Non-elementary Sylow subgroups of the multiplicative group

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

Abstract. Let $U(\mathbb{Z}/n\mathbb{Z})$ denote the group of units modulo n. In [**11**], Pollack proved that the number $W(n)$ of primes p for which the Sylow p-subgroup of $U(\mathbb{Z}/n\mathbb{Z})$ is noncyclic has normal order

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

These subgroups may be either elementary abelian or non-elementary. It was shown in [**8**] that the function $E_k(n)$, which counts noncyclic, elementary abelian Sylow p-subgroups of $U(\mathbb{Z}/n\mathbb{Z})$ of rank $k \geq 2$, has normal order

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

This paper studies non-elementary Sylow p-subgroups of the multiplicative group $U(\mathbb{Z}/n\mathbb{Z})$. We prove that the set of integers n for which there exists a Sylow p-subgroup of $U(\mathbb{Z}/n\mathbb{Z})$ isomorphic to a group of the form $\mathbb{Z}/p^{e_1}\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/p^{e_k}\mathbb{Z}$, for a fixed partition $e_1 \geq \cdots \geq e_k$ of an integer r with $e_1 \geq 2$, has asymptotic density zero. We further show that, for a fixed partition, there is an asymptotic formula for the number of integers n admitting a Sylow p-subgroup with this structure, which depends only on the number of parts of the partition, the multiplicities of repeated parts, and the integer r.

1. Introduction

For any positive integer n, we write $U(\mathbb{Z}/n\mathbb{Z})$ for the group of units modulo n, also known as the multiplicative group modulo n. This group has order $\varphi(n)$, where φ denotes Euler's totient function. While many authors have looked at properties of $U(\mathbb{Z}/n\mathbb{Z})$, in this paper, we focus specifically on its Sylow p-subgroups[1], which are fundamental in group theory and number theory. If a prime p divides $\varphi(n)$, then the multiplicative group modulo n has a Sylow p-subgroup, which can be either cyclic or noncyclic. A Sylow p-subgroup of order p^r is noncyclic if it is isomorphic to a group $F = \mathbb{Z}/p^{e_1}\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/p^{e_k}\mathbb{Z}$ where $r = e_1 + \cdots + e_k$ and $k \geq 2$.

Several papers have examined the Sylow p-subgroups of the group of units modulo n. Erdős and Pomerance [**3**], as well as Murty and Murty [**9**], independently showed that the function $\omega(\varphi(n))$, which counts the total number of Sylow p-subgroups of the group of units modulo n, has normal order $\frac{1}{2}(\log\log n)^2$. The function $W(n)$, counting the noncyclic Sylow p-subgroups of the group of units modulo n, has also been studied. In 2021, Pollack [**11**] showed, using the second-moment method, that this function has normal order $\log\log n/\log\log\log n$. Similarly, in [**8**], the authors study the function $E_F(n)$, which counts the Sylow p-subgroups of the multiplicative group modulo n that are isomorphic to a group of the form $F = \mathbb{Z}/p^{e_1}\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/p^{e_k}\mathbb{Z}$, where $e_1 \geq \cdots \geq e_k$ are positive integers. They showed that the normal order of $E_F(n)$ is $\frac{1}{k(k-1)}\frac{\log\log n}{\log\log\log n}$ when $e_1 = \cdots = e_k = 1$. Other statistical questions related to the structure of multiplicative groups are addressed in [**1, 2, 7, 6, 12**].

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

[1]Let G be a finite group and p a prime. A subgroup H of G is called a Sylow p-subgroup if $|H| = p^s$ and s is maximal; equivalently, if K is any p-subgroup of G with $|K| = p^t$, then $t \leq s$.

Since

$$\sum_{k=2}^{\infty} \frac{1}{k(k-1)} = 1,$$

it follows that, if we sum the normal orders of all functions $E_F(n)$ with $e_1 = \cdots = e_k = 1$, we obtain the normal order of $W(n)$. This suggests that the typical size of $E_F(n)$ when $e_1 \geq 2$ should be $o\left(\frac{\log\log n}{\log\log\log n}\right)$. Indeed, in this work, we will show that for any $F = \mathbb{Z}/p^{e_1}\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/p^{e_k}\mathbb{Z}$ with $e_1 \geq 2$, the set of positive integers n such that $E_F(n) > 0$ has asymptotic density zero. Specifically, we prove the following theorem.

THEOREM 1.1. *Let $F = \mathbb{Z}/p^{e_1}\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/p^{e_k}\mathbb{Z}$ where $e_1 \geq \cdots \geq e_k$ are positive integers with $e_1 \geq 2$. Then*

$$\#\{n \leq x : E_F(n) > 0\} \sim \left(\frac{(r-2)!}{|A_1|! \cdots |A_{e_1}|!}\right) \frac{x}{\log_3 x (\log_2 x)^{r-k-1}},$$

where $r = e_1 + \cdots + e_k$ and $A_\alpha = \{i \in \{1, \ldots, k\} : e_i = \alpha\}$.

EXAMPLE. *To illustrate the result of Theorem 1.1, consider the following example. Let*

$$F = \mathbb{Z}/p^7\mathbb{Z} \oplus \mathbb{Z}/p^5\mathbb{Z} \oplus \mathbb{Z}/p^5\mathbb{Z} \oplus \mathbb{Z}/p^5\mathbb{Z} \oplus \mathbb{Z}/p\mathbb{Z} \oplus \mathbb{Z}/p\mathbb{Z}.$$

In this case, $k = 6$ and $r = 7+5+5+5+1+1 = 24$. Moreover, $|A_7| = 1, \quad |A_5| = 3, \quad |A_1| = 2$, whereas $|A_\alpha| = 0$ for $\alpha \in \{2, 3, 4, 6\}$. Consequently, Theorem 1.1 implies that

$$\#\{n \leq x : E_F(n) > 0\} \sim \frac{(24-2)!}{2!\,3!\,1!} \frac{x}{\log_3 x\,(\log_2 x)^{24-6-1}} = \frac{22!}{12} \frac{x}{\log_3 x\,(\log_2 x)^{17}}.$$

This result shows, in particular, how the parameters r and k affect the asymptotic order of the proportion of positive integers $n \leq x$ for which some Sylow p-subgroup of $U(\mathbb{Z}/n\mathbb{Z})$ is isomorphic to F. Among groups with the same value of k, a larger value of r makes such integers n less frequent. On the other hand, among groups with the same value of r, a larger value of k makes such integers more frequent. Finally, for fixed r and k, repeated entries in the sequence $e_1, \ldots, e_k$ also affect the asymptotic formula: the larger the product $\prod_{\alpha=1}^{e_1} |A_\alpha|!$, the less frequent such integers n become.

Notation

The letters ℓ, p, and q (possibly with subscripts or other decorations) are always reserved for primes. We write $\log_k$ for the k-fold iterated natural logarithm. We use $\mathbf{1}_C$ for the indicator function of the condition C; for example, $\mathbf{1}_{d|n}$ takes the value 1 when $d \mid n$ and the value 0 otherwise. Implied constants are usually absolute, but in proofs involving a fixed parameter A we allow such constants to depend on A.

2. Lemmata

One of the key facts used in this research is Gauss's lemma on the structure of the group of units modulo n.

LEMMA 2.1. *Let n be a positive integer, and write $n = \prod_{p|n} p^{v_p}$. Then $U(\mathbb{Z}/n\mathbb{Z}) \cong \prod_{p|n} U(\mathbb{Z}/p^{v_p}\mathbb{Z})$. If p is odd or if $v_p \leq 2$, then $U(\mathbb{Z}/p^{v_p}\mathbb{Z}) \cong \mathbb{Z}/\varphi(p^{v_p})\mathbb{Z}$. If $p = 2$ and $v_2 \geq 3$, we have $U(\mathbb{Z}/2^{v_2}\mathbb{Z}) \cong \mathbb{Z}/2^{v_2-2}\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$.*

For our estimates when proving Theorem 1.1, we will use the following variant of the Fundamental Lemma of Sieve Theory [**4**, p. 209, thm. 7.2].

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

$$\#\{n \leq x : p \nmid n \text{ for every } p \in \mathcal{P}\} = \left(x \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right)\right) \left(1 + O(e^{-u/2})\right),$$

where $u = \frac{\log x}{\log z}$.

The following result, obtained independently by Pomerance and Norton **[13, 10]**, will also be useful.

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

$$S(x;m) = \sum_{\substack{\ell \leq x \\ \ell \equiv 1 \pmod m}} \frac{1}{\ell}.$$

Then

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

3. Preparation for the Proof of Theorem 1.1

Throughout this section and the next, write

$$r = e_1 + \cdots + e_k, \qquad y = x^{1/(2\log_3 x)}.$$

Also put

$$z_1 = \frac{\log_2 x}{\log((\log_2 x)^r)}, \qquad z_2 = \log_2 x \log_3 x.$$

We first translate the condition that the Sylow p-subgroup of $U(\mathbb{Z}/n\mathbb{Z})$ is isomorphic to $F = \mathbb{Z}/p^{e_1}\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/p^{e_k}\mathbb{Z}$. By Lemma 2.1, for odd p, this happens if and only if one of the following two conditions holds.

(I) We have $p^2 \nmid n$, and n satisfies the following k**-prime condition**: there are k distinct primes $q_1, \ldots, q_k$ dividing n which are congruent to 1 (mod p), and furthermore, for each $i \in \{1, \ldots, k\}$

$$q_i \equiv 1 \pmod{p^{e_i}}, \qquad q_i \not\equiv 1 \pmod{p^{e_i+1}}.$$

In addition, there are no other primes congruent to 1 (mod p) dividing n besides these $q_1, \ldots, q_k$.

(II) There is an index $i \in \{1, \ldots, k\}$ such that $p^{e_i+1} \| n$, and there are $k-1$ distinct primes q_j dividing n which are congruent to 1 (mod p), indexed by $j \in \{1, \ldots, k\} \setminus \{i\}$, such that

$$q_j \equiv 1 \pmod{p^{e_j}}, \qquad q_j \not\equiv 1 \pmod{p^{e_j+1}}.$$

Furthermore, there are no other primes congruent to 1 (mod p) dividing n besides these $\{q_j\}_{j \in \{1,\ldots,k\} \setminus \{i\}}$.

When $p = 2$, this description must be modified because of the exceptional structure of $U(\mathbb{Z}/2^a\mathbb{Z})$. If $e_k \neq 1$, the Sylow 2-subgroup of $U(\mathbb{Z}/n\mathbb{Z})$ is isomorphic to F if and only if condition (I) holds. If $e_k = 1$, it is isomorphic to F if and only if one of the following holds: condition (I), condition (II) with $p = 2$ and $e_i = 1$, or condition (III).

(III) There is an index $i \in \{1, \ldots, k-1\}$ such that $2^{e_i+2} \| n$, and there are $k-2$ distinct primes $q_j \mid n$, indexed by $j \in \{1, \ldots, k-1\} \setminus \{i\}$, such that

$$q_j \equiv 1 \pmod{2^{e_j}}, \qquad q_j \not\equiv 1 \pmod{2^{e_j+1}}.$$

Moreover, there are no other primes congruent to 1 (mod 2) dividing n besides these $\{q_j\}_{j \in \{1,\ldots,k-1\} \setminus \{i\}}$.

In the remainder of the paper, we will refer to the k-prime condition specified in case (I) above as the main k-prime configuration or the main k-prime condition. The reason for this is that, as we will see later, the main contribution comes precisely from such integers n.

Now define $\tilde{E}_F(n)$ to count only the primes p in the range $\mathcal{I} = (z_1, z_2]$ for which the main k-prime condition occurs in a truncated manner, i.e., using only small primes. More precisely, $\tilde{E}_F(n)$ is the number of primes $p \in \mathcal{I}$ such that

I′: n has k distinct prime divisors $q_1, \dots, q_k \le y$ such that $q_i \equiv 1 \pmod{p^{e_i}}$, and $q_i \not\equiv 1 \pmod{p^{e_i+1}}$ for every $i \in \{1, \dots, k\}$. Further, there are no other primes $q \le y$ congruent to 1 (mod p) dividing n besides these $q_1, \dots, q_k$.

The goal of this section is to show that replacing $E_F(n)$ by $\tilde{E}_F(n)$ introduces only a negligible error.

LEMMA 3.1. *For all large x,*

$$\sum_{n \le x} \left| E_F(n) - \tilde{E}_F(n) \right| = O\left(\frac{x}{(\log_2 x)^{r-k-1} (\log_3 x)^r} \right),$$

where $r = e_1 + \cdots + e_k$.

PROOF. Let $\mathcal{E}_F(n)$ be the set of primes p counted by $E_F(n)$. Thus $p \in \mathcal{E}_F(n)$ if and only if p satisfies one of the structural conditions (I), (II), or (III) above. Similarly, let $\tilde{\mathcal{E}}_F(n)$ be the set of primes $p \in \mathcal{I}$ counted by $\tilde{E}_F(n)$. Thus

$$E_F(n) = \#\mathcal{E}_F(n), \qquad \tilde{E}_F(n) = \#\tilde{\mathcal{E}}_F(n).$$

Put

$$\Delta(n) = \left| E_F(n) - \tilde{E}_F(n) \right|.$$

Then it follows that

$$\Delta(n) \le \# \left(\mathcal{E}_F(n) \mathbin{\dot{-}} \tilde{\mathcal{E}}_F(n) \right),$$

where $\dot{-}$ denotes the symmetric difference. Therefore, in order to prove Lemma 3.1, it is enough to appropriately bound the sum

$$\sum_{n \le x} \# \left(\mathcal{E}_F(n) \mathbin{\dot{-}} \tilde{\mathcal{E}}_F(n) \right).$$

Hence we will focus on counting the primes that are included in one of the two sets but not the other. There are six possible ways this can happen.

Bad cases coming from $\mathcal{E}_F(n) \setminus \tilde{\mathcal{E}}_F(n)$. First suppose that $p \in \mathcal{E}_F(n) \setminus \tilde{\mathcal{E}}_F(n)$. Then p gives a genuine Sylow subgroup of type F, but it is not counted by the truncated function $\tilde{E}_F(n)$. This can occur in one of the following four ways.

Case 1. The prime p satisfies the condition (I), but lies outside the chosen interval $\mathcal{I}$ ($p \le z_1$ or $p > z_2$).

Case 2. The prime p satisfies condition (II).

Case 3. The exceptional $p = 2$ condition (III) occurs.

Case 4. The prime $p \in \mathcal{I}$ satisfies condition (I), but at least one of the required primes is too large. In other words, for some i, $q_i > y = x^{1/(2 \log_3 x)}$.

Bad cases coming from $\tilde{\mathcal{E}}_F(n) \setminus \mathcal{E}_F(n)$. Now suppose that $p \in \tilde{\mathcal{E}}_F(n) \setminus \mathcal{E}_F(n)$. Then $p \in \mathcal{I}$ satisfies the truncated condition I′ defining $\tilde{E}_F(n)$, but it does not give a Sylow subgroup of type F. This can happen in either of the following two ways.

Case 5. The truncated condition I′ holds, but there is an additional large prime divisor $q_{k+1} > y$ of n such that $q_{k+1} \equiv 1 \pmod{p}$. This extra prime contributes an additional p-component and therefore prevents the Sylow p-subgroup from being isomorphic to F.

Case 6. The truncated condition I′ holds, but $p^2 \mid n$. Again, this changes the structure of the Sylow p-subgroup, so p is counted by $\tilde{E}_F(n)$ but not by $E_F(n)$.

Let $\Delta_j(n)$ denote the number of primes p satisfying Case j. Then

$$\Delta(n) \le \sum_{j=1}^{6} \Delta_j(n).$$

It remains to estimate the total contribution of these six cases.

Auxiliary notation. For a prime p and a positive integer Q, write

$$M_p(Q) = \#\left\{m \le \frac{x}{Q} : \text{no prime divisor } q \mid m \text{ satisfies } q \equiv 1 \pmod p\right\}$$

and

$$\tilde{M}_p(Q) = \#\left\{m \le \frac{x}{Q} : m \text{ is free of prime factors } q \le y \text{ with } q \equiv 1 \pmod p\right\}.$$

This notation will be used repeatedly: the integer Q records the forced prime divisors of n, while m records the remaining part of n, which is required to have no further prime divisor congruent to $1 \pmod p$.

Estimate for Case 1. Let $\mathcal{Q}_1(p)$ be the set of integers Q formed as products of powers of k distinct primes $q_1, \dots, q_k \le x$, where $q_i \equiv 1 \pmod{p^{e_i}}$ for each $i \in \{1, \dots, k\}$. The sum

$$\sum_{n \le x} \Delta_1(n)$$

is the number of pairs (n, p) where p is a prime with $p \le z_1$ or $p > z_2$, and $n \le x$ satisfies the condition (I) for p. For each fixed p, each corresponding n has the form Qm, where $Q \in \mathcal{Q}_1(p)$ and m is not divisible by any prime congruent to $1 \pmod p$. Therefore,

$$\sum_{n \le x} \Delta_1(n) \ll \sum_{p \le z_1} \sum_{Q \in \mathcal{Q}_1(p)} M_p(Q) + \sum_{p > z_2} \sum_{Q \in \mathcal{Q}_1(p)} M_p(Q).$$

For $p \le z_1$, we split the sum according as $Q \le x^{1/2}$ or $Q > x^{1/2}$. If $Q \le x^{1/2}$, then Lemma 2.2 gives

$$M_p(Q) \ll \frac{x}{Q} \prod_{\substack{q \le x/Q \\ q \equiv 1 \pmod p}} \left(1 - \frac{1}{q}\right) \ll \frac{x}{Q} \exp\left(\sum_{\substack{q \le \frac{x}{Q} \\ q \equiv 1 \pmod p}} \log\left(1 - \frac{1}{q}\right)\right). \tag{1}$$

Using the facts that $\log\left(1 - \frac{1}{q}\right) = -\frac{1}{q} + O\left(\frac{1}{q^2}\right)$, and that $\sum_q \frac{1}{q^2} \ll 1$, we have that

$$\exp\left(\sum_{\substack{q \le \frac{x}{Q} \\ q \equiv 1 \pmod p}} \log\left(1 - \frac{1}{q}\right)\right) \ll \exp\left(-\sum_{\substack{q \le \frac{x}{Q} \\ q \equiv 1 \pmod p}} \frac{1}{q}\right). \tag{2}$$

Substituting (2) into (1), we obtain

$$M_p(Q) \ll \frac{x}{Q} \exp\left(-\sum_{\substack{q \le \frac{x}{Q} \\ q \equiv 1 \pmod p}} \frac{1}{q}\right).$$

Using Lemma 2.3, along with the facts that $\frac{x}{Q} \geq x^{\frac{1}{2}}$ and $p \leq \log_2 x / \log\left((\log_2 x)^r\right)$, this becomes

$$M_p(Q) \ll \frac{x}{Q} \exp\left(-\frac{\log_2(\frac{x}{Q})}{p} + O\left(\frac{\log(2p)}{p}\right)\right) \ll \frac{x}{Q(\log_2 x)^r}.$$

Also, for fixed p, taking into account that $\frac{1}{q_i} + \frac{1}{q_i^2} + \cdots = \frac{1}{q_i - 1} \ll \frac{1}{q_i}$, applying the Brun–Titchmarsh inequality, which can be found in [**14**, p. 83, thm. 4.16], together with partial summation, we have

$$\begin{aligned}(3)\quad \sum_{Q \in \mathcal{Q}_1(p)} \frac{1}{Q} &\leq \left(\sum_{\substack{q_1 \leq x \\ q_1 \equiv 1 \pmod{p^{e_1}}}} \left(\frac{1}{q_1} + \frac{1}{q_1^2} + \cdots\right)\right) \cdots \left(\sum_{\substack{q_k \leq x \\ q_k \equiv 1 \pmod{p^{e_k}}}} \left(\frac{1}{q_k} + \frac{1}{q_k^2} + \cdots\right)\right) \\ &\ll \prod_{i=1}^{k} \frac{\log_2 x}{p^{e_i}} \\ &\ll \frac{(\log_2 x)^k}{p^r}.\end{aligned}$$

Since $r > 1$, the sum $\sum_{p \leq z_1} \frac{1}{p^r}$ is bounded above by an absolute constant. Thus,

$$(4)\qquad \sum_{p \leq z_1} \sum_{\substack{Q \leq x^{1/2} \\ Q \in \mathcal{Q}_1(p)}} M_p(Q) \ll x \sum_{p \leq z_1} \frac{(\log_2 x)^k}{p^r (\log_2 x)^r} \ll \frac{x}{(\log_2 x)^{r-k}}.$$

We now focus on the case when $Q > x^{1/2}$. To evaluate the double sum corresponding to (4) in this case, we use Landau's bound on the number of integers with a given number of prime factors found in [**5**, p. 491, thm. 437]. This double sum gives

$$\begin{aligned}\sum_{p \leq z_1} \sum_{\substack{Q > x^{1/2} \\ Q \in \mathcal{Q}_1(p)}} M_p(Q) &\ll \sum_{p \leq z_1} \sum_{\substack{m \leq x^{\frac{1}{2}} \\ q \mid m \to q \not\equiv 1 \pmod{p}}} \frac{x(\log_2(\frac{x}{m}))^{k-1}}{m \log(\frac{x}{m})} \\ &\ll \sum_{p \leq z_1} \sum_{\substack{m \leq x^{\frac{1}{2}} \\ q \mid m \to q \not\equiv 1 \pmod{p}}} \frac{x(\log_2 x)^{k-1}}{m \log(x)}.\end{aligned}$$

We have

$$(5)\qquad \sum_{\substack{m \leq x^{\frac{1}{2}} \\ q \mid m \to q \not\equiv 1 \pmod{p}}} \frac{1}{m} \ll \prod_{\substack{q \leq x \\ q \not\equiv 1 \pmod{p}}} \left(1 + \frac{1}{q} + \frac{1}{q^2} + \cdots\right) \ll \frac{\prod_{q \leq x} \left(1 + \frac{1}{q} + \frac{1}{q^2} + \cdots\right)}{\prod_{\substack{q \leq x \\ q \equiv 1 \pmod{p}}} \left(1 + \frac{1}{q} + \frac{1}{q^2} + \cdots\right)}.$$

The numerator in the last term of expression (5) is $O(\log x)$ by Mertens' third theorem. To estimate the denominator in expression (5), using elementary estimates, we have

$$\begin{aligned}\prod_{\substack{q\le x\\ q\equiv 1 \pmod p}}\left(1+\frac{1}{q}+\frac{1}{q^2}+\cdots\right) &= \exp\left(-\sum_{\substack{q\le x\\ q\equiv 1 \pmod p}}\log\left(1-\frac{1}{q}\right)\right)\\ &\gg \exp\left(\sum_{\substack{q\le x\\ q\equiv 1 \pmod p}}\frac{1}{q}\right)\\ &\gg \exp\left(\frac{\log_2 x}{p}\right)\\ &\gg (\log_2 x)^r.\end{aligned}$$

Then,

$$\sum_{\substack{m\le x^{\frac{1}{2}}\\ q\mid m\to q\not\equiv 1 \pmod p}}\frac{x(\log_2 x)^{k-1}}{m\log x}\ll\left(\frac{x(\log_2 x)^{k-1}}{\log x}\right)\left(\frac{\log x}{(\log_2 x)^r}\right)\ll\frac{x}{(\log_2 x)^{r+1-k}}.$$

Then, summing over all $p\le z_1$, we have

$$\sum_{p\le z_1}\sum_{\substack{Q>x^{1/2}\\ Q\in\mathcal{Q}_1(p)}}M_p(Q)\ll\frac{x}{(\log_2 x)^{r-k}}.$$

For $p>z_2$, the trivial bound $M_p(Q)\le x/Q$ together with the result of equation (3) gives

$$\sum_{p>z_2}\sum_{Q\in\mathcal{Q}_1(p)}M_p(Q)\ll x(\log_2 x)^k\sum_{p>z_2}\frac{1}{p^r}\ll\frac{x}{(\log_2 x)^{r-k-1}(\log_3 x)^r}.$$

Thus

$$\sum_{n\le x}\Delta_1(n)\ll\frac{x}{(\log_2 x)^{r-k-1}(\log_3 x)^r}.$$

Estimate for Case 2. In this case n contains a factor p^{e_i+1} for some $i\in\{1,\dots,k\}$, together with $k-1$ auxiliary primes. Let $\mathcal{Q}_2(p)$ denote the corresponding set of forced factors Q. Then

$$\sum_{n\le x}\Delta_2(n)\ll\sum_{p\le z_1}\sum_{\substack{Q\le x^{\frac{1}{2}}\\ Q\in\mathcal{Q}_2(p)}}M_p(Q)+\sum_{p\le z_1}\sum_{\substack{Q>x^{\frac{1}{2}}\\ Q\in\mathcal{Q}_2(p)}}M_p(Q)+\sum_{p>z_1}\sum_{Q\in\mathcal{Q}_2(p)}M_p(Q).$$

In the same way as in equation (4), and taking into account the factor p^{e_i+1} in Q, we obtain

$$\sum_{p\le z_1}\sum_{\substack{Q\le x^{\frac{1}{2}}\\ Q\in\mathcal{Q}_2(p)}}M_p(Q)\ll x\sum_{p\le z_1}\frac{(\log_2 x)^{k-1}}{p^{r+1}(\log_2 x)^r}\ll\frac{x}{(\log_2 x)^{r-k+1}}$$

and using Landau's bound, as in Case 1, it follows that

$$\sum_{p\le z_1}\sum_{\substack{Q>x^{\frac{1}{2}}\\ Q\in\mathcal{Q}_2(p)}}M_p(Q)\ll\frac{x}{(\log_2 x)^{r-k}}.$$

For $p > z_1$, we have

$$\sum_{p>z_1} \sum_{Q\in\mathcal{Q}_2(p)} M_p(Q) \ll \sum_{p>z_1} \sum_{Q\in\mathcal{Q}_2(p)} \frac{x}{Q} \ll \sum_{p>z_1} x\frac{(\log_2 x)^{k-1}}{p^{r+1}} \ll \frac{x(\log_3 x)^{r-1}}{(\log_2 x)^{r+1-k}},$$

and hence

$$\sum_{n\le x} \Delta_2(n) \ll \frac{x}{(\log_2 x)^{r-k}}.$$

This is negligible compared with the claimed error term.

Estimate for Case 3. This is the exceptional case $p = 2$. The integers n counted by $\Delta_3(n)$ are of the form $n = 2^{e_i+2}Q$ for some $i \in \{1, \dots, k-1\}$ and Q is a product of powers of $k-2$ distinct odd primes. For $k \ge 3$, Landau's bound for the number of integers less than x having exactly $k-2$ distinct prime factors gives

$$\sum_{n\le x} \Delta_3(n) \ll \frac{x(\log_2 x)^{k-3}}{\log x}.$$

For $k = 2$, there are no odd prime factors in Q, and hence the contribution is $O(1)$.

Estimate for Case 4. Here $p \in \mathcal{I}$, but one of the required primes q_i is larger than y. Let $\mathcal{Q}_4(p)$ be the set of the corresponding forced factors. Then

$$\sum_{n\le x} \Delta_4(n) \ll \sum_{p\in\mathcal{I}} \sum_{Q\in\mathcal{Q}_4(p)} M_p(Q) \ll \sum_{p\in\mathcal{I}} \sum_{Q\in\mathcal{Q}_4(p)} \frac{x}{Q}.$$

For each integer counted in Case 4, choose an index i such that $q_i > y$. For $j \ne i$, we have

$$\sum_{\substack{q_j\le x\\ q_j\equiv 1 \pmod{p^{e_j}}}} \frac{1}{q_j} \ll \frac{\log_2 x}{p^{e_j}},$$

while the chosen large prime factor contributes

$$\sum_{\substack{y<q_i\le x\\ q_i\equiv 1 \pmod{p^{e_i}}}} \frac{1}{q_i} \ll \frac{\log_4 x}{p^{e_i}}.$$

Here we have applied the Brun–Titchmarsh inequality and partial summation to estimate the sum over q_i. Hence

$$\sum_{Q\in\mathcal{Q}_4(p)} \frac{1}{Q} \ll \frac{(\log_2 x)^{k-1}\log_4 x}{p^r}.$$

Since $z_1 < p \le z_2$ for every $p \in \mathcal{I}$, replacing $1/p^r$ by its upper bound $1/z_1^r$ and summing over all primes $p \le z_2$ gives

$$\sum_{n\le x} \Delta_4(n) \ll x \sum_{p\in\mathcal{I}} \frac{(\log_2 x)^{k-1}\log_4 x}{p^r} \ll x \sum_{p\le z_2} \frac{(\log_3 x)^r \log_4 x}{(\log_2 x)^{r-k+1}} \ll \frac{x(\log_3 x)^{r+1}}{(\log_2 x)^{r-k}}.$$

Estimate for Case 5. Here the main k-prime configuration occurs, and it is also truncated $(q_1, \dots, q_k \le y)$, but there is an additional large prime q_{k+1} satisfying

$$q_{k+1} > y, \qquad q_{k+1} \equiv 1 \pmod{p}.$$

Let $\mathcal{Q}_5(p)$ be the set of corresponding forced factors. Then,

$$\sum_{n\le x} \Delta_5(n) \ll \sum_{p\in\mathcal{I}} \sum_{Q\in\mathcal{Q}_5(p)} \tilde{M}_p(Q) \ll \sum_{p>z_1} \sum_{Q\in\mathcal{Q}_5(p)} \frac{x}{Q}.$$

Similarly to Case 4, for a fixed prime p, we have

$$\sum_{Q\in\mathcal{Q}_5(p)}\frac{1}{Q}\ll\frac{(\log_2 x)^k\log_4 x}{p^{r+1}}.$$

Since

$$z_1=\frac{\log_2 x}{\log((\log_2 x)^r)},$$

this gives

$$\sum_{n\le x}\Delta_5(n)\ll x\frac{\log_4 x(\log_3 x)^{r-1}}{(\log_2 x)^{r-k}}\ll\frac{x}{(\log_2 x)^{r-k-1}(\log_3 x)^r}.$$

Estimate for Case 6. Finally, suppose the truncated k-prime condition holds but $p^2\mid n$. We proceed as in Case 1, but now we obtain an additional factor $\frac{1}{p^2}$, yielding

$$\sum_{n\le x}\Delta_6(n)\ll x\sum_{p\in\mathcal{I}}\frac{(\log_2 x)^k}{p^{r+2}}\ll\frac{x}{(\log_2 x)^{r-k}}.$$

Combining the six estimates, we obtain

$$\sum_{n\le x}\left|E_F(n)-\tilde{E}_F(n)\right|\le\sum_{j=1}^{6}\sum_{n\le x}\Delta_j(n)\ll\frac{x}{(\log_2 x)^{r-k-1}(\log_3 x)^r}.$$

This proves Lemma 3.1. □

4. Proof of Theorem 1.1

Proof. By Lemma 3.1, it is enough to prove that

$$\#\{n\le x:\tilde{E}_F(n)>0\}\sim\frac{(r-2)!}{|A_1|!\cdots|A_{e_1}|!}\cdot\frac{x}{\log_3 x\,(\log_2 x)^{r-k-1}},\tag{6}$$

where $r=e_1+\cdots+e_k$ and $A_\alpha=\{i\in\{1,\dots,k\}:e_i=\alpha\}$.

For every nonnegative integer a, we have

$$2a-a^2\le\mathbf{1}_{a>0}\le a.$$

Applying this with $a=\tilde{E}_F(n)$ and summing over $n\le x$, we obtain

$$\sum_{n\le x}\left(2\tilde{E}_F(n)-\tilde{E}_F(n)^2\right)\le\#\{n\le x:\tilde{E}_F(n)>0\}\le\sum_{n\le x}\tilde{E}_F(n).$$

Thus, (6) follows from the next two lemmas.

Lemma 4.1. *As $x\to\infty$,*

$$\frac{1}{x}\sum_{n\le x}\tilde{E}_F(n)=(1+o(1))\,\frac{(r-2)!}{|A_1|!\cdots|A_{e_1}|!}\cdot\frac{1}{\log_3 x\,(\log_2 x)^{r-k-1}},$$

where $r=e_1+\cdots+e_k$ and $A_\alpha=\{i\in\{1,\dots,k\}:e_i=\alpha\}$.

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

$$\frac{1}{x}\sum_{n\le x}\tilde{E}_F(n)^2=(1+o(1))\,\frac{(r-2)!}{|A_1|!\cdots|A_{e_1}|!}\cdot\frac{1}{\log_3 x\,(\log_2 x)^{r-k-1}},$$

where $r=e_1+\cdots+e_k$ and $A_\alpha=\{i\in\{1,\dots,k\}:e_i=\alpha\}$.

Proof of Lemma 4.1. For a prime $p \in \mathcal{I}$ and $i \in \{1, \dots, k\}$, define

$$\mathcal{Q}(p,i) = \left\{q \le y : \text{ with } q \equiv 1 \pmod{p^{e_i}} \text{ and } q \not\equiv 1 \pmod{p^{e_i+1}}\right\}.$$

Let $N_p(x)$ denote the number of integers $n \le x$ for which there exist k distinct primes $q_1, \dots, q_k \le y$ dividing n, with $q_i \in \mathcal{Q}(p,i)$ for every $i \in \{1, \dots, k\}$, and such that no other prime $q \le y$ dividing n is congruent to 1 (mod p). By the definition of $\tilde{E}_F(n)$,

$$\sum_{n \le x} \tilde{E}_F(n) = \sum_{p \in \mathcal{I}} N_p(x).$$

Each integer n counted by $N_p(x)$ satisfies the following condition.

* The integer n may be written as $n = Qm$, where $Q = q_1^{a_1} \cdots q_k^{a_k}$, with $a_i \ge 1$, the primes $q_1, \dots, q_k$ are distinct and satisfy $q_i \in \mathcal{Q}(p,i)$ for all $i \in \{1, \dots, k\}$, and $m \le x/Q$ is not divisible by any prime $q \le y$ with $q \equiv 1 \pmod{p}$.

We partition the integers Q into squarefree and non-squarefree ones. Define the set of products given by

$$\mathcal{Q}_6(p) = \{q_1 q_2 \cdots q_k : q_i \in \mathcal{Q}(p,i), \text{ with } q_1, \dots, q_k \text{ distinct}\},$$

and let $\mathcal{Q}_7(p)$ be the corresponding set of products $q_1^{a_1} \cdots q_k^{a_k}$ for which at least one exponent a_i is greater than 1. Then,

$$\frac{1}{x}\sum_{p \in \mathcal{I}} N_p(x) = \frac{1}{x} \sum_{\substack{p \in \mathcal{I} \\ Q \in \mathcal{Q}_6(p)}} \#\left\{m \le \frac{x}{Q} : \gcd\left(m, \prod_{\substack{q' \le y \\ q' \equiv 1 \pmod{p}}} q'\right) = 1\right\} + \frac{1}{x} \sum_{\substack{p \in \mathcal{I} \\ Q \in \mathcal{Q}_7(p)}} \#\left\{m \le \frac{x}{Q} : \gcd\left(m, \prod_{\substack{q' \le y \\ q' \equiv 1 \pmod{p}}} q'\right) = 1\right\}. \tag{7}$$

The contribution from the second summand on the right-hand side of equation (7) is negligible. Indeed,

$$\frac{1}{x} \sum_{\substack{p \in \mathcal{I} \\ Q \in \mathcal{Q}_7(p)}} \#\left\{m \le \frac{x}{Q} : \gcd\left(m, \prod_{\substack{q' \le y \\ q' \equiv 1 \pmod{p}}} q'\right) = 1\right\} \ll \sum_{p \in \mathcal{I}} \sum_{Q \in \mathcal{Q}_7(p)} \frac{1}{Q} \ll \frac{(\log_3 x)^{r+1}}{(\log_2 x)^{r-k}}. \tag{8}$$

To see this, suppose, for example, that $a_1 \ge 2$. Since $\displaystyle\sum_{i=2}^{\infty} \frac{1}{q_j^i} \ll \frac{1}{q_j^2}$ and $\displaystyle\sum_{i=1}^{\infty} \frac{1}{q_j^i} \ll \frac{1}{q_j}$ for all primes q_j, it follows, taking into account that $\frac{1}{q_1} < \frac{1}{p}$ and applying Lemma 2.3 that

$$\sum_{\substack{q_1 \le y \\ q_1 \equiv 1 \pmod{p^{e_1}}}} \sum_{a_1=2}^{\infty} \frac{1}{q_1^{a_1}} \ll \sum_{\substack{q_1 \le y \\ q_1 \equiv 1 \pmod{p^{e_1}}}} \frac{1}{q_1^2} \ll \sum_{\substack{q_1 \le y \\ q_1 \equiv 1 \pmod{p^{e_1}}}} \frac{1}{pq_1} \ll \frac{\log_2 x}{p^{e_1+1}},$$

whereas, for $j \ge 2$,

$$\sum_{\substack{q_j \le y \\ q_j \equiv 1 \pmod{p^{e_j}}}} \sum_{a_j=1}^{\infty} \frac{1}{q_j^{a_j}} \ll \sum_{\substack{q_j \le y \\ q_j \equiv 1 \pmod{p^{e_j}}}} \frac{1}{q_j} \ll \frac{\log_2 x}{p^{e_j}}.$$

Then,

$$\sum_{Q\in\mathcal{Q}_7(p)}\frac{1}{Q}\ll \sum_{\substack{q_1\le y\\ q_1\equiv 1 \pmod{p^{e_1}}}}\left(\sum_{i=2}^{\infty}\frac{1}{q_1^i}\right)\sum_{\substack{q_2\le y\\ q_2\equiv 1 \pmod{p^{e_2}}}}\left(\sum_{i=1}^{\infty}\frac{1}{q_2^i}\right)\cdots\sum_{\substack{q_k\le y\\ q_k\equiv 1 \pmod{p^{e_k}}}}\left(\sum_{i=1}^{\infty}\frac{1}{q_k^i}\right)$$
$$\ll \frac{(\log_2 x)^k}{p^{r+1}}.$$

The same argument applies when any other exponent a_i exceeds 1. Summing over $p\in\mathcal{I}$ gives (8).

It remains to estimate the squarefree contribution. For $Q\in\mathcal{Q}_6(p)$, we have $Q\le y^k$, and hence $\frac{\log(x/Q)}{\log y}\ge 2\log_3 x-k$. The Fundamental Lemma of Sieve Theory therefore gives

$$\#\left\{m\le\frac{x}{Q}:\gcd\left(m,\prod_{\substack{q'\le y\\ q'\equiv 1 \pmod p}}q'\right)=1\right\}=\frac{x}{Q}\prod_{\substack{q\le y\\ q\equiv 1 \pmod p}}\left(1-\frac{1}{q}\right)\left(1+O\left(\frac{1}{\log_2 x}\right)\right)$$
$$=\frac{x}{Q}\exp\left(-\sum_{\substack{q\le y\\ q\equiv 1 \pmod p}}\frac{1}{q}\right)\left(1+O\left(\frac{1}{\log_2 x}\right)\right).$$

We used above the fact that $\log\left(1-\frac{1}{q}\right)=-\frac{1}{q}+O\left(\frac{1}{q^2}\right)$, and that $\sum_{q\equiv 1 \pmod p}\frac{1}{q^2}<\sum_{q>p}\frac{1}{q^2}\ll\frac{1}{p\log p}\ll\frac{1}{\log_2 x}$. By Lemma 2.3,

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

for $p\in\mathcal{I}$. Then,

(9)
$$\#\left\{m\le\frac{x}{Q}:\gcd\left(m,\prod_{\substack{q'\le y\\ q'\equiv 1 \pmod p}}q'\right)=1\right\}=\frac{x}{Q}\exp\left(-\frac{\log_2 x}{p}\right)\left(1+O\left(\frac{(\log_3 x)^2}{\log_2 x}\right)\right).$$

We next estimate the sum of $1/Q$ over the Q's of interest that are squarefree. We are going to use the factor $C=|A_1|!\cdots|A_{e_1}|!$ with $A_\alpha=\{i\in\{1,\dots,k\}:e_i=\alpha\}$ to account for permutations among the indices i having the same exponent e_i. Then,

$$\sum_{Q\in\mathcal{Q}_6(p)}\frac{1}{Q}=\frac{1}{C}\sum_{q_1\in\mathcal{Q}(p,1)}\left(\cdots\left(\sum_{\substack{q_{k-1}\in\mathcal{Q}(p,k-1)\\ q_{k-1}\notin\{q_1,\dots,q_{k-2}\}}}\left(\sum_{\substack{q_k\in\mathcal{Q}(p,k)\\ q_k\notin\{q_1,\dots,q_{k-1}\}}}\frac{1}{q_1\cdots q_k}\right)\right)\right).$$

For the innermost sum, we have

$$\sum_{\substack{q_k\in\mathcal{Q}(p,k)\\ q_k\notin\{q_1,\dots,q_{k-1}\}}} \frac{1}{q_1\cdots q_k}$$

$$= \frac{1}{q_1\cdots q_{k-1}}\left(\sum_{\substack{q_k\le y\\ q_k\equiv 1 \pmod{p^{e_k}}}}\frac{1}{q_k} - \sum_{\substack{q_k\le y\\ q_k\equiv 1 \pmod{p^{e_k+1}}\\ q_k\notin\{q_1,\dots,q_{k-1}\}}}\frac{1}{q_k} - \sum_{\substack{q_k\le y\\ q_k\equiv 1 \pmod{p^{e_k}}\\ q_k\in\{q_1,\dots,q_{k-1}\}}}\frac{1}{q_k}\right)$$

$$= \frac{1}{q_1\cdots q_{k-1}}\left(\frac{\log_2 x}{p^{e_k}} + O\left(\frac{\log(2p^{e_k})}{\varphi(p^{e_k})} + \frac{\log_2 x}{p^{e_k+1}} + \frac{k-1}{p^{e_k}}\right)\right)$$

$$= \frac{1}{q_1\cdots q_{k-1}}\left(\frac{\log_2 x}{p^{e_k}} + O\left(\frac{(\log_3 x)^{e_k+1}}{(\log_2 x)^{e_k}}\right)\right).$$

Applying the same procedure to the remaining sums, we obtain the following estimate:

$$\sum_{Q\in\mathcal{Q}_6(p)}\frac{1}{Q} = \frac{1}{C}\left(\frac{\log_2 x}{p^{e_1}} + O\left(\frac{(\log_3 x)^{e_1+1}}{(\log_2 x)^{e_1}}\right)\right)\cdots\left(\frac{\log_2 x}{p^{e_k}} + O\left(\frac{(\log_3 x)^{e_k+1}}{(\log_2 x)^{e_k}}\right)\right). \tag{10}$$

Since $e_1 \ge \cdots \ge e_k$, the error term in the last factor is of order greater than or equal to the error terms in the preceding factors. Moreover, in each factor the main term is of higher order than its corresponding error term. To estimate the error arising from the product, for each $j \in \{1,\dots,k\}$, the contribution obtained by taking the error term in the j-th factor and the main terms in all the remaining factors is

$$\ll \frac{(\log_3 x)^{e_j+1}}{(\log_2 x)^{e_j}}\frac{(\log_2 x)^{k-1}}{p^{r-e_j}}.$$

Since $p \gg \frac{\log_2 x}{\log_3 x}$, each such contribution is

$$\ll \frac{(\log_3 x)^{r+1}}{(\log_2 x)^{r+1-k}}.$$

The terms involving two or more error factors are of smaller order. Therefore,

$$\sum_{Q\in\mathcal{Q}_6(p)}\frac{1}{Q} = \frac{1}{C}\frac{(\log_2 x)^k}{p^r} + O\left(\frac{(\log_3 x)^{r+1}}{(\log_2 x)^{r+1-k}}\right). \tag{11}$$

Combining (9)–(11), we find that

$$\frac{1}{x}\sum_{n\le x}\tilde{E}_F(n) = \frac{1}{C}\sum_{p\in\mathcal{I}}\frac{(\log_2 x)^k}{p^r}\exp\left(-\frac{\log_2 x}{p}\right) + O\left(\frac{(\log_3 x)^{r+2}}{(\log_2 x)^{r-k}}\right). \tag{12}$$

By Lemma 4.3 stated below, we have

$$\sum_{p\in\mathcal{I}}\frac{(\log_2 x)^k}{p^r}\exp\left(-\frac{\log_2 x}{p}\right) = (1+o(1))\frac{(r-2)!}{\log_3 x\,(\log_2 x)^{r-k-1}}.$$

Substituting this expression into (12) completes the proof of Lemma 4.1. □

LEMMA 4.3. *For any fixed integer $s \ge 2$ and any fixed $c > 0$,*

$$\sum_{p\in\mathcal{I}}\frac{(\log_2 x)^c}{p^s}\exp\left(-\frac{\log_2 x}{p}\right) = (1+o(1))\frac{(s-2)!}{\log_3 x\,(\log_2 x)^{s-c-1}}.$$

PROOF OF LEMMA 4.3. Write $\pi(t) = \int_2^t \frac{dz}{\log z} + E(t)$, where $E(t) \ll \frac{t}{(\log t)^A}$ for any A and all $t \geq 2$. Summing by parts, we have

$$\sum_{p\in\mathcal{I}} \frac{(\log_2 x)^c}{p^s} \exp\left(-\frac{\log_2 x}{p}\right) = \int_{z_1}^{z_2} \frac{(\log_2 x)^c}{t^s} \exp\left(-\frac{\log_2 x}{t}\right) \frac{1}{\log t} dt + \int_{z_1}^{z_2} \frac{(\log_2 x)^c}{t^s} \exp\left(-\frac{\log_2 x}{t}\right) dE(t). \tag{13}$$

For the second integral, we have

$$\int_{z_1}^{z_2} \frac{(\log_2 x)^c}{t^s} \exp\left(-\frac{\log_2 x}{t}\right) dE(t) = E(t)\frac{(\log_2 x)^c}{t^s} \exp\left(-\frac{\log_2 x}{t}\right)\Big|_{z_1}^{z_2} - \int_{z_1}^{z_2} E(t) \left(\frac{d}{dt}\left(\frac{(\log_2 x)^c}{t^s} \exp\left(-\frac{\log_2 x}{t}\right)\right)\right) dt,$$

which is bounded by

$$\sup_{t\in\mathcal{I}} |E(t)| \left(|f(z_1)| + |f(z_2)| + \int_{z_1}^{z_2} |f'(t)|\, dt\right),$$

where

$$f(t) = \frac{(\log_2 x)^c}{t^s} \exp\left(-\frac{\log_2 x}{t}\right).$$

Since

$$f'(t) = f(t)\frac{\log_2 x - st}{t^2},$$

the function f is increasing up to $t = \log_2 x/s$ and decreasing thereafter. Hence

$$\int_{z_1}^{z_2} |f'(t)|\, dt \ll \max_{z_1\leq t\leq z_2} f(t) \ll (\log_2 x)^{c-s}.$$

Therefore,

$$\int_{z_1}^{z_2} \frac{(\log_2 x)^c}{t^s} \exp\left(-\frac{\log_2 x}{t}\right) dE(t) \ll \frac{1}{(\log_2 x)^{s-c-1}(\log_3 x)^A}.$$

The first integral on the right-hand side of equation (13) is equal to

$$\frac{1+o(1)}{\log_3 x} \int_{z_1}^{z_2} \frac{(\log_2 x)^c}{t^s} \exp\left(-\frac{\log_2 x}{t}\right) dt.$$

We recognize the integrand in the latter integral as the derivative of

$$\frac{(s-2)!}{(\log_2 x)^{s-c-1}} \left(\sum_{i=0}^{s-2} \frac{1}{i!} \left(\frac{\log_2 x}{t}\right)^i\right) \exp\left(-\frac{\log_2 x}{t}\right).$$

The contribution at the upper endpoint is

$$\frac{(s-2)!}{(\log_2 x)^{s-c-1}} (1+o(1)),$$

whereas the contribution at the lower endpoint is negligible. Consequently,

$$\sum_{p\in\mathcal{I}} \frac{(\log_2 x)^c}{p^s} \exp\left(-\frac{\log_2 x}{p}\right) = (1+o(1)) \frac{(s-2)!}{\log_3 x\, (\log_2 x)^{s-c-1}}. \tag{14}$$

□

PROOF OF LEMMA 4.2. For $p \in \mathcal{I}$, let $I(p)$ denote the condition that there exist k distinct primes $\{q_1, \dots, q_k\}$ dividing n with $q_i \in \mathcal{Q}(p,i)$ for all $i \in \{1,\dots,k\}$ and there are no other primes $q \le y$ congruent to 1 (mod p) dividing n besides these $q_1, \dots, q_k$. To ease notation, we denote

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

Therefore,

$$\sum_{n \le x} (\tilde{E}_F(n))^2 = \sum_{p_1,p_2 \in \mathcal{I}} N_{p_1,p_2}(x). \tag{15}$$

We split the sum in (15) into its diagonal and off-diagonal parts. For the diagonal terms, where $p_1 = p_2 = p$, we have

$$N_{p,p}(x) = N_p(x).$$

Hence, by Lemma 4.1,

$$\frac{1}{x} \sum_{p \in \mathcal{I}} N_{p,p}(x) = (1 + o(1)) \, \frac{(r-2)!}{C} \frac{1}{\log_3 x \, (\log_2 x)^{r-k-1}}. \tag{16}$$

We now consider the off-diagonal terms. We will show that

$$\sum_{\substack{p_1,p_2 \in \mathcal{I} \\ p_1 \neq p_2}} N_{p_1,p_2}(x) = o\left(\frac{x}{\log_3 x \, (\log_2 x)^{r-k-1}} \right). \tag{17}$$

Fix distinct primes $p_1, p_2 \in \mathcal{I}$. We distinguish two types of integers counted by $N_{p_1,p_2}(x)$. We say that n is of the first kind (with respect to p_1 and p_2) if it satisfies the following:

* The integer n has a factorization $n = QLm$, where $Q = q_1^{a_1} \cdots q_k^{a_k}$, and $L = \ell_1^{b_1} \cdots \ell_k^{b_k}$, and these satisfy the congruence patterns

  $$q_i \equiv 1 \pmod{p_1^{e_i}}, \qquad q_i \not\equiv 1 \pmod{p_1^{e_i+1}}, \qquad q_i \not\equiv 1 \pmod{p_2},$$

  and

  $$\ell_i \equiv 1 \pmod{p_2^{e_i}}, \qquad \ell_i \not\equiv 1 \pmod{p_2^{e_i+1}}, \qquad \ell_i \not\equiv 1 \pmod{p_1}.$$

  In addition, m is free of prime factors $q \le y$ satisfying
  $q \equiv 1 \pmod{p_1}$ or $q \equiv 1 \pmod{p_2}$.

All remaining integers counted by $N_{p_1,p_2}(x)$ are said to be of the second kind. Every such integer is divisible by a prime $q \le y$ satisfying $q \equiv 1 \pmod{p_1 p_2}$. We first estimate the integers of the first kind. As in the proof of Lemma 4.1, the contribution from terms in which at least one prime factor of Q or L appears with exponent greater than 1 is $o\left(\frac{x}{\log_3 x (\log_2 x)^{r-k-1}}\right)$. We may therefore restrict our attention to the squarefree products $Q = q_1 \cdots q_k, \qquad L = \ell_1 \cdots \ell_k$. Define

$$\mathcal{Q}_8(p_1,p_2) = \left\{ q_1 \cdots q_k : \begin{array}{l} q_1, \dots, q_k \text{ are distinct}, q_i \in \mathcal{Q}(p_1,i) \\ q_i \not\equiv 1 \pmod{p_2} \text{ for every } i \end{array} \right\},$$

and

$$\tilde{M}_{p_1,p_2}(QL) = \# \left\{ m \le \frac{x}{QL} : \begin{array}{l} m \text{ is free of prime factors } q \le y \text{ with} \\ q \equiv 1 \pmod{p_1} \text{ or } q \equiv 1 \pmod{p_2} \end{array} \right\}.$$

Thus, the contribution of the integers of the first kind to $\sum_{p_1,p_2 \in \mathcal{I}} N_{p_1,p_2}(x)$ is

$$\ll \sum_{p_1,p_2 \in \mathcal{I}} \sum_{\substack{Q \in \mathcal{Q}_8(p_1,p_2) \\ L \in \mathcal{Q}_8(p_2,p_1)}} \tilde{M}_{p_1,p_2}(QL).$$

Since $QL \le y^{2k}$, the sieve estimate applies. We obtain that the quantity

$$\tilde{M}_{p_1,p_2}(QL) = \frac{x}{QL} \prod_{\substack{q\le y \\ q\equiv 1 \pmod{p_1} \\ \text{or } q\equiv 1 \pmod{p_2}}} \left(1-\frac{1}{q}\right)\left(1+O\left(\frac{1}{\log_2 x}\right)\right)$$

$$= \frac{x}{QL}\exp\left(-\sum_{\substack{q\le y \\ q\equiv 1 \pmod{p_1} \\ \text{or } q\equiv 1 \pmod{p_2}}} \frac{1}{q}\right)\left(1+O\left(\frac{1}{\log_2 x}\right)\right).$$

By Lemma 2.3,

$$\sum_{\substack{q\le y \\ q\equiv 1 \pmod{p_1} \\ \text{or } q\equiv 1 \pmod{p_2}}} \frac{1}{q} = \frac{\log_2 x}{p_1-1} + \frac{\log_2 x}{p_2-1} - \frac{\log_2 x}{(p_1-1)(p_2-1)} + O\left(\frac{\log_3 x}{p_1}+\frac{\log_3 x}{p_2}\right)$$

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

Then,

$$\tilde{M}_{p_1,p_2}(QL) = \frac{x}{QL}\exp\left(-\frac{\log_2 x}{p_1}\right)\exp\left(-\frac{\log_2 x}{p_2}\right)\left(1+O\left(\frac{(\log_3 x)^2}{\log_2 x}\right)\right).$$

Using Lemma 2.3,

$$\sum_{\substack{Q\in\mathcal{Q}_8(p_1,p_2) \\ L\in\mathcal{Q}_8(p_2,p_1)}} \frac{1}{QL}$$

$$\ll \left(\sum_{\substack{q_1\le y \\ q_1\equiv 1 \pmod{p_1^{e_1}}}} \frac{1}{q_1}\cdots \sum_{\substack{q_k\le y \\ q_k\equiv 1 \pmod{p_1^{e_k}}}} \frac{1}{q_k}\right)\left(\sum_{\substack{\ell_1\le y \\ \ell_1\equiv 1 \pmod{p_2^{e_1}}}} \frac{1}{\ell_1}\cdots \sum_{\substack{\ell_k\le y \\ \ell_k\equiv 1 \pmod{p_2^{e_k}}}} \frac{1}{\ell_k}\right)$$

$$\ll \frac{(\log_2(x))^k}{p_1^r}\frac{(\log_2(x))^k}{p_2^r}.$$

Summing over the primes $p_1, p_2 \in \mathcal{I}$ and using Lemma 4.3, we find that the total contribution from integers of the first kind is

$$\ll x\left(\sum_{p\in\mathcal{I}} \frac{(\log_2 x)^k}{p^r}\exp\left(-\frac{\log_2 x}{p}\right)\right)^2 \ll \frac{x}{(\log_3 x)^2\,(\log_2 x)^{2(r-k-1)}}. \tag{18}$$

Since $r > k$, the last term in (18) is

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

It remains to estimate the integers of the second kind. If n is of the second kind with respect to p_1 and p_2, then n is divisible by at least one prime $q \le y$ satisfying $q \equiv 1 \pmod{p_1p_2}$. On the other hand, n has exactly k distinct prime divisors $q_1, \dots, q_k \le y$ that are congruent to 1 modulo p_1 and exactly k distinct prime divisors $\ell_1, \dots, \ell_k \le y$ that are congruent to 1 modulo p_2. Consequently, any prime $q \equiv 1 \pmod{p_1p_2}$ dividing n must coincide with one of the q_i and, at the same time, with one of the ℓ_j; that is, $q = q_i = \ell_j$ for some $i, j \in \{1, \dots, k\}$.

Furthermore, by the definition of integers of the second kind, $q_i \equiv 1 \pmod{p_1^{e_i}}$ and $\ell_j \equiv 1 \pmod{p_2^{e_j}}$. Hence, $q \equiv 1 \pmod{p_1^{e_i}}$ and $q \equiv 1 \pmod{p_2^{e_j}}$ for some $i, j \in \{1, \dots, k\}$. The prime q need not be unique; there may be several primes with these properties. This motivates the introduction of two index sets B_1 and B_2. The set B_1 records the indices i of the primes q_i that are also congruent to 1 (mod p_2), whereas B_2 records the corresponding indices j in the family $\ell_1, \dots, \ell_k$. These sets allow us to describe the general form of an integer n of the second kind with respect to p_1 and p_2. Let B_1, B_2 be nonempty subsets of $\{1, \dots, k\}$ having the same cardinality, say $|B_1| = |B_2| = b \geq 1$. Consider a bijection f from B_1 to B_2. We denote by $N^f_{p_1,p_2}(x)$ the set of second-kind integers admitting a decomposition of the form $n = mQL$, where the following conditions hold:

1. The integer m is free of prime factors $q \leq y$ such that $q \equiv 1 \pmod{p_1}$ or $q \equiv 1 \pmod{p_2}$.
2. The integer Q is a product of powers of the k distinct primes $q_1, \dots, q_k \leq y$, satisfying $q_i \equiv 1 \pmod{p_1^{e_i}}$ and $q_i \not\equiv 1 \pmod{p_1^{e_i+1}}$ for all $i \in \{1, \dots, k\}$. Additionally, $q_i \equiv 1 \pmod{p_2^{e_{f(i)}}}$ and $q_i \not\equiv 1 \pmod{p_2^{e_{f(i)}+1}}$ for all $i \in B_1$, and $q_i \not\equiv 1 \pmod{p_2}$ for all $i \in \{1, \dots, k\} \setminus B_1$.
3. The integer L is a product of powers of the $k - b$ distinct primes $\{\ell_i\}_{i \in \{1,\dots,k\} \setminus B_2}$, where $\ell_i \leq y$, $\ell_i \equiv 1 \pmod{p_2^{e_i}}$ and $\ell_i \not\equiv 1 \pmod{p_2^{e_i+1}}$ for all $i \in \{1, \dots, k\} \setminus B_2$. Furthermore, $\ell_i \not\equiv 1 \pmod{p_1}$ for all $i \in \{1, \dots, k\} \setminus B_2$. When $b = k$, we use the convention $L = 1$.

Thus, the shared primes are included only in Q, whereas L contains the $k - b$ remaining primes associated with p_2. Consequently, $\gcd(Q, L) = 1$. Each shared prime q_i with $i \in B_1$ introduces an additional congruence condition involving both p_1 and p_2. For each $i \in B_1$, set $g_i = e_{f(i)}$. We define $\mathcal{Q}_f(p_1, p_2)$ to be the set of integers Q satisfying condition (2), and $\mathcal{L}_f(p_1, p_2)$ to be the set of integers L satisfying condition (3). Then,

$$\sum_{\substack{p_1,p_2 \in \mathcal{I} \\ p_1 \neq p_2}} \#N^f_{p_1,p_2}(x) \leq \sum_{\substack{p_1,p_2 \in \mathcal{I} \\ p_1 \neq p_2 \\ Q \in \mathcal{Q}_f(p_1,p_2) \\ L \in \mathcal{L}_f(p_1,p_2)}} \# \left\{ m \leq \frac{x}{QL} : \gcd\left(m, \prod_{\substack{q \leq y \\ q \equiv 1 \pmod{p_1} \text{ or} \\ q \equiv 1 \pmod{p_2}}} q \right) = 1 \right\}$$

$$\ll \sum_{\substack{p_1,p_2 \in \mathcal{I} \\ p_1 \neq p_2 \\ Q \in \mathcal{Q}_f(p_1,p_2) \\ L \in \mathcal{L}_f(p_1,p_2)}} \frac{x}{QL}.$$

A similar argument to that used in (3) yields the bound

$$\sum_{\substack{Q\in\mathcal{Q}_f(p_1,p_2)\\ L\in\mathcal{L}_f(p_1,p_2)}} \frac{1}{QL}$$

$$\ll \prod_{i\in B_1}\left(\sum_{\substack{q_i\le y\\ q_i\equiv 1 \pmod{p_1^{e_i}p_2^{g_i}}}}\frac{1}{q_i}\right)\prod_{i\in\{1,\dots,k\}\setminus B_1}\left(\sum_{\substack{q_i\le y\\ q_i\equiv 1 \pmod{p_1^{e_i}}}}\frac{1}{q_i}\right)\prod_{i\in\{1,\dots,k\}\setminus B_2}\left(\sum_{\substack{\ell_i\le y\\ \ell_i\equiv 1 \pmod{p_2^{e_i}}}}\frac{1}{\ell_i}\right)$$

$$\ll \prod_{i\in B_1}\left(\frac{\log_2 x}{p_1^{e_i}p_2^{g_i}}\right)\prod_{i\in\{1,\dots,k\}\setminus B_1}\left(\frac{\log_2 x}{p_1^{e_i}}\right)\prod_{i\in\{1,\dots,k\}\setminus B_2}\left(\frac{\log_2 x}{p_2^{e_i}}\right)$$

$$\ll \frac{(\log_2 x)^{2k-b}}{p_1^r p_2^r}.$$

Since $p_1, p_2 \gg \frac{\log_2 x}{\log_3 x}$, summing over $p_1, p_2 \in \mathcal{I}$ yields

$$\sum_{\substack{p_1,p_2\in\mathcal{I}\\ p_1\neq p_2}} \#N^f_{p_1,p_2}(x) \ll \frac{x(\log_3 x)^{2r}}{(\log_2 x)^{2r-2k-2+b}}.$$

Since $r > k$ and $b \ge 1$, it follows that

$$\sum_{\substack{p_1,p_2\in\mathcal{I}\\ p_1\neq p_2}} \#N^f_{p_1,p_2}(x) = o\left(\frac{x}{\log_3 x\,(\log_2 x)^{r-k-1}}\right). \tag{19}$$

There are only finitely many pairs (B_1, B_2) and, for each such pair, finitely many bijections $f : B_1 \to B_2$. Moreover, every integer n of the second kind belongs to a set $N^f_{p_1,p_2}(x)$ for some bijection $f : B_1 \to B_2$ and some distinct primes p_1, p_2. Hence, the total contribution from integers of the second kind is also

$$o\left(\frac{x}{\log_3 x\,(\log_2 x)^{r-k-1}}\right). \tag{20}$$

Combining (18) and (20) proves (17). Together with the diagonal estimate (16), this establishes Lemma 4.2. □

Consequently, taking into account the discussion at the beginning of the proof, this completes the proof of Theorem 1.1.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

The mathematical arguments in the paper are due to the authors. The manuscript was edited for style and presentation and "pre-refereed" with the help of ChatGPT 5.5 and 5.6.

Acknowledgements

Research on this paper was conducted with the following support:

(i) The first and second named authors (S.M. and G.P.) were supported by FONDOCyT, grant number 2022-1D1-085.

(ii) The second named author (G.P.) was supported by NSF award DMS-2316986.

Data Availability

No datasets were generated or analyzed during the current study.

References

1. B. Chang and G. Martin, *The smallest invariant factor of the multiplicative group*, Int. J. Number Theory **16** (2020), 1377–1405.
2. J. Downey and G. Martin, *Counting multiplicative groups with prescribed subgroups*, Int. J. Number Theory **17** (2021), 2087–2112.
3. P. Erdős, C. Pomerance, and E. Schmutz, *Carmichael's lambda function*, Acta Arith. **58** (1991), 363–385.
4. H. Halberstam and H.-E. Richert, *Sieve methods*, London Mathematical Society Monographs, no. 4, Academic Press, 1974.
5. G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 6 ed., Oxford University Press, Oxford, 2008.
6. G. Martin and R. M. Simpson, *The universal profile of the invariant factors of* $(\mathbb{Z}/n\mathbb{Z})^\times$, Int. J. Number Theory **22** (2026), 1441–1523.
7. 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.
8. S. Morales, G. Polanco, and P. Pollack, *Elementary abelian Sylow subgroups of the multiplicative group*, J. Number Theory **281** (2026), 205–223.
9. M. R. Murty and V. K. Murty, *Prime divisors of Fourier coefficients of modular forms*, Duke Math. J. **51** (1984), 57–76.
10. K. K. Norton, *On the number of restricted prime factors of an integer. I*, Illinois J. Math. **20** (1976), 681–705.
11. P. Pollack, *The number of non-cyclic Sylow subgroups of the multiplicative group modulo n*, Canad. Math. Bull. **64** (2021), 204–215.
12. P. Pollack and C. Pomerance, *Phi, Primorials, and Poisson*, Illinois J. Math. **64** (2020), 319–330.
13. C. Pomerance, *On the distribution of amicable numbers*, J. Reine Angew. Math. **293/294** (1977), 217–222.
14. G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, 3 ed., Graduate Studies in Mathematics, vol. 163, American Mathematical Society, Providence, RI, 2015.

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