ON CANDIDATES TO BE PRIMITIVE ROOTS

ANTONELLA PERUCCA AND CARLO SANNA

Abstract. The design of efficient algorithms to find a primitive root modulo a prime number p is a classic important topic in computational number theory. Many algorithms proceed by repeatedly selecting a "candidate" integer g and checking whether g is a primitive root modulo p or not, until a primitive root is found. This essay provides some results and considerations on how to select candidates g that have a high probability of being primitive roots modulo p, under the assumption that p is a random prime number. These methods can easily improve the performance of current software implementations.

1. Introduction

The design of efficient algorithms to find a primitive root modulo a given prime number is a classic topic in computational number theory. It goes back at least to Gauss and is still an active area of research [3, 7]. A basic result says that an integer g is a primitive root modulo a prime number p if and only if

$$g^{(p-1)/q} \not\equiv 1 \pmod{p} \tag{1}$$

for every prime factor q of $p-1$. The core idea of many algorithms is selecting (randomly or deterministically) a "candidate" g and using (1) to check whether g is a primitive root modulo p or not. In case of a negative answer, a new candidate is selected and the procedure is repeated, until a primitive root is found. Depending on the specific design of the algorithm, a theoretical argument guarantees the success after a certain number of attempts, thus providing an estimate of the complexity of the algorithm. For instance, under the Generalized Riemann Hypothesis (GRH), Shoup [12, Theorem 3] proved that the least primitive root modulo p is $O\big((\log p)^6\big)$. Hence, simply trying $g = 2, 3, 4, \dots$ guarantees termination after at most $O\big((\log p)^6\big)$ attempts. This (plus some optimizations) is the approach taken by the software libraries PARI/GP [9] and SymPy [15].

This essay concerns the choice of candidate integers g that have a high probability of being primitive roots modulo p, under the assumption that p is a random prime number: Section 2 provides the basic notation and conventions; Section 3 recalls the Artin conjecture and provides the best candidate g without assumptions on p; Section 4 concerns a strategy that chooses g depending on the residue class of p modulo a fixed integer f; Section 5 focuses on the sign of g and on the primitive root modulo p having the least absolute value; Section 6 explores alternative choices for primitive roots; finally, Section 7 collects the proofs of all the results.

Acknowledgments. The authors thank Fritz Hörmann for providing C code that improved the speed of some of the computations reported in the paper. C. Sanna is a member of GNSAGA of INdAM and of CrypTO, the group of Cryptography and Number Theory of Politecnico di Torino.

2. Notation and conventions

For integers a and b, let "$a \mid b$" mean "a divides b", and let (a, b) denote the greatest common divisor of a and b. The letters p and q are reserved for prime numbers. Given a finite set $\mathcal{S}$, let $|\mathcal{S}|$ denote the cardinality of $\mathcal{S}$. The *density* of a set of prime numbers $\mathcal{P}$ is the limit

$$\lim_{x \to +\infty} \frac{\big|\{p \leq x : p \in \mathcal{P}\}\big|}{\big|\{p \leq x\}\big|}$$

whenever this exists. Hereafter, with a slight abuse of language, a property P holds for a "random prime number p with probability δ" if the set of prime numbers satisfying P has density δ. (Note that this is not truly a probability, since the density does not satisfy countable additivity.) Other notation is introduced when first needed.

3. Artin's conjecture and the best candidate

Let g be an integer that is neither 0, -1, or a square. In fact, considering that number smaller in absolute value are preferred, and that a power of g is a primitive root only if g is a primitive root, assume that g is not a perfect power.

The probability that g is a primitive root modulo a random prime number p is given by the Artin conjecture on primitive roots, which Hooley [4] proved under the GRH. Hereafter, all results are implicitly assuming the GRH. The *Artin constant* is

$$A := \prod_q \left(1 - \frac{1}{q(q-1)}\right) = 0.37395581...$$

The result of Hooley is the following.

Theorem 1. *Let g be an integer that is neither 0, -1, nor a perfect power; and let Δ be the discriminant of $\mathbb{Q}(\sqrt{g})$. Then, for a random prime number p, the probability that g is a primitive root modulo p is equal to*

$$A \cdot \begin{cases} 1 & \text{if } \Delta \not\equiv 1 \bmod 4, \\ 1 - \mu(|\Delta|) \prod_{q \mid \Delta} \dfrac{1}{q^2 - q - 1} & \text{otherwise}, \end{cases}$$

where μ denotes the Möbius function.

By inspecting the formula of Theorem 1, the integer g that maximizes the probability is -3, and for $g = -3$ the probability is $\frac{6}{5}A \approx 45\%$. However, if the condition of being a non-quadratic residue is ignored (that is, test (1) ignores $q = 2$) then the differences between the candidates for being primitive roots disappear and the probability is always $2A \approx 75\%$ (see Theorem 15 by taking $F = \mathbb{Q}$ and letting S be the set of the odd primes, and see also Theorem 5).

4. Selecting a candidate by the residue class of p modulo f

Let f be a positive integer. In general, prime numbers p such that g is a primitive root modulo p are *not* uniformly distributed modulo f. For instance, by the quadratic reciprocity law, if $p \equiv \pm 1 \pmod 8$ then 2 is a square modulo p; and consequently 2 cannot be a primitive root modulo p. For every integer a such that $1 \leq a \leq f$ and $(a, f) = 1$, let $\mathcal{P}_{a,f,g}$ be the set of prime numbers p such that $p \equiv a \pmod f$ and g is a primitive root modulo g. Under GRH, Moree [8] proved

TABLE 1. Probability P_k that the candidate g selected by the modulo $2p_1 \cdots p_k$ is a primitive root.

k	P_k/A	P_k	k	P_k/A	P_k
1	1	0.373955...	4	1522/779	0.730629...
2	9/5	0.673120...	5	167878/84911	0.739350...
3	181/95	0.712484...	6	26172722/13161205	0.743658...

TABLE 2. Choice of g depending on $p \bmod 2^2 \cdot 3 \cdot 5 \cdot 7$.

$p \bmod 4$	$p \bmod 3$	$p \bmod 5$	$p \bmod 7$	g
1	1	1 or 4	1, 2, or 4	2
1	1	1 or 4	3, 5, or 6	7
1	1	2 or 3	any	5
1	2	any	any	3
3	1	any	any	3
3	2	any	any	-3

an explicit formula for the density of $\mathcal{P}_{a,f,g}$ (see Theorem 13 below). The formula shows that, for fixed a and f, the density $\mathcal{P}_{a,f,g}$ can be higher or lower depending on the choice of g. An extreme example is the following: if $p \equiv 2 \bmod 3$, then the conditional probability of -3 being a primitive root is $\frac{12}{5}A \approx 90\%$.

This suggests the following procedure for selecting a good candidate g to be tested as a primitive root modulo p. As a precomputation phase, fix a modulo f and for each integer a such that $1 \leq a \leq f$ and $(a, f) = 1$, compute an integer $g = g(a, f)$ that is neither $0, \pm 1$ nor a perfect power and for which the density of $\mathcal{P}_{a,f,g}$ is maximal. Then, when given a random prime number p, select as a candidate $g(a, f)$, where a is determined by $p \equiv a \pmod f$. The following result provides an upper bound for the probability of success of this procedure.

Theorem 2. *For a random prime number p, the probability that the candidate g selected by the above procedure is a primitive root modulo p is at most $2A = 0.747911...$*

It seems likely that for the modulo $f_k = 2p_1p_2 \cdots p_k$, where $p_1, \ldots, p_k$ are the first prime numbers, the probability of Theorem 2 approaches $2A$ as $k \to +\infty$, see the values in Table 1. However, employing a large modulo is clearly impractical. The next result provides a practical instantiation of the modulo, which gives a probability close to the upper bound of Theorem 2.

Proposition 3. *For a random prime number p, select $g = g(p)$ according to Table 2 (which corresponds to the choice of the modulo $f = 2^2 \cdot 3 \cdot 5 \cdot 7$). Then the probability that g is a primitive root modulo p is $\frac{1522}{779}A = 0.730629...$ (cf. Table 1).*

The following proposition focuses on the case of prime numbers $p \equiv 3 \pmod 4$ and shows that the candidates g maximizing the probability of being primitive roots modulo p have an easy characterization.

Proposition 4. *For a random prime number $p \equiv 3 \bmod 4$, the numbers g maximizing the probability of being a primitive root are those of the form $g = -t^2$ where*

t is a positive integer that is not a q-th power for any odd prime q. For them, the conditional probability is $2A$.

The last theorem of this section shows that, for a fixed modulo f, the differences for the probabilities that the various candidates g are primitive roots modulo $p \equiv a \pmod f$ are only due to the probability of being a non-quadratic residue and they would disappear if restricting to candidates that are non-quadratic residues modulo p.

Theorem 5. *Let a, f be integers such that $1 \le a \le f$, $f > 2$, and $(a, f) = 1$. Let g be and integer that is neither $0, -1$, nor a perfect power. For a random prime number $p \equiv a \bmod f$, the conditional probability that the index of $g \bmod p$ is a power of 2 (including 2^0) is independent of g and equal to*

$$2A \cdot \prod_{\substack{q \mid f \\ q>2}} \Big(1 - \frac{1}{q(q-1)}\Big)^{-1} \cdot \prod_{q \mid (a-1, f)} \Big(1 - \frac{1}{q}\Big).$$

5. The least primitive root in absolute value

5.1. **Theoretical results.** For a prime number $p \equiv 1 \bmod 4$, it is easy to show that an integer g is a primitive root modulo p if and only if the same holds for its absolute value $|g|$, so there is no need of considering negative numbers. This section focuses on prime numbers $p \equiv 3 \bmod 4$ and describes the advantage of considering also negative numbers that are small in absolute value as candidates for being primitive roots modulo p. Perucca and Tholl [10] already studied this idea of considering negative candidates.

Note that -1 is a non-quadratic residue modulo p. Let $g \neq 0, \pm 1$ be a rational number that is not a perfect power. Then g is a quadratic residue modulo p if and only if $-g$ is a non-quadratic residue modulo p. Thus, up to taking the correct sign, to ensure that g is a primitive root modulo p it suffices to ensure that its index modulo p has no odd prime factors. Observe that the condition that either g or $-g$ is a primitive root modulo p is equivalent to $-g^2$ being a primitive root (this number is surely a non-quadratic residue modulo p).

Proposition 6. *Let g be an integer that is neither 0, -1, nor a perfect power. For a random prime number $p \equiv 3 \pmod 4$, the conditional probability that g is a primitive root modulo p up to a sign is $2A \approx 75\%$.*

Non-zero rational numbers $g_1, \dots, g_r$ are *strongly multiplicatively independent* if, for $i = 1, \dots, r$, each representant of g_i modulo the subgroup of $\mathbb{Q}^\times$ generated by the g_j's for $j \neq i$ is not a perfect power.

Theorem 7. *Suppose that $g_1, \dots, g_r$ are non-zero rational numbers that are strongly multiplicative independent. Then, for a random prime number $p \equiv 3 \bmod 4$, the numbers $g_1, \dots, g_r$ are all primitive roots up to a sign with conditional probability*

$$C_r := \prod_{q \neq 2} \Big(1 - \frac{1 - (1 - \frac{1}{q})^r}{q - 1}\Big).$$

Corollary 8. *Let $(g_n)_{n \ge 1}$ be a sequence of distinct non-zero rational numbers such that any finite subset of its elements consists is strongly multiplicative independent.*

Then, for a random prime number $p \equiv 3 \bmod 4$, the expected value of the least index i such that g_i is a primitive root up to a sign is

$$1 + \sum_{k=1}^{\infty} \sum_{r=0}^{k} (-1)^r \binom{k}{r} C_r \,.$$

5.2. **Experimental data on the least primitive root up to a sign.** The following experimental data concerns the prime numbers $p \equiv 3 \bmod 4$ up to 10^9 and suggests that there is a considerable advantage in allowing negative primitive roots. The probabilities to have found a primitive root, respectively a primitive root up to a sign by testing all candidates with absolute value up to a given bound are collected in Table 3

TABLE 3. Probabilities of finding a primitive root vs a primitive root up to sign.

Bound	primitive root	primitive root up to a sign
2	37.396%	74.790%
3	58.241%	90.639%
5	74.616%	96.210%
6	80.209%	97.161%
7	87.813%	98.818%
10	89.274%	99.064%
11	93.382%	99.605%
12	93.953%	99.619%
13	96.257%	99.835%
14	96.826%	99.867%
15	97.216%	99.879%

Out of the 25 million primes tested, only for 31 thousands of them the least primitive root up to a sign is larger than 15, compared to the 708 thousands for which the least primitive root is larger than 15 (these figures are rounded).

5.3. **Experimental data on the sign of the least primitive root.** For each prime number $p \equiv 3 \pmod 4$, define

$$\mathrm{lpr}^+(p) := \min\{g > 1 : g \text{ is a primitive root modulo } p\},$$
$$\mathrm{lpr}^-(p) := -\max\{g < 1 : g \text{ is a primitive root modulo } p\}.$$

Note that $\mathrm{lpr}^+(p) \neq \mathrm{lpr}^-(p)$. From a computation involving prime numbers $p \equiv 3 \bmod 4$ up to 10^7, it seems likely that $\mathrm{lpr}^+(p) < \mathrm{lpr}^-(p)$ holds for about 53.4% of the primes, see Figure 1. Thus the least primitive root in absolute value seems to be more often positive than negative. This fact may seem counter-intuitive because there is a bias for positive residues being quadratic residues [1, Theorem 4 p.346], the bias being the class number h of the field $\mathbb{Q}(\sqrt{-p})$ (respectively, $3h$) for $p \equiv 7 \bmod 8$ (respectively, $p \equiv 3 \bmod 8$). This class number h grows asymptotically like $\sqrt{p}$, as proven in [14]. However, as it is very seldom that the least primitive root in absolute value is large, it may be that by considering only the first, say, 20 candidates it is possible to establish that $\mathrm{lpr}^+(p) < \mathrm{lpr}^-(p)$ more often then not.

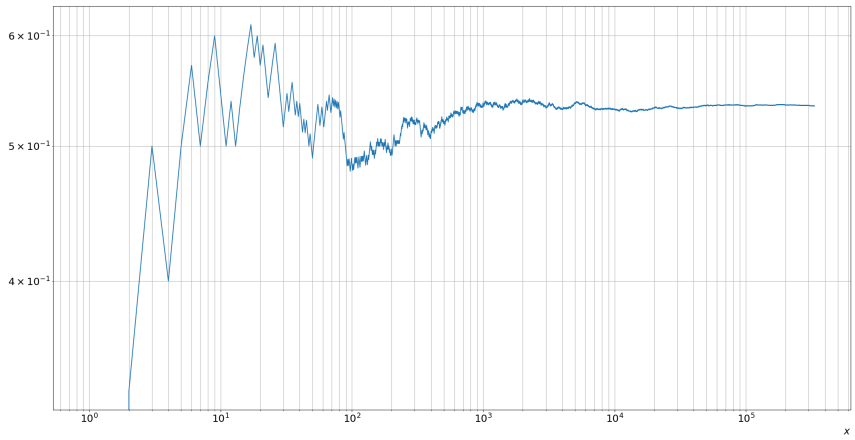

FIGURE 1. The proportion of the prime numbers $p \leq x$, $p \equiv 3 \pmod 4$ such that $\mathrm{lpr}^+(p) < \mathrm{lpr}^-(p)$.

For $p \equiv 3 \bmod 4$, the conditional probability that -3 (respectively, 3) is a primitive root is $\frac{6}{5}A \approx 45\%$ (respectively, $\frac{2}{3} \cdot \frac{6}{5}A = \frac{4}{5}A \approx 30\%$) by Theorem 15. The next proposition explains the following apparently counter-intuitive fact: it is more likely for 3 than for -3 to be the least primitive root in absolute value (for primes $p \equiv 3 \bmod 4$ up to 10^9, the probabilities are respectively 11.8% and 4.1%).

Proposition 9. *The ratio between the probability for -3 to be a primitive root or a primitive root up to a sign is*

$$\frac{\frac{6}{5}A - \frac{9}{13}C_2}{2A - C_2} \approx \frac{1}{4}.$$

A reasoning similarly to the proof of Proposition 9 could give in general a bias towards $\pm q$ being the least primitive root in absolute value, where q is an odd prime and the sign is chosen so that $\sqrt{\pm q} \notin \mathbb{Q}(\zeta_q)$ (hereafter, ζ_n denotes a primitive n-th root of unity). Indeed, such a bias can be observed for $3, -5, 7, 11, -13$ considering the primes p up to 10^9.

6. Beyond testing candidates for primitive roots

This section concerns some alternative candidates for being primitive roots with respect to the least primitive root (or the least primitive root in absolute value).

6.1. Safe primes and generalizations thereof.

Safe primes and their natural generalization are a specific class of prime numbers for which the probabilities coming from Artin's conjecture would be misleading.

Let p be a prime number, and suppose that $p - 1 = 2^t q$, where t is a small positive integer and q is a prime number. If $t = 1$ then p is called a *safe prime*. Then an integer $a \neq 0, \pm 1$ is a primitive root modulo p if and only if it is a non-quadratic residue modulo p and it is not a q-th power modulo p. Suppose that the former condition is satisfied. Then the latter condition is satisfied if and only if $a^{2^t} \not\equiv 1 \bmod p$ (this holds in particular if $a \leq \frac{1}{t} \log p / \log 2$).

If $t = 1$ then -1 is not a quadratic residue modulo p, hence either a or $-a$ is a non-quadratic residue. Thus, for $p \geq 5$ a safe prime, either 2 or -2 is a primitive

root, and in any case -4 is a primitive root. If $t > 1$ then $p \equiv 1 \bmod 4$. If $t = 2$ then $p \not\equiv 1 \bmod 8$ hence 2 is a non-quadratic residue modulo p. In fact 2 is a primitive root because $p \nmid (2^4 - 1)$ (see [2, Corollary 3.1]).

Now suppose that $t \geq 3$. Then 2 is a quadratic residue because $p \equiv 1 \bmod 8$. If $q \neq 3$ then 3 is a non-quadratic residue modulo p (see [2, Lemma 3]). Hence it is most likely that 3 is a primitive root modulo p (and this is surely the case if $p > 3^{2^t}$). If $t = 3$ and $r \neq 5$, or $t = 4$, then 3 is a primitive root (see [2, Corollary 4.1 and 4.2]).

6.2. **A suitable modification of** 2. There is a primitive root for which exponentiation basically amounts to the exponentiation of 2.

Consider a prime number p such that $p - 1 = sQ$, where Q is a power of a prime number q that is large compared to s. It is possible to determine the smallest positive integer m that is not a q-th power modulo p, and to construct the primitive root $\tilde{r} := m\eta$, where $\eta \bmod p$ is a root of unity of order dividing s.

If t is a prime divisor of s, the order of said root of unity can be taken coprime to t if m is not a t-th power and must be taken divisible by $t^{v_t(s)}$ otherwise. For practical purposes, it may also be possible not to compute the needed root of unity, but just work with m that is not a q-th power. Indeed, by the Chinese remainder theorem, the multiplicative group modulo p is the direct product of a group of order q, for which a generator is known, and a small group of order s which could be easy to handle even without an explicit generator.

The probability that a random prime number congruent to 1 mod q splits in the field $\mathbb{Q}(\zeta_q, \sqrt[q]{2})$ is $1/q$, hence probability that $m = 2$ can be estimated as $1 - 1/q$. Moreover, if $p > 2^s$ then surely $m = 2$.

Example 10. *If $s = 6$ then for $0 < g < \sqrt[6]{p}$ it is only needed to ensure that g is neither a quadratic residue nor a cubic residue to make sure that g is a primitive root modulo p. If g is a cube modulo p, then g can be multiplied by a root of unity of order 3 (and, if it is a square, by -1) to obtain a primitive root modulo p. So for any $p > 64$ of the above form, a primitive root is given by an element of the form $\pm 2\eta$, where $\eta \bmod p$ has order dividing 3.*

6.3. **Primitive roots with small prime divisors.** Another possibility is finding a, possibly rational, primitive root that is the product of small prime numbers (with integer exponents).

Different candidates for being primitive roots can be tested in parallel: beyond increasing the speed of finding a primitive root (respectively, the least positive primitive root) this feature could be exploited to construct a primitive root that consists of a product of powers of small prime numbers.

Consider the prime factors $q_1, \ldots, q_f$ of $p - 1$. For each tested candidate a_j, record the primes q_i for which a_j is not a q_i-th power. Then, by comparing these information, build a primitive root a that is of the form $a = \prod_j a_j^{e_j}$ where the exponents e_j are, if possible, 0, and otherwise small (or small in absolute value, if rational numbers as primitive roots are allowed). If $p - 1 = sQ$, where Q is a power of a large prime q not dividing s, then most likely it suffices to take $e_j \neq 0$ to ensure that a is not a q_i-th power for the prime divisors of s because a will most likely be not a q-th power.

Testing n pairwise coprime candidates $a_1, \ldots, a_n$ that are no perfect powers, the probability that all of them are q_i-th powers for one fixed i (without additional

information which could create a bias) is merely $\frac{1}{q_i^n}$. So it is expected that with very few candidates a_j's one may produce a primitive root as described. Notice that this procedure is not determining roots of unity modulo p whose orders are the largest possible power of the primes q_i's, as done for example in the efficient probabilistic algorithm in the book by Shoup [13, Section 11.1].

Remark 11. *Restricting to the primes $p \equiv 1 \bmod 4$ up to 10^9, considering the products (possibly with repeated factors) of previously tested prime candidates is more promising than testing the next prime because of the following conditional probabilities for being a primitive root:*

2 *or* 3 *or* 5	*73.3%*
a product of 2,3,5	*85.9%*
2 *or* 3 *or* 5 *or* 7	*81.2%*
a product of 2,3,5,7	*93.2%*
2 *or* 3 *or* 5 *or* 7 *or* 11	*86.6%*
a product of 2,3,5,7,11	*96.7%*.

Remark 12. *Restrict to the primes $p \equiv 3 \bmod 4$: by possibly changing the sign, the candidates can be assumed to be non-quadratic residues. For the primes up to 10^9, considering the products (possibly with repeated factors) of previously tested prime candidates is less promising than testing the next prime because of the following conditional probabilities for being a primitive root:*

± 2 *or* ± 3 *or* ± 5	*96.2%*
$\pm$ *a product of* 2,3,5	*97.9%*
± 2 *or* ± 3 *or* ± 5 *or* ± 7	*98.4%*
$\pm$ *a product of* 2,3,5,7	*99.3%*
± 2 *or* ± 3 *or* ± 5 *or* ± 7 *or* ± 11	*99.3%*
$\pm$ *a product of* 2,3,5,7,11	*99.8%*.

7. Proofs

The proofs require some preliminary results. The first is the aforementioned formula for the density of $\mathcal{P}_{a,g,f}$ due to Moree [8, Theorem 1.2 with $h = 1$].

Theorem 13. *Let g be an integer that is neither equal to 0, -1, nor a perfect power, and let Δ denote the discriminant of the quadratic field $\mathbb{Q}(\sqrt{g})$. Let a and f be relatively prime integers such that $1 \le a \le f$. Set $b := \Delta/(f, \Delta)$,*

$$\gamma := \begin{cases} (-1)^{(b-1)/2}(f, \Delta) & \text{if } b \text{ is odd;} \\ 1 & \text{otherwise;} \end{cases}$$

and

$$A(a, f) := \prod_{q \mid (a-1, f)} \left(1 - \frac{1}{q}\right) \prod_{q \mid f} \left(1 - \frac{1}{q(q-1)}\right)^{-1} A.$$

Then the density of $\mathcal{P}_{a,f,g}$ is equal to

$$\frac{A(a, f)}{\varphi(f)} \left(1 + \left(\frac{\gamma}{a}\right) \frac{\mu(2|b|)}{\prod_{q \mid b}(q^2 - q - 1)}\right),$$

where φ is the Euler totient function and $\left(\frac{\cdot}{\cdot}\right)$ denotes the Kronecker symbol.

The next result is due to Kesava Menon [6, Theorem 1] (cf. the survey by Tóth [16, Theorem 4]).

Theorem 14. *Let ϑ be a multiplicative arithmetic function and define the arithmetic function ϑ^* by*

$$\vartheta^*(n) := \sum_{\substack{1\le a\le n\\(a,n)=1}} \vartheta\big((a-1,n)\big)$$

for every positive integer n. Then ϑ^* is a multiplicative arithmetic function satisfying

$$\vartheta^*(p^v) = -p^{v-1} + \sum_{j=0}^{v} \varphi(p^{v-j})\vartheta(p^j)$$

for every prime p and positive integer v.

7.1. **Proof of Theorem 2.** From Theorem 13, the probability is asymptotically at most

(2)
$$2\sum_{\substack{1\le a\le f\\(a,f)=1}} \frac{A(a,f)}{\varphi(f)} = \frac{2A}{f}\prod_{p\mid f}\left(1-\frac{1}{p}\right)^{-1}\left(1-\frac{1}{p(p-1)}\right)^{-1}\sum_{\substack{1\le a\le f\\(a,f)=1}}\prod_{p\mid(a-1,f)}\left(1-\frac{1}{p}\right)$$
$$= \frac{2A}{f}\prod_{p\mid f}\left(1-\frac{1}{p}-\frac{1}{p^2}\right)^{-1}\sum_{\substack{1\le a\le f\\(a,f)=1}}\prod_{p\mid(a-1,f)}\left(1-\frac{1}{p}\right).$$

Let ϑ be the arithmetic function defined by $\vartheta(n) := \varphi(n)/n$ for every positive integer n. Then

$$-p^{v-1} + \sum_{j=0}^{v}\varphi(p^{v-j})\vartheta(p^j) = -p^{v-1} + \varphi(p^v) + \vartheta(p^v) + \sum_{j=1}^{v-1}\varphi(p^{v-j})\vartheta(p^j)$$
$$= -p^{v-1} + \varphi(p^v) + \vartheta(p^v) + \sum_{j=1}^{v-1} p^{v-j}\left(1-\frac{1}{p}\right)^2$$
$$= -p^{v-1} + (p^v+1)\left(1-\frac{1}{p}\right) + \frac{p^v-p}{p-1}\left(1-\frac{1}{p}\right)^2$$
$$= p^v\left(1-\frac{1}{p}-\frac{1}{p^2}\right),$$

for every positive integer v. Hence, Theorem 14 implies that

$$\sum_{\substack{1\le a\le f\\(a,f)=1}}\prod_{p\mid(a-1,f)}\left(1-\frac{1}{p}\right) = f\prod_{p\mid f}\left(1-\frac{1}{p}-\frac{1}{p^2}\right), \tag{3}$$

for every positive integer f. Combining (2) and (3) yields that the probability is at most $2A$, as claimed. □

7.2. **Proof of Proposition 3.** The claim follows by applying Theorem 13 to compute the sum of the densities of the sets $\mathcal{P}_{a,f,g}$, where $f = 2^2\cdot 3\cdot 5\cdot 7$, $1\le a\le f$ with $(a,f)=1$, and g is selected according to Table 2. Note that these sets $\mathcal{P}_{a,f,g}$ are pairwise disjoint because of the condition $p\equiv a \pmod f$. Hence the density of their union is the sum of their densities. □

7.3. **Interlude.** The next proofs require a result that is a consequence of the Master theorem on Artin type problems [5, Theorem 4.1], proven by Järviniemi and Perucca assuming GRH. Observe that the fields $\mathbb{Q}(\zeta_q, \sqrt[q]{g})$ are linearly disjoint over $\mathbb{Q}$ for $q \neq 2$ by Schinzel's Theorem on abelian radical extensions [11, Theorem 2] (and they have degree $q(q-1)$ because g is not a perfect power).

Theorem 15. *Let $F/\mathbb{Q}$ be a finite Galois extension, let C be a non-empty conjugacy stable subset of $\mathrm{Gal}(F/\mathbb{Q})$, and consider a set S of odd primes such that F is linearly disjoint from the compositum of the fields $\mathbb{Q}(\zeta_q, \sqrt[q]{g})$ for $q \in S$. Then the density of the primes p such that the index of $(g \bmod p)$ is coprime to the primes in S and whose Frobenius conjugacy class in $F/\mathbb{Q}$ is contained in C is well-defined and equals*

$$A \cdot \frac{|C|}{|\mathrm{Gal}(F/\mathbb{Q})|} \cdot \prod_{q \notin S} \left(1 - \frac{1}{q(q-1)}\right)^{-1}.$$

Remark 16. *In Theorem 15, if F contains $\mathbb{Q}(\zeta_q, \sqrt[q]{g})$ for some odd prime $q \notin S$, then it is possible to impose that the index of $g \bmod p$ is coprime to q by requiring that C consists of automorphisms that are not the identity on that subfield. For $q = 2$, it is also needed the condition that the primes $q \notin S$ include the prime divisors of Δ, to ensure the independence with the coprimality conditions for $q \in S$.*

7.4. **Proof of Proposition 4.** The field $\mathbb{Q}(\zeta_4)$ is linearly disjoint from the compositum of the fields $\mathbb{Q}(\zeta_q, \sqrt[q]{z})$ for q odd, for any rational number z. Supposing that $z \neq 0, \pm 1$ is not a perfect power, the probability that $z \bmod p$ has an index without odd prime factors is $2A$ by Theorem 15. Considering that -1 is not a square modulo $p \equiv 3 \bmod 4$, the numbers $-t^2$ are non-quadratic residues, but that cannot be said with certainty for the other numbers.

7.5. **Proof of Theorem 5.** The claim follows from an application of Theorem 15, considering that the arithmetic progression condition concerns the Frobenius in the field $F := \mathbb{Q}(\zeta_f)$ and letting S be the set of the odd primes that do not divide f. □

7.6. **Proof of Proposition 6.** Apply Theorem 15, taking $F = \mathbb{Q}(\zeta_4)$ to account for the condition $p \equiv 3 \bmod 4$ and letting S be the set of all odd prime numbers. □

7.7. **Proof of Theorem 7.** Assume GRH and apply [5, Theorem 4.1] with the Frobenius condition at $\mathbb{Q}(\zeta_4)$ to account for $p \equiv 3 \bmod 4$. For q odd, the extensions $L_q := \mathbb{Q}(\zeta_q, \sqrt[q]{g_1}, \dots, \sqrt[q]{g_r})$ are linearly disjoint over $\mathbb{Q}$ (among themselves and from $\mathbb{Q}(\zeta_4)$). Thus the requested probability is the product of the probabilities requiring the coprimality of the index with one fixed odd prime q. The coprimality condition of the index with an odd prime q is the proportion of the automorphisms in $\mathrm{Gal}(L_q/\mathbb{Q})$ that are not the identity on any of the subfields $\mathbb{Q}(\zeta_q, \sqrt[q]{g_i})$'s. Observe that

$$[L_q : \mathbb{Q}] = q^r(q-1).$$

The suitable automorphisms are those $q^r(q-2)$ that are not the identity on $\mathbb{Q}(\zeta_q)$ and those $(q-1)^r$ that are the identity on $\mathbb{Q}(\zeta_q)$ but do not fix any of the $\sqrt[q]{g_i}$'s. The requested conditional probability is then

$$\frac{(q-2)\cdot q^r + (q-1)^r}{(q-1)q^r} = 1 - \frac{1-(1-\frac{1}{q})^r}{q-1},$$

as desired. □

7.8. **Proof of Corollary 8.** It suffices to apply the following basic probability lemma by setting $P_r = C_r$.

Lemma 17. *Let $E_1, E_2, \ldots$ be some events and let t be the minimum positive integer such that E_t occurs ($t := +\infty$ if none of the events occurs). The expected value of t is equal to*

$$\mathbb{E}[t] = 1 + \sum_{k=1}^{\infty} \mathbb{P}\left[\bigwedge_{i=1}^{k} \overline{E_i}\right]$$

Furthermore, if for every set $\mathcal{S}$ of r positive integers the probability $P_r := \mathbb{P}\big[\bigwedge_{i\in\mathcal{S}} E_i\big]$ depends only on r, then

$$\mathbb{E}[t] = 1 + \sum_{k=1}^{\infty} \sum_{r=0}^{k} (-1)^r \binom{k}{r} P_r$$

Proof. Put $F_k := \bigwedge_{i=1}^{k} \overline{E_i}$, with the convention that $\mathbb{P}[F_0] := 1$. By the definitions of expected value and t,

$$\begin{aligned}
\mathbb{E}[t] &= \sum_{k=1}^{\infty} k\, \mathbb{P}[F_{k-1} \wedge E_k] \\
&= \sum_{k=1}^{\infty} k\, (\mathbb{P}[F_{k-1}] - \mathbb{P}[F_k]) \\
&= 1 + \sum_{k=1}^{\infty} ((k+1) - k)\, \mathbb{P}[F_k] \\
&= 1 + \sum_{k=1}^{\infty} \mathbb{P}[F_k],
\end{aligned}$$

as claimed. Under the second hypothesis, the inclusion-exclusion principle implies that

$$\begin{aligned}
\mathbb{P}[F_k] &= \sum_{\mathcal{S}\subseteq\{1,\ldots,k\}} (-1)^{|\mathcal{S}|} \mathbb{P}\left[\bigwedge_{i\in\mathcal{S}} E_i\right] \\
&= \sum_{r=0}^{k} (-1)^r \binom{k}{r} P_r,
\end{aligned}$$

and so

$$\mathbb{E}[t] = 1 + \sum_{k=1}^{\infty} \sum_{r=0}^{k} (-1)^r \binom{k}{r} P_r,$$

as claimed. □

7.9. **Proof of Proposition 9.** By Theorem 7, the probability that 3 or -3 is a primitive root is $2A$, and the probability that additionally ± 2 are not primitive roots is $2A - C_2 \sim 16\%$. Now suppose that -3 is a non-quadratic residue (which happens with probability $\frac{1}{2}$). In this case, the only difference is made by the prime factors $q > 3$ for the index of ($-3 \bmod p$) and of ($\pm 2 \bmod p$). By a straightforward modification of Theorem 7 (namely, removing the factor at $q = 3$), the conditional

probability that -3 is a primitive root is $\frac{12}{5}A$, and the conditional probability that additionally ± 2 are not primitive roots is $\frac{12}{5}A - \frac{18}{13}C_2 \sim 8\%$. □

References

1. Zenon I. Borevich and Igor R. Shafarevich, *Number theory*, Pure and Applied Mathematics, vol. 20, Academic Press, New York and London, 1966.
2. Andreas Ecker, *On primitive roots*, Elem. Math. **37** (1982), 103–108.
3. Ofer Grossman, *Finding primitive roots pseudo-deterministically*, Electronic Colloquium on Computational Complexity (2015), TR15–207.
4. Christopher Hooley, *On Artin's conjecture*, J. Reine Angew. Math. **225** (1967), 209–220.
5. Olli Järviniemi and Antonella Perucca, *Unified treatment of Artin-type problems*, Res. Number Theory **9** (2023), no. 1, Paper No. 10, 20.
6. P. Kesava Menon, *On the sum* $\sum (a-1, n)$, $[(a, n) = 1]$, J. Indian Math. Soc. (N.S.) **29** (1965), 155–163.
7. Kevin J. McGown and Jonathan P. Sorenson, *Computation of the least primitive root*, Math. Comp. **94** (2025), no. 352, 909–917.
8. Pieter Moree, *On primes in arithmetic progression having a prescribed primitive root. II*, Funct. Approx. Comment. Math. **39** (2008), 133–144.
9. The PARI Group, Univ. Bordeaux, *PARI/GP v. 2.17.3*, 2025, available from http://pari.math.u-bordeaux.fr.
10. Antonella Perucca and Mia Tholl, *Computing primitive roots according to Artin's conjecture*, JP Journal of Algebra, Number Theory and Applications **63** (2024), no. 5, 435–445.
11. Andrzej Schinzel, *Abelian binomials, power residues and exponential congruences*, Acta Arith. **32** (1977), 245–274.
12. Victor Shoup, *Searching for primitive roots in finite fields*, Math. Comp. **58** (1992), no. 197, 369–380.
13. ———, *A computational introduction to number theory and algebra*, second ed., Cambridge University Press, Cambridge, 2009.
14. Carl Ludwig Siegel, *Über die Classenzahl quadratischer Zahlkörper*, Acta Arith. **1** (1935), 83–86.
15. SymPy Development Team, *SymPy v. 1.14.0*, 2025, available from https://www.sympy.org.
16. László Tóth, *Proofs, generalizations and analogs of Menon's identity: a survey*, Acta Univ. Sapientiae Math. **15** (2023), no. 1, 142–197.

Department of Mathematics, University of Luxembourg
6, avenue de la Fonte L-4364 Esch-sur-Alzette, Luxembourg
Email address: antonella.perucca@uni.lu

Department of Mathematical Sciences, Politecnico di Torino
Corso Duca degli Abruzzi 24, 10129 Torino, Italy
Email address: carlo.sanna@polito.it