On a variance associated with the distribution of real sequences in arithmetic progressions

by

PENGYONG DING (Jinan)

1. Introduction. First, define $A(x; q, a)$ as the sum of a real sequence $\{a_n\}$ over an arithmetic progression of indices:

$$A(x; q, a) = \sum_{n \leq x, \ n \equiv a \ (\text{mod} \ q)} a_n.$$  

Sometimes, there exist real functions $f(q, a)$ and $M(x)$, which approximately reflect the local and global properties of the sequence respectively, such that

$$A(x; q, a) \sim f(q, a)M(x).$$

To measure how far the sums over arithmetic progressions of indices are spread out from their approximation, we are interested in the variance

$$V(x, Q) = \sum_{q \leq Q} \sum_{a=1}^q |A(x; q, a) - f(q, a)M(x)|^2.$$ 

We always assume $1 \leq Q \leq x$.

The difficulty of the problem largely depends on the sequence $\{a_n\}$, especially on whether the sum and the second moment sum satisfy

$$\sum_{n \leq x} a_n \ll x, \quad \sum_{n \leq x} a_n^2 \ll x,$$

or equivalently, whether the mean value and the mean square are bounded or not.

The first example of the variance (1.3) was studied by Barban [Bar] with $a_n = \Lambda(n)$, $f(q, a) = 1/\phi(q)$, and $M(x) = x$, with the condition $(a, q) = 1$.

2020 Mathematics Subject Classification: Primary 11N64; Secondary 11P55.

Key words and phrases: Hardy–Littlewood circle method, asymptotic, distribution, sequence, variance, residue classes.

Received 3 October 2022; revised 3 April 2023.

Published online 30 October 2023.
added in the sum over $a$ in (1.3). This case was then studied by Davenport and Halberstam [DH], Gallagher [Ga], Montgomery [Mon] and Hooley [Ho1]. Montgomery established an asymptotic formula, and Hooley simplified the proof and gave a more refined result. He showed that if $a_n = \Lambda(n)$, then for any positive constant $A$,

$$(1.4) \quad V(x, x) = x^2 \log x + D_1 x^2 + O\left(\frac{x^2}{(\log x)^A}\right),$$

and also

$$(1.5) \quad V(x, Q) = Q x \log Q + D_2 Q x + O(Q^{5/4} x^{3/4}) + O\left(\frac{x^2}{(\log x)^A}\right),$$

where $D_1$ and $D_2$ are constants. And if one assumes the generalized Riemann hypothesis, then the last error terms in (1.4) and (1.5) can be replaced by $O(x^{11/7+\epsilon}).$

Later, Hooley [Ho2, Ho3, Ho7–Ho11] developed the subject further by studying far reaching generalizations. In particular, he established that an asymptotic formula can be obtained for a wide class of $\{a_n\}$ if one has

$$A(x; q, a) \sim x f(q, a)$$

with a reasonable error term, analogous to the Siegel–Walfisz theorem, and some control over the behavior of the mean square of $\{a_n\}$.

In the case of $\Lambda(n)$, further refinements occur in Friedlander and Goldston [FG], and in Vaughan and Goldston [GV], the latter paper showing that some advantage could be accrued by applying a version of the Hardy–Littlewood circle method. In detail, suppose that the generalized Riemann hypothesis holds and let

$$U(x, Q) = V(x, Q) - Q x \log Q - D_2 Q x.$$ 

Then for any $\epsilon > 0$,

$$U(x, Q) \ll Q^2 (x/Q)^{1/4+\epsilon} + x^{3/2} (\log 2x)^{5/2} (\log \log 3x)^2.$$ 

Having shown the usefulness of this method in that specific question, Vaughan [RV2, RV3] then considered the more general problem. Suppose that $M(x) = x$, $f(q, a)$ only depends on $q$ and $(q, a)$, and $\{a_n\}$ satisfies

$$\sum_{n \leq x} a_n^2 \ll x.$$ 

For convenience write $f(q, (q, a))$ for $f(q, a)$ and suppose further that

$$A(x; q, a) = x f(q, (q, a)) + O\left(\frac{x}{\Psi(x)}\right),$$

where $\Psi(x)$ is an increasing function with $\Psi(x) > \log x$ for all large $x$, $\Psi(1) > 0$
and
\[
\int_{1}^{x} \Psi(y)^{-1} dy \ll x \Psi(x)^{-1}.
\]

Then it was shown that
\[
V(x, Q) = Q \sum_{n \leq x} a_n^2 - Qx \sum_{q=1}^{\infty} g(q) + U(x, Q),
\]
where
\[
U(x, Q) \ll x^{3/2} \log x + x^2(\log 2x)^{9/2} \Psi(x)^{-1}
\]
\[+ x^2(\log x)^{4/3} \Psi(x)^{-2/3} + Q^2 E(x/Q),
\]
\[E(z) = \sum_{0 < q > y} g(q) dy,
\]
\[g(q) = \phi(q) \left( \sum_{r|q} f(q, r) \mu(q/r) \right)^2.
\]

Another example of a sequence whose mean value is a constant but whose mean square is unbounded is \(a_n = r(n)\), the number of ways of writing a positive integer \(n\) as the sum of two squares. This was studied in the context of the above variance by Dancs [Dan]. Therefore, it is of some interest to consider examples in which both the mean square and the mean value of \(\{a_n\}\) are unbounded. An example of such a sequence is \(a_n = d(n)\). The above variance has been studied in this special case by Pongsriiam [P] and Pongsriiam and Vaughan [PV]. They proved that if \(a_n = d(n)\), then

\[
V(x, x) = \frac{x^2}{\pi^2} (\log x)^3 + g_2 x^2 (\log x)^2 + g_1 x^2 \log x + g_0 x^2 + O(x^{5/3+\epsilon}),
\]

and for \(1 \leq Q < x\) and \(0 < \Theta < 1\),

\[
V(x, Q) = \frac{Qx}{\pi^2} \left( \log \frac{Q^2}{x} \right)^3 + h_3 Qx \left( \log \frac{Q^2}{x} \right)^2 + h_2 Qx \log \frac{Q^2}{x}
\]
\[+ h_1 Qx \log x + h_0 Qx + O(x^{5/3+\epsilon} + Q^{1+\Theta} x^{1-\Theta} (\log x)^2),
\]

where \(g_i (i = 0, 1, 2)\) and \(h_j (j = 0, 1, 2, 3)\) are constants.

Note that in all of the special cases studied hitherto, the behavior of the mean square is at least well understood even if it is not bounded. However, this is not the case when
\[
(1.6) \quad a_n = r_3(n) = \sum_{x_1, x_2, x_3 \mid x_1^3 + x_2^3 + x_3^3 = n} 1,
\]
the number of (ordered) representations of \(n\) as the sum of three cubes. In spite of significant work by Hooley, this function is only poorly understood.
The best known results about the mean value and the mean square of $r_3(n)$ (Vaughan [RV4]) are

\[
\sum_{n \leq x} r_3(n) = \Gamma(4/3)^3 x - \frac{\Gamma(4/3)^2}{2\Gamma(5/3)} x^{2/3} + O(x^{5/9}(\log x)^{1/3}),
\]

(1.7)

\[
\sum_{n \leq x} r_3(n)^2 \ll x^{7/6}(\log x)^{\varepsilon - 5/2}.
\]

(1.8)

In other words, the size of its mean square is unknown, and we are not sure whether the mean square is bounded or not. Therefore, this case is more difficult than the previous ones.

There are several conjectures on the mean square of $r_3(n)$. Hooley [Ho6] has shown that

\[
\sum_{n \leq x} r_3(n)^2 \ll x^{1+\varepsilon}
\]

assuming that a certain Hasse–Weil $L$-function satisfies the Riemann Hypothesis. It may be true that

\[
\sum_{n \leq x} r_3(n)^2 \sim Cx
\]

for some positive constant $C$, but we only know a lower bound of this order with a value of $C$ larger than the obvious guess. See Hooley [Ho5].

Also, by refining the deep work of Hooley [Ho4], Vaughan [RV4] has shown that for every $\varepsilon > 0$ there exists a constant $\delta > 0$ such that for all $Q \leq x$,

\[
\sum_{n \leq x} \max_{q \leq Q} \sup_{y \leq x} |\mathcal{Y}(y; q, a) - \Gamma(4/3)^3 y \rho(q, a) q^{-3}| \ll \varepsilon x^{8/9+\varepsilon} + x^{1/3} Q^{2/9} (Q^{10/9} + x^{5/9})(\log x)^{-\delta},
\]

(1.9)

where

\[
\mathcal{Y}(x; q, a) = \sum_{n \equiv a (\mod q)} r_3(n),
\]

(1.10)

and $\rho(q, a)$ denotes the number of solutions of the congruence $l_1^3 + l_2^3 + l_3^3 \equiv a \pmod{q}$.

When comparing (1.7), (1.9) and (1.10) with (1.1) and (1.2), we find that if $a_n = r_3(n)$, then

\[
A(x; q, a) = \mathcal{Y}(x; q, a),
\]

(1.11)

\[
f(q, a) = \rho(q, a) q^{-3},
\]

(1.12)

\[
M(x) = \Gamma(4/3)^3 x,
\]

(1.13)

and the variance (1.3) becomes
\[ V(x, Q) = \sum_{q \leq Q} \sum_{a=1}^{q} \left\lfloor \frac{x}{q} \right\rfloor \rho(q,a)q^{-3}. \]

In this article, we will prove the following result on that variance:

**Main Theorem 1.1.** Assume that \( \varepsilon \) and \( \delta \) are positive numbers satisfying \( 0 < \delta < 1/3 \) and let

\[ U_0(x, Q) = V(x, Q) - Q \sum_{n \leq x} r_3(n)^2 + A_1 Qx - A_2 Q^{5/3} x^{1/3}, \]

where \( A_1 \) and \( A_2 \) are appropriate positive constants. When \( x^{1/2} \log x \leq Q \leq x \) one has

\[ U_0(x, Q) \ll x^{10/9 + \frac{2}{3} A^+ + \varepsilon} + Q^{2-\delta} x^\delta, \]

where \( A^+ \) is the upper exponent of \( x \) in the representation of \( \sum_{n \leq x} r_3(n)^2 \), and

\[ U_0(x, Q) \ll x^{10/9 + \varepsilon} \left( \sum_{n \leq x} r_3(n)^2 \right)^{2/3} + Q^{2-\delta} x^\delta \]

if the exponent of \( x \) in the representation of \( \sum_{n \leq x} r_3(n)^2 \) exists.

The terms “upper exponent” and “exponent” will be defined in (3.6) and (3.8), and the positive constants \( A_1 \) and \( A_2 \) will be defined in (10.19) and (10.20) respectively.

For large \( Q \), we provide a more precise conclusion:

**Main Theorem 1.2.** Assume that \( \varepsilon \) is a positive number and \( k \geq 1 \) is an integer, and let

\[ U_0(x, Q) = V(x, Q) - Q \sum_{n \leq x} r_3(n)^2 - x^2 \left( C_0 \log \frac{x}{Q} + C_0 C_2 - C_1 \right) \]

\[ - \Gamma(4/3)^6 \sum_{l \leq k} \frac{h(l)}{l} (x - lQ)^2, \]

where the function \( h(l) \) only depends on \( l \), and \( C_0, C_1 \) and \( C_2 \) are appropriate constants. When

\[ \frac{x}{k+1} < Q \leq \frac{x}{k} \]

one has

\[ U_0(x, Q) \ll x^{10/9 + \frac{2}{3} A^+ + \varepsilon}, \]

where \( A^+ \) is the upper exponent of \( x \) in the representation of \( \sum_{n \leq x} r_3(n)^2 \), and

\[ U_0(x, Q) \ll x^{10/9 + \varepsilon} \left( \sum_{n \leq x} r_3(n)^2 \right)^{2/3} \]

if the exponent of \( x \) in the representation of \( \sum_{n \leq x} r_3(n)^2 \) exists.
To be more precise, the function $h(l)$ will be defined in (9.1), the constants $C_0$ and $C_1$ will be defined in (5.5) and (5.6), and $C_2$ will be given in (11.3).

2. The standard initial procedure. We start from the general case of (1.3). By (1.1) and the binomial theorem, we have

$$V(x, Q) = 2S_1 - 2S_2 + S_3 + [Q] \sum_{n \leq x} a_n^2,$$

where

$$S_1 = \sum_{q \leq Q} \sum_{\substack{m < n \leq x \\mod q}} a_m a_n,$$

$$S_2 = M(x) \sum_{q \leq Q} \sum_{a=1}^{q} f(q, a) \sum_{n \leq x} a_n,$$

$$S_3 = M(x)^2 \sum_{q \leq Q} \sum_{a=1}^{q} f(q, a)^2.$$

Here $[Q]$ is the integer part of $Q$. To avoid the case when $Q$ is not an integer, we may rewrite the last term of (2.1) as

$$Q \sum_{n \leq x} a_n^2 + O(\sum_{n \leq x} a_n^2).$$

The most important part is to estimate $S_1$. We give the main conclusions following §3 of Goldston and Vaughan [GV], §4 of Vaughan [RV2], and §4 of Vaughan [RV3]. First, let

$$F(\alpha) = \sum_{q \leq Q} \sum_{r \leq x/q} e(\alpha qr),$$

$$G(\alpha) = \sum_{n \leq x} a_n e(n\alpha).$$

Then for any $r_0 \in \mathbb{R}$,

$$S_1 = \int_{r_0}^{1+r_0} F(\alpha)|G(\alpha)|^2 \, d\alpha.$$

We also have

$$F(\alpha) = F_q(\alpha) + H_q(\alpha),$$

where

$$F_q(\alpha) = \sum_{l \leq \sqrt{x}} \left( \sum_{\substack{m \leq x/l \\mod \sqrt{x} \leq m \leq \min(Q,x/l) \\mod q \not| l}} e(\alpha ml) \right)$$

and $H_q(\alpha)$ is the corresponding multiple sum with $q \not| l$. 
Suppose that $x$ is sufficiently large, and $R$ satisfies
\begin{equation}
2\sqrt{x} \leq R \leq \frac{1}{2}x.
\end{equation}

To construct the major arcs and the minor arcs, we need to use the Farey sequence of order $R$, which is defined as the sequence of all fractions $a/q$ such that $a, q \in \mathbb{Z}, 1 \leq q \leq R$ and $(a, q) = 1$, listed in order of their size (see Niven, Zuckerman and Montgomery [NZM, p. 300]).

Now for any fraction $a/q$ in the Farey sequence of order $R$, let $a_-/q_-$ and $a_+/q_+$ be the preceding and succeeding terms, and define
\begin{equation}
\mathcal{M}(q, a) = \left( \frac{a + a_-}{q + q_-}, \frac{a + a_+}{q + q_+} \right).
\end{equation}

It is obvious that all $\mathcal{M}(q, a)$ such that $a, q \in \mathbb{Z}, 1 \leq a \leq q \leq R$ and $(a, q) = 1$ form a partition of the unit interval
\begin{equation}
\left( \frac{1}{[R] + 1}, \frac{[R] + 2}{[R] + 1} \right),
\end{equation}
so letting
\begin{equation}
r_0 = \frac{1}{[R] + 1}
\end{equation}
in (2.7), we have
\begin{equation}
S_1 = \sum_{q \leq R} \sum_{a=1}^{q} \int_{\mathcal{M}(q, a)} F_q(\alpha)|G(\alpha)|^2 \, d\alpha
+ \sum_{q \leq R} \sum_{a=1}^{q} \int_{\mathcal{M}(q, a)} H_q(\alpha)|G(\alpha)|^2 \, d\alpha.
\end{equation}

Note that the outer sum of the first term on the RHS can be rewritten as taken over $q \leq \sqrt{x}$, since by definition (2.8), $F_q(\alpha) = 0$ when $q > \sqrt{x}$.

Now let $\beta = \alpha - a/q$ when $\alpha \in \mathcal{M}(q, a)$. To measure the length of $\mathcal{M}(q, a)$ and estimate $\beta$, we need the following lemma, which can be proved similarly to [NZM Theorem 6.7]:

**Lemma 2.1.** If $a, q \in \mathbb{Z}, 1 \leq a \leq q \leq R$ and $(a, q) = 1$, $a_-/q_-$ and $a_+/q_+$ are the terms preceding and succeeding $a/q$ in the Farey sequence of order $R$ respectively, then
\begin{equation}
\frac{1}{2qR} \leq \frac{a}{q} - \frac{a + a_-}{q + q_-} < \frac{1}{qR},
\end{equation}
and
\begin{equation}
\frac{1}{2qR} \leq \frac{a + a_+}{q + q_+} - \frac{a}{q} < \frac{1}{qR}.
\end{equation}
From (2.9), (2.10) and the lemma above,

$$|\beta| < \frac{1}{qR} \leq \frac{1}{2q\sqrt{x}},$$

so

(2.13) \hspace{1cm} F_q(\alpha) \ll \frac{x \log(2\sqrt{x}/q)}{q + qx|\beta|} \quad \text{if} \quad q \leq \sqrt{x},

and

(2.14) \hspace{1cm} H_q(\alpha) \ll (\sqrt{x} + q) \log 2q,

which implies that \( H_q(\alpha) \ll R \log x \). Hence the second term on the RHS of (2.12) is

$$\ll R(\log x) \sum_{q \leq R} \sum_{a=1}^{q} \int_{\mathfrak{M}(q,a)} |G(\alpha)|^2 \, d\alpha = R(\log x) \sum_{n \leq x} a_n^2,$$

which dominates the error term

$$O\left( \sum_{n \leq x} a_n^2 \right)$$

since \( R \log x \) is sufficiently large.

Lemma 2.1 also shows that the interval

(2.15) \hspace{1cm} \mathfrak{N}(q,a) = \left[ \frac{a}{q} - \frac{1}{2qR} \frac{a}{q} + \frac{1}{2qR} \right]

is a subinterval of \( \mathfrak{M}(q,a) \). By (2.13), \( F_q(\alpha) \ll R \log x \) if \( \alpha \in \mathfrak{M}(q,a) \setminus \mathfrak{N}(q,a) \) and \( q \leq \sqrt{x} \), or if \( \alpha \in \mathfrak{N}(q,a) \) and \( x/R < q \leq \sqrt{x} \). So the first term on the RHS of (2.12) is

$$S_4 + O\left( R(\log x) \sum_{n \leq x} a_n^2 \right),$$

where

(2.16) \hspace{1cm} S_4 = \sum_{q \leq x/R} \sum_{a=1}^{q} \int_{\mathfrak{N}(q,a)} F_q(\alpha)|G(\alpha)|^2 \, d\alpha

$$= \sum_{q \leq x/R} \int_{1/(2qR)} F_q(\beta) \sum_{a=1}^{q} \left| G\left( \beta + \frac{a}{q} \right) \right|^2 \, d\beta.$$

Therefore,

(2.17) \hspace{1cm} S_1 = S_4 + O\left( R(\log x) \sum_{n \leq x} a_n^2 \right).
Hence we have

\[(2.18) \quad V(x, Q) = 2S_4 - S_3 + 2(S_3 - S_2) + Q \sum_{n \leq x} a_n^2 + O \left( R(\log x) \sum_{n \leq x} a_n^2 \right). \]

Note that

\[ Q \sum_{n \leq x} a_n^2 \]

can dominate

\[ O \left( R(\log x) \sum_{n \leq x} a_n^2 \right) \]

only when \( Q/(R \log x) \) is large, otherwise the result is trivial. So from now on, we always assume

\[(2.19) \quad x^{1/2} \log x \leq Q \leq x. \]

The treatment of \( S_4 \) depends on obtaining an approximation to the value \( G(\beta + a/q) \) which can be considered a product of local factors. Sometimes, it is possible to estimate \( G(\beta + a/q) \) as \( \nu(q)J(\beta) \) for some \( \nu(q) \) and \( J(\beta) \) that are independent of \( a \), at least when \( (a, q) = 1 \), so the integral will be changed to

\[(2.20) \quad \int_{-1/(2qR)}^{1/(2qR)} F_q(\beta)|J(\beta)|^2 \, d\beta \]

when factoring out \( \phi(q)\nu(q)^2 \). In other situations, we have to deal with more complex approximations, for example, \( \nu(q, a)J(\beta) \). Nevertheless, in this case, it is possible to factor out

\[ \sum_{\substack{a=1 \\ (a,q)=1}}^q |\nu(q, a)|^2 \]

for a given \( q \), and the integral will still be changed to \(2.20 \), which is likely to be close to

\[ \int_{-1/2}^{1/2} F_q(\beta)|J(\beta)|^2 \, d\beta. \]

3. Preliminary results when \( a_n = r_3(n) \). The method introduced in the previous section still works when \( a_n = r_3(n) \). For example, we still use \(2.18 \) to calculate the variance, where \( S_2, S_3, S_4, F_q(\alpha) \) and \( G(\alpha) \) are defined by \(2.3 \), \(2.4 \), \(2.16 \), \(2.8 \) and \(2.6 \) respectively. To figure out what these functions look like when \( a_n = r_3(n) \), we rewrite the expressions for \( V(x, Q) \),
$S_2$, $S_3$ and $G(\alpha)$ as follows:

\begin{equation}
V(x, Q) = 2S_4 - S_3 + 2(S_3 - S_2) + Q \sum_{n \leq x} r_3(n)^2 \\
+ O \left( R(\log x) \sum_{n \leq x} r_3(n)^2 \right),
\end{equation}

\begin{equation}
S_2 = \Gamma(4/3)^3 x \sum_{q \leq Q} \sum_{a=1}^{q} \frac{\rho(q, a)}{q^3} \sum_{n \leq x} r_3(n),
\end{equation}

\begin{equation}
S_3 = \Gamma(4/3)^6 x^2 \sum_{q \leq Q} \sum_{a=1}^{q} \frac{\rho(q, a)^2}{q^6},
\end{equation}

\begin{equation}
G(\alpha) = \sum_{n \leq x} r_3(n)e(n\alpha).
\end{equation}

Meanwhile, there is no need to rewrite the expressions for $S_4$ and $F_q(\alpha)$. The restrictions on the variables also hold. For example, the order of the Farey sequence $R$ still satisfies (2.9), and $Q$ still follows the assumption (2.19).

To calculate the variance, first, we see that the main term and the error term of $V(x, Q)$ largely depend on the second-moment sum of $r_3(n)$,

\begin{equation}
\sum_{n \leq x} r_3(n)^2,
\end{equation}

which is only poorly understood. Therefore, to measure the size of the main term and the error term, we need to redefine the exponent of $x$ in an expression $P(x)$. First, define the upper exponent $E^+$ and lower exponent $E^-$ of $x$ in $P(x)$ as follows:

\begin{equation}
E^+ = \limsup_{x \to \infty} \frac{\log P(x)}{\log x},
\end{equation}

\begin{equation}
E^- = \liminf_{x \to \infty} \frac{\log P(x)}{\log x}.
\end{equation}

If $E^+ = E^-$, then we define the exponent $E$ of $x$ in $P(x)$ as

\begin{equation}
E = E^+ = E^-.
\end{equation}

For example, if $A^+, A^-$ and $A$ are the upper exponent, the lower exponent and the exponent of $x$ in (3.5), then by (1.7) and (1.8), we have

\begin{equation}
1 \leq A^- \leq A^+ \leq 7/6.
\end{equation}

Later, we will show that the size of the main term and the error term depend on $A^+$, which is still unknown. Once its value is found after some research on (3.5), the results mentioned above will be finalized accordingly.
4. Several lemmata. Here we prove some lemmata. First, define the exponential sum $S(q, a)$ as follows:

\[(4.1)\quad S(q, a) = \sum_{m=1}^{q} e\left(\frac{am^3}{q}\right).\]

The following few lemmas show the properties of $S(q, a)$.

**Lemma 4.1.** Let $\rho(q, a)$ denote the number of solutions of the congruence $l_1^3 + l_2^3 + l_3^3 \equiv a \pmod{q}$. Then

\[(4.2)\quad \rho(q, a) = \frac{1}{q} \sum_{b=1}^{q} e\left(-\frac{ba}{q}\right) S(q, b)^3,\]

\[(4.3)\quad \sum_{a=1}^{q} \rho(q, a)^2 = \frac{1}{q} \sum_{b=1}^{q} |S(q, b)|^6 = q^5 \sum_{r|q} \frac{1}{r^6} \sum_{\substack{c=1 \atop (c,r)=1}}^{r} |S(r, c)|^6.\]

**Proof.** The value of

\[1 \sum_{q}^q e\left(\frac{b(l_1^3 + l_2^3 + l_3^3 - a)}{q}\right) = \frac{1}{q} \sum_{b=1}^{q} e\left(-\frac{ba}{q}\right) e\left(\frac{bl_1^3}{q}\right) e\left(\frac{bl_2^3}{q}\right) e\left(\frac{bl_3^3}{q}\right),\]

is 1 if $(l_1, l_2, l_3) \pmod{q}$ is a solution of the congruence $l_1^3 + l_2^3 + l_3^3 \equiv a \pmod{q}$, and 0 otherwise. So the number of all such solutions is

\[
\rho(q, a) = \sum_{l_1=1}^{q} \sum_{l_2=1}^{q} \sum_{l_3=1}^{q} \left(\frac{1}{q} \sum_{b=1}^{q} e\left(-\frac{ba}{q}\right) e\left(\frac{bl_1^3}{q}\right) e\left(\frac{bl_2^3}{q}\right) e\left(\frac{bl_3^3}{q}\right)\right)
\]

\[= \frac{1}{q} \sum_{b=1}^{q} e\left(-\frac{ba}{q}\right) S(q, b)^3.\]

As $\rho(q, a)$ is a nonnegative integer, we have $\rho(q, a) = \overline{\rho(q, a)} = |\rho(q, a)|$, so

\[
\sum_{a=1}^{q} \rho(q, a)^2 = \sum_{a=1}^{q} \left(\frac{1}{q} \sum_{b_1=1}^{q} e\left(-\frac{b_1a}{q}\right) S(q, b_1)^3\right) \left(\frac{1}{q} \sum_{b_2=1}^{q} e\left(\frac{b_2a}{q}\right) S(q, b_2)^3\right)
\]

\[= \frac{1}{q} \sum_{b=1}^{q} |S(q, b)|^6.\]

To complete the proof, note that if $c/r$ is the simplest form of $b/q$, in other words, if $c/r = b/q$ and $(c, r) = 1$, then $S(q, b) = (q/r)S(r, c)$. So

\[
\sum_{b=1}^{q} |S(q, b)|^6 = \sum_{r|q} \sum_{c=1 \atop (c,r)=1}^{r} \left|\frac{q}{r} S(r, c)\right|^6 = q^6 \sum_{r|q} \frac{1}{r^6} \sum_{c=1 \atop (c,r)=1}^{r} |S(r, c)|^6.\]

Lemma 4.2. Suppose that \( p \) is a prime number and \( p \nmid a \). Then

\[
S(p, a) = \sum_{\chi \in A} \overline{\chi}(a)\tau(\chi),
\]

where \( \overline{\chi} \) represents the congruent of \( \chi \), \( A \) denotes the set of nonprincipal characters \( \chi \) modulo \( p \) for which \( \chi^3 \) is principal, \( \text{card} \ A = (3, p - 1) - 1 \) and \( \tau(\chi) \) satisfies \( |\tau(\chi)| = p^{1/2} \).

Proof. See Vaughan [RV1, pp. 45–46, Lemma 4.3], and let \( k = 3 \). ■

Corollary 4.3. Suppose that \( p \) is a prime number and \( p \nmid a \). Then

(i) \( S(p, a) = 0 \) if \( p = 3 \) and \( p \equiv 2 \pmod{3} \),
(ii) \( |S(p, a)| \leq 2p^{1/2} \) if \( p \equiv 1 \pmod{3} \).

Proof. If \( p = 3 \) or \( p \equiv 2 \pmod{3} \), then \( A = \emptyset \), so \( S(p, a) = 0 \).

If \( p \equiv 1 \pmod{3} \), then \( \text{card} \ A = 2 \), say \( A = \{\chi_1, \chi_2\} \). By (4.4),

\[
S(p, a) = \overline{\chi_1}(a)\tau(\chi_1) + \overline{\chi_2}(a)\tau(\chi_2),
\]

where

\[
|\tau(\chi_1)| = |\tau(\chi_2)| = p^{1/2}.
\]

Hence \( |S(p, a)| \leq 2p^{1/2} \). ■

Lemma 4.4. Suppose that \( p \) is a prime number, \( p \nmid a \) and \( l \) is an integer. Then

\[
S(p^l, a) = \begin{cases} 
 p & \text{when } l = 2 \text{ and } p \neq 3, \\
 p^2 & \text{when } l = 3, \\
 p^2S(p^{l-3}, a) & \text{when } l > 3.
\end{cases}
\]

Proof. See Vaughan [RV1, p. 46, Lemma 4.4] and let \( k = 3 \). Note that \( \gamma = 2 \) when \( p = 3 \), and \( \gamma = 1 \) otherwise (see [RV1, p. 22]). ■

Corollary 4.5. Suppose that \( p \) is a prime number, \( p \nmid a \), \( u \) is a non-negative integer and \( v = 1, 2 \text{ or } 3 \). Then \( S(p^{3u+v}, a) = p^{2u}S(p^v, a) \).

Proof. Use Lemma 4.4 and induction on \( u \). ■

Lemma 4.6. Suppose that \((q, r) = (qr, a) = 1\). Then

\[
S(qr, a) = S(q, ar^2)S(r, aq^2).
\]

Proof. See Vaughan [RV1, p. 47, Lemma 4.5], and let \( k = 3 \). ■

Lemma 4.7. Suppose that \((q, a) = 1\), then

\[
S(q, a) \ll q^{2/3}.
\]

Proof. See Vaughan [RV1, p. 47, Theorem 4.2], and let \( k = 3 \). ■
By prime factorization, every positive integer $r$ can be uniquely written as

$$r = r_1 r_2^2 r_3^3,$$

where $r_1$ and $r_2$ are squarefree, and $(r_1, r_2) = 1$. Then the following lemma holds.

**Lemma 4.8.** If $(r, a) = 1$, then for every $\varepsilon > 0$,

$$|S(r, a)| \ll r_1^{1/2} r_2 r_3^2,$$

where $r_1, r_2$ and $r_3$ are given by (4.8).

**Proof.** The result is trivial when $r = 1$, so suppose that $r \geq 2$.

First consider the case when $r = p^\alpha$, where $p$ is a prime number and $\alpha$ is a positive integer. Let $\alpha = 3u + v$, where $u$ is a nonnegative integer and $v = 1, 2$ or 3.

(i) If $v = 3$, then $r_1 = 1, r_2 = 1, r_3 = p^{u+1}$. By Corollary 4.5 and Lemma 4.4,

$$S(p^\alpha, a) = S(p^{3u+3}, a) = p^{2u+2} = r_1^{1/2} r_2 r_3^2.$$

(ii) If $v = 2$, then $r_1 = 1, r_2 = p, r_3 = p^u$. By Corollary 4.5,

$$S(p^\alpha, a) = S(p^{3u+2}, a) = p^{2u} S(p^2, a).$$

If $p \neq 3$, then by Lemma 4.4,

$$S(p^\alpha, a) = p^{2u+1} = r_1^{1/2} r_2 r_3^2.$$

If $p = 3$, then an immediate calculation gives $|S(p^2, a)| \leq 9 (= 3p)$. So

$$|S(p^\alpha, a)| \leq 3p^{2u+1} = 3r_1^{1/2} r_2 r_3^2.$$

(iii) If $v = 1$, then $r_1 = p, r_2 = 1, r_3 = p^u$. By Corollaries 4.5 and 4.3,

$$|S(p^\alpha, a)| = |S(p^{3u+1}, a)| = p^{2u} |S(p, a)| \leq 2p^{2u+1/2} = 2r_1^{1/2} r_2 r_3^2.$$

To sum up, we always have

$$|S(p^\alpha, a)| \leq 3r_1^{1/2} r_2 r_3^2.$$  

Finally, for every integer $r \geq 2$, use prime factorization: $r = \prod_{i=1}^{t} p_i^{\alpha_i}$, where $p_i$ are distinct primes and $\alpha_i$ are positive integers for $1 \leq i \leq t$. Then by Lemma 4.6 and induction, there exist $a_i$ for $1 \leq i \leq t$ such that $(p_i, a_i) = 1$ and

$$S(r, a) = \prod_{i=1}^{t} S(p_i^{\alpha_i}, a_i).$$

From the definition, $r_1, r_2$ and $r_3$ are multiplicative as functions of $r$. Therefore, by (4.10),

$$|S(r, a)| \leq 3^{\omega(r)} r_1^{1/2} r_2 r_3^2,$$

where $r_1, r_2$ and $r_3$ are given by (4.8).
Theorem 2.10 in Montgomery and Vaughan [MV] tells us that $3^{\omega(r)} \ll r^\varepsilon$ for every $\varepsilon > 0$. Hence the lemma follows. ■

The following two lemmas concern the function $T(r)$ defined by

\begin{equation}
T(r) = \frac{1}{r^7} \sum_{\substack{c=1 \\ (c,r)=1}}^r |S(r,c)|^6.
\end{equation}

**Lemma 4.9.** We have $T(r) \ll r^{-2}$.

*Proof.* By Lemma 4.7 $|S(r,c)| \ll r^{2/3}$ if $(c, r) = 1$. So

$$T(r) \ll \frac{1}{r^7} \sum_{\substack{c=1 \\ (c,r)=1}}^r (r^{2/3})^6 = \frac{\phi(r)}{r^3} \leq \frac{1}{r^2}.$$ ■

**Lemma 4.10.** $T(r)$ is a multiplicative function.

*Proof.* Obviously $T(1) = 1$, so it is sufficient to prove that if $(r_1, r_2) = 1$, then $T(r_1 r_2) = T(r_1) T(r_2)$. By (4.11) and Lemma 4.6

$$T(r_1 r_2) = \frac{1}{r_1^7 r_2^7} \sum_{\substack{c_1=1 \\ (c_1, r_1)=1}}^{r_1} \sum_{\substack{c_2=1 \\ (c_2, r_2)=1}}^{r_2} |S(r_1, c_1 r_2^2)|^6 |S(r_2, c_2 r_1^2)|^6.$$ From elementary number theory, we know that if $1 \leq c \leq r_1 r_2$, $(c, r_1 r_2) = 1$ and $(r_1, r_2) = 1$, then there exists a unique pair of numbers $c_1$ and $c_2$, such that $c \equiv c_2 r_1 + c_1 r_2 \pmod{r_1 r_2}$, $1 \leq c_i \leq r_i$ and $(c_i, r_i) = 1$ where $i = 1$ or 2. In this situation, we have $S(r_1, c_1 r_2^2) = S(r_1, c_1)$ and $S(r_2, c_2 r_1^2) = S(r_2, c_2)$ by (4.1). So the identity above becomes

$$T(r_1 r_2) = \frac{1}{r_1^7 r_2^7} \sum_{\substack{c_1=1 \\ (c_1, r_1)=1}}^{r_1} \sum_{\substack{c_2=1 \\ (c_2, r_2)=1}}^{r_2} |S(r_1, c_1)|^6 |S(r_2, c_2)|^6 = T(r_1) T(r_2).$$ ■

To estimate $S_2 - S_3$, we need the following two lemmas.

**Lemma 4.11.** Suppose that $(q, a) = 1$. Then for every $\varepsilon > 0$,

\begin{equation}
\sum_{x \leq n^{1/3}} e\left(\frac{ax^3}{q}\right) = q^{-1} S(q,a)n^{1/3} + O(q^{1/2+\varepsilon}).
\end{equation}

*Proof.* See Vaughan [RV], p. 43, Theorem 4.1. Let $\beta = 0$ (thus $\alpha = a/q$) and $k = 3$ there, and use the refined version for $V(\alpha, q, a)$, i.e. replace $v(\beta)$ by $v_1(\beta)$; then the lemma follows. ■

**Lemma 4.12.** Suppose that $(q, a) = 1$. Then for every $\varepsilon > 0$,

\begin{equation}
\sum_{n \leq x} r_3(n) e\left(\frac{an}{q}\right) = \Gamma(4/3)^3 x q^{-3} S(q,a)^3 + O(x^{2/3} q^{1/2+\varepsilon}).
\end{equation}
Proof. By (1.6),
\[
\sum_{n \leq x} r_3(n) e\left(\frac{an}{q}\right) = \sum_{x_1, x_2, x_3 \leq x} e\left(\frac{a(x_1^3 + x_2^3)}{q}\right) \sum_{x_3 \leq (x-x_1^3-x_2^3)^{1/3}} e\left(\frac{ax_3^3}{q}\right).
\]
Consider the innermost sum. By Lemma 4.11,
\[
\sum_{x_3 \leq (x-x_1^3-x_2^3)^{1/3}} e\left(\frac{ax_3^3}{q}\right) = q^{-1}S(q, a)(x - x_1^3 - x_2^3)^{1/3} + O(q^{1/2+\varepsilon}).
\]
So
\[
(4.14) \quad \sum_{n \leq x} r_3(n) e\left(\frac{an}{q}\right) = q^{-1}S(q, a) \sum_{x_1, x_2 \leq x, x_1^3 + x_2^3 \leq x} e\left(\frac{a(x_1^3 + x_2^3)}{q}\right) (x - x_1^3 - x_2^3)^{1/3} + O\left(\sum_{x_1, x_2 \leq x, x_1^3 + x_2^3 \leq x} q^{1/2+\varepsilon}\right).
\]
The error term above is
\[
\ll q^{1/2+\varepsilon} \sum_{x_1, x_2 \leq x, x_1^3 + x_2^3 \leq x} 1 \leq q^{1/2+\varepsilon} x^{2/3}.
\]
For the main term, repeat this argument for the sum over \(x_2\): it is
\[
= \frac{S(q, a)}{q} \sum_{x_1 \leq x^{1/3}} e\left(\frac{ax_1^3}{q}\right) \sum_{x_2 \leq (x-x_1^3)^{1/3}} e\left(\frac{ax_2^3}{q}\right) (x - x_1^3 - x_2^3)^{1/3}
\]
\[
= \frac{S(q, a)}{q} \sum_{x_1 \leq x^{1/3}} e\left(\frac{ax_1^3}{q}\right) \sum_{x_2 \leq (x-x_1^3)^{1/3}} e\left(\frac{ax_2^3}{q}\right) \int_0^{(x-x_1^3)^{1/3}} (x - x_1^3 - y^3)^{-2/3} y^2 \left(\sum_{x_2 \leq y} e\left(\frac{ax_2^3}{q}\right)\right) dy
\]
by changing the order of the sum and the integral. By Lemma 4.11, the innermost sum in the parenthesis is
\[
\sum_{x_2 \leq y} e\left(\frac{ax_2^3}{q}\right) = q^{-1}S(q, a)y + O(q^{1/2+\varepsilon}).
\]
Hence the main term of (4.14) is
\[
(4.15) \quad q^{-2}S(q, a)^2 \sum_{x_1 \leq x^{1/3}} e\left(\frac{ax_1^3}{q}\right) \int_0^{(x-x_1^3)^{1/3}} (x - x_1^3 - y^3)^{-2/3} y^3 dy
\]
\[
+ O\left(q^{-1}|S(q, a)| \sum_{x_1 \leq x^{1/3}} \int_0^{(x-x_1^3)^{1/3}} (x - x_1^3 - y^3)^{-2/3} y^2 q^{1/2+\varepsilon} dy\right).
\]
Note that $S(q, a) \ll q$, so again the error term above is
\[
\ll \sum_{x_1 \leq x^{1/3}} \int_0^{(x-x_1^3)^{1/3}} (x-x_1^3-y^3)^{-2/3} y^{2\epsilon} q^{-1/2} \, dy \leq q^{1/2+\epsilon} x^{2/3}.
\]
Finally, calculating the integral using the beta function $B(a, b)$ and repeating the arguments, we find that the main term is
\[
= \frac{1}{3} B(4/3, 1/3) q^{-2} S(q, a)^2 \sum_{x_1 \leq x^{1/3}} e\left(\frac{ax_1^3}{q}\right) (x-x_1^3)^{2/3}
= \frac{2}{3} B(4/3, 1/3) q^{-2} S(q, a)^2 \sum_{x_1 \leq x^{1/3}} e\left(\frac{ax_1^3}{q}\right) \int_{x_1}^{x^{1/3}} (x-y^3)^{-1/3} y^2 \, dy
= \frac{2}{3} B(4/3, 1/3) q^{-2} S(q, a)^2 \int_0^{x^{1/3}} (x-y^3)^{-1/3} y^2 (q^{-1} S(q, a) y + O(q^{1/2+\epsilon})) \, dy
= \frac{2}{3} B(4/3, 1/3) q^{-3} S(q, a)^3 \cdot \frac{1}{3} B(4/3, 2/3) x + O(x^{2/3} q^{1/2+\epsilon}),
\]
where
\[
\frac{2}{3} B(4/3, 1/3) \cdot \frac{1}{3} B(4/3, 2/3) = \Gamma(4/3)^3.
\]
Therefore, by (4.14) and (4.15), the lemma follows.

Now define
\[
(4.16) \quad \nu(q, a) = \frac{S(q, a)^3}{q^3},
\]
\[
(4.17) \quad J(\beta) = \Gamma(4/3)^3 \sum_{n \leq x} e(\beta n).
\]
We have $\nu(q, a) \ll 1$ since $S(q, a) \ll q$. More precisely, by Lemma 4.7
\[
(4.18) \quad \nu(q, a) \ll 1/q \quad \text{if} \quad (q, a) = 1.
\]
There are two estimates on $J(\beta)$:
\[
(4.19) \quad J(\beta) \ll \sum_{n \leq x} 1 \leq x,
\]
and
\[
(4.20) \quad J(\beta) \ll 1/\|\beta\| = 1/|\beta| \quad \text{when} \quad \beta \in [-1/2, 1/2] \setminus \{0\}.
\]
So if $x \in [-1/2, 1/2]$, then
\[
(4.21) \quad J(\beta) \ll \frac{x}{1+x|\beta|}.
\]
We also have

\[(4.22) \int_{-1/2}^{1/2} |J(\beta)|^2 \, d\beta = \Gamma(4/3)^6 [x].\]

The lemma below tells us that \(\nu(q, a)J(\beta)\) is a good approximation of \(G(\beta + a/q)\).

**Lemma 4.13.** Suppose that \((q, a) = 1\) and \(x \geq 1\). Then for every \(\varepsilon > 0\),

\[(4.23) \quad G\left(\beta + \frac{a}{q}\right) = \nu(q, a)J(\beta) + O(x^{2/3}q^{1/2+\varepsilon}(1 + x|\beta|)),\]

where \(G(\alpha)\), \(\nu(q, a)\) and \(J(\beta)\) are defined by (3.4), (4.16) and (4.17) respectively.

**Proof.** By definition,

\[G\left(\beta + \frac{a}{q}\right) = \sum_{n \leq x} r_3(n) e\left(\frac{an}{q}\right) e(\beta n)\]

\[= \sum_{n \leq x} r_3(n) e\left(\frac{an}{q}\right) \left( e(\beta x) - \int_0^x 2\pi i\beta e(\beta t) \, dt \right)\]

\[= e(\beta x) G\left(\frac{a}{q}\right) - \int_0^x 2\pi i\beta e(\beta t) \left( \sum_{n \leq t} r_3(n) e\left(\frac{an}{q}\right) \right) dt.\]

By Lemma 4.12, the above is

\[= \Gamma(4/3)^3 x q^{-3} S(q, a)^3 e(\beta x) + O(x^{2/3}q^{1/2+\varepsilon})\]

\[- \int_0^x 2\pi i\beta e(\beta t) \Gamma(4/3)^3 t q^{-3} S(q, a)^3 \, dt + O\left( \int_0^x |\beta|^2 t^{2/3} q^{1/2+\varepsilon} \, dt \right)\]

\[= \Gamma(4/3)^3 q^{-3} S(q, a)^3 \left( xe(\beta x) - \int_0^x 2\pi i\beta e(\beta t) t \, dt \right) + O(x^{2/3}q^{1/2+\varepsilon}(1 + x|\beta|)).\]

Integrating by parts gives

\[\int_0^x 2\pi i\beta e(\beta t) t \, dt = xe(\beta x) - \int_0^x e(\beta t) \, dt,\]

so

\[G\left(\beta + \frac{a}{q}\right) = \Gamma(4/3)^3 q^{-3} S(q, a)^3 \int_0^x e(\beta t) \, dt + O(x^{2/3}q^{1/2+\varepsilon}(1 + x|\beta|))\]

\[= \nu(q, a)J(\beta) + \Gamma(4/3)^3 \nu(q, a) \left( \int_0^x e(\beta t) \, dt - \sum_{n \leq x} e(\beta n) \right)\]

\[+ O(x^{2/3}q^{1/2+\varepsilon}(1 + x|\beta|)).\]
Finally, we have
\[
\sum_{n \leq x} e(\beta n) = \sum_{n \leq x} \left( e(\beta x) - \int_{n}^{x} 2\pi i \beta e(\beta t) \, dt \right)
\]
\[
= e(\beta x) \sum_{n \leq x} 1 - \int_{0}^{x} 2\pi i \beta e(\beta t) \left( \sum_{n \leq t} 1 \right) \, dt
\]
\[
= xe(\beta x) - \int_{0}^{x} 2\pi i \beta e(\beta t) t \, dt + O(1 + x|\beta|)
\]
\[
= \int_{0}^{x} e(\beta t) \, dt + O(1 + x|\beta|).
\]
Hence by (4.18), the lemma is proved. ■

Now define
\[
\Delta(q, a, \beta) = G\left( \beta + \frac{a}{q} \right) - \nu(q, a) J(\beta).
\]

Lemma 4.13 says that if \( (q, a) = 1 \) and \( x \geq 1 \), then for every \( \varepsilon > 0 \),
\[
\Delta(q, a, \beta) \ll x^{2/3} q^{1/2+\varepsilon} (1 + x|\beta|).
\]

5. The main term \( S_3 \). Now we calculate the variance \( V(x, Q) \) according to (3.1). We start from the main term \( S_3 \). By (3.3), Lemma 4.1 and (4.11),
\[
\Gamma(4/3) - 6x^{-2} S_3 = \sum_{q \leq Q} \frac{1}{q} \sum_{r | q} r T(r).
\]
Changing the order of summation, we have
\[
\Gamma(4/3) - 6x^{-2} S_3 = \sum_{r \leq Q} \sum_{\ell \leq Q/r} \frac{1}{\ell r} \cdot r T(r)
\]
\[
= (\log Q + \gamma) \sum_{r \leq Q} T(r) - \sum_{r \leq Q} T(r) \log r + O\left( \frac{1}{Q} \sum_{r \leq Q} r T(r) \right).
\]
By Lemma 4.9, \( T(r) \ll r^{-2} \), so
\[
\sum_{r > Q} T(r) \ll \frac{1}{Q},
\]
\[
\sum_{r > Q} T(r) \log r \ll \frac{\log Q}{Q},
\]
\[
\sum_{r \leq Q} r T(r) \ll \log Q.
\]
and the constants
\begin{align}
(5.5) \quad C_0 &= \Gamma(4/3)^6 \sum_{r=1}^{\infty} T(r), \\
(5.6) \quad C_1 &= \Gamma(4/3)^6 \sum_{r=1}^{\infty} T(r) \log r
\end{align}
are well-defined since the series are convergent. Therefore,
\begin{equation}
(5.7) \quad S_3 = C_0 x^2 \log Q + (\gamma C_0 - C_1) x^2 + O\left(\frac{x^2 \log Q}{Q}\right).
\end{equation}

6. The error term $S_2 - S_3$. We are concerned with the error term
\begin{equation}
(6.1) \quad S_2 - S_3 = \Gamma(4/3)^3 x \sum_{q \leq Q} \rho(q, a) \frac{\rho(q, a)^2}{q^6} \left( \sum_{n \leq x, n \equiv a \pmod{q}} r_3(n) \right) - \frac{\rho(q, a)}{q^3} \Gamma(4/3)^3 x.
\end{equation}
By the proof of (4.3), we have
\begin{equation}
\sum_{a=1}^{q} \rho(q, a)^2 \frac{\Gamma(4/3)^3 x}{q^6} = \frac{1}{q^7} \sum_{b=1}^{q} S(q, b)^3 S(q, b)^3 \Gamma(4/3)^3 x,
\end{equation}
and by (4.2),
\begin{align}
&\sum_{a=1}^{q} \frac{\rho(q, a)}{q^3} \left( \sum_{n \leq x, n \equiv a \pmod{q}} r_3(n) \right) = \sum_{a=1}^{q} \frac{1}{q^4} \sum_{b=1}^{q} e\left(\frac{-ba}{q}\right) S(q, b)^3 \sum_{n \leq x, n \equiv a \pmod{q}} r_3(n) \\
&= \frac{1}{q^4} \sum_{b=1}^{q} S(q, b)^3 \sum_{n \leq x} e\left(\frac{-bn}{q}\right) r_3(n),
\end{align}
so
\begin{equation}
\sum_{a=1}^{q} \frac{\rho(q, a)}{q^3} \left( \sum_{n \leq x, n \equiv a \pmod{q}} r_3(n) \right) - \frac{\rho(q, a)}{q^3} \Gamma(4/3)^3 x
\end{equation}
\begin{align}
&= \frac{1}{q^4} \sum_{b=1}^{q} S(q, b)^3 \left( \sum_{n \leq x} e\left(\frac{-bn}{q}\right) r_3(n) \right) - \Gamma(4/3)^3 x q^{-3} S(q, -b)^3.
\end{align}
By taking out the common factor $(q, b)$, we can observe that the expression above is
\begin{equation}
\frac{1}{q} \sum_{r \mid q} \sum_{(a, r) = 1}^{r} S(r, a)^3 \frac{r^3}{r^3} \left( \sum_{n \leq x} e\left(\frac{-an}{r}\right) r_3(n) \right) - \Gamma(4/3)^3 x r^{-3} S(r, -a)^3.
\end{equation}
By Lemma 4.12, this is
\[
\ll \frac{1}{q} \sum_{r \mid q} \sum_{\substack{a = 1 \\ (a, r) = 1}}^r \frac{|S(r, a)|^3}{r^3} x^{2/3} r^{1/2 + \epsilon},
\]
so by (6.1) and Lemma 4.8 for any \( \epsilon > 0 \), we have
\[
S_2 - S_3 \ll x^{5/3} \sum_{q \leq Q} \frac{1}{q} \sum_{r \mid q} \sum_{\substack{a = 1 \\ (a, r) = 1}}^r |S(r, a)|^3 r^{-5/2 + \epsilon}
\]
\[
= x^{5/3} \sum_{r \leq Q} \left( \sum_{l \leq Q/r} \frac{1}{l} \right) \sum_{\substack{a = 1 \\ (a, r) = 1}}^r |S(r, a)|^3 r^{-7/2 + \epsilon}
\]
\[
\ll x^{5/3} \sum_{r \leq Q} (\log Q) r^{-5/2 + 4\epsilon} r_1^{3/2} r_2^3 r_3^6
\]
\[
\ll x^{5/3} Q\epsilon \left( \sum_{r_1 \leq Q} r_1^{4\epsilon - 1} \right) \left( \sum_{r_2 = 1}^\infty r_2^{8\epsilon - 2} \right) \left( \sum_{r_3 = 1}^\infty r_3^{12\epsilon - 3/2} \right),
\]
where \( r_1, r_2 \) and \( r_3 \) are given by (4.8). If \( \epsilon \) is small enough, then the series inside the second and third parentheses are convergent. Also,
\[
\sum_{r_1 \leq Q} r_1^{4\epsilon - 1} \leq Q^{4\epsilon} \sum_{r_1 \leq Q} r_1^{-1} \ll Q^{5\epsilon}.
\]
Hence \( S_2 - S_3 \ll x^{5/3} Q^{6\epsilon} \). Finally, since \( \epsilon \) represents an arbitrary positive real number, then \( 6\epsilon \) also has the same meaning, so we can rewrite \( 6\epsilon \) as \( \epsilon \). Therefore,
\[
(6.2) \quad S_2 - S_3 = O(x^{5/3} Q^\epsilon).
\]

7. The major arcs. By (4.24), we have
\[
\left| G \left( \beta + \frac{a}{q} \right) \right|^2 = |\nu(q, a)|^2 |J(\beta)|^2 + \Delta_1 + \Delta_2,
\]
where
\[
\Delta_1 = 2\Re \left( \nu(q, a) \cdot J(\beta) \cdot \Delta(q, a, \beta) \right),
\]
\[
\Delta_2 = |\Delta(q, a, \beta)|^2.
\]
So if we define
\[
(7.1) \quad S_5 = \sum_{q \leq x/R} \left\{ \sum_{q \leq x/R}^{1/(2qR)} F_q(\beta)|J(\beta)|^2 \sum_{\substack{a = 1 \\ (a, q) = 1}}^q |\nu(q, a)|^2 \right. d\beta,
\]
then from (2.16) we have

\[(7.2) \quad S_4 = S_5 + \Sigma_1 + \Sigma_2,\]

where

\[
\Sigma_1 = \sum_{q \leq x/R} \int_{-1/(2qR)}^{1/(2qR)} F_q(\beta) \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \Delta_1 \, d\beta,
\]

\[
\Sigma_2 = \sum_{q \leq x/R} \int_{-1/(2qR)}^{1/(2qR)} F_q(\beta) \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \Delta_2 \, d\beta.
\]

By (2.13) and (4.25), we have

\[
\Sigma_2 \ll \sum_{q \leq x/R} \int_{-1/(2qR)}^{1/(2qR)} \frac{x \log(2\sqrt{x}/q)}{q + qx|\beta|} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} (x^{4/3} q^{1+2\varepsilon} (1 + x|\beta|)^2) \, d\beta
\]

\[
\ll \sum_{q \leq x/R} \int_{-1/(2qR)}^{1/(2qR)} x^{7/3} (\log x) q^{1+2\varepsilon} (1 + x|\beta|) \, d\beta
\]

\[
\ll x^{10/3+2\varepsilon} (\log x)^2 R^{-2}.
\]

Moreover, if we also apply (4.18) and (4.21), we have

\[
\Sigma_1 \ll \sum_{q \leq x/R} \int_{-1/(2qR)}^{1/(2qR)} \frac{x \log(2\sqrt{x}/q)}{q + qx|\beta|} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \frac{x \cdot x^{2/3} q^{1/2+\varepsilon} (1 + x|\beta|)}{(1 + x|\beta|)q} \, d\beta
\]

\[
\ll \sum_{q \leq x/R} \int_{-1/(2qR)}^{1/(2qR)} x^{8/3} (\log x) q^{-1/2+\varepsilon} \cdot \frac{d\beta}{1 + x|\beta|}
\]

\[
\ll x^{13/6+\varepsilon} (\log x)^2 R^{-1/2}.
\]

Hence by (7.2),

\[(7.3) \quad S_4 = S_5 + O(x^{10/3+2\varepsilon} (\log x)^2 R^{-2} + x^{13/6+\varepsilon} (\log x)^2 R^{-1/2}).\]

To estimate \(S_5\), note that by (4.16) and (4.11),

\[
\sum_{\substack{a=1 \\ (a,q)=1}}^{q} |\nu(q,a)|^2 = \frac{1}{q^6} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |S(q,a)|^6 = qT(q),
\]

so

\[
S_5 = \sum_{q \leq x/R} q^{1/(2qR)} \int_{-1/(2qR)}^{1/(2qR)} F_q(\beta) |J(\beta)|^2 \, d\beta.
\]
Define
\[ S_{11/2} = \sum_{q \leq x/R} \left( \int_{-1/2}^{1/2} qT(q) F_q(\beta) |J(\beta)|^2 d\beta \right)^{1/2}, \]
\[ S_6 = \sum_{q \leq \sqrt{x}} \left( \int_{-1/2}^{1/2} qT(q) F_q(\beta) |J(\beta)|^2 d\beta \right)^{1/2}. \]

By (2.13), we have a crude estimate \( F_q(\beta) \ll q^{-1} x \log x \). Thus by Lemma 4.9 (4.20) and (4.22),
\[ |S_{11/2} - S_5| \leq \sum_{q \leq x/R} \left( \int_{-1/2}^{1/2} qT(q) \left( \int_{-1/2}^{1/2} |F_q(\beta)| |J(\beta)|^2 d\beta \right)^{1/2} \right)^{1/2} \ll x(\log x) \sum_{q \leq x/R} q^{-2} \int_{1/(2qR)}^{1} \frac{1}{\beta^2} d\beta \ll Rx(\log x)^2, \]
and
\[ |S_6 - S_{11/2}| \leq \sum_{x/R < q \leq \sqrt{x}} qT(q) \left( \int_{-1/2}^{1/2} |F_q(\beta)| |J(\beta)|^2 d\beta \right)^{1/2} \ll x(\log x) \sum_{x/R < q \leq \sqrt{x}} q^{-2} [x] \ll Rx(\log x)^2. \]

Hence
\[ S_5 = S_6 + O(Rx(\log x)^2). \]

Therefore, by (3.1), (5.7), (6.2), (7.3) and (7.6),
\[ V(x, Q) = 2S_6 - C_0 x^2 \log Q + (C_1 - \gamma C_0) x^2 + Q \sum_{n \leq x} r_3(n)^2 + U(x, Q), \]
where
\[ U(x, Q) \ll R(\log x) \sum_{n \leq x} r_3(n)^2 + \frac{x^{10/3 + 2\varepsilon}(\log x)^2}{R^2} + \frac{x^{13/6 + \varepsilon}(\log x)^2}{R^{1/2}} + Rx(\log x)^2 + x^5 Q^{\varepsilon} + \frac{x^2 \log Q}{Q}. \]
8. The optimal choice for $R$. By (7.5) and (7.7), the main term of $V(x, Q)$ does not depend on $R$. Therefore, to find the optimal choice for $R$, we only consider the error term $U(x, Q)$ such that the RHS of (7.8) has the minimum size. In other words, the error term has the minimum upper exponent of $x$ defined by (3.6), which equals the minimum exponent of $x$ defined by (3.8) if it exists.

Assume that the exponent of $x$ in $R$ exists and equals $B$. Then the upper exponents of $x$ in the first five terms on the RHS of (7.8) are $B + A^+$, $10/3 - 2B$, $13/6 - B/2$, $B + 1$ and $5/3$ respectively, where $A^+$ is the upper exponent of (3.5) with the range given by (3.9). The last term is dominated by the fifth due to the assumption (2.19):

$$
\frac{x^2 \log Q}{Q} \leq \frac{x^2 \log x}{x^{1/2} \log x} \ll x^{5/3} Q^\varepsilon.
$$

Hence the upper exponent of $x$ on RHS of (7.8) is

$$I(B) = \max \{B + A^+, 10/3 - 2B, 13/6 - B/2, B + 1, 5/3\}.$$

If $A^+ \in [1, 7/6]$, then the function $I(B)$ has the minimum value when

(8.1) \quad B = 10/9 - A^+/3.

By (3.9), if $B$ satisfies (8.1) and $x$ is large enough, then $R$ fulfills the general assumption (2.9). As the behavior of (3.5) and the values of $A^+$ and $A$ if it exists are still unknown, instead of finding a precise expression, we compare different expressions for $R$ whose exponent satisfies (8.1). One expression is the exponential function

(8.2) \quad R = x^{10/9 - A^+/3}.

From the definition of upper exponent, for any $\varepsilon > 0$ we have

$$\sum_{n \leq x} r_3(n)^2 \ll x^{A^+ + \varepsilon},$$

so a straightforward calculation shows that

(8.3) \quad U(x, Q) \ll x^{10/9 + 2A^+/3 + 4\varepsilon}.

Generally, we use $\varepsilon$ instead of $4\varepsilon$ in the expression above, since both represent an arbitrary positive number. If the exponent of $x$ in (3.5) exists, then $A = A^+$, and we can replace $A^+$ by $A$ in (8.2) and (8.3).

To avoid using $A$ when it exists, we replace $x^{A^+} = x^A$ by the second moment sum (3.5) in (8.2) and (8.3) to get another expression:

(8.4) \quad R = x^{10/9} \left( \sum_{n \leq x} r_3(n)^2 \right)^{-1/3}.
and the error term satisfies

\begin{equation}
U(x, Q) \ll x^{10/9 + \varepsilon} \left( \sum_{n \leq x} r_3(n)^2 \right)^{2/3}.
\end{equation}

Both expressions have their own advantages: \((8.5)\) shows the relationship between the error term and the second moment sum, while the other expression does not show it directly. However, \((8.3)\) is always true, whether \(A\) exists or not, while the other expression might be wrong if \(A\) does not exist. Therefore, both expressions will be included in the conclusions. In the future, if we know how the second moment sum \((3.5)\) behaves, we can choose the better expression between them.

9. The main term \(S_6\). To finalize the conclusion, we need to calculate the integral \((7.5)\). By \((2.8)\) and \((4.17)\),

\[\int_{-1/2}^{1/2} F_q(\beta) |J(\beta)|^2 \, d\beta = \Gamma(4/3)^6 \sum_{q \mid l} \left( \sum_{m \leq x/l} + \sum_{\sqrt{x} < m \leq \min(Q, x/l)} \right) ([x] - lm),\]

and a straightforward calculation shows that

\[\left( \sum_{m \leq x/l} + \sum_{\sqrt{x} < m \leq \min(Q, x/l)} \right) ([x] - lm)\]

is

\[\frac{x^2}{2l} + \frac{x}{2l} (\sqrt{x} - l)^2 - \frac{Q^2}{2l} \left( \frac{x}{Q} - l \right)^2 + O(x)\]

when \(l \leq x/Q\), and is

\[\frac{x^2}{2l} + \frac{x}{2l} (\sqrt{x} - l)^2 + O(x)\]

when \(x/Q < l \leq \sqrt{x}\). For convenience, we write it as \(K(x,l,Q)\); then by \((7.5)\),

\[S_6 = \Gamma(4/3)^6 \sum_{q \leq \sqrt{x}} qT(q) \sum_{l \leq \sqrt{x}} K(x, l, Q)\]

\[= \Gamma(4/3)^6 \sum_{l \leq \sqrt{x}} \left( \sum_{q \mid l} qT(q) \right) K(x, l, Q)\]

Define

\begin{equation}
h(l) = \sum_{q \mid l} qT(q).
\end{equation}
Then
\[ 2S_6 = 2\Gamma(4/3)^6 \left( \sum_{l \leq x/Q} h(l)K(x, l, Q) + \sum_{x/Q < l \leq \sqrt{x}} h(l)K(x, l, Q) \right) \]
\[ = \Gamma(4/3)^6 \left( x^2 \sum_{l \leq \sqrt{x}} \frac{h(l)}{l} + xW(\sqrt{x}) - Q^2W\left(\frac{x}{Q}\right) \right) \]
\[ + O \left( x \sum_{l \leq \sqrt{x}} h(l) \right), \]
where
\[ W(X) = \sum_{l \leq X} \frac{h(l)}{l} (X - l)^2. \]

By Lemma 4.9 and (9.1), the error term is
\[ \ll x \sum_{l \leq \sqrt{x}} \sum_{q | l} qT(q) = x \sum_{q \leq \sqrt{x}} qT(q) \sum_{l \leq \sqrt{x}} 1 \leq x^{3/2} \sum_{q \leq \sqrt{x}} T(q) \ll x^{3/2}. \]

Moreover, the first term is
\[ \Gamma(4/3)^6 x^2 \sum_{l \leq \sqrt{x}} \frac{h(l)}{l} = \Gamma(4/3)^6 x^2 \sum_{l \leq \sqrt{x}} \frac{1}{l} \sum_{q | l} qT(q), \]
which looks similar to (5.1), an expression that contains $S_3$. To calculate this, we repeat a similar argument to calculating $S_3$, and we have
\[ \Gamma(4/3)^6 x^2 \sum_{l \leq \sqrt{x}} \frac{h(l)}{l} = \frac{1}{2} C_0 x^2 \log x + (\gamma C_0 - C_1) x^2 + O(x^{3/2} \log x), \]
where the constants $C_0$ and $C_1$ are defined by (5.5) and (5.6). Hence by (9.2) and (9.4),
\[ 2S_6 = \frac{1}{2} C_0 x^2 \log x + (\gamma C_0 - C_1) x^2 \]
\[ + \Gamma(4/3)^6 \left( xW(\sqrt{x}) - Q^2W\left(\frac{x}{Q}\right) \right) + O(x^{3/2} \log x). \]

Therefore, we need to calculate $W(X)$ when $X = \sqrt{x}$ and $X = x/Q$.

10. Conclusion. We start from the following lemma.

**Lemma 10.1.** Assume that $s = \sigma + it$ is a complex number, where $\sigma, t \in \mathbb{R}$ and $\sigma > -2$. Then the following results hold:

(i) If $p \neq 3$ is a prime number, then
\[ \sum_{k=0}^{\infty} \frac{T(p^k)}{p^{ks}} = \frac{1}{1 - p^{-(3s+6)}} \left( 1 + \frac{1}{p^{s+\gamma}} \sum_{c=1}^{p-1} |S(p, c)|^6 + \frac{p-1}{p^{2s+\gamma}} - \frac{1}{p^{3s+\gamma}} \right). \]
(ii) We have
\[
\sum_{k=0}^{\infty} \frac{T(3^k)}{3^{ks}} = \frac{1}{1 - 3^{-(3s+6)}} \left( 1 + \frac{1}{3^{2s+14}} \sum_{c=1}^{9} |S(9, c)|^6 - \frac{1}{3^{3s+7}} \right).
\]

Proof. By the proof of Lemma 4.8, if \( p \) is prime, \( (p, a) = 1 \) and \( u \in \mathbb{Z} \), \( u > 0 \), then
\[
S(p^{3u+3}, a) = p^{2u+2},
\]
\[
S(p^{3u+2}, a) = \begin{cases} p^{2u+1} & \text{when } p \neq 3, \\ 3^{2u} S(9, a) & \text{when } p = 3, \end{cases}
\]
\[
S(p^{3u+1}, a) = p^{2u} S(p, a).
\]

So by (4.11), we have
\[
T(p^{3u+3}) = \frac{p - 1}{p^{6u+7}},
\]
\[
T(p^{3u+2}) = \begin{cases} p - 1 & \text{when } p \neq 3, \\ \frac{1}{3^{6u+14}} \sum_{c=1}^{9} |S(9, c)|^6 & \text{when } p = 3, \end{cases}
\]
\[
T(p^{3u+1}) = \frac{1}{p^{6u+7}} \sum_{c=1}^{p-1} |S(p, c)|^6,
\]
in particular \( T(3^{3u+1}) = 0 \) since \( S(3, 1) = S(3, 2) = 0 \) by Corollary 4.3. Note that the identity
\[
(10.1) \quad \sum_{k=0}^{\infty} T(p^k) / p^{ks} = 1 + \sum_{u=0}^{\infty} T(p^{3u+1}) / p^{(3u+1)s} + \sum_{u=0}^{\infty} T(p^{3u+2}) / p^{(3u+2)s} + \sum_{u=0}^{\infty} T(p^{3u+3}) / p^{(3u+3)s}
\]
holds when the series on the RHS are convergent. First assume that \( p \neq 3 \). Then from the discussion above, the RHS of (10.1) becomes
\[
1 + \left( \frac{1}{p^{s+7}} \sum_{c=1}^{p-1} |S(p, c)|^6 + \frac{p - 1}{p^{2s+7}} + \frac{p - 1}{p^{3s+7}} \right) \sum_{u=0}^{\infty} \frac{1}{p^{(3s+6)u}}.
\]
If \( \sigma = \Re s > -2 \), then the series above is absolutely convergent since \( |p^{3s+6}| = p^{3\sigma+6} > 1 \), and it converges to \( (1 - p^{-3(s+6)})^{-1} \). Hence the first result is proved.
Now assume that $p = 3$. Then the RHS of (10.1) becomes
\[
1 + \left( \frac{1}{3^{2s+14}} \sum_{c=1}^{9} \frac{|S(9, c)|^6}{(c, 3) = 1} + \frac{2}{3^{3s+7}} \right) \sum_{u=0}^{\infty} \frac{1}{3^{(3s+6)u}}.
\]
Similarly, the series above converges to $(1 - 3^{-(3s+6)})^{-1}$ when $\sigma > -2$, and thus the second result is also proved.

Now define
\[
D(s) = \sum_{l=1}^{\infty} \frac{h(l)}{l} \cdot \frac{1}{l^s} = \sum_{l=1}^{\infty} h(l) \cdot \frac{1}{l^{s+1}}
\]
if $\sigma = \Re s > 0$. Then by (9.1) and Lemmata 4.10 and 10.1 we have
\[
D(s) = \sum_{l=1}^{\infty} \frac{1}{l^{s+1}} \sum_{q|l} qT(q)
\]
(10.3)
\[
= \zeta(s + 1) \sum_{q=1}^{\infty} \frac{T(q)}{q^s}
\]
\[
= \zeta(s + 1) \prod_p \left( \sum_{k=0}^{\infty} \frac{T(p^k)}{p^{ks}} \right)
\]
(10.4)
\[
= \zeta(s + 1) \zeta(3s + 6) D_0(s),
\]
where
\[
D_0(s) = \prod_{p \neq 3} \left( 1 + \frac{1}{p^{s+7}} \sum_{c=1}^{p-1} |S(p, c)|^6 + \frac{p - 1}{p^{2s+7}} - \frac{1}{p^{3s+7}} \right)
\]
(10.5)
\[
\cdot \left( 1 + \frac{1}{3^{2s+14}} \sum_{c=1}^{9} \frac{|S(9, c)|^6}{(c, 3) = 1} \right).
\]
By Corollary 4.3, we have
\[
|D_0(s)| \leq \prod_{p \neq 3} \left( 1 + \frac{64}{p^{\sigma+3}} + \frac{1}{p^{2\sigma+6}} + \frac{1}{p^{3\sigma+7}} \right) \cdot \left( 1 + \frac{2}{3^{2\sigma+1}} + \frac{1}{3^{3\sigma+7}} \right).
\]
The RHS above converges locally uniformly when $\sigma > -2$, so $D_0(s)$ converges absolutely and locally uniformly for $\Re s > -2$. Note that Lemma 10.1 also holds when $\Re s > -2$, so (10.4) affords an analytic continuation of $D(s)$ to that open half-plane. Hence $D(s)$ is meromorphic in $\Re s > -2$ with simple poles at $s = 0$ and $s = -5/3$. Moreover, if $\Re s > -5/3$, then when comparing
with \( \zeta(3)s+6 \) \( D_0(s) \).

This identity is especially useful when \( s = -1 \) or \( s = 0 \), which shows that the series
\[
\sum_{q=1}^{\infty} qT(q) \quad \text{and} \quad \sum_{q=1}^{\infty} T(q)
\]
converge to \( \zeta(3)D_0(-1) \) and \( \zeta(6)D_0(0) \) respectively.

The following identity shows the relationship between \( W(X) \) and \( D(s) \):
\[
\frac{1}{2} W(X) = \frac{1}{2\pi i} \int_{\sigma_0-i\infty}^{\sigma_0+i\infty} D(s) \cdot \frac{X^{s+2}}{s(s+1)(s+2)} \, ds,
\]
with a real constant \( \sigma_0 > 0 \), since \( D(s) \) has an abscissa of convergence \( \sigma_c \leq 0 \). We calculate the integral on the RHS. Note that the integrand is holomorphic in the open half-plane \( \Re s > -2 \) except for simple poles at both \( s = -1 \) and \( s = -5/3 \), and a double pole at \( s = 0 \). Let \( 0 < \delta < 1/3 \), and let \( \mathcal{C}(t_1, t_2) \) be the counterclockwise rectangular contour with vertices \( \sigma_0 - it_1, \sigma_0 + it_2, -2 + \delta + it_2 \) and \( -2 + \delta - it_1 \) where \( t_1, t_2 > 1 \). Then we have
\[
\frac{1}{2\pi i} \int_{\mathcal{C}(t_1, t_2)} D(s)X^{s+2} \frac{ds}{s(s+1)(s+2)} = \text{Res}(f, 0) + \text{Res}(f, -1) + \text{Res}(f, -5/3),
\]
where \( f \) represents the integrand. The residues are
\[
\text{Res}(f, 0) = \lim_{s \to 0} \frac{d}{ds} \left( \frac{s^2 D(s)X^{s+2}}{s(s+1)(s+2)} \right) = D_1X^2 \log X + D_2X^2,
\]
\[
\text{Res}(f, -1) = \lim_{s \to -1} \frac{(s+1)D(s)X^{s+2}}{s(s+1)(s+2)} = D_3X,
\]
\[
\text{Res} \left( f, -\frac{5}{3} \right) = \lim_{s \to -5/3} \frac{(s+5/3)D(s)X^{s+2}}{s(s+1)(s+2)} = D_4X^{1/3},
\]
where \( D_1, D_2, D_3, D_4 \) are constants. We calculate them, except \( D_2 \), from \( \zeta(3)s+6 \) \( D_0(s) \):
\[
D_1 = \lim_{s \to 0} \frac{sD(s)}{(s+1)(s+2)} = \frac{1}{2} \zeta(6)D_0(0) = \frac{1}{2} \sum_{q=1}^{\infty} T(q),
\]
\[
D_3 = \lim_{s \to -1} \frac{D(s)}{s(s+2)} = -\zeta(0)\zeta(3)D_0(-1) = \frac{1}{2} \sum_{q=1}^{\infty} qT(q),
\]
\[
D_4 = \lim_{s \to -5/3} \frac{(s+5/3)D(s)}{s(s+1)(s+2)} = \frac{9}{10} \zeta(-2/3)D_0(-5/3).
\]
Hence the RHS of (10.8) is

\[(10.9)\quad \frac{1}{2} \left( \sum_{q=1}^{\infty} T(q) \right) X^2 \log X + D_2 X^2 + \frac{1}{2} \left( \sum_{q=1}^{\infty} q T(q) \right) X + \frac{9}{10} \zeta(-2/3) D_0(-5/3) X^{1/3}.\]

Now consider the LHS. First, by (10.4), we rewrite the integrand as

\[(10.10)\quad \frac{\zeta(s+1) \zeta(3s+6) D_0(s) X^{s+2}}{s(s+1)(s+2)}.\]

Note that \(D_0(s)\) is uniformly bounded when \(\Re s \geq -2 + \delta\), so the behavior of the integrand on the contour \(C(t_1, t_2)\) largely depends on \(\zeta(s+1)\) and \(\zeta(3s+6)\). Therefore, we need the following lemma, which can be proved from the results of Titchmarsh \([T1, \text{Chapter 5, Section 5.1, pp. 95–96}]\):

**Lemma 10.2.** If \(w\) is a complex number, then for any \(\varepsilon > 0\),

\[(10.11)\quad \zeta(w) - \frac{1}{w-1} \ll \tau^{\lambda(u)+\varepsilon},\]

where \(u = \Re w, v = \Im w, \tau = 4 + |v|,\) and

\[(10.12)\quad \lambda(u) = \begin{cases} 0 & \text{when } u > 1, \\ \frac{1}{2} - \frac{1}{2}u & \text{when } 0 < u \leq 1, \\ \frac{1}{2} - u & \text{when } u \leq 0. \end{cases}\]

Consider the integral along the line segment from \(\sigma_0 + it_2\) to \(-2 + \delta + it_2\). If \(s\) lies on this segment, then \(\Re(s+1) > -1, \Re(3s+6) > 0,\) and both \(|(s+1) - 1|^{-1}\) and \(|(3s+6) - 1|^{-1}\) are bounded by 1. So by Lemma 10.2

\[\zeta(s+1) \ll (4 + t_2)^{3/2+\varepsilon},\]

\[\zeta(3s+6) \ll (4 + 3t_2)^{1/2+\varepsilon}.\]

Therefore, as \(t_2 > 1,\) we have

\[\frac{\zeta(s+1) \zeta(3s+6) D_0(s) X^{s+2}}{s(s+1)(s+2)} \ll t_2^{-1+2\varepsilon} X^{\sigma_0+2}.\]

Hence if \(\varepsilon\) is small enough and \(X, \sigma_0\) are fixed, then the integral along this line segment tends to 0 when \(t_2 \to \infty.\) Similarly, with the same assumptions, the integral along the line segment from \(-2 + \delta - it_1\) to \(\sigma_0 - it_1\) also tends to 0 when \(t_1 \to \infty.\)

Now consider the integral along the line segment from \(-2 + \delta + it_2\) to \(-2 + \delta - it_1\). We have
\begin{equation}
(10.13) \quad \frac{1}{2\pi i} \int_{-2+\delta+it_2}^{-2+\delta-it_1} \frac{\zeta(s+1)\zeta(3s+6)D_0(s)X^{s+2}}{s(s+1)(s+2)} \, ds
= -\frac{1}{2\pi} X^\delta \left( \int_{-t_1}^{-1} f_0(t) \, dt + \int_{-1}^{1} f_0(t) \, dt + \int_{1}^{t_2} f_0(t) \, dt \right),
\end{equation}

where

\begin{equation}
(10.14) \quad f_0(t) = \frac{\zeta(-1+\delta+it)\zeta(3\delta+3it)D_0(-2+\delta+it)X^{it}}{(-2+\delta+it)(-1+\delta+it)(\delta+it)}.
\end{equation}

For convenience, let \( I_1, I_2 \) and \( I_3 \) be the integrals on the RHS of (10.13). If \(|t| \geq 1\), then by Lemma 10.2, \( \zeta(-1+\delta+it) \ll (4+|t|)^{3/2-\delta+\epsilon}, \quad \zeta(3\delta+3it) \ll (4+3|t|)^{1/2-3\delta/2+\epsilon} \).

Moreover, if we let \( \epsilon = \delta \), then (10.14) shows that

\[ f_0(t) \ll |t|^{-1-\delta/2}, \]

since \( D_0(-2+\delta+it) \) is bounded. So for a fixed \( \delta \), we have

\[ I_1 \ll 1 - t_1^{-\delta/2}, \quad I_3 \ll 1 - t_2^{-\delta/2}. \]

If \(|t| \leq 1\), then \( f_0(t) \) is bounded, so

\[ I_2 \ll 1. \]

Hence by (10.13), we have

\begin{equation}
(10.15) \quad \lim_{t_1 \to \infty} \lim_{t_2 \to \infty} \frac{1}{2\pi i} \int_{-2+\delta+it_2}^{-2+\delta-it_1} \frac{\zeta(s+1)\zeta(3s+6)D_0(s)X^{s+2}}{s(s+1)(s+2)} \, ds = O(X^\delta).
\end{equation}

Letting \( t_1, t_2 \to \infty \) on both sides of (10.8), by (10.9), (10.7) and (10.15), we have

\begin{equation}
(10.16) \quad W(X) = \left( \sum_{q=1}^{\infty} T(q) \right) X^2 \log X + 2D_2X^2 + \left( \sum_{q=1}^{\infty} qT(q) \right) X
+ \frac{9}{5} \zeta(-2/3)D_0(-5/3)X^{1/3} + O(X^\delta).
\end{equation}

Therefore, applying \( X = \sqrt{x} \) and \( X = x/Q \), we have

\begin{equation}
(10.17) \quad \Gamma(4/3)^6 x W(\sqrt{x}) = \frac{1}{2} C_0 x^2 \log x + 2\Gamma(4/3)^6 D_2 x^2 + O(x^{3/2}),
\end{equation}

and

\begin{equation}
(10.18) \quad \Gamma(4/3)^6 Q^2 W\left( \frac{x}{Q} \right) = C_0 x^2 \log \frac{x}{Q} + 2\Gamma(4/3)^6 D_2 x^2
+ A_1 Q x - A_2 Q^{5/3} x^{1/3} + O(Q^{2-\delta} x^{\delta}),
\end{equation}

where

\[(10.19) \quad A_1 = \Gamma(4/3)^6 \sum_{q=1}^{\infty} qT(q),\]

\[(10.20) \quad A_2 = -\frac{9}{8} \Gamma(4/3)^6 \zeta(-2/3)D_0(-5/3)\]

are positive constants, and \(C_0\) is defined by (5.5). Finally, by (7.7), (9.5), (10.17), (10.18), and two approximations for the error term \(U(x, Q)\), namely (8.3) and (8.5), Theorem 1.1 is proved, where the positive constants \(A_1\) and \(A_2\) in the theorem are defined above.

11. Results for large \(Q\). Theorem 1.1 gives us a general result. However, if \(Q\) is very close to \(x\), say \(Q \approx x\), then \(Q^{2-\delta}x^\delta\) has the same size as the main terms, thus the conclusion is trivial. Therefore, it is essential to provide a more accurate conclusion for large \(Q\).

To avoid the error term \(O(Q^{2-\delta}x^\delta)\), we use a direct calculation instead of applying (10.18). Assume that

\[
\frac{x}{k+1} < Q \leq \frac{x}{k},
\]

where \(k \geq 1\) is an integer constant. Then by definition (9.3),

\[(11.1) \quad Q^2 W\left(\frac{x}{Q}\right) = \sum_{l \leq k} \frac{h(l)}{l} (x-lQ)^2.\]

Note that for a fixed integer \(k\), the number of terms on the RHS is limited. To calculate \(xW(\sqrt{x})\) and reach the conclusion, we may still apply (10.17). However, the constant \(D_2\) cannot be cancelled in this situation, so we have to calculate it, which is not easy if we still apply the residue theorem. Therefore, we follow a result provided by Vaughan [RV3]:

**Lemma 11.1.** If \(Y \geq 1\), then

\[(11.2) \quad \sum_{m \leq Y} \frac{1}{m} (Y - m)^2 = Y^2 \log Y + C_2 Y^2 + Y + O(1),\]

where

\[(11.3) \quad C_2 = \frac{-11}{12} - 2 \int_{1}^{\infty} \frac{B_2(u)}{u^3} \, du,\]

\[(11.4) \quad B_2(u) = \frac{1}{2} (u - \lfloor u \rfloor)^2 - \frac{1}{2} (u - \lfloor u \rfloor) + \frac{1}{12}.\]

Let \(Y = \sqrt{x}/q\). Then by Lemma 4.9 and definitions (9.3) and (9.1), we have
\[
W(\sqrt{x}) = \sum_{l \leq \sqrt{x}} \frac{1}{l}(\sqrt{x} - l)^2 \sum_{q | l} q T(q)
\]
\[
= \sum_{q \leq \sqrt{x}} q^2 T(q) \sum_{r \leq \sqrt{x}/q} \frac{1}{r} \left( \frac{\sqrt{x}}{q} - r \right)^2
\]
\[
= \left( \frac{1}{2} x \log x + C_2 x \right) \sum_{q \leq \sqrt{x}} T(q) - x \sum_{q \leq \sqrt{x}} T(q) \log q
\]
\[
+ \sqrt{x} \sum_{q \leq \sqrt{x}} q T(q) + O(\sqrt{x}).
\]

Replacing \(Q\) by \(\sqrt{x}\) in (5.2)–(5.4), and repeating a similar argument to calculating \(S_3\), we have

\[
W(\sqrt{x}) = \left( \frac{1}{2} x \log x + C_2 x \right) \sum_{q=1}^{\infty} T(q) - x \sum_{q=1}^{\infty} T(q) \log q + O(\sqrt{x} \log x).
\]

Hence

(11.5) \( \Gamma(4/3)^6 x W(\sqrt{x}) = \frac{1}{2} C_0 x^2 \log x + (C_0 C_2 - C_1) x^2 + O(x^{3/2} \log x). \)

Note that comparing (11.5) with (10.17), we have

(11.6) \( D_2 = \frac{1}{2} \left( C_2 \sum_{q=1}^{\infty} T(q) - \sum_{q=1}^{\infty} T(q) \log q \right). \)

Finally, by (7.7), (9.5), (11.1), (11.5), and approximations (8.3) and (8.5), we can prove Theorem 1.2 where \(h(l)\) is defined by (9.1), and the constants \(C_0, C_1\) and \(C_2\) are defined by (5.5), (5.6) and (11.3) respectively. Note that by (3.9), we have

\[
U_0(x, Q) \ll x^{10/9+\varepsilon} (x^{7/6})^{2/3} = x^{17/9+\varepsilon},
\]

hence Theorem 1.2 is more accurate than Theorem 1.1 when \(Q\) is large.

The corollaries below can be proved by straightforward calculations. First, if \(Q\) and \(x\) are linearly dependent, then we have:

**Corollary 11.2.** Let \(h(l)\) be defined by (9.1), and let \(C_0, C_1\) and \(C_2\) be the constants defined by (5.5), (5.6) and (11.3) respectively. Then

\[
V\left( x, \frac{x}{m} \right) = x^2 \left( C_0 \log m + C_0 C_2 - C_1 - \Gamma(4/3)^6 \frac{1}{m^2} \sum_{l \leq m} \frac{h(l)}{l} (m - l)^2 \right)
\]
\[\quad + \frac{x}{m} \sum_{n \leq x} r_3(n)^2 + U_0(x),\]

where \(m \geq 1\) is a constant, and \(U_0(x)\) has the same property as \(U_0(x, Q)\) in Theorem 1.2.
Sometimes, we need to consider the case when $Q$ is close to $x$, for example, when $Q > x/3$. We have:

**Corollary 11.3.** Let $C_0$, $C_1$ and $C_2$ be constants defined by (5.5), (5.6) and (11.3) respectively. Then

(i) if $x/3 < Q \leq x/2$ one has

$$V(x, Q) = Q \sum_{n \leq x} r_3(n)^2 + x^2 \left( C_0 \log \frac{x}{Q} + C_0 C_2 - C_1 - \frac{3}{2} \Gamma(4/3)^6 \right) + 4 \Gamma(4/3)^6 Q x - 3 \Gamma(4/3)^6 Q^2 + U_0(x, Q),$$

(ii) if $x/2 < Q \leq x$ one has

$$V(x, Q) = Q \sum_{n \leq x} r_3(n)^2 + x^2 \left( C_0 \log \frac{x}{Q} + C_0 C_2 - C_1 - \Gamma(4/3)^6 \right) + 2 \Gamma(4/3)^6 Q x - \Gamma(4/3)^6 Q^2 + U_0(x, Q),$$

where $U_0(x, Q)$ has the same property as in Theorem 1.2.

Finally, let $Q = x$, the largest possible value. We have:

**Corollary 11.4.** Let $C_0$, $C_1$ and $C_2$ be the constants defined by (5.5), (5.6) and (11.3) respectively. Then

$$V(x, x) = x \sum_{n \leq x} r_3(n)^2 + x^2 (C_0 C_2 - C_1) + U_0(x),$$

where $U_0(x)$ has the same property as $U_0(x, Q)$ in Theorem 1.2.

**Acknowledgements.** This paper is the main part of my 2021 Ph.D. thesis, written under the supervision of Robert Vaughan at Pennsylvania State University. I would like to thank Professor Vaughan for his help with this work, as well as Professor Ae Ja Yee and Professor Wen-Ching Winnie Li for their suggestions on finalizing my thesis.

**References**


Pengyong Ding
School of Mathematics
Shandong University
250100 Jinan, P.R. China
E-mail: dlfw3@sdu.edu.cn