ON SUMS OF THREE SQUARES

S.K.K. CHOI, A.V. KUMCHEV, AND R. OSBURN

Abstract. Let \( r_3(n) \) be the number of representations of a positive integer \( n \) as a sum of three squares of integers. We give two distinct alternative proofs of a conjecture of Wagon concerning the asymptotic value of the mean square of \( r_3(n) \).

1. Introduction

Problems concerning sums of three squares have a rich history. It is a classical result of Gauss that

\[ n = x_1^2 + x_2^2 + x_3^2 \]

has a solution in integers if and only if \( n \) is not of the form \( 4^a(8k + 7) \) with \( a, k \in \mathbb{Z} \). Let \( r_3(n) \) be the number of representations of \( n \) as a sum of three squares (counting signs and order). It was conjectured by Hardy and proved by Bateman [1] that

\[ r_3(n) = 4\pi n^{1/2} \Xi_3(n), \quad (1) \]

where the singular series \( \Xi_3(n) \) is given by (16) with \( Q = \infty \).

While in principle this exact formula can be used to answer almost any question concerning \( r_3(n) \), the ensuing calculations can be tricky because of the slow convergence of the singular series \( \Xi_3(n) \). Thus, one often sidesteps (1) and attacks problems involving \( r_3(n) \) directly. For example, concerning the mean value of \( r_3(n) \), one can adapt the method of solution of the circle problem to obtain the following

\[ \sum_{n \leq x} r_3(n) \sim \frac{4}{3} \pi x^{3/2}. \]

Moreover, such a direct approach enables us to bound the error term in this asymptotic formula. An application of a result of Landau [9, pp. 200–218] yields

\[ \sum_{n \leq x} r_3(n) = \frac{4}{3} \pi x^{3/2} + O(x^{3/4+\epsilon}) \]

for all \( \epsilon > 0 \), and subsequent improvements on the error term have been obtained by Vinogradov [19], Chamizo and Iwaniec [3], and Heath-Brown [6].

In this note we consider the mean square of \( r_3(n) \). The following asymptotic formula was conjectured by Wagon and proved by Crandall; see [4] (see also [2]).
Theorem. Let \( r_3(n) \) be the number of representations of a positive integer \( n \) as a sum of three squares of integers. Then

\[
\sum_{n \leq x} r_3(n)^2 \sim \frac{8\pi^4}{21\zeta(3)} x^2.
\]

Apparently, at the time they proposed this conjecture Crandall and Wagon were unaware of the earlier work of Müller [11, 12]. He obtained a more general result which, in a special case, gives

\[
\sum_{n \leq x} r_3(n)^2 = Bx^2 + O(x^{14/9}),
\]

where \( B \) is a constant. However, since in Müller’s work \( B \) arises as a specialization of a more general (and more complicated) quantity, it is not immediately clear that \( B = \frac{8}{21}\pi^4/\zeta(3) \). The purpose of this paper is to give two distinct proofs of this fact: one that evaluates \( B \) in the form given by Müller and a direct proof using the Hardy–Littlewood circle method.

2. A direct proof: the circle method

Our first proof exploits the observation that the left side of (2) counts solutions of the equation

\[
m_1^2 + m_2^2 + m_3^2 = m_4^2 + m_5^2 + m_6^2
\]
in integers \( m_1, \ldots, m_6 \) with \(|m_j| \leq x\). This is exactly the kind of problem that the circle method was designed for.

Set \( N = \sqrt{x} \) and define

\[
f(\alpha) = \sum_{m \leq N} e(\alpha m^2),
\]

where \( e(z) = e^{2\pi iz} \). Then for an integer \( n \leq x \), the number \( r^*(n) \) of representations of \( n \) as a sum of three squares of positive integers is

\[
r^*(n) = \int_0^1 f(\alpha)^3 e(-an) d\alpha.
\]

Since \( r_3(n) = 8r^*(n) + O(r_2(n)) \), where \( r_2(n) \) is the number of representations of \( n \) as a sum of two squares, we have

\[
\sum_{n \leq x} r_3(n)^2 = 64 \sum_{n \leq x} r^*(n)^2 + O(x^{3/2+\epsilon}).
\]

Therefore, it suffices to evaluate the mean square of \( r^*(n) \). Let

\[
P = N/4 \quad \text{and} \quad Q = N^{1/2}.
\]

We introduce the sets

\[\mathcal{M}(q, a) = \{ \alpha \in [PN^{-2}, 1 + PN^{-2}] : |q\alpha - a| \leq PN^{-2} \}\]

and

\[\mathcal{M} = \bigcup_{q \leq Q} \bigcup_{1 \leq a \leq q,(a,q)=1} \mathcal{M}(q,a), \quad m = [PN^{-2}, 1 + PN^{-2}] \setminus \mathcal{M}.
\]
We have
\[ r^*(n) = \left( \int_{\mathbb{R}} + \int_{\mathbb{M}} \right) f(\alpha)^3 e(-\alpha n) \, d\alpha = r^*(n, \mathbb{M}) + r^*(n, m), \quad \text{say.} \]

We now proceed to approximate the mean square of \( r^*(n) \) by that of \( r^*(n, \mathbb{M}) \). By (4) and Cauchy’s inequality,
\[ \sum_{n \leq x} r^*(n)^2 = \sum_{n \leq x} r^*(n, \mathbb{M})^2 + O\left( (\Sigma_1 \Sigma_2)^{1/2} + \Sigma_2 \right), \]
where
\[ \Sigma_1 = \sum_{n \leq x} |r^*(n, \mathbb{M})|^2, \quad \Sigma_2 = \sum_{n \leq x} |r^*(n, m)|^2. \]

By Bessel’s inequality,
\[ |\Sigma_2| = \sum_{n \leq x} \left| \int_{\mathbb{M}} f(\alpha)^3 e(-\alpha n) \, d\alpha \right|^2 \leq \int_{\mathbb{M}} |f(\alpha)|^6 \, d\alpha. \]

By Dirichlet’s theorem of diophantine approximation, we can write any real \( \alpha \) as \( \alpha = a/q + \beta \), where
\[ 1 \leq q \leq N^2 P^{-1}, \quad (a, q) = 1, \quad |\beta| \leq P/(qN^2). \]
When \( \alpha \in \mathbb{M} \), we have \( q \geq Q \), and hence Weyl’s inequality (see Vaughan [18, Lemma 2.4]) yields
\[ |f(\alpha)| \ll N^{1+\epsilon} (q^{-1} + N^{-1} + qN^{-2})^{1/2} \ll N^{1+\epsilon} Q^{-1/2}. \]

Furthermore, we have
\[ \int_0^1 |f(\alpha)|^4 \, d\alpha \ll N^{2+\epsilon}, \]
because the integral on the right equals the number of solutions of
\[ m_1^2 + m_2^2 = m_3^2 + m_4^2 \]
in integers \( m_1, \ldots, m_4 \leq N \). For each choice of \( m_1 \) and \( m_2 \), this equation has \( \ll N^\epsilon \) solutions.

Combining (6)–(8) and replacing \( \epsilon \) by \( \epsilon/3 \), we obtain
\[ \Sigma_2 \ll N^{4+\epsilon} Q^{-1}. \]

Furthermore, another appeal to Bessel’s inequality and appeals to (8) and to the trivial estimate \( |f(\alpha)| \leq N \) yield
\[ \Sigma_1 \leq \int_{\mathbb{M}} |f(\alpha)|^6 \, d\alpha \leq \int_0^1 |f(\alpha)|^6 \, d\alpha \ll N^{4+\epsilon}. \]

We now define a function \( f^* \) on \( \mathbb{M} \) by setting
\[ f^*(\alpha) = q^{-1} S(q, a) v(\alpha - a/q) \quad \text{for } \alpha \in \mathbb{M}(q, a) \subseteq \mathbb{M}; \]
here
\[ S(q, a) = \sum_{1 \leq h \leq q} e(ah^2/q), \quad v(\beta) = \frac{1}{2} \sum_{m \leq \beta} m^{-1/2} e(\beta m). \]
Our next goal is to approximate the mean square of \( r^*(n, \mathfrak{M}) \) by the mean square of the integral

\[
R^*(n) = \int_{\mathfrak{M}} f^*(\alpha)^3 e(-an) \, d\alpha.
\]

Similarly to (5),

\[
(11) \quad \sum_{n \leq x} r^*(n, \mathfrak{M})^2 = \sum_{n \leq x} R^*(n)^2 + O(\Sigma_3 + (\Sigma_4 \Sigma_5)^{1/2}),
\]

where

\[
(12) \quad \Sigma_3 = \sum_{n \leq x} \left| \int_{\mathfrak{M}} [f(\alpha)^3 - f^*(\alpha)^3] e(-an) \, d\alpha \right|^2 \leq \int_{\mathfrak{M}} |f(\alpha)^3 - f^*(\alpha)^3|^2 \, d\alpha,
\]

after yet another appeal to Bessel's inequality. By [18, Theorem 4.1], when \( \alpha \in \mathfrak{M}(q, a) \),

\[
f(\alpha) = f^*(\alpha) + O(q^{1/2+\varepsilon}).
\]

Thus,

\[
\int_{\mathfrak{M}(q,a)} |f(\alpha)^3 - f^*(\alpha)^3|^2 \, d\alpha \ll q^{1+2\varepsilon} \int_{\mathfrak{M}(q,a)} (|f(\alpha)|^4 + q^{2+4\varepsilon}) \, d\alpha,
\]

whence

\[
\int_{\mathfrak{M}} |f(\alpha)^3 - f^*(\alpha)^3|^2 \, d\alpha \ll Q^{1+2\varepsilon} \int_0^1 |f(\alpha)|^4 \, d\alpha + PQ^{4+6\varepsilon}N^{-2}.
\]

Bounding the last integral using (8) and substituting the ensuing estimate into (12), we obtain

\[
(13) \quad \Sigma_3 \ll QN^{2+2\varepsilon} + PQ^{4}N^{-2+3\varepsilon} \ll QN^{2+2\varepsilon}.
\]

Combining (5), (9)–(11), and (13), we deduce that

\[
(14) \quad \sum_{n \leq x} r^*(n)^2 = \sum_{n \leq x} R^*(n)^2 + O\left(N^{4+\varepsilon}Q^{-1/2} + N^{3+\varepsilon}Q^{1/2}\right).
\]

We now proceed to evaluate the main term in (14). We have

\[
\int_{\mathfrak{M}(q,a)} f^*(\alpha)^3 e(-an) \, d\alpha = q^{-3} S(q, a)^3 e(-an/q) \int_{\mathfrak{M}(q,0)} \nu(\beta)^3 e(-\beta n) \, d\beta,
\]

so

\[
R^*(n) = \sum_{q \leq Q} A(q, n) I(q, n),
\]

where

\[
A(q, n) = \sum_{\substack{1 \leq a \leq q \atop (a, q) = 1}} q^{-3} S(q, a)^3 e(-an/q), \quad I(q, n) = \int_{\mathfrak{M}(q,0)} \nu(\beta)^3 e(-\beta n) \, d\beta.
\]

Hence,

\[
(15) \quad \sum_{n \leq x} R^*(n)^2 = \sum_{n \leq x} I(n)^2 \Xi_3(n, Q)^2 + O\left((\Sigma_4 \Sigma_5)^{1/2} + \Sigma_5\right),
\]
We then use (17) and (18) to replace

\[ I(n) = \int_{-1/2}^{1/2} \sqrt{\beta} e(-\beta n) \, d\beta, \]

where

\[ \Xi_3(n, Q) = \sum_{q \leq Q} A(q, n), \quad I(n) = \int_{-1/2}^{1/2} \sqrt{\beta} e(-\beta n) \, d\beta, \]

\[ \Sigma_4 = \sum_{n \leq x} (\sum_{q \leq Q} |A(q, n)|^2)^2, \quad \Sigma_3 = \sum_{n \leq x} \left( \sum_{q \leq Q} |A(q, n)(I(n) - I(q, n))| \right)^2. \]

By [18, Theorem 2.3] and [18, Theorem 4.2],

\[ I(n) = \Gamma(3/2)^2 \sqrt{n} + O(1) = \frac{\pi}{4} \sqrt{n} + O(1), \quad A(q, n) \ll q^{-1/2}. \]

Furthermore, since \( A(q, n) \) is multiplicative in \( q \), [18, Lemma 4.7] yields

\[ \sum_{q \leq Q} |A(q, n)| \leq \prod_{p \leq Q} \left( 1 + |A(p, n)| + |A(p^2, n)| + \cdots \right) \ll \prod_{p \leq Q} \left( 1 + c_1(p, n)p^{-3/2} + 3c_1p^{-1} \right) \ll (nQ)^{e}, \]

where \( c_1 > 0 \) is an absolute constant. In particular, we have

\[ \Sigma_4 \ll N^{4+\epsilon}. \]

We now turn to the estimation of \( \Sigma_3 \). By Cauchy’s inequality and the second bound in (17),

\[ \Sigma_3 \ll (\log Q) \sum_{n \leq x} \sum_{q \leq Q} |I(n) - I(n, q)|^2 \]

Another application of Bessel’s inequality gives

\[ \sum_{n \leq x} |I(n) - I(n, q)|^2 \leq \int_{\beta \leq L} |v(\beta)|^6 \, d\beta. \]

Using [18, Lemma 2.8] to estimate the last integral, we deduce that

\[ \Sigma_3 \ll \log Q \sum_{q \leq Q} (q^2N^4p^{-2} + 1) \ll N^2Q^{3+\epsilon}. \]

Substituting this inequality and (19) into (15), we conclude that

\[ \sum_{n \leq x} R^*(n)^2 = \sum_{n \leq x} I(n)^2 \Xi_3(n, Q)^2 + O(N^{3+\epsilon}Q^{3/2}). \]

We then use (17) and (18) to replace \( I(n) \) on the right side of (20) by \( \frac{\pi}{4} \sqrt{n} \). We get

\[ \sum_{n \leq x} I(n)^2 \Xi_3(n, Q)^2 = \frac{\pi^2}{16} \sum_{n \leq x} n \Xi_3(n, Q)^2 + O(N^{3+\epsilon}). \]

Together with (14) and (20), this leads to the asymptotic formula

\[ \sum_{n \leq x} r^*(n)^2 = \frac{\pi^2}{16} \sum_{n \leq x} n \Xi_3(n, Q)^2 + O(N^{4+\epsilon}Q^{-1/2} + N^{3+\epsilon}Q^{3/2}). \]
Finally, we evaluate the sum on the right side of (21). On observing that \( S_3(n) \) is in fact a real number, we have

\[
\sum_{n \leq t} S_3(n, Q)^2 = \sum_{q_1, q_2 \leq Q} \sum_{1 \leq a_1 \leq q_1} \sum_{1 \leq a_2 \leq q_2} (q_1 q_2)^{-3} S(q_1, a_1)^3 S(q_2, -a_2)^3 \sum_{n \leq t} e((a_1/q_1 - a_2/q_2)n).
\]

As the sum over \( n \) equals \( t + O(1) \) when \( a_1 = a_2 \) and \( q_1 = q_2 \) and \( O(q_1 q_2) \) otherwise, we get

\[
\sum_{n \leq t} S_3(n, Q)^2 = t \sum_{q \leq Q} \sum_{1 \leq a \leq q} q^{-6} |S(q, a)|^6 + O(\Sigma_6^2),
\]

where

\[
\Sigma_6 = \sum_{q \leq Q} \sum_{1 \leq a \leq q} q^{-2} |S(q, a)|^3 \ll Q^{3/2}.
\]

We find that

\[
\sum_{n \leq t} S_3(n, Q)^2 = B_1 t + O(t^{-1} + Q),
\]

with

\[
B_1 = \sum_{q=1}^{\infty} \sum_{1 \leq a \leq q} q^{-6} |S(q, a)|^6.
\]

Thus, by partial summation,

\[
\sum_{n \leq t} n S_3(n, Q)^2 = (B_1/2) x^2 + O(x^{-1} + Q^3).
\]

Combining this asymptotic formula with (21), we deduce that

\[
\sum_{n \leq t} r^2(n)^2 = \frac{\pi^2}{32} B_1 x^2 + O(x^{15/8+\epsilon}).
\]

Recalling (3), we see that (2) will follow if we show that

\[
B_1 = \frac{8 \zeta(2)}{7 \zeta(3)}.
\]

This, however, follows easily from the well-known formula (see [7, §7.5])

\[
|S(q, a)| = \begin{cases} \sqrt{q} & \text{if } q \equiv 1 \pmod{2}, \\ \sqrt{2q} & \text{if } q \equiv 0 \pmod{4}, \\ 0 & \text{if } q \equiv 2 \pmod{4}. \end{cases}
\]

Indeed, (22) yields

\[
B_1 = \frac{4}{3} \sum_{q \text{ odd}} q^{-3} \phi(q) = \frac{8 \zeta(2)}{7 \zeta(3)},
\]

where the last step uses the Euler product of \( \zeta(s) \). This completes the proof of our theorem.
Rankin [13] and Selberg [17] independently introduced an important method which allows one to study the analytic behavior of the Dirichlet series
\[ \sum_{n=1}^{\infty} \frac{a(n)}{n^s} \]
where \( a(n) \) are Fourier coefficients of a holomorphic cusp form for some congruence subgroup of \( \Gamma = SL_2(\mathbb{Z}) \). Originally the method was for holomorphic cusp forms. Zagier [20] extended the method to cover forms that are not cuspidal and may not decay rapidly at infinity. Müller [11, 12] considered the case where \( a(n) \) is the Fourier coefficient of non-holomorphic cusp or non-cusp form of real weight with respect to a Fuchsian group of the first kind. It is this last approach we wish to discuss. Note that if we apply a Tauberian theorem to the above Dirichlet series, we then gain information on the asymptotic behavior of the partial sum
\[ \sum_{n \leq x} a(n). \]

We now discuss Müller’s elegant work. For details regarding discontinuous groups and automorphic forms, see [8, 10, 11, 14, 15, 16]. Let \( \mathbb{H} = \{ z \in \mathbb{C} : \Im(z) > 0 \} \) denote the upper half plane and \( G = SL(2, \mathbb{R}) \) the special linear group of all \( 2 \times 2 \) matrices with determinant 1. \( G \) acts on \( \mathbb{H} \) by
\[ z \mapsto gz = \frac{az + b}{cz + d} \]
for \( g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G \). We write \( y = y(z) = \Im(z) \). Thus we have
\[ y(gz) = \frac{y}{|cz + d|^2}. \]
Let \( dx \, dy \) denote the Lebesgue measure in the plane. Then the measure
\[ d\mu = \frac{dx \, dy}{y^2} \]
is invariant under the action of \( G \) on \( \mathbb{H} \). A discrete subgroup \( \Gamma \) of \( G \) is called a Fuchsian group of the first kind if its fundamental domain \( \Gamma \backslash \mathbb{H} \) has finite volume. Let \( \Gamma \) be a Fuchsian group of the first kind containing \( \pm I \) where \( I \) is the identity matrix. Let \( \mathcal{F}(\Gamma, \chi, k, \lambda) \) denote the space of (non-holomorphic) automorphic forms of real weight \( k \), eigenvalue \( \lambda = \frac{1}{4} - \rho^2 \), \( \Re(\rho) \geq 0 \), and multiplier system \( \chi \). For \( k \in \mathbb{R}, g \in SL(2, \mathbb{R}) \) and \( f : \mathbb{H} \to \mathbb{C} \), we define the stroke operator \( \mid_k \) by
\[ (f \mid_k g)(z) := \left( \frac{cz + d}{|cz + d|} \right)^{-k} f(gz) \]
where \( g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma \). The transformation law for \( f \in \mathcal{F}(\Gamma, \chi, k, \lambda) \) is then
\[ (f \mid_k g)(z) = \chi(g) f(z) \]
for all \( g \in \Gamma \). Automorphic forms \( f \in \mathcal{F}(\Gamma, \chi, k, \lambda) \) have a Fourier expansion at every cusp \( \kappa \) of \( \Gamma \), namely
\[ A_{\kappa,0}(y) + \sum_{n \neq 0} a_{\kappa,n} W_{(\text{sign} n) \rho} (4\pi |n + \mu_k| y) e((n + \mu_k)x), \]
where \( \mu_k \) is the cusp parameter and \( a_{\kappa,n} \) are the Fourier coefficients of \( f \) at \( \kappa \). The functions \( W_{\alpha,\rho} \) are Whittaker functions (see [11, §3]), \( A_{\kappa,0}(y) = 0 \) if \( \mu_k \neq 0 \) and
An automorphic form $f$ is called a cusp form if $a_{k,0} = b_{k,0} = 0$ for all cusps $\kappa$ of $\Gamma$. Now consider the Dirichlet series

$$S_{\kappa}(f, s) = \sum_{n \geq 0} |a_{\kappa,n}|^2 / (n + \mu_\kappa)^s.$$  

This series is absolutely convergent for $\Re(s) > 2\Re(\rho)$ and has been shown [12] to have meromorphic continuation in the entire complex plane. In what follows, we will only be interested in the case $f$ is not a cusp form. If $f$ is not a cusp form and $\Re(\rho) > 0$, then $S_{\kappa}(f, s)$ has a simple pole at $s = 2\Re(\rho)$ with residue

$$\beta_\kappa(f) = \res_{s = 2\Re(\rho)} S_{\kappa}(f, s) = (4\pi)^{2\Re(\rho)}b^+(k/2, \rho) \sum_{\kappa \in K} \varphi_{\kappa,i}(1 + 2\Re(\rho))|a_{\kappa,0}|^2,$$

where $K$ denotes a complete set of $\Gamma$-inequivalent cusps, $\varphi_{\kappa,i}(1 + 2\Re(\rho)) > 0$ and $b^+(k/2, \rho) > 0$ if $\rho + 1 \pm \frac{k}{2}$ is a non-negative integer. For the definition of the functions $\varphi_{\kappa,i}$ and $b^+$, see Lemma 3.6 and (69) in [12]. This result (23) and a Tauberian argument then provide the asymptotic behaviour of the summatory function

$$\sum_{n \leq x} |a_{\kappa,n}|^2 |n + \mu_\kappa|^{\gamma}.$$

Precisely, we have (see [11, Theorem 2.1] or [12, Theorem 5.2]) that

$$\sum_{n \leq x} |a_{\kappa,n}|^2 |n + \mu_\kappa|^{\gamma} \sim \sum_{s \in \mathbb{R}} \res_{s = 2\Re(\rho)} S_{\kappa}(f, s) \frac{x^{r+s}}{r+s} + O(x^{r+2\Re(\rho)}(\log x)^b),$$

where $2\Re(\rho) + r \geq 0$, $R = \{ \pm 2\Re(\rho), \pm 2\Im(\rho), 0, -r \}$, $\gamma = (2 + 8\Re(\rho))(5 + 16\Re(\rho))^{-1}$, and $g = \max(0, b - 1)$; $b$ denotes the order of the pole of $S_{\kappa}(f, s)(r + s)^{-1}x^{s+r}$ at $s = 2\Re(\rho) (0 \leq b \leq 5)$.

We now consider an application of (24). Let $Q \in \mathbb{Z}^{m \times m}$ be a non-singular symmetric matrix with even diagonal entries and $q(x) = \frac{1}{2}Q[x] = \frac{1}{2}x^T Qx$, $x \in \mathbb{Z}^m$, the associated quadratic form in $m \geq 3$ variables. Here we assume that $q(x)$ is positive definite. Let $r(Q, n)$ denote the number of representations of $n$ by the quadratic form $Q$. Now consider the theta function

$$\theta_Q(z) = \sum_{x \in \mathbb{Z}^m} e^{\pi i Q[x]}.$$  

By [11, Lemma 6.1], the Dirichlet series associated with the automorphic form $\theta_Q$ is

$$(4\pi)^{-m/4} \zeta_Q\left(\frac{m}{4} + s\right)$$

where

$$\zeta_Q(s) = \sum_{n=1}^{\infty} \frac{r(Q, n)}{n^s} = \sum_{x \in \mathbb{Z}^m \setminus \{0\}} q(x)^{-s}$$

for $\Re(s) > m/2$. Using (24), Müller proved the following (see [11, Theorem 6.1])
Theorem (Müller). Let \( q(x) = \frac{1}{2}Q[x] = \frac{1}{2}x^TQx, \ x \in \mathbb{Z}^m \) be a primitive positive definite quadratic form in \( m \geq 3 \) variables with integral coefficients. Then
\[
\sum_{n \leq x} r(Q,n)^2 = Bx^{m-1} + O\left(x^{(m-1)\frac{3}{m+3}}\right)
\]
where
\[
B = (4\pi)^{m/2}B_\infty(\theta_Q)
\]
and \( B_\infty(\theta_Q) \) is given by (23).

We are now in a position to prove our theorem in page 1.

Proof. We are interested in the case \( q(x) = x_1^2 + x_2^2 + x_3^2 \) and so \( r(Q,n) = r_3(n) \) counts the number of representations of \( n \) as a sum of three squares. By Müller’s Theorem above,
\[
\sum_{n \leq x} r_3(n)^2 = Bx^2 + O\left(x^{14/9}\right)
\]
where \( B \) is a computable constant. Specifically, we have by (23) (with \( k = 3/2 \) and \( \rho = 1/4 \))
\[
B = \frac{4\pi^2}{3 - 1}b^+(3/4, 1/4)\sum_{\iota \in K} \varphi_{\infty,\iota}(3/2)|a_{\iota,0}|^2,
\]
where \( K \) denotes a complete set of \( \Gamma_0(4) \)-inequivalent cusps and \( a_{\iota,0} \) is the 0-th Fourier coefficient of \( \theta_Q(z) \) at a rational cusp \( \iota \). Choose \( K = \{1, \frac{1}{2}, \frac{1}{3}\} \). Then by p. 145 and (67) in [11], we have
\[
|a_{\iota,0}|^2 = W^2\xi_j(S,\iota)^2
\]
where \( \iota = u/w, (u, w) = 1, w \geq 1, W_i \) is width of the cusp \( \iota \), and
\[
|\xi_j(S,\iota)|^2 = 2^{-3}w^{-3}\sum_{x=1}^{w} \left| e\left(\frac{u}{w}x^2\right)\right|^6.
\]

As \( W_{1/4} = W_{1/2} = 1, W_1 = 4 \), we have \( |a_{1,0,0}|^2 = 1, |a_{1/2,0}|^2 = 0, \) and \( |a_{1/4,0}|^2 = 1 \). An explicit description of the functions \( \varphi_{\infty,\iota}(s) \) in the case \( \Gamma_0(4) \) is given by (see (1.17) and p. 247 in [5])
\[
\varphi_{\infty,1/4}(s) = 2^{-4s}(1 - 2^{-2s})^{-1}\Gamma(1/2)\Gamma(s - 1/2)\zeta(2s - 1)\zeta(2s),
\]
\[
\varphi_{\infty,1/2}(s) = \varphi_{\infty,1}(s) = 2^{-2s}(1 - 2^{-2s})^{-1}(1 - 2^{-1-2s})\Gamma(1/2)\Gamma(s - 1/2)\zeta(2s - 1)\zeta(2s).
\]
Thus for \( s = 3/2 \), we have
\[
\varphi_{\infty,1/4}(3/2) = 2^{-5}(1 - 2^{-3})^{-1}\pi^2\frac{\zeta(2)}{\Gamma(3/2)\zeta(3)},
\]
\[
\varphi_{\infty,1/2}(3/2) = \varphi_{\infty,1}(3/2) = 2^{-3}(1 - 2^{-3})^{-1}(1 - 2^{-2})\pi^2\frac{\zeta(2)}{\Gamma(3/2)\zeta(3)}.
\]

Now, from p. 65 in [12], we have
\[
b^+(3/4, 1/4) = G^*_{1/4,1/4}(3/2).
\]
By Lemma 3.3 and (16) in [12],
\[ G_{1/4,1/4}^{*}(s) = \Gamma(s + 1/2)^{-1} \]
and so \( b^*(3/4, 1/4) = \Gamma(2)^{-1} \). In total,

\[
B = \frac{(4\pi)^2}{(3 - 1) \Gamma(2)} \left( 2^{-3}(1 - 2^{-3})^{-1}(1 - 2^{-2})\pi^{1/2} \frac{\zeta(2)}{\Gamma(3/2)\zeta(3)} \right.
\]
\[ + 2^{-5}(1 - 2^{-3})^{-1}\pi^{1/2} \frac{\zeta(2)}{\Gamma(3/2)\zeta(3)} \left) \right. = \frac{8\pi^4}{21\zeta(3)} . \]

Thus

\[
\sum_{n \leq x} r_3(n)^2 \sim \frac{8\pi^4}{21\zeta(3)} x^2 .
\]

\[ \square \]

**Remark.** Müller’s Theorem can also be used to obtain the mean square value of sums of \( N > 3 \) squares. Precisely, if \( r_N(n) \) is the number of representations of \( n \) by \( N > 3 \) squares, then a calculation similar to the second proof of our theorem yields (compare with Theorem 3.3 in [2])

\[
\sum_{n \leq x} r_N(n)^2 = W_N x^{N-1} + O\left(x^{(N-1)\frac{N-2}{N}}\right)
\]

where

\[
W_N = \frac{1}{(N - 1)(1 - 2^{-N}) \Gamma(N/2)^2} \frac{\pi^N}{\zeta(N)} \frac{\zeta(N - 1)}{\zeta(N)} .
\]

**Acknowledgments**

The authors would like to thank Wolfgang Müller for his comments regarding the second proof of the theorem. The third author would like to thank the Max-Planck-Institut für Mathematik for their hospitality and support during the preparation of this paper.

**References**


DEPARTMENT OF MATHEMATICS, SIMON FRASER UNIVERSITY, BURNABY, BRITISH COLUMBIA, CANADA V5A 1S6
E-mail address: kkchoi@cecm.sfu.ca

DEPARTMENT OF MATHEMATICS, 1 UNIVERSITY STATION C1200, THE UNIVERSITY OF TEXAS AT AUSTIN, AUSTIN, TX 78712, U.S.A.
E-mail address: kumchev@math.utexas.edu

MAX-PLANCK INSTITUT FÜR MATHEMATIK, VIVATSGASSE 7, 53111 BONN, GERMANY
E-mail address: osburn@mpim-bonn.mpg.de