Algebra & Number Theory

Volume 7

2013

No. 4

On a problem of Arnold: The average multiplicative order of a given integer

Pär Kurlberg and Carl Pomerance

ALGEBRA AND NUMBER THEORY 7:4 (2013)
dx.doi.org/10.2140/ant.2013.7.981

On a problem of Arnold: The average multiplicative order of a given integer

Pär Kurlberg and Carl Pomerance

For coprime integers g and n, let $\ell_g(n)$ denote the multiplicative order of g modulo n. Motivated by a conjecture of Arnold, we study the average of $\ell_g(n)$ as $n \le x$ ranges over integers coprime to g, and x tending to infinity. Assuming the generalized Riemann Hypothesis, we show that this average is essentially as large as the average of the Carmichael lambda function. We also determine the asymptotics of the average of $\ell_g(p)$ as $p \le x$ ranges over primes.

1. Introduction

Given coprime integers g and n with $n > 0$ and $|g| > 1$, let $\ell_g(n)$ denote the multiplicative order of g modulo n, that is, the smallest integer $k \ge 1$ such that $g^k \equiv 1 \bmod n$. For $x \ge 1$ an integer, let

$$T_g(x) := \frac{1}{x} \sum_{\substack{n \le x \\ (n,g)=1}} \ell_g(n),$$

essentially the average multiplicative order of g. Arnold [2005] conjectured that if $|g| > 1$, then

$$T_g(x) \sim c(g) \frac{x}{\log x},$$

as $x \to \infty$, for some constant $c(g) > 0$. However, Shparlinski [2007] showed that if the generalized Riemann Hypothesis[1] (GRH) is true, then

$$T_g(x) \gg \frac{x}{\log x} \exp\bigl(C(g)(\log\log\log x)^{3/2}\bigr),$$

Kurlberg was partially supported by grants from the Göran Gustafsson Foundation, the Knut and Alice Wallenberg foundation, the Royal Swedish Academy of Sciences, and the Swedish Research Council. Pomerance was supported by NSF grants, numbers DMS-0703850 and DMS-1001180.
MSC2010: 11N37.
Keywords: average multiplicative order.

[1] What is needed is that the Riemann Hypothesis holds for Dedekind zeta functions $\zeta_{K_n}(s)$ for all $n > 1$, where K_n is the Kummer extension $\mathbb{Q}(e^{2\pi i/n}, g^{1/n})$.

where $C(g) > 0$. He also suggested that it should be possible to obtain, again assuming GRH, a lower bound of the form

$$T_g(x) \geq \frac{x}{\log x} \exp\bigl((\log\log\log x)^{2+o(1)}\bigr) \quad \text{as } x \to \infty.$$

Let

$$B = \mathrm{e}^{-\gamma} \prod_p \Bigl(1 - \frac{1}{(p-1)^2(p+1)}\Bigr) = 0.3453720641\ldots, \tag{1}$$

the product being over primes, and where γ is the Euler–Mascheroni constant. The principal aim of this paper is to prove the following result.

Theorem 1. *Assuming the GRH,*

$$T_g(x) = \frac{x}{\log x} \exp\biggl(\frac{B \log\log x}{\log\log\log x}(1+o(1))\biggr) \quad \textit{as } x \to \infty,$$

uniformly in g with $1 < |g| \leq \log x$. The upper bound implicit in this result holds unconditionally.

Let $\lambda(n)$ denote the exponent of the group $(\mathbb{Z}/n\mathbb{Z})^\times$, which is commonly known as Carmichael's function. We have $\ell_g(n) \leq \lambda(n)$ when $(g, n) = 1$, so we immediately obtain that

$$T_g(x) \leq \frac{1}{x} \sum_{n \leq x} \lambda(n),$$

and it is via this inequality that we are able to unconditionally establish the upper bound implicit in Theorem 1. Indeed, Erdős, Pomerance, and Schmutz [Erdős et al. 1991] determined the average order of $\lambda(n)$ showing that, as $x \to \infty$,

$$\frac{1}{x} \sum_{n \leq x} \lambda(n) = \frac{x}{\log x} \exp\biggl(\frac{B \log\log x}{\log\log\log x}(1+o(1))\biggr). \tag{2}$$

Theorem 1 thus shows under assumption of the GRH that the mean values of $\lambda(n)$ and $\ell_g(n)$ are of a similar order of magnitude. We know, on assuming the GRH, that $\lambda(n)/\ell_g(n)$ is very small for almost all n (for instance, see [Kurlberg 2003; Li and Pomerance 2003]; in the latter paper it was in fact shown that $\lambda(n)/\ell_g(n) \leq (\log n)^{o(\log\log\log n)}$ as $n \to \infty$ on a set of relative asymptotic density 1 among integers coprime to g), so perhaps Theorem 1 is not very surprising. *However*, in [Erdős et al. 1991] it was also shown that the normal order of $\lambda(n)$ is quite a bit smaller than the average order: There exists a subset S of the positive integers of asymptotic density 1 such that for $n \in S$ and $n \to \infty$,

$$\lambda(n) = \frac{n}{(\log n)^{\log\log\log n + A + (\log\log\log n)^{-1+o(1)}}},$$

where $A > 0$ is an explicit constant. Thus the main contribution to the average of $\lambda(n)$ comes from a *density-zero subset* of the integers, and to obtain our result on the average multiplicative order, we must show that $\ell_g(n)$ is large for many n for which $\lambda(n)$ is large.

If one averages over g as well, then a result like our Theorem 1 holds unconditionally. In particular, it follows from [Luca and Shparlinski 2003, Theorem 6] that

$$\frac{1}{x^2}\sum_{n\le x}\sum_{\substack{1<g<n\\(g,n)=1}}\ell_g(n)=\frac{x}{\log x}\exp\biggl(\frac{B\log\log x}{\log\log\log x}(1+o(1))\biggr)\quad\text{as } x\to\infty.$$

We also note that our methods give that Theorem 1 still holds for $g = a/b$ a rational number, with uniform error for $|a|, |b| \le \log x$, and n ranging over integers coprime to ab.

1.1. ***Averaging over prime moduli.*** We shall always have the letters p, q denoting prime numbers. Given a rational number $g \ne 0, \pm 1$ and a prime p not dividing the numerator or denominator of g, let $\ell_g(p)$ denote the multiplicative order of g modulo p. For simplicity, when p does divide the numerator or denominator of g, we let $\ell_g(p) = 1$.

Further, given $k \in \mathbb{Z}^+$, let

$$D_g(k) := \bigl[\mathbb{Q}(g^{1/k}, \mathrm{e}^{2\pi i/k}) : \mathbb{Q}\bigr]$$

denote the degree of the Kummer extension obtained by taking the splitting field of $X^k - g$. Let $\operatorname{rad}(k)$ denote the largest squarefree divisor of k and let $\omega(k)$ be the number of primes dividing $\operatorname{rad}(k)$.

Theorem 2. *Given $g \in \mathbb{Q}$, $g \ne 0, \pm 1$, define*

$$c_g := \sum_{k=1}^{\infty}\frac{\phi(k)\operatorname{rad}(k)(-1)^{\omega(k)}}{k^2 D_g(k)}.$$

The series for c_g converges absolutely, and, assuming the GRH,

$$\frac{1}{\pi(x)}\sum_{p\le x}\ell_g(p)=\tfrac{1}{2}c_g\cdot x+O\biggl(\frac{x}{(\log x)^{2-4/\log\log\log x}}\biggr).$$

Furthermore, with $g = a/b$, where $a, b \in \mathbb{Z}$, the error estimate holds uniformly for $|a|, |b| \le x$.

At the heart of our claims of uniformity, both in Theorems 1 and 2, is our Theorem 6 in Section 2.

Though perhaps not obvious from the definition, $c_g > 0$ for all $g \neq 0, \pm 1$. In order to determine c_g, define

$$c := \prod_p \Bigl(1 - \frac{p}{p^3 - 1}\Bigr) = 0.5759599689\ldots,$$

the product being over primes; c_g turns out to be a positive *rational* multiple of c. To sum the series that defines c_g we will need some further notation. For p a prime and $\alpha \in \mathbb{Q}^*$, let $v_p(\alpha)$ be the exponential p-valuation at α, that is, it is the integer for which $p^{-v_p(\alpha)}\alpha$ is invertible modulo p. Write $g = \pm g_0^h$, where h is a positive integer and $g_0 > 0$ is not an exact power of a rational number, and write $g_0 = g_1 g_2^2$, where g_1 is a squarefree integer and g_2 is a rational. Let $e = v_2(h)$ and define $\Delta(g) = g_1$ if $g_1 \equiv 1 \bmod 4$, and $\Delta(g) = 4g_1$ if $g_1 \equiv 2$ or $3 \mod 4$. For $g > 0$, define $n = \operatorname{lcm}(2^{e+1}, \Delta(g))$. For $g < 0$, define $n = 2g_1$ if $e = 0$ and $g_1 \equiv 3 \mod 4$, or $e = 1$ and $g_1 \equiv 2 \mod 4$; let $n = \operatorname{lcm}(2^{e+2}, \Delta(g))$ otherwise.

Consider the multiplicative function $f(k) = (-1)^{\omega(k)}\operatorname{rad}(k)(h,k)/k^3$. We note that for p prime and $j \geq 1$,

$$f(p^j) = -p^{1-3j+\min(j, v_p(h)))}.$$

Given an integer $t \geq 1$, define $F(p,t)$ and $F(p)$ by

$$F(p,t) := \sum_{j=0}^{t-1} f(p^j) \quad \text{and} \quad F(p) := \sum_{j=0}^{\infty} f(p^j).$$

In particular, we note that if $p \nmid h$, then

$$F(p) = 1 - \sum_{j=1}^{\infty} p^{1-3j} = 1 - \frac{p}{p^3-1}. \tag{3}$$

Proposition 3. *With notation as above, if $g < 0$ and $e > 0$, we have*

$$c_g = c \cdot \prod_{p \mid h} \frac{F(p)}{1 - \frac{p}{p^3-1}} \cdot \biggl(1 - \frac{F(2, e+1) - 1}{2F(2)} + \prod_{p \mid n}\Bigl(1 - \frac{F(p, v_p(n))}{F(p)}\Bigr)\biggr);$$

otherwise

$$c_g = c \cdot \prod_{p \mid h} \frac{F(p)}{1 - \frac{p}{p^3-1}} \cdot \biggl(1 + \prod_{p \mid n}\Bigl(1 - \frac{F(p, v_p(n))}{F(p)}\Bigr)\biggr).$$

For example, if $g = 2$, then $h = 1$, $e = 0$, and $n = 8$. Thus

$$c_2 = c \cdot \Bigl(1 + 1 - \frac{F(2,3)}{F(2)}\Bigr) = c \cdot \Bigl(2 - \frac{1 - 2/(2^1)^3 - 2/(2^2)^3}{1 - 2/(8-1)}\Bigr) = c \cdot \tfrac{159}{160}.$$

We remark that the universal constant

$$c=\prod_p\Bigl(1-\frac{p}{p^3-1}\Bigr)$$

is already present in the work of Stephens on prime divisors of recurrence sequences. Motivated by a conjecture of Laxton, Stephens [1976] showed that on GRH, the limit

$$\lim_{x\to\infty}\frac{1}{\pi(x)}\sum_{p\le x}\frac{\ell_g(p)}{p-1}$$

exists and equals c times a rational correction factor depending on g. In fact, from the result it is easy deduce our Theorem 2 with a somewhat better error term. However, Stephens only treats integral g that are not powers, the error term is not uniform in g, and, as noted by Moree and Stevenhagen [2000], the correction factors must be adjusted in certain cases.

Theorem 2 might also be compared with the work of Pappalardi [1995]. In fact, his method suggests an alternate route to our Theorem 2, and would allow the upper bound

$$\frac{1}{\pi(x)}\sum_{p\le x}\ell_g(p)\le\tfrac{1}{2}(c_g+o(1))x,$$

as $x\to\infty$ to be established unconditionally. The advantage of our method is that it avoids computing the density of those primes for which g has a given index.

Finally, Theorem 2 should also be contrasted with the unconditional result of Luca [2005] that

$$\frac{1}{\pi(x)}\sum_{p\le x}\frac{1}{(p-1)^2}\sum_{g=1}^{p-1}\ell_g(p)=c+O(1/(\log x)^{\kappa})$$

for any fixed $\kappa>0$. By partial summation one can then obtain

$$\frac{1}{\pi(x)}\sum_{p\le x}\frac{1}{p-1}\sum_{g=1}^{p-1}\ell_g(p)\sim\tfrac{1}{2}c\cdot x\quad\text{as } x\to\infty,$$

a result that is more comparable to Theorem 2.

2. Some preliminary results

For an integer $m\ge 2$, we let $P(m)$ denote the largest prime dividing m, and we let $P(1)=1$.

Given a rational number $g\ne 0,\pm1$, we recall the notation h, e, n described in Section 1.1, and for a positive integer k, we recall that $D_g(k)$ is the degree of the splitting field of X^k-g over $\mathbb{Q}$. We record a result of Wagstaff on $D_g(k)$; see

Proposition 4.1 and the second paragraph in the proof of Theorem 2.2 in [Wagstaff 1982].

Proposition 4. *With notation as above*,

$$D_g(k)=\frac{\phi(k)\cdot k}{(k,h)\cdot\epsilon_g(k)},\tag{4}$$

where ϕ is Euler's function and $\epsilon_g(k)$ is defined as follows: *If $g>0$, then*

$$\epsilon_g(k):=\begin{cases}2 & \text{if } n\mid k,\\ 1 & \text{if } n\nmid k.\end{cases}$$

If $g<0$, then

$$\epsilon_g(k):=\begin{cases}2 & \text{if } n\mid k,\\ \frac{1}{2} & \text{if } 2\mid k \text{ and } 2^{e+1}\nmid k,\\ 1 & \text{otherwise.}\end{cases}$$

We also record a GRH-conditional version of the Chebotarev density theorem for Kummerian fields over $\mathbb{Q}$; see [Hooley 1967, Section 5; Lagarias and Odlyzko 1977, Theorem 1]. Let $i_g(p)=(p-1)/\ell_g(p)$, the index of $\langle g\rangle$ in $(\mathbb{Z}/p\mathbb{Z})^*$ when $g\in(\mathbb{Z}/p\mathbb{Z})^*$.

Theorem 5. *Assume the GRH. Suppose $g=a/b\neq 0,\pm1$, where a, b are integers of absolute value at most x. For each integer $k\le x$, the number of primes $p\le x$ for which $k\mid i_g(p)$ is*

$$\frac{1}{D_g(k)}\pi(x)+O(x^{1/2}\log x).$$

Note that $k\mid i_g(p)$ if and only if x^k-g splits completely modulo p. Also note that the trivial bound x/k is majorized by the error term in Theorem 5 when $k\ge x^{1/2}/\log x$. In fact, the error term majorizes the main term for $k\ge x^{1/4}$.

We will need the following *uniform* version of [Kurlberg and Pomerance 2005, Theorem 23].

Theorem 6. *If the GRH is true*, *then for x, L with $1\le L\le\log x$ and $g=a/b\neq 0,\pm1$, where a, b are integers with $|a|, |b|\le x$, we have*

$$\left|\left\{p\le x:\ell_g(p)\le\frac{p-1}{L}\right\}\right|\ll\frac{\pi(x)}{L}\cdot\frac{h\tau(h)}{\phi(h)}+\frac{x\log\log x}{\log^2 x}$$

uniformly, *where $\tau(h)$ is the number of divisors of h.*

Proof. Since the proof is rather similar to the proofs of the main theorem in [Hooley 1967], Theorem 2 in [Kurlberg 2003], and Theorem 23 in [Kurlberg and Pomerance 2005], we only give a brief outline. We see that $\ell_g(p)\le(p-1)/L$ implies that $i_g(p)\ge L$. Further, in the case that $p\mid ab$, where we are defining $\ell_g(p)=1$ and hence $i_g(p)=p-1$, the number of primes p is $O(\log x)$. So we assume that $p\nmid ab$.

First step: Consider primes $p \le x$ such that $i_g(p) > x^{1/2} \log^2 x$. Such a prime p divides $a^k - b^k$ for some positive integer $k < x^{1/2}/\log^2 x$. Since $\omega(|a^k - b^k|) \ll k \log x$, it follows that the number of primes p in this case is

$$O((x^{1/2}/\log^2 x)^2 \log x) = O(x/\log^3 x).$$

Second step: Consider primes p such that $q \mid i_g(p)$ for some prime q in the interval $I := [x^{1/2}/\log^2 x, x^{1/2} \log^2 x]$. We may bound this by considering primes $p \le x$ such that $p \equiv 1 \mod q$ for some prime $q \in I$. The Brun–Titchmarsh inequality then gives that the number of such primes p is at most a constant times

$$\sum_{q \in I} \frac{x}{\phi(q) \log(x/q)} \ll \frac{x}{\log x} \sum_{q \in I} \frac{1}{q} \ll \frac{x \log\log x}{\log^2 x}.$$

Third step: Now consider primes p such that $q \mid i_g(p)$ for some prime q in the interval $[L, x^{1/2}/\log^2 x)$. In this range we use Proposition 4 and Theorem 5 to get on the GRH that

$$|\{p \le x : q \mid i_g(p)\}| \ll \frac{\pi(x)(q,h)}{q\phi(q)} + x^{1/2} \log x.$$

Summing over primes q, we find that the number of such p is bounded by a constant times

$$\sum_{q \in [L, x^{1/2}/\log^2 x)} \left(\frac{\pi(x)(q,h)}{q^2} + x^{1/2} \log x \right) \ll \frac{\pi(x)\omega(h)}{L} + \frac{x}{\log^2 x}.$$

Fourth step: For the remaining primes p, any prime divisor $q \mid i_g(p)$ is smaller than L. Hence $i_g(p)$ must be divisible by some integer d in the interval $[L, L^2]$. By Proposition 4 and Theorem 5, assuming the GRH, we have

$$\left|\{p \le x : d \mid i_g(p)\}\right| \le 2\frac{\pi(x)(d,h)}{d\phi(d)} + O(x^{1/2} \log x). \tag{5}$$

Hence the total number of such p is bounded by

$$\sum_{d \in [L, L^2]} \left(2\frac{\pi(x)(d,h)}{d\phi(d)} + O(x^{1/2} \log x) \right) \ll \frac{\pi(x)}{L} \frac{h\tau(h)}{\phi(h)},$$

where the last estimate follows from

$$\begin{aligned} \sum_{d \in [L, L^2]} \frac{(d,h)}{d\phi(d)} &\le \sum_{m \mid h} \sum_{\substack{d \in [L, L^2] \\ m \mid d}} \frac{m}{d\phi(d)} \le \sum_{m \mid h} \sum_{k \ge L/m} \frac{1}{\phi(m) k \phi(k)} \\ &\ll \sum_{m \mid h} \frac{m}{L\phi(m)} = \frac{h}{L\phi(h)} \sum_{m \mid h} \frac{m}{\phi(m)} \cdot \frac{\phi(h)}{h} \le \frac{h\tau(h)}{L\phi(h)}. \end{aligned} \tag{6}$$

Here we used the bound

$$\sum_{k\ge T}\frac{1}{k\phi(k)}\ll 1/T$$

for $T>0$, which follows by an elementary argument from the bound

$$\sum_{k\ge T}\frac{1}{k^2}\ll 1/T$$

and the identity

$$k/\phi(k)=\sum_{j\,|\,k}\frac{\mu^2(j)}{\phi(j)}.$$

Indeed,

$$\sum_{k\ge T}\frac{1}{k\phi(k)}=\sum_{j}\frac{\mu^2(j)}{\phi(j)j^2}\sum_{l\ge T/j}\frac{1}{l^2}\ll\frac{1}{T}\sum_{j\le T}\frac{1}{\phi(j)j}+\sum_{j>T}\frac{1}{j^2}\ll\frac{1}{T}. \qquad \square$$

Corollary 7. *Assume the GRH is true. Let $m\ge 2$ be an integer and $x\ge 10^7$ a real number. Let $y=\log\log x$ and assume that $m\le\log y/\log\log y$. Let $g=a/b\ne 0,\pm 1$, where a,b are integers with $|a|,|b|\le\exp((\log x)^{3/m})$, and let h be as above. Then uniformly,*

$$\sum_{\substack{p\le x\\ P(i_g(p))>m}}\frac{1}{p}\ll y\Bigg(\frac{1}{m}+\sum_{\substack{q\,|\,h\\ q>m}}\frac{1}{q}\Bigg).$$

Proof. This result is more a corollary of the proof of Theorem 6 than its statement. We consider intervals $I_j:=(\mathrm{e}^j,\mathrm{e}^{j+1}]$ for $j\le\log x$, with j a nonnegative integer. The sum of reciprocals of all primes $p\le\exp((\log x)^{1/m})$ is $y/m+O(1)$, so this contribution to the sum is under control. We thus may restrict to the consideration of primes $p\in I_j$ for $j>(\log x)^{1/m}$. For such an integer j, let $t=\mathrm{e}^{j+1}$. If $q\,|\,i_g(p)$ for some prime $q>t^{1/2}\log^2 t$, then $\ell_g(p)\le t^{1/2}/\log^2 t$, and the number of such primes is

$$O\Bigg(\sum_{k\le t^{1/2}/\log^2 t}k\log|ab|\Bigg)=O(t\log|ab|/\log^4 t),$$

so that the sum of their reciprocals is $O(\log|ab|/\log^4 t)=O((\log x)^{3/m}/j^4)$. Summing this for $j>(\log x)^{1/m}$, we get $O(1)$, which is acceptable.

For $J:=(t^{1/2}/\log^2 t, t^{1/2}\log^2 t]$, with $t=\mathrm{e}^{j+1}$, we have that the reciprocal sum of the primes $p\in I_j$ with some $q\in J$ dividing $i_g(p)$ (so that $q\,|\,p-1$) is $O(\log\log t/\log^2 t)=O(\log j/j^2)$. Summing this for $j>(\log x)^{1/m}$ is $o(1)$ as $x\to\infty$ and is acceptable.

For $q\le t^{1/2}/\log^2 t$ we need the GRH. As in the proof of Theorem 6, the number of primes $p\in I_j$ with $q\,|\,i_g(p)$ is bounded by a constant times

$$\frac{t}{\log t}\frac{(q,h)}{q^2}+t^{1/2}\log t.$$

Thus, the reciprocal sum of these primes p is

$$O\biggl(\frac{(q,h)}{q^2\log t}+\frac{\log t}{t^{1/2}}\biggr)=O\biggl(\frac{(q,h)}{q^2 j}+\frac{j}{\mathrm{e}^{j/2}}\biggr).$$

We sum this expression over primes q with $m<q\ll \mathrm{e}^{j/2}/j^2$, getting

$$O\biggl(\frac{1}{jm\log m}+\frac{1}{j}\sum_{q\,|\,h,q>m}\frac{1}{q}+\frac{1}{j^2}\biggr).$$

Summing on $j\le\log x$ completes the proof. □

3. Proof of Theorem 1

Let x be large and let g be an integer with $1<|g|\le\log x$. Define

$$y=\log\log x,\quad m=\lfloor y/\log^3 y\rfloor,\quad D=m!\,,\quad S_k=\{p\le x:(p-1,\,D)=2k\}.$$

Then $S_1, S_2,\dots, S_{D/2}$ are disjoint sets of primes whose union equals $\{2<p\le x\}$. Let

$$\tilde S_k=\Bigl\{p\in S_k: p\nmid g,\ \frac{p-1}{2k}\,\Bigm|\,\ell_g(p)\Bigr\} \tag{7}$$

be the subset of S_k, where $\ell_g(p)$ is "large". Note that if $k\le\log y,\ \ p\in S_k\setminus\tilde S_k$, and $p\nmid g$, there is some prime $q>m$ with $q\mid(p-1)/\ell_g(p)$, so that $P(i_g(p))>m$. Indeed, for x sufficiently large, we have $\log y\le m/2$, and thus $k\le\log y$ implies that each prime dividing D also divides $D/(2k)$, so that $(p-1,\,D)=2k$ implies that the least prime factor of $(p-1)/(2k)$ exceeds m.

Thus, from Theorem 6,

$$|S_k\setminus\tilde S_k|\le\bigl|\{p\le x:\ell_g(p)<p/m\}\bigr|+\sum_{p\mid g}1\ll\frac{\pi(x)}{m}\cdot\frac{h\tau(h)}{\phi(h)}$$

uniformly for $k\le\log y$. Using this it is easy to see that S_k and $\tilde S_k$ are of similar size when k is small. However, we shall essentially measure the "size" of S_k or $\tilde S_k$ by the sum of the reciprocals of its members and for this we will use Corollary 7. We define

$$E_k:=\sum_{\substack{p\in S_k\\ 1<p^\alpha\le x}}\frac{1}{p^\alpha}\quad\text{and}\quad \tilde E_k:=\sum_{\substack{p\in\tilde S_k\\ 1<p^\alpha\le x}}\frac{1}{p^\alpha}.$$

By Lemma 1 of [Erdős et al. 1991],

$$E_k=\frac{y}{\log y}\cdot P_k\cdot(1+o(1)) \tag{8}$$

uniformly for $k \le \log^2 y$, where

$$P_k = \frac{\mathrm{e}^{-\gamma}}{k} \prod_{q>2}\Bigl(1 - \frac{1}{(q-1)^2}\Bigr) \prod_{q \mid k,\, q>2} \frac{q-1}{q-2}. \tag{9}$$

Note that, with B given by (1),

$$\sum_{k=1}^{\infty} \frac{P_k}{2k} = B. \tag{10}$$

The next lemma shows that not much is lost when restricting to primes $p \in \tilde{S}_k$.

Lemma 8. *For $k \le \log y$, we uniformly have*

$$\tilde{E}_k = E_k \cdot \Bigl(1 + O\Bigl(\frac{\log^5 y}{y}\Bigr)\Bigr).$$

Proof. By (8) and (9), we have

$$E_k \gg \frac{y}{k \log y} \ge \frac{y}{\log^2 y}, \tag{11}$$

and it is thus sufficient to show that $\sum_{p \in S_k \setminus \tilde{S}_k} 1/p \ll \log^3 y$ since the contribution from prime powers p^α for $\alpha \ge 2$ is $O(1)$. As we have seen, if $k \le \log y$ and $p \in S_k \setminus \tilde{S}_k$, then either $p \mid g$ or $P(i_g(p)) > m$. Hence, using Corollary 7 and noting that the hypothesis $|g| \le \log x$ implies that $h \ll y$ and so h has at most one prime factor $q > m$, we have

$$\sum_{p \in E_k \setminus \tilde{E}_k} \frac{1}{p} \ll \frac{y}{m} = \frac{y}{\lfloor y/\log^3 y \rfloor} \ll \log^3 y. \qquad \square$$

Lemma 9. *We have*

$$\sum_{k \le \log y} \frac{E_k}{2k} = \frac{By}{\log y}(1 + o(1)),$$

where B is given by (1).

Proof. This follows immediately from (8), (9), and (10). □

Given a vector $j = (j_1, j_2, \ldots, j_{D/2})$ with each $j_i \in \mathbb{Z}_{\ge 0}$, let

$$\|j\| := j_1 + j_2 + \cdots + j_{D/2}.$$

Paralleling the notation $\Omega_i(x;\, j)$ from [Erdős et al. 1991], we let

- $\tilde{\Omega}_1(x;\, \boldsymbol{j})$ be the set of integers that can be formed by taking products of $v = \|j\|$ distinct primes $p_1, p_2, \ldots, p_v$ so that
  - for each i, $p_i < x^{1/y^3}$, and
  - the first j_1 primes are in $\tilde{S}_1$, the next j_2 are in $\tilde{S}_2$, etc.;

- $\tilde{\Omega}_2(x;\, \boldsymbol{j})$ be the set of integers $u = p_1 p_2 \cdots p_v \in \tilde{\Omega}_1(x;\, \boldsymbol{j})$ where $(p_i - 1,\, p_j - 1)$ divides D for all $i \neq j$;
- $\tilde{\Omega}_3(x;\, \boldsymbol{j})$ be the set of integers of the form $n = up$, where $u \in \tilde{\Omega}_2(x;\, \boldsymbol{j})$ and p satisfies $(p-1,\, D) = 2$, $\max(x/(2u),\, x^{1/y}) < p \leq x/u$ and $\ell_g(p) > p/y^2$;
- $\tilde{\Omega}_4(x;\, \boldsymbol{j})$ be the set of integers $n = (p_1 p_2 \cdots p_v)p$ in $\tilde{\Omega}_3(x;\, \boldsymbol{j})$ with the additional property that $(p-1,\, p_i - 1) = 2$ for all i.

(In the third bullet, note that the max is not strictly necessary since when x is sufficiently large, $x/(2u) > x^{1/y}$.)

3.1. ***Some lemmas.*** We shall also need the following analogues of [Erdős et al. 1991, Lemmas 2–4]. Let

$$J := \{j : 0 \leq j_k \leq E_k/k \text{ for } k \leq \log y, \text{ and } j_k = 0 \text{ for } k > \log y\}.$$

Lemma 10. *If $j \in J$, $n \in \tilde{\Omega}_4(x;\, j)$, and $x \geq x_1$, then*

$$\ell_g(n) \geq c_1 \frac{x}{y^3} \prod_{k \leq \log y} (2k)^{-j_k},$$

where $x_1, c_1 > 0$ are absolute constants.

Proof. Suppose that $n = (p_1 p_2 \cdots p_v)p \in \tilde{\Omega}_4(x;\, j)$. Let $d_i = (p_i - 1,\, D)$, and let $u_i := (p_i - 1)/d_i$. By (7), u_i divides $\ell_g(p_i)$ for all i, and by the definition of $\tilde{\Omega}_3(x;\, j)$ we also have $\ell_g(p) > p/y^2$. Since $(p-1)/2$ is coprime to $(p_i - 1)/2$ for each i and each $(p_i - 1,\, p_j - 1) \mid D$ for $i \neq j$, we have $u_1, \ldots, u_v$, $p-1$ pairwise coprime. But

$$\ell_g(n) = \operatorname{lcm}[\ell_g(p_1), \ell_g(p_2), \ldots, \ell_g(p_v), \ell_g(p)],$$

so we find that, using the minimal order of Euler's function and $\ell_g(p) > p/y^2$,

$$\begin{aligned}\ell_g(n) \geq u_1 u_2 \cdots u_v \ell_g(p) &\geq \frac{\phi(n)}{y^2 \cdot \prod_{i=1}^{v} d_i} \\ &\gg \frac{n}{y^2 \cdot \log\log n \cdot \prod_{k=1}^{l} (2k)^{j_k}} \gg \frac{x}{y^3 \cdot \prod_{k=1}^{l} (2k)^{j_k}},\end{aligned}$$

where we recall that $d_i = (p_i - 1,\, D) = 2k$ if $p_i \in \tilde{S}_k$, and that $n \in \tilde{\Omega}_4(x;\, j)$ implies that $n > x/2$. □

Lemma 11. *If $j \in J$, $u \in \tilde{\Omega}_2(x;\, j)$, and $x \geq x_2$, then*

$$\left|\{p : up \in \tilde{\Omega}_4(x;\, j)\}\right| > c_2 x/(uy \log x),$$

where $x_2, c_2 > 0$ are absolute constants.

Proof. Note that $\|j\| \le \sum_{k=1}^{l} E_k/k \ll y/\log y$ for $j \in J$, by (8) and (9). For such vectors j, Lemma 3 of [Erdős et al. 1991] implies that the number of primes p with $\max(x/2u, x^{1/y}) < p \le x/u,\ (p-1, D) = 2$, and $(p-1, p_i - 1) = 2$ for all $p_i \mid u$ is $\gg x/(uy\log x)$. Thus it suffices to show that

$$\bigl|\{p \le x/u : (p-1, D) = 2,\ \ell_g(p) \le p/y^2\}\bigr| = o(x/(uy\log x)).$$

As we have seen, $\|j\| \ll y/\log y$ for $j \in J$, so that $u \in \tilde{\Omega}_2(x; j)$ has $u \le x^{1/y^2}$ for all large x. Thus, Theorem 6 implies that

$$\sum_{\substack{p \le x/u \\ \ell_g(p) \le p/y^2}} 1 \ll \frac{\pi(x/u)}{y^2} \ll \frac{x}{uy^2\log x} = o\Bigl(\frac{x}{uy\log x}\Bigr).$$

The result follows. □

Lemma 12. *If $j \in J$, then for $x \ge x_3$,*

$$\sum_{u \in \tilde{\Omega}_2(x;j)} \frac{1}{u} > \exp\Bigl(\frac{-c_3 y\log\log y}{\log^2 y}\Bigr) \prod_{k \le \log y} \frac{E_k^{j_k}}{j_k!},$$

where $x_3, c_3 > 0$ are absolute constants.

Proof. The sum in the lemma is equal to

$$\frac{1}{j_1! j_2! \cdots j_{\lfloor \log y\rfloor}!} \sum_{\langle p_1, p_2, \dots, p_v\rangle} \frac{1}{p_1 p_2 \cdots p_v},$$

where the sum is over sequences of distinct primes for which the first j_1 are in $\tilde{S}_1$, the next j_2 are in $\tilde{S}_2$, and so on, and also each $(p_i - 1, p_j - 1) \mid D$ for $i \ne j$. Such a sum is estimated from below in Lemma 4 of [Erdős et al. 1991] but without the extra conditions that differentiate $\tilde{S}_k$ from S_k. The key prime reciprocal sum there is estimated on pages 381–383 to be

$$E_k\Bigl(1 + O\Bigl(\frac{\log\log y}{\log y}\Bigr)\Bigr).$$

In our case we have the extra conditions that $p \nmid g$ and $(p-1)/2k \mid \ell_g(p)$, which alters the sum by a factor of $1 + O(\log^5 y/y)$ by Lemma 8. But the factor $1 + O(\log^5 y/y)$ is negligible compared with the factor $1 + O(\log\log y/\log y)$, so we have exactly the same expression in our current case. □

3.2. *Conclusion.*

For brevity, let $l = \lfloor \log y\rfloor$. We clearly have

$$T_g(x) \ge \frac{1}{x} \sum_{j \in J} \sum_{n \in \tilde{\Omega}_4(x;j)} \ell_g(n).$$

By Lemma 10, we thus have

$$T_g(x) \gg \frac{1}{y^3} \sum_{j \in J} \prod_{k=1}^{l} (2k)^{-j_k} \sum_{n \in \tilde{\Omega}_4(x;j)} 1.$$

Now,

$$\sum_{n \in \tilde{\Omega}_4(x;j)} 1 = \sum_{u \in \tilde{\Omega}_2(x;j)} \sum_{up \in \tilde{\Omega}_4(x;j)} 1,$$

and by Lemma 11, this is

$$\gg \sum_{u \in \tilde{\Omega}_2(x;j)} \frac{x}{uy \log x},$$

which in turn by Lemma 12 is

$$\gg \frac{x}{y \log x} \exp\left(\frac{-c_3 y \log\log y}{\log^2 y} \right) \prod_{k=1}^{l} \frac{E_k^{j_k}}{j_k!}.$$

Hence

$$T_g(x) \gg \frac{x}{y^4 \log x} \exp\left(\frac{-c_3 y \log\log y}{\log^2 y} \right) \sum_{j \in J} \prod_{k=1}^{l} (2k)^{-j_k} \frac{E_k^{j_k}}{j_k!}.$$

Now,

$$\sum_{j \in J} \prod_{k=1}^{l} (2k)^{-j_k} \frac{E_k^{j_k}}{j_k!} = \prod_{k=1}^{l} \left(\sum_{j_k=0}^{[E_k/k]} \frac{(E_k/2k)^{j_k}}{j_k!} \right).$$

Note that $\sum_{j=0}^{2w} w^j/j! > \mathrm{e}^w/2$ for $w \geq 1$ and also that $E_k/2k \geq 1$ for x sufficiently large, as $E_k \gg y/(k \log y)$ by (11). Thus,

$$\sum_{j \in J} \prod_{k=1}^{l} (2k)^{-j_k} \frac{E_k^{j_k}}{j_k!} > 2^{-l} \exp\left(\sum_{k=1}^{l} \frac{E_k}{2k} \right).$$

Hence

$$T_g(x) \gg \frac{x}{y^4 \log x} \exp\left(\frac{-c_3 y \log\log y}{\log^2 y} \right) 2^{-l} \exp\left(\sum_{k=1}^{l} \frac{E_k}{2k} \right).$$

By Lemma 9 we thus have the lower bound in the theorem. The proof is concluded.

4. Averaging over prime moduli — the proofs

4.1. ***Proof of Theorem 2.*** Let $z = \log x$ and abbreviate $\ell_g(p)$ and $i_g(p)$ by $\ell(p)$ and $i(p)$, respectively. We have

$$\sum_{p\le x} \ell(p) = \sum_{\substack{p\le x\\ i(p)\le z}} \ell(p) + \sum_{\substack{p\le x\\ i(p)>z}} \ell(p) = A + E,$$

say. Writing $\ell(p) = (p-1)/i(p)$ and using the identity $1/i(p) = \sum_{uv|i(p)} \mu(v)/u$, we find that

$$\begin{aligned} A &= \sum_{\substack{p\le x\\ i(p)\le z}} (p-1) \sum_{uv\,|\,i(p)} \frac{\mu(v)}{u} \\ &= \sum_{p\le x} (p-1) \sum_{\substack{uv\,|\,i(p)\\ uv\le z}} \frac{\mu(v)}{u} - \sum_{\substack{p\le x\\ i(p)>z}} (p-1) \sum_{\substack{uv\,|\,i(p)\\ uv\le z}} \frac{\mu(v)}{u} \\ &= A_1 - E_1, \end{aligned}$$

say. The main term A_1 is

$$A_1 = \sum_{uv\le z} \frac{\mu(v)}{u} \sum_{\substack{p\le x\\ uv|i(p)}} (p-1).$$

By a simple partial summation using Theorem 5, the inner sum here is

$$\frac{\mathrm{Li}(x^2)}{D_g(uv)} + O(x^{3/2}\log x)$$

assuming the GRH. Thus,

$$A_1 = \mathrm{Li}(x^2)\Bigg(\sum_{uv\le z} \frac{\mu(v)}{uD_g(uv)}\Bigg) + O\Bigg(x^{3/2}\log x \sum_{n\le z}\Bigg|\sum_{uv=n}\frac{\mu(v)}{u}\Bigg|\Bigg).$$

The inner sum in the O-term is bounded by $\phi(n)/n$, so the O-term is $O(x^{3/2}\log^2 x)$. Recalling that $\mathrm{rad}(n)$ denotes the largest squarefree divisor of n, we note that $\sum_{v|k}\mu(v)v = \prod_{p|k}(1-p) = (-1)^{\omega(k)}\phi(\mathrm{rad}(k))$, and hence

$$\sum_{uv=k} \frac{\mu(v)}{uD_g(uv)} = \sum_{v|k} \frac{\mu(v)v}{D_g(k)k} = \frac{(-1)^{\omega(k)}\phi(\mathrm{rad}(k))}{D_g(k)k}.$$

On noting that $\phi(\mathrm{rad}(k)) = \phi(k)\mathrm{rad}(k)/k$, we have

$$\sum_{u,v} \frac{\mu(v)}{uD_g(uv)} = \sum_{k\ge 1} \frac{(-1)^{\omega(k)}\mathrm{rad}(k)\phi(k)}{D_g(k)k^2} = c_g.$$

Thus, with $\psi(h) := h\tau(h)/\phi(h)$,

$$\sum_{uv\le z} \frac{\mu(v)}{uvD_g(uv)} = c_g - \sum_{k>z} \frac{(-1)^{\omega(k)}\mathrm{rad}(k)\phi(k)}{D_g(k)k^2} = c_g + O(\psi(h)/z),$$

by Proposition 4 and the same argument as in the fourth step of the proof of Theorem 6 (in particular, see (6)). It now follows that

$$A_1 = \mathrm{Li}(x^2)(c_g + O(\psi(h)/z)).$$

It remains to estimate the two error terms E, E_1. Using Theorem 6, we have

$$E \ll \frac{x}{z} \cdot \frac{x \log\log x}{\log^2 x}\psi(h) \ll \frac{x^2\psi(h)}{\log^2 x}.$$

Toward estimating E_1, we note that

$$f_z(n) := \left|\sum_{\substack{uv\,|\,n\\ uv\le z}} \frac{\mu(v)}{u}\right| \le \sum_{\substack{d\,|\,n\\ d\le z}} \left|\sum_{v|d} \frac{\mu(v)v}{d}\right| = \sum_{\substack{d\,|\,n\\ d\le z}} \frac{\phi(\mathrm{rad}(d))}{d} \le z.$$

Further, from the last sum we get

$$f_z(n) \le \prod_{\substack{p^a\,\|\,n\\ p\le z}} \left(1 + \frac{p-1}{p} + \cdots + \frac{p-1}{p^a}\right) < 2^{\omega(n_z)},$$

where n_z denotes the largest divisor of n composed of primes in $[1, z]$. We have

$$|E_1| \le \sum_{\substack{p\le x\\ i(p)>z}} (p-1) f_z(i(p)) \le x \sum_{\substack{p\le x\\ i(p)>z}} f_z(i(p)).$$

Let $w := 4\log z/\log\log z$. We break the sum above into three possibly overlapping parts:

$$E_{1,1} := x \sum_{\substack{p\le x\\ i(p)>z\\ \omega(i(p)_z)\le w}} f_z(i(p)), \quad E_{1,2} := x \sum_{\substack{p\le x\\ z<i(p)\le x^{1/2}\log^2 x\\ \omega(i(p)_z)>w}} f_z(i(p)),$$

$$E_{1,3} := x \sum_{\substack{p\le x\\ i(p)>x^{1/2}\log^2 x}} f_z(i(p)).$$

Using Theorem 6, we have

$$E_{1,1} \le x2^w \sum_{\substack{p\le x\\ i(p)>z}} 1 \ll 2^w \psi(h)\, \frac{x^2 \log\log x}{\log^2 x}.$$

The estimate for $E_{1,3}$ is similarly brief, this time using the "first step" in the proof of Theorem 6. We have

$$E_{1,3} \le xz \sum_{\substack{p\le x\\ i(p)>x^{1/2}\log^2 x}} 1 \ll \frac{x^2}{\log^2 x}.$$

The estimate for $E_{1,2}$ takes a little work. By the Brun–Titchmarsh inequality,

$$\begin{aligned} E_{1,2} &\le xz \sum_{\substack{z<n\le x^{1/2}\log^2 x\\ \omega(n_z)>w}} \pi(x;n,1) \ll \frac{x^2 z}{\log x} \sum_{\substack{z<n\le x^{1/2}\log^2 x\\ \omega(n_z)>w}} \frac{1}{\phi(n)} \\ &\le x^2 \sum_{\substack{P(m)\le z\\ \omega(m)>w}} \frac{1}{\phi(m)} \sum_{n\le x^{1/2}\log^2 x} \frac{1}{\phi(n)} \ll x^2 \log x \sum_{\substack{P(m)\le z\\ \omega(m)>w}} \frac{1}{\phi(m)}. \end{aligned}$$

This last sum is smaller than

$$\begin{aligned} \sum_{k>w} \frac{1}{k!}\Big(\sum_{p\le z}\Big(\frac{1}{p-1}+\frac{1}{p(p-1)}+\cdots\Big)\Big)^k &= \sum_{k>w}\frac{1}{k!}\Big(\sum_{p\le z}\frac{p}{(p-1)^2}\Big)^k \\ &= \sum_{k>w}\frac{1}{k!}\big(\log\log z + O(1)\big)^k. \end{aligned}$$

The terms in this series are decaying at least geometrically by a large factor, so by a weak form of Stirling's formula, we have

$$\sum_{\substack{P(m)\le z\\ \omega(m)>w}} \frac{1}{m} \ll \exp\big(w\log\log\log z - w\log w + w + O(w/\log\log z)\big).$$

By our choice for w, this last expression is smaller than $\exp(-3\log z) = (\log x)^{-3}$ for all large values of x. Hence, $E_{1,2} \ll x^2/\log^2 x$.

Noting that $\psi(h) \ll \tau(h)\log\log x$, we conclude that

$$\begin{aligned} \sum_{p\le x} \ell(p) &= A + E = A_1 + E + O\big(E_{1,1}+E_{1,2}+E_{1,3}\big) \\ &= c_g \mathrm{Li}(x^2) + O\Big(\frac{x^2}{\log^2 x}\big(\psi(h) + 2^w\psi(h)\log\log x + 1 + 1\big)\Big) \\ &= c_g \mathrm{Li}(x^2) + O\Big(2^w\tau(h)\cdot\frac{x^2(\log\log x)^2}{\log^2 x}\Big) \\ &= \tfrac{1}{2}c_g x\pi(x) + O\Big(\frac{x^2}{(\log x)^{2-4/\log\log\log x}}\Big), \end{aligned}$$

using that $\mathrm{Li}(x^2) = \frac{1}{2}x\pi(x) + O(x^2/\log^2 x)$, the definition of w, and $h \le \log x$ together with Wigert's theorem for the maximal order of the divisor function $\tau(h)$. This completes the proof.

4.2. ***Proof of Proposition 3.*** We begin with the cases $g > 0$, or $g < 0$ and $e = 0$. Recalling that $D_g(k) = \phi(k)k/(\epsilon_g(k)(k,h))$, we find that

$$c_g = \sum_{k\ge1} \frac{(-1)^{\omega(k)}\mathrm{rad}(k)\phi(k)}{D_g(k)k^2} = \sum_{k\ge1} \frac{(-1)^{\omega(k)}\mathrm{rad}(k)(k,h)\epsilon_g(k)}{k^3}. \tag{12}$$

Now, since $\epsilon_g(k)$ equals 1 if $n \nmid k$, and 2 otherwise, (12) equals

$$\sum_{k\ge1} \frac{(-1)^{\omega(k)}\mathrm{rad}(k)(h,k)}{k^3} + \sum_{n\,|\,k} \frac{(-1)^{\omega(k)}\mathrm{rad}(k)(h,k)}{k^3} = \sum_{k\ge1}(f(k)+f(kn)), \tag{13}$$

where the function $f(k) = (-1)^{\omega(k)}\mathrm{rad}(k)(h,k)/k^3$ is multiplicative.

If $p \nmid h$ and $j \ge 1$, we have $f(p^j) = -p/p^{3j}$. On the other hand, writing $h = \prod_{p|h} p^{e_{h,p}}$ we have $f(p^j) = -p^{1+\min(j,e_{h,p})}/p^{3j}$ for $p \mid h$ and $j \ge 1$. Since f is multiplicative,

$$\sum_{k\ge1}(f(k)+f(kn)) = \sum_{k:\mathrm{rad}(k)\,|\,hn} (f(k)+f(kn)) \cdot \sum_{(k,hn)=1} f(k).$$

Now, for $p \nmid h$ and $j \ge 1$, we have $f(p^j) = -\mathrm{rad}(p^j)/p^{3j} = -p/p^{3j}$; hence

$$\sum_{j\ge0} f(p^j) = 1 - \frac{p}{p^3(1-1/p^3)} = 1 - \frac{p}{p^3-1}$$

and thus

$$\sum_{(k,hn)=1} f(k) = \prod_{p\nmid hn} F(p) = \prod_{p\nmid hn}\Big(1 - \frac{p}{p^3-1}\Big) = \frac{c}{\prod_{p\,|\,hn}(1-\frac{p}{p^3-1})}.$$

Similarly, $\sum_{\mathrm{rad}(k)\,|\,hn} f(k) = \prod_{p\,|\,hn} F(p)$ and

$$\sum_{\mathrm{rad}(k)\,|\,hn} f(kn) = \prod_{p\,|\,hn}\Big(\sum_{j\ge e_{n,p}} f(p^j)\Big) = \prod_{p\,|\,hn}\big(F(p) - F(p,e_{n,p})\big).$$

Hence

$$\begin{aligned}\sum_{\mathrm{rad}(k)\,|\,hn} f(k) + \sum_{\mathrm{rad}(k)\,|\,hn} f(kn) &= \prod_{p\,|\,hn} F(p) + \prod_{p|hn}\big(F(p) - F(p,e_{n,p})\big)\\ &= \prod_{p|hn} F(p)\cdot\Big(1 + \prod_{p|hn}\Big(1 - \frac{F(p,e_{n,p})}{F(p)}\Big)\Big).\end{aligned}$$

Thus

$$c_g = \frac{c}{\prod_{p\,|\,hn}(1-\frac{p}{p^3-1})} \cdot \prod_{p\,|\,hn} F(p) \cdot \Bigl(1 + \prod_{p\,|\,hn}\Bigl(1 - \frac{F(p, e_{n,p})}{F(p)}\Bigr)\Bigr),$$

which, by (3), simplifies to

$$c_g = c \cdot \prod_{p|h} \frac{F(p)}{1-\frac{p}{p^3-1}} \cdot \Bigl(1 + \prod_{p\,|\,hn}\Bigl(1 - \frac{F(p, e_{n,p})}{F(p)}\Bigr)\Bigr).$$

The case $g < 0$ and $e > 0$ is similar: using the multiplicativity of f together with the definition of $\epsilon_g(k)$, we find that

$$\begin{aligned} c_g &= \sum_{k\geq 1}(f(k) + f(kn)) - \tfrac{1}{2}\sum_{j=1}^{e}\sum_{(k,2)=1} f(2^j k) \\ &= \prod_p F(p) + \prod_p (F(p) - F(p, e_{n,p})) - \tfrac{1}{2}\cdot(F(2, e+1) - 1)\cdot \prod_{p>2} F(p) \\ &= \prod_p F(p)\Bigl(1 + \prod_{p\,|\,n}\Bigl(1 - \frac{F(p, e_{n,p})}{F(p)}\Bigr) - \frac{F(2, e+1) - 1}{2F(2)}\Bigr). \end{aligned}$$

Again using the fact that

$$\prod_p F(p) = \prod_{p\nmid h}\Bigl(1 - \frac{p}{p^3+1}\Bigr)\prod_{p\,|\,h} F(p) = c \cdot \prod_{p\,|\,h} \frac{F(p)}{1 - p/(p^3+1)},$$

the proof is concluded.

Acknowledgments

Part of this work was done while the authors visited MSRI, as part of the semester program “Arithmetic Statistics”. We thank MSRI for their support, funded through the NSF. We are very grateful to Michel Balazard for suggesting Arnold’s conjecture to us. In addition we thank Pieter Moree, Adam Felix, and the two anonymous referees for helpful comments.

References

[Arnold 2005] V. Arnold, “Number-theoretical turbulence in Fermat–Euler arithmetics and large Young diagrams geometry statistics”, *J. Mathematical Fluid Mech.* **7**:suppl. 1 (2005), S4–S50. MR 2006g:11199 Zbl 1134.11344

[Erdős et al. 1991] P. Erdős, C. Pomerance, and E. Schmutz, “Carmichael’s lambda function”, *Acta Arith.* **58**:4 (1991), 363–385. MR 92g:11093 Zbl 0734.11047

[Hooley 1967] C. Hooley, “On Artin’s conjecture”, *J. Reine Angew. Math.* **225** (1967), 209–220. MR 34 #7445 Zbl 0221.10048

[Kurlberg 2003] P. Kurlberg, "On the order of unimodular matrices modulo integers", *Acta Arith.* **110**:2 (2003), 141–151. MR 2005a:11146 Zbl 1030.11048

[Kurlberg and Pomerance 2005] P. Kurlberg and C. Pomerance, "On the periods of the linear congruential and power generators", *Acta Arith.* **119**:2 (2005), 149–169. MR 2006k:11153 Zbl 1080.11059

[Lagarias and Odlyzko 1977] J. C. Lagarias and A. M. Odlyzko, "Effective versions of the Chebotarev density theorem", pp. 409–464 in *Algebraic number fields: L-functions and Galois properties* (Durham, 1975), edited by A. Fröhlich, Academic Press, London, 1977. MR 56 #5506 Zbl 0362.12011

[Li and Pomerance 2003] S. Li and C. Pomerance, "On generalizing Artin's conjecture on primitive roots to composite moduli", *J. Reine Angew. Math.* **556** (2003), 205–224. MR 2004c:11177 Zbl 1022.11049

[Luca 2005] F. Luca, "Some mean values related to average multiplicative orders of elements in finite fields", *Ramanujan J.* **9**:1-2 (2005), 33–44. MR 2006i:11111 Zbl 1155.11344

[Luca and Shparlinski 2003] F. Luca and I. E. Shparlinski, "Average multiplicative orders of elements modulo *n*", *Acta Arith.* **109**:4 (2003), 387–411. MR 2004i:11113 Zbl 1043.11067

[Moree and Stevenhagen 2000] P. Moree and P. Stevenhagen, "A two-variable Artin conjecture", *J. Number Theory* **85**:2 (2000), 291–304. MR 2001k:11188 Zbl 0966.11042

[Pappalardi 1995] F. Pappalardi, "On Hooley's theorem with weights", *Rend. Sem. Mat. Univ. Politec. Torino* **53**:4 (1995), 375–388. MR 98c:11102 Zbl 0883.11042

[Shparlinski 2007] I. E. Shparlinski, "On some dynamical systems in finite fields and residue rings", *Discrete Contin. Dyn. Syst.* **17**:4 (2007), 901–917. MR 2007j:11098 Zbl 1127.11052

[Stephens 1976] P. J. Stephens, "Prime divisors of second-order linear recurrences, I", *J. Number Theory* **8**:3 (1976), 313–332. MR 54 #5142 Zbl 0334.10018

[Wagstaff 1982] S. S. Wagstaff, Jr., "Pseudoprimes and a generalization of Artin's conjecture", *Acta Arith.* **41**:2 (1982), 141–150. MR 83m:10004 Zbl 0496.10001

Communicated by Andrew Granville
Received 2011-08-25 Revised 2012-03-03 Accepted 2012-05-24

kurlberg@math.kth.se
Department of Mathematics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden
http://www.math.kth.se/~kurlberg/

carl.pomerance@dartmouth.edu
Mathematics Department, Kemeny Hall, Dartmouth College, Hanover NH 03755, United States
www.math.dartmouth.edu/~carlp

Algebra & Number Theory

msp.org/ant

EDITORS

MANAGING EDITOR	EDITORIAL BOARD CHAIR
Bjorn Poonen Massachusetts Institute of Technology Cambridge, USA	David Eisenbud University of California Berkeley, USA

BOARD OF EDITORS

Georgia Benkart	University of Wisconsin, Madison, USA	Susan Montgomery	University of Southern California, USA
Dave Benson	University of Aberdeen, Scotland	Shigefumi Mori	RIMS, Kyoto University, Japan
Richard E. Borcherds	University of California, Berkeley, USA	Raman Parimala	Emory University, USA
John H. Coates	University of Cambridge, UK	Jonathan Pila	University of Oxford, UK
J-L. Colliot-Thélène	CNRS, Université Paris-Sud, France	Victor Reiner	University of Minnesota, USA
Brian D. Conrad	University of Michigan, USA	Karl Rubin	University of California, Irvine, USA
Hélène Esnault	Freie Universität Berlin, Germany	Peter Sarnak	Princeton University, USA
Hubert Flenner	Ruhr-Universität, Germany	Joseph H. Silverman	Brown University, USA
Edward Frenkel	University of California, Berkeley, USA	Michael Singer	North Carolina State University, USA
Andrew Granville	Université de Montréal, Canada	Vasudevan Srinivas	Tata Inst. of Fund. Research, India
Joseph Gubeladze	San Francisco State University, USA	J. Toby Stafford	University of Michigan, USA
Ehud Hrushovski	Hebrew University, Israel	Bernd Sturmfels	University of California, Berkeley, USA
Craig Huneke	University of Virginia, USA	Richard Taylor	Harvard University, USA
Mikhail Kapranov	Yale University, USA	Ravi Vakil	Stanford University, USA
Yujiro Kawamata	University of Tokyo, Japan	Michel van den Bergh	Hasselt University, Belgium
János Kollár	Princeton University, USA	Marie-France Vignéras	Université Paris VII, France
Yuri Manin	Northwestern University, USA	Kei-Ichi Watanabe	Nihon University, Japan
Barry Mazur	Harvard University, USA	Efim Zelmanov	University of California, San Diego, USA
Philippe Michel	École Polytechnique Fédérale de Lausanne		

PRODUCTION

production@msp.org

Silvio Levy, Scientific Editor

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2013 is US $200/year for the electronic version, and $350/year (+$40, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscribers address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

ANT peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

© 2013 Mathematical Sciences Publishers

Algebra & Number Theory

Volume 7 No. 4 2013