**Ingela Mennema**

# The distribution of consecutive square-free numbers

**Master Thesis**

**Supervisor: Dr. E. Sofos**

**Co-supervisor: Dr. J.-H. Evertse**

**16 May 2017**

**Department of Mathematics**
**Leiden University**

# Contents



# 1 Introduction

Dirichlet proved in 1837 that for any two positive integers $a$ and $q$ there are infinitely many primes $p$ such that $p \equiv a \bmod q$. Let

$$\pi(x;q,a) := \#\{p \le x : p \equiv a \bmod q\}.$$

The Prime number theorem in arithmetic progressions states

$$\pi(x;q,a) \sim \frac{1}{\phi(q)} \frac{x}{\log x} \text{ as } x \to \infty,$$

where $\phi$ is the Euler phi function. Chebyshev observed in 1853 that primes congruent to 3 modulo 4 are more frequent, in a specific quantitative sense, than primes congruent to 1 modulo 4. Specifically, one would expect that the inequalities $\pi(x;4,1) > \pi(x;4,3)$ and $\pi(x;4,1) < \pi(x;4,3)$ each occur half of the time, since asymptotically half of the primes are of the form $4k+1$ and half of the form $4k+3$. However, numerical evidence shows that the inequality $\pi(x;4,1) < \pi(x;4,3)$ occurs much more frequently. This phenomenon, which holds in greater generality, is known as *Chebyshev's bias* and has been studied extensively, see for example the work of Rubinstein and Sarnak [10].
Let $p_n$ denote the $n$-th prime and for integers $q, a, b$ with $\gcd(q,a) = \gcd(q,b) = 1$ let

$$\pi(x;q,(a,b)) := \#\big\{p_n \le x : p_n \equiv a \bmod q,\ p_{n+1} \equiv b \bmod q\big\}.$$

This quantity is related to the Chebyshev's bias, because it considers the patterns of residues modulo $q$ among consecutive primes. It was recently studied by Lemke Oliver and Soundararajan [6]. They proved certain asymptotic estimates for it, conditionally on the, still open, Hardy–Littlewood conjecture which regards prime values of collections of polynomials. A very special case of their results is that for all large enough $x$ one has

$$\pi(x;q,(a,b)) > \pi\left(x;q,(a',b')\right)$$

whenever $a \not\equiv b \bmod q$ and $a' \equiv b' \bmod q$.

Our aim in this thesis is to investigate similar phenomena for the sequence of square-free integers. Recall that a positive integer is called *square-free* if it is not divisible by the square of a prime. This sequence is easier to handle and one might hope that such bias can be proved without resorting to an unproved hypothesis.
Let $S(x)$ denote the number of square-free integers up to $x$, then it is well-known that

$$\lim_{x\to\infty} \frac{S(x)}{x} = \prod_p \left(1 - \frac{1}{p^2}\right) = \frac{6}{\pi^2}.$$

Let $a \in \mathbb{Z}$, $q \in \mathbb{N}$ and let $S(x; q, a)$ be the number of square-free numbers $s \leq x$ such that $s \equiv a \bmod q$. The proof of the following result is standard and can be found in Proposition 2.12,

$$S(x; q, a) = \frac{x}{q} \prod_{\substack{p \\ \gcd\left(p^2, q\right) \mid a}} \left(1 - \frac{\gcd\left(p^2, q\right)}{p^2}\right) + O(\sqrt{x}),$$

where the definition of the Landau's $O-$notation is given in the notation passage at the end of the introduction. Denote by $s_n$ the $n$-th square-free number and define for $x \in \mathbb{R}_{\geq 1}$ and $a, b, q$ integers with $q > 0$ the quantity

$$S\left(x; q, (a, b)\right) := \#\left\{s_n \leq x : s_n \equiv a \bmod q,\ s_{n+1} \equiv b \bmod q\right\}.$$

The main theorem in this thesis is about the probability that out of all square-free integers $s_n \leq x$ one has $s_n \equiv a \bmod q$ and $s_{n+1} \equiv b \bmod q$, this probability will be denoted by

$$\mathbb{P}_x(q; (a, b)) := \frac{S\left(x; q, (a, b)\right)}{x}.$$

The proof of the following result is in §4.

**Theorem 1.1.** *For any integers $q, a, b$ satisfying*

$$0 \leq a < q \quad \text{and} \quad a < b \leq q + a$$

*the following limit exists,*

$$\lim_{x\to+\infty} \mathbb{P}_x(q; (a, b)).$$

*Furthermore, denoting its value by $\ell(q; (a, b))$, the following rate of convergence holds:*

$$\mathbb{P}_x(q; (a, b)) = \ell(q; (a, b)) + O\left(\frac{1}{\log\log x}\right),$$

*where the implied constant depends on $q, a$ and $b$.*

The size of gaps between consecutive square-free integers has been a topic of intensive research; we shall draw upon work related to the entity

$$M_\alpha(x) := \sum_{s_{n+1}\le x} (s_{n+1}-s_n)^\alpha, \ (\alpha \in \mathbb{R}_{\ge 0}).$$

If one wishes to show that many of the gaps $s_{n+1}-s_n$ are small then upper bounds on $M_\alpha(x)$ for a single $\alpha > 0$ and all large enough $x$ are required. It may be however that quite rarely there are very large gaps, these large gaps will make $M_\alpha(x)$ explode when $\alpha$ is a large positive integer. The first to study $M_\alpha(x)$ was Erdős [1] who showed in 1951 that for all $0 \le \alpha \le 2$ there exists a constant $C(\alpha)$ such that

$$\lim_{x\to+\infty} \frac{M_\alpha(x)}{x} = C(\alpha). \tag{1.1}$$

Several advances were made on the problem or proving (1.1) for larger values of $\alpha$ later. We give a table of all related developments here:

| | |
|---|---|
| $0 \le \alpha \le 2$ | Erdős [1],1951 |
| $0 \le \alpha \le 3$ | Hooley [4],1973 |
| $0 \le \alpha \le 3.22\ldots$ | Filaseta [2],1993 |
| $0 \le \alpha \le \frac{11}{3} = 3.66\ldots$ | Huxley [5],1995. |

To prove Theorem 1.1 we shall partition the pairs $(s_n, s_{n+1})$ in $S(x;q,(a,b))$ according to the size of the gap $s_{n+1}-s_n$. Big gaps will be dealt with using the estimate of Erdős above, while small gaps will be dealt with by proving Theorem 1.2 below. This theorem will be proven in §3 and forms the technical core of the present thesis. We will not do so here, but we should mention that using the estimate of Huxley instead of the estimate of Erdős one can improve the error term in Theorem 1.1 to $O\left(\frac{1}{(\log\log x)^\beta}\right)$ for some $\beta > 1$. The *Möbius function* is defined by

$$\mu(n) := \begin{cases} (-1)^t, & \text{if } n = p_1\cdots p_t \text{ where } p_1,...,p_t \text{ are distinct primes} \\ 0, & \text{if } n \text{ is not square-free.} \end{cases}$$

Note that $\mu(n)^2$ is an indicator function for square-free numbers $n$. For each small gap we will use the inclusion exclusion-principle. For example if $q = 4$, $a = 1$ and $b = 3$ we have that

$$\#\{n \le z : qn+1 \text{ is square-free and the next square-free is } qn+3\}$$

equals

$$\sum_{n\le z} \mu(4n+1)^2\mu(4n+3)^2 - \sum_{n\le z} \mu(4n+1)^2\mu(4n+2)^2\mu(4n+3)^2.$$

The following theorem provides an expression for each summation that appears if one applies the inclusion-exclusion principle on every small gap.

**Theorem 1.2.** *Let $q \neq 0, a_1, ..., a_k$ be integers with $0 \leq a_1 < a_2 < \cdots < a_k$ and define the Euler product $\gamma = \prod_p (1 - \omega(p) p^{-2})$, where*

$$\omega(p) = \# \left\{ 1 \leq m \leq p^2 : \exists\ 1 \leq i \leq k \text{ with } qm + a_i \equiv 0 \bmod p^2 \right\}.$$

*Then we have for $z \geq a_k$,*

$$\sum_{1 \leq n \leq z} \prod_{i=1}^{k} \mu(qn + a_i)^2 = \gamma z + O\left( z^{2/3} e^{2\sqrt{a_k}} (2k + \log z)^{O(k)} (2qz)^{O\left(\frac{k}{\log \log z}\right)} \right),$$

*where all implied constants depend at most on $q$.*

For the proof of this theorem we adopt the approach of Mirsky [7]. Tsang [11] and Reuss [8] also studied the sum in Theorem 1.2, however all previous work has focused on improving the dependence of the error term on $z$ but does not provide uniformity of the error term with respect to $k$. Making the dependence on $k$ explicit is the new element of Theorem 1.2. This feature is important for us because in our proof of the asymptotic for $S(x; q, (a, b))$ in Theorem 1.1 we need to study square-free values of many linear polynomials, the cardinality of which will be a function of $x$ that tends to infinity.

The constant $\ell = \ell(q; (a, b))$ appearing in Theorem 1.1 is unusually complicated, because it is an infinite sum of finite alternating sums of Euler products as is shown in (4.6). However, we can at least characterize all cases where $\ell(q; (a, b))$ is non-zero, the proof of the following result is in §5.

**Theorem 1.3.** *The limit $\ell(q; (a, b))$ in Theorem 1.1 is strictly positive if and only if there exist square-free integers $s, t$ fulfilling $s \equiv a \mod q$ and $t \equiv b \mod q$.*

**Notation.** Throughout this thesis we define $\mathbb{N} := \mathbb{Z}_{>0}$ and $(n, m) := \gcd(n, m)$. Furthermore $\nu_p(n)$ will refer to the usual $p$-adic valuation and we write $p^k \| n$ if $k = \nu_p(n)$. We use the notation $f(x) = O(g(x))$ and the notation $f(x) \ll g(x)$ if for all $x \geq 1$ there exists a positive real $C$ such that $|f(x)| \leq C g(x)$. All the implied constants in Landau's $O$-notation and in Vinogradov's $\ll$-notation may depend on the entities $a, b$ and $q$ that appear in Theorem 1.1, but not on other variables.

# 2 Preliminaries

This sections is devoted to some auxiliary results that are well-known.

**Definition 2.1.** An *arithmetic function* is a function

$$f : \mathbb{N} \to \mathbb{C}.$$

The associated *Dirichlet series* is defined as follows

$$L_f(s) := \sum_{n=1}^{\infty} f(n)n^{-s}$$

for every $s \in \mathbb{C}$ for which the series converges.

**Definition 2.2.** A *multiplicative function* is an arithmetic function $f : \mathbb{N} \to \mathbb{C}$ such that $f \not\equiv 0$ and $f(mn) = f(m)f(n)$ for every $m, n \in \mathbb{N}$ with $(m, n) = 1$.
A *strongly multiplicative function* is an arithmetic function $f : \mathbb{N} \to \mathbb{C}$ such that $f \not\equiv 0$ and $f(mn) = f(m)f(n)$ for every $m, n \in \mathbb{N}$.

**Lemma 2.3** (Euler)**.** *Let $f$ be a multiplicative function and let $s \in \mathbb{C}$ be such that $L_f(s)$ converges absolutely. Then*

$$L_f(s) = \prod_p \left( \sum_{j=0}^{\infty} f(p^j)\, p^{-js} \right).$$

**Corollary 2.4.** *Let $f$ be a strongly multiplicative function and let $s \in \mathbb{C}$ be such that $L_f(s)$ converges absolutely. Then*

$$L_f(s) = \prod_p \frac{1}{1 - f(p)p^{-s}}.$$

**Definition 2.5.** The *Riemann zeta function* is given by

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s},$$

for $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$.

**Lemma 2.6.** *For every $s \in \mathbb{C}$ with Re$(s) > 1$ we have*

$$\zeta(s) = \prod_p \frac{1}{1 - p^{-s}}.$$

**Remark 2.7.** The Möbius function is a multiplicative function.

**Lemma 2.8.** *For every $s \in \mathbb{C}$ with Re$(s) > 1$ we have*

$$\frac{1}{\zeta(s)} = L_\mu(s).$$

*Proof.* Note that $L_\mu(s)$ converges absolutely for every $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$. Combined with Lemma 2.3 and Remark 2.7 we obtain

$$L_\mu(s) = \prod_p \left( \sum_{j=0}^{\infty} \mu(p^j) p^{-js} \right) = \prod_p \left( 1 - \frac{1}{p^s} \right)$$

for every $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$. From Lemma 2.6 it follows that for every $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$ we have

$$\frac{1}{\zeta(s)} = L_\mu(s).$$

□

**Lemma 2.9.** *Let $n \in \mathbb{N}$. Then*

$$\sum_{d^2 | n} \mu(d) = \mu^2(n).$$

*Proof.* Write $n = p_1^{\alpha_1} \cdots p_t^{\alpha_t}$ where $p_1, ..., p_t$ are distinct primes and $\alpha_1, ..., \alpha_t \geq 1$. We have

$$\sum_{d^2 | n} \mu(d) = \prod_p \left( \sum_{\substack{k \in \mathbb{Z} \\ p^{2k} | n}} \mu(p^k) \right) = \prod_{i=1}^{m} \left( \sum_{\substack{0 \leq k \leq 1 \\ 2k \leq \alpha_i}} (-1)^k \right).$$

The inner sum equals 1 if $\alpha_i = 1$ and 0 if $\alpha_i \geq 2$. Hence $\sum_{d^2|n} \mu(d)$ equals 1 if $n$ is square-free and 0 otherwise. This implies that the sum is equal to $\mu(n)^2$. □

We shall repeatedly make use of the following standard fact.

**Lemma 2.10.** *Let $a, b \in \mathbb{Z}$ and $q \in \mathbb{N}$. There exists $n \in \mathbb{N}$ satisfying $bn \equiv a$ mod $q$ if and only if $(b, q)$ divides $a$.*

**Lemma 2.11.** *Let $a, b \in \mathbb{Z}$ and $q \in \mathbb{N}$ and $x \in \mathbb{R}$ with $x \geq 1$. Then*

$$\# \{1 \leq n \leq x : bn \equiv a \text{ mod } q\} = \begin{cases} \frac{x}{q}(b,q) + O(1) & \text{if } (b,q) \mid a \\ O(1) & \text{otherwise,} \end{cases}$$

*where the implied constants are absolute.*

*Proof.* Let $(b, q) \nmid a$, then from Lemma 2.10 it is clear that

$$\# \{1 \leq n \leq x : bn \equiv a \text{ mod } q\} = O(1).$$

Assume that $(b, q) \mid a$. By Lemma 2.10 there exists an integer $n$ such that $bn \equiv a$ mod $q$. Equivalently, for this $n$ the congruence $\frac{b}{(b,q)} n \equiv \frac{a}{(b,q)}$ mod $\frac{q}{(b,q)}$ holds. Since $\frac{b}{(b,q)}$ and $\frac{q}{(b,q)}$ are coprime, this congruence has a unique solution mod $\frac{q}{(b,q)}$, say $r$. Hence

$$\# \{1 \leq n \leq x : bn \equiv a \text{ mod } q\} = \# \left\{ 1 \leq n \leq x : n \equiv r \text{ mod } \tfrac{q}{(b,q)} \right\}.$$

Dividing the positive integers into blocks of length $\frac{q}{(b,q)}$, we obtain

$$\#\left\{1 \leq n \leq x : n \equiv r \bmod \frac{q}{(b,q)}\right\} = \frac{x}{q}(b,q) + O(1).$$

□

**Proposition 2.12.** *Let $a \in \mathbb{Z}$ and $q \in \mathbb{N}$. Then*

$$\sum_{\substack{1\leq n\leq x \\ n\equiv a \bmod q}} \mu(n)^2 = \frac{x}{q} \prod_{\substack{p \\ (p^2,q)|a}} \left(1 - \frac{(p^2,q)}{p^2}\right) + O(\sqrt{x}),$$

*where the implied constant is absolute.*

*Proof.* Using Lemma 2.9 we obtain

$$\sum_{\substack{1\leq n\leq x \\ n\equiv a \bmod q}} \mu(n)^2 = \sum_{\substack{1\leq n\leq x \\ n\equiv a \bmod q}} \sum_{d^2|n} \mu(d).$$

Reordering the summation gives

$$\sum_{\substack{1\leq n\leq x \\ n\equiv a \bmod q}} \sum_{d^2|n} \mu(d) = \sum_{d\leq\sqrt{x}} \mu(d) \sum_{\substack{1\leq n\leq x \\ n\equiv a \bmod q \\ d^2|n}} 1 = \sum_{d\leq\sqrt{x}} \mu(d) \sum_{\substack{k\leq \frac{x}{d^2} \\ kd^2\equiv a \bmod q}} 1.$$

By Lemma 2.10 the congruence $kd^2 \equiv a \bmod q$ has a solution if and only if $(d^2,q)$ divides $a$. Together with Lemma 2.11 it follows that the sum equals

$$\frac{x}{q} \sum_{\substack{d\leq\sqrt{x} \\ (d^2,q)|a}} \frac{\mu(d)}{d^2}(d^2,q) + O\left(\sum_{\substack{d\leq\sqrt{x} \\ (d^2,q)|a}} 1\right).$$

Clearly, $\sum_{\substack{d\leq\sqrt{x} \\ (d^2,q)|a}} 1 = O\left(\sqrt{x}\right)$. Furthermore,

$$\sum_{\substack{d\leq\sqrt{x} \\ (d^2,q)|a}} \frac{\mu(d)}{d^2}(d^2,q) = \sum_{\substack{d\in\mathbb{N} \\ (d^2,q)|a}} \frac{\mu(d)}{d^2}(d^2,q) + O\left(\sum_{\substack{d>\sqrt{x} \\ (d^2,q)|a}} \frac{(d^2,q)}{d^2}\right).$$

Note that $(d^2,q)$ is a multiplicative function with respect to $d$. By Lemma 2.3 it follows that

$$\sum_{\substack{d\in\mathbb{N} \\ (d^2,q)|a}} \frac{\mu(d)}{d^2}(d^2,q) = \prod_{\substack{p \\ (p^2,q)|a}} \left(1 - \frac{(p^2,q)}{p^2}\right).$$

Further,

$$\frac{x}{q}\sum_{\substack{d>\sqrt{x}\\(d^2,q)|a}}\frac{(d^2,q)}{d^2}\leq x\sum_{d>\sqrt{x}}\frac{1}{d^2},$$

where we have used

$$\sum_{d>\sqrt{x}}\frac{1}{d^2}=O\left(\frac{1}{\sqrt{x}}\right).$$

This proves the proposition. □

The following result was proved in a slightly stronger form by Tsang [11, Th. 3].

**Theorem 2.13** (Tsang). *Let $f_1,...,f_r \in \mathbb{Z}[x]$ such that $f_i$ is irreducible over $\mathbb{Z}$ and $\deg f_i \leq 3$ for all $i$. Then*

$$\sum_{1\leq n\leq x}\prod_{i=1}^{r}\mu(f_i(n))^2=cx+O(x^{3/4}),$$

*where $c=\prod_p(1-\omega_{\mathbf{f}}(p)p^{-2})$ with*

$$\omega_{\mathbf{f}}(p):=\#\left\{1\leq n\leq p^2:\exists\ 1\leq i\leq r \text{ such that } f_i(n)\equiv 0 \bmod p^2\right\}.$$

*The implied constant depends on $r$ as well as on the coefficients of each $f_i$.*

We will use the following upper bound for the number of primes up to $x$, a proof of which can be found in [9, Cor.1, Eq.(3.6)].

**Lemma 2.14.** *For all $x>1$ we have*

$$\pi(x)<1.25506\frac{x}{\log x}.$$

**Definition 2.15.** Let $s\in\mathbb{N}$. Define $\tau_s(d)$ to be the number of vectors $(d_1,...,d_s)\in\mathbb{N}^s$ with $d_1\cdots d_s=d$.

We shall henceforth use the following notational convention

$$\tau(d):=\tau_2(d)=\sum_{d|n}1.$$

**Lemma 2.16.** *Let $s\in\mathbb{Z}_{\geq 1}$. Then $\tau_{s+1}(d)\leq\tau(d)^s$.*

*Proof.* The proof is by induction. For $s = 1$, it is clear. Let $s \geq 1$ and assume that for all $k < s$ the theorem holds. Consider

$$\tau_{s+1}(d) = \sum_{d_2 | d} \tau_s\left(\frac{d}{d_2}\right).$$

Using $\frac{d}{d_2} \mid d$ for all $d_2 \mid d$, the induction hypothesis gives

$$\tau_{s+1}(d) \leq \tau_s(d) \sum_{d_2 | d} 1 = \tau(d)\tau_s(d) \leq \tau(d)^s.$$

□

**Lemma 2.17.** *For every $\epsilon > 0$ there exists a constant $C_\epsilon > 0$ such that $\tau(n) \leq C_\epsilon n^\epsilon$ for all $n \geq 1$.*

*Proof.* Clearly, $\tau\big(p^k\big) = k + 1$ for any prime $p$ and integer $k \geq 0$, hence

$$\frac{\tau(n)}{n^\epsilon} = \prod_{p^\alpha \| n} \frac{\alpha + 1}{p^{\alpha\epsilon}}.$$

We divide the product into two parts, one part with the primes $p < 2^{1/\epsilon}$ and the other part with the primes $p \geq 2^{1/\epsilon}$. First, let $p \geq 2^{1/\epsilon}$, so that $p^{\alpha\epsilon} \geq 2^\alpha$. Furthermore using the binomial formula we find

$$2^\alpha \geq \alpha + 1.$$

Hence

$$\frac{\alpha + 1}{p^{\alpha\epsilon}} \leq \frac{\alpha + 1}{2^\alpha} \leq 1.$$

Next, let $p < 2^{1/\epsilon}$. We have

$$p^{\alpha\epsilon} \geq 2^{\alpha\epsilon} = e^{\alpha\epsilon \log 2} \geq \alpha\epsilon \log 2.$$

Hence

$$\frac{\alpha + 1}{p^{\alpha\epsilon}} = \frac{\alpha}{p^{\alpha\epsilon}} + \frac{1}{p^{\alpha\epsilon}} \leq \frac{1}{\epsilon \log 2} + 1.$$

Combining all of the above gives

$$\begin{aligned}
\frac{\tau(n)}{n^\epsilon} &= \prod_{\substack{p^\alpha \| n \\ p < 2^{1/\epsilon}}} \frac{\alpha + 1}{p^{\alpha\epsilon}} \prod_{\substack{p^\alpha \| n \\ p \geq 2^{1/\epsilon}}} \frac{\alpha + 1}{p^{\alpha\epsilon}} \\
&\leq \prod_{p < 2^{1/\epsilon}} \left(1 + \frac{1}{\epsilon \log 2}\right).
\end{aligned}$$

Hence with $C_\epsilon = \prod_{p<2^{1/\epsilon}} \left(1 + \frac{1}{\epsilon \log 2}\right)$, we get

$$\tau(n) \leq C_\epsilon n^\epsilon.$$

□

**Lemma 2.18.** *For every $n, s \in \mathbb{N}$ we have $\tau_s(n) \leq n^{O\left(\frac{s}{\log\log 3n}\right)}$, where the implied constant is absolute.*

*Proof.* Applying the inequality $1 + x \leq \mathrm{e}^x$ for $x \in \mathbb{R}$ and using the proof of Lemma 2.17 we get

$$C_\epsilon \leq \prod_{p<2^{1/\epsilon}} \mathrm{e}^{\frac{1}{\epsilon \log 2}},$$

equivalently

$$\log(C_\epsilon) \leq \sum_{p \leq 2^{1/\epsilon}} \frac{1}{\epsilon \log 2}.$$

By Lemma 2.14 it follows that

$$\sum_{p \leq 2^{1/\epsilon}} \frac{1}{\epsilon \log 2} \leq \frac{2}{(\log 2)^2} 2^{1/\epsilon}.$$

With $\epsilon = \frac{2 \log 2}{\log\log 3n}$. We note that $2^{1/\epsilon} = (\log 3n)^{1/2}$, therefore

$$\log(C_\epsilon) \leq \frac{2}{(\log 2)^2} (\log 3n)^{1/2}.$$

This yields the following inequality

$$C_\epsilon \leq n^{\frac{2}{(\log 2)^2 (\log 3n)^{1/2}}},$$

hence by Lemma 2.16 and 2.17 we obtain

$$\tau_s(n) \leq \tau(n)^s \leq n^{O\left(\frac{s}{\log\log 3n}\right)}.$$

□

**Lemma 2.19.** *There exists $C_1 > 1$ such that for all $r \in \mathbb{N}_{\geq 2}$ and for all $x \in \mathbb{R}_{\geq 1}$ we have*

$$\sum_{n \leq x} \mu(n)^2 \tau_r(n) \leq C_1^r x (\log x)^{r-1}.$$

*Proof.* Letting $f(n) = \mu(n)^2 \tau_r(n)$ one easily sees that

$$f(p^i) = \begin{cases} 1 & i = 0 \\ r & i = 1 \\ 0 & i > 1 \end{cases}.$$

Thus we infer

$$\sum_{p \le y} f(p) \log p = r \sum_{p \le y} \log p \le r\pi(y) \log y$$

and Lemma 2.14 reveals that

$$\sum_{p \le y} f(p) \log p \le 2ry.$$

Furthermore

$$\sum_{p} \sum_{i>1} \frac{f(p^i)}{p^i} \log p^i = 0.$$

By [3, Th. 01] with $A = 2r$ and $B = 0$, we find

$$\sum_{n \le x} f(n) \le (2r+1) \frac{x}{\log x} \sum_{n \le x} \frac{f(n)}{n}$$

Note that

$$\sum_{n \le x} \frac{f(n)}{n} \le \prod_{p \le x} \left(1 + \tfrac{r}{p}\right) \le \left( \prod_{p \le x} \left(1 + \tfrac{1}{p}\right) \right)^r,$$

where the second inequality follows from the binomial formula. Note that

$$\prod_{p \le x} \left(1 + \tfrac{1}{p}\right) = \prod_{p \le x} \left(1 - \tfrac{1}{p}\right)^{-1} \prod_{p \le x} \left(1 - \tfrac{1}{p^2}\right).$$

Mertens' third theorem gives us that

$$\prod_{p \le x} \left(1 - \tfrac{1}{p}\right)^{-1} = O(\log x).$$

Furthermore by Lemma 2.6 it follows that

$$\lim_{x \to \infty} \prod_{p \le x} \left(1 - \frac{1}{p^2}\right) = \zeta(2) = \frac{6}{\pi^2}$$

The equality $\zeta(2) = \frac{6}{\pi^2}$ is a very well-known theorem by Euler. Hence there exists $C_1 > 1$ such that

$$\prod_{p \le x} \left(1 + \frac{1}{p}\right) \le \frac{C_1}{3} \log x.$$

We now have

$$\sum_{n\leq x} f(n) \leq C_1^r x(\log x)^{r-1},$$

since $(2r+1)\left(\frac{C_1}{3}\right)^r \leq C_1^r$ for every $r \geq 2$. $\square$

**Lemma 2.20.** *There exists $C_2 > 1$ such that for all $r \in \mathbb{N}_{\geq 2}$ and for all $x \in \mathbb{R}_{\geq 1}$ we have*

$$\sum_{d>x} \frac{\mu(d)^2\tau_r(d)}{d^2} \leq \frac{C_2^r(2(r-1)+\log x)^{r-1}}{x}.$$

*Proof.* We partition the interval $(x,\infty)$ as follows

$$(x,\infty) = \bigcup_{i=0}^{\infty}(\mathrm{e}^i x, \mathrm{e}^{i+1}x],$$

therefore

$$\sum_{d>x} \frac{\mu(d)^2\tau_r(d)}{d^2} = \sum_{i=0}^{\infty} \sum_{\mathrm{e}^i x<d\leq \mathrm{e}^{i+1}x} \frac{\mu(d)^2\tau_r(d)}{d^2}.$$

By Lemma 2.19 there exists a $C_1 > 1$ such that for all $r \in \mathbb{N}_{\geq 2}$ and for all $x \in \mathbb{R}_{\geq 1}$ we have

$$\sum_{i=0}^{\infty} \sum_{\mathrm{e}^i x<d\leq \mathrm{e}^{i+1}x} \frac{\mu(d)^2\tau_r(d)}{d^2} \quad\leq\quad \frac{C_1^r}{x}\sum_{i=0}^{\infty}\frac{1}{\mathrm{e}^i}(i+1+\log x)^{r-1}.$$

Let $z := 2(r-1)+\log x$. Then

$$(i+1+\log x)^{r-1} = (i+1+z-2(r-1))^{r-1} \leq (i+z)^{r-1} = z^{r-1}\left(1+\frac{i}{z}\right)^{r-1}.$$

This implies

$$\sum_{i=0}^{\infty}\frac{1}{\mathrm{e}^i}(i+1+\log x)^{r-1} \leq z^{r-1}\sum_{i=0}^{\infty}\frac{1}{\mathrm{e}^i}\left(1+\frac{i}{z}\right)^{r-1}.$$

Note that $z \geq 2(r-1)$, equivalently $\left(1+\frac{i}{z}\right)^{r-1} \leq \left(1+\frac{i}{2(r-1)}\right)^{r-1}$. Furthermore

$$1+\frac{i}{2(r-1)} \leq \mathrm{e}^{\frac{i}{2(r-1)}},$$

hence

$$\frac{1}{\mathrm{e}^i}\left(1+\frac{i}{2(r-1)}\right)^{r-1} \leq \frac{1}{\mathrm{e}^i}\left(\mathrm{e}^{\frac{i}{2(r-1)}}\right)^{r-1} = \mathrm{e}^{-i/2}.$$

There exists $C_2 > 1$ such that

$$\frac{C_1^r}{x} z^{r-1} \sum_{i=0}^{\infty} \mathrm{e}^{-i/2} \leq \frac{C_2^r}{x} z^{r-1},$$

since the series converges, this concludes the proof. $\square$

Erdős proved the following lemma in 1951 [1]. Its meaning is that big gaps between consecutive square-free numbers are rare.

**Lemma 2.21** (Erdős)**.** *Let $t, x \in \mathbb{R}_{>0}$. Then*

$$\#\{s_{n+1} \leq x : s_{n+1} - s_n > t\} \ll \frac{x}{t^2(\log t)^2},$$

*where the implied constant is absolute.*

*Proof.* Let $s_{n+1} - s_n = r > t$ and $A = \left\{s_n < m < s_{n+1} \ : \ \exists p > \frac{t \log t}{100} \text{ such that } p^2 \mid m\right\}$. We want to show that $|A| \geq \frac{r}{16}$ for every $r > 1$.
Assume $t < 16$, then we have $\frac{t \log t}{100} < 1$. This implies that $|A| = r - 1$, since every integer $m \in \mathbb{N} \cap (s_n, s_{n+1})$ is divisible by the square of a prime. Note that $r - 1 > \frac{r}{16}$ for $r > 1$, which implies $|A| \geq \frac{r}{16}$. Assume that $t \geq 16$, then it follows that $r > 16$. We first look at the number of integers between $s_n$ and $s_{n+1}$ that are divisible by the square of a prime less than $\frac{t \log t}{100}$. Let

$$B = \left\{s_n < m < s_{n+1} \mid \exists p \leq \tfrac{t \log t}{100} : p^2 \mid m\right\}.$$

Note that

$$|B| \leq \sum_{p \leq \frac{t \log t}{100}} \left[\frac{r}{p^2}\right] + 1 \leq r \sum_p \frac{1}{p^2} + \pi\left(\frac{t \log t}{100}\right).$$

Calculations show that $\sum_p \frac{1}{p^2} < 0.46$. Hence

$$r \sum_p \frac{1}{p^2} \leq \frac{1}{2} r.$$

We want to show that $\pi\left(\frac{t \log t}{100}\right)$ is smaller than $\frac{3}{8}t$ for $t \geq 16$. For $t \leq 50$ we have $\frac{t \log t}{100} < 2$, hence $\pi\left(\frac{t \log t}{100}\right) = 0 \leq \frac{3}{8}t$.
Assume $t > 50$. By Lemma 2.14 we obtain $\pi\left(\frac{t \log t}{100}\right) \leq \frac{t \log t}{50 \log(t \log t / 100)}$. Consider the function

$$f(t) = \frac{\frac{t \log t}{50 \log(t \log t/100)}}{\frac{3}{8} t}.$$

We want to show that $f(t) \leq 1$. Note that

$$f(t) = \frac{\frac{8}{3}\log t}{50(\log t + \log\log t - \log 100)}.$$

Clearly, we have $f(t) \leq 1$ if and only if $50\log t + 50\log\log t - 50\log 100 \geq \frac{8}{3}\log t$. We infer that $f(1) \leq 1$ if and only if $t^{\frac{71}{75}}\log t \geq 100$. The latter holds for $t = 50$ and hence for $t > 50$, since the function $t \mapsto t^{\frac{71}{75}}\log t$ is increasing for $t > 50$.
Hence $|B| \leq \frac{7}{8}r$. Furthermore $|A| + |B| \geq r - 1$, since every integer $m \in \mathbb{N} \cap (s_n, s_{n+1})$ is divisible by the square of a prime. This implies $|A| \geq r - 1 - \frac{7}{8}r = \frac{r}{8} - 1 > \frac{r}{16}$, since we have assumed that $\frac{r}{16} > 1$.
It follows that

$$\#\left\{1 < m < x \mid \exists p > \tfrac{t\log t}{100} \text{ such that } p^2 \mid m\right\} \geq \frac{t}{16}\#\{s_{n+1} \leq x : s_{n+1} - s_n \geq t\}.$$

On the other hand

$$\#\left\{1 < m < x \mid \exists p > \tfrac{t\log t}{100} \text{ such that } p^2 \mid m\right\} \leq \sum_{p > \frac{t\log t}{100}} \left[\frac{x}{p^2}\right] \leq x \sum_{p > \frac{t\log t}{100}} \frac{1}{p^2}.$$

Using Lemma 2.14 and the fact that the function $1/p^2$ is decreasing we obtain

$$x \sum_{p > \frac{t\log t}{100}} \frac{1}{p^2} \ll \frac{x}{\frac{t\log t}{100}} \frac{1}{\log(\frac{t\log t}{100})}.$$

Observe that $\log\left(\frac{t\log t}{100}\right) = \log t + \log\log t - 100$ and furthermore that $\log\log t = O(\log t)$ and also $\log 100 = O(\log t)$. This implies

$$\#\left\{1 < m < x \mid \exists p > \tfrac{t\log t}{100} : p^2 \mid m\right\} \ll \frac{x}{t(\log t)^2}$$

and concludes the proof. □

# 3 Proof of Theorem 1.2

We would like to write each term of the summation in Theorem 1.2 as one Möbius function. For this to happen we need all $qn + a_i$ to be coprime in pairs. For a lot of choices of $a_1, ..., a_k$ this does not happen. We introduce the two following functions and study their properties.

**Definition 3.1.** Let $W \in \mathbb{N}$ and define $\mu_W$ and $\widetilde{\mu}$ by

$$\mu_W(n) := \mu\left(\prod_{p|W} p^{\nu_p(n)}\right)$$

and

$$\widetilde{\mu}(n) := \mu\left(\prod_{p\nmid W} p^{\nu_p(n)}\right).$$

**Lemma 3.2.** *Let $n, W \in \mathbb{N}$. Then $\mu(n) = \mu_W(n)\widetilde{\mu}(n)$.*

*Proof.* Obvious. □

**Lemma 3.3.** *Let $n, m \in \mathbb{N}$ be such that $n \equiv m \bmod W$ and assume that $W$ is an integer square. Then $\mu_W(n) = \mu_W(m)$.*

*Proof.* Note that $p \mid W$ if and only if $p^2|W$. Let $\mu_W(n) = 0$. Then there exists a $p \mid W$ such that $n \equiv 0 \bmod p^2$. It follows that $m \equiv 0 \mod p^2$, since $n \equiv m \bmod p^2$. Hence

$$\mu_W(n) = 0 \text{ if and only if } \mu_W(m) = 0,$$

since the relation congruence modulo $W$ is symmetric. Let $p \mid W$. Then the congruence $n \equiv m \bmod p$ implies that $p \mid n$ if and only if $p \mid m$. This implies

$$\#\{p \text{ prime} : p \mid W, p \mid n\} = \#\{p \text{ prime} : p \mid W, p \mid m\}.$$

Let $p$ be a prime such that $p \mid W$ and $p \mid n$, but $p^2 \nmid n$. The congruence $n \equiv m \bmod p^2$ implies $p^2 \nmid m$. Hence

$$\left\{p \text{ prime} : p \mid W \text{ and } p \mid n \text{ implies } p^2 \nmid n\right\} = \left\{p \text{ prime} : p \mid W \text{ and } p \mid m \text{ implies } p^2 \nmid m\right\}.$$

This concludes the proof. □

**Lemma 3.4.** *Let $n, W \in \mathbb{N}$. Then*

$$\widetilde{\mu}(n)^2 = \sum_{\substack{d^2|n \\ (d,W)=1}} \mu(d).$$

*Proof.* Let $n = p_1^{\alpha_1} \cdots p_t^{\alpha_t} p_{t+1}^{\alpha_{t+1}} \cdots p_s^{\alpha_s}$ where $p_i \mid W$ for all $1 \le i \le t$ and $p_j \nmid W$ for all $t < j \le s$. Define $n_1 = p_1^{\alpha_1} \cdots p_t^{\alpha_t}$ and $n_2 = p_{t+1}^{\alpha_{t+1}} \cdots p_s^{\alpha_s}$. Clearly $(n_1, n_2) = 1$ and $n_1 n_2 = n$. Furthermore $\widetilde{\mu}(n)^2 = \mu(n_1)^2$ and by Lemma 2.9 we find $\mu(n_1)^2 = \sum_{d^2|n_1} \mu(d)$. Note that

$$\{d \in \mathbb{N} : d^2|n_1\} = \{d \in \mathbb{N} : d^2|n, (d, W) = 1\}.$$

Hence it follows that

$$\widetilde{\mu}(n)^2 = \sum_{\substack{d^2 | n \\ (d,W)=1}} \mu(d).$$

□

For the rest of this thesis we will have

$$W := q^2 \prod_{p^2 \le a_k - a_1} p^2.$$

**Lemma 3.5.** *Let $q \neq 0, n, a_1, ..., a_k$ be integers with $0 \le a_1 < a_2 < \cdots < a_k$. If $d_i^2 \mid qn + a_i$, $(d_i, W) = 1$, $d_j^2 \mid qn + a_j$ and $(d_j, W) = 1$, then $(d_i, d_j) = 1$ for all $j \neq i$.*

*Proof.* Assume that $d_i$ and $d_j$ are not coprime. Then there exists a prime $p$ such that $p \mid d_i$ and $p \mid d_j$. So $p^2 \mid d_i^2$ and $p^2 \mid d_j^2$. Therefore $p^2 \mid qn + a_i$ and $p^2 \mid qn + a_j$. We infer that $p^2 \mid a_j - a_i$, since $p^2$ divides the difference of $(qn + a_j)$ and $(qn + a_i)$. This implies $p^2 \le |a_j - a_i|$. Furthermore $(d_i, W) = 1$ and $(d_j, W) = 1$ imply $p^2 > a_k - a_1$. Hence $(d_i, d_j) = 1$, since $a_k - a_1 \ge |a_j - a_i|$ for all $j \neq i$. □

Define

$$\widetilde{T}_m(z) = \sum_{\substack{1 \le n \le z \\ n \equiv m \bmod W}} \prod_{i=1}^{k} \widetilde{\mu}(qn + a_i)^2.$$

By Lemma 3.2 and Lemma 3.3 it follows that the sum in Theorem 1.2 equals

$$\sum_{1 \le m \le W} \prod_{i=1}^{k} \mu_W(qm + a_i)^2 \widetilde{T}_m(z).$$

Define

$$M = \left\{ 1 \le m \le W : \prod_{i=1}^{k} \mu_W(qm + a_i)^2 = 1 \right\}.$$

Hence the sum in Theorem 1.2 equals

$$\sum_{m \in M} \widetilde{T}_m(z), \tag{3.1}$$

since $\prod_{i=1}^{k} \mu_W(qm + a_i)^2$ is either zero or one.

**Proposition 3.6.** *Let $q \neq 0, a_1, ..., a_k$ be integers with $0 \leq a_1 < a_2 < \cdots < a_k$ and let $m$ be an integer with $1 \leq m \leq W$. Then there exists a constant $\delta := \delta(q, a_1, ..., a_k, W, m)$ such that for all $z \geq a_k$,*

$$\widetilde{T}_m(z) = \delta z + O\left( z^{2/3}(2k + \log z)^{O(k)}(2qz)^{O\left(\frac{k}{\log\log z}\right)} \right),$$

*where the implied constants depend at most on $q$.*

**Remark 3.7.** Before we start the proof, we describe its main idea.
First we use Lemma 3.4 which provides the equality

$$\widetilde{T}_m(z) = \sum_{\substack{1 \leq n \leq z \\ n \equiv m \bmod W}} \prod_{i=1}^{k} \left( \sum_{\substack{d_i^2 | qn+a_i \\ (d_i, W)=1}} \mu(d_i) \right). \tag{3.2}$$

Next we partition the summation domain into specific parts as follows.
Let $A_1, ..., A_k$ be positive reals to be chosen later. Assume that $d_1 \cdots d_k > A_1$ for some integers $d_1, ..., d_k$. By demanding $d_1 \leq \cdots \leq d_k$ we get an extra factor $k^k$ in the error term, since the number of ways to order $d_1, ..., d_k$ is equal to $k! \leq k^k$. For all $2 \leq m \leq k$ define

$$D_m := \left\{ (d_1, ..., d_k) \in \mathbb{N}^k : \begin{array}{ll} d_1 \leq \cdots \leq d_k, & (d_i, W) = 1 \text{ for all } i \\ d_m \cdots d_k \leq A_m, & d_{m-1} \cdots d_k > A_{m-1} \end{array} \right\}.$$

We have $z \geq a_k$, therefore $d_i^2 \leq qn + a_i \leq qz + a_i \leq 2qz$ for all $i$, hence $d_k \leq \sqrt{2qz}$. Take $A_k := \sqrt{2qz}$.
Let $d_1 \cdots d_k > A_1$. If $d_2 \cdots d_k \leq A_2$, then $\mathbf{d} \in D_2$. Let $d_2 \cdots d_k > A_2$. If $d_3 \cdots d_k \leq A_3$, then $\mathbf{d} \in D_3$. Let $d_3 \cdots d_k > A_3$. Continue this way until $d_{k-1}d_k > A_{k-1}$. Since $d_k \leq \sqrt{2qz} = A_k$, we obtain $\mathbf{d} \in D_k$. This proves that whenever $d_1 \cdots d_k > A_1$ then there exists an integer $m$ such that $\mathbf{d} \in D_m$.

*Proof of Proposition 3.6.* We begin by studying the contribution of $d_i$ in (3.2) with $d_1 \cdots d_k \leq A_1$. This contribution will provide our main term. It is equal to

$$\widetilde{B}_m(z) := \sum_{\substack{1 \leq n \leq z \\ n \equiv m \bmod W}} \prod_{i=1}^{k} \sum_{\substack{d_i^2 | qn+a_i \\ (d_i, W)=1 \\ d_1 \cdots d_k \leq A_1}} \mu(d_i) = \sum_{\substack{d_1, ..., d_k \in \mathbb{N} \\ (d_i, W)=1, \\ d_1 \cdots d_k \leq A_1}} \prod_{i=1}^{k} \mu(d_i) \sum_{\substack{1 \leq n \leq z \\ qn+a_i \equiv 0 \bmod d_i^2 \\ n \equiv m \bmod W}} 1.$$

By Lemma 3.5 we find $\prod_{i=1}^{k} \mu(d_i) = \mu(d_1 \cdots d_k)$. Note that $(d_i^2, q) = 1$, since $(d_i, W) = 1$, therefore $q$ has a multiplicative inverse modulo $d_i^2$. Hence $qn+a_i \equiv 0 \bmod d_i^2$ is equivalent to $n \equiv -\frac{a_i}{q} \mod d_i^2$. By the Chinese remainder theorem there exists an integer $n_0$ such that $n \equiv n_0 \mod W d_1^2 \cdots d_k^2$. Thus, we get

$$\widetilde{B}_m(z) = \sum_{\substack{d_1,\dots,d_k \in \mathbb{N} \\ (d_i,W)=1, \\ d_1\cdots d_k \le A_1}} \mu(d_1\cdots d_k) \sum_{\substack{1\le n\le z \\ n\equiv n_0 \mod W d_1^2\cdots d_k^2}} 1 \tag{3.3}$$

$$= \sum_{\substack{d_1,\dots,d_k \in \mathbb{N} \\ (d_i,W)=1, \\ d_1\cdots d_k \le A_1}} \mu(d_1\cdots d_k) \left(\frac{z}{W(d_1\cdots d_k)^2} + O(1)\right). \tag{3.4}$$

Let $d = d_1 \cdots d_k$. Clearly $(d, W) = 1$, since $(d_i, W) = 1$ for all $i$. Recalling Definition 2.15 we have

$$\sum_{\substack{d_1,\dots,d_k \in \mathbb{N} \\ (d_i,W)=1, \\ d_1\cdots d_k \le A_1}} \mu(d_1\cdots d_k)\left(\frac{z}{W(d_1\cdots d_k)^2} + O(1)\right) = \sum_{\substack{d \in \mathbb{N} \\ (d,W)=1, \\ d\le A_1}} \mu(d)\left(\frac{z}{Wd^2} + O(1)\right)\tau_k(d)$$

$$= \frac{z}{W} \sum_{\substack{d \in \mathbb{N} \\ (d,W)=1, \\ d\le A_1}} \frac{\mu(d)\tau_k(d)}{d^2} + O\left(\sum_{\substack{d \in \mathbb{N} \\ (d,W)=1, \\ d\le A_1}} \mu(d)^2\tau_k(d)\right).$$

Completing the series gives

$$\sum_{\substack{d \in \mathbb{N} \\ (d,W)=1, \\ d\le A_1}} \frac{\mu(d)\tau_k(d)}{d^2} = \sum_{\substack{d=1 \\ (d,W)=1}}^{\infty} \frac{\mu(d)\tau_k(d)}{d^2} + O\left(\sum_{\substack{d \in \mathbb{N} \\ (d,W)=1, \\ d> A_1}} \frac{\mu(d)^2\tau_k(d)}{d^2}\right).$$

Since $\mu(d)\tau_k(d)$ is a multiplicative function, we obtain by Lemma 2.3 that

$$\sum_{\substack{d=1 \\ (d,W)=1}}^{\infty} \frac{\mu(d)\tau_k(d)}{d^2} = \prod_{p\nmid W}\left(1 - \frac{k}{p^2}\right).$$

By Lemma 2.20 we have

$$\sum_{\substack{d \in \mathbb{N} \\ (d,W)=1, \\ d> A_1}} \frac{\mu(d)^2\tau_k(d)}{d^2} \le \frac{C_2^k(2(k-1)+\log A_1)^{k-1}}{A_1}$$

for some $C_2 > 1$. Furthermore, by Lemma 2.19 we obtain

$$\sum_{\substack{d\in\mathbb{N}\\(d,W)=1,\\d\leq A_1}} \mu(d)^2\tau_k(d) \leq C_1^k A_1(\log A_1)^{k-1}$$

for some $C_1 > 1$. Define

$$\delta := \delta(q, a_1, ..., a_k, W, m) = \frac{1}{W} \sum_{\substack{d=1\\(d,W)=1}}^{\infty} \frac{\mu(d)\tau_k(d)}{d^2}.$$

It follows that $\widetilde{B}_m(z)$ equals

$$\delta z + O\left(\frac{z}{W}\frac{C_2^k(2(k-1)+\log A_1)^{k-1}}{A_1}\right) + O\left(C_1^k A_1(\log A_1)^{k-1}\right).$$

Define $C := \max\{C_1, C_2\}$. The choice $A_1 = \sqrt{z}$ makes the error terms to be

$$O\left(\sqrt{z}C^k(2k+\log z)^{k-1}\right).$$

Let $\mathbf{d} = (d_1, ..., d_k) \in \mathbb{N}^k$ with $d_1 \cdots d_k > A_1$ and let $m$ be an integer with $2 \leq m \leq k$ such that $\mathbf{d} \in D_m$. We have shown in Remark 3.7 that such $m$ exists for every $(d_1, ..., d_k) \in \mathbb{N}^k$ with $d_1 \cdots d_k > A_1$.
Then $d_m \cdots d_k \leq A_m$ and $d_{m-1} \cdots d_k > A_{m-1}$. Furthermore $\frac{d_{m-1}\cdots d_k}{d_j} \leq A_m$ for all $j$ with $m-1 \leq j \leq k$, since $d_{m-1} \leq d_j$ for all such $j$. Multiplying all these inequalities implies $\frac{(d_{m-1}\cdots d_k)^{k-m+2}}{d_{m-1}\cdots d_k} \leq A_m^{k-m+2}$. Hence $d_{m-1}\cdots d_k \leq A_m^{\frac{k-m+2}{k-m+1}}$. Define $\alpha_m := \frac{k-m+2}{k-m+1}$. The contribution of $\mathbf{d} \in D_m$ is given by

$$\left|\sum_{1\leq n\leq z}\prod_{i=1}^{k}\sum_{\substack{d_i^2|qn+a_i\\(d_i,W)=1\\A_{m-1}<d_{m-1}\cdots d_k\leq A_m^{\alpha_m}}}\mu(d_i)\right| \leq \sum_{\substack{d_{m-1},...,d_k\in\mathbb{N}\\(d_i,W)=1,\\A_{m-1}<d_{m-1}\cdots d_k\leq A_m^{\alpha_m}}}|\mu(d_{m-1}\cdots d_k)| \sum_{\substack{1\leq n\leq z\\d_{m-1}^2|qn+a_{m-1}\\\cdots\\d_k^2|qn+a_k}}\prod_{i=1}^{m-2}\sum_{\substack{d_i\in\mathbb{N}\\d_i^2|qn+a_i}}1.$$

Clearly,

$$\sum_{\substack{d_i\in\mathbb{N}\\d_i^2|qn+a_i}} 1 \leq \tau(qn+a_i).$$

Note that $qn + a_i \leq 2qz$ for every $i$. By Lemma 2.18 it follows that

$$\tau(qn+a_i) \leq (2qz)^{O\left(\frac{1}{\log\log 6qz}\right)}.$$

We infer that

$$\prod_{i=1}^{m-2} \sum_{\substack{d_i \in \mathbb{N} \\ d_i^2 | qn+a_i}} 1 \leq (2qz)^{O\left(\frac{k}{\log\log z}\right)},$$

since $m-2 \leq k$ for every $m$. Following the same steps as in (3.3) and (3.4) we obtain

$$\sum_{\substack{d_{m-1},\dots,d_k \in \mathbb{N} \\ (d_i,W)=1, \\ A_{m-1} < d_{m-1}\cdots d_k \leq A_m^{\alpha_m}}} |\mu(d_{m-1}\cdots d_k)| \sum_{\substack{1 \leq n \leq z \\ d_{m-1}^2 | qn+a_{m-1} \\ \cdots \\ d_k^2 | qn+a_k}} 1 \ll \sum_{\substack{d_1,\dots,d_k \in \mathbb{N} \\ (d_i,W)=1, \\ A_{m-1} < d_{m-1}\cdots d_k \leq A_m^{\alpha_m}}} |\mu(d_{m-1}\cdots d_k)| \left( \frac{z}{d_{m-1}^2 \cdots d_k^2} + 1 \right).$$

This equals

$$z \sum_{\substack{d \in \mathbb{N} \\ (d,W)=1, \\ A_{m-1} < d \leq A_m^{\alpha_m}}} \frac{\mu(d)^2 \tau_{k-m+2}(d)}{d^2} + \sum_{\substack{d \in \mathbb{N} \\ (d,W)=1, \\ A_{m-1} < d \leq A_m^{\alpha_m}}} \mu(d)^2 \tau_{k-m+2}(d).$$

By Lemma 2.19 and Lemma 2.20 this is less than or equal to

$$z \frac{C_2^k (2k + \log A_{m-1})^k}{A_{m-1}} + C_1^k A_m^{\alpha_m} (\log(A_m^{\alpha_m}))^k.$$

We take

$$A_m := z^{\frac{1}{1+\alpha_m}} = z^{\frac{k-m+1}{2k-2m+3}} \text{ for } m \leq k-1. \tag{3.5}$$

We note that $\alpha_m \geq 1$, therefore $A_m \leq z^{1/2}$. Observe that

$$\frac{z}{A_m} = z^{1-\frac{k-m+1}{2k-2m+3}} = z^{\frac{k-m+2}{2k-2m+3}} \text{ for every } m \text{ with } 2 \leq m < k.$$

Furthermore

$$\frac{k-m+2}{2k-2m+3} \leq \frac{2}{3} \text{ for every } m \text{ with } 2 \leq m \leq k.$$

Using our previous choice $A_k = \sqrt{2qz}$ we get for each $2 \leq m \leq k$ and every $\mathbf{d} \in D_m$ that

$$\sum_{1 \leq n \leq z} \prod_{i=1}^{k} \sum_{\substack{d_i^2 | qn+a_i \\ (d_i,W)=1 \\ A_{m-1} < d_{m-1}\cdots d_k \leq A_m^{\alpha_m}}} \mu(d_i) \ll z^{2/3} C^k k^k (2k + \log z)^k (2qz)^{O\left(\frac{k}{\log\log z}\right)}.$$

Finally using $C^k k^k \leq 2k^{O(k)}$, we obtain

$$\widetilde{T}_m(z) = \delta z + O\left( z^{2/3} (2k + \log z)^{O(k)} (2qz)^{O\left(\frac{k}{\log\log z}\right)} \right).$$

□

*Proof of Theorem 1.2.* By Proposition 3.6 there exists $\delta_m := \delta(q, a_1, ..., a_k)$ such that

$$\widetilde{T}_m(z) = \delta_m z + O\left( z^{2/3}(2k + \log z)^{O(k)}(2qz)^{O\left(\frac{k}{\log\log z}\right)} \right),$$

where the implied constant depends at most on $q$. Define

$$\gamma := \gamma(q, a_1, .., a_k) = \sum_{m \in M} \delta_m.$$

By (3.1) it follows that the sum in Theorem 1.2 equals

$$\gamma z + O\left( W z^{2/3}(2k + \log z)^{O(k)}(2qz)^{O\left(\frac{k}{\log\log z}\right)} \right).$$

To obtain an upper bound for $W$, observe that

$$\log W \quad \leq \quad 2\log q + 2 \sum_{p < \sqrt{a_k - a_1}} \log p,$$

By Lemma 2.14 we find

$$\sum_{p < \sqrt{a_k - a_1}} \log p \leq 2 \frac{\sqrt{a_k}}{\log\left(\sqrt{a_k}\right)} \log\left(\sqrt{a_k}\right).$$

We infer that $W \ll q^2 e^{2\sqrt{a_k}}$.
From Theorem 2.13 it follows that

$$\sum_{1 \leq n \leq x} \prod_{i=1}^{r} \mu(qn + a_i)^2 = cx + O\left(x^{3/4}\right),$$

where $c = \prod_p (1 - \omega(p)p^{-2})$ and

$$\omega(p) := \#\left\{1 \leq n \leq p^2 : \exists\ 1 \leq i \leq r \text{ such that } qn + a_i \equiv 0 \bmod p^2\right\}.$$

This concludes the proof. □

# 4 Proof of Theorem 1.1

Let $x$ be a positive real and $a, b, q$ positive integers and define

$$S(x; q, (a, b)) := \#\left\{s_n \leq x : s_n \equiv a \bmod q, s_{n+1} \equiv b \bmod q\right\}. \tag{4.1}$$

From now on $a, b, q$ are fixed integers such that $0 \leq a < q$ and $a < b \leq q + a$. Observe that

$$S(x; q, (a,b)) = \sum_{h=1}^{\infty} S_h(x; q, (a,b)), \tag{4.2}$$

where

$$S_h(x; q, (a,b)) = \# \left\{ n \leq \tfrac{x-a}{q} : \begin{array}{l} qn + a \text{ is square-free} \\ \text{and the next square-free is } q(n+h-1)+b \end{array} \right\}. \tag{4.3}$$

The main idea of the proof of Theorem 1.1 is to partition the summation (4.2) into the two ranges

$$1 \leq h \leq H \text{ and } h > H,$$

where

$$H = \alpha \log\log x$$

for some positive $\alpha$ that will be optimized later.

*Proof of Theorem 1.1.* By the inclusion-exclusion principle

$$S_h(x; q; (a,b)) = \sum_{c=0}^{l_h} (-1)^c \sum_{a < i_1 < \cdots < i_c < q(h-1)+b} T(\tfrac{x-a}{q}; i_1, ..., i_c, h), \tag{4.4}$$

where $l_h = b + q(h-1) - a - 1$ and

$$T(z; i_1, ..., i_c, h) = \sum_{n \leq z} \mu(qn+a)^2 \left( \prod_{j=1}^{c} \mu(qn+i_j)^2 \right) \mu(q(n+h-1)+b)^2.$$

Let $a_1 = a$, $a_2 = i_1$,..., $a_{k-1} = i_c$ and $a_k = b + q(h-1)$. By Theorem 1.2 there exists $\gamma$ such that

$$T(\tfrac{x-a}{q}; i_1, ..., i_c, h) = \gamma \frac{x-a}{q} + O\left( x^{2/3} e^{2\sqrt{b+q(h-1)}} (2c + \log x)^{O(c)} (2(x-a))^{O\left(\frac{c}{\log\log x}\right)} \right).$$

Define

$$\sigma_h := \sigma(q, a, b, W, h) = -\frac{a}{q} \sum_{c=0}^{l_h} (-1)^c \sum_{a < i_1 < \cdots < i_c < q(h-1)+b} \gamma.$$

Observe that $\left| \sum_{c=0}^{l_h} (-1)^c \sum_{a < i_1 < \cdots < i_c < q(h-1)+b} 1 \right|$ is equal to the number of possible sequences of length $l_h$ where each element is either zero or one. We obtain that

$$\left| \sum_{c=0}^{l_h} (-1)^c \sum_{a < i_1 < \cdots < i_c < q(h-1)+b} 1 \right| = 2^{b+q(h-1)-a-1} \leq 2^{O(h)}.$$

The last inequality holds, since $a, b$ and $q$ are fixed. Note that $c \leq b + q(h-1) - a - 1$ for every $c$. This implies $c \ll h$, since $a, b, q$ are fixed. Furthermore $(x - a) \leq x$ and $\frac{h}{\log\log x} \leq h$. It follows that

$$S_h(x; q, (a, b)) \quad = \quad \sigma_h \frac{x}{q} + O\left( x^{2/3} \mathrm{e}^{O\left(\sqrt{h}\right)} (2h + \log x)^{O(h)} x^{O\left(\frac{h}{\log\log x}\right)} \right).$$

Let $H$ be a function of $x$ to be chosen later. We have

$$S(x; q, (a, b)) = \sum_{1 \leq h \leq H} S_h(x; q, (a, b)) + \sum_{h > H} S_h(x; q, (a, b)).$$

Note that

$$S_h(x; q, (a, b)) \leq \#\{s_n \leq x : s_{n+1} - s_n = qh - q + b - a\},$$

furthermore for any $H \geq 2q$ and for every $h > H$ we have

$$qh - q + b - a > qh - q > qH - q \geq \frac{H}{2}.$$

By Lemma 2.21 we get

$$\sum_{h > H} S_h(x; q, (a, b)) \ll \frac{x}{\left(\frac{H}{2}\right)^2 \left(\log \frac{H}{2}\right)^2} \ll \frac{x}{H^2}.$$

Define $\theta(q, a, b, H) := \frac{1}{q} \sum_{1 \leq h \leq H} \sigma(q, a, b, W, h)$. It follows that

$$\sum_{1 \leq h \leq H} S_h(x; q, (a, b)) \quad = \quad \theta(q, a, b, H)x + O\left( x^{2/3} H \mathrm{e}^{O\left(\sqrt{h}\right)} (2H + \log x)^{O(H)} x^{O\left(\frac{H}{\log\log x}\right)} \right).$$

Clearly, $H\mathrm{e}^{O\left(\sqrt{h}\right)} \leq (2H + \log z)^{O(h)}$. Note that the implied constant for $x^{O\left(\frac{H}{\log\log x}\right)}$ depends at most on $q$, say $c_q > 0$. Let $H := \alpha \log\log x$ where $c_q \alpha \in (0, \frac{1}{3})$. Define

$$L := (2H + \log x)^{O(H)}$$

and note that

$$\log L = O\left((\log\log x)^2\right).$$

Hence $\log L \leq \frac{1}{4} \log x$ for $x$ sufficiently large. It follows that $L \leq x^{1/4}$. This implies $x^{2/3} L x^{c_q \alpha} \leq x^{2/3 + 1/4 + c_q \alpha}$. The choice of $\alpha$ with $c_q \alpha \in (0, \frac{1}{12})$ provides us with

$$S(x; q, (a, b)) = \theta(q; (a, b), H)x + O\left( \frac{x}{(\log\log x)^2} \right).$$

Define

$$\ell(q;(a,b)) := \frac{1}{q}\sum_{h=1}^{\infty}\sigma(q,a,b,W,h) \text{ and } c_h(q,a,b) := \lim_{x\to\infty}\frac{S_h(x;q,(a,b))}{x}.$$

Reformulating the definition of $S_h(x;q,(a,b))$ gives

$$S_h(x;q,(a,b)) = \#\{s_n \le x : s_n \equiv a \bmod q \text{ and } s_{n+1} = s_n - a + b + q(h-1)\}.$$

Clearly this is less than or equal to

$$\#\{s_n \le x : s_{n+1} - s_n > b - a + q(h-1) - 1\}.$$

By Lemma 2.21 we obtain

$$\frac{S_h(x;q,(a,b))}{x} \ll \frac{1}{h^2}$$

for every $h > 1$. Observe that

$$\theta(q;(a,b),H) = \ell(q;(a,b)) + O\left(\sum_{h>H} c_h(q;(a,b))\right). \tag{4.5}$$

It follows that

$$\sum_{h>H} c_h(q;(a,b)) \ll \sum_{h>H}\frac{1}{h^2} \ll \frac{1}{H}.$$

Hence

$$S(x;q,(a,b)) = \ell(q;(a,b))x + O\left(\frac{x}{\log\log x}\right).$$

□

We now proceed to provide an explicit expression for $\ell(q;(a,b))$. Combining (4.2) and (4.4) we obtain

$$S(x;q,(a,b)) = \sum_{h=1}^{\infty}\sum_{c=0}^{l_h}(-1)^c \sum_{a<i_1<\cdots<i_c<q(h-1)+b} T(\tfrac{x-a}{q}; i_1, ..., i_c, h)$$

where $l_h = b + q(h-1) - a - 1$. Recall that

$$T(\tfrac{x-a}{q}; i_1, ..., i_c, h) = \sum_{n\le\frac{x-a}{q}} \mu(qn+a)^2\left(\prod_{j=1}^{c}\mu(qn+i_j)^2\right)\mu(q(n+h-1)+b)^2.$$

By Theorem 2.13, $T(\frac{x-a}{q}; i_1, ..., i_c, h)$ equals

$$\prod_p\left(1 - \omega_{i_1,\ldots,i_c}(p)p^{-2}\right)x + O(x^{3/4}),$$

where $\omega_{i_1,\dots,i_c}(p)$ is the number of residue classes $n$ modulo $p^2$ such that at least one of the quantities $qn+a,\ qn+i_1,\dots,qn+i_c\ ,\ qn+q(h-1)+n$ is divisible by $p^2$. Hence

$$\ell(q;(a,b)) = \sum_{h=1}^{\infty}\sum_{c=0}^{l_h}(-1)^c \sum_{a<i_1<\cdots<i_c<q(h-1)+b}\prod_p \left(1-\omega_{i_1,\dots,i_c}(p)p^{-2}\right). \tag{4.6}$$

# 5 Proof of Theorem 1.3

For $q \in \mathbb{N}$ we define

$$G_q := \left\{a \in \mathbb{Z}\cap[1,q] : (a,q) = \text{square-free}\right\}.$$

**Remark 5.1.** Note that $1 \in G_q$ for every $q$. Furthermore $G_q = [1,q]$ if $q$ is square-free.

**Proposition 5.2.** *Let $a, q \in \mathbb{Z}$. Then the following statements are equivalent.*

1. *$a \in G_q$.*
2. *The progression $a$ mod $q$ contains infinitely many square-free numbers.*
3. *The progression $a$ mod $q$ contains at least one square-free number.*

*Proof.* (**1** $\Rightarrow$ **2**) Assume $a \in G_q$. By Proposition 2.12 it is sufficient to prove that the following absolutely convergent product

$$\prod_{\substack{p\\(p^2,q)|a}}\left(1-\frac{(p^2,q)}{p^2}\right)$$

does not vanish. Indeed, if it vanishes then there exists a prime $p$ such that $(p^2,q)=p^2$ and $(p^2,q)|a$. There two conditions however imply that $p^2|(q,a)$, thus contradicting our assumption that $a \in G_q$.
(**2** $\Rightarrow$ **3**) Obvious.
(**3** $\Rightarrow$ **1**) Assume that there exists a square-free integer $s$ such that $s \equiv a$ mod $q$. Hence, there exists $m$ such that $s = mq + a$. Writing

$$s = (a,q)\left(\frac{q}{(a,q)}m + \frac{a}{(a,q)}\right)$$

we deduce that the factor $(a,q)$ must also be square-free. ☐

**Lemma 5.3.** *Let $a, b \in \mathbb{Z}$ and define*

$$N_p(a,b) := \#\left\{1 \le x \le p^2 : ax+b \equiv 0 \ mod\ p^2\right\}.$$

*If $p^2|(a,b)$ then $N_p(a,b) = p^2$ and in the opposite case we have $N_p(a,b) \le p$.*

*Proof.* If $p^2 \mid a$ and $p^2 \mid b$ then the congruence $ax + b \equiv 0 \bmod p^2$ holds for every integer $x$. Hence $N_p(a, b) = p^2$.

Assume that $p^2 \nmid a$ or $p^2 \nmid b$. If $p \nmid a$ then the equation $ax \equiv b \bmod p^2$ is equivalent to $x \equiv b/a \bmod p^2$, thus $N_p(a, b) = 1$. The two remaining cases are $p \| a$ and $p^2 \mid a$. Let $p \| a$. If $p \nmid b$ then by Lemma 2.10 we obtain $N_p = 0$. If $p \mid b$, then there exists $a_0$ and $b_0$ with $(p, a_0) = 1$ such that $a_0 x + b_0 \equiv 0 \bmod p^2$. This implies that there exists a unique solution mod $p$, since $(p, a_0) \mid b_0$. It follows that $N_p \leq p$. Let $p^2 \mid a$. Then $N_p(a, b)$ is either 0 or $p^2$, according to if $p^2 \nmid b$ or $p^2 \mid b$ respectively. The assumption $p^2 \nmid (a, b)$ allows us to rule our the second case, thus leaving us with the desired bound. □

**Corollary 5.4.** *Let $a_1, b_1, a_2, b_2 \in \mathbb{Z}$ and define*

$$M_p(a_1, b_1; a_2, b_2) := \# \left\{1 \leq x \leq p^2 : a_1 x + b_1 \equiv 0 \textit{ mod } p^2 \textit{ or } a_2 x + b_2 \equiv 0 \textit{ mod } p^2\right\}.$$

*Then $M_p(a_1, b_1; a_2, b_2) \leq 2p$ except if $p^2 \mid (a_1, b_1)$ or $p^2 \mid (a_2, b_2)$.*

*Proof.* Clearly, $M_p(a_1, b_1; a_2, b_2) \leq N_p(a_1, b_1) + N_p(a_2, b_2)$. The proof then follows directly from Lemma 5.3. □

**Lemma 5.5.** *Let $a_1, b_1, a_2, b_2 \in \mathbb{Z}$ and assume that $4 \nmid (a_1, b_1)$ and $4 \nmid (a_2, b_2)$. If we have $M_2(a_1, b_1; a_2, b_2) = 4$ then $2 \mid (a_1, b_1, a_2, b_2)$.*

*Proof.* Using Lemma 5.3 for $p = 2$ shows that we have

$$N_2(a_1, b_1) \leq 2 \text{ and } N_2(a_2, b_2) \leq 2,$$

which, in addition to

$$M_2(a_1, b_1; a_2, b_2) \leq N_2(a_1, b_1) + N_2(a_2, b_2),$$

shows that if $N_2(a_1, b_1) < 2$ or $N_2(a_2, b_2) < 2$ then we have a contradiction. Therefore we can safely deduce that

$$N_2(a_1, b_1) = 2 \text{ and } N_2(a_2, b_2) = 2.$$

This shows that $2|(a_1, a_2)$, since otherwise at least one of the two last quantities would be 1. Of course, if $2 \nmid b_1$ then $N_2(a_1, b_1) = 0$, and similarly, we must have $2 \mid b_2$. □

**Lemma 5.6.** *Let $a, q$ be integers such that $(a, q)$ is square-free. Then there exists a positive integer $k_0$ such that*

$$(k_0, (a, q)) = 1 \textit{ and } k_0 \equiv \frac{a}{(a, q)} \textit{ mod } \frac{q}{(a, q)}.$$

*Proof.* By the Chinese remainder theorem our lemma is equivalent to the following: for each prime $p$ that divides $\frac{q}{(a,q)}$ there exists $k_p \in \mathbb{Z}$ such that

$$\left(k_p, p^{\min\{\nu_p(a),\nu_p(q)\}}\right) = 1 \text{ and } k_p \equiv \frac{a}{(a,q)} \bmod p^{\nu_p(q)-\min\{\nu_p(a),\nu_p(q)\}}. \tag{5.1}$$

Clearly, this holds if $\min\{\nu_p(a), \nu_p(q)\} = 0$. If $\min\{\nu_p(a), \nu_p(q)\} \neq 0$, then the minimum must equal 1, since $(a, q)$ is square-free.

There are two cases to consider: $\nu_p(q) = 1$ and $\nu_p(q) \geq 2$. If $\nu_p(q) = 1$, then (5.1) is equivalent only to $p \nmid k_p$; such a $k_p$ can obviously always be found. If $\nu_p(q) \geq 2$, then our assumptions on $a, q$ guarantee that $\nu_p(a) = 1$. Hence (5.1) is equivalent to

$$p \nmid k_p \text{ and } k_p \equiv \frac{a}{(a,q)} \bmod p^{1+(\nu_p(q)-2)}.$$

The integer $k_p = \frac{a}{(a,q)}$ is admissible here, because its $p$-adic valuation equals

$$\nu_p(a) - \min\{\nu_p(a), \nu_p(q)\} = 0.$$

□

*Proof of Theorem 1.3.* One direction is obvious: if $\ell(q; (a, b)) > 0$ then the progressions $a \bmod q$ and $b \bmod q$ contain at least one square-free number.

For the other direction assume that the progressions $a \bmod q$ and $b \bmod q$ contain at least one square-free number, thus Proposition 5.2 provides us with

$$\mu^2((a, q)) = 1 = \mu^2((b, q)).$$

Let $r := b - a - 1$ and choose $r$ distinct primes $p_1, ..., p_r$ with the property that

$$2q < p_1 < p_2 < \ldots < p_r. \tag{5.2}$$

Recalling equation (4.1) allows us to observe the following simple inequality,

$$S(x; q, (a, b)) \geq \sum_{\substack{m \leq x, m \equiv a \bmod q \\ m+i \equiv 0 \bmod p_i^2 \ (1 \leq i \leq r)}} \mu(m)^2 \mu(m + b - a)^2.$$

The congruence conditions imposed on $m$ may imply that no such $m$ exist, in this case the sum above is taken to be zero. However, we shall now see that there are in fact plenty of $m$ satisfying all congruences. Let $m = k(a, q)$ for some $k \in \mathbb{N}$. Note that $m$ is square-free if and only if $\mu^2(k) = 1$ and $\gcd(k, \gcd(a, q)) = 1$, thus getting

$$S(x; q, (a, b)) \geq \sum_{\substack{k \leq \frac{x}{(a,q)}, (k,(a,q))=1 \\ k \equiv \frac{a}{(a,q)} \bmod \frac{q}{(a,q)} \\ k(a,q)+i \equiv 0 \bmod p_i^2 \ \forall i}} \mu(k)^2 \mu\left(k(a, q) + b - a\right)^2.$$

By the Chinese remainder theorem, the fact that each $p_i$ is coprime to $q$ and Lemma 5.6, there exists an integer $k_0$ that satisfies

$$(k_0,(a,q))=1,\ k_0\equiv \frac{a}{(a,q)} \bmod \frac{q}{(a,q)},\ k_0(a,q)+i\equiv 0 \bmod p_i^2\ \forall 1\le i\le r. \tag{5.3}$$

Letting

$$D:=\frac{q}{(a,q)}\prod_{i=1}^{r}p_i^2$$

and writing $k=k_0+nD$ we are therefore led to

$$S(x;q,(a,b))\ge T(x):=\sum_{1\le n\le \frac{x}{D(a,q)}-\frac{k_0}{D}}\mu(k_0+nD)^2\mu\left((k_0+nD)(a,q)+b-a\right)^2.$$

We may now invoke Theorem 2.13 to obtain

$$\lim_{x\to+\infty}\frac{T(x)}{x}=\prod_p\left(1-\frac{\omega(p)}{p^2}\right),$$

where $\omega(p)$ is given by

$$\#\left\{1\le n\le p^2: k_0+nD\equiv 0 \bmod p^2 \ \text{ or }\ nD(a,q)+k_0(a,q)+b-a\equiv 0 \mod p^2\right\}.$$

It will therefore be sufficient to show that the limit above is non-zero. If $p$ is coprime to $D\cdot D(a,q)$, then $\omega(p)\le 2$, therefore the last infinite product converges absolutely. Hence it is enough to show that each individual term is non-vanishing and we dedicate the rest of the proof solely to this task.

We shall now prove that for each prime $p$ one has

$$p^2\nmid (k_0,D) \text{ and } p^2\nmid\left(q\prod_{i=1}^{r}p_i^2, k_0(a,q)+b-a\right). \tag{5.4}$$

By Corollary 5.4 this shows that $\omega(p)\le 2p$, hence the required inequality $\omega(p)\ne p^2$ holds as long as $p\ne 2$; we shall prove that $\omega(2)\ne 4$ later.

We prove (5.4) via contradiction. First, assume that there exists $p$ such that

$$p^2\mid (k_0,D).$$

If $p=p_i$ for some $1\le i\le r$, then $p_i^2|k_0$, hence $k_0(a,q)+i\equiv 0 \bmod p_i^2$ implies $p_i^2$ divides $i$. This is a contradiction, since

$$i\le r=b-a-1<b-a\le q$$

and we have chosen all $p_i$ such that $p_i > 2q$. If $p \neq p_i$, then $p^2|D$, hence $p^2$ must divide $q/(a,q)$ and by (5.3) combined with $p^2|k_0$, we deduce that $p^2|a/(a,q)$. This is a contradiction, since the integers $\left(\frac{a}{(a,q)}, \frac{q}{(a,q)}\right)$ are coprime.

It remains to prove that for each prime $p$ one cannot have

$$p^2 | \left( q \prod_{i=1}^{r} p_i^2, k_0(a,q) + b - a \right).$$

If $p = p_i$ for some $1 \leq i \leq r$, then (5.3) provides us with

$$b - a - i \equiv 0 \bmod p_i^2.$$

This is clearly not possible because of (5.2) and $b - a - i < q$. If $p \neq p_i$, then $p^2$ divides $(q, k_0(a,q) + b - a)$. The congruence $k_0(a,q) \equiv a \bmod q$ is implied by (5.3) and it shows that $p^2|b$. This is a contradiction, since by assumption $b \in G_q$.

The last task in our proof is to prove that $\omega(2) \neq 4$. If we had $\omega(2) = 4$ then Lemma 5.5 and (5.4) would imply that

$$2 \mid \left( D, k_0, q \prod_{i=1}^{r} p_i^2, k_0(a,q) + b - a \right).$$

In light of (5.2) we know that $2 \nmid \prod_i p_i$, hence we may thus infer

$$2 \mid \left( \frac{q}{(a,q)}, k_0, b - a \right).$$

However, (5.3) supplies us with $k_0 \equiv \frac{a}{(a,q)} \bmod \frac{q}{(a,q)}$, thus $a/(a,q)$ must be even. This is a contradiction since $q/(a,q)$ and $a/(a,q)$ are always coprime but here we found that they are both even. This remark concludes our proof. □

# References


[1] P. Erdős. Some problems and results in elementary number theory. *Publ. Math. Debrecen*, 2:103–109, 1951.

[2] M. Filaseta. On the distribution of gaps between squarefree numbers. *Mathematika*, 40:88–101, 1993.

[3] R. R. Hall and G. Tenenbaum. *Divisors*, volume 90 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1988.

[4] C. Hooley. On the distribution of square-free numbers. *Can. J. Math.*, 25:1216–1223, 1973.

[5] M. N. Huxley. Moments of differences between square-free numbers. *Cambridge Univ. Press*, pages 187–204, 1997.

[6] R. J. Lemke Oliver and K. Soundararajan. Unexpected biases in the distribution of consecutive primes. *Proc. Natl. Acad. Sci. USA*, 113(31):4446–4454, 2016.

[7] L. Mirsky. Note on an asymptotic formula connected with $r$-free integers. *Quart. J. Math., Oxford Ser.*, 18:178–182, 1947.

[8] T. Reuss. Pairs of $k$-free numbers, consecutive square-full numbers. *arXiv preprint: 1212.3150.*

[9] J. B. Rosser and L. Schoenfeld. Approximate formulas for some functions of prime numbers. *Illinois J. Math.*, 6:64–94, 1962.

[10] M. Rubinstein and P. Sarnak. Chebyshev's bias. *Experiment. Math.*, 3(3):173–197, 1994.

[11] K. M. Tsang. The distribution of $r$-tuples of squarefree numbers. *Mathematatika*, 32(2):265–275, 1985.