COUNTING PRIMES WITH A GIVEN PRIMITIVE ROOT, UNIFORMLY

KAI (STEVE) FAN AND PAUL POLLACK

For Greg Martin on his retirement.

Abstract. The celebrated Artin conjecture on primitive roots asserts that given any integer g which is neither -1 nor a perfect square, there is an explicit constant $A(g) > 0$ such that the number $\Pi(x; g)$ of primes $p \leq x$ for which g is a primitive root is asymptotically $A(g)\pi(x)$ as $x \to \infty$, where $\pi(x)$ counts the number of primes not exceeding x. Artin's conjecture has remained unsolved since its formulation 98 years ago. Nevertheless, Hooley demonstrated in 1967 that Artin's conjecture is a consequence of the Generalized Riemann Hypothesis (GRH) for Dedekind zeta functions of certain Kummer extensions over $\mathbb{Q}$. In this paper, we establish the Artin–Hooley asymptotic formula, under GRH, whenever $\log x/\log\log 2|g| \to \infty$. Under GRH, we also show that the least prime p_g possessing g as a primitive root satisfies the upper bound $p_g = O(\log^{19}(2|g|))$ uniformly for all non-square $g \neq -1$. We conclude with an application to the average value of p_g as well as discussion of an analogue concerning the least "almost-primitive" root p_g^*.

1. Introduction

It is a classical result, due to Gauss, that the multiplicative group modulo a prime p is always cyclic. That is, given any prime number p, there is an integer g whose reduction mod p generates the group $(\mathbb{Z}/p\mathbb{Z})^\times$; following tradition, we call such an integer g a primitive root modulo p. On the other hand, if we start with a given $g \in \mathbb{Z}$, there need not be any prime p with g a primitive root mod p. For instance, $g = 4$ is not a primitive root modulo any prime, and the same holds for all even square values of g.

The distribution of primes p possessing a prescribed integer g as a primitive root is the subject of a celebrated 1927 conjecture of Emil Artin, formulated during a visit of Artin to Hasse (consult [1, §17.2] for the history, and see [12] for a comprehensive survey of related developments). For real $x > 0$ and integers g, let

$$\Pi(x; g) = \#\{\text{primes } p \leq x : g \text{ is a primitive root mod } p\}.$$

Let

$$\mathcal{G} = \{g \in \mathbb{Z} : |g| > 1, g \text{ not a square}\}.$$

Artin's primitive root conjecture predicts that for each $g \in \mathcal{G}$,

$$\Pi(x; g) \sim A(g)\pi(x), \quad \text{as} \quad x \to \infty, \tag{1}$$

for an explicitly given $A(g) > 0$.

The conjectured form of $A(g)$ depends on the arithmetic nature of g. For each $g \in \mathcal{G}$, let g_1 denote the unique squarefree integer with $g \in g_1(\mathbb{Q}^\times)^2$, and let h be the largest positive integer for which

---

2020 *Mathematics Subject Classification.* Primary 11N05; Secondary 11A07, 11N36.

$g \in (\mathbb{Q}^\times)^h$. Since g is not a square, h is odd. Put

$$A_0(g) = \prod_{q \mid h} \left(1 - \frac{1}{q-1}\right) \prod_{q \nmid h} \left(1 - \frac{1}{q(q-1)}\right). \tag{2}$$

If $g_1 \equiv 1 \pmod 4$, put

$$A_1(g) = 1 - \mu(|g_1|) \prod_{\substack{q \mid h \\ q \mid g_1}} \frac{1}{q-2} \prod_{\substack{q \nmid h \\ q \mid g_1}} \frac{1}{q^2 - q - 1}; \tag{3}$$

otherwise, set $A_1(g) = 1$. Finally, put

$$A(g) = A_0(g) A_1(g).$$

It is this value of $A(g)$ for which Artin predicts (1).[1]

Artin's conjecture remains unresolved. In fact, to this day there is not a single value of g for which we can show even the weaker assertion that $\Pi(x; g) \to \infty$. (However, work of Heath-Brown [7] implies this holds for at least one of $g = 2$, 3, or 5.) The most important progress in this direction is a 1967 theorem of Hooley [8], asserting that the full asymptotic relation (1) follows from the Generalized Riemann Hypothesis (GRH).[2]

Hooley states and proves his asymptotic formula for *fixed* $g \in \mathcal{G}$. Our main result makes the dependence on g explicit.

Theorem 1.1 (assuming GRH)**.** *The asymptotic formula $\pi(x; g) \sim A(g)\Pi(x; g)$ holds whenever $\log x / \log\log 2|g| \to \infty$. More precisely, there is an absolute constant $x_0 > 0$ for which the following holds: If $g \in \mathcal{G}$ and $x \geq \max\{x_0, \log^3(2|g|)\}$, then*

$$\Pi(x; g) = A(g)\pi(x) \left(1 + O\left(\frac{\log\log x}{\log x} + \frac{\log\log 2|g|}{\log x}\right)\right). \tag{4}$$

The proof of Theorem 1.1, presented in §2, broadly proceeds along the same course as Hooley's, but care and caution are required to ensure the final estimate is nontrivial in a wide range of x and g. In particular, the fact that the positive constant $A(g)$ can be arbitrarily small causes substantial complications.

Let p_g denote the least prime p possessing g as a primitive root, where we set $p_g = \infty$ when no such p exists. Theorem 1.1 implies immediately that for all $g \in \mathcal{G}$,

$$p_g \ll \log^B(2|g|), \tag{5}$$

for a certain absolute constant B. Indeed, if K is an admissible value of the implied constant in (4), then (5) holds for any $B > K$. In our next theorem, we pinpoint a numerically explicit value of B.

Theorem 1.2 (assuming GRH)**.** *The upper bound* (5) *holds with $B = 19$.*

[1] Artin's original 1927 formulation was missing the factor of $A_1(g)$. Artin realized the need for $A_1(g)$ after learning of computations carried out by the Lehmers. See Stevenhagen's discussion in [16].

[2] Here and below, GRH means the Riemann Hypothesis for all Dedekind zeta functions.

Usually p_g is quite small. For instance, $p_g = 2$ whenever g is odd, while for even g, one has $p_g = 3$ one-third of the time (whenever $3 \mid g+1$). Proceeding more generally, there are $\varphi(p-1)$ primitive roots modulo the prime p. So by the Chinese remainder theorem, for each fixed p a random g satisfies $p_g > p$ with probability $\prod_{r\le p}(1-\frac{\varphi(r-1)}{r})$. To make the term "probability" here rigorous, we can interpret it as limiting frequency, with g sampled from integers satisfying $|g| \le x$, where $x \to \infty$.

This probabilistic viewpoint can be used to formulate a conjecture on the upper order of p_g. While $\varphi(r-1)/r$ fluctuates as the prime r varies, for the sake of estimating the above product on r, we can treat the terms $1-\frac{\varphi(r-1)}{r}$ as constant. More precisely, there is a certain real number $\varrho > 1$ such that $\prod_{r\le r_k}(1-\frac{\varphi(r-1)}{r}) = \varrho^{-(1+o(1))k}$ as $k\to\infty$, where r_k denotes the kth prime in the usual order. (We do not prove this here, but a related result appears as Lemma 4.1 below.) Since $2x\varrho^{-k} < 1$ once $k > k_0(x) := \frac{\log 2x}{\log \varrho}$, it is tempting to conjecture that $\max_{|g|\le x} p_g$ is never more than about $p_{k_0(x)}$. (This requires "pretending" that our probabilities, which were given rigorous meaning only when fixing k and sending x to infinity, can be interpreted uniformly in k and x.) This cannot be quite right, as $p_g = \infty$ for even square values of g ! Nevertheless, it seems sensible to guess that $p_g \ll (\log 2|g|)(\log\log 2|g|)$ for all $g \notin \mathcal{G}$. If correct, this is sharp: In [13], Pomerance and Shparlinski report a a construction of Soundararajan yielding an infinite sequence of positive integers g that (a) are all products of two distinct primes and (b) are squares modulo every odd prime $p \le 0.7(\log g)(\log\log g)$.[3] These g satisfy $p_{4g} \gg \log(4g)\log\log(4g)$.

This same perspective suggests that the "probability" $p_g > p$ is given by

$$\delta_p := \frac{\varphi(p-1)}{p}\prod_{r<p}\left(1-\frac{\varphi(r-1)}{r}\right). \tag{6}$$

Taking this for granted and proceeding formally, $\mathbb{E}[p_g] = \sum_p p\delta_p$. Using Theorem 1.2, we give a GRH-conditional proof that this sum represents the honest average of p_g.

Corollary 1.3. *We have that $\sum_p p\delta_p < \infty$. Furthermore, assuming GRH,*

$$\lim_{x\to\infty}\frac{1}{2x}\sum_{g\in\mathcal{G},\,|g|\le x} 1 = \sum_p p\delta_p. \tag{7}$$

Here δ_p is as defined in (6).

(We divide by $2x$, as there are $2x + O(x^{1/2})$ integers $g \in \mathcal{G}$ with $|g| \le x$.) There seems no hope at present of proving Corollary 1.3 unconditionally. If $p_g = \infty$ for even a single value of g, then the average becomes meaningless, and we know of no way to rule this out. Infinite values of p_g are not the only enemy: Having $p_g > x\log x$ for some g, $|g| \le x$ (along a sequence of x tending to infinity) is enough to doom (7).

In an attempt to salvage the situation, one might tamp down the large values of p_g by averaging $\min\{p_g, \psi(x)\}$ for a threshold function ψ. In our final theorem, established in §5, we show that this strategy succeeds for $\psi(x) = x^\eta$, for any positive $\eta < \frac{1}{2}$.

[3] Here 0.7 can be replaced with any constant smaller than $1/\log 4$.

Theorem 1.4. *Fix a positive real number $\eta < \frac{1}{2}$. Then*

$$\lim_{x\to\infty} \frac{1}{2x} \sum_{g\in\mathcal{G},\, |g|\le x} \min\{p_g, x^\eta\} = \sum_p p\delta_p.$$

It would be interesting to prove Theorem 1.4 with a less stringent condition on η, such as $\eta < 1$. But a substantial new idea would seem required to take η past 1/2. As we explain in §6, the problem becomes easier if we look instead at almost-primitive roots, meaning numbers g which generate a subgroup of index at most two inside $(\mathbb{Z}/p\mathbb{Z})^\times$.

The problems we have taken up about p_g are dual to those classically considered for g_p, the least primitive root modulo the prime p. Burgess [2] and Wang [18] have shown unconditionally that $g_p \ll p^{\frac{1}{4}+\varepsilon}$ for all primes p, while Shoup [15] (sharpening an earlier, qualitatively similar result of Wang, op. cit.) has proved under GRH that $g_p \ll r^4(1+\log r)^4 \log^2 p$, where $r = \omega(p-1)$. Shoup's upper bound is of size $\log^{2+o(1)} p$ for most primes p and is always $O(\log^6 p)$. These pointwise results are stronger than those known for p_g, but the story for average values is different. While g_p is conjectured to have a finite, limiting mean value, this has not been established even assuming GRH (that is, the analogue of Corollary 1.3 remains open). In fact, GRH has not yielded a stronger upper bound for $\pi(x)^{-1}\sum_{p\le x} g_p$ than $(\log x)(\log\log x)^{1+o(1)}$ (as $x\to\infty$); this last estimate is due to Elliott and Murata [3].

2. A uniform variant of Hooley's formula: Proof of Theorem 1.1

The following lemma encodes the input of GRH to the proof. It will be of vital importance both in this section and the next.

Lemma 2.1 (assuming GRH)**.** *Let g be a nonzero integer. For each real number $x \ge 2$ and each $d \in \mathbb{N}$, the count of primes $p \le x$ for which*

$$p \equiv 1 \pmod d \quad \text{and} \quad g^{(p-1)/d} \equiv 1 \pmod p \tag{8}$$

is

$$\frac{\pi(x)}{[\mathbb{Q}(\zeta_d, \sqrt[d]{g}) : \mathbb{Q}]} + O(x^{1/2}\log(|g|dx)).$$

Here the implied constant is absolute.

Proof. Apart from making explicit the dependence on g, this result is well-known and present already in [8]. Since dependence on g is crucial for our purposes, we sketch a proof. We first throw out primes dividing $d|g|$; there are only $O(\log(|g|d))$ of these, a quantity subsumed by our error term. For the remaining primes p,

$$\begin{aligned} (8) \text{ holds } &\iff x^d - g \text{ has } d \text{ distinct roots over } \mathbb{F}_p \\ &\iff x^d - g \text{ factors over } \mathbb{F}_p \text{ into } d \text{ distinct monic linear polynomials} \\ &\iff p \text{ splits completely in } \mathbb{Q}(\zeta_d, \sqrt[d]{g}). \end{aligned}$$

To count primes up to x satisfying this last condition, we apply the GRH-conditional Chebotarev density theorem in the form (20_R) of [14] (in the notation of [14], take $K = \mathbb{Q}$, $E = \mathbb{Q}(\zeta_d, \sqrt[d]{g})$, $C = \{\text{id}\}$, and keep in mind that all primes ramifying in E divide gd). $\square$

We now turn to the proof proper. We follow Hooley's strategy, but keep a more watchful eye on g-dependence in the error terms.

Let p be a prime not dividing g. For each prime number ℓ, we say that p fails the ℓ-test if

$$p \equiv 1 \pmod{\ell} \quad \text{and} \quad g^{(p-1)/\ell} \equiv 1 \pmod{p};$$

otherwise, we say p passes the ℓ-test. Then g is a primitive root modulo p precisely when p passes the ℓ-test for all primes p. In particular, if we define

$$\Pi_0(x;g) = \#\{p \le x : p \nmid g, p \text{ passes all } \ell\text{-tests for } \ell \le \log x\},$$

then

$$\Pi(x;g) \le \Pi_0(x;g).$$

For each squarefree $d \in \mathbb{N}$, let N_d denote the count of primes $p \le x$ which fail the ℓ-test for each prime $\ell \mid d$. These are precisely the primes $p \le x$ for which (8) holds, so that by Lemma 2.1 and inclusion-exclusion,

$$\begin{aligned} \Pi_0(x;g) &= \sum_{d\colon P^+(d) \le \log x} \mu(d) N_d \\ &= \pi(x) \sum_{d\colon P^+(d)\le \log x} \frac{\mu(d)}{[\mathbb{Q}(\zeta_d, \sqrt[d]{g})]} + O\Bigg(x^{1/2} \sum_{d\colon P^+(d)\le\log x} \log(|g|dx)\Bigg). \end{aligned} \tag{9}$$

(Throughout, we use $P^+(\cdot)$ for the largest prime factor, with the convention that $P^+(1) = 1$.) The error term is readily handled: Each squarefree d with $P^+(d) \le \log x$ satisfies $d \le \prod_{r \le \log x} r \le x^2$, and there are $2^{\pi(x)} = \exp(O(\log x/\log\log x))$ such values of d. Hence,

$$x^{1/2} \sum_{d\colon P^+(d)\le\log x} \log(|g|dx) \ll x^{1/2}\log(|g|x) \cdot \exp(O(\log x/\log\log x)) \ll x^{3/5}\log|g|. \tag{10}$$

Turning to the main term, we extract from [8, pp. 213–214] that for each squarefree $d \in \mathbb{N}$,

$$[\mathbb{Q}(\zeta_d, \sqrt[d]{g}) : \mathbb{Q}] = \frac{d\varphi(d)}{\varepsilon(d)\gcd(d,h)}, \quad \text{where} \quad \varepsilon(d) = \begin{cases} 2 & \text{if } 2g_1 \mid d \text{ and } g_1 \equiv 1 \pmod 4, \\ 1 & \text{otherwise.} \end{cases}$$

(Actually, what Hooley computes in [8] is the degree of $\mathbb{Q}(\zeta_d, \sqrt[d_1]{g})$, where $d_1 := d/\gcd(d,h)$. But this is the same field as $\mathbb{Q}(\zeta_d, \sqrt[d]{g})$, by Kummer theory, since g and $g^{\gcd(d,h)}$ generate the same subgroup of $\mathbb{Q}(\zeta_d)^\times/(\mathbb{Q}(\zeta_d)^\times)^d$.) From this, Hooley deduces in [8] that

$$\sum_d \frac{\mu(d)}{[\mathbb{Q}(\zeta_d) : \mathbb{Q}]} = A(g),$$

where the sum is over all $d \in \mathbb{N}$. We would like to plug this result into (9), but the corresponding sum in (9) is restricted to $(\log x)$-smooth values of d.

Let us examine the error incurred by replacing the sum over all d by the sum over $(\log x)$-smooth d. If $g_1 \not\equiv 1 \pmod 4$, then

$$\begin{aligned}\sum_{d:\, P^+(d)>\log x} \frac{\mu(d)}{[\mathbb{Q}(\zeta_d):\mathbb{Q}]} &= \sum_{d:\, P^+(d)>\log x} \mu(d)\frac{(d,h)}{d\varphi(d)} = -\sum_{\ell>\log x} \frac{(\ell,h)}{\ell\varphi(\ell)} \sum_{d:\, P^+(d)<\ell} \mu(d)\frac{(d,h)}{d\varphi(d)} \\ &= -\sum_{\ell>\log x} \frac{(\ell,h)}{\ell(\ell-1)} \prod_{\substack{r<\ell\\ r\nmid h}}\left(1-\frac{1}{r(r-1)}\right)\prod_{\substack{r<\ell\\ r\mid h}}\left(1-\frac{1}{r-1}\right) \\ &\ll \sum_{\ell>\log x} \frac{(\ell,h)}{\ell(\ell-1)}\frac{\varphi(h)}{h}\prod_{\substack{r\mid h\\ r\geq \ell}}\left(1+\frac{1}{r}\right).\end{aligned}$$

Each r appearing in this last expression has $r>\log x$. Furthermore,

$$\prod_{\substack{r\mid h\\ r>\log x}}\left(1+\frac{1}{r}\right) \leq \exp\Bigg(\sum_{\substack{r\mid h\\ r>\log x}}\frac{1}{r}\Bigg) \leq \exp\Bigg(\frac{1}{\log x}\sum_{\substack{r\mid h\\ r>\log x}} 1\Bigg) \leq \exp\left(\frac{\log h}{\log x\cdot\log\log x}\right) \ll 1, \qquad (11)$$

noting that

$$h \leq \frac{\log|g|}{\log 2} < \log^3(2|g|) \leq x$$

in the last step. Hence, $\prod_{r\mid h,\ r\geq\ell}(1+1/r) \ll 1$, and

$$\begin{aligned}\sum_{d:\, P^+(d)>\log x} \frac{\mu(d)}{[\mathbb{Q}(\zeta_d):\mathbb{Q}]} &\ll \frac{\varphi(h)}{h}\Bigg(\sum_{\substack{\ell>\log x\\ \ell\mid h}}\frac{1}{\ell} + \sum_{\substack{\ell>\log x\\ \ell\nmid h}}\frac{1}{\ell^2}\Bigg) \\ &\ll \frac{\varphi(h)}{h}\left(\frac{1}{\log x}\frac{\log h}{\log\log x} + \frac{1}{\log x}\right) \\ &\ll \frac{\varphi(h)}{h}\cdot\frac{\log\log 2|g|}{\log x},\end{aligned}$$

where we take from the last display that $\log h \ll \log\log 2|g|$.

When $g_1 \equiv 1 \pmod 4$, the argument is similar, but the details are slightly more involved. In this case,

$$\sum_{d:\, P^+(d)>\log x} \frac{\mu(d)}{[\mathbb{Q}(\zeta_d):\mathbb{Q}]} = \sum_{d:\, P^+(d)>\log x} \mu(d)\frac{(d,h)}{d\varphi(d)} + \sum_{\substack{d:\, P^+(d)>\log x\\ 2g_1\mid d}} \mu(d)\frac{(d,h)}{d\varphi(d)}.$$

The first-right hand sum has already been shown to be $O(\frac{\varphi(h)}{h}\frac{\log\log 2|g|}{\log x})$. The second obeys the same bound: In this case,

$$\sum_{\substack{d:\, P^+(d)>\log x\\ 2g_1\mid d}} \mu(d)\frac{(d,h)}{d\varphi(d)} = -\sum_{\ell>\log x}\frac{(\ell,h)}{\ell\varphi(\ell)}\sum_{\substack{d:\, P^+(d)<\ell\\ 2g_1\mid \ell d}} \mu(d)\frac{(d,h)}{d\varphi(d)}. \qquad (12)$$

The right-hand sum on d is empty if $2g_1/(2g_1,\ell)$ has a prime factor at least ℓ. In all other cases,

$$\sum_{\substack{d\colon P^+(d)<\ell\\ 2g_1\mid \ell d}} \mu(d)\frac{(d,h)}{d\varphi(d)} = \prod_{r\mid \frac{2g_1}{(2g_1,\ell)}} -\frac{(r,h)}{r(r-1)} \prod_{\substack{r<\ell\\ r\nmid \frac{2g_1}{(2g_1,\ell)}}} \left(1-\frac{(r,h)}{r(r-1)}\right).$$

Keeping in mind that h is odd, we observe that $\frac{(r,h)}{r(r-1)} \le \frac{1}{2}$ for each prime r, so that $\frac{(r,h)}{r(r-1)} \le 1-\frac{(r,h)}{r(r-1)}$. Therefore,

$$\left|\sum_{\substack{d\colon P^+(d)<\ell\\ 2g_1\mid \ell d}} \mu(d)\frac{(d,h)}{d\varphi(d)}\right| \le \prod_{r<\ell}\left(1-\frac{(r,h)}{r(r-1)}\right) \le \prod_{\substack{r<\ell\\ r\mid h}}\left(1-\frac{1}{r-1}\right) \ll \frac{\varphi(h)}{h}\prod_{\substack{r\mid h\\ r\ge \ell}}\left(1+\frac{1}{r}\right),$$

and referring back to (12),

$$\sum_{\substack{d\colon P^+(d)>\log x\\ 2g_1\mid d}} \mu(d)\frac{(d,h)}{d\varphi(d)} \ll \sum_{\ell>\log x} \frac{(\ell,h)}{\ell(\ell-1)}\frac{\varphi(h)}{h}\prod_{\substack{r\mid h\\ r\ge\ell}}\left(1+\frac{1}{r}\right).$$

But the right-hand side was estimated above as $O(\frac{\varphi(h)}{h}\frac{\log\log 2|g|}{\log x})$.

We conclude that in every case,

$$\sum_{d\colon P^+(d)>\log x} \frac{\mu(d)}{[\mathbb{Q}(\zeta_d):\mathbb{Q}]} \ll \frac{\varphi(h)}{h}\frac{\log\log 2|g|}{\log x}. \tag{13}$$

Combining (9), (10), and (13), we arrive at the estimate

$$\Pi_0(x;g) = A(g)\pi(x) + O\left(\frac{\varphi(h)}{h}\frac{\log\log 2|g|}{\log x}\pi(x) + x^{3/5}\log|g|\right).$$

We will need the error term in "multiplicative form". Notice that $A_0(g)$, as defined in (2), satisfies $A_0(g) \asymp \varphi(h)/h$. Recalling the definition (3) of $A_1(g)$ in the case when $g_1 \equiv 1 \pmod 4$, we see that the subtracted term in (3) always has absolute value at most 1. In fact, that absolute value is at most 1/3 unless $g_1 = -3$, in which case $\mu(|g_1|) = -1$. Hence, $\frac{2}{3} \le A_1(g) \le 2$, and

$$A(g) = A_0(g)A_1(g) \asymp \frac{\varphi(h)}{h}.$$

Therefore,

$$\begin{aligned}\frac{\varphi(h)}{h}\pi(x)\cdot\frac{\log\log 2|g|}{\log x} + x^{3/5}\log|g| &\ll A(g)\pi(x)\left(\frac{\log\log 2|g|}{\log x} + \frac{(h/\varphi(h))\log|g|}{x^{3/8}}\right)\\ &\ll A(g)\pi(x)\left(\frac{\log\log 2|g|}{\log x} + \frac{\log\log x}{x^{1/24}}\right)\\ &\ll A(g)\pi(x)\frac{\log\log 2|g|}{\log x}.\end{aligned}$$

Here in going from the first line to the second, we use that $h/\varphi(h) \ll \log\log 3h \ll \log\log x$ and that $\log|g| \ll x^{1/3} = x^{3/8}/x^{1/24}$. We conclude that

$$\Pi_0(x;g) = A(g)\pi(x)\left(1 + O\left(\frac{\log\log 2|g|}{\log x}\right)\right). \tag{14}$$

Next, we investigate the difference $\Pi_0(x;g)-\Pi(x;g)$. If the prime $p\le x$ is counted by $\Pi_0(x;g)$ but not $\Pi(x;g)$, then p passes the ℓ-tests for all $\ell\le\log x$ but fails the ℓ-test for some $\ell>\log x$. Set

$$x_1=\log x,\quad x_2=x^{1/2}(\log x)^{-2}(\log|g|)^{-1},\quad x_3=x^{1/2}(\log x)^2\log|g|,$$

and put

$$I_1=(x_1,x_2],\quad I_2=(x_2,x_3],\quad I_3=(x_3,\infty).$$

For $j\in\{1,2,3\}$, let E_j denote the count of primes $p\le x$, $p\nmid g$, which fail the ℓ-test for the first time for an $\ell\in I_j$. Then

$$\Pi_0(x;g)\ge\Pi(x;g)\ge\Pi_0(x;g)-E_1-E_2-E_3. \tag{15}$$

We proceed to estimate the E_j in turn.

Invoking again Lemma 2.1,

$$\begin{aligned}E_1\le\sum_{\ell\in I_1}N_\ell&\ll\sum_{\ell\in I_1}\left(\pi(x)\frac{(\ell,h)}{\ell^2}+x^{1/2}\log(|g|\ell x)\right)\\&\ll\pi(x)\left(\sum_{\ell>\log x}\frac{1}{\ell^2}+\sum_{\substack{\ell>\log x\\ \ell\mid h}}\frac{1}{\ell}\right)+x^{1/2}\log(|g|x)\cdot\pi(x_2).\end{aligned}$$

Since $h<x$ and $\log h\ll\log\log 2|g|$,

$$\begin{aligned}\sum_{\ell>\log x}\frac{1}{\ell^2}+\sum_{\substack{\ell>\log x\\ \ell\mid h}}\frac{1}{\ell}&\ll\frac{1}{\log x\cdot\log\log x}+\frac{1}{\log x}\frac{\log h}{\log\log x}\ll\frac{\log\log 2|g|}{\log x\cdot\log\log x}\\&=\frac{\varphi(h)}{h}\left(\frac{h/\varphi(h)}{\log x}\frac{\log\log 2|g|}{\log\log x}\right)\ll\frac{\varphi(h)}{h}\left(\frac{\log\log x}{\log x}\frac{\log\log 2|g|}{\log\log x}\right)=\frac{\varphi(h)}{h}\frac{\log\log 2|g|}{\log x},\end{aligned}$$

so that

$$\pi(x)\left(\sum_{\ell>\log x}\frac{1}{\ell^2}+\sum_{\substack{\ell>\log x\\ \ell\mid h}}\frac{1}{\ell}\right)\ll A(g)\pi(x)\cdot\frac{\log\log 2|g|}{\log x}.$$

We are assuming that $x\ge(\log 2|g|)^3$. Hence,

$$x_2\ge x^{1/6}(\log x)^{-2}>x^{1/7}$$

for all x exceeding a certain absolute constant, and $\log x_2\gg\log x$. Thus, $\pi(x_2)\ll x_2(\log x)^{-1}=x^{1/2}(\log x)^{-3}(\log|g|)^{-1}$, and

$$\begin{aligned}x^{1/2}\log(|g|x)\cdot\pi(x_2)&\ll\frac{x}{(\log x)^3\log|g|}(\log|g|x)\ll\pi(x)\frac{\log|g|x}{(\log x)^2\log|g|}\\&\ll\frac{\pi(x)}{\log x}=\frac{\varphi(h)}{h}\pi(x)\cdot\frac{h/\varphi(h)}{\log x}\ll A(g)\pi(x)\frac{\log\log x}{\log x}.\end{aligned}$$

Collecting our results,

$$E_1\ll A(g)\pi(x)\left(\frac{\log\log 2|g|}{\log x}+\frac{\log\log x}{\log x}\right). \tag{16}$$

We turn now to E_2. Let ℓ be a prime dividing h. Then every prime $p\equiv 1\pmod{\ell}$, with p not dividing g, satisfies

$$g^{(p-1)/\ell}\equiv 1\pmod{p},$$

as g is an ℓth power. Hence, in order for a prime p (not dividing g) to pass the ℓ-test, it must be that $p \not\equiv 1 \pmod{\ell}$. By assumption, the primes counted in E_2 pass the ℓ-test for all $\ell \le x_2$, and hence for all $\ell \le x^{1/7}$. So if we let h' denote the $x^{1/7}$-smooth part of h, then each prime p counted in E_2 has $(p-1, h') = 1$. Since p also fails the ℓ-test for some $p \in I_2$,

$$E_2 \le \sum_{\ell \in I_2} \sum_{\substack{p \le x \\ (p-1,h')=1 \\ p \equiv 1 \ (\mathrm{mod}\ \ell)}} 1.$$

Each prime p counted by the inner sum has the form $p = 1 + \ell m$. Here $0 < m < x/\ell$, and m avoids the residue classes $0 \bmod r$ for all primes $r \mid h$, $r \le x^{1/7}$, as well as the residue classes of $-1/\ell \bmod r$ for each prime $r < \ell$. Moreover, for each $\ell \in I_2$, we have $\ell > x_2 > x^{1/7}$ as well as $x/\ell \ge x/x_3 = x_2 > x^{1/7}$. Applying Brun's sieve,

$$\sum_{\substack{p \le x \\ (p-1,h')=1 \\ p \equiv 1 \ (\mathrm{mod}\ \ell)}} 1 \ll \frac{x}{\ell} \prod_{r \le x^{1/7}} \left(1 - \frac{1 + 1_{r|h}}{r}\right) \ll \frac{x}{\ell \log x} \prod_{\substack{r \le x^{1/7} \\ r|h}} \left(1 - \frac{1}{r}\right) \ll \frac{\pi(x)}{\ell} \frac{\varphi(h)}{h} \prod_{\substack{r > x^{1/7} \\ r|h}} \left(1 + \frac{1}{r}\right).$$

We have from (11) that the final product on r is $O(1)$. Thus,

$$\sum_{\ell \in I_2} \sum_{\substack{p \le x \\ (p-1,h')=1 \\ p \equiv 1 \ (\mathrm{mod}\ \ell)}} 1 \ll \pi(x) \frac{\varphi(h)}{h} \sum_{\ell \in I_2} \frac{1}{\ell} \ll \pi(x) \frac{\varphi(h)}{h} \left(\frac{\log\log x}{\log x} + \frac{\log\log 2|g|}{\log x} \right),$$

using Mertens' theorem to estimate the sum on ℓ. As $A(g) \asymp \varphi(h)/h$, we conclude that

$$E_2 \ll \pi(x) A(g) \left(\frac{\log\log x}{\log x} + \frac{\log\log 2|g|}{\log x} \right). \tag{17}$$

Finally we consider E_3. Each p counted in E_3 has $g^{(p-1)/\ell} \equiv 1 \pmod{p}$ for some $\ell > x_3$. Thus, the order of g mod p is smaller than $x/x_3 = x_2$, and p divides $g^m - 1$ for some natural number $m < x_2$. The number of distinct prime factors of $g^m - 1$ is $O(m \log |g|)$, and so

$$E_3 \ll \log|g| \sum_{m < x_2} m \ll x_2^2 \log|g| = \frac{x}{(\log^4 x)(\log|g|)}.$$

In particular,

$$E_3 \ll \frac{\pi(x)}{\log x} = \frac{\varphi(h)}{h} \pi(x) \cdot \frac{h/\varphi(h)}{\log x} \ll A(g)\pi(x) \cdot \frac{\log\log x}{\log x}. \tag{18}$$

Combining (14), (15), (16), (17), and (18),

$$\Pi(x; g) = A(g)\pi(x) \left(1 + O\left(\frac{\log\log x}{\log x} + \frac{\log\log 2|g|}{\log x} \right)\right);$$

this completes the proof of Theorem 1.1.

3. An explicit upper bound for the least Artin prime p_g: Proof of Theorem 1.2

Now we turn to the proof of Theorem 1.2. We may assume that $|g|$ is sufficiently large. Let $x = \log^B |g|$ with $B = 19$, and put $W = \prod_{2<p\le\log x} p$. Denote by $\mathcal{S}$ the set of primes $p \le x$ with $(g/p) = -1$ and $\gcd(p-1, W) = 1$.

First of all, let us estimate the number of elements in $\mathcal{S}$. By inclusion-exclusion,

$$\begin{aligned}\#\mathcal{S} = \frac{1}{2}\sum_{\substack{p\le x,\, p\nmid g\\ (p-1,W)=1}} (1-(g/p)) &= \frac{1}{2}\sum_{\substack{p\le x\\ (p-1,W)=1}} (1-(g/p)) + O(\omega(g))\\ &= \frac{1}{2}\sum_{p\le x}(1-(g/p))\sum_{\substack{d\mid p-1\\ d\mid W}}\mu(d) + O(\log|g|)\\ &= \frac{1}{2}\sum_{d\mid W}\mu(d)\sum_{\substack{p\le x\\ p\equiv 1 \ (\mathrm{mod}\ d)}}(1-(g/p)) + O(\log|g|)\\ &= \frac{1}{2}\sum_{d\mid W}\mu(d)\pi(x;d,1) - \frac{1}{2}\sum_{d\mid W}\mu(d)\sum_{\substack{p\le x\\ p\equiv 1 \ (\mathrm{mod}\ d)}}(g/p) + O(\log|g|),\end{aligned}$$

where $\pi(x;d,1)$ denotes the number of primes $p \le x$ with $p \equiv 1 \pmod d$. To estimate the first sum above, we appeal to [11, Corollary 13.8], the GRH-conditional prime number theorem for primes in arithmetic progressions, to obtain

$$\begin{aligned}\frac{1}{2}\sum_{d\mid W}\mu(d)\pi(x;d,1) = \frac{1}{2}\sum_{d\mid W}\mu(d)\left(\frac{\mathrm{Li}(x)}{\varphi(d)} + O\left(x^{1/2}\log x\right)\right) &= \frac{\tilde{A}_0(g)}{2}\mathrm{Li}(x) + O\left(2^{\pi(\log x)}x^{1/2}\log x\right)\\ &= \frac{\tilde{A}_0(g)}{2}\mathrm{Li}(x) + O\left(x^{1/2+o(1)}\right),\end{aligned}$$

where

$$\tilde{A}_0(g) = \sum_{d\mid W}\frac{\mu(d)}{\varphi(d)} = \prod_{2<q\le\log x}\left(1-\frac{1}{q-1}\right).$$

In addition, we can rewrite the second sum above as

$$\frac{1}{2}\sum_{d\mid W}\mu(d)\sum_{\chi\ (\mathrm{mod}\ d)}\sum_{\substack{p\le x\\ p\nmid g}}\chi(p)(g/p),$$

by the orthogonality relations of Dirichlet characters, where the second summation in the triple sum runs over all Dirichlet characters $\chi \pmod d$. It follows that

$$\#\mathcal{S} = \frac{\tilde{A}_0(g)}{2}\mathrm{Li}(x) - \frac{1}{2}\sum_{d\mid W}\frac{\mu(d)}{\varphi(d)}\sum_{\chi\ (\mathrm{mod}\ d)}\sum_{\substack{p\le x\\ p\nmid g}}\chi(p)(g/p) + O\left(x^{1/2+o(1)}\right). \tag{19}$$

To estimate the triple sum in (19), we recall that $\mathbb{Q}(\sqrt{g}) = \mathbb{Q}(\sqrt{g_1})$, where $g_1 \ne 1$ is the unique squarefree integer with $g_1(\mathbb{Q}^\times)^2 = g(\mathbb{Q}^\times)^2$. Let Δ be the discriminant of $\mathbb{Q}(\sqrt{g_1})$. Then $(g/p) = (\Delta/p)$ for all odd primes p not dividing g, and for these primes p, $\chi(p)(g/p)$ can be viewed as the value at p of a character $\psi_{\chi,g} \pmod{|\Delta|d}$. The character $\psi_{\chi,g}$ is non-principal unless χ is induced by the primitive character $(\Delta/\cdot) \pmod{|\Delta|}$. For that to occur, one needs $\Delta \mid d$; in that eventuality, to each d there corresponds exactly one character $\chi \pmod d$ for which $\psi_{\chi,g}$ is trivial. All of the d appearing above are odd, squarefree, and divide W, so in order for Δ to divide d we need Δ to be a squarefree divisor of W. This forces $\Delta = g_1 \equiv 1 \pmod 4$ and requires that $g_1 \mid W$. By [11, Theorem 13.7], the GRH-conditional estimates for character sums over primes, we have

$$\begin{aligned}
\frac{1}{2}\sum_{d|W}\frac{\mu(d)}{\varphi(d)}\sum_{\chi \pmod d}\sum_{\substack{p\le x\\ p\nmid g}}\chi(p)(g/p) &= \frac{1}{2}\sum_{d|W}\frac{\mu(d)}{\varphi(d)}\left(1_{g_1|d}\cdot 1_{4|(g_1-1)}\mathrm{Li}(x) + O\left(\varphi(d)x^{1/2}\log(dx)\right)\right)\\
&= \frac{1_{4|(g_1-1),\, g_1|W}}{2}\mathrm{Li}(x)\sum_{g_1|d|W}\frac{\mu(d)}{\varphi(d)} + O\left(2^{\pi(\log x)}x^{1/2}\log x\right)\\
&= \frac{1_{4|(g_1-1),\, g_1|W}}{2}\cdot\frac{\mu(g_1)}{\varphi(g_1)}\mathrm{Li}(x)\sum_{d|W/g_1}\frac{\mu(d)}{\varphi(d)} + O\left(x^{1/2+o(1)}\right)\\
&= \frac{1_{4|(g_1-1),\, g_1|W}}{2}\cdot\frac{\mu(g_1)}{\varphi(g_1)}\mathrm{Li}(x)\prod_{q|W/g_1}\left(1-\frac{1}{q-1}\right) + O\left(x^{1/2+o(1)}\right)\\
&= \frac{\tilde{A}_0(g)(1-\tilde{A}_1(g))}{2}\mathrm{Li}(x) + O\left(x^{1/2+o(1)}\right),
\end{aligned}$$

where

$$\tilde{A}_1(g) := 1 - 1_{4|(g_1-1),\, g_1|W}\frac{\mu(g_1)}{\varphi(g_1)}\prod_{q|g_1}\left(1-\frac{1}{q-1}\right)^{-1} = 1 - 1_{4|(g_1-1),\, g_1|W}\prod_{q|g_1}\frac{-1}{q-2}.$$

Inserting this estimate in (19) yields

$$\#\mathcal{S} = \frac{\tilde{A}_0(g)\tilde{A}_1(g)}{2}\mathrm{Li}(x) + O\left(x^{1/2+o(1)}\right). \tag{20}$$

It is worth noting that

$$\begin{aligned}
\tilde{A}_0(g) = \prod_{2<q\le\log x}\left(1-\frac{1}{q-1}\right) &= \prod_{2<q\le\log x}\left(1-\frac{1}{q}\right)\prod_{2<q\le\log x}\left(1-\frac{1}{q-1}\right)\left(1-\frac{1}{q}\right)^{-1}\\
&= \left(1+O\left(\frac{1}{\log\log x}\right)\right)\frac{2C_2e^{-\gamma}}{\log\log x}
\end{aligned}$$

and that

$$\frac{2}{3} = \tilde{A}_1(-15) \le \tilde{A}_1(g) \le \tilde{A}_1(-3) = 2,$$

where

$$C_2 := \prod_{q>2}\left(1-\frac{1}{q-1}\right)\left(1-\frac{1}{q}\right)^{-1} = \prod_{q>2}\left(1-\frac{1}{(q-1)^2}\right)$$

is the twin prime constant. Thus, the main term in (20) is of order $\mathrm{Li}(x)/\log\log x$.

Next, we estimate the number of $p \in \mathcal{S}$ modulo which g is not a primitive root. To this end, we count those $p \in \mathcal{S}$ which fail the ℓ-test for some $\ell > \log x$. Such an ℓ falls necessarily into one of the following four intervals:

$$J_1 := (\log x, y_1], \qquad J_2 := (y_1, y_2],$$
$$J_3 := (y_2, x^\alpha], \qquad J_4 := (x^\alpha, \infty],$$

where $\alpha \in (10/19, 1)$ is fixed, and

$$y_1 := \frac{x^{1/2}}{(\log |g|) \log^2 x},$$
$$y_2 := x^{1/2 - 1/\log\log x},$$

We start with J_1. Suppose first that $\ell \nmid h$. Applying Lemma 2.1 as in the proof of Theorem 1.1, we see that the count of $p \in \mathcal{S}$ that fail the ℓ-test for some $\ell \in J_1$ is

$$\ll \sum_{\ell \in J_1} \left(\frac{\mathrm{Li}(x)}{\ell^2} + x^{1/2} \log(|g|\ell x) \right) \ll \mathrm{Li}(x) \sum_{\ell > \log x} \frac{1}{\ell^2} + x^{1/2} \pi(y_1) \log(|g|) \ll \frac{\mathrm{Li}(x)}{\log x},$$

which is negligible compared to the main term in (20). In the case where $\ell \mid h$, we observe that a prime $p \le x$ failing the ℓ-test satisfies $p \equiv 1 \pmod{\ell}$ and $\gcd(p-1, W) = 1$. For each $\ell \in J_1$, the number of such $p \le x$ is

$$\sum_{\substack{p \le x \\ p \equiv 1 \ (\mathrm{mod}\ \ell) \\ (p-1, W) = 1}} 1 \le x^{1/3} + \sum_{\substack{m \le x/\ell \\ (m, W) = 1 \\ P^-(\ell m + 1) > x^{1/3}}} 1 \ll x^{1/3} + \frac{x}{\ell} \prod_{q \le x^{1/3}} \left(1 - \frac{1_{q \mid W} + 1_{q \ne \ell}}{q}\right)$$
$$\ll x^{1/3} + \frac{x}{\ell} \prod_{q \mid W} \left(1 - \frac{1}{q}\right) \prod_{\substack{q \le x^{1/3} \\ q \ne \ell}} \left(1 - \frac{1}{q}\right)$$
$$\ll \frac{\mathrm{Li}(x)}{\ell \log\log x},$$

by Brun's sieve. Summing this on $\ell > \log x$ with $\ell \mid h$ gives

$$\ll \frac{\mathrm{Li}(x)}{\log\log x} \sum_{\substack{\ell > \log x \\ \ell \mid h}} \frac{1}{\ell} \ll \frac{\mathrm{Li}(x)}{(\log x) \log\log x} \sum_{\substack{\ell > \log x \\ \ell \mid h}} 1 \ll \frac{\mathrm{Li}(x)}{(\log x) \log\log x} \cdot \frac{\log h}{\log\log x}.$$

Since $h \ll \log |g| = x^{1/B}$, this is $\ll \mathrm{Li}(x)/(\log\log x)^2$, which is also negligible compared to the main term in (20).

Moving on to J_2, we seek to bound the number of primes $p \in \mathcal{S}$ failing the ℓ-test for some $\ell \in J_2$. Such a prime p certainly satisfies $p \le x$, $\gcd(p-1, W) = 1$, and $p \equiv 1 \pmod{\ell}$. Using inclusion-exclusion and invoking [11, Corollary 13.8] again, we find that for each $\ell \in J_2$, the number of such p is

$$\sum_{\substack{p\le x\\ p\equiv 1 \pmod{\ell}\\ (p-1,W)=1}} 1 = \sum_{d\mid W}\mu(d)\pi(x;\ell d,1) = \sum_{d\mid W}\mu(d)\left(\frac{\mathrm{Li}(x)}{\varphi(\ell d)} + O\left(x^{1/2}\log x\right)\right)$$

$$= \frac{\mathrm{Li}(x)}{\varphi(\ell)}\sum_{d\mid W}\frac{\mu(d)}{\varphi(d)} + O\left(2^{\pi(\log x)}x^{1/2}\log x\right)$$

$$= \frac{\tilde{A}_0(g)}{\ell-1}\mathrm{Li}(x) + O\left(2^{\pi(\log x)}x^{1/2}\log x\right).$$

Summing this on $\ell\in J_2$ shows that the number of primes $p\in\mathcal{S}$ failing the ℓ-test for some $\ell\in J_2$ is

$$\sum_{\ell\in J_2}\left(\frac{\tilde{A}_0(g)}{\ell-1}\mathrm{Li}(x) + O\left(2^{\pi(\log x)}\sqrt{x}\log x\right)\right)$$

$$= \tilde{A}_0(g)\mathrm{Li}(x)\left(\log\frac{\log y_2}{\log y_1} + O\left(\frac{1}{\log y_1}\right)\right) + O\left(2^{\pi(\log x)}\pi(y_2)\sqrt{x}\log x\right)$$

$$= \left(\tilde{A}_0(g)\log\frac{B}{B-2} + O\left(\frac{1}{\log\log x}\right)\right)\mathrm{Li}(x) + O\left(x^{1-(1-\log 2+o(1))/\log\log x}\right)$$

$$= \left(\tilde{A}_0(g)\log\frac{B}{B-2} + O\left(\frac{1}{\log\log x}\right)\right)\mathrm{Li}(x),$$

where we have made use of the prime number theorem, Mertens' theorem, and the relation $x=\log^B|g|$.

Now we turn to J_3. As in the treatment of J_2, we shall only use that a prime $p\in\mathcal{S}$ failing the ℓ-test satisfies $p\equiv 1 \pmod{\ell}$ and that $\gcd(p-1,W)=1$. However, [11, Corollary 13.8] loses its strength in this case, for most $\ell\in J_3$ go way beyond $x^{1/2}$. To get around this issue, we resort to the following "arithmetic large sieve" inequality due to Montgomery (see [10, Chapter 3] and [4, §9.4]) to obtain an asymptotically explicit upper bound for the number of primes $p\le x$ satisfying $p\equiv 1 \pmod{\ell}$ and $\gcd(p-1,W)=1$, rather than pursue an asymptotic formula for this count.

Arithmetic large sieve. *Let Q be a positive integer. To each prime $p\le Q$, associate $\nu(p)<p$ residue classes modulo p. For every pair of integers M,N, with $N>0$, the number of integers in $[M+1,M+N]$ avoiding the distinguished residue classes mod p for all primes $p\le Q$ is bounded above by*

$$\frac{N+Q^2}{J},\quad \textit{where}\quad J:=\sum_{n\le Q}\mu^2(n)\prod_{p\mid n}\frac{\nu(p)}{p-\nu(p)}.$$

By the large sieve, the count of $p\le x$ corresponding to a given $\ell\in J_3$ is at most

$$\sum_{\substack{m\le x/\ell\\ (m,V)=1\\ P^-(\ell m+1)>(x/\ell)^\beta}} 1 \le \left(\frac{x}{\ell}+\left(\frac{x}{\ell}\right)^{2\beta}\right)\left(\sum_{n\le(x/\ell)^\beta}\mu(n)^2\prod_{q\mid n}\frac{\nu(q)}{q-\nu(q)}\right)^{-1} \tag{21}$$

where $\beta=\beta(x)=1/2-1/\log\log x$, V is the product of all odd primes not exceeding $\log x/\log\log x$, and $\nu(q)=1_{q\mid V}+1$. Here we have exploited the facts that $V\mid W$ and that $(x/\ell)^\beta<\ell$ for every $\ell\in J_3$.

To handle the sum on the right-hand side, we observe that $V = x^{(1+o(1))/\log\log x} = (x/\ell)^{O(1/\log\log x)}$ and that

$$\sum_{n\le (x/\ell)^\beta} \mu(n)^2 \prod_{q\mid n} \frac{\nu(q)}{q-\nu(q)} \ge \left(\sum_{d\mid V} \mu(d)^2 \prod_{q\mid d} \frac{2}{q-2}\right)\left(\sum_{\substack{m\le (x/\ell)^\beta/V\\ (m,V)=1}} \mu(m)^2 \prod_{q\mid m} \frac{1}{q-1}\right). \tag{22}$$

It is easy to see that

$$\sum_{d\mid V} \mu(d)^2 \prod_{q\mid d} \frac{2}{q-2} = \prod_{q\mid V}\left(1+\frac{2}{q-2}\right) = \left(1+O\left(\frac{\log\log\log x}{\log\log x}\right)\right)\prod_{q\mid W}\left(1+\frac{2}{q-2}\right). \tag{23}$$

In addition, we have

$$\sum_{\substack{m\le (x/\ell)^\beta/V\\ (m,V)=1}} \mu(m)^2 \prod_{q\mid m} \frac{1}{q-1} = \sum_{\substack{m\le (x/\ell)^\beta/V\\ (m,V)=1}} \frac{\mu(m)^2}{\varphi(m)} \ge \frac{\varphi(V)}{V} \sum_{m\le (x/\ell)^\beta/V} \frac{\mu(m)^2}{\varphi(m)},$$

where the last inequality follows from

$$\sum_{n\le z} \frac{\mu(n)^2}{\varphi(n)} \le \left(\sum_{d\mid a} \frac{\mu(d)^2}{\varphi(d)}\right)\left(\sum_{\substack{m\le z\\ (m,a)=1}} \frac{\mu(m)^2}{\varphi(m)}\right)$$

and

$$\sum_{d\mid a} \frac{\mu(d)^2}{\varphi(d)} = \frac{a}{\varphi(a)}$$

for all $z\ge 1$ and $a\in\mathbb{N}$. Since an application of [11, Eq. (3.18)] yields

$$\sum_{m\le (x/\ell)^\beta/V} \frac{\mu(m)^2}{\varphi(m)} > \log\left((x/\ell)^\beta/V\right) = \left(\frac{1}{2}+O\left(\frac{1}{\log\log x}\right)\right)\log(x/\ell),$$

we obtain

$$\begin{aligned}\sum_{\substack{m\le (x/\ell)^\beta/V\\ (m,V)=1}} \mu(m)^2 \prod_{q\mid m} \frac{1}{q-1} &\ge \left(\frac{1}{2}+O\left(\frac{1}{\log\log x}\right)\right)\frac{\varphi(V)}{V}\log(x/\ell)\\ &= \left(\frac{1}{2}+O\left(\frac{\log\log\log x}{\log\log x}\right)\right)\frac{\varphi(W)}{W}\log(x/\ell).\end{aligned}$$

Inserting this and (23) in (22) yields

$$\begin{aligned}\sum_{n\le (x/\ell)^\beta} \mu(n)^2 \prod_{q\mid n} \frac{\nu(q)}{q-\nu(q)} &\ge \left(\frac{1}{2}+O\left(\frac{\log\log\log x}{\log\log x}\right)\right)\frac{\varphi(W)}{W}\log(x/\ell)\prod_{q\mid W}\left(1+\frac{2}{q-2}\right)\\ &= \left(\frac{1}{2}+O\left(\frac{\log\log\log x}{\log\log x}\right)\right)\tilde{A}_0(g)^{-1}\log(x/\ell).\end{aligned}$$

Combining the above with (21), we find that the count of $p\le x$ corresponding to a given $\ell\in J_3$ is at most

$$\left(2+O\left(\frac{\log\log\log x}{\log\log x}\right)\right)\tilde{A}_0(g)\frac{x}{\ell\log(x/\ell)} = \left(2+O\left(\frac{\log\log\log x}{\log\log x}\right)\right)\tilde{A}_0(g)\frac{\mathrm{Li}(x)\log x}{\ell\log(x/\ell)}.$$

Summing this on $\ell \in J_3$, we see that the count of $p \le x$ in consideration is at most

$$\left(2 + O\left(\frac{\log\log\log x}{\log\log x}\right)\right) \tilde{A}_0(g)\mathrm{Li}(x)\log x \sum_{\ell \in J_3} \frac{1}{\ell \log(x/\ell)}.$$

By Mertens' theorem and partial summation, we have

$$\begin{aligned}
\sum_{\ell \in J_3} \frac{1}{\ell \log(x/\ell)} &= \int_{t \in J_3} \frac{1}{\log(x/t)} \,\mathrm{d}\left(\sum_{\ell \le t} \frac{1}{\ell}\right) \\
&= \int_{t \in J_3} \frac{\mathrm{d}t}{t(\log t)\log(x/t)} + \int_{t \in J_3} \frac{1}{\log(x/t)} \,\mathrm{d}\left(O\left(\frac{1}{\log t}\right)\right) \\
&= \frac{1}{\log x} \int_{1/2 - 1/\log\log x}^{\alpha} \frac{\mathrm{d}u}{u(1-u)} + O\left(\frac{1}{(\log x)^2}\right) \\
&= \frac{1}{\log x} \int_{1/2}^{\alpha} \frac{\mathrm{d}u}{u(1-u)} + O\left(\frac{1}{(\log x)\log\log x}\right) \\
&= \left(\log \frac{\alpha}{1-\alpha} + O\left(\frac{1}{\log\log x}\right)\right) \frac{1}{\log x}.
\end{aligned}$$

Hence, the count of $p \le x$ in consideration is at most

$$\left(2\log \frac{\alpha}{1-\alpha} + O\left(\frac{\log\log\log x}{\log\log x}\right)\right) \tilde{A}_0(g)\mathrm{Li}(x).$$

Finally, it remains to estimate the number of primes $p \in \mathcal{S}$ failing the ℓ-test for some $\ell \in J_4$. For each such p, the order of $g \pmod p$ is smaller than $x^{1-\alpha}$. Thus, $p \mid (g^m - 1)$ for some positive integer $m \le x^{1-\alpha}$. The number of distinct prime factors of $g^m - 1$ is $O(m \log |g|)$. Hence, the number of primes $p \in \mathcal{S}$ failing the ℓ-test for some $\ell \in J_4$ is at most

$$\sum_{m \le x^{1-\alpha}} m \log|g| \ll x^{2-2\alpha} \log|g| = x^{2-2\alpha+1/B}.$$

Since $\alpha \in (10/19, 1)$, we have $2 - 2\alpha + 1/B < 1$. Thus, $x^{2-2\alpha}\log|g|$ is of smaller order than the main term in (20).

Putting everything together, we deduce that the number of $p \in \mathcal{S}$ having g as a primitive root is at least

$$\left(\frac{\tilde{A}_1(g)}{2} - \log\frac{B}{B-2} - 2\log\frac{\alpha}{1-\alpha} + o(1)\right) \tilde{A}_0(g)\mathrm{Li}(x).$$

Since $\tilde{A}_1(g) \ge 2/3$, our choice of B guarantees that

$$\frac{\tilde{A}_1(g)}{2} - \log\frac{B}{B-2} - 2\log\frac{\alpha}{1-\alpha} \ge \frac{1}{3} - \log\frac{B}{B-2} - 2\log\frac{\alpha}{1-\alpha} > 0,$$

provided that $\alpha \in (10/19, 1)$ is sufficiently close to $10/19$. This proves that $p_g \le x = \log^B |g|$ with $B = 19$ for sufficiently large $|g|$.

Remark. Since $\tilde{A}_1(g) \ge \tilde{A}_1(21) = 4/5$ for $g > 1$, the proof of Theorem 1.2 shows that the exponent $B = 19$ can be improved to 16 if we focus merely on positive $g \in \mathcal{G}$. Besides, if the square factor of g has size $o(|g|)$ or $g_1 \not\equiv 1 \pmod 4$, we have $\tilde{A}_1(g) = 1 + o(1)$ for g with $|g|$ sufficiently large.

Consequently, our proof of Theorem 1.2 yields $p_g \ll \log^{13}(2|g|)$ for these $g \in \mathcal{G}$. In particular, this inequality holds for all squarefree $g \in \mathcal{G}$.

4. The average value of p_g: Proof of Corollary 1.3

The following lemma is due to Vaughan (see Theorem 4.1 in [17]).

Lemma 4.1. *For a certain constant $\alpha > 0$, we have*

$$\sum_{2<p\le y} \frac{\varphi(p-1)}{p-\varphi(p-1)} = (\alpha + o(1))\frac{y}{\log y}, \qquad \text{as } y \to \infty.$$

Put $L = \log x/\log\log x$. Let δ_p be defined as in (6), and put $M_p = \prod_{r\le p} r$. Then $p_g = p$ precisely when g belongs to one of $\delta_p M_p$ residue classes modulo M_p. Since $M_p \ll 3^p$,

$$\#\{g : |g| \le x : p_g = p\} = 2\delta_p x + O(3^p).$$

As $\#[-x,x] \setminus \mathcal{G} \ll x^{1/2}$, it follows that

$$\begin{aligned} \sum_{\substack{g\in\mathcal{G}\\ |g|\le x\\ p_g\le L}} p_g = \sum_{p\le L} p \sum_{\substack{g\in\mathcal{G}\\ |g|\le x\\ p_g=p}} 1 &= 2x\sum_{p\le L} p\delta_p + O\Big(\sum_{p\le L} p(3^p + x^{1/2})\Big) \\ &= 2x\sum_{p\le L} p\delta_p + O(x^{1/2}L^2). \end{aligned} \tag{24}$$

We now extend the sum on p to infinity, using Lemma 4.1 to estimate the resulting error. Observe that

$$\delta_p < \prod_{r<p}\left(1 - \frac{\varphi(r-1)}{r}\right) = \prod_{r<p}\left(1 + \frac{\varphi(r-1)}{r-\varphi(r-1)}\right)^{-1}.$$

If $r > 2$, then $r-1$ is even, and $\varphi(r-1) \le \frac{r-1}{2}$. Hence, $r - \varphi(r-1) > \varphi(r-1)$, and the ratio $\frac{\varphi(r-1)}{r-\varphi(r-1)} < 1$. Using the inequality $1 + u \ge \exp(u/2)$ valid when $0 \le u \le 1$, we conclude from Lemma 4.1 that for all sufficiently large p,

$$\prod_{r<p}\left(1 + \frac{\varphi(r-1)}{r-\varphi(r-1)}\right) \ge \exp\left(\frac{1}{2}\sum_{2<r<p} \frac{\varphi(r-1)}{r-\varphi(r-1)}\right) \ge \exp(cp/\log p),$$

for $c := \frac{1}{3}\alpha$. As a consequence, $\delta_p \ll \exp(-cp/\log p)$ for all primes p, and

$$\sum_{p>L} p\delta_p \ll \exp\big(-\frac{c}{2}L/\log L\big) \ll \exp(-(\log x)^{1+o(1)}).$$

Referring back to (24), we deduce that

$$\sum_{\substack{g\in\mathcal{G}\\ |g|\le x\\ p_g\le L}} p_g = 2x\sum_{p} p\delta_p + O(x\exp(-(\log x)^{1+o(1)})).$$

Next, we bound the sum of the p_g taken over $g \in \mathcal{G}$, $|g| \le x$, having $p_g > L$. If $p_g > L$, then g belongs to one of γM residue classes mod M, where

$$M := \prod_{r \le L} r, \quad \text{and} \quad \gamma := \prod_{r \le L} \left(1 - \frac{\varphi(r-1)}{r}\right).$$

The number of such g with $|g| \le x$ is $\ll \gamma(x + M) \ll \gamma x$, noting that $M \le 3^L = x^{o(1)}$. Moreover, essentially the same work used to estimate γ_p shows that $\gamma \le \exp(-cL/\log L)$. (All of this is being claimed for large enough values of x.) So by Theorem 1.2,

$$\sum_{\substack{g \in \mathcal{G} \\ |g| \le x \\ p_g > L}} p_g \le (\max_{\substack{g \in \mathcal{G} \\ |g| \le x}} p_g) \sum_{\substack{g \in \mathcal{G} \\ |g| \le x \\ p_g > L}} 1 \ll (\log x)^{16} (x \exp(-cL/\log L)) \ll x \exp(-(\log x)^{1+o(1)}).$$

Putting together the pieces,

$$\sum_{\substack{g \in \mathcal{G} \\ |g| \le x}} p_g = 2x \sum_p p\delta_p + O(x/\exp((\log x)^{1+o(1)})).$$

Corollary 1.3 follows; in fact, the ratio appearing on the left in (7) is $\sum_p p\delta_p + O(\exp(-(\log x)^{1+o(1)}))$.

5. An unconditional tamed average: Proof of Theorem 1.4

Our main tool for this proof will be Montgomery's "arithmetic large sieve" inequality introduced in Section 3. Using Montgomery's sieve, Vaughan showed [17] that $p_g \le N^{1/2}$ for all $g \in [M+1, M+N]$, apart from $O(N^{1/2}(\log N)^{1-\alpha})$ exceptions, where α is the constant of Lemma 4.1. Earlier Gallagher [5] had shown such a result with 1 in place of $1-\alpha$. The next proposition implies that $N^{1/2}$ can be replaced by a large power of $\log N$, if one is willing to slightly inflate the exponent $1/2$ on N in the size of the exceptional set.

Proposition 5.1. *Let $M, N \in \mathbb{Z}$ with $N > 100$. Let Y be a real number satisfying*

$$\log^2 N \le Y \le \exp\left(\log N \frac{\log\log\log N}{\log\log N}\right).$$

The count of integers g in $[M+1, M+N]$ with $p_g > Y$ does not exceed

$$N^{1/2} \exp\left(O\left(\log N \frac{\log\log\log N}{\log\log N}\right)\right) \cdot \exp(u \log u),$$

where $u := \frac{1}{2}\frac{\log N}{\log Y}$. Here the O-constant is absolute.

Note that if $Y = \log^K N$ for a fixed $K \ge 1$, then the upper bound in the conclusion of Proposition 5.1 assumes the form $N^{\frac{1}{2}(1+1/K)+o(1)}$, as $N \to \infty$.

Proof of Proposition 5.1. We may assume N is sufficiently large. We apply the arithmetic large sieve with $Q = N^{1/2}$, taking $\nu(p) = \varphi(p-1)$ for $p \le Y$, and $\nu(p) = 0$ for $Y < p \le Q$. It suffices to show that with these choices of parameters, the denominator

$$J = \sum_{n \le N^{1/2}} \mu^2(n) \prod_{p \mid n} \frac{\varphi(p-1)}{p - \varphi(p-1)}. \tag{25}$$

in the sieve bound satisfies

$$J \geq N^{1/2} \exp\left(O\left(\log N \frac{\log\log\log N}{\log\log N}\right)\right) \cdot \exp(-u\log u). \tag{26}$$

Let R be the number of primes $p \in [\frac{1}{2}Y, Y]$ for which the smallest prime factor of $\frac{p-1}{2}$ exceeds $Y^{1/5}$. By the linear sieve and the Bombieri–Vinogradov theorem,

$$R > Y/\log^3 Y.$$

(We need for the application of Bombieri–Vinogradov that $\frac{1}{5} < \frac{1}{4}$.) For each such p, the ratio $\frac{\varphi(p-1)}{p-1} = \frac{1}{2}\prod_{\ell \mid p-1,\ \ell>2}(1-1/\ell) > \frac{1}{2}(1-y^{-1/5})^4 > 2/5$ (say). Hence, $\frac{\varphi(p-1)}{p} > \frac{1}{3}$, and $\frac{\varphi(p-1)}{p-\varphi(p-1)} > \frac{1}{2}$. Let $u_0 = \lfloor \log(N^{1/2})/\log Y \rfloor$ (so that $u_0 = \lfloor u \rfloor$, with u as in the proposition). By considering the contribution to the right-hand side of (25) from products of u_0 distinct primes p of the above kind, we see that $J \geq 2^{-u_0}\binom{R}{u_0}$. Now $R > Y/(\log Y)^3 > (\log N)^{3/2} > u_0$. Since $\binom{n}{k} \geq (n/k)^k$ for each pair of integers n, k with $n \geq k > 0$, we conclude that

$$\frac{1}{2^{u_0}}\binom{R}{u_0} \geq (R/2u_0)^{u_0} \geq (R/2)^{u_0}\exp(-u\log u).$$

Furthermore,

$$(R/2)^{u_0} \geq (R/2)^{u-1} \geq Y^{u-1}(2(\log Y)^3)^{-u} = N^{1/2}Y^{-1}(2(\log Y)^3)^{-u}.$$

The assumed bounds on Y ensure that $Y^{-1}(2(\log Y)^3)^{-u} = \exp\left(O\left(\log N \frac{\log\log\log N}{\log\log N}\right)\right)$. Our desired lower estimate (26) follows by combining the last two displays. □

Proof of Theorem 1.4. Fix $K \geq 2$ with $\eta + \frac{1}{2}(1+1/K) < 1$. We first consider the contribution of $g \in \mathcal{G}$, $|g| \leq x$, having $p_g \leq \log^K(3x)$. Note that the corresponding summand $\min\{p_g, x^{\frac{1}{2}-\varepsilon}\} = p_g$ for these values of g (once x exceeds a certain constant depending only on K).

In the course of proving Corollary 1.3, we showed that with $L = \log x/\log\log x$,

$$\sum_{\substack{g\in\mathcal{G},\ |g|\leq x\\ p_g \leq L}} p_g = 2x\sum_p p\delta_p + O(x\exp(-(\log x)^{1+o(1)})).$$

Furthermore, the count of $g \in \mathcal{G}$, $|g| \leq x$ with $p_g > L$ is $O(x\exp(-cL/\log L))$. Hence,

$$\sum_{\substack{g\in\mathcal{G},\ |g|\leq x\\ L<p_g\leq \log^K(3x)}} p_g \ll x\log^K(3x)\exp(-cL/\log L) \ll x\exp(-(\log x)^{1+o(1)}).$$

Therefore, the theorem will be proved if is shown that

$$\sum_{\substack{g\in\mathcal{G},\ |g|\leq x\\ p_g>\log^K(3x)}} \min\{p_g, x^{\frac{1}{2}-\varepsilon}\} = o(x),$$

as $x \to \infty$. For this we apply Proposition 5.1. Choose M and N with $M+1 = -\lfloor x \rfloor$ and $M+N = \lfloor x \rfloor$; then $[M+1, M+N]$ is the set of all integers g with $|g| \leq x$, and $N = 2\lfloor x \rfloor + 1 < 3x$. Thus, if $p_g > \log^K(3x)$, then $p_g > \log^K N$. By Proposition 5.1, the number of such g, $|g| \leq x$, is at most $x^{\frac{1}{2}(1+1/K)+o(1)}$. It follows that the sum appearing in the last display is bounded above by $x^\eta \cdot x^{\frac{1}{2}(1+1/K)+o(1)}$, which is $o(x)$ by our choice of K. □

6. Almost-primitive roots

In the statement and proof of Theorem 1.4, there is no need to restrict g to $\mathcal{G}$; all the arguments work just as well if we average $\min\{p_g, x^{\frac{1}{2}-\varepsilon}\}$ over all g, $|g| \le x$. For this unrestricted average, the exponent $\frac{1}{2}$ in the cutoff is optimal, in that even, square values of g push the average of $\min\{p_g, x^{\frac{1}{2}+\varepsilon}\}$ to infinity. One could hope to transcend $\frac{1}{2}$ after restoring the condition $g \in \mathcal{G}$, but it is not clear how to work that $g \in \mathcal{G}$ into the proof of a result like Proposition 5.1.

Recall from the introduction that g is called an almost-primitive root mod p when g generates a subgroup of $(\mathbb{Z}/p\mathbb{Z})^\times$ of index at most 2. Define p_g^* analogously to p_g but with "almost-primitive root" in place of "primitive root." We then expect that $p_g^* < \infty$ for *every* nonzero $g \in \mathbb{Z}$. This seems difficult to establish unconditionally, but it can be seen to follow from GRH by a slight modification of Hooley's argument.

Our final theorem is an upper bound on the frequency of large values of p_g^*, strong enough to imply that $\min\{p_g^*, x^{1-\varepsilon}\}$ has its expected mean value.

Theorem 6.1. *For all $x \ge 2$, there are $O(\log^3 x)$ integers g, $|g| \le x$, with $p_g^* > \log^4 x$.*

Most of this section will be devoted to the proof of Theorem 6.1, but we start with a few words about the application of this theorem to the average of p_g^*. Put $F(p) = 1_{p>2}\varphi(\frac{p-1}{2}) + \varphi(p-1)$, so that $F(p)$ is the number of almost primitive roots mod p. Let

$$\delta_p^* = \frac{F(p)}{p}\prod_{r<p}\left(1 - \frac{F(r)}{r}\right).$$

Reasoning as in the introduction, we expect p_g^* to have mean value $\sum_p p\delta_p^*$. Under GRH this could be proved analogously to our Corollary 1.3. Using Theorem 6.1, we obtain (unconditionally) that for each positive $\varepsilon \in (0,1)$, the average of $\min\{p_g^*, x^{1-\varepsilon}\}$ tends to $\sum_p p\delta_p^*$. For this, follow the argument for Theorem 1.4 but plug in Theorem 6.1 in place of Proposition 5.1.

We turn now to the proof of Theorem 6.1. This requires a new ingredient, Gallagher's "larger sieve" (see [6] or [4, §9.7]).

Larger sieve. *Let $N \in \mathbb{N}$, and let $\mathcal{S}$ be a finite set of prime powers. Suppose that all but $\overline{\nu}(q)$ residue classes mod q are removed for each $q \in \mathcal{S}$. Then among any N consecutive integers, the number remaining unsieved does not exceed*

$$\left(\sum_{q\in\mathcal{S}}\Lambda(q) - \log N\right)\Big/\left(\sum_{q\in\mathcal{S}}\frac{\Lambda(q)}{\overline{\nu}(q)} - \log N\right), \tag{27}$$

as long as the denominator is positive.

We call $\theta \in (0,1)$ admissible if, for all large enough values of Y, we have

$$\#\{p \le Y : P^-(\frac{p-1}{2}) > Y^\theta\} \gg Y/\log^2 Y.$$

(The implied constant here is allowed to depend on θ.) As remarked in the proof of Proposition 5.1, the Bombieri–Vinogradov theorem in conjunction with the linear sieve implies that any $\theta < \frac{1}{4}$ is admissible. It is known how to do a little better; for instance, [4, Theorem 25.11] shows that

$\theta = 3/11$ is admissible. Any admissible $\theta > 0$ would yield a version of Theorem 6.1; to obtain the clean exponents of 3 and 4, we will use the existence of an admissible $\theta > \frac{1}{4}$.

Proof of Theorem 6.1. Fix an admissible $\theta > \frac{1}{4}$. We let x be large, and we sieve the $N := 2\lfloor x \rfloor + 1$ integers in the interval $[-x, x]$. With $y := \log^4 x$, let

$$\mathcal{S} = \left\{ \text{primes } p : 3 < p \le y, P^-\left(\frac{p-1}{2}\right) > y^\theta \right\},$$

so that

$$\#\mathcal{S} \gg \frac{y}{\log^2 y}.$$

(Here $P^-(\cdot)$ denotes the smallest prime factor.) For each $p \in \mathcal{S}$, we remove every residue class *except* 0 mod p and the classes corresponding to integers whose order mod p does not exceed

$$z := y^{1-\theta}.$$

Then, in the notation of the larger sieve,

$$\overline{\nu}(p) = 1 + \sum_{\substack{f \mid p-1 \\ f \le z}} \varphi(f). \tag{28}$$

Suppose the integer g, $|g| \le x$, is removed in the sieve. Then there is a prime $p \in \mathcal{S}$ not dividing g for which the order ℓ (say) of g mod p exceeds z. Then $\frac{p-1}{\ell} < y/z = y^\theta$, while every odd divisor of $p-1$ exceeds y^θ. Thus (keeping in mind that $p \equiv 3 \pmod 4$), $\ell = \frac{p-1}{2}$ or $p-1$, meaning that g is an almost-primitive root mod p. In particular, $p_g^* \le y$.

Hence, the number of g, $|g| \le x$, with $p_g^* > y$ is bounded above by the count of unsieved integers, which can be approached with the larger sieve. The arguments below draw inspiration from Gallagher's proof of Theorem 2 in [6].

By the Cauchy–Schwarz inequality,

$$\left(\sum_{p \in \mathcal{S}} \frac{\log p}{\overline{\nu}(p)}\right)\left(\sum_{p \in \mathcal{S}} \overline{\nu}(p) \log p\right) \ge \left(\sum_{p \in \mathcal{S}} \log p\right)^2 \gg (\log(y) \cdot \#\mathcal{S})^2 \gg y^2/\log^2 y.$$

(We use here that $\log p \gg \log y$ for each $p \in \mathcal{S}$, which follows from $P^-(\frac{p-1}{2}) > y^\theta$.) On the other hand, referring back to (28),

$$\begin{aligned} \sum_{p \in \mathcal{S}} \overline{\nu}(p) \log p &\le \sum_{p \in \mathcal{S}} \log p + \sum_{f \le z} \varphi(f) \sum_{\substack{p \in \mathcal{S} \\ p \equiv 1 \pmod f}} \log p \\ &\ll (\log y)\#\mathcal{S} + \log y \sum_{f \le z} \varphi(f) \#\{p \in \mathcal{S} : p \equiv 1 \pmod f\}. \end{aligned}$$

Brun's sieve implies that $\#\mathcal{S} \ll y/\log^2 y$. Brun's sieve also handles the counts in appearing in the sum on f: If $p \in \mathcal{S}$, $p \equiv 1 \pmod f$, and $p > y^\theta$, then $t := \frac{p-1}{f} < y/f$, and both $tf + 1, t$ have no odd prime factors up to y^θ. The sieve shows that the number of such t is

$$\ll \frac{y}{f} \prod_{2 < r \le y^\theta} \left(1 - \frac{1 + 1_{r \nmid f}}{r}\right) \ll \frac{y}{f \log^2 y} \prod_{r \mid f} \left(1 - \frac{1}{r}\right)^{-1} = \frac{y}{\varphi(f) \log^2 y}.$$

Since there are trivially at most y^θ/f primes up to y^θ in the residue class 1 mod f,

$$\#\{p \in \mathcal{S} : p \equiv 1 \pmod f\} \ll \frac{y}{\varphi(f)\log^2 y},$$

and

$$\log y \sum_{f \le z} \varphi(f) \#\{p \in \mathcal{S} : p \equiv 1 \pmod f\} \ll yz/\log y.$$

We conclude that

$$\sum_{p \in \mathcal{S}} \overline{\nu}(p) \log p \ll yz/\log y,$$

and hence

$$\sum_{p \in \mathcal{S}} \frac{\log p}{\overline{\nu}(p)} \gg \frac{y^2/\log^2 y}{yz/\log y} = \frac{y}{z}\frac{1}{\log y} = \frac{y^\theta}{\log y}.$$

Since $y = \log^4 x$ and $\theta > \frac{1}{4}$, this last expression is of larger order than $\log N$, and the denominator in (27) is $\gg y^\theta/\log y$. The numerator in (27) is $\ll (\log y)\#\mathcal{S} \ll y/\log y$. Therefore, the number of unsieved g, $|g| \le x$, is

$$\ll \frac{y/\log y}{y^\theta/\log y} = y^{1-\theta}.$$

By our choices of y and θ, this last expression is $o(\log^3 x)$. □

Remark. Fix an admissible $\theta \in (0,1)$, and set $y = ((\log x)(\log\log x)^2)^{1/\theta}$. A slight tweak to the above argument shows that there are $O(y^{1-\theta})$ integers g, $|g| \le x$, with $p_g^* > y$. Taking $\theta = 3/11$, we can replace the exponents 3 and 4 in Theorem 6.1 with $8/3 + o(1)$ and $11/3 + o(1)$, respectively. It is probably the case that every $\theta \in (0,1)$ is admissible; if so, those exponents can be brought arbitrarily close to 0 and 1.

Acknowledgments

P.P. is supported by NSF award DMS-2001581. The authors thank Igor Shparlinski for a stimulating email message that prompted these investigations. A reply of Andrew Granville on MathOverflow was also useful [9].

References

1. E. Artin and H. Hasse, *Emil Artin und Helmut Hasse: Die Korrespondenz 1923–1934*, Universitätsverlag Göttingen, Göttingen, 2008, edited, with commentary and introduction in English by Günter Frei and Peter Roquette in collaboration with Franz Lemmermeyer.
2. D. A. Burgess, *On character sums and primitive roots*, Proc. London Math. Soc. (3) **12** (1962), 179–192.
3. P. D. T. A. Elliott and L. Murata, *On the average of the least primitive root modulo p*, J. London Math. Soc. (2) **56** (1997), 435–454.
4. J. Friedlander and H. Iwaniec, *Opera de cribro*, American Mathematical Society Colloquium Publications, vol. 57, American Mathematical Society, Providence, RI, 2010.
5. P. X. Gallagher, *The large sieve*, Mathematika **14** (1967), 14–20.
6. ———, *A larger sieve*, Acta Arith. **18** (1971), 77–81.
7. D. R. Heath-Brown, *Artin's conjecture for primitive roots*, Quart. J. Math. Oxford Ser. (2) **37** (1986), 27–38.
8. C. Hooley, *On Artin's conjecture*, J. Reine Angew. Math. **225** (1967), 209–220.
9. A. Granville (https://mathoverflow.net/users/70432/andrew granville), *Least prime in Artin's primitive root conjecture*, MathOverflow, URL:https://mathoverflow.net/q/386952 (version: 2021-03-19).
10. H. L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Mathematics, vol. Vol. 227, Springer-Verlag, Berlin-New York, 1971.

11. H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2006.
12. P. Moree, *Artin's primitive root conjecture—a survey*, Integers **12** (2012), 1305–1416.
13. C. Pomerance and I. E. Shparlinski, *On pseudosquares and pseudopowers*, Combinatorial number theory, Walter de Gruyter, Berlin, 2009, pp. 171–184.
14. J.-P. Serre, *Quelques applications du théorème de densité de Chebotarev*, Inst. Hautes Études Sci. Publ. Math. (1981), no. 54, 323–401.
15. V. Shoup, *Searching for primitive roots in finite fields*, Math. Comp. **58** (1992), 369–380.
16. P. Stevenhagen, *The correction factor in Artin's primitive root conjecture*, J. Théor. Nombres Bordeaux **15** (2003), 383–391.
17. R. C. Vaughan, *Some applications of Montgomery's sieve*, J. Number Theory **5** (1973), 64–79.
18. Y. Wang, *On the least primitive root of a prime*, Acta Math. Sinica **9** (1959), 432–441.

Department of Mathematics, University of Georgia, Athens, GA 30602

Email address: `Steve.Fan@uga.edu`

Email address: `pollack@uga.edu`