RESIDUE CLASSES CONTAINING AN UNEXPECTED NUMBER OF PRIMES

DANIEL FIORILLI

Abstract. We fix a non-zero integer a and consider arithmetic progressions $a \bmod q$, with q varying over a given range. We show that for certain specific values of a, the arithmetic progressions $a \bmod q$ contain, on average, significantly fewer primes than expected.

1. Introduction

The prime number theorem for arithmetic progressions asserts that

$$\psi(x;q,a) \sim \psi(x)/\phi(q)$$

for any a and q such that $(a,q)=1$. Another way to say this is that the primes are equidistributed in the $\phi(q)$ arithmetic progressions $a \bmod q$ with $(a,q)=1$.

Fix an integer $a \neq 0$. We will be interested in the number of primes in the arithmetic progressions $a \bmod q$ with q varying in certain ranges, and we will show that for specific values of a, there are significantly fewer primes in these arithmetic progressions than in typical arithmetic progressions. Consider the average value of $\psi(x;q,a) - \psi(x)/\phi(q)$ over q. One might expect that no matter what the value of a is, the cancellations in these oscillating terms will force the average to be very small. However it turns out that the average is highly dependent on the arithmetical properties of a.

Here is the main result of the paper.

Theorem 1.1. *Fix an integer $a \neq 0$ and let $M = M(x) \leq \log^B x$ where $B > 0$ is a fixed real number. The average error term in the usual approximation for the number of primes $p \equiv a \bmod q$ with $p \leq x$, where $(q,a)=1$ and $q \leq x/M$, is*

$$\frac{1}{\frac{\phi(a)}{a}\frac{x}{M}} \sum_{\substack{q \leq \frac{x}{M} \\ (q,a)=1}} \left(\psi(x;q,a) - \frac{\psi(x)}{\phi(q)} - \Lambda(a) \right) = \mu(a,M) + O_{a,\epsilon,B}\left(\frac{1}{M^{\frac{205}{538}-\epsilon}} \right) \tag{1}$$

with

$$\mu(a,M) := \begin{cases} 0 & \text{if } \omega(a) \geq 2 \\ -\frac{1}{2}\log p & \text{if } a = \pm p^e \\ -\frac{1}{2}\log M - C_5 & \text{if } a = \pm 1 \end{cases}$$

where

$$C_5 := \frac{1}{2}\left(\log 2\pi + \gamma + \sum_p \frac{\log p}{p(p-1)} + 1 \right).$$

Remark 1.2. Assuming Lindelöf's hypothesis, we can replace the error term in (1) by $O_{a,\epsilon,B}\left(\frac{1}{M^{1/2-\epsilon}}\right)$.

Remark 1.3. We substracted $\Lambda(a)$ from $\psi(x;q,a)$ in (1) because the arithmetic progression $a \bmod q$ contains the prime power p^e for all q if $a = p^e$.

Remark 1.4. It may be preferable to replace $\psi(x)$ by $\psi(x,\chi_0)$ in Theorem 1.1, since the quantity

$$\psi(x;q,a) - \psi(x,\chi_0)/\phi(q)$$

is the discrepancy (with signs) of the sequence of primes in the residue classes mod q. One can do this with a negligible error term.

2. Past Results

The study of the discrepancy $\psi(x;q,a) - x/\phi(q)$ on average has been a fruitful subject over the past decades. For example, the celebrated theorem of Bombieri-Vinogradov gives a bound on the sum of the mean absolute value of the maximum of this discrepancy over all $1 \leq a < q$ with $(a,q) = 1$, summed over $q \leq x^{1/2-o(1)}$. The Hooley-Montgomery refinement of the Barban-Davenport-Halberstam Theorem gives an estimation of the variance of $\psi(x;q,a) - x/\phi(q)$, again for all values of a in the range $1 \leq a < q$ with $(a,q) = 1$ for $q < x/\log^A x$. The mean value of $\psi(x;q,a) - x/\phi(q)$ was studied for fixed values of a for $q \geq x^{1/2}$ (see [1],[3] or [4]), and bounds on this mean value turned out to be applicable to Titchmarsh's divisor problem, first solved by Linnik. The best result so far for this problem was obtained by Friedlander and Granville.

Theorem 2.1 (Friedlander, Granville)**.** *Let $0 < \lambda < 1/4$, $A > 0$ be given. Then uniformly for $0 < |a| < x^\lambda$, $2 \leq Q \leq x/3$ we have*

$$\sum_{\substack{Q<q\leq 2Q\\(q,a)=1}} \left(\psi(x;q,a) - \frac{x}{\phi(q)}\right) \ll_{\lambda,A} 2^{\omega(a)} Q\log(x/Q) + \frac{x}{\log^A x} + Q\log|a|. \tag{2}$$

Remark 2.2. If a is not a prime power the term $Q\log|a|$ may be deleted.

(Note that the $\log\log(x/Q)$ in the original paper is a misprint which was corrected to $\log(x/Q)$ in Proposition 2.1 of [5].) The method used in the present paper is a refinement of the method used to prove Proposition 2.1 in [5], which is essentially the Hooley-Montgomery "switching divisor" technique (see [7]).

3. Main results

A number of constants will appear throughout the paper. We will denote by γ the Euler-Mascheroni constant.

Definition 3.1. *We define*

$$C_1(a) := \frac{\zeta(2)\zeta(3)}{\zeta(6)}\frac{\phi(a)}{a}\prod_{p|a}\left(1 - \frac{1}{p^2-p+1}\right),$$

$$C_3(a) := C_1(a)\left(\gamma - 1 - \sum_p \frac{\log p}{p^2-p+1} + \sum_{p|a}\frac{p^2\log p}{(p-1)(p^2-p+1)}\right),$$

$$C_5 := \frac{1}{2}\left(\log 2\pi + \gamma + \sum_p \frac{\log p}{p(p-1)} + 1\right).$$

From now on, $\lambda < 1/4$ (from Theorem 2.1) is going to be a fixed, positive real number and every constant can possibly depend on λ. We will denote by $\omega(a)$ the number of prime factors of a. Note that there are $\sim \frac{\phi(a)}{a}\frac{x}{M}$ terms in the sum over $1 \le q \le x/M$ with the condition $(q,a) = 1$.

We will see later on that Theorem 1.1 can be made uniform for a in some range, which is the object of the more technical Theorem 7.2. A consequence of this is the following.

Theorem 3.2. *Let $M := \log^B x$ where $B > 4\log 4$ is a fixed real number. The proportion of integers a in the range $0 < |a| \le x^\lambda$ for which*

$$\frac{1}{\frac{\phi(a)}{a}\frac{x}{M}} \sum_{\substack{q \le \frac{x}{M} \\ (q,a)=1}} \left(\psi(x;q,a) - \frac{\psi(x)}{\phi(q)}\right) = O_B\left(\frac{1}{M^{1/4 - \frac{\log 4}{B}}}\right) \tag{3}$$

is at least $\frac{4}{15}e^{-\gamma}$.

Remark 3.3. One can show more generally that for any fixed $\epsilon > 0$, there exists $\delta > 0$ such that for $B > B_\epsilon$, the proportion of integers a in the range $0 < |a| \le x^\lambda$ for which we get an error term of $O_B\left(\frac{1}{M^\delta}\right)$ in (3) is at least $e^{-\gamma} - \epsilon$.

One can consider a different range for q.

Proposition 3.4. *Fix $A > 0$. For a in the range $0 < |a| \le x^\lambda$ such that $\omega(a) \le 10\log\log x$ we have*

$$\frac{1}{\frac{\phi(a)}{a}x} \sum_{\substack{q \le x \\ (q,a)=1}} \left(\psi(x;q,a) - \frac{\psi(x)}{\phi(q)} - \Lambda(a)\right) = \frac{a}{\phi(a)}C_3(a) + O_A\left(\frac{1}{\log^A x}\right). \tag{4}$$

More generally, for $M \ge 1$ a fixed integer we have

$$\frac{1}{\frac{\phi(a)}{a}\frac{x}{M}} \sum_{\substack{q \le \frac{x}{M} \\ (q,a)=1}} \left(\psi(x;q,a) - \frac{\psi(x)}{\phi(q)} - \Lambda(a)\right) = \mu'(a,M) + O_A\left(\frac{1}{\log^A x}\right) \tag{5}$$

where

$$\mu'(a,M) := \frac{a}{\phi(a)}M\left(C_1(a)\log M + C_3(a) - \sum_{\substack{r \le M \\ (r,a)=1}} \frac{1}{\phi(r)}\left(1 - \frac{r}{M}\right)\right).$$

(Note that $\mu'(a,1) = \frac{a}{\phi(a)}C_3(a)$.)

A corollary of Proposition 3.4 (which can be deduced from Theorem 1 of [4]) is Titchmarsh's divisor problem.

Corollary 3.5 (Titchmarsh's divisor problem)**.** *Fix $A > 0$. For a in the range $0 < |a| \le x^\lambda$ such that $\omega(a) \le 10\log\log x$ we have*

$$\sum_{|a| < n \le x} \Lambda(n)\tau(n-a) = C_1(a)x\log x + (2C_3(a) + C_1(a))x + O_A\left(\frac{x}{\log^A x}\right). \tag{6}$$

Note that the constant 10 in Proposition 3.4 and Corollary 3.5 can be replaced by an arbitrary large real number.

4. Acknowledgements

I would like thank my supervisor Andrew Granville for his advice, as well as my colleagues Farzad Aryan, Mohammad Bardestani, Tristan Freiberg and Kevin Henriot for many fruitful conversations. L'auteur est titulaire d'une bourse doctorale du Conseil de recherches en sciences naturelles et en génie du Canada.

5. Notation

Definition 5.1. *For $n \neq 0$ an integer (possibly negative), we define*

$$\Lambda(n) := \begin{cases} \log p & \text{if } n = p^e \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

$$\vartheta(n) := \begin{cases} \log p & \text{if } n = p \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

Definition 5.2.

$$\psi(x; q, a) := \sum_{\substack{n \leq x \\ n \equiv a \bmod q}} \Lambda(n) \tag{9}$$

$$\theta(x; q, a) := \sum_{\substack{n \leq x \\ n \equiv a \bmod q}} \vartheta(n) \tag{10}$$

The following definition is non-standard but will be useful in the proofs.

Definition 5.3.

$$\psi^*(x; q, a) := \sum_{\substack{|a| < n \leq x \\ n \equiv a \bmod q}} \Lambda(n) \tag{11}$$

$$\theta^*(x; q, a) := \sum_{\substack{|a| < n \leq x \\ n \equiv a \bmod q}} \vartheta(n) \tag{12}$$

We will need to consider the prime divisors of a which are less than M.

Definition 5.4. *For a an integer and $M > 0$ a real number, we define*

$$a_M := \prod_{\substack{p \mid a \\ p \leq M}} p. \tag{13}$$

The error term $E(M, a)$ will be defined depending on the context, so one has to pay attention to its definition in every statement.

6. Lemmas

We begin by the recalling the Hooley-Montgomery technique.

Lemma 6.1 (Switching divisors). *Let a be an integer such that $0 < |a| \leq x^{1/4}$ and let $M = M(x)$ such that $1 \leq M < x$. We have*

$$\sum_{\substack{\frac{x}{M} < q \leq x \\ (q,a)=1}} \sum_{\substack{|a| < p \leq x \\ p \equiv a \bmod q}} \log p = \sum_{\substack{1 \leq r < (x-a)\frac{M}{x} \\ (r,a)=1}} \sum_{\substack{r\frac{x}{M} + a < p \leq x \\ p \equiv a \bmod r \\ p > |a|}} \log p + O(|a| \log x). \tag{14}$$

Proof. Clearly,

$$\sum_{\substack{\frac{x}{M} < q \leq x \\ (q,a)=1}} \sum_{\substack{|a| < p \leq x \\ p \equiv a \bmod q}} \log p = \sum_{\substack{\frac{x}{M} < q \leq x \\ (q,a)=1}} \sum_{\substack{|a| < p \leq x \\ p \equiv a \bmod q \\ p > a + \frac{x}{M}}} \log p. \tag{15}$$

Now we will apply Hooley's technique which is somewhat similar to Dirichlet's hyperbola method. Setting $p - a = rq$ in (15), one can sum over r instead of summing over q. Now $(r, a) = 1$, else $(p, a) > 1$ so $p \mid a$, but this is impossible since $p > |a|$. Taking $a > 0$ for now, we get that (15) is equal to

$$\sum_{\substack{\frac{x}{M} < q \leq x-a \\ (q,a)=1}} \sum_{\substack{|a| < p \leq x \\ p \equiv a \bmod q \\ p > a + \frac{x}{M}}} \log p = \sum_{\substack{1 \leq r < (x-a)\frac{M}{x} \\ (r,a)=1}} \sum_{\substack{r\frac{x}{M} + a < p \leq x \\ p \equiv a \bmod r \\ p > |a|}} \log p. \tag{16}$$

If we had $a < 0$, the minor difference in (16) would be the $r = 1$ term where the additional condition $p \leq x + a$ is needed. Moreover, additional terms would be needed in passing from the right hand side of (15) to the left hand side of (16). Both these modifications can be made at the cost of adding an error term of size $\ll |a| \log x$.

□

Lemma 6.2. *We have the following estimates.*

$$\sum_{\substack{n \leq M \\ (n,a)=1}} \frac{1}{\phi(n)} = C_1(a) \log M + C_1(a) + C_3(a) + O\left(2^{\omega(a)} \frac{\log M}{M}\right) \tag{17}$$

$$\sum_{\substack{n \leq M \\ (n,a)=1}} \frac{n}{\phi(n)} = C_1(a) M + O\left(2^{\omega(a)} \log M\right) \tag{18}$$

Note that without loss of generality, we can replace a by a_M on the right side of (17) and (18).

Proof. The proof of (17) is very similar to the proof of Lemma 13.1 in [5]. One first has to prove the following estimate:

$$\sum_{\substack{n \leq M \\ (n,a)=1}} \frac{1}{n} = \frac{\phi(a)}{a} \left(\log M + \gamma + \sum_{p \mid a} \frac{\log p}{p-1}\right) + O\left(\frac{2^{\omega(a)}}{M}\right). \tag{19}$$

One then writes

$$\sum_{\substack{n\le M\\(n,a)=1}}\frac{1}{\phi(n)}=\sum_{\substack{n\le M\\(n,a)=1}}\frac{1}{n}\sum_{d|n}\frac{\mu^2(d)}{\phi(d)}=\sum_{\substack{d\le M\\(d,a)=1}}\frac{\mu^2(d)}{d\phi(d)}\sum_{\substack{r\le M/d\\(r,a)=1}}\frac{1}{r} \tag{20}$$

and then inserts the estimate (19) into (20). The final step is to bound the tail of the sums and to compute the following constants:

$$\sum_{(d,a)=1}\frac{\mu^2(d)}{d\phi(d)}=\frac{\zeta(2)\zeta(3)}{\zeta(6)}\prod_{p|a}\left(1-\frac{1}{p^2-p+1}\right),$$

$$\sum_{(d,a)=1}\frac{\mu^2(d)}{d\phi(d)}\log d=\frac{\zeta(2)\zeta(3)}{\zeta(6)}\prod_{p|a}\left(1-\frac{1}{p^2-p+1}\right)\sum_{p\nmid a}\frac{\log p}{p^2-p+1}.$$

The proof of (18) goes along the same lines. □

Our very delicate analysis forces us to give some details about the "trivial" estimates for the prime counting functions.

Lemma 6.3. *We have for any $\epsilon>0$ that*

$$\sum_{\substack{q\le x\\(q,a)=1}}(\psi^*(x;q,a)-\theta^*(x;q,a))\ll_\epsilon x^{1/2+\epsilon}. \tag{21}$$

Proof.

$$\begin{aligned}&\sum_{\substack{q\le x\\(q,a)=1}}(\psi^*(x;q,a)-\theta^*(x;q,a))\\&\qquad\le\sum_{q\le x}\sum_{\substack{|a|<p^e\le x\\p^e\equiv a \bmod q\\e\ge2}}\log p\le\sum_{2\le e\le\frac{\log x}{\log 2}}\sum_{p\le x^{1/e}}\sum_{\substack{q\le x\\q|p^e-a}}\log p\\&\qquad\le\log x\sum_{2\le e\le\frac{\log x}{\log 2}}\sum_{p\le x^{1/e}}\tau(p^e-a)\ll_\epsilon x^{\epsilon/2}\sum_{2\le e\le\frac{\log x}{\log 2}}\pi(x^{1/e})\\&\qquad\ll_\epsilon x^{1/2+\epsilon}.\end{aligned}$$

□

Lemma 6.4. *Let $a\neq0$ be an integer and $1\le Q\le x$. We have*

$$\sum_{\substack{q\le Q\\(q,a)=1}}(\psi(x;q,a)-\Lambda(a)-\psi^*(x;q,a))=O(|a|\log^2|a|), \tag{22}$$

$$\sum_{\substack{q\le Q\\(q,a)=1}}(\theta(x;q,a)-\vartheta(a)-\theta^*(x;q,a))=O(|a|\log^2|a|). \tag{23}$$

Proof. Note that as soon as $q > 2|a|$, there are no integers congruent to $a \bmod q$ in the interval $[1, |a|)$. We then have

$$\sum_{\substack{q\le Q\\(q,a)=1}} (\psi(x;q,a) - \Lambda(a) - \psi^*(x;q,a)) = \sum_{\substack{q\le Q\\(q,a)=1}} \sum_{\substack{1\le n<|a|\\ n\equiv a \bmod q}} \Lambda(n)$$
$$\le \sum_{\substack{q\le 2|a|\\(q,a)=1}} \log|a| \sum_{\substack{1\le n<|a|\\ n\equiv a \bmod q}} 1 \ll \sum_{\substack{q\le 2|a|\\(q,a)=1}} \log|a| \left(\frac{|a|}{q}\right) \ll |a|\log^2|a|.$$

The proof for θ and θ^* is similar. □

Lemma 6.5. *Let $I \subset [1,x] \cap \mathbb{N}$. We have*

$$\sum_{q\in I} \left(\frac{x}{\phi(q)} - \frac{\psi(x)}{\phi(q)}\right) \ll xe^{-C\sqrt{\log x}} \tag{24}$$

where C is an absolute positive constant.

Proof. This follows from Lemma 6.2 and the prime number theorem. □

To prove Lemma 6.9 we will need bounds on $\zeta(s)$.

Lemma 6.6. *Define $\theta := \frac{32}{205}$ and take any $\epsilon > 0$. In the region $|\sigma + it - 1| > \frac{1}{10}$, we have*

$$\zeta(\sigma+it) \ll_\epsilon (|t|+1)^{\mu(\sigma)+\epsilon}$$

where

$$\mu(\sigma) = \begin{cases} 1/2-\sigma & \text{if } \sigma \le 0\\ 1/2+(2\theta-1)\sigma & \text{if } 0\le\sigma\le 1/2\\ 2\theta(1-\sigma) & \text{if } 1/2\le\sigma\le 1\\ 0 & \text{if } \sigma\ge 1.\end{cases}$$

Proof. For the values outside the critical strip, see for example section II.3.4 of [9]. In the critical strip, we use an estimate due to Huxley [8], which showed that $\zeta(1/2+it) \ll_\epsilon (|t|+1)^{\frac{32}{205}+\epsilon}$. The lemma then follows by convexity of μ.

□

Remark 6.7. Under Lindelöf's hypothesis, the conclusion of Lemma 6.6 holds with $\theta = 0$.

Lemma 6.8 (Perron's formula)**.** *Let $0 < \kappa < 1$, $y > 0$ and define*

$$h(y) := \begin{cases} 0 & \text{if } 0<y<1\\ 1-\frac{1}{y} & \text{if } y\ge 1.\end{cases} \tag{25}$$

We have

$$h(y) = \frac{1}{2\pi i}\int_{(\kappa)} \frac{y^s}{s(s+1)}ds. \tag{26}$$

Moreover, for $T \ge 1$ and positive $y \ne 1$, we have the estimate

$$h(y) = \frac{1}{2\pi i}\int_{\kappa-iT}^{\kappa+iT} \frac{y^s}{s(s+1)}ds + O\left(\frac{y^\kappa}{T^2|\log y|}\right). \tag{27}$$

Finally, for $y = 1$,

$$0 = h(1) = \frac{1}{2\pi i} \int_{\kappa - iT}^{\kappa + iT} \frac{ds}{s(s+1)} + O\left(\frac{1}{T}\right). \tag{28}$$

Proof. The first assertion is an easy application of the residue theorem.

Now take $y > 1$. We have again by the residue theorem that for any large integer $K \geq 3$ and for $T \geq 1$,

$$\begin{aligned} \frac{1}{2\pi i} \int_{\kappa - iT}^{\kappa + iT} \frac{y^s}{s(s+1)} ds - h(y) &= \frac{1}{2\pi i} \left(\int_{\kappa - iT}^{\kappa - K - iT} + \int_{\kappa - K - iT}^{\kappa - K + iT} + \int_{\kappa - K + iT}^{\kappa + iT} \right) \frac{y^s}{s(s+1)} ds \\ &\ll \frac{1}{T^2} \int_{\kappa - K}^{\kappa} y^{\sigma} d\sigma + \frac{y^{\kappa - K}}{|\kappa - K|^2} \int_{\kappa - K - iT}^{\kappa - K + iT} |ds| \ll \frac{y^{\kappa}}{T^2 |\log y|} + T \frac{y^{\kappa - K}}{|\kappa - K|^2}. \end{aligned}$$

We deduce the second assertion of the lemma by letting K tend to infinity. The proof is similar in the case $0 < y < 1$.

The last case remaining is for $y = 1$. We have

$$\begin{aligned} \frac{1}{2\pi i} \int_{\kappa - iT}^{\kappa + iT} \frac{ds}{s(s+1)} &= \frac{1}{2\pi i} \log \left(\frac{1 + \frac{1}{\kappa - iT}}{1 + \frac{1}{\kappa + iT}} \right) \\ &= \frac{1}{2\pi i} \log \left(1 + O\left(\frac{1}{T}\right) \right) \\ &= O\left(\frac{1}{T}\right) \end{aligned}$$

which concludes the proof.

□

The following is a crucial lemma estimating a weighted sum of the inverse of the totient function.

Lemma 6.9. *Let $a \neq 0$ be an integer and $M \geq 1$ be a real number.*

If $\omega(a_M) \geq 1$,

$$\sum_{\substack{n \leq M \\ (n,a)=1}} \frac{1}{\phi(n)} \left(1 - \frac{n}{M}\right) = C_1(a_M) \log M + C_3(a_M) + \frac{\phi(a_M)}{a_M} \frac{\Lambda(a_M)}{2M} + E(M,a). \tag{29}$$

If $a_M = 1$,

$$\sum_{\substack{n \leq M \\ (n,a)=1}} \frac{1}{\phi(n)} \left(1 - \frac{n}{M}\right) = C_1(1) \log M + C_3(1) + \frac{1}{2} \frac{\log M}{M} + \frac{C_5}{M} + E(M,a). \tag{30}$$

The error term $E(M,a)$ satisfies

$$\frac{a_M}{\phi(a_M)} E(M,a) \ll_{\epsilon} \frac{2^{\omega(a_M)}}{M} \left(\frac{a_M}{M}\right)^{\frac{205}{538} - \epsilon}. \tag{31}$$

Remark 6.10. Under Lindelöf's hypothesis,

$$\frac{a_M}{\phi(a_M)}E(M,a) \ll_\epsilon \frac{2^{\omega(a_M)}}{M}\left(\frac{a_M}{M}\right)^{1/2-\epsilon}. \tag{32}$$

Proof. Note first that we need only to consider the prime factors of a less or equal to M, since for $1 \le n \le M$, $(n,a)=1 \;\Leftrightarrow (n,a_M)=1$.

To calculate our sum we will write it as a contour integral and shift contours, showing that the contribution of the shifted contours is negligible and obtaining the main terms from the residues at the poles.

Setting $\kappa = \frac{1}{\log M}$ in Lemma 6.8,

$$\begin{aligned}\sum_{\substack{n\le M\\(n,a)=1}} \frac{1}{\phi(n)}\left(1-\frac{n}{M}\right) &= \sum_{(n,a_M)=1}\frac{1}{\phi(n)}h\left(\frac{M}{n}\right)\\ &= \sum_{(n,a_M)=1}\frac{1}{\phi(n)}\frac{1}{2\pi i}\int_{\kappa-iT}^{\kappa+iT}\left(\frac{M}{n}\right)^s\frac{ds}{s(s+1)}\\ &\qquad + O\left(\frac{1}{T^2}\sum_{n\ne M}\frac{1}{\phi(n)|\log M/n|}\left(\frac{M}{n}\right)^\kappa + \frac{\log M}{TM}\right)\\ &= \frac{1}{2\pi i}\int_{\kappa-iT}^{\kappa+iT}\left(\sum_{(n,a_M)=1}\frac{1}{n^s\phi(n)}\right)\frac{M^s}{s(s+1)}ds + O_M\left(\frac{1}{T}\right).\end{aligned} \tag{33}$$

In the last step we used the elementary estimates

$$\sum_{n\le M}\frac{1}{\phi(n)} \ll \log M \quad \text{and} \quad \sum_{n>M}\frac{1}{n^\kappa\phi(n)} \ll \log^2 M.$$

Now taking Euler products we compute that

$$\sum_{(n,a_M)=1}\frac{1}{n^s\phi(n)} = \mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1) \tag{34}$$

where

$$\mathfrak{S}_{a_M}(s+1) := \prod_{\substack{p|a\\p\le M}}\left(1-\frac{1}{p^{s+1}}\right)\left(1+\frac{1}{(p-1)p^{s+1}}\right)^{-1} \tag{35}$$

and

$$Z_2(s+1) := \prod_p\left(1+\frac{1}{p(p-1)}\left(\frac{1}{p^{s+1}}-\frac{1}{p^{2s+2}}\right)\right) \tag{36}$$

which converges for $\Re s > -3/2$.

Therefore, (33) becomes

$$\sum_{\substack{n\le M\\(n,a)=1}} \frac{1}{\phi(n)}\left(1-\frac{n}{M}\right) = \frac{1}{2\pi i}\int_{\kappa-iT}^{\kappa+iT} \mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\frac{M^s}{s(s+1)}ds + O_M\left(\frac{1}{T}\right). \quad (37)$$

The different results for different values of $\omega(a_M)$ come from the pole at $s=-1$. We see that $\mathfrak{S}_{a_M}(s+1)$ has a zero of order $\omega(a_M)$ at $s=-1$ whereas

$$\frac{\zeta(s+1)\zeta(s+2)}{s(s+1)}$$

has a pole of order two at $s=-1$. Hence the product has no pole if $\omega(a_M)\ge 2$, a pole of order one if $\omega(a_M)=1$, and a pole of order two if $\omega(a_M)=0$. We now shift the contour of integration to the left until the line $\Re(s)=\sigma$, where $-1-\frac{1}{2+4\theta}<\sigma<-1$ and $\theta := \frac{32}{205}$. The right hand side of (37) becomes

$$= P_T + \frac{1}{2\pi i}\int_{\sigma-iT}^{\sigma+iT} \mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\frac{M^s}{s(s+1)}ds + O_M\left(\frac{1}{T}\right) + O_a\left(\frac{\log^2 T}{T^2}\left(T^{1/6}+\frac{T^{1/2}}{M^{1/2}}+\frac{T^{7/6}}{M}\right)\right). \quad (38)$$

Here, P_T denotes the sum of all residues in the box $\sigma\le\Re s\le\frac{1}{\log M}$ and $|\Im s|\le T$. The second error term in (38) comes from the horizontal integrals which we have bounded using Lemma 6.6 (note that $\theta<1/6$). Taking $T\to\infty$ yields

$$\sum_{\substack{n\le M\\(n,a)=1}} \frac{1}{\phi(n)}\left(1-\frac{n}{M}\right) = P_\infty + E(M,a) \quad (39)$$

where

$$E(M,a) := \frac{1}{2\pi i}\int_{(\sigma)} \mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\frac{M^s}{s(s+1)}ds.$$

Now on the line $\Re s=\sigma$ we have the bound (note that $0<-1-\sigma<\frac{1}{2+4\theta}$)

$$\mathfrak{S}_{a_M}(s+1) \ll \left(\frac{3}{2}\right)^{\omega(a_M)} \prod_{\substack{p\mid a\\p\le M}} p^{-1-\sigma} = \left(\frac{3}{2}\right)^{\omega(a_M)} a_M^{-1-\sigma}. \quad (40)$$

Combining this with Lemma 6.6 and the trivial estimate $\frac{a_M}{\phi(a_M)} \ll \left(\frac{4}{3}\right)^{\omega(a_M)}$ yields

$$\begin{aligned}\frac{a_M}{\phi(a_M)}E(M,a) &\ll_\sigma \left(\frac{4}{3}\cdot\frac{3}{2}\right)^{\omega(a_M)} a_M^{-1-\sigma}\int_{-\infty}^{\infty}|\zeta(\sigma+1+it)||\zeta(\sigma+2+it)|\frac{M^\sigma}{(|t|+1)^2}dt \\ &\ll 2^{\omega(a_M)}\frac{1}{M}\left(\frac{a_M}{M}\right)^{-1-\sigma}\int_{-\infty}^{\infty}\frac{(|t|+1)^{1/2-(\sigma+1)}(|t|+1)^{2\theta(1-(\sigma+2))}}{(|t|+1)^2}dt \qquad (41)\\ &\ll_\sigma 2^{\omega(a_M)}\frac{1}{M}\left(\frac{a_M}{M}\right)^{-1-\sigma}\end{aligned}$$

since $1/2-(\sigma+1)+2\theta(1-(\sigma+2))<1$ by our choice of σ. The claimed bound on $E(M,a)$ then follows by taking $\sigma := -1-\frac{1}{2+4\theta}+\epsilon$ in (41).

It remains to compute P_∞ which is the sum of the residues of $\mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\frac{M^s}{s(s+1)}$ in the region $\sigma \le \Re s \le \frac{1}{\log M}$. Note that $\mathfrak{S}_{a_M}(s+1)$ has poles on the lines $\Re s = -1-\frac{\log(p-1)}{\log p}$, however these poles are cancelled by the zeros of $Z_2(s+1)$. Thus the only possible singularities of $\mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\frac{M^s}{s(s+1)}$ in the region in question are at the points $s=0$ and $s=-1$.

Now a lengthy but straightfoward computation shows that we have a double pole at $s=0$ with residue equal to $C_1(a_M)\log M + C_3(a_M)$.

As for $s=-1$, we have to consider three cases.

If $\omega(a_M)\ge 2$, then $\mathfrak{S}_{a_M}(s+1) = O((s+1)^2)$ around $s=-1$, so $\mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\frac{M^s}{s(s+1)}$ is holomorphic and we don't have any residue.

If $\omega(a_M)=1$, then $\mathfrak{S}_{a_M}(s+1)$ has a simple zero at $s=-1$ and thus $\mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\frac{M^s}{s(s+1)}$ has a simple pole with residue equal to $\frac{\phi(a_M)}{a_M}\frac{\Lambda(a_M)}{2M}$.

Finally, if $a_M=1$, then $\mathfrak{S}_{a_M}(s)\equiv 1$ and thus $\mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\frac{M^s}{s(s+1)}$ has a double pole at $s=-1$ with residue equal to $\frac{1}{2}\frac{\log M}{M}+\frac{C_5}{M}$

□

7. Further results and proofs

We will start by giving the fundamental result of this paper which works for M fixed as well as for M varying with x under the condition $M \le (\log x)^{O(1)}$.

Proposition 7.1. *Fix $A > B > 0$ two positive real numbers. Let $M = M(x)$ be an integer such that $1 \le M(x) \le \log^B x$. For a in the range $0 < |a| \le x^\lambda$ we have that*

$$\sum_{\substack{q\le \frac{x}{M}\\(q,a)=1}}\left(\psi(x;q,a)-\frac{\psi(x)}{\phi(q)}-\Lambda(a)\right) = x\left(C_1(a)\log M + C_3(a) - \sum_{\substack{r\le M\\(r,a)=1}}\frac{1}{\phi(r)}\left(1-\frac{r}{M}\right)\right) + O_{A,B}\left(\frac{\phi(a)}{a}2^{\omega(a)}\frac{x}{\log^A x}\right). \qquad (42)$$

We can remove the condition of M being an integer at the cost of adding the error term $O(x\log\log M/M^2)$.

Proof. We will prove that

$$\sum_{\substack{q\le \frac{x}{M}\\(q,a)=1}}\left(\psi(x;q,a)-\frac{x}{\phi(q)}-\Lambda(a)\right)=x\left(C_1(a)\log M+C_3(a)-\sum_{\substack{r\le M\\(r,a)=1}}\frac{1}{\phi(r)}\left(1-\frac{r}{M}\right)\right)$$
$$+O_{A,B}\left(\frac{\phi(a)}{a}2^{\omega(a)}\frac{x}{\log^A x}\right). \quad (43)$$

From this we can deduce the proposition since by Lemma 6.5 the difference between the left hand side of (42) and that of (43) is negligible.

Define $L:=\log^{A+3}x$. Partitioning the sum into dyadic intervals and applying Theorem 2.1 gives

$$\sum_{\substack{q\le \frac{x}{L}\\(q,a)=1}}\left(\psi(x;q,a)-\frac{x}{\phi(q)}-\Lambda(a)\right)=O_A\left(2^{\omega(a)}\frac{x}{\log^{A+1}x}\right). \quad (44)$$

Therefore, we need to compute

$$\sum_{\substack{\frac{x}{L}<q\le \frac{x}{M}\\(q,a)=1}}\left(\psi(x;q,a)-\frac{x}{\phi(q)}-\Lambda(a)\right)$$

which by lemmas 6.3 and 6.4 is equal to

$$\sum_{\substack{\frac{x}{L}<q\le \frac{x}{M}\\(q,a)=1}}\left(\theta^*(x;q,a)-\frac{x}{\phi(q)}\right)+O(x^{2/3}+|a|\log^2|a|). \quad (45)$$

We split the sum in (45) in three distrinct sums as following:

$$\sum_{\substack{\frac{x}{L}<q\le x\\(q,a)=1}}\theta^*(x;q,a)-\sum_{\substack{\frac{x}{M}<q\le x\\(q,a)=1}}\theta^*(x;q,a)-x\sum_{\substack{\frac{x}{L}<q\le \frac{x}{M}\\(q,a)=1}}\frac{1}{\phi(q)}=I-II-III.$$

The third sum is easily treated using Lemma 6.2:

$$III=x\sum_{\substack{\frac{x}{L}<q\le \frac{x}{M}\\(q,a)=1}}\frac{1}{\phi(q)}=x\left(C_1(a)\log(x/M)+C_1(a)+C_3(a)+O\left(2^{\omega(a)}\frac{\log(x/M)}{x/M}\right)\right.$$
$$\left.-\left(C_1(a)\log(x/L)+C_1(a)+C_3(a)+O\left(2^{\omega(a)}\frac{\log(x/L)}{x/L}\right)\right)\right)$$
$$=C_1(a)x\log(L/M)+O\left(2^{\omega(a)}L\log x\right). \quad (46)$$

For the first sum, we have

$$I=\sum_{\substack{\frac{x}{L}<q\le x\\(q,a)=1}}\theta^*(x;q,a)=\sum_{\substack{\frac{x}{L}<q\le x\\(q,a)=1}}\sum_{\substack{|a|<p\le x\\p\equiv a \bmod q}}\log p.$$

Using Lemma 6.1,

$$I = \sum_{\substack{1\le r<(x-a)\frac{L}{x}\\(r,a)=1}} \sum_{\substack{r\frac{x}{L}+a<p\le x\\ p\equiv a \bmod r\\ p>|a|}} \log p + O(|a|\log x)$$

$$= \sum_{\substack{1\le r<(x-a)\frac{L}{x}\\(r,a)=1}} \left(\theta^*(x;r,a) - \theta^*\left(r\frac{x}{L}+a;r,a\right)\right) + O(x^{1/3})$$

$$= \sum_{\substack{1\le r<L-a\frac{L}{x}\\(r,a)=1}} \left(\frac{x}{\phi(r)} - \frac{rx}{\phi(r)L}\right) + O(|a|L\log^2 x)) + O_A\left(L\frac{x}{L\log^{A+1} x}\right).$$

In the last step we used Lemma 6.4 (with $|a| < x^{\lambda}$) combined with the Siegel-Walfisz theorem in the form $\theta(x;r,a) - \frac{x}{\phi(r)} \ll_A \frac{x}{L\log^{A+1}x}$ for $r \le 2L$, as well as the estimate $\sum_{r\le R} \frac{|a|}{\phi(r)} \ll |a|\log R$. As $\left|a\frac{L}{x}\right| \le 1$ for x large enough and $\phi(r) \gg r/\log\log r$, this gives

$$I = x\sum_{\substack{r<L\\(r,a)=1}} \frac{1}{\phi(r)}\left(1-\frac{r}{L}\right) + O_A\left(\frac{x}{\log^{A+1} x}\right) + O\left(\frac{x\log\log L}{L}\right)$$

$$= x\sum_{\substack{r<L\\(r,a)=1}} \frac{1}{\phi(r)}\left(1-\frac{r}{L}\right) + O_A\left(\frac{x}{\log^{A+1} x}\right).$$

We conclude the evaluation of I by applying Lemma 6.2:

$$I = x\left(C_1(a)\log L + C_3(a) + O_A\left(\frac{2^{\omega(a)}}{\log^{A+1} x}\right)\right). \tag{47}$$

Now with a similar computation using lemmas 6.1 and 6.4 as well as the Siegel-Walfisz theorem in the form $\theta(x;r,a) - \frac{x}{\phi(r)} \ll_{A,B} \frac{x}{M\log^{A+1}x}$ for $r \le 2M$, we show that

$$II = \sum_{\substack{1\le r<(x-a)\frac{M}{x}\\(r,a)=1}} \left(\theta^*(x;r,a) - \theta^*\left(r\frac{x}{M}+a,r,a\right)\right) + O(|a|\log x)$$

$$= \sum_{\substack{1\le r<M-a\frac{M}{x}\\(r,a)=1}} \left(\frac{x}{\phi(r)} - \frac{rx}{\phi(r)M}\right) + O(|a|M\log^2 x) + O_{A,B}\left(M\frac{x}{M\log^{A+1} x}\right)$$

$$= x\sum_{\substack{1\le r\le M\\(r,a)=1}} \frac{1}{\phi(r)}\left(1-\frac{r}{M}\right) + O_{A,B}\left(\frac{x}{\log^{A+1} x}\right).$$

In the last step we have used that M is an integer so the term for $r = M$ in the sum is given by $\frac{1}{\phi(M)}\left(1-\frac{M}{M}\right) = 0$. If $M \notin \mathbb{N}$, we have to add the error term $\frac{x}{\phi(\lfloor M\rfloor)}\left(1-\frac{\lfloor M\rfloor}{M}\right) = O(x\log\log M/M^2)$.

We conclude the proof by combining our estimates for I, II and III with the bound $\frac{a}{\phi(a)} \ll \log\log x$. □

Proof of Proposition 3.4. Follows from Proposition 7.1. □

For larger values of M, we will need to use Lemma 6.9. We will denote by $P_M^-(n)$ the least prime factor of n which is less or equal to M.

Theorem 7.2. *Let $M = M(x) \le \log^B x$ (not necessarily an integer) where $B > 0$ is a fixed real number. Let $a \neq 0$ be an integer such that $|a| \le x^\lambda$.*

If $\omega(a_M) \ge 1$,

$$\frac{1}{\frac{\phi(a)}{a}\frac{x}{M}} \sum_{\substack{q \le \frac{x}{M} \\ (q,a)=1}} \left(\psi(x;q,a) - \frac{\psi(x)}{\phi(q)} - \Lambda(a) \right) = -\frac{1}{2}\Lambda(a_M) + E(M,a). \tag{48}$$

If $a_M = 1$,

$$\frac{1}{\frac{\phi(a)}{a}\frac{x}{M}} \sum_{\substack{q \le \frac{x}{M} \\ (q,a)=1}} \left(\psi(x;q,a) - \frac{\psi(x)}{\phi(q)} - \Lambda(a) \right) = -\frac{1}{2}\log M - C_5 + E(M,a). \tag{49}$$

The error term $E(M,a)$ satisfies

$$E(M,a) \ll_{\epsilon,B} 2^{\omega(a)} \left(\frac{a_M}{M}\right)^{\frac{205}{538}-\epsilon} + M\log M \frac{2^{\omega(a)}}{P_M^-(a)}. \tag{50}$$

Remark 7.3. The term $M\log M\frac{2^{\omega(a)}}{P_M^-(a)}$ in the bound for $E(M,a)$ can be deleted if a has no prime factor greater than M.

Remark 7.4. If we assume Lindelöf's hypothesis, we can replace the first term of the right hand side of (50) by $2^{\omega(a)}\left(\frac{a_M}{M}\right)^{1/2-\epsilon}$.

Proof of Theorem 7.2. Our starting point will be to set $A := 2B$ in Proposition 7.1:

$$\sum_{\substack{q \le \frac{x}{M} \\ (q,a)=1}} \left(\psi(x;q,a) - \frac{\psi(x)}{\phi(q)} - \Lambda(a) \right) = x\left(C_1(a)\log M + C_3(a) - \sum_{\substack{r \le M \\ (r,a)=1}} \frac{1}{\phi(r)}\left(1 - \frac{r}{M}\right)\right)$$
$$+ O_B\left(\frac{\phi(a)}{a} 2^{\omega(a)} \frac{x}{M^2} + x\frac{\log\log M}{M^2}\right). \tag{51}$$

We now consider three cases depending on the number of prime factors of a_M.

Case 1: $\omega(a_M) \ge 2$, which implies $\omega(a) \ge 2$. Applying Lemma 6.9 gives

$$\sum_{\substack{q \le \frac{x}{M} \\ (q,a)=1}} \left(\psi(x;q,a) - \frac{\psi(x)}{\phi(q)} - \Lambda(a) \right) \ll_B x(|C_1(a_M) - C_1(a)|\log M + |C_3(a_M) - C_3(a)|$$
$$+ |E(M,a)|) + \frac{\phi(a)}{a} 2^{\omega(a)} \frac{x}{M^2} + x\frac{\log\log M}{M^2} \tag{52}$$

with

$$\frac{a_M}{\phi(a_M)} E(M,a) \ll_\epsilon \frac{2^{\omega(a_M)}}{M} \left(\frac{a_M}{M}\right)^{\frac{205}{538}-\epsilon}.$$

This last bound implies that

$$\frac{a}{\phi(a)}E(M,a) \ll_\epsilon \frac{2^{\omega(a)}}{M}\left(\frac{a_M}{M}\right)^{\frac{205}{538}-\epsilon}.$$

If all the prime factors of a are less or equal to M, then $|C_i(a_M) - C_i(a)| = 0$. If not, we need upper bounds.

By the definition of $C_1(a)$,

$$|C_1(a_M) - C_1(a)| = C_1(a)\left(\prod_{\substack{p|a\\p>M}}\left(1-\frac{1}{p}\right)^{-1}\left(1-\frac{1}{p^2-p+1}\right)^{-1} - 1\right). \tag{53}$$

For M large enough, the terms in the product are $\le 1 + \frac{2}{p}$ (since $p > M$), so this is

$$\ll \frac{\phi(a)}{a}\left(\prod_{\substack{p|a\\p>M}}\left(1+\frac{2}{p}\right) - 1\right) \le \frac{\phi(a)}{a}2^{\omega(a)}\frac{2}{P_M^-(a)} \tag{54}$$

since the number of terms in the expanded product is at most $2^{\omega(a)}$. Using this bound, we compute that

$$\begin{aligned} C_3(a_M) - C_3(a) &= \left(C_1(a) + O\left(\frac{\phi(a)}{a}\frac{2^{\omega(a)}}{P_M^-(a)}\right)\right)\frac{C_3(a_M)}{C_1(a_M)} - C_3(a) \\ &= C_1(a)\left(\frac{C_3(a_M)}{C_1(a_M)} - \frac{C_3(a)}{C_1(a)}\right) + O\left(\frac{\phi(a)}{a}\frac{2^{\omega(a)}}{P_M^-(a)}\log M\right) \end{aligned}$$

since $\sum_{p|a_M}\frac{p^2\log p}{(p-1)(p^2-p+1)} \ll \sum_{p\le M}\frac{\log p}{p} \ll \log M$. Thus,

$$\begin{aligned} |C_3(a_M) - C_3(a)| &\ll \frac{\phi(a)}{a}\sum_{\substack{p|a\\p>M}}\frac{p^2\log p}{(p-1)(p^2-p+1)} + \frac{\phi(a)}{a}\frac{2^{\omega(a)}}{P_M^-(a)}\log M \\ &\ll \frac{\phi(a)}{a}\frac{2^{\omega(a)}}{P_M^-(a)}\log M. \end{aligned}$$

Putting all this together, dividing by $\frac{x}{M}\frac{\phi(a)}{a}$ and using the bounds $\frac{a}{\phi(a)} \ll \log\log x$ and $\frac{a}{\phi(a)} \ll 2^{\omega(a)}$ gives the claimed estimate.

Case 2: $\omega(a_M) = 1$. The calculation is similar, however Lemma 6.9 gives a contribution of

$$-\frac{1}{2}\frac{a}{\phi(a)}\frac{\phi(a_M)}{a_M}\Lambda(a_M) = -\frac{1}{2}\Lambda(a_M) + O\left(M\frac{2^{\omega(a)}}{P_M^-(a)}\right)$$

since one can show by the prime number theorem that $\log|a_M| \ll M$.

Case 3: $a_M = 1$. The contribution of

$$\frac{a}{\phi(a)}\left(-\frac{1}{2}\log M - C_5\right) = -\frac{1}{2}\log M - C_5 + O\left(\log M\frac{2^{\omega(a)}}{P_M^-(a)}\right)$$

comes from Lemma 6.9.

□

Proof of Theorem 1.1. It is a particular case of Theorem 7.2. Note that for a fixed and for x large enough we get that $a_M = a$, so by Remark 7.3 the error term in Theorem 7.2 satisfies $E(M,a) \ll_{a,\epsilon,B} \left(\frac{a_M}{M}\right)^{\frac{205}{538}-\epsilon}$. □

The final step is to count for how many integers a the error term in Theorem 7.2 is small.

Lemma 7.5. *Fix $B > 4\log 4$, set $X := x^\lambda$, $M := \log^B x$ and let*

$$S(X) := \left\{1 \le a \le X : \omega(a_M) \ge 2, E(M,a) \ll_B \frac{1}{M^{1/4-\frac{\log 4}{B}}}\right\}$$

where $E(M,a)$ is defined as in Theorem 7.2. We have that $|S(X)| \ge \frac{4}{15}e^{-\gamma}X + o_B(X)$.

Proof. Define the set $T(X)$ consisting of the numbers of the form dr such that d is an integer in the interval $1 \le d \le M^{1/3}$ having at least two prime factors and r is a $M^{5/4}$-rough integer in the interval $M^{5/4} \le r \le \frac{X}{d}$. (We call an integer y-rough if all of its prime factors are greater or equal to y.) Define also $U(X) := \{1 \le a \le X : \omega(a) \le \frac{3}{2}\log\log X\}$.

The bound on $E(M,a)$ given in Theorem 7.2 shows that for $a \in T(X) \cap U(X)$ we have $E(M,a) \ll_B M^{\frac{\log 4}{B}-1/4}$, thus $T(X) \cap U(X) \subset S(X)$. The Hardy-Ramanujan theorem (see [6]) gives the estimate $|U(X)| = X(1+o(1))$. To give a lower bound for $T(X)$ we write

$$|T(X)| = \sum_{\substack{1\le d\le M^{1/3} \\ \omega(d)\ge 2}} \Phi\left(\frac{X}{d}, M^{5/4}\right)$$

where $\Phi(x,y)$ is the number of y-rough integers up to x. Equation (1.13) of [2] shows that in our range of d and for x large enough we have

$$\Phi\left(\frac{X}{d}, M^{5/4}\right) = \frac{X}{d}\prod_{p\le M^{5/4}}\left(1-\frac{1}{p}\right)\left(1+O_B\left(\exp\left(-\frac{\lambda}{3}M^{\frac{1}{2B}}\right)\right)\right).$$

Summing this for $d \le M^{1/3}$ such that $\omega(d) \ge 2$ we get that $|T(X)| \ge \frac{4}{15}e^{-\gamma}X(1+o_B(1))$. We finish our proof by putting all the pieces together:

$$\begin{aligned} |S(X)| \ge |T(X)\cap U(X)| &= |T(X)| + |U(X)| - |T(X)\cup U(X)| \\ &\ge \frac{4}{15}e^{-\gamma}X(1+o_B(1)) + X(1+o(1)) - X \\ &= \frac{4}{15}e^{-\gamma}X(1+o_B(1)). \end{aligned} \tag{55}$$

□

Proof of Theorem 3.2. This is a consequence of Theorem 7.2 and Lemma 7.5. □

8. Concluding remarks

Remark 8.1. One could ask if the results of Theorem 1.1 are intrinsic to the sequence of prime numbers or if they are just a result of the artificial weight $\Lambda(n)$ in the prime counting functions. However one can see that if we replace $\psi(x;q,a)$ by $\pi(x;q,a)$ and $\psi(x)$ by $\pi(x)$, the proof of Proposition 7.1 will go through with x replaced by $Li(x)$ and an additionnal error term of

$$O\left(x\frac{\log\log x}{\log^2 x}\right).$$

One has to prove the analogue of Theorem 2.1 which can be done using the Bombieri-Vinogradov theorem and a very delicate summation by parts. We conclude that an analogue of Theorem 1.1 holds with the natural prime counting functions in the range $M \leq \sqrt{\log x}$.

Theorem 8.2. *Fix an integer $a \neq 0$ and let $M = M(x) \leq \sqrt{\log x}$. We have*

$$\frac{1}{\frac{\phi(a)}{a}\frac{Li(x)}{M}} \sum_{\substack{q \leq \frac{x}{M} \\ (q,a)=1}} \left(\pi(x;q,a) - \frac{\pi(x)}{\phi(q)} - \frac{\vartheta(a)}{\log a}\right) = \mu(a,M) + O_{a,\epsilon}\left(\frac{1}{M^{\frac{205}{538}-\epsilon}}\right) \tag{56}$$

where $\mu(a,M)$ is defined as in Theorem 1.1.

Remark 8.3. Below we sketch an argument showing that the proportion of integers $a \leq x^\lambda$ for which the first term of the right hand side of (50) is ≤ 1 is not more than $e^{-\gamma}$. Setting $X := x^\lambda$ and using that $M \ll \log X$, we compute

$$\#\Big\{a \leq X : \prod_{\substack{p|a \\ p\leq M}} p \leq M\Big\} = \sum_{\substack{n\leq M \\ \mu^2(n)=1}} \#\Big\{a \leq X : n \mid a, \Big(\frac{a}{n}, \prod_{\substack{p\leq M \\ p\nmid n}} p\Big) = 1\Big\}$$

which by the fundamental lemma of the combinatorial sieve is

$$\sim \sum_{\substack{n\leq M \\ \mu^2(n)=1}} \frac{X}{n} \prod_{\substack{p\leq M \\ p\nmid n}} \left(1 - \frac{1}{p}\right) \sim \frac{Xe^{-\gamma}}{\log M} \sum_{n\leq M} \frac{\mu^2(n)}{\phi(n)} \sim Xe^{-\gamma}.$$

Theorefore, to extend the proportion of integers a for which we get an admissible error term in Remark 3.3, one would need to improve the bound (50) for $E(M,a)$.

References

[1] Enrico Bombieri, John B. Friedlander and Henryk Iwaniec, *Primes in arithmetic progressions to large moduli.* Acta Math. 156 (1986), no. 3-4, 203–251.

[2] Nicolaas Govert De Bruijn, *On the number of uncancelled elements in the sieve of Eratosthenes.* Nederl. Akad. Wetensch., Proc. 53, (1950) 803–812

[3] Étienne Fouvry, *Sur le problème des diviseurs de Titchmarsh.* J. Reine Angew. Math. 357 (1985), 51–76.

[4] John B. Friedlander, Andrew Granville, *Relevance of the residue class to the abundance of primes.* Proceedings of the Amalfi Conference on Analytic Number Theory (Maiori, 1989), 95–103, Univ. Salerno, Salerno, 1992.

[5] John B. Friedlander, Andrew Granville, Adolf Hildebrand, Helmut Maier, *Oscillation theorems for primes in arithmetic progressions and for sifting functions.* J. Amer. Math. Soc. 4 (1991), no. 1, 25–86.

[6] Godfrey Harold Hardy, Srinivasa Ramanujan, *The normal number of prime factors of a number n.* Quart. J. Math. 48 (1917), 76-92.

[7] Christopher Hooley, *On the Barban-Davenport-Halberstam theorem. I.* Collection of articles dedicated to Helmut Hasse on his seventy-fifth birthday, III. J. Reine Angew. Math. 274/275 (1975), 206–223
[8] Martin N. Huxley, *Exponential sums and the Riemann zeta function. V.* Proc. London Math. Soc. (3) 90 (2005), no. 1, 1–41.
[9] Gérald Tenenbaum, *Introduction à la théorie analytique et probabiliste des nombres.* Deuxième édition, Cours Spécialisés, 1. Société Mathématique de France, Paris, 1995. xv+457 pp.

Département de mathématiques et de statistique, Université de Montréal, CP 6128, succ. Centre-ville, Montréal, QC, Canada H3C 3J7
E-mail address: fiorilli@dms.umontreal.ca