors of order p^{λ_p} has order p^{λ_p}. The chance that one particular projection does not have this order, that is, it is killed by the exponent $p^{\lambda_p - 1}$, is

$1/p$. Thus, the fraction of elements g for which each of the ν_p projections is killed by the exponent $p^{\lambda_p - 1}$ is $p^{-\nu_p}$, so the fraction for which at least one projection has order p^{λ_p} is $1 - p^{-\nu_p}$. The assertion follows.

To apply this result to $G = (\mathbf{Z}/n\mathbf{Z})^*$ we must compute the numbers ν_p for this group. By the Chinese remainder theorem, G has a decomposition into the product of the groups $(\mathbf{Z}/q^a\mathbf{Z})^*$, where q is prime and $q^a \| n$. Further, the groups $(\mathbf{Z}/q^a\mathbf{Z})^*$ are themselves cyclic unless $q = 2$ and $a \geq 3$, in which case $(\mathbf{Z}/2^a\mathbf{Z})^*$ is the product of a cyclic group of order 2 and a cyclic group of order 2^{a-2}. It is thus a simple task to further refine the decomposition afforded by the Chinese remainder theorem into a factorization of $(\mathbf{Z}/n\mathbf{Z})^*$ into cyclic groups of prime power order. We thus can work out a formula, albeit not so simple, for $R(n)$. For the sake of completeness, we record this formula: If q^a is a prime power, let $\lambda(q^a)$ be the order of the largest cyclic subgroup of $(\mathbf{Z}/q^a\mathbf{Z})^*$; thus, $\lambda(q^a) = \varphi(q^a)$ if q is odd or if $q = 2$ and $a < 3$, while if $q = 2$ and $a \geq 3$, then $\lambda(2^a) = \frac{1}{2}\varphi(2^a) = 2^{a-2}$. If the prime factorization of n is $\prod_{i=1}^k q_i^{a_i}$, and p is a prime with $p|\varphi(n)$, let λ_p be the largest number such that $p^{\lambda_p}|\lambda(q_i^{a_i})$ for some i. If p is odd, let ν_p be the number of i's with $p^{\lambda_p}|\lambda(q_i^{a_i})$. If $p = 2$ and either $\lambda_2 > 1$ or $n \not\equiv 8 \pmod{16}$, the definition of ν_2 is the same. If $p = 2$, $\lambda_2 = 1$, and $n \equiv 8 \pmod{16}$, then $\nu_2 = k+1$. Then

$$R(n) = \varphi(n) \prod_{p|\varphi(n)} \left(1 - p^{-\nu_p}\right).$$

In analogy with Kátai's theorem about $R(p)/(p-1)$, one might ask if $R(n)/\varphi(n)$, has a distribution function. That is, for a given real number u does the set

$$\mathcal{R}_u := \{n \,:\, R(n)/\varphi(n) \leq u\}$$

have a natural density? Our first surprise is that the answer is no. It is shown by the first author in [7], [8] that there are values of u so that $\mathcal{R}_u$ does not have a natural density. In fact, there is a small positive number δ such that for every $u > 0$, $\mathcal{R}_u$ has upper density at least δ, but the lower density tends to 0 as $u \to 0$.

There are other naturally occurring sets in number theory where there is no natural density. For example, consider the set of integers n with an even number of decimal digits. While the natural density does not exist (the fraction of numbers in the set at 10^{2n} is at least $9/10$ while the fraction in the set at 10^{2n+1} is at most $1/10$), note that this set does have a logarithmic density. That is, the sum of the reciprocals of the numbers in the set that are $\leq x$, when divided by $\ln x$, approaches a limit, namely $1/2$. (It is interesting to note that logarithmic density is equivalent to the concept of Dirichlet density from analytic number theory.) Well, perhaps the set of numbers with an even number of decimal digits is not so natural a concept. But also consider the set of integers n with $\pi(n) > \mathrm{li}\,(n)$. (Here, $\pi(x)$ is the number of primes in the interval $[1, x]$ and $\mathrm{li}\,(x) = \int_0^x dt/\ln t$, where the principal value is taken for the singularity at $t = 1$.) It was once thought that there should be no values of n with $\pi(n) > \mathrm{li}\,(n)$, until Littlewood showed that there are infinitely many with the inequality holding, and also infinitely many with the reverse inequality. It is shown in Rubinstein and Sarnak [13] that assuming reasonable conjectures concerning the zeroes of the Riemann zeta function, the set of integers n with $\pi(n) > \mathrm{li}\,(n)$ does not have a natural density, but it does have a logarithmic density. Similar results pertain to the set of integers n with $\pi(n, 4, 1) > \pi(n, 4, 3)$, where $\pi(x, k, l)$ denotes the number of primes p in $[1, x]$ that are in the residue class $l \pmod{k}$.

So maybe the sets $\mathcal{R}_u$ have a logarithmic density? Alas, the answer is again no, as is shown in [7]. In fact, the oscillation persists at even the double logarithmic density (where one sums $1/a \ln a$ for members a of the set that are in $[2, x]$ and divides the sum by $\ln \ln x$). Maybe the *triple* logarithmic density exists: In [7] it is shown that at the triple level, $\mathcal{R}_u$ has upper density tending to 0 as $u \to 0$.

The source of the oscillation

Where does this surprising oscillation come from? The answer lies in the numbers ν_p described above. Consider a game played with n coins: We give you n coins, and at the end of the game you will either have given us back all n of the coins, or you will have given us back $n - 1$ coins, keeping one for yourself. Here's how the game is played. You flip the n coins (assume they are all fair coins with a $1/2$ probability of landing heads—the front of the coin—and a $1/2$ probability of landing tails—the back of the coin.), returning to us all of the coins that land tails. If there is more than one coin left, you repeat the process. If at any time you have exactly one coin left, you get to keep it. What is the probability P_n that you win the game by getting to keep a coin? It is not so hard to work out an expression for P_n, it is

$$P_n = \sum_{k=1}^{\infty} n 2^{-k} \left(1 - 2^{1-k}\right)^{n-1}.$$

Indeed, if one keeps flipping until no coins are left, and the last coin leaves on round k, with the other $n - 1$ coins leaving on earlier rounds, then the probability of this is $n2^{-k}\left(1 - 2^{1-k}\right)^{n-1}$. (There are n choices for the "last" coin, the probability it falls heads $k - 1$ straight times followed by a tails is 2^{-k}, and the probability that each of the other

$n-1$ coins has at least one tails in the first $k-1$ flips is $\left(1-2^{1-k}\right)^{n-1}$.) But more interestingly, one can ask:

$$\text{What is } \lim_{n\to\infty} P_n?$$

It is easy to convince oneself that when n is large, the biggest contribution to the sum for P_n is from the terms k with $2^k \approx n$. Suppose $0 \le \alpha < 1$ and S_α is an infinite set of natural numbers n such that the fractional part of the base-2 logarithm for $n \in S_\alpha$ converges to α modulo 1. For example, if $\alpha = 0$, then we might take S_0 as the set of powers of 2. Or we might also throw in the numbers of the form $2^m - 1$ and numbers of the form $2^m + m^2$. Then

$$\lim_{n\to\infty,\, n\in S_\alpha} P_n = \sum_{j=-\infty}^{\infty} 2^{-\alpha-j} e^{-2^{1-\alpha-j}}.$$

From this result it surely looks like the limiting value of P_n actually depends on α, the limiting value of the fractional part of the base-2 logarithm of n. That is, it looks like $\lim_{n\to\infty} P_n$ does not exist!

And this is indeed the case, though the oscillation in P_n is very gentle. We have $\limsup P_n \approx 0.72135465$ which is achieved when $\alpha \approx 0.139$, and $\liminf P_n \approx 0.72134039$ which is achieved when $\alpha \approx 0.639$. That is, the oscillation is only in the fifth decimal place! (For more on this kind of oscillation in probability theory, see [1] and the references in the acknowledgment of priority therein, and [6].)

It may be unclear what this game has to do with $R(n)$. Consider the number ν_2: If 2^{λ_2} is the highest power of 2 dividing the order of an element modulo n, then ν_2 is the number of cyclic factors of order 2^{λ_2} in $(\mathbf{Z}/n\mathbf{Z})^*$. We might ask where these ν_2 factors come from. But for a set of numbers n of density 0 we have ν_2 equal to the number of primes $p|n$ with $p \equiv 1 \pmod{2^{\lambda_2}}$. Now think of the odd primes dividing n as the coins in the game. Those primes $p \equiv 3 \pmod 4$ are the coins that turn tails on the first round and are returned. Those primes $p \equiv 5 \pmod 8$ are returned on the second round of flips, and so on. The number of primes that are alive in the last round is ν_p, and from our coin experience, we see that there is some oscillation for the probability that $\nu_p = 1$. But what corresponds to the number of coins? This is the number of odd prime factors of n, which is normally about $\ln\ln n$. Thus the limiting probability should depend on the fractional part of $\ln\ln\ln n/\ln 2$. With all of these iterated logarithms, it may begin to be clear why the density oscillation persists at logarithmic and double logarithmic levels.

But we noticed that the oscillation for the coin game is very slight. To see why there are great oscillations in the normal value of $R(n)$, we need to bring the other numbers ν_p into play for higher values of p. This then suggests a game played with unfair coins, where the probability of landing heads is $1/p$. An analysis of this game shows that there is again oscillation for the probability of winning, and as p tends to infinity, the ratio of the limsup of the probability to the liminf of the probability tends to infinity. In particular, if x tends to infinity in such a way that the fractional part of $\ln\ln\ln x/\ln p$ is very nearly 1 for all small primes p, then for most numbers n up to x, the values ν_p will frequently be 1 for these small primes p, so that then $R(n) = o(\varphi(n))$ for most numbers n up to x. But if x tends to infinity in such away that $\ln\ln\ln x/\ln p$ has fractional part about $1/2$ for all small primes p, then the values ν_p will mostly be > 1, so that $R(n)/\varphi(n)$ is bounded away from 0 for most numbers n up to x.

Artin's conjecture

Rather than fixing the modulus and asking for the number of primitive roots, as we have been doing, we may do the reverse: Fix an integer a and ask for the number of primes (or integers) for which a is a primitive root. Artin's famous conjecture deals with primes, and gives a supposedly necessary and sufficient condition on when there are infinitely many primes p with primitive root a. For example, take $a = 10$. (Note that in the special case $a = 10$, Gauss already had conjectured that there are infinitely many primes p with primitive root 10.) Note that 10 is a primitive root for a prime $p \neq 2, 5$ if and only if the length of the period for the decimal for $1/p$ has length $p-1$. Thus, the Artin–Gauss problem might even be understandable to a school child.

Since primes $p > 2$ are all odd, the groups $(\mathbf{Z}/p\mathbf{Z})^*$ all have even order, so that squares cannot be cyclic generators. Clearly too, the number -1 has order dividing 2 in $(\mathbf{Z}/p\mathbf{Z})^*$, so that -1 cannot be a cyclic generator when $p > 3$. Thus, a necessary condition on a for there to be infinitely many primes p with primitive root a is that a should not be a square and that a should not be -1. Artin's conjecture is that these trivially necessary conditions are also sufficient:

Artin's conjecture. *If the integer a is not a square and not -1, then there are infinitely many primes with primitive root a.*

Artin also formulated a strong form of this conjecture:

Artin's conjecture, strong form. *If the integer a is not a square and not -1, then there is a positive number $A(a)$ such that the number of primes $p \leq x$ with primitive root a is $\sim A(a)\pi(x)$.*

Artin gave a heuristic argument for a formula for the numbers $A(a)$ appearing in the conjecture, but, as reported in [4], after some numerical experiments of the Lehmers which cast some doubt on Artin's formula, Heilbronn revised Artin's heuristic argument and came up with a formula which agreed better with the numerical experiments. Let

$$A = \prod_{q \text{ prime}} \left(1 - \frac{1}{q(q-1)}\right) = 0.3739558136\ldots,$$

the number known as *Artin's constant.* Write a as $a_1 a_2^2$, where a_1 is squarefree. We are assuming that a is not a square, but it might be some other power. Let h be the largest integer for which a is an h-th power, so that necessarily h is odd. In the case that $h = 1$, that is, that a is not any power higher than the first power, the Artin–Heilbronn formula for $A(a)$ is fairly simple; it is

$$A(a) = \begin{cases} A, & \text{if } a_1 \not\equiv 1 \pmod 4 \text{ and } h = 1 \\ A\Big(1 - \prod_{q|a_1} \dfrac{1}{1+q-q^2}\Big), & \text{if } a_1 \equiv 1 \pmod 4 \text{ and } h = 1. \end{cases}$$

In particular, if $h = 1$, then $A(a)$ is uniformly bounded away from 0. In the case that $h > 1$, the formula is more complicated:

$$A(a) = \begin{cases} A \prod_{q|h} \dfrac{q^2-2q}{q^2-q-1}, & \text{if } a_1 \not\equiv 1 \pmod 4 \\ A \prod_{q|h} \dfrac{q^2-2q}{q^2-q-1} \Big(1 - \prod_{q|a_1} \dfrac{1}{1+q-q^2} \prod_{q|(a_1,h)} \dfrac{q^2-q-1}{q-2}\Big), & \\ \qquad \text{if } a_1 \equiv 1 \pmod 4. \end{cases}$$

(This more complicated formula reduces to the earlier one in the case that $h = 1$.)

Where do these formulas come from? Understanding at least the appearance of Artin's constant is relatively simple. Assume that $h = 1$, that is, assume that a is not a nontrivial power. For a to be a primitive root modulo a prime p, it must be the case that for each prime q that divides $p - 1$ (namely, the order of the group $(\mathbf{Z}/p\mathbf{Z})^*$), a is not a q-th power modulo p. These conditions are not only necessary, they are sufficient. Say that p "passes the q-test" if either q does not divide $p-1$ or $q|p-1$ and a is not a q-th power modulo p. (Passing the q-test for a prime p is equivalent to q not dividing the index of the subgroup generated by

a in $(\mathbf{Z}/p\mathbf{Z})^*$). By the Chebotarev density theorem, the proportion of primes p with $p \equiv 1 \pmod q$ and a is a q-th power modulo p is $1/q(q-1)$. Thus, the proportion of primes p that pass the q-test is $1 - 1/q(q-1)$. Assuming "independence", the product of these expressions, which is Artin's constant, should then give the density of primes p for which a is a primitive root.

But are the events independent? In fact, if a_1, the squarefree part of a, is not $1 \pmod 4$, then it can be shown by the Chebotarev theorem that for fixed primes q, the q-tests are independent. And in the general case, the correct joint densities may be computed.

So why then is the strong form of Artin's conjecture not a theorem? The answer lies in the tail of the inclusion-exclusion. One can prove rigorously that if $\psi(x)$ tends to infinity very slowly with x, then the proportion of primes p for which the index of the subgroup generated by a in $(\mathbf{Z}/p\mathbf{Z})^*$ is not divisible by any prime $q \leq \psi(x)$ is indeed asymptotically $A(a)$. To complete the proof one needs to exclude those primes p which fail the q-test for some prime $q > \psi(x)$. We would only need a crude upper bound for these counts, such as $\leq c/q^2$ of all primes, or even $\leq c/q \ln q$ of all primes. However, we have nothing better than $\leq c/q$ afforded by the Brun–Titchmarsh inequality. And so, the strong form of Artin's conjecture remains just that, a conjecture.

Hooley [4] however, has made the above heuristic into a rigorous proof under the assumption of the Generalized Riemann Hypothesis. This hypothesis allows a stronger form of the Chebotarev theorem which gives an estimate of $O(\pi(x)/q^2 + x^{1/2} \ln x)$ primes up to x which fail the q-test, uniformly for $q \leq x^{1/2}/\ln^2 x$. Larger primes q may then be handled by an elementary argument that does not involve the GRH.

A parallel with another problem may be instructive here. Let $S(x)$ be the number of primes $p \leq x$ with $p-1$ squarefree. Here the "q-test" is that we should not have $q^2 | p-1$. The proportion of primes p which pass this q-test is then $1 - 1/q(q-1)$, by the prime number theorem for arithmetic progressions. The Chinese remainder theorem implies we have independence, without any exceptional cases, so that we may conjecture that $S(x) \sim A\pi(x)$. However, in this case, the heuristic may be turned into a rigorous and unconditional proof, since Brun–Titchmarsh allows a good uniform upper estimate for the distribution of primes p which fail the q-test for the primes $q < x^\epsilon$, and a trivial argument can be used for larger primes q. (Actually, we can use the Page–Siegel–Walfisz theorem instead of Brun–Titchmarsh.) The difference here is that we have tools for handling large primes q that are not readily available in the Artin context.

It is interesting that not only do we not have a proof of the strong form of Artin's conjecture, we do not have a proof of the weak form either, not for any single number a. However, if several numbers a are thrown in together, there are theorems. The most intriguing perhaps is the result of Heath-Brown [3] (based on earlier work of Gupta and Murty [2]) that there are at most two *prime* values of a for which the weak form of Artin's conjecture is false. Nevertheless, we repeat, we do not know a single value of a for which the conjecture is true.

Allowing more values of a, we can even show the strong form of Artin's conjecture unconditionally on average. Let $P_a(x)$ denote the number of primes $p \le x$ which have a as a primitive root. It is relatively easy to estimate $\frac{1}{x \ln x} \sum_{1 \le a \le x \ln x} P_a(x)$, showing it to be $\sim A\pi(x)$. (Note that the average of the numbers $A(a)$ is $\sim A$, so that this result for $P_a(x)$ on average is consistent with the strong form of Artin's conjecture.) The sum of $P_a(x)$ may be thought of as the number of pairs a, p with $1 \le a \le x \ln x$, p a prime with $p \le x$, and a is a primitive root modulo p. Thus, the sum may be reorganized as a sum over primes p, and then we may use the trivial result that there are $\varphi(p-1)$ primitive roots modulo p in every interval of p consecutive integers. Far less trivial is to get an average estimate with a running over an interval of the shape $[1, x^\epsilon]$. The champion theorem here is due to Stephens [14] (improving on earlier work of Goldfeld), and ϵ may be taken as $4(\ln\ln x/\ln x)^{1/2}$.

Artin's conjecture for composite moduli

We saw that it makes perfectly good sense to consider primitive roots for composite moduli, namely, a is a primitive root for n if the order of a in $(\mathbf{Z}/n\mathbf{Z})^*$ is as large as possible. Let $N_a(x)$ denote the number of integers n in $[1, x]$ with primitive root a. In analogy with Artin's conjecture for primes, it is tempting to conjecture that if a does not lie in some exceptional set, yet to be determined, then there is a positive constant $B(a)$ with $N_a(x) \sim B(a)x$. However, the experience above with the normal value of $R(n)$, the number of primitive roots modulo n, shows that we might be wary of such a conjecture.

To gain some further insight, we might begin by first considering the kind of result on average that was relatively easy in the case of primes. Namely, what can be said about $\frac{1}{x} \sum_{1 \le a \le x} N_a(x)$? Now the problem is not as easy as before, but the same sort of trick works, namely reorganizing the sum, so that now we are summing over integers $n \le x$, and for each n we would like to know how many primitive roots it has in $[1, x]$. This estimate was worked out by the first author in [10], and sure enough, there is an oscillation. It is shown that

$$\liminf_{x\to\infty} \frac{1}{x^2} \sum_{1\le a\le x} N_a(x) = 0, \quad \limsup_{x\to\infty} \frac{1}{x^2} \sum_{1\le a\le x} N_a(x) > 0.$$

(The reason for the extra factor of $1/x$ is that it is natural to begin with the assumption that each term $N_a(x)$ is of order of magnitude x.)

Thus, while this result is not inconsistent with the assertion that $N_a(x) \sim B(a)x$, it certainly causes some serious doubt. In addition, the first author of this survey believes he may be able to generalize the Goldfeld–Stephens argument and achieve results like the ones above, but with the average taken over an interval of a-values of the form $[1, x^\epsilon]$.

Before proceeding, we note that there are certain numbers a for which we always have $N_a(x) = o(x)$. Namely if a is a nontrivial power, or if a is a square times -1 or a square times ± 2, then $N_a(x) = o(x)$. To get the idea of this, consider for example the case of $a = 2$, which is not exceptional at all when one considers prime moduli, but is exceptional for composite moduli. For all odd numbers n but for a set of density 0, the highest power of 2 which divides an order of an element in $(\mathbf{Z}/n\mathbf{Z})^*$, as before call it 2^{λ_2}, has $\lambda_2 \ge 3$. (That is, almost all numbers are divisible by a prime that is 1 (mod 8).) If $p|n$ where $p \equiv 1 \pmod{2^{\lambda_2}}$ (at least one such prime must divide n), then necessarily, since $p \equiv 1 \pmod 8$, we have that 2 is a quadratic residue modulo p. Thus, 2^{λ_2} cannot divide the order of 2 in $(\mathbf{Z}/n\mathbf{Z})^*$ and so 2 cannot be a primitive root modulo n. The number of exceptional numbers $n \le x$ where this argument is not valid is $O(x/(\ln x)^{1/4})$, which is $o(x)$ as claimed.

Let $\mathcal{E}$ denote the set of integers a such that either a is a nontrivial power, or a is a square times -1 or a square times ± 2. Thus, if $a \in \mathcal{E}$, then $N_a(x) = o(x)$. The set $\mathcal{E}$ should then stand as a candidate for the exceptional set in a generalization of Artin's conjecture for composite moduli.

But beyond this exceptional set, the first author in [9] was able to show that for *any* integer a, we have

$$\liminf_{x\to\infty} \frac{1}{x} N_a(x) = 0. \tag{1}$$

Moreover, this result was obtained on a set of real numbers x that is independent of the choice of a, in some sense. That is, there is an unbounded set S of positive reals such that for every integer a,

$$\lim_{x\to\infty,\, x\in S} \frac{1}{x} N_a(x) = 0.$$

So, we definitely do *not* have $N_a(x) \sim B(a)x$ for a positive number $B(a)$, not for any integer a.

With these thoughts in place, the first author in [7] made the conjecture that if a is a fixed integer not in $\mathcal{E}$, then there is a positive number $B(a)$ with

$$\limsup_{x\to\infty} \frac{1}{x} N_a(x) = B(a).$$

Recently, see [11], we have been able to prove this conjecture, under assumption of the GRH. In fact, we have been able to show that there is an unbounded set S' of positive reals and a positive constant c such that for each integer $a \notin \mathcal{E}$,

$$\limsup_{x\to\infty,\, x\in S'} \frac{1}{x} N_a(x) \geq c\frac{\varphi(|a|)}{|a|}. \tag{2}$$

One might ask about the weak Artin conjecture for composite moduli. Actually on this question, it is indeed possible to unconditionally prove that there are infinitely many n with primitive root a for many values of a. For example, take $a = 2$. We have that 2 is a primitive root for all of the numbers 3^j. In general, if a is a primitive root for p^2, where p is an odd prime, then a is a primitive root for p^j for every j. Other examples: Any number $a \equiv \pm 3 \pmod 8$ is a primitive root for all of the numbers 2^j. What is still unsolved, and may be tractable without the GRH: Given an integer a that is not a square nor -1, are there infinitely many squarefree integers n with primitive root a?

Local densities and a stronger conjecture

Let us first consider an easier question. Given a fixed prime q and a fixed integer $a \notin \mathcal{E}$, what is the distribution of the set of natural numbers n coprime to a such that the power of q in the order of a in the group $(\mathbf{Z}/n\mathbf{Z})^*$ is as large as possible over all elements in the group? Say the number of such integers $n \leq x$ is $N_a^q(x)$. This problem, see [11], can be analyzed unconditionally, giving

$$N_a^q(x) = (1 + o(1))\frac{\varphi(|a|)}{|a|} x\left(1 - F_q(x)\right),$$

where

$$F_q(x) \;=\; \sum_{j=0}^{\infty}\left(\exp\Big(-\Big(\frac{1}{\varphi(q^j)} - \frac{1}{q^{j+1}}\Big)\ln\ln x\Big) - \exp\Big(-\frac{1}{\varphi(q^j)}\ln\ln x\Big)\right).$$

As with the coin-flip problem, the density $1 - F_q(x)$ does not tend to a limit as $x \to \infty$. It is possible to show that

$$\liminf_{x\to\infty} F_q(x) \sim \frac{\ln q}{q^2}, \quad \limsup_{x\to\infty} F_q(x) \sim \frac{1}{eq},$$

as $q \to \infty$. By choosing a sequence of x-values where $F_q(x) \ge c/q$ occurs for many small primes q, it is possible to prove (1). It is also possible to choose a sequence of x-values where $F_q(x) = O(1/q \ln q)$ for most small primes q, but this is not sufficient for (2), since larger primes q can spoil the result. To show that larger primes usually do not pose too great an influence, the GRH comes into play.

Let

$$F_q \;=\; \liminf_{x\to\infty} F_q(x) \;=\; \inf_{t>0} \sum_{j=-\infty}^{\infty} \frac{\exp(t/q^{j+1}) - 1}{\exp(t/(q^j - q^{j-1}))}. \tag{3}$$

(Notice that the function of t is invariant under $t \mapsto tq$.) It seems reasonable to conjecture that the upper density $B(a)$ may be taken as $\alpha\varphi(|a|)/|a|$, where

$$\alpha := \prod_q (1 - F_q) \approx 0.326,$$

a conjecture made in [11]. That is, it is conjectured that for every integer a not in $\mathcal{E}$,

$$\limsup_{x\to\infty} \frac{1}{x} N_a(x) = \alpha \frac{\varphi(|a|)}{|a|}, \tag{4}$$

and that this limsup is attained on a set of positive reals independent of the choice of a. Note that for the number c in (2), we do have $c \le \alpha$; in fact this is unconditional. That is, for every integer $a \ne 0$,

$$\limsup_{x\to\infty} \frac{1}{x} N_a(x) \le \alpha \frac{\varphi(|a|)}{|a|}.$$

Let t_q be a value of t in $[1, q)$ where the infimum in (3) occurs. Then $t_q = \ln q + \ln\ln q + o(1)$ as $q \to \infty$. If $x \to \infty$ in such a way that the fractional part of $\ln\ln\ln x/\ln q$ tends to the fractional part of $\ln t_q/\ln q$, then $F_q(x) \to F_q$. Part of the problem in showing the conjecture (4) is to show that there is an unbounded sequence of values of x such that simultaneously, for all small primes q, the fractional part of $\ln\ln\ln x/\ln q$ approaches the fractional part of $\ln t_q/\ln q$. That such a sequence of x-values exists follows from Schanuel's conjecture in transcendental number theory. Indeed, from this conjecture, it follows that if $q_1, \ldots, q_k$ are distinct primes, then the real numbers $\ln q_1, \ldots, \ln q_k$ are algebraically independent. It would follow that the real numbers $1/\ln q_1, \ldots, 1/\ln q_k$ are linearly independent over the rationals, allowing simultaneous diophantine approximation of the quantities $\ln\ln\ln x/\ln q_1, \ldots, \ln\ln\ln x/\ln q_k$ modulo 1. However, even with Schanuel's conjecture and the GRH, there still seems to be some difficulties with the stronger conjecture.

Perhaps somewhat more tractable may be the conjecture from [9] that for a fixed integer a_0 not in $\mathcal{E}$, the individual count $N_{a_0}(x)$ is asymptotically equal to the average count over all integers a in $[1,x]$. That is, as $x \to \infty$,

$$N_{a_0}(x) = (1+o(1))\frac{1}{x}\sum_{1\le a\le x} N_a(x).$$

We close with another conjecture that is perhaps tractable:

$$\limsup_{x\to\infty}\frac{1}{x^2}\sum_{1\le a\le x} N_a(x) = \frac{6\alpha}{\pi^2}.$$

References

[1] J. S. Athreya and L. M. Fidkowski, *Number theory, balls in boxes, and the asymptotic uniqueness of maximal discrete order statistics*, Integers—The Electronic Journal of Combinatorial Number Theory (`http://www.integers-ejcnt.org`) **0** (2000), article A3.

[2] R. Gupta and M. R. Murty, *A remark on Artin's conjecture*, Invent. Math. **78** (1984), 127–130.

[3] D. R. Heath-Brown, *Artin's conjecture for primitive roots*, Quart. J. Math. Oxford Ser. (2) **37** (1986), 27–38.

[4] C. Hooley, *On Artin's conjecture*, J. reine angew. Math. **225** (1967), 209–220.

[5] I. Kátai, *On distribution of arithmetical functions on the set prime plus one*, Compositio Math. **19** (1968), 278–289.

[6] P. Kirschenhofer and H. Prodinger, *The number of winners in a discrete geometrically distributed random sample*, Ann. Appl. Probab. **6** (1996), 687–694; *Addendum*, ibid. **8** (1998), 647.

[7] S. Li, *On Artin's conjecture for composite moduli*, Ph. D. dissertation, University of Georgia, 1998.

[8] S. Li, *On the number of elements with maximal order in the multiplicative group modulo n*, Acta Arith. **86** (1998), 113–132.

[9] S. Li, *On extending Artin's conjecture to composite moduli*, Mathematika **46** (1999), 373–390.

[10] S. Li, *Artin's conjecture on average for composite moduli*, J. Number Theory **84** (2000), 93–118.

[11] S. Li and C. Pomerance, *On generalizing Artin's conjecture on primitive roots to composite moduli*, J. reine angew. Math., to appear.

[12] M. R. Murty, *Artin's conjecture for primitive roots*, Math. Intelligencer **10** (1988), no. 4, 59–67.

[13] M. Rubinstein and P. Sarnak, *Chebyshev's bias*, Experiment. Math. **3** (1994), 173–197.

[14] P. J. Stephens, *An average result for Artin's conjecture*, Mathematika **16** (1969), 178–188.