Elementary abelian Sylow subgroups of the multiplicative group

S. Morales, G. Polanco, and P. Pollack

Abstract. Erdős and Pomerance have shown that $\varphi(n)$ typically has about $\frac{1}{2}(\log\log n)^2$ distinct prime factors. More precisely, $\omega(\varphi(n))$ has normal order $\frac{1}{2}(\log\log n)^2$. Since $\varphi(n)$ is the size of the multiplicative group $(\mathbf{Z}/n\mathbf{Z})^\times$, this result also gives the normal number of Sylow subgroups of $(\mathbf{Z}/n\mathbf{Z})^\times$. Recently, Pollack considered specifically noncyclic Sylow subgroups of $(\mathbf{Z}/n\mathbf{Z})^\times$, showing that the count of those has normal order $\log\log n/\log\log\log n$. We prove that the count of noncyclic Sylow subgroups that are elementary abelian of a fixed rank $k \geq 2$ has normal order $\frac{1}{k(k-1)}\log\log n/\log\log\log n$. So for example, (typically) among the primes p for which the p-primary component of $(\mathbf{Z}/n\mathbf{Z})^\times$ is noncyclic, this component is $\mathbf{Z}/p\mathbf{Z}\oplus\mathbf{Z}/p\mathbf{Z}$ about half the time. Additionally, we show that the count of p for which the p-Sylow subgroup of $(\mathbf{Z}/n\mathbf{Z})^\times$ is not elementary abelian has normal order $2\sqrt{\pi}\sqrt{\log\log n}/\log\log\log n$.

1. Introduction

The multiplicative groups $(\mathbf{Z}/n\mathbf{Z})^\times$ are objects of fundamental interest in both algebra and number theory. For any given n, the structure of $(\mathbf{Z}/n\mathbf{Z})^\times$ was determined by Gauss two centuries ago, in his *Disquisitiones* (see Lemma 2.1 below). From a statistical perspective, it is natural to ask how this structure varies with n. For example, one might look for features that hold for almost all values of n. Here and below, we say a property holds for almost all positive integers n if the count of exceptional $n \leq x$ is $o(x)$, as $x \to \infty$.

A notable result in this direction is due to Erdős and Pomerance in [**5**] (see also the paper [**11**] of Murty and Murty for a generalization). Recall that the arithmetic function $f(n)$ has normal order $g(n)$ if, for each fixed $\epsilon > 0$, we have

$$|f(n) - g(n)| < \epsilon g(n)$$

for almost all n. The theorem of Erdős and Pomerance is that $\omega(\varphi(n))$, the number of distinct prime factors of $\varphi(n) := \#(\mathbf{Z}/n\mathbf{Z})^\times$, has normal order $\frac{1}{2}(\log\log n)^2$.[1] As $(\mathbf{Z}/n\mathbf{Z})^\times$ has exactly one p-Sylow subgroup for every prime $p \mid \varphi(n)$, one can interpret the Erdős–Pomerance result group-theoretically: It provides the normal order for the number of Sylow subgroups of $(\mathbf{Z}/n\mathbf{Z})^\times$.

In [**1**], Banks, Luca, and Shparlinski investigate several problems concerning the ratios $\varphi(n)/\lambda(n)$, where $\lambda(n)$ is Carmichael's lambda function, the exponent[2] of $(\mathbf{Z}/n\mathbf{Z})^\times$. They observe that $\omega(\frac{\varphi(n)}{\lambda(n)})$ counts noncyclic Sylow subgroups of $(\mathbf{Z}/n\mathbf{Z})^\times$, and they prove that for

2020 *Mathematics Subject Classification.* Primary 11N37; Secondary 11N36.

[1] In fact, both [**5**] and [**11**] establish Gaussian laws: Loosely speaking, $\omega(\varphi(n))$ is normally distributed with mean $\frac{1}{2}(\log\log n)^2$ and variance $\frac{1}{3}(\log\log n)^3$. This is stronger than having normal order $\frac{1}{2}(\log\log n)^2$. The theorems of [**5, 11**] are analogues of a celebrated result of Erdős and Kac implying (among other things) that $\omega(n)$ follows a normal distribution [**4**].

[2] The exponent of a finite group G is the smallest positive integer h for which $g^h = e$ (the identity) for all $g \in G$.

almost all n,

$$\frac{\log\log n}{(\log\log\log n)^2} \ll \omega\left(\frac{\varphi(n)}{\lambda(n)}\right) \ll \log\log n.$$

This was sharpened in [**13**], where it was shown that $\omega(\frac{\varphi(n)}{\lambda(n)})$ has normal order

$$\frac{\log\log n}{\log\log\log n}. \tag{1}$$

In this paper we examine a related question: Fix a positive integer k. How often is the p-part of $(\mathbf{Z}/n\mathbf{Z})^\times$ elementary abelian of rank k, i.e., isomorphic to $\mathbf{Z}/p\mathbf{Z} \oplus \cdots \oplus \mathbf{Z}/p\mathbf{Z}$ (with k factors)? Our first theorem addresses this question for $k > 1$. Specifically, if $E_k(n)$ denotes the number of Sylow subgroups of $(\mathbf{Z}/n\mathbf{Z})^\times$ that are elementary abelian of rank k, then we have the following result.

Theorem 1.1. *Fix an integer $k \geq 2$. Then $E_k(n)$ has normal order*

$$\frac{1}{k(k-1)}\frac{\log\log n}{\log\log\log n}. \tag{2}$$

Of course, an elementary abelian group of rank larger than 1 is not cyclic. Comparing (2) with (1), and noting that $\sum_{k\geq 2}\frac{1}{k(k-1)} = 1$, we see that the elementary abelian Sylow subgroups of rank larger than 1 essentially "fill out" the space of noncyclic Sylow subgroups.

What about $E_1(n)$? By the results of [**5, 11**] and [**13**], the count of cyclic Sylow subgroups of $(\mathbf{Z}/n\mathbf{Z})^\times$ has normal order $\frac{1}{2}(\log\log n)^2$. A cyclic Sylow subgroup that is elementary abelian is counted by $E_1(n)$. That $E_1(n)$ has normal order $\frac{1}{2}(\log\log n)^2$ now follows from our next result, which also seems to be of independent interest.

Theorem 1.2. *Let $F(n)$ denote the count of primes p for which the p-Sylow subgroup of $(\mathbf{Z}/n\mathbf{Z})^\times$ is* not *elementary abelian. Then $F(n)$ has normal order*

$$2\sqrt{\pi}\frac{\sqrt{\log\log n}}{\log\log\log n}.$$

The proofs of Theorems 1.1 and 1.2 are variants of the argument used in [**13**]. Theorem 1.1 is proved in §2–§3, while Theorem 1.2 is demonstrated in §4–§5.

We take this opportunity to inform the reader that several further statistical problems concerning the groups $(\mathbf{Z}/n\mathbf{Z})^\times$ have been considered in papers of Martin and collaborators; see [**2**], [**10**], [**3**], [**8**], [**7**], and [**9**].

Notation. In what follows, the kth iterate of the natural logarithm function will be abbreviated to $\log_k$. The variables l, p, and q are always to be understood as running only over primes.

2. Preparation for the proof of Theorem 1.1

The following result of Gauss was alluded to already in the introduction.

Lemma 2.1. *Let n be a positive integer, and factor $n = \prod_p p^{v_p}$. Then $(\mathbf{Z}/n\mathbf{Z})^\times \cong \prod_{p\mid n}(\mathbf{Z}/p^{v_p}\mathbf{Z})^\times$. If p is odd, or if $v_p \leq 2$, then $(\mathbf{Z}/p^{v_p}\mathbf{Z})^\times \cong \mathbf{Z}/\varphi(p^{v_p})\mathbf{Z}$. When $p = 2$ and $v_2 \geq 3$, we have $(\mathbf{Z}/2^{v_2}\mathbf{Z})^\times \cong \mathbf{Z}/2^{v_2-2}\mathbf{Z} \oplus \mathbf{Z}/2\mathbf{Z}$.*

Lemma 2.1 implies that there are three possible ways the elementary abelian p-group of rank k can appear as the p-Sylow subgroup of $(\mathbf{Z}/n\mathbf{Z})^\times$:

I. $p^2 \parallel n$, and there are exactly $k-1$ distinct primes $q \mid n$ such that $q \equiv 1 \pmod{p}$, where furthermore none of these q satisfy $q \equiv 1 \pmod{p^2}$.

II. $p^2 \nmid n$, and there are exactly k distinct primes $q \mid n$ such that $q \equiv 1 \pmod{p}$, where furthermore none of these q satisfy $q \equiv 1 \pmod{p^2}$.

III. $p = 2$, $2^3 \parallel n$, and there are exactly $k-2$ distinct odd primes $q \mid n$, all of which satisfy $q \equiv 3 \pmod 4$.

Hence, $E_k(n)$ is the count of primes p satisfying one of (I)–(III).

The next two lemmas are the primary technical input to our arguments (excepting "standard" results). The first is a consequence of the Fundamental Lemma of the Sieve [**6**, Lemma 7.2], while the second is an estimate for sums of reciprocals of primes in progressions, located in papers of Pomerance [**14**, p. 219] and Norton [**12**, p. 699].

Lemma 2.2. *Let $x \geq z \geq 2$. If $\mathcal{P}$ is any set of primes not exceeding z, then*

$$\#\{n \leq x : p \mid n \Rightarrow p \notin \mathcal{P}\} = \left(x \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right)\right) \left(1 + O\big(\mathrm{e}^{-\frac{u}{2}}\big)\right)$$

where $u = \frac{\log x}{\log z}$. Here the implied constant is absolute.

Lemma 2.3. *Let m be a positive integer, and let $x \geq 3$. With*

$$S(x; m) := \sum_{\substack{l \leq x \\ l \equiv 1 \pmod m}} \frac{1}{l},$$

we have

$$S(x; m) = \frac{\log_2 x}{\varphi(m)} + O\left(\frac{\log(2m)}{\varphi(m)}\right).$$

Again, the implied constant is absolute.

Instead of directly studying the function $E_k(n)$, we follow [**13**] and introduce a more tractable proxy function. Let

$$\mathcal{I} = (\log_2 x / \log_3 x, \log_2 x \log_3 x]$$

and define $\tilde{E}_k(n)$ as the count of primes $p \in \mathcal{I}$ for which condition I' below is satisfied.

I$'$: there are precisely k distinct primes $q \leq x^{\frac{1}{2\log_3 x}}$ dividing n for which $q \equiv 1 \pmod p$, and furthermore, none of these q satisfy $q \equiv 1 \pmod{p^2}$.

Lemma 2.4. *For all sufficiently large values of x,*

$$\sum_{n \leq x} |E_k(n) - \tilde{E}_k(n)| = O\left(x \log_2 x / (\log_3 x)^2\right).$$

Thus, if $\xi(x)$ is any function tending to infinity, then

$$|E_k(n) - \tilde{E}_k(n)| < \xi(x) \frac{1}{k(k-1)} \log_2 x / (\log_3 x)^2$$

for almost all $n \leq x$.

Before giving the proof, we say a few words about how to interpret Lemma 2.4. Condition (III), by its very definition, can be satisfied only for $p = 2$. Furthermore, condition (I) can be satisfied only when $p^2 \mid n$, and for most n there are very few such p. Thus, one expects the dominant contribution to $E_k(n)$ to come from primes p satisfying condition (II). Roughly speaking, Lemma 2.4 will tell us that — on average over $n \leq x$ — $E_k(n)$ is governed by the count of $p \in \mathcal{I}$ for which (II) is satisfied with primes q not too close to x.

Proof. Let $\mathcal{E}_k(n)$ denote the set of primes p satisfying one of (I), (II), or (III), and let $\tilde{\mathcal{E}}_k(n)$ denote the set of primes satisfying (I$'$); then $E_k(n) = \#\mathcal{E}_k(n)$ while $\tilde{E}_k(n) = \#\tilde{\mathcal{E}}_k(n)$. Put $\Delta(n) := |E_k(n) - \tilde{E}_k(n)|$. Then

$$|\Delta(n)| \leq \#(\mathcal{E}_k(n) \dot{-} \tilde{\mathcal{E}}_k(n)),$$

where $\dot{-}$ denotes setwise symmetric difference. If p belongs to the right-hand symmetric difference, then p satisfies at least one of the following four conditions:

(i) $p \le \log_2 x/\log_3 x$,
(ii) $p > \log_2 x/\log_3 x$ and $p^2 \mid n$,
(iii) $p \in \mathcal{I}$ and $q \mid n$ for some prime $q \equiv 1 \pmod p$, $q > x^{1/2\log_3 x}$,
(iv) $p > \log_2 x \log_3 x$ and $qq' \mid n$ for distinct primes $q, q' \equiv 1 \pmod p$.

Let Δ_0 denote the total count of primes $p \le \log_2 x/\log_3 x$, and let $\Delta_1(n)$ be the count of primes satisfying one of conditions (ii)—(iv). By the prime number theorem, $\Delta_0 \ll \log_2 x/(\log_3 x)^2$. On the other hand, it is shown on [**13**, p. 207] that

$$\sum_{n\le x} \Delta_1(n) \ll \frac{x\log_2 x}{(\log_3 x)^2}.$$

Thus, $\sum_{n\le x} |\Delta(n)| \le \sum_{n\le x}(\Delta_0 + \Delta_1(n)) \ll x\log_2 x/(\log_3 x)^2$. ☐

To complete the proof of Theorem 1.1, it suffices to show that

$$\sum_{n\le x}\left(\tilde{E}_k(n) - \frac{1}{k(k-1)}\frac{\log_2 x}{\log_3 x}\right)^2 = o\left(x\left(\frac{\log_2 x}{\log_3 x}\right)^2\right). \tag{3}$$

Once equality (3) is proved, it follows immediately that for any fixed $\epsilon > 0$, almost all $n \le x$ have

$$\left|\tilde{E}_k(n) - \frac{1}{k(k-1)}\frac{\log_2 x}{\log_3 x}\right| < \epsilon\frac{1}{k(k-1)}\frac{\log_2(x)}{\log_3(x)}.$$

Lemma 2.4 then allows us to replace $\tilde{E}_k(n)$ here with $E_k(n)$. The resulting statement is equivalent to Theorem 1.1, upon observing that $\frac{\log_2 n}{\log_3 n} = \frac{\log_2 x}{\log_3 x} + o(1)$ for $\sqrt{x} < n \le x$ (as $x \to \infty$).

3. Proof of Theorem 1.1

The following two lemmas suffice to establish (3).

Lemma 3.1. *As $x \to \infty$,*

$$\frac{1}{x}\sum_{n\le x}\tilde{E}_k(n) \sim \frac{1}{k(k-1)}\frac{\log_2 x}{\log_3 x}.$$

Lemma 3.2. *As $x \to \infty$,*

$$\frac{1}{x}\sum_{n\le x}\tilde{E}_k(n)^2 \sim \left(\frac{1}{k(k-1)}\frac{\log_2 x}{\log_3 x}\right)^2.$$

Before proceeding to the proofs of the lemmas, we warn the reader that implied constants below may always depend on the fixed parameter k.

Proof of Lemma 3.1. For each prime $p \in \mathcal{I}$ and each large real number x, define the set

$$\mathcal{N}_p(x) = \{n \le x : \text{ there are exactly } k \text{ distinct primes } q \le x^{\frac{1}{2\log_3 x}}, \\ q \equiv 1 \pmod p \text{ dividing } n, \text{ and none of these have } q \equiv 1 \pmod{p^2}\}.$$

Recalling the definition of condition (I′), we have after inverting the order of summation that

$$\sum_{n\le x}\tilde{E}_k(n) = \sum_{p\in\mathcal{I}}\#\mathcal{N}_p(x).$$

If $n \in \mathcal{N}_p(x)$, then there are precisely k primes $q \le x^{1/2\log_3 x}$, $q \equiv 1 \pmod p$, dividing n. Furthermore, $p \parallel q-1$ for each of these q. We decompose

$$\mathcal{N}_p(x) = \mathcal{N}_p^{(0)}(x) \cup \mathcal{N}_p^{(1)}(x),$$

where $\mathcal{N}_p^{(0)}(x)$ contains those $n \in \mathcal{N}_p(x)$ for which each of these k primes q appears to the first power only, and where $\mathcal{N}_p^{(1)}(x)$ contains all other elements of $\mathcal{N}_p(x)$. Note that $q \equiv 1 \pmod p$ implies that $q > p > \log_2 x/\log_3 x$. Thus,

$$\begin{aligned}\sum_{p\in\mathcal{I}} \#\mathcal{N}_p^{(1)}(x) &\ll \sum_{p\in\mathcal{I}} \#\{n \le x : q^2 \mid n \text{ for some } q > \log_2 x/\log_3 x\}\\ &\ll \sum_{p\in\mathcal{I}} \sum_{q>\log_2 x/\log_3 x} \frac{x}{q^2} \ll \sum_{p\in\mathcal{I}} \frac{x}{\log_2 x} \ll x,\end{aligned}$$

which is $o(x\log_2 x/\log_3 x)$. Thus, Lemma 3.1 will be proved once we show that

$$\sum_{p\in\mathcal{I}} \#\mathcal{N}_p^{(0)}(x) \sim \frac{x}{k(k-1)}\frac{\log_2 x}{\log_3 x}.$$

For every prime $p \in \mathcal{I}$,

(4) $\#\mathcal{N}_p^{(0)}(x)$

$$= \sum_{\substack{q_1<\dots<q_k\le x^{1/2\log_3 x}\\ p\,\parallel\, q_i-1 \text{ for each } i}} \#\left\{m \le \frac{x}{q_1\cdots q_k} : \text{no prime } q \equiv 1 \pmod p,\ q \le x^{\frac{1}{2\log_3 x}} \text{ divides } m\right\}.$$

The right-hand sum may be estimated using the sieve (Lemma 2.2). Noting that $\frac{\log(x/q_1\cdots q_k)}{\log(x^{1/2\log_3 x})} \ge 2\log_3 x - k$ for large x, Lemma 2.2 gives us that

$$\begin{aligned}&\#\{m \le x/q_1\cdots q_k\colon \text{no prime } q \equiv 1 \pmod p,\ q \le x^{1/2\log_3 x} \text{ divides } m\}\\ &\quad= \left(\frac{x}{q_1\cdots q_k}\prod_{\substack{q\le x^{\frac{1}{2\log_3 x}}\\ q\equiv 1 \pmod p}}\left(1-\frac{1}{q}\right)\right)\left(1+O\left(\frac{1}{\log_2 x}\right)\right)\\ &\quad= \frac{x}{q_1\cdots q_k}\exp\left(-\sum_{\substack{q\le x^{\frac{1}{2\log_3 x}}\\ q\equiv 1 \pmod p}}\frac{1}{q}\right)\left(1+O\left(\frac{1}{\log_2 x}\right)\right).\end{aligned}$$

(We used that $\log(1-1/q) = -1/q + O(1/q^2)$, and that $\sum_{q\equiv 1 \pmod p} 1/q^2 \le \sum_{q>\frac{\log_2 x}{\log_3 x}} 1/q^2 \ll 1/\log_2 x$.) By Lemma 2.3,

$$\begin{aligned}\sum_{\substack{q\le x^{\frac{1}{2\log_3 x}}\\ q\equiv 1 \pmod p}}\frac{1}{q} &= \frac{\log_2(x^{1/2\log_3 x})}{p-1} + O\left(\frac{\log(2p)}{p-1}\right)\\ &= \frac{\log_2 x}{p} + O\left(\frac{(\log_3 x)^2}{\log_2 x}\right).\end{aligned}$$

Therefore,

(5) $\#\{m \le x/q_1\cdots q_k\colon \text{no prime } q \equiv 1 \pmod p,\ q \le x^{1/2\log_3 x} \text{ divides } m\}$

$$= \frac{x}{q_1\cdots q_k}\exp\left(-\frac{\log_2 x}{p}\right)\left(1+O\left(\frac{(\log_3 x)^2}{\log_2 x}\right)\right).$$

We insert this estimate back into (4) and sum on $q_1,\dots,q_k$.

At the cost of later inserting a factor of $1/k!$, we may replace the condition in (4) that $q_1 < \dots < q_k$ by the simpler condition that $q_1, \dots, q_k$ are distinct. If $1 \le i \le k$, and we have already chosen $q_1, \dots, q_{i-1}$, then

$$\begin{aligned}\sum_{\substack{q_i \le x^{1/2\log_3 x} \\ p \| q_i - 1 \\ q_i \text{ is distinct from } q_1, \dots, q_{i-1}}} \frac{1}{q_i} &= \sum_{\substack{q \le x^{1/2\log_3 x} \\ p \| q-1}} \frac{1}{q} + O\left(i/p\right) \\ &= \frac{\log_2(x^{1/2\log_3 x})}{\varphi(p)} - \frac{\log_2(x^{1/2\log_3 x})}{\varphi(p^2)} + O\left(\frac{\log(2p)}{p}\right) \\ &= \frac{\log_2 x}{p} + O\left(\frac{(\log_3 x)^2}{\log_2 x}\right).\end{aligned}$$

(In the first line, we use that each of $q_1, \dots, q_{i-1} \equiv 1 \pmod p$, hence each q_i is larger than p. This implies that $1/q_1 + \dots + 1/q_{i-1} \le (i-1)/p < i/p$.) Summing successively on $q_1, q_2, \dots, q_k$, we conclude that

$$\sum_{\substack{q_1, \dots, q_k \text{ distinct} \\ p \| q_i - 1 \text{ for each } i}} \frac{1}{q_1 \cdots q_k} = \frac{(\log_2 x)^k}{p^k} \left(1 + O\left(p \frac{(\log_3 x)^2}{(\log_2 x)^2}\right)\right),$$

and thus (being mindful that $p \in \mathcal{I}$)

$$\sum_{\substack{q_1 < q_2 < \dots < q_k \\ p \| q_i - 1 \text{ for each } i}} \frac{1}{q_1 \cdots q_k} = \frac{1}{k!} \frac{(\log_2 x)^k}{p^k} \left(1 + O\left(\frac{(\log_3 x)^3}{\log_2 x}\right)\right). \tag{6}$$

It follows from (4), (5), and (6) that

$$\frac{1}{x} \# \mathcal{N}_p^{(0)}(x) = \frac{1}{k!} \frac{(\log_2 x)^k}{p^k} \exp\left(-\frac{\log_2 x}{p}\right) \left(1 + O\left(\frac{(\log_3 x)^3}{\log_2 x}\right)\right). \tag{7}$$

We are at last in a position to sum over $p \in \mathcal{I}$. We write $\pi(t) = \int_2^t \mathrm{d}v / \log v + E(t)$, so that $E(t) \ll_A t/(\log t)^A$ for any fixed A and all $t \ge 2$. Setting

$$z_1 := \frac{\log_2 x}{\log_3 x} \quad \text{and} \quad z_2 := \log_2 x \log_3 x,$$

we have

$$\begin{aligned}&\sum_{p \in \mathcal{I}} \left(\frac{\log_2 x}{p}\right)^k \exp\left(-\frac{\log_2 x}{p}\right) \\ &\qquad = \int_{z_1}^{z_2} \left(\frac{\log_2 x}{t}\right)^k \exp\left(-\frac{\log_2 x}{t}\right) \frac{\mathrm{d}t}{\log t} + \int_{z_1}^{z_2} \left(\frac{\log_2 x}{t}\right)^k \exp\left(-\frac{\log_2 x}{t}\right) \mathrm{d}E(t).\end{aligned}$$

Integrating by parts bounds the second integral as

$$\ll \left(\sup_{t \in \mathcal{I}} |E(t)|\right) \left(1 + \int_{z_1}^{z_2} \left| \frac{\mathrm{d}}{\mathrm{d}t} \left(\left(\frac{\log_2 x}{t}\right)^k \exp\left(-\frac{\log_2 x}{t}\right)\right) \right| \mathrm{d}t \right) \ll \frac{\log_2 x}{(\log_3 x)^A},$$

for any fixed A. Turning to the first integral, we observe that $\log t = (1 + o(1)) \log_3 x$ uniformly for $t \in [z_1, z_2]$, so that

$$\int_{z_1}^{z_2} \left(\frac{\log_2 x}{t}\right)^k \exp\left(-\frac{\log_2 x}{t}\right) \frac{\mathrm{d}t}{\log t} = \frac{(1 + o(1))}{\log_3 x} \int_{z_1}^{z_2} \left(\frac{\log_2 x}{t}\right)^k \exp\left(-\frac{\log_2 x}{t}\right) \mathrm{d}t.$$

The right-hand integrand has antiderivative

$$(\log_2 x)(k-2)!\exp\left(-\frac{\log_2 x}{t}\right)\sum_{j=0}^{k-2}\frac{1}{j!}\left(\frac{\log_2 x}{t}\right)^j,$$

so that (making simple computations)

$$\int_{z_1}^{z_2}\left(\frac{\log_2 x}{t}\right)^k \exp\left(-\frac{\log_2 x}{t}\right)\mathrm{d}t \sim (k-2)!\log_2 x.$$

Hence,

$$\sum_{p\in\mathcal{I}}\left(\frac{\log_2 x}{p}\right)^k \exp\left(-\frac{\log_2 x}{p}\right) \sim (k-2)!\frac{\log_2 x}{\log_3 x}. \tag{8}$$

We now conclude from (7) that

$$\frac{1}{x}\sum_{p\in\mathcal{I}}\#\mathcal{N}_p^{(0)}(x) \sim \frac{(k-2)!}{k!}\frac{\log_2 x}{\log_3 x} = \frac{1}{k(k-1)}\frac{\log_2 x}{\log_3 x}.$$

As discussed earlier, this estimate completes the proof of the lemma. □

Proof of Lemma 3.2. For each prime $p\in\mathcal{I}$, we let $\mathrm{I}(p)$ denote the following condition on a positive integer n.

$\mathrm{I}(p)$: there are precisely k distinct primes $q\le x^{1/2\log_3 x}$ dividing n for which $q\equiv 1 \pmod p$, and furthermore, none of these satisfy $q\equiv 1 \pmod{p^2}$.

That is, n satisfies $I(p)$ precisely when p belongs to $\tilde{E}_k(n)$. Note that condition $I(p)$ was labeled I' in §2; for our present purposes it is important to track the dependence on p. With this in mind, we have

$$\sum_{n\le x}\tilde{E}_k(n)^2 = \sum_{n\le x}\left(\sum_{p\in\mathcal{I}}\mathbf{1}_{\mathrm{I}(p)}(n)\right)^2 = \sum_{p_1,p_2\in\mathcal{I}}\sum_{n\le x}\mathbf{1}_{\mathrm{I}(p_1)}(n)\cdot\mathbf{1}_{\mathrm{I}(p_2)}(n).$$

To ease notation, let

$$N_{p_1,p_2}(x) := \sum_{n\le x}\mathbf{1}_{\mathrm{I}(p_1)}(n)\cdot\mathbf{1}_{\mathrm{I}(p_2)}(n),$$

so that our task is to estimate $\sum_{p_1,p_2\in\mathcal{I}}N_{p_1,p_2}(x)$. separaremos las contribuciones $p_1=p_2$ de las de los $p_1\ne p_2$.

Case 1: $p_1=p_2$**.** We have that

$$\sum_{p\in\mathcal{I}}N_{p,p}(x) = \sum_{p\in\mathcal{I}}\sum_{n\le x}\mathbf{1}_{\mathrm{I}(p)}(n) = \sum_{n\le x}\tilde{E}_k(n) \ll x\log_2 x/\log_3 x,$$

by Lemma 3.1. This is $o(x(\log_2 x/\log_3 x)^2)$, and so (referring back to the statement of Lemma 3.2) may be considered negligible.

Case 2: $p_1\ne p_2$**.** If $p_1\ne p_2$, then the values of n counted by $N_{p_1,p_2}(x)$ include all $n=q_1\cdots q_k l_1\cdots l_k m\le x$ where

- $q_1,\dots,q_k,l_1,\dots,l_k\le x^{\frac{1}{2\log_3 x}}$ are distinct primes,
- $p_1 \parallel q_1-1,\dots,q_k-1$, and $p_2\nmid q_1-1,\dots,q_k-1$,
- $p_2 \parallel l_1-1,\dots,l_k-1$, and $p_1\nmid l_1-1,\dots,l_k-1$,
- m is free of prime factors $\equiv 1 \pmod{p_1}$ or $\equiv 1 \pmod{p_2}$ not exceeding $x^{1/2\log_3 x}$.

We refer to values of n of this form as being of the **first kind** (with respect to p_1, p_2), and we say that all other n counted by $N_{p_1,p_2}(x)$ are of the **second kind**. If n is of the second kind, then either $q^2 \mid n$ for some prime $q > \log_2 x/\log_3 x$, or $q \mid n$ for some prime $q \equiv 1 \pmod{p_1 p_2}$. The number of n satisfying at least one of these second kind conditions is

$$\ll x \sum_{q > \log_2 x/\log_3 x} \frac{1}{q^2} + x \sum_{\substack{q \le x \\ q \equiv 1 \pmod{p_1 p_2}}} \frac{1}{q} \ll \frac{x}{\log_2 x} + \frac{x \log_2 x}{p_1 p_2}.$$

Summing on $p_1, p_2 \in \mathcal{I}$, we see that the n's of the second kind contribute $O(x \log_2 x)$ to $\sum_{p_1,p_2 \in \mathcal{I}} N_{p_1,p_2}(x)$. This is $o(x(\log_2 x/\log_3 x)^2)$, and so is negligible for us.

Now we move our attention over to the case when n is of the first kind. Given p_1, p_2 and given $q_1, \dots, q_k$, $l_1, \dots, l_k$, applying Lemma 2.2 in the same manner as the proof of Lemma 3.1 yields that the count of corresponding n is

$$\left(\frac{x}{q_1 \cdots q_k l_1 \cdots l_k} \prod_{\substack{q \le x^{\frac{1}{2\log_3 x}} \\ q \equiv 1 \pmod{p_1} \text{ or} \\ q \equiv 1 \pmod{p_2}}} \left(1 - \frac{1}{q}\right) \right) \left(1 + O\left(\frac{1}{\log_2 x}\right)\right) \tag{9}$$
$$= \left(\frac{x}{q_1 \cdots q_k l_1 \dots l_k} \exp\left(- \sum_{\substack{q \le x^{\frac{1}{2\log_3 x}} \\ q \equiv 1 \pmod{p_1} \text{ or} \\ q \equiv 1 \pmod{p_2}}} \frac{1}{q} \right) \right) \left(1 + O\left(\frac{1}{\log_2 x}\right)\right).$$

For $p_1, p_2 \in \mathcal{I}$ with $p_1 \neq p_2$,

$$\sum_{\substack{q \le x^{1/2\log_3 x} \\ q \equiv 1 \pmod{p_1} \text{ or} \\ q \equiv 1 \pmod{p_2}}} \frac{1}{q} = \frac{\log_2 (x^{1/2\log_3 x})}{p_1 - 1} + \frac{\log_2 (x^{1/2\log_3 x})}{p_2 - 1} - \frac{\log_2 (x^{1/2\log_3 x})}{(p_1-1)(p_2-1)}$$
$$+ O\left(\frac{\log(2p)}{p} + \frac{\log(2q)}{q} + \frac{\log(2pq)}{pq}\right) = \frac{\log_2 x}{p_1} + \frac{\log_2 x}{p_2} + O\left(\frac{(\log_3 x)^2}{\log_2 x}\right).$$

Putting this back into (9), we see that the n's of the first kind corresponding to $p_1, p_2 \in \mathcal{I}$ and prescribed primes $q_1, \dots, q_k, l_1, \dots, l_k$ is

$$\frac{x}{q_1 \cdots q_k l_1 \cdots l_k} \exp\left(-\frac{\log_2 x}{p_1}\right) \exp\left(-\frac{\log_2 x}{p_2}\right) \left(1 + O\left(\frac{(\log_3 x)^2}{\log_2 x}\right)\right).$$

We now sum on $q_1, \dots, q_k$, and $l_1, \dots, l_k$. If we have chosen $q_1, \dots, q_{i-1}$ already, then

$$\sum_{\substack{q_i \le x^{1/2\log_3 x} \\ p_1 \| q_i - 1 \\ q_i \not\equiv 1 \pmod{p_2} \\ q_i \text{ is distinct from } q_1, \dots, q_{i-1}}} \frac{1}{q_i} = \sum_{\substack{q \le x^{1/2\log_3 x} \\ q \equiv 1 \pmod{p_1}}} \frac{1}{q} + O\left(\sum_{\substack{q \le x \\ q \equiv 1 \pmod{p_1^2}}} \frac{1}{q} + \sum_{\substack{q \le x \\ q \equiv 1 \pmod{p_1 p_2}}} \frac{1}{q} + \frac{i}{p_1} \right)$$
$$= \frac{\log_2(x^{1/2\log_3 x})}{p_1 - 1} + O\left(\frac{\log_2 x}{p_1^2} + \frac{\log_2 x}{p_1 p_2} + \frac{i}{p_1}\right)$$
$$= \frac{\log_2 x}{p_1} + O\left(\frac{(\log_3 x)^2}{\log_2 x}\right).$$

An analogous estimate holds for the reciprocal sums of the primes l_i. Proceeding as in the proof of Lemma 3.1, we conclude that the contribution to $N_{p_1,p_2}(x)$ from all n's of the first kind is

$$x\left(\frac{1}{k!}\left(\frac{\log_2 x}{p_1}\right)^k \exp\left(-\frac{\log_2 x}{p_1}\right)\right)\left(\frac{1}{k!}\left(\frac{\log_2 x}{p_2}\right)^k \exp\left(-\frac{\log_2 x}{p_2}\right)\right)\left(1+O\left(\frac{(\log_3 x)^3}{\log_2 x}\right)\right).$$

(The factors of $k!$ are explained by our freedom to permute the lists of primes $q_1, \dots, q_k$ and $l_1, \dots, l_k$, independently.)

Finally, we sum this last expression on $p_1, p_2 \in \mathcal{I}$ with $p_1 \neq p_2$. If we insert the terms with $p_1 = p_2$, the sum increases by $O(x \log_2 x)$, which is $o(x(\log_2 x/\log_3 x)^2)$ and so is negligible for us. Once these terms are included, the sum factors as

$$(1+o(1))\frac{x}{k!^2}\left(\sum_{p\in\mathcal{I}}\left(\frac{\log_2 x}{p}\right)^k \exp\left(-\frac{\log_2 x}{p}\right)\right)^2.$$

The remaining sum on p was found in (8) to be $(1+o(1))(k-2)!\frac{\log_2 x}{\log_3 x}$. We conclude that the contribution to $\sum_{p_1,p_2\in\mathcal{I},\, p_1\neq p_2} N_{p_1,p_2}(x)$ of all n's of the first kind is

$$(1+o(1))x\left(\frac{1}{k(k-1)}\frac{\log_2 x}{\log_3 x}\right)^2.$$

Collecting estimates completes the proof of the lemma. □

4. Preparation for the proof of Theorem 1.2

We proceed similarly to Theorem 1.1. We let J(p) denote the following condition on a positive integer n:

J(p): There is a prime $q \mid n$, $q \leq x^{1/2\log_3 x}$ for which $p^2 \mid q-1$.

We also define $\mathcal{J}$ as the interval

$$\mathcal{J} = (\sqrt{\log_2 x/\log_3 x}, \sqrt{\log_2 x \log_3 x}],$$

and we let

$$\tilde{F}(n) = \sum_{p\in\mathcal{J}} \mathbf{1}_{\mathrm{J}(p)}(n).$$

The following is the appropriate analogue of Lemma 2.4.

Lemma 4.1. *For all large values of x,*

$$\sum_{n\leq x} |F(n) - \tilde{F}(n)| = O(x\sqrt{\log_2 x/(\log_3 x)^3}).$$

In particular, whenever $\xi(x) \to \infty$, almost all $n \leq x$ satisfy

$$|F(n) - \tilde{F}(n)| < \xi(x)\sqrt{\log_2 x/(\log_3 x)^3}.$$

Proof. Let p be a prime. If $p^2 \mid q-1$ for some $q \mid n$, then the p-Sylow subgroup of $(\mathbf{Z}/n\mathbf{Z})^\times$ is not elementary abelian. Hence, $\tilde{F}(n) \leq F(n)$ always.

On the other hand, if the p-Sylow subgroup of $(\mathbf{Z}/n\mathbf{Z})^\times$ is not elementary abelian, then either $p^3 \mid n$ or there is a prime q dividing n for which $p^2 \mid q-1$. Thus, if p is counted by $F(n)$ but not by $\tilde{F}(n)$, then one of the following holds:

(i) $p \leq \sqrt{\log_2 x/\log_3 x}$,
(ii) $p > \sqrt{\log_2 x/\log_3 x}$ and $p^3 \mid n$,
(iii) $p \in \mathcal{J}$ and there is a prime $q \mid n$, $q > x^{1/2\log_3 x}$ for which $p^2 \mid q-1$,
(iv) $p > \sqrt{\log_2 x \log_3 x}$ and there is a prime $q \mid n$ with $p^2 \mid q-1$.

Let Δ_0 denote the total count of primes $p \le \sqrt{\log_2 x/\log_3 x}$, and let $\Delta_1(n)$ denote the count of primes satisfying one of (ii)–(iv). Then $\Delta_0 \ll \sqrt{\log_2 x/(\log_3 x)^3}$. On the other hand,

$$\sum_{n\le x}\Delta_1(n) \le \sum_{n\le x}\sum_{p>\sqrt{\log_2 x/\log_3 x}}\Bigg(\mathbf{1}_{p^3\mid n} + \mathbf{1}_{p\le\sqrt{\log_2 x\log_3 x}}\sum_{\substack{x^{1/2\log_3 x}<q\le x\\ q\equiv 1 \pmod{p^2}}}\mathbf{1}_{q\mid n}$$
$$+\mathbf{1}_{p>\sqrt{\log_2 x\log_3 x}}\sum_{\substack{q\le x\\ q\equiv 1 \pmod{p^2}}}\mathbf{1}_{q\mid n}\Bigg).$$

This is

$$\begin{aligned}
&\le \sum_{p>\sqrt{\log_2 x/\log_3 x}}\frac{x}{p^3} + \sum_{p\in\mathcal{J}}\sum_{\substack{x^{1/2\log_3 x}<q\le x\\ q\equiv 1 \pmod{p^2}}}\frac{x}{q} + \sum_{p>\sqrt{\log_2 x\log_3 x}}\sum_{\substack{q\le x\\ q\equiv 1 \pmod{p^2}}}\frac{x}{q}\\
&\ll \frac{x}{\log_2 x} + x\log_4 x\sum_{p\in\mathcal{J}}\frac{1}{p^2} + x\log_2 x\sum_{p>\sqrt{\log_2 x\log_3 x}}\frac{1}{p^2}\\
&\ll \frac{x}{\log_2 x} + \frac{x\log_4 x}{\sqrt{\log_2 x\log_3 x}} + \frac{x\log_2 x}{\sqrt{\log_2 x(\log_3 x)^3}}\\
&\ll x\sqrt{\log_2 x/(\log_3 x)^3}.
\end{aligned}$$

(In going from the first line of this display to the second, the sums on q have been estimated using the Brun–Titchmarsh theorem and partial summation.) Therefore, $\sum_{n\le x}|F(n)-\tilde{F}(n)| \le \sum_{n\le x}(\Delta_0+\Delta_1(n)) \ll x\sqrt{\log_2 x/(\log_3 x)^3}$, as desired. $\square$

Note that $\sqrt{\log_2 x/(\log_3 x)^3}$ is of smaller order than $2\sqrt{\pi}\sqrt{\log_2 x}/\log_3 x$. Mimicking the proof of Theorem 1.1, we see that Theorem 1.2 is a consequence of the following analogues of Lemmas 3.1 and 3.2.

Lemma 4.2. *As $x\to\infty$,*

$$\frac{1}{x}\sum_{n\le x}\tilde{F}(n) \sim 2\sqrt{\pi}\frac{\sqrt{\log_2 x}}{\log_3 x}.$$

Lemma 4.3. *As $x\to\infty$,*

$$\frac{1}{x}\sum_{n\le x}\tilde{F}(n)^2 \sim \left(2\sqrt{\pi}\frac{\sqrt{\log_2 x}}{\log_3 x}\right)^2.$$

We prove these in the next section, by arguments paralleling those given for Lemmas 3.1 and 3.2.

5. Proofs of Lemmas 4.2 and 4.3

Proof of Lemma 4.2. Reversing the order of summation in the definition of $\tilde{F}(n)$, we find that $\sum_{n\le x}\tilde{F}(n) = \sum_{p\in\mathcal{J}}\sum_{n\le x}\mathbf{1}_{\mathrm{J}(p)}(n) = \sum_{p\in\mathcal{J}}\sum_{n\le x}(1-\mathbf{1}_{\bar{\mathrm{J}}(p)})$, where $\bar{\mathrm{J}}(p)$ denotes the negation of the condition $\mathrm{J}(p)$. By the sieve (Lemma 2.2),

$$\sum_{n\le x}(1-\mathbf{1}_{\bar{\mathrm{J}}(p)})) = \lfloor x\rfloor - \Bigg(x\prod_{\substack{q\le x^{1/2\log_3 x}\\ q\equiv 1 \pmod{p^2}}}\left(1-\frac{1}{q}\right)\Bigg)\left(1+O\left(\frac{1}{\log_2 x}\right)\right).$$

For each $p \in \mathcal{J}$,

$$\prod_{\substack{q \le x^{1/2\log_3 x}\\ q \equiv 1 \pmod{p^2}}} \left(1 - \frac{1}{q}\right) = \exp\Bigg(-\sum_{\substack{q \le x^{1/2\log_3 x}\\ q \equiv 1 \pmod{p^2}}} \frac{1}{q} + O\Bigg(\sum_{q > \log_2 x/\log_3 x} \frac{1}{q^2}\Bigg)\Bigg)$$
$$= \exp\Bigg(-\sum_{\substack{q \le x^{1/2\log_3 x}\\ q \equiv 1 \pmod{p^2}}} \frac{1}{q}\Bigg)\left(1 + O\left(\frac{1}{\log_2 x}\right)\right).$$

Continuing, we have from Lemma 2.3 that

$$\sum_{\substack{q \le x^{1/2\log_3 x}\\ q \equiv 1 \pmod{p^2}}} \frac{1}{q} = \frac{\log_2(x^{1/2\log_3 x})}{p(p-1)} + O\left(\frac{\log p}{p^2}\right)$$
$$= \frac{\log_2 x}{p(p-1)} + O\left(\frac{(\log_3 x)^2}{\log_2 x}\right)$$
$$= \frac{\log_2 x}{p^2} + O\left(\frac{(\log_3 x)^{3/2}}{(\log_2 x)^{1/2}}\right).$$

Plugging these estimates in above,

$$\sum_{n \le x} (1 - \mathbf{1}_{\mathrm{J}(p)}) = \lfloor x \rfloor - x \exp\left(-\frac{\log_2 x}{p^2}\right)\left(1 + O\left(\frac{(\log_3 x)^{3/2}}{(\log_2 x)^{1/2}}\right)\right)$$
$$= x\left(1 - \exp\left(-\frac{\log_2 x}{p^2}\right)\right) + O\left(x\frac{(\log_3 x)^{3/2}}{(\log_2 x)^{1/2}}\right). \tag{10}$$

Summing the O-expression over $p \in \mathcal{J}$ will result in an error term of size $O(x \log_3 x)$, which is of smaller order than $x\sqrt{\log_2 x}/\log_3 x$. Thus, to complete the proof of Lemma 4.2, it is enough to show that (as $x \to \infty$)

$$\sum_{p \in \mathcal{J}} \left(1 - \exp\left(-\frac{\log_2 x}{p^2}\right)\right) \sim 2\sqrt{\pi}\frac{\sqrt{\log_2 x}}{\log_3 x}.$$

Again, we write $\pi(t) = \int_2^t \mathrm{d}v/\log v + E(t)$. (Re)setting z_1 and z_2 to

$$z_1 := \sqrt{\log_2 x/\log_3 x} \quad \text{and} \quad z_2 := \sqrt{\log_2 x \log_3 x},$$

we have that

$$\sum_{p \in \mathcal{J}} \left(1 - \exp\left(-\frac{\log_2 x}{p^2}\right)\right) = \int_{z_1}^{z_2} \frac{1 - \exp\left(-\frac{\log_2 x}{t^2}\right)}{\log t}\,\mathrm{d}t + \int_{z_1}^{z_2} \left(1 - \exp\left(-\frac{\log_2 x}{t^2}\right)\right)\mathrm{d}E(t).$$

Integrating by parts, we straightforwardly bound the $\mathrm{d}E(t)$-integral as $O(\sqrt{\log_2 x}/(\log_3 x)^A)$, for any fixed A. Turning to the first integral,

$$\int_{z_1}^{z_2} \frac{1 - \exp\left(-\frac{\log_2 x}{t^2}\right)}{\log t}\,\mathrm{d}t = \frac{(1+o(1))}{\frac{1}{2}\log_3 x}\int_{z_1}^{z_2} \left(1 - \exp\left(-\frac{\log_2 x}{t^2}\right)\right)\mathrm{d}t.$$

We change variables, setting $u = \frac{1}{t^2}\log_2 x$. Then $t^2 = \frac{1}{u}\log_2 x$, $2t\,\mathrm{d}t = -\frac{1}{u^2}\log_2 x\,\mathrm{d}u$, and $\mathrm{d}t = -\frac{\sqrt{\log_2 x}}{2u^{3/2}}\,\mathrm{d}u$. Under this substitution,

$$\int_{z_1}^{z_2}\left(1-\exp\left(-\frac{\log_2 x}{t^2}\right)\right)\mathrm{d}t = \frac{1}{2}\sqrt{\log_2 x}\int_{1/\log_3 x}^{\log_3 x}(1-\exp(-u))u^{-3/2}\,\mathrm{d}u.$$

We extend the final integral from $u=0$ to $u=\infty$, noting that $(1-\exp(-u))u^{-3/2}$ is $O(u^{-1/2})$ for u near 0 and $O(u^{-3/2})$ as $u\to\infty$. In this way, we find that

$$\begin{aligned}\int_{1/\log_3 x}^{\log_3 x}(1-\exp(-u))u^{-3/2}\,\mathrm{d}u &= \int_0^\infty(1-\exp(-u))u^{-3/2}\,\mathrm{d}u + O((\log_3 x)^{-1/2})\\ &= 2\int_0^\infty u^{-1/2}\exp(-u)\,\mathrm{d}u + O((\log_3 x)^{-1/2})\\ &= 2\Gamma(1/2) + O((\log_3 x)^{-1/2}).\end{aligned}$$

(We integrated by parts to pass from the first line to the second.) Collecting the last several estimates, we conclude that

$$\begin{aligned}\sum_{p\in\mathcal{J}}\left(1-\exp\left(-\frac{\log_2 x}{p^2}\right)\right) &\sim \frac{\sqrt{\log_2 x}}{\frac{1}{2}\log_3 x}\cdot\frac{1}{2}\cdot 2\Gamma(1/2)\\ &= 2\sqrt{\pi}\frac{\sqrt{\log_2 x}}{\log_3 x}.\end{aligned}\tag{11}$$

As noted above, the lemma follows. □

Proof of Lemma 4.3. Continuing to use $\bar{\mathrm{J}}(p)$ for the negation of the condition $\mathrm{J}(p)$, we have that

$$\begin{aligned}\sum_{n\le x}\tilde{F}(n)^2 - \sum_{n\le x}\tilde{F}(n) &= \sum_{\substack{p_1,p_2\in\mathcal{J}\\ p_1\ne p_2}}\sum_{n\le x}(1-\mathbf{1}_{\bar{\mathrm{J}}(p_1)}(n))(1-\mathbf{1}_{\bar{\mathrm{J}}(p_2)}(n))\\ &= \sum_{\substack{p_1,p_2\in\mathcal{J}\\ p_1\ne p_2}}\sum_{n\le x}\left(1-\mathbf{1}_{\bar{\mathrm{J}}(p_1)}(n)-\mathbf{1}_{\bar{\mathrm{J}}(p_2)}(n)+\mathbf{1}_{\bar{\mathrm{J}}(p_1)}(n)\mathbf{1}_{\bar{\mathrm{J}}(p_2)}(n)\right)\\ &= \sum_{\substack{p_1,p_2\in\mathcal{J}\\ p_1\ne p_2}}(S_{00}(p_1,p_2)+S_{10}(p_1,p_2)+S_{01}(p_1,p_2)+S_{11}(p_1,p_2)),\end{aligned}$$

where $S_{ij}(p_1,p_2) := \sum_{n\le x}(-1)^{i+j}\mathbf{1}_{\bar{\mathrm{J}}(p_1)}(n)^i\mathbf{1}_{\bar{\mathrm{J}}(p_2)}(n)^j$. Trivially, $S_{00}(p_1,p_2) = \lfloor x\rfloor$, while estimates for $S_{01}(p_1,p_2)$ and $S_{10}(p_1,p_2)$ were obtained in the course of proving Lemma 4.2. It remains to estimate $S_{11}(p_1,p_2)$. For each $p_1,p_2\in\mathcal{J}$, Lemma 2.2 gives that

$$S_{11}(p_1,p_2) = \left(x\prod_{\substack{q\le x^{1/2\log_3 x}\\ q\equiv 1 \pmod{p_1^2}\text{ or}\\ q\equiv 1 \pmod{p_2^2}}}\left(1-\frac{1}{q}\right)\right)\left(1+O\left(\frac{1}{\log_2 x}\right)\right).$$

Now

$$\prod_{\substack{q\le x^{1/2\log_3 x}\\ q\equiv 1 \pmod{p_1^2} \text{ or}\\ q\equiv 1 \pmod{p_2^2}}} \left(1-\frac{1}{q}\right) = \exp\Bigg(-\sum_{\substack{q\le x^{1/2\log_3 x}\\ q\equiv 1 \pmod{p_1^2} \text{ or}\\ q\equiv 1 \pmod{p_2^2}}} \frac{1}{q} + O\Bigg(\sum_{q>\log_2 x/\log_3 x} \frac{1}{q^2}\Bigg)\Bigg)$$

$$= \exp\Bigg(-\sum_{\substack{q\le x^{1/2\log_3 x}\\ q\equiv 1 \pmod{p_1^2} \text{ or}\\ q\equiv 1 \pmod{p_2^2}}} \frac{1}{q}\Bigg)\left(1+O\left(\frac{1}{\log_2 x}\right)\right),$$

while (for $p_1 \neq p_2$)

$$\sum_{\substack{q\le x^{1/2\log_3 x}\\ q\equiv 1 \pmod{p_1^2} \text{ or}\\ q\equiv 1 \pmod{p_2^2}}} \frac{1}{q} = \frac{\log_2(x^{1/2\log_3 x})}{p_1(p_1-1)} + \frac{\log_2(x^{1/2\log_3 x})}{p_2(p_2-1)}$$

$$+ O\left(\frac{\log_2 x}{p_1^2p_2^2} + \frac{\log p_1}{p_1^2} + \frac{\log p_2}{p_2^2} + \frac{\log(p_1p_2)}{p_1^2p_2^2}\right).$$

Estimating terms as in the proof of Lemma 4.2, we see that

$$\sum_{\substack{q\le x^{1/2\log_3 x}\\ q\equiv 1 \pmod{p_1^2} \text{ or}\\ q\equiv 1 \pmod{p_2^2}}} \frac{1}{q} = \frac{\log_2 x}{p_1^2} + \frac{\log_2 x}{p_2^2} + O\left(\frac{(\log_3 x)^{3/2}}{(\log_2 x)^{1/2}}\right).$$

Therefore,

$$S_{11}(p_1,p_2) = x\exp\left(-\frac{\log_2 x}{p_1^2}\right)\exp\left(-\frac{\log_2 x}{p_2^2}\right)\left(1+O\left(\frac{(\log_3 x)^{3/2}}{(\log_2 x)^{1/2}}\right)\right)$$

$$= x\exp\left(-\frac{\log_2 x}{p_1^2}\right)\exp\left(-\frac{\log_2 x}{p_2^2}\right) + O\left(x\frac{(\log_3 x)^{3/2}}{(\log_2 x)^{1/2}}\right).$$

From the proof of Lemma 4.2 (see specifically (10)), both S_{10} and S_{01} have the form $x\exp(-\frac{\log_2 x}{p^2}) + O\big(x\frac{(\log_3 x)^{3/2}}{(\log_2 x)^{1/2}}\big)$, where $p = p_1$ for S_{10} and $p = p_2$ for S_{01}. Therefore,

$$\sum_{n\le x}\tilde{F}(n)^2 - \sum_{n\le x}\tilde{F}(n) = S_{00}(p_1,p_2) + S_{10}(p_1,p_2) + S_{01}(p_1,p_2) + S_{11}(p_1,p_2)$$

$$= x\left(1-\exp\left(-\frac{\log_2 x}{p_1}\right)\right)\left(1-\exp\left(-\frac{\log_2 x}{p_2}\right)\right) + O\left(x\frac{(\log_3 x)^{3/2}}{(\log_2 x)^{1/2}}\right).$$

If we sum over $p_1, p_2 \in \mathcal{J}$, $p_1 \neq p_2$, the O-term is $O(x\sqrt{\log_2 x \log_3 x})$, which is of smaller order than $x\log_2 x/(\log_3 x)^2$. Since $\sum_{n\le x}\tilde{F}(n)$ is also $o(x\log_2 x/(\log_3 x)^2)$ (by Lemma 4.2), to finish the proof of Lemma 4.3 it is enough to show that

$$\sum_{\substack{p_1,p_2\in\mathcal{J}\\ p_1\neq p_2}} \left(1-\exp\left(-\frac{\log_2 x}{p_1}\right)\right)\left(1-\exp\left(-\frac{\log_2 x}{p_2}\right)\right) \sim \left(2\sqrt{\pi}\frac{\sqrt{\log_2 x}}{\log_3 x}\right)^2.$$

Were it not for the condition $p_1 \neq p_2$, the sum would factor, and the result would follow from (11). But adding in the terms $p_1 = p_2$ changes the left-hand side by $O(\sqrt{\log_2 x/\log_3 x})$, which is $o(\log_2 x/(\log_3 x)^2)$ and so is negligible. This completes the proof of Lemma 4.3. $\square$

Acknowledgements

Research on this paper was conducted with the following support:

(i) First and second named authors (S.M and G.P.) supported by FONDOCyT, grant number 2022-1D1-085.
(ii) Second named author (G.P.) supported by NSF award DMS-2316986.
(iii) Third named author (P.P.) supported by NSF award DMS-2001581.

We thank Steve Fan and the anonymous referee for helpful comments and corrections.

References

1. W. D. Banks, F. Luca, and I. E. Shparlinski, *Arithmetic properties of $\varphi(n)/\lambda(n)$ and the structure of the multiplicative group modulo n*, Comment. Math. Helv. **81** (2006), 1–22.
2. B. Chang and G. Martin, *The smallest invariant factor of the multiplicative group*, Int. J. Number Theory **16** (2020), 1377–1405.
3. J. Downey and G. Martin, *Counting multiplicative groups with prescribed subgroups*, Int. J. Number Theory **17** (2021), 2087–2112.
4. P. Erdős and M. Kac, *The Gaussian law of errors in the theory of additive number theoretic functions*, Amer. J. Math. **62** (1940), 738–742.
5. P. Erdős and C. Pomerance, *On the normal number of prime factors of $\varphi(n)$*, Rocky Mountain J. Math. **15** (1985), 343–352.
6. H. Halberstam and H.-E. Richert, *Sieve methods*, London Mathematical Society Monographs, no. 4, Academic Press, 1974.
7. M. Hannesson and G. Martin, *Multiplicative groups avoiding a fixed group*, Int. J. Number Theory **21** (2025), 1337–1359.
8. G. Martin and C. Nguyen, *The least primary factor of the multiplicative group*, Int. J. Number Theory **20** (2024), 2509–2527.
9. G. Martin and R. M. Simpson, *The universal profile of the invariant factors of $(\mathbb{Z}/n\mathbb{Z})^\times$*, arXiv preprint: `https://arxiv.org/abs/2504.11452`, 2025.
10. G. Martin and L. Troupe, *The distribution of the number of subgroups of the multiplicative group*, J. Aust. Math. Soc. **108** (2020), no. 1, 46–97.
11. M. R. Murty and V. K. Murty, *Prime divisors of Fourier coefficients of modular forms*, Duke Math. J. **51** (1984), 57–76.
12. K. K. Norton, *On the number of restricted prime factors of an integer. I*, Illinois J. Math. **20** (1976), 681–705.
13. P. Pollack, *The number of non-cyclic Sylow subgroups of the multiplicative group modulo n*, Canad. Math. Bull. **64** (2021), 204–215.
14. C. Pomerance, *On the distribution of amicable numbers*, J. Reine Angew. Math. **293(294)** (1977), 217–222.

Departamento de Ciencias Basicas, Instituto tecnologico de Santo Domingo (INTEC), Santo Domingo 10602, Dominican Republic.

Escuela de Matematica, Universidad Autonoma de Santo Domingo(UASD), Santo Domingo 10105, Dominican Republic.
Email address: `1053240@est.intec.edu.do smorales48@uasd.edu.do`

Department of Mathematics, Smith College, Northampton, MA, 01007
Email address: `gpolanco@smith.edu`

Department of Mathematics, University of Georgia, Athens, GA 30602
Email address: `pollack@uga.edu`