On generalizing Artin's conjecture on primitive roots to composite moduli

Shuguang Li and Carl Pomerance

1. Introduction

For a given integer a, a natural question is whether there are infinitely many primes p with a as a *primitive root*; that is, the residue a generates the cyclic multiplicative group $(\mathbf{Z}/p\mathbf{Z})^*$. Let $P_a(x)$ denote the number of primes $p \le x$ which have a as a primitive root. Clearly, a necessary condition on a for $P_a(x)$ to be unbounded is that $a \ne -1$ and a is not a square. Artin conjectured that these are the only exceptional values of a, and in fact, if a is not -1 nor a square, then there is a positive proportion of primes p with primitive root a. That is, there is a positive number $A(a)$ such that

$$P_a(x) \sim A(a)\pi(x).$$

After a clarification of Heilbronn suggested by some numerical experiments of Lehmer, there is even a precise formula for the conjectured density $A(a)$. In a remarkable paper, Hooley [5] showed that this strong form of Artin's conjecture follows from the Generalized Riemann Hypothesis (GRH), in particular from the Riemann Hypothesis for the Galois closures of Kummerian fields of the form $\mathbf{Q}(a^{1/d})$.

For an arbitrary positive integer n, the group $(\mathbf{Z}/n\mathbf{Z})^*$ is not cyclic in general, but we may naturally generalize the notion of primitive root to an element which generates a maximal cyclic subgroup; that is, an element whose multiplicative order is as large as possible. This notion was suggested by Carmichael, and we use his notation $\lambda(n)$ for the order of the largest cyclic subgroup of $(\mathbf{Z}/n\mathbf{Z})^*$. Further, if $(a,n)=1$, we use the notation $l_a(n)$ for the order of a in $(\mathbf{Z}/n\mathbf{Z})^*$. Thus a is a primitive root for n if $l_a(n)=\lambda(n)$.

Let $N_a(x)$ denote the number of integers $n \le x$ such that $(a,n)=1$ and a is a primitive root for n. In analogy to Artin's conjecture, one might guess that if a is not in some exceptional set, then there is a positive number $B(a)$ such that

$$N_a(x) \sim B(a)x.$$

Let $\mathcal{E}$ denote the set of integers which are a power with an exponent larger than 1, or a square times either -1 or ± 2. It is shown in Li [9] that if $a \in \mathcal{E}$, then $N_a(x) = o(x)$. (In particular, if a is an h-th power, then on a set of integers n of asymptotic density 1, either $(a,n)>1$ or $l_a(n) \mid \lambda(n)/h$, and if a is a square times -1 or ± 2, then on a set of asymptotic density 1, either $(a,n)>1$ or $l_a(n) \mid \lambda(n)/2$.) Further, it is shown in the same paper that for *any* integer a,

$$\liminf_{x\to\infty} N_a(x)/x = 0.$$

So, the above guess of an analogue to Artin's conjecture is not true. However, the conjecture was made in [9] that if $a \notin \mathcal{E}$, then

$$\limsup_{x\to\infty} N_a(x)/x > 0.$$

It is the goal of this paper to prove this conjecture on assumption of the GRH.

For a fixed prime q such that a is not a q-th power, it follows from the Chebotarev density theorem that the relative density of the primes p such that $p \equiv 1 \pmod q$ and a is a q-th power in $(\mathbf{Z}/p\mathbf{Z})^*$ is $1/(q(q-1))$. So if a is not a q-th power for any q, we would heuristically expect that the relative density of the primes p such that a is a primitive root for p is $\prod_q (1 - 1/(q(q-1)))$. In fact, this infinite product, known as Artin's constant, tells most of the story for the number $A(a)$ in Artin's conjecture, and this heuristic forms the basis of Hooley's GRH-conditional proof. However, the GRH is not needed for small primes q. That is, if $\psi(x)$ tends to infinity very slowly, it can indeed be shown via the Chebotarev theorem that $A(a)$ gives the correct relative density of the set of primes p such that $(p-1)/l_a(p)$ has no prime factor $\le \psi(p)$.

In principle, it should not take much to finish the proof, since $\sum_{q>\psi(x)} 1/(q(q-1)) \to 0$. All one would need is an estimate of $\ll \pi(x)/q^2$ (or even a weaker estimate such as $\ll \pi(x)/(q \ln q)$) for the number of primes $p \le x$ with $q|(p-1)/l_a(p)$, uniformly for primes q with $\psi(x) < q \le x^{1/2-o(1)}$. Indeed, larger q's can be handled by other means. Hooley uses the GRH to obtain such an estimate and complete the proof of Artin's conjecture.

In analogy, let $F(q,x)$ denote the relative proportion, among those integers $n \le x$ that are coprime to a, for which $q|\lambda(n)/l_a(n)$. Then heuristically we should have

$$N_a(x) \sim \frac{\varphi(|a|)}{|a|} x \prod_q (1 - F(q,x)).$$

It should be noted though that $F(q,x)$ does not tend to a limit as $x \to \infty$. Rather, it oscillates between a dangerously large order of $1/q$ and a safely small order of $1/(q \ln q)$ and smaller. It is through this oscillation that Li [9] was able to show that $\liminf N_a(x)/x = 0$. This unconditional result can be achieved by looking at the combined affect of small primes, where one can get by with the Chebotarev theorem. To show that $\limsup N_a(x)/x$ is positive when a is not in $\mathcal{E}$, we have to show too that the larger primes do not pose too great an influence, and as with Hooley, this can be shown conditionally on the GRH. To complete the proof, as with Hooley, we use a sieve on the small primes, though this part of our argument (which is unconditional) is the most intricate. However, in one respect, the situation is now simpler. We are able to show (unconditionally) a principle of "separation of powers" for most numbers n: For all numbers n, but for a set of asymptotic density 0, if q_1, q_2 are different fixed primes, and if $q_1^{j_1} \| \lambda(n)$, $q_2^{j_2} \| \lambda(n)$, then n is not divisible by any prime $p \equiv 1 \pmod{q_1^{j_1} q_2^{j_2}}$. This principle suggests that there is a universal "Artin constant" for all numbers $a \notin \mathcal{E}$ rather than some varying $A(a)$ as in Artin's conjecture for primes.

We prove the following theorem.

Theorem *On assumption of the GRH, there is a positive number A such that if a is an integer with $a \notin \mathcal{E}$, then*

$$\limsup_{x\to\infty} N_a(x)/x \ge A\varphi(|a|)/|a|.$$

In particular there is an unbounded set S of positive reals such that for any $a \notin \mathcal{E}$,

$$\liminf_{x\to\infty,\, x\in S} N_a(x)/x \ge A\varphi(|a|)/|a|.$$

We make as a conjecture a somewhat stronger assertion:

Conjecture *For each prime q, let*

$$F_q \;=\; \liminf_{t\to\infty} \sum_{j=0}^{\infty} \frac{\exp(tq^{-j-1})-1}{\exp(t\varphi(q^j)^{-1})},$$

and let

$$\alpha \;=\; \prod_q (1-F_q).$$

If a is an integer with $a \notin \mathcal{E}$, then

$$\limsup_{x\to\infty} N_a(x)/x = \alpha\varphi(|a|)/|a|,$$

and this $\limsup$ *is attained on a set of numbers x that is independent of the choice of a.*

It is not so hard to show that the number α in the conjecture is positive. We have ascertained that $\alpha \approx 0.326$. One might wonder if the conjecture can be proved on assumption of the GRH, but we have not been successful in this regard. It may be possible to prove the conjecture on assumption of the GRH and the assertion that if $p_1, p_2, \ldots, p_k$ are distinct primes, then the numbers $1/\ln p_1, 1/\ln p_2, \ldots, 1/\ln p_k$ are linearly independent over the field of rational numbers. This assertion is itself a corollary of Schanuel's conjecture in transcendental number theory: If the complex numbers $z_1, \ldots, z_k$ are linearly independent over the rationals, then the set $\{z_1, \ldots, z_k, e^{z_1}, \ldots, e^{z_k}\}$ has transcendence degree k. Applied to $z_j = \ln p_j$, we get that $\ln p_1, \ldots, \ln p_k$ are algebraically independent, so that $1/\ln p_1, \ldots, 1/\ln p_k$ are linearly independent.

Though the function $l_a(n)$ is natural, ubiquitous, and useful, there are not very many nontrivial results concerning it. We mention a few. In addition to [9], the first author considers in [8], the number of residues a modulo n with $l_a(n) = \lambda(n)$, and considers in [10], the average order of $N_a(x)$ as a varies. Martin [11] considers the least prime primitive root modulo n. The second author, in [15], shows that on a set of integers n with upper asymptotic density 1, we have $n\sum_{l_a(d)=n} 1/d = o(1)$, disproving a conjecture of Erdős. It follows from a 1934 paper of Romanoff that the series $\sum 1/(nl_a(n))$ is convergent for any fixed integer a with $|a| > 1$. In [12], Murty, Rosen and Silverman obtain a strengthening of Romanoff's theorem. Among other results, Murty and Saidak [13] and Saidak [16] show on the GRH that there is an Erdős–Kac theorem for the number of prime divisors of $l_a(n)$, with mean $\frac{1}{2}(\ln\ln n)^2$ and standard deviation $\frac{1}{\sqrt{3}}(\ln\ln n)^{3/2}$, thus conditionally settling a conjecture in Erdős and Pomerance [1]. (We show in the next section that this Murty–Saidak theorem follows as a corollary of a theorem in [1] and some of the tools we develop for our main theorem.) Kurlberg [6] has recently shown on the GRH that for a given integer a with $|a| > 1$, the set of integers n which are coprime to a and for which $l_a(n) \le n^{1-\epsilon}$, has density 0, for each fixed $\epsilon > 0$. (Again, in the next section we show a somewhat stronger statement as a corollary of the tools we develop and the paper [2].) One possible setting where primitive roots for composite moduli have some use lies in pseudorandom number generators, for example see [3].

In the sequel, we denote by $\ln_k x$ the k-fold iteration of the natural logarithm applied to the number x, when x is sufficiently large that this value exceeds 1; and we let $\ln_k x = 1$ otherwise. We suppose that $\psi(x)$ is an arbitrary function that tends to infinity with x, but that $\psi(x) = o(\ln_4 x)$. In the following results, implicit constants depend at most on the choice of a. The letters p, q will always denote primes. As we have already been doing, we often use the Vinogradov order notation $f(x) \ll g(x)$ when it is the case that $f(x) = O(g(x))$.

2. Large primes

Fix an integer a with $|a| > 1$.

Lemma 1 *The number of integers $n \le x$ divisible by a prime $p > \psi(x)$ with $l_a(p) < p^{1/2}/\ln p$ is $\ll x/\ln\psi(x)$.*

Proof Consider primes p in $(T, 2T]$ with $l_a(p) < p^{1/2}/\ln p$. Each of these primes divides some number $a^j - 1$ with $j < 2T^{1/2}/\ln T$. But the number of distinct prime factors of $a^j - 1$ is $< 2j\ln|a|$, so that the number of such primes is

$$< \sum_{j<2T^{1/2}/\ln T} 2j\ln|a| < \frac{4T\ln|a|}{\ln^2 T}.$$

Hence the number of integers $n \le x$ that are divisible by such a prime in $(T, 2T]$ is $< 4x(\ln|a|)/\ln^2 T$. Now summing dyadically, that is letting T run through the numbers $2^i\psi(x)$ for $i = 0, 1, \dots$ and summing, we get the lemma.

Lemma 2 *The number of integers $n \le x$ divisible by a prime $p \equiv 1 \pmod{q}$ with*

$$\frac{q^2}{4\ln^2 q} < p \le q^2\ln^4 q,$$

is $\ll x(\ln_2 q)/(q\ln q)$.

Proof By the Brun-Titchmarsh inequality, the number of primes $p \le y$ with $p \equiv 1 \pmod{q}$ is $\ll y/(q\ln y)$ when $q < y^{2/3}$. So, if $q^2/(4\ln^2 q) < T < q^2\ln^4 q$, we get that the number of $n \le x$ divisible by a prime $p \in (T, 2T]$ with $p \equiv 1 \pmod{q}$ is $\ll x/(q\ln T)$. There are only $\ll \ln_2 q$ values of T to consider, so we have the lemma.

Lemma 3 (GRH) *Suppose that q is an odd prime and that a is not a q-th power. Let A_q denote the set of primes $p \equiv 1 \pmod q$ with $a^{(p-1)/q} \equiv 1 \pmod p$. The number of integers $n \le x$ divisible by a prime $p \in A_q$ with $p \ge q^2 \ln^4 q$ is $\ll x/(q \ln q) + x(\ln_2 x)/q^2$.*

Proof Using equation (28) in Hooley [5] (which relies on the GRH), we have that the number of primes $p \in A_q$ with $p \le y$ is

$$\frac{\mathrm{li}\,(y)}{q\varphi(q)} + O(y^{1/2}\ln(qy)).$$

For $q^2 \ln^4 q < T < q^4 \ln^4 q$, the number of primes $p \in A_q \cap (T, 2T]$ is therefore $\ll T^{1/2} \ln q$, so that the sum of their reciprocals is $\ll T^{-1/2} \ln q$. Summing dyadically over choices for T, we get that the number of $n \le x$ divisible by a prime $p \in A_q \cap (q^2 \ln^4 q, q^4 \ln^4 q]$ is $\ll x/(q \ln q)$.

If $T > q^4 \ln^4 q$, the number of primes $p \in A_q \cap (T, 2T]$ is $\ll T/(q^2 \ln T)$. So summing dyadically for T up to x, we get that the number of $n \le x$ divisible by a prime $p \in A_q$ with $p > q^4 \ln^4 q$ is $\ll x(\ln_2 x)/q^2$.

Let $\mathrm{P}(m)$ denote the largest prime factor of m, when m is an integer greater than 1, and let $\mathrm{P}(1) = 1$.

Proposition 1 (GRH) *Let a be an integer with $|a| > 1$. The number of integers $n \le x$ with $(a, n) = 1$ and*

$$\mathrm{P}\left(\frac{\lambda(n)}{l_a(n)}\right) \ge \ln_2 x$$

is $o(x)$.

Proof This result follows from Lemmas 1–3. Suppose $n \le x$, $(a, n) = 1$, and $q = \mathrm{P}(\lambda(n)/l_a(n)) \ge \ln_2 x$. Since we may assume that x is large, it thus follows that a is not a q-th power. Then either $q^2 | n$ or $p | n$ for some $p \in A_q$, where A_q is defined in Lemma 3. The number of $n \le x$ divisible by the square of a prime that is at least $\ln_2 x$ is $\ll x/(\ln_2 x \ln_3 x) = o(x)$, so we may assume that n is divisible by a prime $p \in A_q$. By Lemma 1, we may assume that $l_a(p) \ge p^{1/2}/\ln p$. But $l_a(p) \le (p-1)/q$, so that $p > q^2/(4 \ln^2 q)$. Thus, by Lemmas 2 and 3, the number of remaining values of $n \le x$ to be counted is

$$\ll x \sum_{q \ge \ln_2 x} \left(\frac{\ln_2 q}{q \ln q} + \frac{\ln_2 x}{q^2} \right).$$

This expression is $\ll x(\ln_4 x)/\ln_3 x = o(x)$, so the proposition is proved.

We are now in a position to give an alternate proof of the result of Murty and Saidak discussed in the introduction. Let $\omega(m)$ denote the number of distinct prime divisors of the natural number m.

Corollary 1 (GRH) (Murty–Saidak) *Let a be an integer with $|a| > 1$. The Erdős–Kac theorem holds for $\omega(l_a(n))$ with mean $\frac{1}{2}(\ln_2 n)^2$ and standard deviation $\frac{1}{\sqrt{3}}(\ln_2 n)^{3/2}$. That* *is, let*

$$K(x,u) = \frac{1}{2}(\ln_2 x)^2 + \frac{u}{\sqrt{3}}(\ln_2 x)^{3/2},$$
$$G(u) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u} e^{-t^2/2}\,dt.$$

Then for any real number u,

$$\sum_{\substack{n\le x,\ (a,n)=1\\ \omega(l_a(n))\le K(x,u)}} 1 = (1+o(1))G(u)\frac{\varphi(|a|)}{|a|}x,$$

as $x \to \infty$.

Proof The result follows from Proposition 1 and a result in Erdős–Pomerance [1] that

$$(1) \qquad \sum_{\substack{n\le x\\ \omega(\lambda(n))\le K(x,u)}} 1 = (1+o(1))G(u)x.$$

Indeed, it follows from Proposition 1 that, but for a set of integers n of asymptotic density 0,

$$\omega(l_a(n)) = \omega(\lambda(n)) + o(\ln_2 n),$$

as $n \to \infty$ through the set of integers coprime to a. Thus, the corollary will follow if we show that

$$\sum_{\substack{n\le x,\ (a,n)=1\\ \omega(\lambda(n))\le K(x,u)}} 1 = (1+o(1))G(u)\frac{\varphi(|a|)}{|a|}x.$$

But

$$\sum_{\substack{n\le x,\ (a,n)=1\\ \omega(\lambda(n))\le K(x,u)}} 1 = \sum_{d|a}\mu(d) \sum_{\substack{m\le x/d\\ \omega(\lambda(dm))\le K(x,u)}} 1.$$

Now $\omega(\lambda(dm)) = \omega(\lambda(m)) + O(1)$ so by (1),

$$\sum_{\substack{m\le x/d\\ \omega(\lambda(dm))\le K(x,u)}} 1 = (1+o(1))G(u)\frac{x}{d}.$$

Thus,

$$\sum_{\substack{n\le x,\ (a,n)=1\\ \omega(\lambda(n))\le K(x,u)}} 1 = (1+o(1))G(u)x\sum_{d|a}\frac{\mu(d)}{d} = (1+o(1))G(u)\frac{\varphi(|a|)}{|a|}x.$$

This completes the proof of the corollary.

Remark Let $\Omega(m)$ denote the number of prime factors of m with multiplicity, so that

$$\omega(l_a(n)) \le \Omega(l_a(n)) \le \Omega(\lambda(n)).$$

From [1], we have the Erdős–Kac theorem for $\Omega(\lambda(n))$ with the same mean and standard deviation as for $\omega(\lambda(n))$, so the same holds for $\Omega(l_a(n))$.

We now give a second corollary of Proposition 1, strengthing a result of Kurlberg [6].

Corollary 2 (GRH) *For any fixed integer a with $|a| > 1$, the number of positive integers $n \le x$ coprime to a and with $l_a(n) \le x/(\ln x)^{2\ln_3 x}$ is $o(x)$.*

Proof We use, in addition to Proposition 1, both the statement and the proof of Theorem 2 of [2]. This theorem asserts that on a set of integers n of asymptotic density 1, we have $\lambda(n) = n/(\ln n)^{\ln_3 n + O(1)}$ (actually, a somewhat stronger theorem is proved). Let $D_x(n)$ denote the largest divisor of $\lambda(n)$ divisible only by primes in $[1, \ln_2 x]$. From (7) and (19) in [2], we have that the number of integers $n \le x$ for which $D_x(n) > (\ln x)^{2\ln_4 x}$ is $o(x)$. Putting this estimate together with the theorem just quoted, we have that the number of $n \le x$ for which $\lambda(n)/D_x(n) \le x/(\ln x)^{\ln_3 x + 3\ln_4 x}$ is $o(x)$. But Proposition 1 asserts that but for $o(x)$ choices of $n \le x$, if n is coprime to a, we have $\lambda(n)/D_x(n)$ divides $l_a(n)$. This completes the proof.

Remark The proof actually allows us to replace "2" with "$1+\epsilon$" for any $\epsilon > 0$. In this form, the result is best possible, in that it is untrue with $\epsilon = 0$.

3. Intermediate primes

The goal of this section is to extend Proposition 1 to show that the set of integers n with $\mathrm{P}(\lambda(n)/l_a(n)) \ge \psi(n)$ has lower asymptotic density 0. Towards this end, we consider intermediate primes, namely those between $\psi(x)$ and $\ln_2 x$. However, some of our results also hold for the primes smaller than $\psi(x)$, for example, Lemma 4 below.

Fix an integer a not in the exceptional set $\mathcal{E}$. For q prime and k a positive integer, let

$$\begin{aligned} A_{q,k} &= \{p \text{ prime } : \ q^k \| p-1 \text{ and } a^{(p-1)/q} \equiv 1 \pmod{p}\} \\ B_{q,k} &= \{p \text{ prime } : \ q^k | l_a(p) \text{ or } q^{k+1} | p-1\}. \end{aligned}$$

Note that for each pair q, k, we have $A_{q,k}$ disjoint from $B_{q,k}$ and that

$$A_{q,k} \cup B_{q,k} \ = \ C_{q,k} := \ \{p \text{ prime } : q^k | p-1\}.$$

Lemma 4 *For each prime power $q^k < x$ we have*

$$\begin{aligned} \sum_{p \le x,\, p \in A_{q,k}} \frac{1}{p} &= \frac{\ln_2 x}{q^{k+1}} + O\left(\frac{\ln(q^k)}{q^k}\right), \\ \sum_{p \le x,\, p \in B_{q,k}} \frac{1}{p} &= \left(\frac{1}{\varphi(q^k)} - \frac{1}{q^{k+1}}\right) \ln_2 x + O\left(\frac{\ln(q^k)}{q^k}\right). \end{aligned}$$

Proof The first assertion is Corollary 3.5 in [9]. Next, it is shown in [14] that

$$\sum_{\substack{p \le x \\ p \equiv 1 \pmod{m}}} \frac{1}{p} = \frac{\ln_2 x + O(\ln m)}{\varphi(m)}$$

for all $x \ge m \ge 2$, with m integral. Applying this result with $m = q^k$ and using that $C_{q,k}$ is the disjoint union of $A_{q,k}$ and $B_{q,k}$, we get the second assertion of the lemma.

Remarks In this lemma it is important that $a \notin \mathcal{E}$. For example, if a is a square times ± 1 or ± 2 and $k \ge 3$, then $A_{2,k}$ has density 2^{-k} instead of 2^{-k-1}. If these densities were used in Proposition 4 in the next section, the expression $F_2(x)$ defined there would telescope to a quantity very near to 1. Similar telescoping would occur for $F_q(x)$ if a were a q-th power. This phenomenon lies behind our exceptional set $\mathcal{E}$. We remark that by using the GRH-conditional version of the Chebotarev density theorem in [7], we may establish a version of Lemma 4 where the expression $\ln(q^k)$ is replaced with $\ln_2(q^k)$. However, it seems interesting to eschew the use of the GRH where possible.

Lemma 5 *For each prime power q^k let $N(q^k)$ denote the number of integers $n \le x$ such that both*
(i) *n is divisible by some prime p in $A_{q,k}$ with $p \le x^{1-1/\ln_2 x}$,*
(ii) *n is not divisible by any prime in $B_{q,k}$.*
Then

$$N(q^k) \ll x \frac{\ln_2 x}{q^{k+1}} \exp\left(-\frac{\ln_2 x - \ln_3 x}{q^k}\right).$$

Proof Let P denote an arbitrary set of primes. By sieve methods (see Halberstam and Richert [4, Theorem 2.2, p. 68]), the number of integers $m \le y$ that are not divisible by any member of P is $\ll y \exp(-\sum_{p \in P,\, p \le y} 1/p)$. This estimate is uniform over all P and y. Note that a number n counted by $N(q^k)$ is of the form pm, where $p \in A_{q,k}$, $p \le x^{1-1/\ln_2 x}$, and m is not divisible by any member of $B_{q,k}$. Thus,

$$\begin{aligned} N(q^k) &\ll x \sum_{\substack{p \le x^{1-1/\ln_2 x} \\ p \in A_{q,k}}} \frac{1}{p} \exp\left(-\sum_{p' \in B_{q,k},\, p' \le x/p} \frac{1}{p'}\right) \\ &\le x \exp\left(-\sum_{p' \in B_{q,k},\, p' \le x^{1/\ln_2 x}} \frac{1}{p'}\right) \sum_{p \in A_{q,k},\, p \le x} \frac{1}{p}. \end{aligned}$$

By Lemma 4, we have

$$\sum_{p' \in B_{q,k},\, p' \le x^{1/\ln_2 x}} \frac{1}{p'} \ge \frac{\ln_2 x - \ln_3 x}{q^k} + O\left(\frac{\ln(q^k)}{q^k}\right).$$

Thus, by Lemma 4,

$$N(q^k) \ll x\frac{\ln_2 x}{q^{k+1}} \exp\left(-\frac{\ln_2 x - \ln_3 x}{q^k}\right)$$

and the lemma follows.

For each prime $q < \ln_2 x$, let $k_q = k_q(x)$ be the positive integer k which minimizes $|\ln_3 x - \ln(q^k)|$. (If there are two values of k at which the minimum is attained, then let k_q be the smaller of them.) Let

$$E(x) = \{q \text{ prime } : 3 < q < \ln_2 x,\ (\ln_2 x)/\ln q < q^{k_q} < \ln_2 x \ln q\}.$$

Proposition 2 *But for $o(x)$ numbers $n \le x$ with $(a,n) = 1$, every prime factor of $\lambda(n)/l_a(n)$ that lies in the interval $(\psi(x), \ln_2 x)$ is in $E(x)$.*

Proof Suppose $n \le x$ with $(a,n) = 1$. Fix a prime $q \notin E(x)$, and suppose $\ln_2 x > q > \psi(x)$, $q^{k_q} \ge \ln_2 x \ln q$. Thus, $q^{k_q - 1} \le (\ln_2 x)/\ln q$. Suppose $q^k \| \lambda(n)$ and $q | \lambda(n)/l_a(n)$. Then either $q^{k+1} | n$ or n is divisible by a prime in $A_{q,k}$. Further, n cannot be divisible by any prime in $B_{q,k}$. Thus, at least one of the following possibilities must occur:

1. $q^2 | n$;
2. for some $k \ge k_q$, n is divisible by a prime in $A_{q,k}$;
3. $\mathrm{P}(n) > x^{1-1/\ln_2 x}$;
4. for some $k \le k_q - 1$, (i) and (ii) of Lemma 5 hold.

Let N_1, N_2, N_3, N_4 denote the number of n that arise in the respective cases for some q with $\psi(x) < q < \ln_2 x$ and $q^{k_q} \ge \ln_2 x \ln q$.

We have

$$N_1 < \sum_{q > \psi(x)} \frac{x}{q^2} \ll \frac{x}{\psi(x)}.$$

By Lemma 4,

$$\begin{aligned} N_2 &\ll \sum_{q>\psi(x)} \sum_{k \ge k_q} x\left(\frac{\ln_2 x}{q^{k+1}} + \frac{\ln(q^k)}{q^k}\right) \\ &\ll x \sum_{q > \psi(x)} \frac{\ln_2 x}{q^{k_q+1}} + x \sum_{q^k > \ln_2 x} \frac{\ln(q^k)}{q^k} \\ &\ll x \sum_{q>\psi(x)} \frac{1}{q \ln q} + \frac{x}{(\ln_2 x)^{1/2}} \\ &\ll \frac{x}{\ln \psi(x)}. \end{aligned}$$

We have

$$N_3 \le \sum_{x^{1-1/\ln_2 x} < p \le x} \frac{x}{p} \ll \frac{x}{\ln_2 x}.$$

Using Lemma 5,

$$N_4 \ll \sum_{q>\psi(x)} \frac{x}{q} \sum_{k \le k_q - 1} \frac{\ln_2 x}{q^k} \exp\left(-\frac{\ln_2 x - \ln_3 x}{q^k}\right).$$

Let B_k denote the summand in the inner sum. For $2 \le k \le k_q - 1$ we have

$$\frac{B_{k-1}}{B_k} = q \exp\left(-(q-1)\frac{\ln_2 x - \ln_3 x}{q^k}\right).$$

Thus, using $q > \psi(x)$, $q^k \le \ln_2 x / \ln q$, we have $B_{k-1}/B_k < qe^{-q} < e^{-1} < 1$, so that $\sum B_k \ll B_{k_q - 1}$. Hence, using that the function $(1/t)\exp(-B/t)$ is increasing in the variable t when $0 < t \le B$, we have

$$\begin{aligned} N_4 &\ll \sum_{q>\psi(x)} \frac{x}{q} \frac{\ln_2 x}{q^{k_q-1}} \exp\left(-\frac{\ln_2 x - \ln_3 x}{q^{k_q-1}}\right) \\ &\le \sum_{q>\psi(x)} \frac{x}{q} \ln q \exp\left(-\ln q \frac{\ln_2 x - \ln_3 x}{\ln_2 x}\right) \\ &\le \sum_{q>\psi(x)} \frac{x \ln q}{q} \exp\left(-\frac{1}{2}\ln q\right) = \sum_{q>\psi(x)} \frac{x \ln q}{q^{3/2}} \\ &\ll \frac{x}{\psi(x)^{1/2}}. \end{aligned}$$

Thus, $N_1 + N_2 + N_3 + N_4 \ll x/\ln\psi(x) = o(x)$. A parallel argument handles the case when $q^{k_q} \le \ln_2 x/\ln q$, where in N_2 we now have $k \ge k_q + 1$ and in N_4 we have $k \le k_q$. This completes the proof of the proposition.

Lemma 6 *There is a set U of positive integers of asymptotic density 0, with the following property. For each prime q, the number of integers $n \le x$ with $n \notin U$, $(a,n) = 1$, and $q | \lambda(n)/l_a(n)$ is $\ll x/q$.*

Proof Let U be the set of integers n with $\mathrm{P}(n) > n^{1-1/\ln_2 n}$. As in the proof of Proposition 2, we see that U has asymptotic density 0. If $q|\lambda(n)$ then either n is divisible by q^2 or by a prime $p \equiv 1 \pmod{q}$. The number of $n \le x$ divisible by q^2 is $\le x/q^2 < x/q$. Thus, we may assume that n is not divisible by q^2. Let k_0 be the least integer k with $q^k \ge \ln_2 x$. Suppose $n \notin U$, $q^k \| \lambda(n)$, $q|\lambda(n)/l_a(n)$. Then n is divisible by a prime in $A_{q,k}$ and not divisble by any prime in $B_{q,k}$. By Lemma 5, the number of such $n \le x$ is

$$\ll \sum_k x \frac{\ln_2 x}{q^{k+1}} \exp\left(-\frac{\ln_2 x - \ln_3 x}{q^k}\right) = \sum_k B_k, \quad \text{say}.$$

Summing for $k \ge k_0$ we may ignore the exp expression, getting an estimate that is $\ll x/q$. As in the proof of Proposition 2, for $k < k_0$

$$\frac{B_{k-1}}{B_k} = q \exp\left(\frac{\ln_2 x}{q^{k-1}}\left(\frac{1}{q} - 1\right)\right) < q \exp\left(q\left(\frac{1}{q} - 1\right)\right) = \frac{q}{e^{q-1}} \le \frac{3}{e^2} < 1.$$

Thus, the sum of the expressions B_k for $k < k_0$ is $\ll B_{k_0-1}$. But as noticed in the proof of Proposition 2, the function $(1/t)\exp(-B/t)$ is increasing when $0 < t \le B$, so that

$$B_{k_0-1} = \frac{x \ln_2 x}{q} \cdot \frac{1}{q^{k_0-1}} \exp\left(-\frac{\ln_2 x - \ln_3 x}{q^{k_0-1}}\right) \le \frac{x \ln_2 x}{q} \cdot \frac{1}{\ln_2 x - \ln_3 x} \exp(-1) \ll \frac{x}{q}.$$

This completes the proof of the lemma.

Remark It is not important in the sequel, but a natural refinement in the proof of Lemma 6 gives $\ll (\varphi(|a|)/|a|)x/q$ in place of $\ll x/q$, with the implied constant being absolute.

We let $\exp_j$ denote the j-fold iteration of the function exp.

Lemma 7 *Let*

$$f(x) = \sum_{q \in E(x)} \frac{1}{q}, \quad f_\psi(x) = \sum_{q \in E(x),\, q > \psi(x)} \frac{1}{q}.$$

There is a positive number c and an unbounded set S of positive numbers x such that $f(x) < c$ for $x \in S$ and $f_\psi(x) \to 0$ as $x \to \infty$, $x \in S$.

Proof Note that a prime q is in $E(x)$ if and only if $3 < q < \ln_2 x$ and $\|(\ln_3 x)/\ln q\| < (\ln_2 q)/\ln q$, where $\|y\|$ denotes the distance of y from the nearest integer. For a positive integer m let

$$F(m) = f(\exp_3(m)) = \sum_{\| \frac{m}{\ln q} \| < \frac{\ln_2 q}{\ln q},\, 3 < q < e^m} \frac{1}{q},$$

$$F_\psi(m) = f_\psi(\exp_3(m)) = \sum_{\| \frac{m}{\ln q} \| < \frac{\ln_2 q}{\ln q},\, \psi(\exp_3(m)) < q < e^m} \frac{1}{q}.$$

We have that for each even integer $M \ge 2$,

$$\begin{aligned}\sum_{\frac{1}{2}M < m \le M} F(m) &= \sum_{3 < q < e^M} \frac{1}{q} \sum_{\substack{\| \frac{m}{\ln q} \| < \frac{\ln_2 q}{\ln q} \\ \ln q < m,\, \frac{1}{2}M < m \le M}} 1 \\ &\le \sum_{3 < q < e^M} \frac{1}{q} \sum_{\substack{\| \frac{m}{\ln q} \| < \frac{\ln_2 q}{\ln q} \\ \frac{1}{2}M < m \le M}} 1 \\ &\ll M \sum_{3 < q < e^M} \frac{\ln_2 q}{q \ln q},\end{aligned}$$

since the number of integers m satisfying

$$\left| \frac{m}{\ln q} - k \right| < \frac{\ln_2 q}{\ln q},$$

for a given integer k is $< 2\ln_2 q + 1 \ll \ln_2 q$. Since $\sum(\ln_2 q)/(q\ln q) \ll 1$, it follows that there is a number c_1 such that

$$\frac{1}{2M} \sum_{\frac{1}{2}M<m\le M} F(m) \le c_1$$

for all even integers $M \ge 2$. Similarly, there is a number c_2 such that

$$\frac{1}{2M} \sum_{\frac{1}{2}M<m\le M} F_\psi(m) \le c_2 \frac{\ln_2 \psi(\exp_3(M/2))}{\ln \psi(\exp_3(M/2))}.$$

Let $c = 3c_1$. There are at least $\frac{1}{6}M$ integers m in $(\frac{1}{2}M, M]$ with $F(m) < c$ and $F_\psi(m) < 3c_2 \ln_2 \psi(\exp_3(M/2))/\ln\psi(\exp_3(M/2))$. By throwing $\exp_3(m)$ into the set S for these numbers m, and then letting M run through powers of 2, we so create an unbounded set as called for in the lemma.

Proposition 3 *If x is in the set S described in Lemma 7, then but for $o(x)$ integers $n \le x$ with $(a, n) = 1$ we have that $\lambda(n)/l_a(n)$ is not divisible by any prime from the interval $(\psi(x), \ln_2 x)$.*

Proof This result follows immediately from Proposition 2, Lemma 6, and the second part of Lemma 7.

4. Small primes

Fix an integer $a \notin \mathcal{E}$. The interval

$$I(x) = \left(\frac{\ln_2 x}{2\psi(x)^2}, 4^{\psi(x)} \ln_2 x\right)$$

will play a crucial role in what follows.

Lemma 8 *But for $O(x/3^{\psi(x)})$ integers $n \le x$, for each prime $q \le \psi(x)$, if $q^k \| \lambda(n)$, then $q^k \in I(x)$ and $q^{k+1} \nmid n$.*

Proof Let $q \le \psi(x)$ be prime, let k_1 be the largest integer with $q^{k_1} \le (\ln_2 x)/(2\psi(x))$, and let k_2 be the least integer with $q^{k_2} \ge 4^{\psi(x)} \ln_2 x$. Note that $q^{k_1} > (\ln_2 x)/(2\psi(x)^2)$, so that $q^{k_1} \in I(x)$. If $\lambda(n)$ is not divisible by q^{k_1}, then n is not divisible by any prime $p \equiv 1 \pmod{q^{k_1}}$. The number of such $n \le x$ is, uniformly,

$$\begin{aligned} &\ll x \cdot \exp\left(- \sum_{p\le x,\ p\,\equiv\,1\ (\mathrm{mod}\ q^{k_1})} 1/p\right) \\ &\ll x \cdot \exp\left(-\frac{1}{\varphi(q^{k_1})} \ln_2 x\right) \\ &\le x \cdot \exp\left(-\frac{1}{q^{k_1}} \ln_2 x\right) \\ &\le x/e^{2\psi(x)}. \end{aligned}$$

On the other hand, if $q^{k_2}|\lambda(n)$, then n is either divisible by q^{k_2+1} or n is divisible by a prime $p \equiv 1 \pmod{q^{k_2}}$. The number of such $n \le x$ is, uniformly,

$$\begin{aligned}
&\le \frac{x}{q^{k_2+1}} + x \sum_{p \le x,\ p \ \equiv \ 1 \ (\mathrm{mod}\ q^{k_2})} \frac{1}{p} \\
&\ll \frac{x}{q^{k_2+1}} + \frac{x \ln_2 x}{q^{k_2}} \\
&\ll \frac{x}{4^{\psi(x)}}.
\end{aligned}$$

Summing over all choices of q with $q \le \psi(x)$ then shows that but for $O(x/3^{\psi(x)})$ numbers $n \le x$ we have for each prime $q \le \psi(x)$ an integer k such that $q^k \| \lambda(n)$ and $q^k \in I(x)$. Now, the number of $n \le x$ with $q^{k+1}|n$ is $\le x/q^{k+1} < 2x\psi(x)^2/\ln_2 x$. Summing over all possibilities for q^k gives the lemma.

Recall the notation from the start of section 3. Let C denote the union of all pairwise intersections $C_{q_1,j_1} \cap C_{q_2,j_2}$, where q_1, q_2 are different primes $\le \psi(x)$ and $q_1^{j_1}, q_2^{j_2} \in I(x)$. For prime $q \le \psi(x)$ and j with $q^j \in I(x)$, let

$$\begin{aligned}
C'_{q,j} &= C_{q,j} \setminus C \\
A'_{q,j} &= A_{q,j} \setminus C \\
B'_{q,j} &= B_{q,j} \setminus C.
\end{aligned}$$

Note that all of the sets $A'_{q,j}$ are disjoint, and that B'_{q_1,j_2} is disjoint from B'_{q_2,j_2} if $q_1 \ne q_2$, and the same for $C'_{q_1,j_1}, C'_{q_2,j_2}$. As before, $A'_{q,j}$ is disjoint from $B'_{q,j}$ and $A'_{q,j} \cup B'_{q,j} = C'_{q,j}$.

Lemma 9 *We have*

$$\sum_{p \le x,\, p \in C} \frac{1}{p} \ll \frac{\psi(x)^6}{\ln_2 x}.$$

In particular, the number of integers $n \le x$ divisible by a prime in C is $o(x)$.

Proof Let $q_1 < q_2 \le \psi(x)$ be primes and let j_1, j_2 be minimal with $q_1^{j_1}, q_2^{j_2} \in I(x)$. The sum of the reciprocals of the primes $p \equiv 1 \pmod{q_1^{j_1} q_2^{j_2}}$ with $p \le x$ is $\ll (\ln_2 x)/(q_1^{j_1} q_2^{j_2}) \le 4\psi(x)^4/\ln_2 x$, uniformly for all choices of q_1, q_2. Summing over all pairs q_1, q_2, we get the inequality in the lemma.

Let M denote the product of the primes $q \le \psi(x)$. For $d|M$, let $N_{a,d}(x)$ denote the number of integers $n \le x$ such that $(a, n) = 1$ and $d|\lambda(n)/l_a(n)$. And let $N_a^M(x)$ denote the number of integers $n \le x$ such that $(a, n) = 1$ and $\lambda(n)/l_a(n)$ is coprime to M.

Proposition 4 *Let*

$$F_q(x) = \sum_{j\,:\,q^j \in I(x)} \left(\exp\left(-\left(\frac{1}{\varphi(q^j)} - \frac{1}{q^{j+1}} \right) \ln_2 x \right) - \exp\left(-\frac{1}{\varphi(q^j)} \ln_2 x \right) \right).$$

Then

$$N_a^M(x) = \frac{\varphi(|a|)}{|a|} x \prod_{q \le \psi(x)} (1 - F_q(x)) \ + \ O\left(\frac{x}{(1.1)^{\psi(x)}} \right).$$

Proof By inclusion-exclusion, we have

$$N_a^M(x) \ = \ \sum_{d|M} \mu(d) N_{a,d}(x).$$

So, our intermediate goal will be to get a good estimate for $N_{a,d}(x)$.

Suppose $d|M$ has the prime factorization $q_1 q_2 \cdots q_t$. Let $\mathbf{j} = \mathbf{j}(d)$ denote a vector $(j_{q_1}, j_{q_2}, \ldots, j_{q_t})$, where each j_{q_i} is an integer with $q_i^{j_{q_i}} \in I(x)$. For such a vector $\mathbf{j}$, let $N_{a,d,\mathbf{j}}(x)$ denote the contribution to $N_{a,d}(x)$ from those integers n with $q^{j_q} \| \lambda(n)$ for each prime $q|d$. Then, from Lemma 8,

$$N_{a,d}(x) \ = \ \sum_{\mathbf{j}} N_{a,d,\mathbf{j}}(x) \ + \ O(x/3^{\psi(x)}).$$

We now turn to the estimation of an individual term $N_{a,d,\mathbf{j}}(x)$. For a set S of primes, we write $(n, S) = 1$ if n is not divisible by any member of S. Let $B'_{\mathbf{j}} = \cup_{q|d} B'_{q,j_q}$. Clearly, if n is counted by $N_{a,d,\mathbf{j}}(x)$, then $(n, B'_{\mathbf{j}}) = 1$. Further, but for the exceptional numbers n mentioned in the lemmas, n must be divisible by a prime from each A'_{q,j_q} for $q|d$. For $u|d$, let $A'_{u,\mathbf{j}} = \cup_{q|u} A'_{q,j_q}$. By the inclusion-exclusion principle and Lemma 9, we thus have

$$N_{a,d,\mathbf{j}}(x) \ = \ \sum_{u|d} \mu(u) \sum_{\substack{n \le x, \, (a,n)=1 \\ (n, A'_{u,\mathbf{j}} \cup B'_{\mathbf{j}})=1}} 1 \ + \ O(x\psi(x)^6 / \ln_2 x).$$

Let $y = x^{1/\ln_2 x}$. Suppose P is a set of primes none of which divide a. Let $P(y) = P \cap [1, y]$. We have

$$\sum_{\substack{n \le x, \, (a,n)=1 \\ (n,P)=1}} 1 \ = \ \sum_{\substack{n \le x, \, (a,n)=1 \\ (n,P(y))=1}} 1 + O\left(x \sum_{y < p \le x, \, p \in P} \frac{1}{p} \right).$$

Using the fundamental lemma of Brun's sieve, see Theorem 2.5, p. 82 in Halberstam and Richert [4], we have

$$\sum_{\substack{n \le x, \, (a,n)=1 \\ (n,P(y))=1}} 1 \ = \ \frac{\varphi(|a|)}{|a|} x \prod_{p \in P(y)} (1 - 1/p) \ + \ O(x/(\ln_2 x)^{\ln_2 x}),$$

for $P = A'_{u,\mathbf{j}} \cup B'_{\mathbf{j}}$. Further, for this choice of the set P of primes, since $\ln_2 x - \ln_2 y = \ln_3 x$, we have

$$\sum_{y < p \le x, \, p \in P} \frac{1}{p} \ \ll \ \frac{\psi(x)^2 \omega(d) \ln_3 x}{\ln_2 x},$$

where, as before, $\omega(d)$ is the number of distinct prime divisors of d. We thus have

$$\sum_{\substack{n\le x,\,(a,n)=1\\(n,A'_{u,\mathbf{j}}\cup B'_{\mathbf{j}})=1}} 1 =$$

$$\frac{\varphi(|a|)}{|a|}x\prod_{p\in(A'_{u,\mathbf{j}}\cup B'_{\mathbf{j}})(y)}(1-1/p) \;+\; O\left(\frac{\psi(x)^2\omega(d)\ln_3 x}{\ln_2 x}\right)$$

$$=\frac{\varphi(|a|)}{|a|}x\prod_{q|u}\prod_{p\in A'_{q,j_q}(y)}(1-1/p)\cdot\prod_{q|d}\prod_{p\in B'_{q,j_q}(y)}(1-1/p) \;+\; O\left(\frac{\psi(x)^2\omega(d)\ln_3 x}{\ln_2 x}\right).$$

Hence,

$$N_{a,d,\mathbf{j}}(x) =$$

$$\frac{\varphi(|a|)}{|a|}x\prod_{q|d}\prod_{p\in B'_{q,j_q}(y)}(1-1/p)\sum_{u|d}\mu(u)\prod_{q|u}\prod_{p\in A'_{q,j_q}(y)}(1-1/p) \;+\; O(2^{\omega(d)}3^{-\psi(x)}x)$$

$$=\frac{\varphi(|a|)}{|a|}x\prod_{q|d}\left(\prod_{p\in B'_{q,j_q}(y)}(1-1/p)-\prod_{p\in C'_{q,j_q}(y)}(1-1/p)\right) \;+\; O(2^{\omega(d)}3^{-\psi(x)}x).$$

For a given divisor d of M, the number of choices for vectors $\mathbf{j}$ is $<(3\psi(x))^{\omega(d)}$. Thus,

$$N_{a,d}(x) =$$

$$\frac{\varphi(|a|)}{|a|}x\sum_{\mathbf{j}}\prod_{q|d}\left(\prod_{p\in B'_{q,j_q}(y)}(1-1/p)-\prod_{p\in C'_{q,j_q}(y)}(1-1/p)\right) \;+\; O\left(\frac{(6\psi(x))^{\omega(d)}x}{3^{\psi(x)}}\right).$$

Now

$$\prod_{p\in P(y)}(1-1/p) \;=\; \exp\left(-\sum_{p\in P(y)}1/p\right) \;+\; O\left(\sum_{p\in P(y)}1/p^2\right).$$

Also, by Lemmas 4 and 9,

$$\sum_{p\in P(y)}1/p \;=\; \begin{cases}\left(\dfrac{1}{\varphi(q^j)}-\dfrac{1}{q^{j+1}}\right)\ln_2 y+O\left(\dfrac{\psi(x)^2\ln_3 x}{\ln_2 x}\right), & \text{if } P=B'_{q,j_q}\\[2ex] \dfrac{1}{\varphi(q^j)}\ln_2 y+O\left(\dfrac{\psi(x)^2\ln_3 x}{\ln_2 x}\right), & \text{if } P=C'_{q,j_q}.\end{cases}$$

Since $\ln_2 x=\ln_2 y+\ln_3 x$ and

$$\sum_{p\in P(y)}1/p^2 \;\ll\; 1/q^{2j} \;<\; 4\psi(x)^4/(\ln_2 x)^2$$

for $P = B'_{q,j_q}$ or $P = C'_{q,j_q}$, we have

$$\begin{aligned}
\prod_{p\in B'_{q,j_q}(y)} (1-1/p) - &\prod_{p\in C'_{q,j_q}(y)} (1-1/p) \\
&= \exp\left(-\left(\frac{1}{\varphi(q^{j_q})} - \frac{1}{q^{j_q+1}}\right)\ln_2 y + O\left(\psi(x)^2\frac{\ln_3 x}{\ln_2 x}\right)\right) \\
&\quad - \exp\left(-\frac{1}{\varphi(q^{j_q})}\ln_2 y + O\left(\psi(x)^2\frac{\ln_3 x}{\ln_2 x}\right)\right) \ + \ O\left(\frac{\psi(x)^4}{(\ln_2 x)^2}\right) \\
&= \exp\left(-\left(\frac{1}{\varphi(q^{j_q})} - \frac{1}{q^{j_q+1}}\right)\ln_2 y\right) \\
&\quad - \exp\left(-\frac{1}{\varphi(q^{j_q})}\ln_2 y\right) \ + \ O\left(\psi(x)^2\frac{\ln_3 x}{\ln_2 x}\right) \\
&= \exp\left(-\left(\frac{1}{\varphi(q^{j_q})} - \frac{1}{q^{j_q+1}}\right)\ln_2 x\right) \\
&\quad - \exp\left(-\frac{1}{\varphi(q^{j_q})}\ln_2 x\right) \ + \ O\left(\psi(x)^2\frac{\ln_3 x}{\ln_2 x}\right).
\end{aligned}$$

Let

$$F_{q,j}(x) = \exp\left(-\left(\frac{1}{\varphi(q^j)} - \frac{1}{q^{j+1}}\right)\ln_2 x\right) - \exp\left(-\frac{1}{\varphi(q^j)}\ln_2 x\right).$$

Thus, we have

$$\begin{aligned}
N_{a,d}(x) &= \frac{\varphi(|a|)}{|a|}x\sum_{\mathbf{j}}\prod_{q|d} F_{q,j_q}(x) \ + \ O\left(\frac{(6\psi(x))^{\omega(d)}x}{3^{\psi(x)}}\right) \\
&= \frac{\varphi(|a|)}{|a|}x\prod_{q|d}\ \sum_{j\,:\,q^j\in I(x)} F_{q,j}(x) \ + \ O\left(\frac{(6\psi(x))^{\omega(d)}x}{3^{\psi(x)}}\right) \\
&= \frac{\varphi(|a|)}{|a|}x\prod_{q|d} F_q(x) \ + \ O\left(\frac{(6\psi(x))^{\omega(d)}x}{3^{\psi(x)}}\right).
\end{aligned}$$

Then

$$\begin{aligned}
N_a^M(x) &= \sum_{d|M}\mu(d)N_{a,d}(x) \\
&= \frac{\varphi(|a|)}{|a|}x\sum_{d|M}\left(\mu(d)\prod_{q|d}F_q(x) \ + \ O\left(\frac{(6\psi(x))^{\omega(d)}x}{3^{\psi(x)}}\right)\right) \\
&= \frac{\varphi(|a|)}{|a|}x\prod_{q\le\psi(x)}(1-F_q(x)) \ + \ O\left(\frac{(7\psi(x))^{\pi(\psi(x))}x}{3^{\psi(x)}}\right) \\
&= \frac{\varphi(|a|)}{|a|}x\prod_{q\le\psi(x)}(1-F_q(x)) \ + \ O\left(\frac{x}{(1.1)^{\psi(x)}}\right),
\end{aligned}$$

since $3/e > 1.1$. This completes the proof of Proposition 4.

Proposition 5 (GRH) *If x is in the set S described in Lemma 7, then*

$$N_a(x) = \frac{\varphi(|a|)}{|a|} x \prod_{q \le \psi(x)} (1 - F_q(x)) \ + \ o(x).$$

Proof This result is an immediate corollary of Propositions 1, 3, and 4.

It remains now to estimate the product in Proposition 5 when $x \in S$. Towards this end we establish the following lemma.

Lemma 10 *We have $F_2(x) < 0.521$, $F_3(x) < 0.322$, and for $q \ge 5$ we have $F_q(x) < 0.27$. In addition we have $F_q(x) \ll 1/q$ for $q \in E(x)$, and we have $F_q(x) \ll 1/(q \ln q)$ for $q \notin E(x)$.*

Proof Let $f_q(t) = e^{-t}\left(e^{t(q-1)/q^2} - 1\right)$, so that

$$F_q(x) < \sum_{j=0}^{\infty} f_q\left(\varphi(q^j)^{-1} \ln_2 x\right).$$

Note that $f_q(t)$ achieves an absolute maximum on the positive reals at a critical value t_q that satisfies $e^{t_q(q-1)/q^2} = q^2/(q^2 - q + 1)$. The function $f_q(t)$ is increasing on $(0, t_q]$ and decreasing on $[t_q, \infty)$. Further, t_q is in the interval $(1, \ln q)$ for $q \ge 5$.

For $q = 2, 3$ we have

$$F_q(x) < \sum_{l=0}^{\infty} f_q(q^l t_q) + \sum_{l=0}^{\infty} f_q(q^{-l} t_q),$$

so it is then an easy matter to confirm the claimed inequalities for $F_2(x), F_3(x)$.

Now suppose $q \ge 5$. Then

$$F_q(x) \ < \ \sum_{l=0}^{\infty} f_q(q^l \ln q) \ + \ \sum_{l=0}^{\infty} f_q\left(q^{-l} (\ln q)^{-1}\right),$$

if $q \notin E(x)$, and

$$F_q(x) \ < \ f_q(t_q) \ + \ \sum_{l=0}^{\infty} f_q(q^l \ln q) \ + \ \sum_{l=0}^{\infty} f_q\left(q^{-l} (\ln q)^{-1}\right),$$

if $q \in E(x)$. Indeed, if $q \in E(x)$, then there is exactly one value of j with $\varphi(q^j)^{-1} \ln_2 x$ in $((\ln_2 x)/\ln q\,, \ln q \ \ln_2 x)$ and f_q at this value is at most $f_q(t_q)$.

Note that

$$f_q(t_q) \ < \ e^{-1}(q^2/(q^2 - q - 1) - 1) \ = \ (q-1)/(e(q^2 - q - 1).$$

Next note that for an integer $l \geq 0$,

$$f_q\left(\frac{1}{q^l \ln q}\right) \ < \ \exp\left(\frac{q-1}{q^{l+2}\ln q}\right) - 1 \ < \ \frac{1.06(q-1)}{q^{l+2}\ln q},$$

since $e^x - 1 < 1.06x$ when $0 < x \leq 4/(25\ln 5)$. Thus,

$$\sum_{l=0}^{\infty} f_q\left(\frac{1}{q^l \ln q}\right) \ < \ \frac{1.06}{q\ln q}.$$

We have

$$f_q(\ln q) \ = \ \frac{1}{q}\left(\exp\left(\frac{(q-1)\ln q}{q^2}\right) - 1\right) \ < \ \frac{1.141(q-1)\ln q}{q^3},$$

since $e^x - 1 < 1.141$ when $0 < x \leq (4/25)\ln 5$. Further, for an integer $l \geq 1$,

$$f_q\left(q^l \ln q\right) \ < \ \exp\left(\left(\frac{q-1}{q^2} - 1\right) q^l \ln q\right) \ = \ q^{-q^l(1-(q-1)/q^2)}.$$

Note that

$$\frac{q^{-q^l(1-(q-1)/q^2)}}{q^{-q^{l+1}(1-(q-1)/q^2)}} \ = \ q^{q^l(q-1)(1-(q-1)/q^2)} \ \geq \ 5^{84/5} \ > \ 5 \times 10^{11}.$$

Thus,

$$\sum_{l=1}^{\infty} f_q(q^l \ln q) \ < \ \left(1 + 10^{-11}\right) f_q(q\ln q).$$

But

$$f_q(q\ln q) \ = \ e^{-q\ln q}\left(e^{((q-1)/q)\ln q} - 1\right) < q^{-q}(q-1),$$

so that

$$\sum_{l=1}^{\infty} f_q(q^l \ln q) \ < \ \left(1 + 10^{-11}\right)(q-1)/q^q.$$

Thus if $q \geq 5$ and $q \notin E(x)$, then,

$$\begin{aligned} F_q(x) < \ & f_q(\ln q) \ + \ \sum_{l=0}^{\infty} f_q\left(\frac{1}{q^l \ln q}\right) \ + \ \sum_{l=1}^{\infty} f_q(q^l \ln q) \\ < \ & \frac{1.15(q-1)\ln q}{q^3} \ + \ \frac{1.11}{q\ln q} \ + \ \frac{q-1}{q^q}. \end{aligned}$$

We conclude that $F_q(x) \ll 1/(q\ln q)$. And if $q \geq 5$ and $q \in E(x)$, then we merely add $f_q(t_q) < (q-1)/(e(q^2-q+1))$ to this last estimate for $F_q(x)$, getting $F_q(x) \ll 1/q$. Whether or not $q \in E(x)$, we have $F_q(x) < 0.27$, completing the proof of the lemma.

Theorem (GRH) *There is a positive number A such that for each integer $a \notin \mathcal{E}$ there is a number $x_0(a)$ such that if x is in the set S of Lemma 7, $x \geq x_0(a)$, then $N_a(x)/x \geq A\varphi(|a|)/|a|$.*

Proof This result follows immediately from Lemma 7, Proposition 5 and Lemma 10.

Acknowledgements We would like to thank Cam Stewart for bringing the relevance of Schanuel's conjecture to our attention. The second author would like to acknowledge conversations he has had on the general subject of this paper since 1979, with Paul Erdős, Jeff Lagarias, Hendrik Lenstra, Ram Murty and Sam Wagstaff.

References

[1] P. Erdős and C. Pomerance, *On the normal number of prime factors of $\varphi(n)$*, Rocky Mountain J. Math. **15** (1985), 343–352. (Corrigendum in P. Erdős, A. Granville, C. Pomerance and C. Spiro, *On the normal behavior of the iterates of some arithmetic functions*, in Analytic Number Theory, Proc. Bateman Conf., Birkhäuser, Boston, 1990, pp. 165–204.)

[2] P. Erdős, C. Pomerance, and E. Schmutz, *Carmichael's lambda function*, Acta Arith. **58** (1991), 363–385.

[3] J. B. Friedlander, C. Pomerance, and I. E. Shparlinski, *Period of the power generator and small values of Carmichael's function*, Math. Comp. **70** (2001), 1591–1605.

[4] H. Halberstam and H. E. Richert, *Sieve Methods*, Academic Press, London, 1974

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

[6] P. Kurlberg, *On the order of unimodular matrices modulo integers*, preprint available at http://www.math.chalmers.se/~kurlberg/.

[7] J. C. Lagarias and A. M. Odlyzko, *Effective versions of the Chebotarev density theorem*, in Algebraic Number Fields: L-functions and Galois properties (Proc. Sympos. Univ. Durham, Durham, 1975), A. Frohlich, ed., pp. 409–464, Academic Press, London, 1977.

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

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

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

[11] G. Martin, *The least prime primitive root and the shifted sieve*, Acta Arith. **80** (1997), 277–288.

[12] M. R. Murty, M. Rosen, and J. H. Silverman, *Variations on a theme of Romanoff*, Internat. J. Math. **7** (1996), 373–391.

[13] M. R. Murty and F. Saidak, *Non-abelian generalizations of the Erdős–Kac theorem*, preprint, 2001.

[14] C. Pomerance, *On the distribution of amicable numbers*, J. reine angew. Math. **293/294** (1977), 217–222.

[15] C. Pomerance, *On primitive divisors of Mersenne numbers*, Acta Arith. **46** (1986), 355–367.

[16] F. Saidak, *Non-abelian generalizations of the Erdős–Kac theorem*, Ph. D. Thesis, Queen's University, Kingston, Canada, 2001.

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

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