On sums of squares of primes

By GLYN HARMAN

*Department of Mathematics, Royal Holloway University of London, Egham, Surrey TW*20 0*EX, U.K.*
e-mail: G.Harman@rhul.ac.uk

and ANGEL V. KUMCHEV

Department of Mathematics, The University of Texas at Austin, 1 *University Station, C*1200*, Austin, TX* 78712*, U.S.A.*
e-mail: kumchev@math.utexas.edu

(*Received*)

Abstract

In this article we consider the exceptional set of integers, not restricted by elementary congruence conditions, which cannot be represented as sums of three or four squares of primes. Using new exponential sums in tandem with a sieve method we are able to provide stronger "minor arc" estimates than previous authors, thereby improving the saving obtained in the exponent by a factor 8/7.

1. *Introduction*

Let

$$\mathcal{A}_3 = \{n \in \mathbb{N} : n \equiv 3 \pmod{24},\ n \not\equiv 0 \pmod 5\},$$
$$\mathcal{A}_4 = \{n \in \mathbb{N} : n \equiv 4 \pmod{24}\}.$$

It is conjectured that every sufficiently large integer in $\mathcal{A}_j$ can be represented as the sum of j squares of primes. Indeed, there are hypothetical asymptotic formulae for the numbers of such representations:

$$\sum_{p_1^2+p_2^2+p_3^2=n} (\log p_1)(\log p_2)(\log p_3) \ \sim\ \frac{\pi}{4}\mathfrak{S}_3(n)n^{1/2} \tag{1.1}$$

and

$$\sum_{p_1^2+\cdots+p_4^2=n} (\log p_1)\cdots(\log p_4) \ \sim\ \frac{\pi^2}{16}\mathfrak{S}_4(n)n, \tag{1.2}$$

where $\mathfrak{S}_j(n) > 0$ for all large $n \in \mathcal{A}_j$. In 1938 Hua [**9**] proved that almost all $n \in \mathcal{A}_3$ are representable as sums of three squares of primes, from which it instantly follows that almost all $n \in \mathcal{A}_4$ are representable as sums of four squares of primes. The natural question then becomes: how good a bound can we get on the possible exceptional sets? Let $E_j(N)$ denote the number of exceptions up to N for the problem with j squares.

Schwarz [**18**] refined Hua's approach to show that

$$E_3(N) \ll N(\log N)^{-A} \qquad \text{for any } A > 0.$$

In 1993 Leung and Liu [**11**] improved this to $E_3(N) \ll N^{1-\delta}$ for some (very small) fixed $\delta > 0$. Since 1998, when Liu and Zhan [**14**] found a new approach to increase the size of the "major arcs" in the application of the Hardy–Littlewood circle method, there has been a flurry of activity in this area culminating in the results of Liu and Zhan [**16**] and Liu, Wooley and Yu [**13**] who proved respectively that

$$E_3(N) \ll N^{11/12+\epsilon} \qquad \text{and} \qquad E_4(N) \ll N^{3/8+\epsilon}.$$

Recently, the second named author [**10**] established a new exponential sum estimate which leads to a simpler proof of the latter of these bounds and to a further improvement on the former:

$$E_3(N) \ll N^{7/8+\epsilon}.$$

In this note we shall combine this technique with the first named author's sieve method [**6**, **7**] and a new exponential sum estimate from [**8**] to prove the following results.

THEOREM 1. *Let $\epsilon > 0$ be given. Then for all large N we have*

$$E_3(N) \ll N^{6/7+\epsilon}. \tag{1.3}$$

THEOREM 2. *Let $\epsilon > 0$ be given. Then for all large N we have*

$$E_4(N) \ll N^{5/14+\epsilon}. \tag{1.4}$$

Remark. The exponents in (1.3) and (1.4) arise as $1 - \frac{1}{7}$ and $\frac{1}{2} - \frac{1}{7}$ respectively, so we have made an improvement by a factor of $\frac{8}{7}$ over the previous results. In contrast to earlier work, we are unable to show that (1.1) and (1.2) hold except on sets of the size indicated in our theorems. Instead, we obtain lower bounds of the correct order of magnitude for the numbers of solutions, which of course is consistent with our use of a sieve method.

The techniques employed in the proofs of Theorems 1 and 2 can also be used to improve on a recent result of Choi and the second named author [**3**] concerning more general quadratic equations with five prime variables. Consider the Diophantine equation

$$b_1 p_1^2 + \cdots + b_5 p_5^2 = n, \tag{1.5}$$

where the coefficients $b_1, \dots, b_5, n$ satisfy the conditions

$$\gcd(b_i, b_j, b_k) = 1, \qquad 1 \le i < j < k \le 5,$$

and

$$b_1 + \cdots + b_5 \equiv n \pmod{24},$$

and where 5 divides at most two of $b_1, \dots, b_5, n$. Modifying the approach in [**3**] in the same way that we here amend the method of [**16**], one obtains the following theorem.

THEOREM 3. *Let $\epsilon > 0$ be given and let $b_1, \dots, b_5, n$ be integers satisfying the above conditions. Also, assume that $|b_1| \ge \cdots \ge |b_5| \ge 1$. If $b_1, \dots, b_5$ are not all of the same sign, then* (1.5) *has solutions in primes $p_1, \dots, p_5$ satisfying*

$$p_j \ll \sqrt{|n|} + |b_1 \cdots b_4|^{7/4+\epsilon}. \tag{1.6}$$

If $b_1, \dots, b_5$ are all positive, then (1.5) *has solutions provided that*

$$n \gg b_5(b_1 \cdots b_4)^{7/2+\epsilon}. \tag{1.7}$$

2. *Outline and preliminary results*

We shall concentrate on proving Theorem 1. We will describe the straightforward modifications needed for Theorem 2 at the end of the paper and will suppress the proof of Theorem 3 altogether. It suffices to estimate the number of exceptional integers n in the set $\mathcal{B} = \mathcal{A}_3 \cap (\frac{1}{2}N, N]$. Here N is our main parameter, which we assume to be "sufficiently large". We write

$$P = N^{1/2}, \qquad \mathcal{L} = \log P, \qquad \mathcal{I} = \left[\tfrac{1}{3}P, \tfrac{2}{3}P\right).$$

We use c to denote an absolute constant, not necessarily the same at each occurrence; this convention simplifies the estimation of various averages of common number-theoretic functions, as it allows us to write, for example,

$$\sum_{m \le X} d(m)^c \ll X(\log X)^c.$$

Here $d(m)$ is the number of positive divisors of m. The circle method, in the form we require it here, begins with the observation that

$$\sum_{\substack{p_1^2+p_2^2+m^2=n \\ m,p_j \in \mathcal{I}}} (\log p_1)(\log p_2)\rho(m) = \int_0^1 f(\alpha)^2 g(\alpha) e(-\alpha n)\, d\alpha, \tag{2.1}$$

where we write $e(x) = \exp(2\pi i x)$ and

$$f(\alpha) = \sum_{p \in \mathcal{I}} (\log p) e(\alpha p^2), \qquad g(\alpha) = \sum_{m \in \mathcal{I}} \rho(m) e(\alpha m^2). \tag{2.2}$$

Here we will want ρ to satisfy:

$$\rho(m) \le \begin{cases} 1 & \text{if } m \text{ is prime,} \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad \sum_{m \le X} \rho(m) \gg X\mathcal{L}^{-1} \tag{2.3}$$

for $P^{1/2} \le X \le P$. In other words, ρ is a non-trivial lower bound for the characteristic function of the set of primes in $\mathcal{I}$. Further necessary properties of ρ will be introduced when relevant (see [**1**] for a similar situation, in particular the conditions listed in Theorem 2 there).

As explained in [**19**], the circle method rests on a dissection of the interval $[0, 1)$ into "major" and "minor" arcs, on which the exponential sums (2.2) are respectively "large" (and can be expressed in terms of complete sums with an arithmetical interpretation) and "not too large". The main contribution to the right side of (2.1) comes from the major arcs. We define the Hardy–Littlewood dissection as follows. Let $Q = P^{2/7-\epsilon}$ and write (shifting $[0, 1)$ by $\eta = QP^{-2+\epsilon}$ which does not change (2.1))

$$\mathfrak{M} = [\eta, 1+\eta) \cap \bigcup_{1 \le q \le Q} \bigcup_{(a,q)=1} \left[\frac{a}{q} - \frac{\eta}{q}, \frac{a}{q} + \frac{\eta}{q}\right). \tag{2.4}$$

These are the major arcs, and so the minor arcs $\mathfrak{m}$ are given by $\mathfrak{m} = [\eta, 1+\eta) \setminus \mathfrak{M}$. For technical reasons, it is convenient to modify $g(\alpha)$ on the major arcs to remove interference

between possible prime divisors of m (when $\rho(m) < 0$) and approximation denominators. We introduce a function $\theta(m, \alpha)$ which is 1 except when there exist integers a and q such that

$$|q\alpha - a| < \eta, \quad (a, q) = 1, \quad q \le Q, \quad (m, q) \text{ is divisible by a prime } p \ge P^{1/7},$$

in which case $\theta(m, \alpha) = 0$. Write

$$h(\alpha) = \sum_{m \in \mathcal{I}} \rho(m)\theta(m, \alpha)e(\alpha m^2), \qquad k(\alpha) = g(\alpha) - h(\alpha).$$

We note that $h(\alpha) = g(\alpha)$ for $\alpha \in \mathfrak{m}$ and that

$$k(\alpha) \ll P^{6/7} \tag{2.5}$$

for all α.

3. *A minor arc estimate for $g(\alpha)$*

The limit of the method used in earlier work is set by the exponent $\frac{7}{8}$ on the right side of the following bound due to Ghosh [**5**]: if $|q\alpha - a| < q^{-1}$, $(a, q) = 1$, then

$$f(\alpha) \ll P^{7/8+\epsilon} + P^{1+\epsilon}\left(\frac{1}{q} + \frac{q}{P^2}\right)^{1/4}. \tag{3.1}$$

In this section we apply a device developed recently by the first named author [**8**] to show that if ρ has properties (i)–(iii) below, $g(\alpha)$ satisfies a variant of (3.1) in which the exponent $\frac{7}{8}$ is replaced with $\frac{6}{7}$. The quality of our main results is ultimately decided by this improvement. We also make use of an estimate of the second named author [**10**], which allows us (essentially) to replace the exponent $\frac{1}{4}$ on the right side of (3.1) with $\frac{1}{2}$. The reader should note that for the latter improvement we are really using major arc techniques to cover part of the minor arcs.

Henceforth, we write

$$\psi(m, z) = \begin{cases} 1 & \text{if } p \mid m \Rightarrow p \ge z, \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

LEMMA 1. *Suppose that $\alpha \in \mathfrak{m}$ and that the function ρ in (2.2) has the properties:*

(i) *$\rho(m) = 0$ unless $\psi(m, P^{3/14}) = 1$;*

(ii) *$\rho(m)$ is the linear combination of $O(\mathcal{L}^c)$ bilinear sums of the form*

$$\sum_{uv=m} \gamma_u \delta_v, \tag{3.3}$$

where $|\gamma_u| \le d(u)^c$, $|\delta_v| \le d(v)^c$, and either $P^{2/7} \le u \le P^{3/7}$, or $v \ge P^{2/7}$ and $\delta_v = 1$ for all v;

(iii) *$\rho(m)$ is the linear combination of $O(\mathcal{L}^c)$ bilinear sums of the form*

$$\sum_{\substack{uv=m \\ U < u \le 2U}} \gamma_u \psi(v, z), \tag{3.4}$$

where $|\gamma_u| \le d(u)^c$, $U \le P^{1/2}$, and $z \le \sqrt{2P/U}$.

Then

$$g(\alpha) \ll P^{6/7+2\epsilon}. \tag{3.5}$$

Proof. By Dirichlet's theorem in Diophantine approximation, we can find integers a, q with

$$1 \le q \le (P/Q)^2, \quad (a,q) = 1, \quad |q\alpha - a| < (Q/P)^2.$$

Under the assumption of hypothesis (ii), the arguments in [**8**, Sec. 8 and 9] (see [**8**, (34)] in particular) yield the bound

$$g(\alpha) \ll P^{6/7+\epsilon} + P^{1+\epsilon}\left(\frac{1}{q} + \frac{q}{P^2}\right)^{1/4}.$$

This establishes (3.5) when $q \ge Q^2$. On the other hand, when $q \le Q^2$ hypotheses (i) and (iii) ensure that we can appeal to [**10**, Lemma 5.6]. This yields the bound

$$g(\alpha) \ll \frac{P^{1+\epsilon}}{(q + P^2|q\alpha - a|)^{1/2}} + QP^{11/20+\epsilon} + P^{11/14+\epsilon},$$

from which (3.5) follows on noting that for $\alpha \in \mathfrak{m}$ we have

$$q + P^2|q\alpha - a| > Q.$$

4. *The major arcs*

In this section we estimate the contribution from the major arcs. We follow the approach of Liu and Zhan [**16**], with some modifications due to the presence of the sieve weights ρ. Define a function $f^*(\alpha)$ on $\mathfrak{M}$ by setting

$$f^*(\alpha) = \frac{S(\chi_0, a)}{\phi(q)} \sum_{m \in \mathcal{I}} e\left(\beta m^2\right) \qquad \text{if } |q\alpha - a| < \eta.$$

Here χ_0 is the principal character modulo q and

$$S(\chi, a) = \sum_{h=1}^{q} \overline{\chi}(h) e\left(\frac{ah^2}{q}\right).$$

The analysis of the major arcs in [**16**] establishes that on average over $\alpha \in \mathfrak{M}$ one can approximate $f(\alpha)$ by $f^*(\alpha)$. In the present context, we need also an analogue of $f^*(\alpha)$ which approximates $h(\alpha)$ on $\mathfrak{M}$. Let

$$\sum_m \omega(m) e\left(\alpha m^2\right)$$

denote either $f(\alpha)$ or $h(\alpha)$. In particular, we have $\omega(m) = 0$ for $(m, q) > 1$. Then

$$\begin{aligned}\sum_m \omega(m) e\left(\alpha m^2\right) &= \sum_{h=1}^{q} \sum_{m \equiv h \pmod q} \omega(m) e\left(\frac{ah^2}{q} + \beta m^2\right) \\ &= \frac{1}{\phi(q)} \sum_{h=1}^{q} e\left(\frac{ah^2}{q}\right) \sum_m \omega(m) e\left(\beta m^2\right) \sum_{\chi \bmod q} \overline{\chi}(h\overline{m}) \\ &= \frac{1}{\phi(q)} \sum_{\chi \bmod q} S(\chi, a) \sum_m \omega(m)\chi(m) e\left(\beta m^2\right), \end{aligned} \tag{4.1}$$

where $\overline{m}$ is the multiplicative inverse of m modulo q and $\sum_\chi$ denotes a summation over the Dirichlet characters modulo q. (We note that since $\chi(m) = 0$ for $(m, q) > 1$, in the case of $h(\alpha)$ the weight $\theta(m, \alpha)$ becomes superfluous at this point and can be dropped.)

The definition of $f^*(\alpha)$ is suggested by (4.1) and the prime number theorem for arithmetic progressions. In order to have a similar approximation to $h(\alpha)$, we suppose that ρ has the following properties:

(iv) Let $A, B > 0$ be fixed, let χ be a non-principal character mod q, $q \le \mathcal{L}^B$, and let $\mathcal{I}'$ be a subinterval of $\mathcal{I}$. Then

$$\sum_{m \in \mathcal{I}'} \rho(m)\chi(m) \ll P\mathcal{L}^{-A}. \tag{4.2}$$

(v) Let $A > 0$ be fixed and let $\mathcal{I}'$ be a subinterval of $\mathcal{I}$. Then

$$\sum_{m \in \mathcal{I}'} \rho(m) = \sum_{m \in \mathcal{I}'} \varrho(m) + O\big(P\mathcal{L}^{-A}\big) \tag{4.3}$$

$$= \delta|\mathcal{I}'|\mathcal{L}^{-1} + O\big(P\mathcal{L}^{-2}\big), \tag{4.4}$$

where ϱ is a smooth function on $\mathcal{I}$ and $\delta > 0$ is an absolute constant.

Our major arc approximation to $h(\alpha)$ is then given by

$$h^*(\alpha) = \frac{S(\chi_0, a)}{\phi(q)} \sum_{m \in \mathcal{I}} \varrho(m) e\left(\beta m^2\right) \qquad \text{if } |q\alpha - a| < \eta,$$

where ϱ is the function appearing in (4.3). We now proceed to estimate the quantity

$$\int_{\mathfrak{M}} f(\alpha)^2 h(\alpha) e(-\alpha n)\, d\alpha - \int_{\mathfrak{M}} f^*(\alpha)^2 h^*(\alpha) e(-\alpha n)\, d\alpha, \tag{4.5}$$

which we think of as the error of approximation of the integral over $\mathfrak{M}$ by the expected main term. For our purposes, it suffices to show that this quantity is $O(P\mathcal{L}^{-A})$ for any fixed $A > 0$, for example.

Similarly to [**16**, (3.1)], we can use (4.1) to express the difference (4.5) as a linear combination of error terms involving $f^*(\alpha)$, $h^*(\alpha)$, and the sums

$$W(\chi, \beta) = \sum_{p \in \mathcal{I}} (\log p)\chi(p) e\left(\beta p^2\right) - D(\chi) \sum_{m \in \mathcal{I}} e\left(\beta m^2\right),$$

$$W^\sharp(\chi, \beta) = \sum_{m \in \mathcal{I}} \rho(m)\chi(m) e\left(\beta m^2\right) - D(\chi) \sum_{m \in \mathcal{I}} \varrho(m) e\left(\beta m^2\right),$$

where $D(\chi)$ is 1 or 0 according as χ is principal or not. We shall focus on the most troublesome among the error terms that arise, namely the multiple sum

$$\sum_{q \le Q} \sum_{\chi_1 \bmod q} \sum_{\chi_2 \bmod q} \sum_{\chi_3 \bmod q} B(n, q; \chi_1, \chi_2, \chi_3) J(n, q; \chi_1, \chi_2, \chi_3). \tag{4.6}$$

Here

$$B(n, q; \chi_1, \chi_2, \chi_3) = \frac{1}{\phi(q)^3} \sum_{\substack{1 \le a \le q \\ (a,q)=1}} S(\chi_1, a) S(\chi_2, a) S(\chi_3, a) e\left(\frac{-an}{q}\right)$$

and

$$J(n, q; \chi_1, \chi_2, \chi_3) = \int_{-\eta/q}^{\eta/q} W^\sharp(\chi_1, \beta) W(\chi_2, \beta) W(\chi_3, \beta) e(-\beta n)\, d\beta.$$

We first reduce (4.6) to a sum over primitive characters. Suppose that ρ satisfies hypothesis (i) of Lemma 1 and that $\chi_j^* \bmod r_j$, $r_j \mid q$, is the primitive character inducing χ_j. In

general, if $\chi \bmod q$, $q \le Q$, is induced by a primitive character $\chi^* \bmod r$, $r \mid q$, we have

$$W(\chi, \beta) = W(\chi^*, \beta) \tag{4.7}$$

and

$$W^\sharp(\chi, \beta) = W^\sharp(\chi^*, \beta) + O\big(r^{-2}P^{13/14}\big), \tag{4.8}$$

where the error term counts the integers in the set

$$\{m \in \mathcal{I} : (m, q) > 1,\ (m, r) = 1,\ \rho(m) \neq 0\}.$$

(When $r \le QP^{-3/14} < P^{1/14}$, this set contains $\ll P^{11/14} \ll r^{-2}P^{13/14}$ integers; when $r > QP^{-3/14}$, it is empty.) By (4.7) and (4.8),

$$J(n, q; \chi_1, \chi_2, \chi_3) \ll \big(W^\sharp(\chi_1^*) + r_1^{-2}P^{13/14}\big)W(\chi_2^*)W(\chi_3^*), \tag{4.9}$$

where for a character $\chi \bmod r$

$$W^\sharp(\chi) = \max_{|\beta| \le \eta/r} |W^\sharp(\chi, \beta)|, \qquad W(\chi) = \bigg(\int_{-\eta/r}^{\eta/r} |W(\chi, \beta)|^2 d\beta\bigg)^{1/2}.$$

Using (4.9), we can bound (4.6) by

$$\sum_{r_1 \le Q} \sideset{}{^*}\sum_{\chi_1} \sum_{r_2 \le Q} \sideset{}{^*}\sum_{\chi_2} \sum_{r_3 \le Q} \sideset{}{^*}\sum_{\chi_3} \big(W^\sharp(\chi_1) + r_1^{-2}P^{13/14}\big)W(\chi_2)W(\chi_3)B(n; \chi_1, \chi_2, \chi_3).$$

Here $\sum_{r_j} \sideset{}{^*}\sum_{\chi_j}$ denotes a summation over the primitive characters to moduli $r_j \le Q$ and

$$B(n; \chi_1, \chi_2, \chi_3) = \sum_{\substack{q \le Q \\ q_0 \mid q}} |B(n, q; \chi_1\chi_0, \chi_2\chi_0, \chi_3\chi_0)|,$$

with $q_0 = [r_1, r_2, r_3]$ and χ_0 the principal character mod q. By [**15**, Lemma 2.1],

$$B(n; \chi_1, \chi_2, \chi_3) \ll q_0^{-1/2+\epsilon}\mathcal{L}^c,$$

and by [**16**, Lemma 2.4],

$$\sum_{r \le R} \sideset{}{^*}\sum_{\chi} [r, d]^{-1/2+\epsilon} W(\chi) \ll d^{-1/2+\epsilon}\mathcal{L}^c$$

whenever $R \le P^{1/3-\epsilon}$. It thus follows that the sixfold sum above does not exceed

$$\mathcal{L}^c \sum_{r_1 \le Q} \sideset{}{^*}\sum_{\chi_1} \big(r_1^{-1/2+\epsilon} W^\sharp(\chi_1) + r_1^{-5/2+\epsilon}P^{13/14}\big).$$

This reduces the estimation of the error term (4.6) to the following bound: if $\frac{1}{2} \le R \le Q$, then for any fixed $A > 0$ we have

$$\sum_{R < r \le 2R} \sideset{}{^*}\sum_{\chi} W^\sharp(\chi) \ll PR^{1/4}\mathcal{L}^{-A} + P^{6/7}R\mathcal{L}^c. \tag{4.10}$$

We start the proof of (4.10) by replacing the maximum in the definition of $W^\sharp(\chi)$ by a maximum over $\Delta - P^{-2} \le |\beta| \le 2\Delta$, for some $P^{-2} \le \Delta \le \eta R^{-1}$. The price we have to pay for this is an extra logarithmic factor in the final bound. We now write $T_0 = 1 + \Delta P^2$. By standard procedures (see [**16**, (4.8) and (4.9)]), we find that the left side of (4.10) is

bounded above by

$$\mathcal{L}^c(T_0P)^{1/2}\frac{1}{T}\sum_{R<r\le 2R}{\sum_{\chi}}^{*}\int_{-T}^{T}\big|F(1/2+it,\chi)\big|\,dt + R^2P^{1/2}, \tag{4.11}$$

where $T_0 \le T \le P^{10}$ and

$$F(s,\chi)=\sum_{m\in\mathcal{I}}\rho(m)\chi(m)m^{-s}.$$

We now appeal to the mean-value estimate

$$\sum_{R<r\le 2R}{\sum_{\chi}}^{*}\int_{-T}^{T}\big|F(1/2+it,\chi)\big|\,dt \ll \mathcal{L}^c\big(P^{1/2}+RT^{1/2}P^{5/14}+R^2T\big), \tag{4.12}$$

which holds provided that ρ satisfies hypothesis (ii) of Lemma 1. The proof of this inequality is almost identical to the proof of [**12**, Lemma 2.1]. The only significant difference is that we use Dirichlet polynomials with coefficients of the form (3.3) with $P^{2/7} \le u, v \le P^{5/7}$, whereas the coefficients in [**12**] are subject to the more stringent restriction $P^{2/5} \le u, v \le P^{3/5}$. Because of this change, we have a middle term on the right side of (4.12) that is larger than the respective term $RT^{1/2}P^{3/10}$ in [**12**, (2.1)]. Estimating (4.11) by means of (4.12), we get

$$\sum_{R<r\le 2R}{\sum_{\chi}}^{*}W^{\sharp}(\chi) \ll \mathcal{L}^c\big(PT_0^{-1/2}+P^{6/7}R\big),$$

which establishes (4.10), provided that $R \ge \mathcal{L}^B$ or $\Delta \ge \mathcal{L}^BP^{-2}$, with $B(A) > 0$. On the other hand, if $R \le \mathcal{L}^B$ and $\Delta \le \mathcal{L}^BP^{-2}$, partial summation yields

$$\sum_{R<r\le 2R}{\sum_{\chi}}^{*}W^{\sharp}(\chi) \ll \mathcal{L}^B\sum_{R<r\le 2R}{\sum_{\chi}}^{*}\max_{\mathcal{I}'\subset\mathcal{I}}\Big|\sum_{m\in\mathcal{I}'}\rho(m)\chi(m)\Big| \tag{4.13}$$

or

$$\sum_{R<r\le 2R}{\sum_{\chi}}^{*}W^{\sharp}(\chi) \ll \mathcal{L}^B\max_{\mathcal{I}'\subset\mathcal{I}}\Big|\sum_{m\in\mathcal{I}'}\rho(m)-\sum_{m\in\mathcal{I}'}\varrho(m)\Big|, \tag{4.14}$$

according as $R \ge 1$ or $R < 1$. Here the maxima are over all subintervals $\mathcal{I}'$ of $\mathcal{I}$. Since the right sides of (4.13) and (4.14) are $O(P\mathcal{L}^{-A})$ by (4.2) and (4.3) respectively, this completes the proof of (4.10).

We have shown that the multiple sum (4.6) is $O(P\mathcal{L}^{-A})$ for any fixed $A > 0$. Recall that (4.6) was one of several error terms in a representation of (4.5). Since the other error terms in that representation can be estimated by (possibly simpler) versions of the above argument, we conclude that the difference (4.5) is $O(P\mathcal{L}^{-A})$. Finally, standard major arcs techniques yield

$$\int_{\mathfrak{M}} f^*(\alpha)^2h^*(\alpha)e(-\alpha n)\,d\alpha = \mathfrak{S}_3(n,Q)P\mathcal{L}^{-1}\big(\delta'+O(\mathcal{L}^{-1})\big)+O(P^{1-\epsilon}), \tag{4.15}$$

where δ' is a constant multiple of the number δ appearing in (4.4) and

$$\mathfrak{S}_3(n,X)=\sum_{q\le X}\mathfrak{s}(q,n), \qquad \mathfrak{s}(q,n)=\sum_{\substack{1\le a\le q\\(a,q)=1}}\frac{S(\chi_0,a)^3}{\phi(q)^3}e\left(-an/q\right).$$

By [**2**, Lemma 6.1], we have

$$\sum_{N/2<n\le N}\left|\mathfrak{S}_3(n,X)-8\prod_{2<p\le X^{3/2}}(1+\mathfrak{s}(p,n))\right|\ll N^{1+\epsilon/2}X^{-1/2} \tag{4.16}$$

for any X with $N^\epsilon\le X\le N^{1/4}$, and by [**17**, (2.3) and (2.7)],

$$\prod_{2<p\le X^{3/2}}(1+\mathfrak{s}(p,n))\gg(\log X)^{-3} \tag{4.17}$$

for any $n\in\mathcal{B}$. Combining (4.15)–(4.17) we obtain the following result.

Lemma 2. *Suppose that ρ satisfies hypotheses* (i), (ii), (iv), *and* (v) *above. Then for all but $O(N^{6/7+\epsilon})$ integers $n\in\mathcal{B}$,*

$$\int_{\mathfrak{M}}f^2(\alpha)h(\alpha)e(-\alpha n)\,d\alpha\gg P\mathcal{L}^{-4}. \tag{4.18}$$

Remark. Our proof of Lemma 2 follows the "path of least resistance". However, it might leave the reader with the false impression that the choice $Q=P^{2/7-\epsilon}$ in (2.4) is optimal, whereas in reality it is just convenient. For example, if we assume that ρ satisfies hypothesis (iii) of Lemma 1 instead of hypothesis (ii) and then combine the methods in [**12**] and [**10**, Sec. 5], we can replace the term $RT^{1/2}P^{5/14}$ in (4.12) by the term $RT^{1/2}P^{3/10}$ appearing in [**12**, (2.1)]. In particular, the choice $Q=P^{1/3-\epsilon}$ is within the reach of our methods. Note that with such a choice we would match the size of the major arcs obtained by Liu and Zhan [**16**] in the absence of sieve weights.

5. *The sieve method*

We now show how a function ρ having properties (i)–(v) above can be constructed using the sieve method originating in [**6**] and developed in [**7**, **1**]. Essentially the same construction is used in [**8**], although our situation is simpler here since we need only sieve one variable and therefore we have no recourse to the "vector sieve". On the other hand, we have to impose the "Siegel–Walfisz" condition (iv) which was not necessary in [**8**].

We recall (3.2). Of course, $\psi(m,P^{1/2})$ is the characteristic function of the set of primes in $\mathcal{I}$. Buchstab's identity gives

$$\psi(m,P^{1/2})=\psi(m,P^{5/21})-\sum_{P^{5/21}\le p<P^{1/2}}\psi(m/p,p). \tag{5.1}$$

Here and in the sequel we extend $\psi(m,z)$ to all real $m>0$ by setting $\psi(m,z)=0$ for $m\notin\mathbb{N}$, so the sum in (5.1) is really over the prime divisors of m. By (5.1), we have $\psi(m,P^{1/2})\ge\rho(m)$ for any function ρ of the form

$$\rho(m)=\psi(m,P^{5/21})-\sum_{P^{5/21}\le p<P^{1/2}}\psi(m/p,z(p)), \tag{5.2}$$

where $z(p)\le p$. The appropriate choice to make here is

$$z(p)=\begin{cases}P^{5/14}p^{-1/2} & \text{if } p<P^{2/7},\\ p & \text{if } P^{2/7}\le p\le P^{3/7},\\ P^{5/7}p^{-1} & \text{if } p>P^{3/7}.\end{cases}$$

It is straighforward to check that the resulting function satisfies hypotheses (i) and (iii) of

Lemma 1. Furthermore, (iv) and (v) follow by partial summation from the Siegel–Walfisz theorem (in the form of [**4**, §22, (3)]) and from the prime number theorem, respectively. We remark that in order to achieve an error term of the quality stated in (4.3), we have to settle for a rather ungainly function ϱ. As to the value of δ in (4.4), the methods used to evaluate the main terms in [**1**] or [**7**] yield

$$\delta = 1 - \int_{3/7}^{1/2} \int_{(5-7\alpha)/7}^{(1-\alpha)/2} \frac{d\alpha d\beta}{\alpha\beta(1-\alpha-\beta)} - \int_{5/21}^{2/7} \int_{(5-7\alpha)/14}^{\alpha} w\left(\frac{1-\alpha-\beta}{\beta}\right) \frac{d\alpha d\beta}{\alpha\beta^2},$$

which is easily seen to be positive by numerical integration (indeed, we have $\delta > \frac{9}{10}$). Here $w(u)$ is Buchstab's function, defined as the continuous solution of

$$\begin{cases} (uw(u))' = w(u-1) & \text{if } u > 2, \\ w(u) = u^{-1} & \text{if } 1 < u \le 2. \end{cases}$$

Finally, we show that ρ satisfies hypothesis (ii) of Lemma 1. We have

$$\psi(m, P^{5/21}) = \psi(m, P^{1/7}) - \sum_{P^{1/7} \le p < P^{5/21}} \psi(m/p, P^{1/7}) + \sum_{P^{1/7} \le q < p < P^{5/21}} \psi(m/(pq), q). \tag{5.3}$$

We note that in the last sum $pq > P^{2/7}$. Write S^* for the part of this sum with $pq > P^{3/7}$. Then

$$\begin{aligned} S^* &= \sum_{P^{3/7}/p < q < p < P^{5/21}} \psi(m/(pq), q) \\ &= \sum_{P^{3/7}/p < q < p < P^{5/21}} \psi(m/(pq), P^{1/7}) - \sum_{\substack{P^{3/7}/p < q < p < P^{5/21} \\ P^{1/7} \le r < q}} \psi(m/(pqr), r). \end{aligned} \tag{5.4}$$

We observe that in the final sum in (5.4) we have $P^{4/7} < pqr < P^{5/7}$. There is an analogue of our treatment of $\psi(m, P^{5/21})$ for the final term in (5.2). Combining (5.2), (5.3) and (5.4) then yields a decomposition of $\rho(m)$ into functions of the form

$$\sum_{\substack{uv=m \\ P^{2/7} \le u \le P^{3/7}}} \gamma_u \delta_v \tag{5.5}$$

or

$$\sum_{U < u \le 2U} \gamma_u \psi(m/u, P^{1/7}), \tag{5.6}$$

where $U \le P^{4/7}$. The terms of the form (5.5) are admissible in hypothesis (ii) of Lemma 1. Moreover, the method of [**7**] converts each term of the form (5.6) into $O(\mathcal{L})$ sums of the forms appearing in hypothesis (ii). We have thus established all the required properties of ρ.

6. *Proof of Theorem 1*

Let $\mathfrak{Z}$ be the set of integers $n \in \mathcal{B}$ for which (4.18) holds but which are not representable as sums of three squares of primes. We write $|\mathfrak{Z}|$ for the cardinality of $\mathfrak{Z}$ and $Z(\alpha)$ for its

generating function:

$$Z(\alpha) = \sum_{n \in \mathfrak{Z}} e(-\alpha n).$$

Then, by (2.1) and (2.3),

$$\int_0^1 f^2(\alpha) g(\alpha) Z(\alpha)\, d\alpha \le 0,$$

and at the same time

$$\int_{\mathfrak{M}} f^2(\alpha) h(\alpha) Z(\alpha)\, d\alpha \gg |\mathfrak{Z}| P \mathcal{L}^{-4}.$$

Thus,

$$\left| \int_{\mathfrak{m}} f(\alpha)^2 g(\alpha) Z(\alpha)\, d\alpha + \int_{\mathfrak{M}} f(\alpha)^2 k(\alpha) Z(\alpha)\, d\alpha \right| \gg |\mathfrak{Z}| P \mathcal{L}^{-4}.$$

Recalling Lemma 1 and (2.5), we deduce that

$$\begin{aligned} |\mathfrak{Z}| &\ll \mathcal{L}^4 P^{-1} \Big(\int_{\mathfrak{m}} \big| f(\alpha)^2 g(\alpha) Z(\alpha) \big|\, d\alpha + \int_0^1 \big| f(\alpha)^2 k(\alpha) Z(\alpha) \big|\, d\alpha \Big) \\ &\ll P^{-1/7+\epsilon/2} \int_0^1 \big| f(\alpha)^2 Z(\alpha) \big|\, d\alpha. \end{aligned}$$

Finally, using Cauchy's inequality, Parseval's identity and Hua's lemma [**19**, Lemma 2.5], we find that the last integral does not exceed

$$\left(\int_0^1 |Z(\alpha)|^2 d\alpha \right)^{1/2} \left(\int_0^1 |f(\alpha)|^4 d\alpha \right)^{1/2} \ll |\mathfrak{Z}|^{1/2} P^{1+\epsilon/2},$$

whence

$$|\mathfrak{Z}| \ll P^{12/7+2\epsilon} \ll N^{6/7+\epsilon}.$$

Theorem 1 follows from this bound and Lemma 2.

7. *Proof of Theorem 2*

We now outline the modifications necessary to our previous argument. Let $\mathfrak{Z}$ be the set of numbers in $\mathcal{A}_4 \cap (\frac{1}{2}N, N]$ not representable as sums of four squares of primes. We keep the same major and minor arc decomposition. The corresponding variant of Lemma 2 holds for all $n \in \mathcal{A}_4 \cap (\frac{1}{2}N, N]$, so we obtain

$$\int_{\mathfrak{M}} f(\alpha)^3 h(\alpha) Z(\alpha)\, d\alpha \gg |\mathfrak{Z}| P^2 \mathcal{L}^{-1}.$$

(We remark that $\mathfrak{S}_4(n) \gg 1$ for $n \in \mathcal{A}_4$.) Then arguing as in the previous section, we get

$$|\mathfrak{Z}| \ll P^{-8/7+\epsilon/2} \int_0^1 \big| f(\alpha)^3 Z(\alpha) \big|\, d\alpha.$$

To this end we apply the device introduced by Wooley [**20**] which entails using Cauchy's inequality to give

$$\begin{aligned} \int_0^1 \big| f(\alpha)^3 Z(\alpha) \big|\, d\alpha &\le \left(\int_0^1 |f(\alpha) Z(\alpha)|^2 d\alpha \right)^{1/2} \left(\int_0^1 |f(\alpha)|^4 d\alpha \right)^{1/2} \\ &\ll \big(P|\mathfrak{Z}| + |\mathfrak{Z}|^2 \big)^{1/2} P^{1+\epsilon/2} \ll |\mathfrak{Z}|^{1/2} P^{3/2+\epsilon/2}. \end{aligned}$$

From this we quickly obtain $|\mathfrak{Z}| \ll N^{5/14+\epsilon}$ as required.

Acknowledgement. This paper was written while the second named author enjoyed the benefits of a postdoctoral Instructorship at the University of Texas at Austin. He would like to take this opportunity to express his gratitude to the Department of Mathematics for the support and for the excellent working conditions.

REFERENCES

[**1**] R. C. BAKER, G. HARMAN and J. PINTZ. The exceptional set for Goldbach's problem in short intervals. In *Sieve Methods, Exponential Sums and their Applications in Number Theory* LMS Lecture Notes vol. 237 (Cambridge University Press, 1997), pp. 1–54.

[**2**] C. BAUER, M. C. LIU and T. ZHAN. On sums of three prime squares. *J. Number Theory* **85** (2000), 336–359.

[**3**] S. K. K. CHOI and A. V. KUMCHEV. Quadratic equations with five prime unknowns. *J. Number Theory* **107** (2004), 357–367.

[**4**] H. DAVENPORT. *Multiplicative Number Theory (3rd edition, revised by H. L. Montgomery).* Graduate Texts in Math. vol. 74 (Springer–Verlag, 2000).

[**5**] A. GHOSH. The distribution of αp^2 modulo 1. *Proc. London Math. Soc.* (3) **42** (1981), 252–269.

[**6**] G. HARMAN. On the distribution of αp modulo one. *J. London Math. Soc.* (2) **27** (1983), 9–18.

[**7**] G. HARMAN. On the distribution of αp modulo one II. *Proc. London Math. Soc.* (3) **72** (1996), 241–260.

[**8**] G. HARMAN. The values of ternary quadratic forms at prime arguments. To appear in *Mathematika*.

[**9**] L. K. HUA. Some results in additive prime number theory. *Quart. J. Math. Oxford* **9** (1938), 68–80.

[**10**] A. V. KUMCHEV. On Weyl sums over primes and almost primes. To appear.

[**11**] M. C. LEUNG and M. C. LIU. On generalized quadratic equations in three prime variables. *Mh. Math.* **115** (1993), 133–169.

[**12**] J. Y. LIU. On Lagrange's theorem with prime variables. *Quart. J. Math. Oxford* (2) **54** (2003), 453–462.

[**13**] J. Y. LIU, T. D. WOOLEY and G. YU. The quadratic Waring–Goldbach problem. *J. Number Theory* **107** (2004), 298–321.

[**14**] J. Y. LIU and T. ZHAN. Sums of five almost equal prime squares II. *Sci. China* **41** (1998), 710–722.

[**15**] J. Y. LIU and T. ZHAN. Distribution of integers that are sums of three squares of primes. *Acta Arith.* **98** (2001), 207–228.

[**16**] J. Y. LIU and T. ZHAN. An iterative method in the Waring–Goldbach problem. To appear.

[**17**] H. MIKAWA. On sums of three squares of primes. In *Analytic Number Theory* LMS Lecture Notes vol. 247 (Cambridge University Press, 1997), pp. 253–264.

[**18**] W. SCHWARZ. Zur Darstellun von Zahlen durch Summen von Primzahlpotenzen. *J. reine angew. Math.* **206** (1961), 78–112.

[**19**] R. C. VAUGHAN. *The Hardy–Littlewood Method (2nd edition).* Cambridge Tracts in Mathematics vol. 125 (Cambridge University Press, 1997).

[**20**] T. D. WOOLEY. Slim exceptional sets for sums of four squares. *Proc. London Math. Soc.* (3) **85** (2002), 1–21.