COMPOSITIO MATHEMATICA

Sums of smooth squares

V. Blomer, J. Brüdern and R. Dietmann

Compositio Math. **145** (2009), 1401–1441.

doi:10.1112/S0010437X09004254

FOUNDATION COMPOSITIO MATHEMATICA

The London Mathematical Society

Compositio Math. 145 (2009) 1401–1441
doi:10.1112/S0010437X09004254

Sums of smooth squares

V. Blomer, J. Brüdern and R. Dietmann

Abstract

Let $R(n,\theta)$ denote the number of representations of the natural number n as the sum of four squares, each composed only with primes not exceeding $n^{\theta/2}$. When $\theta > e^{-1/3}$ a lower bound for $R(n,\theta)$ of the expected order of magnitude is established, and when $\theta > 365/592$, it is shown that $R(n,\theta) > 0$ holds for large n. A similar result is obtained for sums of three squares. An asymptotic formula is obtained for the related problem of representing an integer as the sum of two squares and two squares composed of small primes, as above, for any fixed $\theta > 0$. This last result is the key to bound $R(n,\theta)$ from below.

1. Introduction

All natural numbers are the sum of four integral squares. The first proof of this very classical result in the theory of numbers is due to Lagrange, and Jacobi and Kloosterman have shown that the number of representations of the natural number n in the proposed manner equals

$$8(2+(-1)^n)\sum_{\substack{d\mid n\\ d\equiv 1 \bmod 2}} d = \pi^2\mathfrak{S}(n)n \tag{1.1}$$

where

$$\mathfrak{S}(n) = \sum_{q=1}^{\infty}\sum_{\substack{a=1\\(a,q)=1}}^{q} q^{-4}\bigg(\sum_{b=1}^{q} e\Big(\frac{ab^2}{q}\Big)\bigg)^4 e\Big(-\frac{an}{q}\Big) \tag{1.2}$$

is the singular series associated with sums of four squares.

Following a suggestion of Sarkőzy, we investigate here the question whether large n are the sum of four smooth squares, that is, of squares composed only of small prime factors. Let $\theta > 0$, and let $R(n,\theta)$ denote the number of solutions of the Diophantine equation

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = n \tag{1.3}$$

in non-zero integers x_1, x_2, x_3, x_4 subject to the constraint that whenever p is a prime with $p \mid x_1x_2x_3x_4$ then $p \leqslant n^{\theta/2}$. Note that $R(n,1)$ is the number of representations of n as the sum of four non-zero squares, and therefore differs from the number described in (1.1) by at most $O(n^{1/2+\varepsilon})$. A formal use of the Hardy–Littlewood method yields an asymptotic formula for $R(n,\theta)$ with a main term that coincides with (1.1) save for a positive constant factor. In particular, for any fixed $\theta > 0$, it is expected that $R(n,\theta) \gg \mathfrak{S}(n)n$, and hence

Received 2 August 2007, accepted in final form 6 March 2009.
2000 Mathematics Subject Classification 11P05 (primary), 11N36, 11E25 (secondary).
Keywords: sums of squares, smooth numbers, enveloping sieve.
The first author acknowledges the support of an NSERC grant.

$R(n, \theta) > 0$ should hold for sufficiently large n. We are able to confirm these conclusions only for rather large values of θ.

Theorem 1. *Let $\theta > e^{-1/3}$. Then $R(n, \theta) \gg_\theta \mathfrak{S}(n)n$.*

Though we fail to establish an asymptotic formula for $R(n, \theta)$, the lower bound is of the expected order of magnitude. At the cost of a weaker lower bound, it is possible to treat somewhat smaller values of θ.

Theorem 2. *Let $\theta > 365/592$. Then $R(n, \theta) > 0$ for all sufficiently large n.*

Note that $e^{-1/3} = 0.716\ldots$ and $365/592 = 0.616\ldots$.

Our results appear to be the first in the literature that yield a representation of all large n as the sum of four smooth squares. Earlier, Harcos [Har99] has shown that almost all natural numbers are the sum of four squares of integers, the greatest prime divisor of which does not exceed $\exp(20(\log n \log\log n)^{1/2})$. Wooley [Woo02] gave a quantitative form of this and proved that for any fixed $\theta > 0$, one has $R(n, \theta) > 0$ for all but $O(N^{1/2-\theta/1600})$ of the natural numbers n not exceeding N.

Many other variants of the four squares theorem with multiplicative constraints on the variables have been considered. Most prominent in this circle of ideas is the Waring–Goldbach problem. All large $n \equiv 4 \bmod 24$ should be the sum of four squares of primes, but again this is only known for n outside a slim set (Wooley [Woo02]). The best approximation known today is due to Tolev [Tol03], improving work of Brüdern and Fouvry [BF94] and Heath-Brown and Tolev [HT03], to the effect that for large n, (1.3) admits solutions in integers with at most 21 prime factors. Perhaps also of interest is a comparison of these results with the current state of the art in a linear analogue of our problem, such as the binary Goldbach problem. We owe to Chen [Che73] the famous proposition that all large even numbers are the sum of a prime and a number with at most two prime factors, and Balog [Bal89] proved that all large numbers n are the sum of two natural numbers whose prime factors do not exceed $n^{4/(9\sqrt{e})}$. As in the quadratic case, the 'smooth' version appears weakish compared to the sieve results with almost-primes. Additive problems with smooth numbers, although so powerful in Waring's problem (see Vaughan and Wooley [VW02]), seem to be intrinsically difficult if one cannot use a suitable mean value estimate of Hua's lemma type.

The methods of proof for Theorems 1 and 2 are very different. The idea for Theorem 2 also appears in Balog [Bal89], and is readily described. We make use of the distribution of smooth numbers in short intervals. Roughly speaking, for any large $m \in \mathbb{N}$ there exists an integer k in a prescribed residue class modulo 8 with its largest prime factor not exceeding $\sqrt{m}$, and $m - 100m^{1/4} < k \leqslant m$, see § 5 for details. Hence, for given $n \in \mathbb{N}$, we can choose x_4 with largest prime factor not exceeding $n^{1/4}$, and $0 \leqslant n - x_4^2 \ll n^{5/8}$, and by adjusting the congruence class of x_4 modulo 8 appropriately, $n - x_4^2$ is a sum of three squares of integers that trivially cannot exceed $n^{5/16}$. This would lead to $R(n, 5/8) > 0$ for large n. Note that the weak point in this argument is that the size of the prime factors of three variables remain uncontrolled. One can do slightly better.

Theorem 3. *Let $\varepsilon > 0$, and let $n \not\equiv 0, 4, 7 \bmod 8$ be sufficiently large. Then n is the sum of three squares of integers whose largest prime factors do not exceed $n^{73/148+\varepsilon}$.*

Theorem 2 follows from Theorem 3 in the manner indicated above. Theorem 3 is a first attempt towards a result on sums of three smooth squares. It follows with little effort from uniform estimates for Fourier coefficients of cusp forms of half integral weight that were obtained by the first author [Blo04]; see also [Blo08] for refinements. The details will be provided in § 5.

The approach to Theorem 1 takes the inclusion–exclusion principle as a starting point. This provides a link between $R(n,\theta)$ and related counting functions for sums of four squares in which only two or one of the variables are smooth; see § 4. It is this part of the argument that forces us to take $\theta > e^{-1/3}$. The new counting functions contain at least two squares that are free of constraints, and we treat them as a binary additive problem by writing them in the form

$$\sum_{m\leqslant n} r(m)f(n-m) \tag{1.4}$$

where $r(m)$ is the number of representations of m as the sum of two integral squares, and f is a suitable function. The main ideas go back to Hooley [Hoo57] who considered the case where f is the indicator function on the set of primes. His ideas were used by Shields [Shi79], Plaksin [Pla81, Pla84] and Kovalchik [Kov82] to obtain an asymptotic formula for the number of representations of n as the sum of two squares and two squares of primes. These methods may be modified so as to cover the case where $f(k)$ is the number of solutions of $y_1^2+y_2^2=k$ with either y_1 or y_1 and y_2 smooth, and this will ultimately yield Theorem 1. Our principal conclusions in this context may be of some independent interest, and therefore we describe them as another theorem.

For $0<\theta\leqslant 1$ and $1\leqslant j\leqslant 4$ let $R_j(n,\theta)$ denote the number of solutions of $x_1^2+x_2^2+x_3^2+x_4^2=n$ with the greatest prime factor of $x_1\cdots x_j$ not exceeding $n^{\theta/2}$. Note that $R_4(n,\theta)=R(n,\theta)$. When $j=1$ or 2, we are able to evaluate $R_j(n,\theta)$ asymptotically. The result features Dickman's function, defined as the continuous function $\varrho:[0,\infty)\to(0,1]$ defined by $\varrho(u)=1$ for $0\leqslant u\leqslant 1$ and the delay equation $\varrho(u-1)+u\varrho'(u)=0$ for $u>1$.

Theorem 4. *Let $0<\theta\leqslant 1$ and $j=1$ or 2. Then, for all n with $4\nmid n$ one has*

$$R_j(n,\theta)=\varrho(1/\theta)^j\pi^2\mathfrak{S}(n)n+O(n(\log n)^{-1/30}).$$

It is certainly possible to impose still stronger smoothness conditions in Theorem 4, with θ tending slowly to zero as n tends to infinity.

It would be interesting to prove a similar result for $R_3(n,\theta)$. Professor Wooley has informed us that he has now shown that for any $\theta>0$ there are no more than $O((\log N)^5)$ integers $n\leqslant N$ for which the expected asymptotic formula for $R_3(n,\theta)$ may fail.

Notational conventions. In this paper we apply the common ε-convention: whenever ε occurs in a statement it is asserted that the statement holds for any fixed real positive value of ε. Implicit constants in the Vinogradov- or Landau-symbols may depend on ε. Note that this convention allows us to conclude from $S\ll X^\varepsilon$, $T\ll X^\varepsilon$ that $ST\ll X^\varepsilon$, for example. We also apply the same convention with the letter A in place of ε. Usually our assertions are meaningful only when ε is very small whereas we use A in situations where we have in mind large values of A. Summations start at 1 unless indicated otherwise. Small latin letters usually denote natural numbers except x,y,z that are integers, and p is reserved for primes. The greatest common factor of a and b is $(a;b)$ whereas (a,b) is a pair; $[a;b]$ is the lowest common multiple, and $[a,b]$ denotes a closed interval of real numbers.

2. Auxiliary tools

2.1 Smooth numbers

We summarize here some results on the distribution of smooth numbers. Let $P(k)$ denote the greatest prime factor of the natural number $k \geqslant 2$, and put $P(1)=1$, $P(0)=\infty$. Following standard notational practice, let

$$\Psi(X,Y;q,a) = \#\{k \leqslant X : k \equiv a \bmod q,\ P(k) \leqslant Y\}, \tag{2.1}$$

$$\Psi_q(X,Y) = \#\{k \leqslant X : (k;q)=1,\ P(k) \leqslant Y\}. \tag{2.2}$$

Fouvry and Tenenbaum have shown that for $(a;q)=1$ one has

$$\Psi(X,Y;q,a) - \frac{\Psi_q(X;Y)}{\varphi(q)} \ll \frac{X}{(\log X)^A}. \tag{2.3}$$

Note that there is no restriction on the ranges of q, X or Y here. This very weak form of a Siegel–Walfisz theorem for smooth numbers follows from [FT91, Théorème 6], for example. In our application, however, a and q may have common factors. Therefore we put

$$E(X,Y;q,a) = \varphi\biggl(\frac{q}{(q;a)}\biggr)^{-1} \Psi_{q/(q;a)}(X/(q;a),Y) - \Psi(X,Y;q,a). \tag{2.4}$$

Lemma 2.1. *Let $0<\delta<1$. Then, whenever $X^\delta \leqslant Y \leqslant X$, one has*

$$E(X,Y;q,a) \ll X(\log X)^{-A}.$$

Proof. When $(a;q)=1$ this is contained in (2.3). Now put $d=(q;a)$ and note that $\Psi(X,Y;q,a) = \Psi(X/d,Y;q/d,a/d)$ unless $P(d)>Y$ in which case $\Psi(X,Y;q,a)=0$. In the first case we apply (2.3) to $\Psi(X/d,Y;q/d,a/d)$ and immediately confirm the conclusion in Lemma 2.1. In the latter case, note that $P(d)>Y$ implies $d>Y$ whence

$$E(X,Y;q,a) = (\varphi(q/d))^{-1}\Psi_{q/d}(X/d,Y) \ll X/d \ll X^{1-\delta}.$$

This completes the proof. □

Lemma 2.2. *Let $0<\delta<1$. Then, whenever $X^\delta \leqslant Y \leqslant X$, one has*

$$\sum_{q\leqslant Q} \sum_{\substack{a=1\\(a;q)=1}}^{q} E(X,Y;q,a)^2 \ll QX + X^2(\log X)^{-A}.$$

This lemma is a weak form of a Barban–Davenport–Halberstam theorem for smooth numbers. It is nowadays routine to deduce such estimates from the corresponding Siegel–Walfisz theorem (our Lemma 2.1) via the large sieve. Nonetheless the only explicit reference for Lemma 2.2 that we are aware of is to the recent dissertation of Neumann [Neu06] where a stronger version of Lemma 2.2 is the starting point for a detailed study of the distribution of smooth numbers in arithmetic progressions. However, the methods of either Hooley [Hoo75, Hoo98] or Vaughan [Vau98] readily yield Lemma 2.2 in its present form, and we may leave it to the reader to give a detailed proof along the indicated lines, or to consult [Neu06].

Before we leave the subject of estimating E we record here the uniform bound

$$E(X+H,Y;q,a) - E(X,Y;q,a) \ll \frac{H}{q} + 1 \tag{2.5}$$

that follows from the definition of E, (2.1), (2.2) and the elementary observation that at most $O(H/q+1)$ integers in an interval of length H belong to an arithmetic progression of modulus q.

Our next theme is a continuous approximation for $\Psi_q(X, Y)$. This features the Dickman function ϱ, Euler's constant γ, and the arithmetical function

$$\Pi(q) = \sum_{p|q} \frac{\log p}{p-1}. \tag{2.6}$$

Then, if $\zeta(s)$ denotes the Riemann zeta function, one readily checks that

$$\frac{s\zeta(1+s)}{1+s} \prod_{p|q} (1 - p^{-1-s}) = \frac{\varphi(q)}{q}(1 + s(\Pi(q) + \gamma - 1) + O(|s|^2)),$$

and hence, by Corollaire 2 of Fouvry and Tenenbaum [FT91] (with $k = 1$), for any fixed $A > 0$, $0 < \delta < 1$,

$$\Psi_q(X, Y) = \frac{\varphi(q)}{q} X\left(\varrho\left(\frac{\log X}{\log Y}\right) + \varrho'\left(\frac{\log X}{\log Y}\right)\frac{\Pi(q) + \gamma - 1}{\log Y}\right) + O\left(X\left(\frac{\log\log qY}{\log Y}\right)^2\right) \tag{2.7}$$

holds uniformly in $q \leqslant X^A$, $X^\delta \leqslant Y \leqslant X$. In particular, we may insert this into (2.4) and then conclude as follows.

LEMMA 2.3. *Let $0 < \delta < 1$. Then, for $X^\delta \leqslant Y \leqslant X$ and any a, q with $q \leqslant X^A$ one has*

$$\begin{aligned}\Psi(X, Y; q, a) = \frac{X}{q}\left(\varrho\left(\frac{\log(X/(q;a))}{\log Y}\right) + \varrho'\left(\frac{\log(X/(q;a))}{\log Y}\right)\frac{\Pi(q/(q;a)) + \gamma - 1}{\log Y}\right) \\ - E(X, Y; q, a) + O\left(\frac{X}{q}\frac{(\log\log X)^3}{(\log X)^2}\right).\end{aligned}$$

We conclude our discussion of sums over smooth numbers with a character sum estimate.

LEMMA 2.4. *Let χ denote the non-principal character modulo 4. Let $X \geqslant 10$ and $Y = \exp(\log X/\log\log X)$. Then, uniformly in $1 \leqslant Z_1 \leqslant Z_2 \leqslant X$, one has*

$$\sum_{\substack{Z_1 \leqslant k \leqslant Z_2 \\ P(k) \leqslant Y}} \frac{\chi(k)}{k} \ll Z_1^{-1} + \exp(-(\log X)^{2/5}).$$

Proof. A crude application of Theorème 4 of Fouvry and Tenenbaum [FT91] shows that there is a constant $c > 0$ such that whenever $Y \leqslant Z \leqslant X$ one has

$$\sum_{\substack{k \leqslant Z \\ P(k) \leqslant Y}} \chi(k) \ll Z \exp(-c\sqrt{\log Y}).$$

When $1 \leqslant Z \leqslant Y$ and $k \leqslant Z$, the condition $P(k) \leqslant Y$ is void, and hence

$$\sum_{\substack{k \leqslant Z \\ P(k) \leqslant Y}} \chi(k) \ll 1.$$

We split the sum over $Z_1 \leqslant k \leqslant Z_2$ into two parts $Z_1 \leqslant k \leqslant Y$, $Y < k \leqslant Z_2$ if necessary, and then apply partial summation to see that

$$\sum_{\substack{Z_1 \leqslant k \leqslant Z_2 \\ P(k) \leqslant Y}} \frac{\chi(k)}{k} \ll Z_1^{-1} + (\log X)\exp(-c\sqrt{\log Y}),$$

which is stronger than asserted in the lemma. □

2.2 Divisor sums

The later part of the proof of Theorem 4 features several divisor sums that are of a more recondite nature. The basic principles derive from work of Hooley [Hoo57, Hoo79]. We begin with a refinement of the classical estimate

$$\sum_{d|n} \frac{\tau(d)^k}{d} \ll (\log\log n)^{2^k} \tag{2.8}$$

that is within the scope of elementary prime number theory.

Lemma 2.5. *Let $0 < \alpha \leqslant 1/4$ and k be a non-negative integer. Then*

$$\sum_{\substack{d|n \\ d>\exp((\log n)^\alpha)}} \frac{\tau(d)^k}{d} \ll \exp\Big(-\frac{1}{2}\alpha(\log n)^\alpha\Big). \tag{2.9}$$

Proof. We work somewhat more generally than necessary for the immediate needs. Let $N \geqslant 1$ and $\alpha \geqslant 0$. Then

$$\sum_{\substack{d|n \\ d>N}} \frac{\tau(d)^k}{d} \leqslant \sum_{d|n} \frac{\tau(d)^k}{d}\Big(\frac{d}{N}\Big)^\alpha \leqslant N^{-\alpha} \prod_{p|n} \sum_{l=0}^{\infty} \tau(p^l)^k p^{(\alpha-1)l}.$$

For $0 \leqslant \alpha \leqslant 1/4$ one has $p^{\alpha-1} \leqslant p^{-3/4} \leqslant 2^{-3/4}$. The previous inequality now simplifies to

$$\sum_{\substack{d|n \\ d>N}} \frac{\tau(d)^k}{d} \leqslant N^{-\alpha} \prod_{p|n}\Big(1 + \frac{2^k}{p^{1-\alpha}} + \frac{B_k}{p^{3/2}}\Big) \leqslant C_k N^{-\alpha} \prod_{p|n}(1 + 2^k p^{\alpha-1}) \tag{2.10}$$

where

$$B_k = \sum_{l=0}^{\infty} (l+3)^k 2^{-(3/4)l}, \quad C_k = \prod_p (1 + B_k p^{-3/2}).$$

The inequality $\log(1+t) \leqslant t$ suffices to bound the product in (2.10) by

$$\log \prod_{p|n}(1 + 2^k p^{\alpha-1}) \leqslant 2^k \sum_{p|n} p^{\alpha-1} \leqslant 2^k \sum_{p \leqslant \log n} p^{\alpha-1} + 2^k \sum_{\substack{p|n \\ p>\log n}} p^{\alpha-1}. \tag{2.11}$$

For the second summand, note that no more than $\log n/\log\log n$ primes $p > \log n$ may divide n. It follows that

$$\sum_{\substack{p|n \\ p>\log n}} p^{\alpha-1} \leqslant (\log n)^\alpha (\log\log n)^{-1}.$$

We now take $\alpha = 0$ and $N = 1$. Then, by one of Mertens' asymptotic formulae, one has

$$\sum_{p \leqslant Y} \frac{1}{p} \leqslant \log\log Y + C, \tag{2.12}$$

where C is a certain positive constant. We use this with $Y = \log n$, and then deduce (2.8) from (2.10) and (2.11). To establish Lemma 2.5, fix $0 < \alpha \leqslant 1/4$ and take $N = \exp((\log n)^\alpha)$. By Chebychev's estimates and partial summation, one finds that there is $C > 0$ (depending on α only) with

$$\sum_{p \leqslant Y} p^{\alpha-1} \leqslant C Y^\alpha (\log Y)^{-1}.$$

Here we take $Y = \log n$ again and then find from (2.11) that

$$\prod_{p|n}(1+2^k p^{\alpha-1}) \leqslant \exp\left(2^k(C+1)\frac{(\log n)^\alpha}{\log\log n}\right)$$

whereas $N^{-\alpha} = \exp(-\alpha(\log n)^\alpha)$. Hence, for n sufficiently large in terms of k, we find from (2.10) that the sum (2.9) does not exceed $N^{-\alpha/2}$, as required to confirm Lemma 2.5. □

We require another variant of (2.8) that underpins much of Hooley's work [Hoo57, Hoo79]. To describe this in detail, fix a constant $0 < \delta \leqslant 1$, and let

$$Y = \exp\left(\frac{\log n}{\log\log n}\right).$$

Then, we shall prove the inequality

$$\sum_{\substack{P(a)\leqslant Y\\ a>n^\delta}} \frac{\mu(a)^2}{a} \ll (\log n)^{-\delta(\log\log\log n)/3}. \tag{2.13}$$

Note that the sum in (2.13) is of the type we currently consider. In fact, if Q denotes the product of all primes $p \leqslant Y$, then

$$\sum_{\substack{P(a)\leqslant Y\\ a>n^\delta}} \frac{\mu(a)^2}{a} = \sum_{\substack{d|Q\\ d>n^\delta}} \frac{1}{d},$$

and we may now argue as in the reasoning leading to (2.10). Then, for any $\alpha > 0$ we find that

$$\sum_{\substack{P(a)\leqslant Y\\ a>n^\delta}} \frac{\mu(a)^2}{a} \leqslant n^{-\delta\alpha} \prod_{p\leqslant Y}(1+p^{\alpha-1}). \tag{2.14}$$

It will be convenient to write temporarily $\ell = (1/2)\log\log\log n$. We take $\alpha = \ell(\log Y)^{-1}$. Then, for $2 \leqslant p \leqslant Y$, one has

$$p^\alpha = \exp\left(\ell\frac{\log p}{\log Y}\right) \leqslant 1 + \frac{\log p}{\log Y}\ell e^\ell,$$

for example by the mean value theorem. Consequently,

$$\sum_{p\leqslant Y} p^{\alpha-1} \leqslant \sum_{p\leqslant Y}\frac{1}{p} + \ell e^\ell(\log Y)^{-1}\sum_{p\leqslant Y}\frac{\log p}{p}.$$

We may now use (2.11) and the other Mertens formula

$$\sum_{p\leqslant Y}\frac{\log p}{p} \leqslant \log Y + C$$

(for suitable $C > 0$) to see that

$$\sum_{p\leqslant Y} p^{\alpha-1} \leqslant \log\log Y + C + 2\ell e^\ell$$

holds for all sufficiently large n. Taking exponentials, we infer that the right-hand side of (2.14) is bounded by a value not exceeding

$$e^C(\log Y)n^{-\delta\alpha}\exp(2\ell e^\ell) \leqslant e^C(\log n)\exp(-\delta\ell(\log\log n) + 2\ell(\log\log n)^{1/2}),$$

and (2.13) follows.

https://doi.org/10.1112/S0010437X09004254 Published online by Cambridge University Press

A slightly weaker form of (2.13) occurs in Plaksin [Pla81, Lemma 13]. Our proof, however, is very different. Unfortunately, (2.13) is not quite sufficient for our purposes, but the following variant will do.

Lemma 2.6. *Let $0<\delta<1$, $k\in\mathbb{N}$, and $Y=\exp(\log n/\log\log n)$. Then*

$$\sum_{\substack{P(a)\leqslant Y\\ a>n^\delta}} \frac{\tau(a)^k}{\varphi(a)} \ll (\log n)^{-A}.$$

Proof. By Cauchy's inequality,

$$\sum_a \frac{\tau(a)^k}{\varphi(a)} \leqslant \bigg(\sum_a \frac{1}{a}\bigg)^{1/2}\bigg(\sum_a \frac{a\tau(a)^{2k}}{\varphi(a)^2}\bigg)^{1/2}$$

where all sums are over integers a with $P(a)\leqslant Y$ and $a>n^\delta$. Now

$$\sum_{P(a)\leqslant Y} \frac{a\tau(a)^{2k}}{\varphi(a)^2} = \prod_{p\leqslant Y}\sum_{h=0}^\infty \frac{p^h(h+1)^{2k}}{\varphi(p^h)^2} \leqslant \prod_{p\leqslant n}\bigg(1+\frac{4^k}{p}+O_k\bigg(\frac{1}{p^2}\bigg)\bigg)$$

so that the second factor in the previous inequality does not exceed $(\log n)^{4k}$. To estimate the first factor, write $a=uv^2$ where $\mu(u)^2=1$; this decomposition is unique. Now

$$\sum_{\substack{P(a)\leqslant Y\\ a>n^\delta}} \frac{1}{a} \leqslant \sum_{v=1}^\infty \frac{1}{v^2}\sum_{\substack{P(u)\leqslant Y\\ uv^2>n^\delta}} \frac{1}{u}. \tag{2.15}$$

First consider the contribution of terms with $u>n^{\delta/2}$. We may then use (2.13) with $\delta/2$ in place of δ to see that these terms contribute to (2.15) at most $O((\log n)^{-4A})$. In the complementary portion, $u\leqslant n^{\delta/2}$, the sum over u in (2.15) is $O(\log n)$, and since now $v>n^{\delta/4}$, the outer sum is $O(n^{-\delta/4})$. This shows that (2.15) is $O((\log n)^{-4A})$, and the lemma follows. □

Our final divisor sum estimate rests on the sparsity of integers that have a divisor in a prescribed small interval. Hooley [Hoo79] made important contributions to this subject, and we base our analysis on powerful and essentially best possible estimates of Tenenbaum [Ten84].

Lemma 2.7. *Let $X(n)=\exp((\log n)^{1/250})$. Let $\mathcal{S}$ denote the set of all integers $s\leqslant n$ that have a divisor in the interval $[\sqrt{n}X^{-1},\sqrt{n}X]$. Then*

$$\sum_{s\in\mathcal{S}} \frac{1}{s} \ll (\log n)^{37/40}.$$

Proof. Let $S=\sqrt{n}\exp((\log n)^{23/25})$. Note that any $s\in\mathcal{S}$ must exceed $\sqrt{n}X^{-1}$, and we have

$$\sum_{\sqrt{n}X^{-1}\leqslant s\leqslant S} \frac{1}{s} \ll \log\frac{SX}{\sqrt{n}} \ll (\log n)^{23/25}. \tag{2.16}$$

Hence, it remains to consider $s\in\mathcal{S}$ with $s>S$. Now let $\Upsilon(x,y)$ denote the number of integers $k\in[x,2x]$ that have a divisor in the interval $[y,2y]$. Since $[\sqrt{n}X^{-1},\sqrt{n}X]$ can be covered by $O((\log n)^{1/250})$ intervals $[y,2y]$, a dyadic splitting-up argument shows

$$\sum_{\substack{s\in\mathcal{S}\\ s>S}} \frac{1}{s} \ll (\log n)^{251/250} \max_{\substack{S\leqslant x\leqslant n\\ \sqrt{n}X^{-1}\leqslant y\leqslant\sqrt{n}X}} \frac{\Upsilon(x,y)}{x}. \tag{2.17}$$

We now invoke the fundamental inequality

$$\Upsilon(x, y) \leqslant x(\log y)^{-0.086} \quad (3 \leqslant y \leqslant \sqrt{2x}); \tag{2.18}$$

this is a crude version of Hall and Tenenbaum [HT88, (2.2)]. In particular, since $S > \sqrt{n}$, for pairs x, y with $y \leqslant \sqrt{2x}$ in (2.17), we have $\Upsilon(x, y) \ll x(\log x)^{-0.086}$. In the opposite case $y > \sqrt{2x}$, we cannot use (2.18) directly, but if $x \leqslant k \leqslant 2x$ and k has a divisor $d \in [y, 2y]$, then k/d also divides k and satisfies $x/2y \leqslant k/d \leqslant 2x/y$. Hence, by (2.18) with y replaced by $x/2y$ and x/y, we see that when $y > \sqrt{2x}$ we have

$$\Upsilon(x, y) \leqslant 2x\left(\log \frac{x}{2y}\right)^{-0.086} \leqslant 2x\left(\log \frac{S}{2\sqrt{n}X}\right)^{-0.086}.$$

One now checks that $\Upsilon(x, y) \ll x(\log x)^{-0.079}$ for $y > \sqrt{2x}$ in (2.17), and the lemma follows from (2.16) and (2.17). $\square$

2.3 Sums of two squares

The next three lemmata concern the arithmetical function $\lambda_n(q)$ defined by

$$q\lambda_n(q) = \#\{1 \leqslant a_1, a_2 \leqslant q : a_1^2 + a_2^2 \equiv n \bmod q\}. \tag{2.19}$$

As a function of q, this expression is multiplicative. Other properties are listed below.

Lemma 2.8. *For any $q \in \mathbb{N}, n \in \mathbb{N}$, the function $\lambda_n(q)$ satisfies the inequalities*

$$\lambda_n(q) \ll \tau(q), \quad \lambda_n(q) \ll \frac{q}{\varphi(q)}\tau((q; n)).$$

If q is odd and $(q; n) = 1$, then $\lambda_n(q) > 0$.

Proof. We begin by examining $\lambda_n(p^l)$ when p is a prime. The reader will readily confirm the elementary inequalities $0 \leqslant \lambda_n(2^l) \leqslant 4$ for all $n \in \mathbb{N}, l \in \mathbb{N}$ with the aid of the fact that an odd number is a square modulo 2^l $(l \geqslant 3)$, if and only if it is congruent to 1 modulo 8. Therefore we may concentrate on the case where p is odd. Define the Gauß sum

$$G(q, a) = \sum_{x=1}^{q} e\left(\frac{ax^2}{q}\right)$$

and note that by (2.19) and orthogonality, one has

$$\lambda_n(q) = q^{-2}\sum_{a=1}^{q} G(q, a)^2 e\left(-\frac{an}{q}\right).$$

When p is odd and $p \nmid a$, then the explicit evaluation of Gauß sums (see Estermann [Est45], for example) shows that

$$\left(\frac{G(p^k, a)}{p^k}\right)^2 = \chi(p)^k p^{-k};$$

here χ is the non-trivial character modulo 4, as before. Moreover, $q^{-1}G(q, b) = \tilde{q}^{-1}G(\tilde{q}, \tilde{b})$ whenever $b/q = \tilde{b}/\tilde{q}$, as one confirms directly from the definition of G. Combining these facts one finds that for any $l \geqslant 1$ one has

$$\lambda_n(p^l) = \sum_{k=0}^{l} \chi(p)^k p^{-k} c_{p^k}(n) \tag{2.20}$$

where

$$c_q(n) = \sum_{\substack{a=1 \\ (a;q)=1}}^{q} e\left(\frac{an}{q}\right)$$

is Ramanujan's sum. By Theorems 67 and 272 of Hardy and Wright [HW79], one has

$$c_q(n) = \varphi(q)\frac{\mu(q/(q;n))}{\varphi(q/(q;n))}. \tag{2.21}$$

Write $n = p^\nu n_0$ with $p \nmid n_0$. By (2.20) and (2.21), we now see that

$$\lambda_n(p^l) = \lambda_n(p^{\nu+1}) \quad (l \geqslant \nu+1), \tag{2.22}$$

and that for $p \equiv 1 \bmod 4$ one has

$$\lambda_n(p^l) = 1 + l\left(1 - \frac{1}{p}\right)\ (1 \leqslant l \leqslant \nu); \quad \lambda_n(p^{\nu+1}) = (\nu+1)\left(1 - \frac{1}{p}\right) \tag{2.23}$$

whereas when $p \equiv 3 \bmod 4$, the formulae

$$\lambda_n(p^l) = \begin{cases} \dfrac{1}{p} & (1 \leqslant l \leqslant \nu, l \text{ odd}), \\ 1 & (1 \leqslant l \leqslant \nu, l \text{ even}), \end{cases} \qquad \lambda_n(p^{\nu+1}) = \begin{cases} 1 + \dfrac{1}{p} & (\nu \text{ even}), \\ 0 & (\nu \text{ odd}), \end{cases} \tag{2.24}$$

are valid. In particular, (2.22), (2.23) and (2.24) imply $\lambda_n(p^l) \leqslant l + 1 = \tau(p^l)$. This confirms the first inequality in Lemma 2.8. When $l \leqslant \nu$, this also gives the inequality $\lambda_n(p^l) \leqslant \tau((p^l;p^\nu))(p^l/\varphi(p^l))$ which is weaker in this case. When $l > \nu$, we conclude from (2.22), (2.23) and (2.24) that

$$\lambda_n(p^l) = \lambda_n(p^{\nu+1}) \leqslant \left(1 + \frac{1}{p}\right)\tau(p^\nu) \leqslant \frac{p^l}{\varphi(p^l)}\tau(p^\nu),$$

and the second inequality in Lemma 2.8 also follows by multiplicativity. The final statement of the lemma is an immediate consequence of (2.22) and (2.24). □

More subtle than the readier accessible $\lambda_n(q)$, we also have to evaluate the quantity

$$\Lambda_d(q;n) = \#\{1 \leqslant x, y \leqslant q : (y;q) = d, x^2 + y^2 \equiv n \bmod q\}. \tag{2.25}$$

By (2.19) one has

$$q\lambda_n(q) = \sum_{d|q} \Lambda_d(q;n), \tag{2.26}$$

and much like the multiplicativity of $\lambda_n(q)$, the Chinese remainder theorem implies a quasi-multiplicativity for Λ. Whenever $d|q$, one finds that

$$\Lambda_d(q;n) = \prod_{p^l \| q} \Lambda_{(d;p^l)}(p^l;n). \tag{2.27}$$

We proceed to compute $\Lambda_{p^h}(p^{h+l};n)$ for all odd primes p and all non-negative integers h, l. The case $h = 0$ will be considered first. Trivially, $\Lambda_1(1;n) = 1$. When $l \geqslant 1$, then $\Lambda_1(p^l;n)$ is the number of solutions of $x^2 + y^2 \equiv n \bmod p^l$ with $p \nmid y$. Since a number is a quadratic residue modulo p^l if and only if this is so modulo p, we see that

$$\Lambda_1(p^l;n) = p^{l-1}\Lambda_1(p;n). \tag{2.28}$$

From (2.25) it is immediate that $\Lambda_p(p;n) = 1 + (\frac{n}{p})$, and (2.26) then shows that $\Lambda_1(p;n) = p\lambda_n(p) - 1 - (\frac{n}{p})$. By (2.20), it follows that

$$\Lambda_1(p;n) = \begin{cases} p - \chi(p) - \left(\dfrac{n}{p}\right) - 1 & (p \nmid n), \\ p + (p-1)\chi(p) - 1 & (p|n). \end{cases} \tag{2.29}$$

When $h \geqslant 1$ there is a similar formula. Write $n = p^\nu n_0$ with $p \nmid n_0$. Then, if $1 \leqslant h \leqslant \nu/2$, one has

$$\Lambda_{p^h}(p^{h+l};n) = \begin{cases} \varphi(p^l)p^{[(h+l)/2]} & (0 \leqslant l \leqslant h), \\ p^{h+l-1}\Lambda_1(p, np^{-2h}) & (l > h) \end{cases} \tag{2.30}$$

whereas for $h > \nu/2$ the corresponding result is

$$\Lambda_{p^h}(p^{h+l};n) = \begin{cases} \varphi(p^l)p^{[(h+l)/2]} & (h+l \leqslant \nu), \\ \varphi(p^l)p^{\nu/2}\left(1 + \left(\dfrac{n_0}{p}\right)\right) & (h+l > \nu, \nu \text{ even}), \\ 0 & (h+l > \nu,\ \nu \text{ odd}). \end{cases} \tag{2.31}$$

For a proof, note that for $h \geqslant 1, l \geqslant 0$, the definition (2.25) shows that

$$\Lambda_{p^h}(p^{h+l};n) = \#\{1 \leqslant x \leqslant p^{h+l}, 1 \leqslant y \leqslant p^l : x^2 + p^{2h}y^2 \equiv n \bmod p^{h+l}, p \nmid y\}. \tag{2.32}$$

In particular, when $h \geqslant l$, the congruence reduces to $x^2 \equiv n \bmod p^{h+l}$, and

$$\Lambda_{p^h}(p^{h+l};n) = \varphi(p^l)\#\{1 \leqslant x \leqslant p^{h+l} : x^2 \equiv n \bmod p^{h+l}\}.$$

By elementary number theory, the congruence $x^2 \equiv p^\nu n_0 \bmod p^t$ has $p^{[t/2]}$ solutions when $t \leqslant \nu$, it has $p^{\nu/2}(1 + (\frac{n_0}{p}))$ solutions when $t > \nu$ and ν is even, and it has no solutions when $t > \nu$ and ν is odd. We use this with $t = h + l$ to confirm (2.30) and (2.31) in all cases where $h \geqslant l$.

Now suppose that $l > h \geqslant 1$ and $2h \leqslant \nu$. Then the congruence in (2.32) implies that $x^2 \equiv 0 \bmod p^{2h}$, and we may substitute $x = p^h z$ to reduce that congruence to $z^2 + y^2 \equiv np^{-2h} \bmod p^{l-h}$. On considering the ranges for z and y implied by (2.32), one finds that $\Lambda_{p^h}(p^{h+l};n) = p^{2h}\Lambda_1(p^{l-h}; np^{-2h})$, and (2.30) follows from (2.28).

It remains to examine the case where $l > h \geqslant 1$ and $2h > \nu$. Here (2.32) implies $x^2 \equiv p^\nu n_0 \bmod p^{2h}$. Hence, if ν is odd, there is no solution, in accordance with (2.31). If ν is even, then $p^{\nu/2}|x$. We write $x = p^{\nu/2}z$ in (2.32). It follows that $\Lambda_{p^h}(p^{h+l};n)$ equals the number of solutions of

$$z^2 + p^{2h-\nu}y^2 \equiv n_0 \bmod p^{h+l-\nu} \tag{2.33}$$

with $1 \leqslant z \leqslant p^{h+l-\nu/2}$ and y as in (2.32). Since $2h > \nu$ in the current context, there is no solution of (2.33) unless $(\frac{n_0}{p}) = 1$. In the latter case, $n_0 - p^{2h-\nu}y^2$ is a quadratic residue modulo p, and hence also modulo $p^{h+l-\nu}$, for each of the $\varphi(p^l)$ values of y in (2.32). For a particular value of y, there are two solutions for z modulo $p^{h+l-\nu}$ in (2.33), and hence $2p^{\nu/2}$ solutions with $1 \leqslant z \leqslant p^{h+l-\nu/2}$. This confirms (2.31).

We require estimates for certain convolutions involving $\lambda_n(q)$ that do not follow from Lemma 2.8 alone. The next lemma is a typical example.

Lemma 2.9. *Let q be odd and $(n;q) = 1$. Then*

$$0 < \lambda_n(q) \sum_{d|q} \frac{\mu(d)}{d\lambda_n(q/d)} \ll \frac{\varphi(q)}{q}.$$

https://doi.org/10.1112/S0010437X09004254 Published online by Cambridge University Press

Proof. If $(n;q)=1$, then $\lambda_n(d)>0$ for any divisor $d|q$, by Lemma 2.8. Since λ_n is multiplicative, it suffices to prove Lemma 2.9 in the special case where $q=p^l$ is a power of an odd prime p, $p\nmid n$. When $l\geqslant 2$, then $\lambda_n(p^l)=\lambda_n(p)$ by (2.22), and hence, in this case,

$$\lambda_n(p^l)\sum_{d|p^l}\frac{\mu(d)}{d\lambda_n(p^l/d)}=1-\frac{1}{p}$$

which already confirms the proposed inequality. When $l=1$, however,

$$\lambda_n(p)\sum_{d|p}\frac{\mu(d)}{d\lambda_n(p/d)}=1-\frac{\lambda_n(p)}{p}=1-\frac{1}{p}+\left(\frac{-1}{p}\right)\frac{1}{p^2}\leqslant\left(1-\frac{1}{p}\right)\left(1+\frac{2}{p^2}\right),$$

and the lemma follows. □

Lemma 2.10. *Let a be odd and $(a;n)=1$. Then the function*

$$h_a(q)=\sum_{d|q}\mu\left(\frac{q}{d}\right)\frac{\lambda_n(ad)}{\lambda_n(a)}$$

is multiplicative in q. When q is odd and $(q;n)=1$, one has

$$|h_a(q)|\leqslant\frac{\mu(q)^2}{q}.$$

Proof. By Lemma 2.8, $\lambda_n(a)\neq 0$, and hence, $q\mapsto\lambda_n(aq)/\lambda_n(a)$ is defined and multiplicative, and so is h_a. It now suffices to establish the proposed inequality when $q=p^l$ is a power of an odd prime $p\nmid n$. Write $a=p^\alpha a_0$ with $p\nmid a_0$. Then

$$h_a(p^l)=\lambda_n(p^\alpha)^{-1}(\lambda_n(p^{l+\alpha})-\lambda_n(p^{l+\alpha-1})).$$

Since $p\nmid n$, we may use (2.22) to see that $h_a(p^l)=0$ unless $l+\alpha=1$. In this last case, we must have $l=1$, $\alpha=0$, and then,

$$h_a(p)=\lambda_n(p)-1=-\chi(p)/p.$$

The lemma follows. □

The function $\lambda_n(q)$ arises in connection with the distribution of sums of two squares in arithmetic progressions. In this direction, the following results suffice for our purposes.

Lemma 2.11. *For any natural numbers a and q, one has*

$$\sum_{\substack{k\leqslant X\\ k\equiv a \bmod q}}r(k)=\frac{\pi\lambda_a(q)}{q}X+O(X^{1/2+\varepsilon}q^{-1/4}(q;a)^{1/4}).$$

Let $0<\delta<1$. Then, uniformly in $X^\delta\leqslant Y\leqslant X$ and $q\leqslant X^{1-\delta}$ one has

$$\sum_{\substack{X-Y<k\leqslant X\\ k\equiv a \bmod q}}r(k)\ll\frac{\tau((q;a))}{\varphi(q)}Y\log X.$$

Proof. For the first statement we may refer to Plaksin [Pla81, Lemma 20], or Smith [Smi68] (with a different proof), at least when $q\leqslant X^{2/3-\varepsilon}$. For $q>X^{2/3-\varepsilon}$ this part of Lemma 2.11 is trivially true if we recall that $r(k)\ll\tau(k)$ and invoke the bound $\lambda_a(q)\ll\tau(q)$ from Lemma 2.8.

The second of the proposed inequalities follows from a result of Shiu [Shi80], applied to $\tau(k)$. □

https://doi.org/10.1112/S0010437X09004254 Published online by Cambridge University Press

LEMMA 2.12. *One has*

$$\sum_{q\leqslant Q}\frac{\lambda_n(q)}{q}\ll(\log Q)(\log\log n)^2. \tag{2.34}$$

Moreover, for any $L\geqslant 1$, one also has

$$\sum_{q\leqslant Q}\sum_{\substack{a_1,a_2=1\\ a_1^2+a_2^2\equiv n \bmod q\\ (a_1;q)>L}}^{q}\frac{1}{q^2}\ll L^{-1/3}(\log Q)^4. \tag{2.35}$$

Proof. For (2.34), we merely invoke Lemma 2.8 and routinely find that

$$\begin{aligned}\sum_{q\leqslant Q}\frac{\lambda_n(q)}{q}&\ll\sum_{q\leqslant Q}\frac{\tau\big((q;n)\big)}{\varphi(q)}=\sum_{\substack{r\mid n\\ r\leqslant Q}}\tau(r)\sum_{\substack{q\leqslant Q\\ (q;n)=r}}\frac{1}{\varphi(q)}\\ &=\sum_{\substack{r\mid n\\ r\leqslant Q}}\tau(r)\sum_{\substack{s\leqslant Q/r\\ (s;n/r)=1}}\frac{1}{\varphi(rs)}\leqslant\sum_{r\mid n}\frac{\tau(r)}{\varphi(r)}\sum_{s\leqslant Q}\frac{1}{\varphi(s)},\end{aligned}$$

and (2.34) follows from standard estimates. For (2.35), we have to work much harder, and the following technical estimate is essential: for natural numbers r, q and n, one has

$$\sum_{\substack{a=1\\ (a;q)=1}}^{q}(q;n-r^2a^2)^{1/2}\ll\tau(q)q(q;r^2;n)^{1/2}. \tag{2.36}$$

We postpone the proof of (2.36) and proceed directly to the derivation of (2.35). From the elementary theory of congruences we know that the number of solutions of a quadratic congruence $x^2\equiv m \bmod q$ does not exceed $\tau(q)(q;m)^{1/2}$. It follows that

$$\sum_{\substack{a_1,a_2=1\\ a_1^2+a_2^2\equiv n \bmod q\\ (a_1;q)>L}}^{q}1\leqslant\sum_{\substack{a=1\\ (a;q)>L}}^{q}\tau(q)(q;n-a^2)^{1/2}=\tau(q)\sum_{\substack{r\mid q\\ r>L}}\sum_{\substack{a=1\\ (a;q/r)=1}}^{q/r}(q;n-r^2a^2)^{1/2}.$$

For simplicity, write $q'=q/r$ and note that

$$(q;n-r^2a^2)\leqslant(r;n-r^2a^2)(q';n-r^2a^2)=(r;n)(q';n-r^2a^2).$$

By (2.36) (with q' in place of q),

$$\sum_{\substack{a_1,a_2=1\\ a_1^2+a_2^2\equiv n \bmod q\\ (a_1;q)>L}}^{q}1\ll\tau(q)^2\sum_{\substack{r\mid q\\ r>L}}(r;n)^{1/2}q'(q';r^2;n)^{1/2},$$

and a routine calculation yields

$$\sum_{q\leqslant Q}\sum_{\substack{a_1,a_2=1\\ a_1^2+a_2^2\equiv n \bmod q\\ (a_1;q)>L}}^{q}\frac{1}{q^2}\ll\sum_{q\leqslant Q}\frac{\tau(q)^2}{q}\sum_{\substack{r\mid q\\ r>L}}\frac{(r;n)^{1/2}}{r}(q/r;r^2;n)^{1/2}$$
$$\ll\sum_{r>L}\sum_{s\leqslant Q/r}\frac{\tau(rs)^2}{r^2s}(r;n)^{1/2}(s;r^2;n)^{1/2}$$
$$\ll\sum_{r>L}\frac{\tau(r)^2}{r^{3/2}}\sum_{s\leqslant Q}\frac{\tau(s)^2}{s}(s;r);$$

so far we have used only the trivial bounds $\tau(rs)\leqslant\tau(r)\tau(s)$ and $(s;r^2)^{1/2}\leqslant(s;r)$. Again by standard methods, the inner sum here is

$$\sum_{s\leqslant Q}\frac{\tau(s)^2}{s}(s;r)\leqslant\sum_{d\mid r}\sum_{s\leqslant Q/d}\frac{\tau(ds)^2}{s}\leqslant\sum_{d\mid r}\tau(d)^2\sum_{s\leqslant Q}\frac{\tau(s)^2}{s}\ll r^{\varepsilon}(\log Q)^4,$$

and (2.35) follows. The reader will notice that even finer estimates are possible. For our purposes, (2.35) is sufficient, and is independent of n.

It remains to prove (2.36). Note that the value of $(q;n-r^2a^2)$ depends only on the residue class of a modulo q. Hence, when $q=q_1q_2$ with $(q_1;q_2)=1$, we write $a=a_1q_2+a_2q_1$ and then see that

$$\sum_{\substack{a=1\\ (a;q)=1}}^{q}(q;n-r^2a^2)^{1/2}=\sum_{\substack{a_1=1\\ (a_1;q_1)=1}}^{q_1}\sum_{\substack{a_2=1\\ (a_2;q_2)=1}}^{q_2}(q_1;n-r^2q_2^2a_1^2)^{1/2}(q_2;n-r^2q_1^2a_2^2)^{1/2}.$$

The substitutions $b_1=q_2a_1$ and $b_2=q_1a_2$ now show that the left-hand side of (2.36) is a multiplicative function of q, for any fixed value of r and n.

We now proceed to prove that when $q=p^l$ is a power of a prime, then

$$\sum_{\substack{a=1\\ p\nmid a}}^{p^l}(p^l;n-r^2a^2)^{1/2}\leqslant\begin{cases}2p^l(p^l;r^2;n)^{1/2} & \text{if } p\geqslant 11,\\ 25p^l(p^l;r^2;n)^{1/2} & \text{for all } p.\end{cases}\tag{2.37}$$

By multiplicativity, this implies (2.36). We write $n=p^{\nu}n_0, r=p^{\varrho}r_0$ with $p\nmid n_0r_0$. Then, we substitute ar_0 for a to confirm the identity

$$\sum_{\substack{a=1\\ p\nmid a}}^{p^l}(p^l;n-r^2a^2)^{1/2}=\sum_{\substack{a=1\\ p\nmid a}}^{p^l}(p^l;p^{\nu}n_0-p^{2\varrho}a^2)^{1/2}.\tag{2.38}$$

Suppose that $\nu=0$, $\varrho>0$, or that $\varrho=0$, $\nu>0$. Then, for any a with $p\nmid a$, one also has $p\nmid p^{\nu}n_0-p^{2\varrho}a^2$, and hence the sum in (2.38) equals $\varphi(p^l)$. This establishes (2.37) in these cases.

Next, consider the case where $\varrho=\nu=0$. Then $n=n_0$, and the sum in (2.38) is

$$\sum_{\substack{a=1\\ p\nmid a}}^{p^l}(p^l;n-a^2)^{1/2}\leqslant\varphi(p^l)+\sum_{k=1}^{l}p^{k/2}\#\{1\leqslant a\leqslant p^l:\ a^2\equiv n \bmod p^k\}.$$

If $p \geqslant 3$, then $a^2 \equiv n \bmod p^k$ has at most two solutions. Hence, the sum in (2.38) does not exceed

$$\varphi(p^l) + 2\sum_{k=1}^{l} p^{l-k/2} \leqslant p^l\left(1 + \frac{2}{\sqrt{p}-1}\right), \tag{2.39}$$

and (2.37) again follows. When $p = 2$, then $a^2 \equiv n \bmod 2^k$ may have up to four solutions, and a consequential adjustment in the previous computation also confirms the claim in (2.37).

Finally, we have to treat the remaining case where $\nu > 0$, $\varrho > 0$. When $\nu \geqslant l$, $2\varrho \geqslant l$, then (2.37) is trivial. Hence, we write $t = \min(\nu, 2\varrho)$ and may assume that $t < l$ so that $(p^l; r^2; n) = p^t$ and $(p^l; p^\nu n_0 - p^{2\varrho} a^2) = p^t(p^{l-t}; p^{\nu-t} n_0 - p^{2\varrho-t} a^2)$. If $\nu \neq 2\varrho$, then $p \nmid p^{\nu-t} n_0 - p^{2\varrho-t} a^2$ whenever $p \nmid a$, and so,

$$(p^l; p^\nu n_0 - p^{2\varrho} a^2) = p^t$$

which shows that the sum in (2.38) is $\varphi(p^l) p^{t/2}$ in accordance with (2.37). If $\nu = 2\varrho$, we find

$$\sum_{\substack{a=1\\ p\nmid a}}^{p^l} (p^l; n - r^2 a^2)^{1/2} = p^{t/2} \sum_{\substack{a=1\\ p\nmid a}}^{p^l} (p^{l-t}; n_0 - a^2)^{1/2} = p^{(3/2)t} \sum_{\substack{a=1\\ p\nmid a}}^{p^{l-t}} (p^{l-t}; n_0 - a^2).$$

This final sum corresponds to the case where $\nu = \varrho = 0$, and, hence, is bounded by $p^{l+(1/2)t}(1 + (2/(\sqrt{p}-1)))$ when $p \geqslant 3$, and similarly when $p = 2$. This completes the proof of (2.37), and of Lemma 2.12. □

3. A binary additive problem

3.1 Initial transformations

This section is devoted to the proof of Theorem 4. We present details only for the case $j = 2$, the case $j = 1$ is similar but simpler. Let $0 < \theta < 1$. As was pointed out in the introduction, we begin by writing $R_2(n, \theta)$ in the form (1.4). Let $t(m, n)$ denote the number of $(x_1, x_2) \in \mathbb{Z}^2$ with $x_1^2 + x_2^2 = m$; $P(x_1 x_2) \leqslant n^{\theta/2}$. Then

$$R_2(n, \theta) = \sum_{m=0}^{n} r(m) t(n-m, n) = t(n, n) + 4\sum_{m=1}^{n} t(n-m, n) \sum_{d \mid m} \chi(d)$$

where for $m \in \mathbb{N}$ we have used the familiar convolution

$$r(m) = 4\sum_{d \mid m} \chi(d).$$

Note that $t(n, n) \leqslant r(n) \ll n^\varepsilon$. We follow Hooley [Hoo57] and his successors and split the inner sum into three parts, according to the size of d. Let

$$D = \sqrt{n}(\log n)^{-100}. \tag{3.1}$$

Then

$$R_2(n, \theta) = 4(M_1(n) + M_2(n) + E(n)) + O(n^\varepsilon) \tag{3.2}$$

where

$$E(n) = \sum_{m=1}^{n} t(n-m, n) \sum_{\substack{d \mid m\\ D < d \leqslant m/D}} \chi(d), \tag{3.3}$$

https://doi.org/10.1112/S0010437X09004254 Published online by Cambridge University Press

and where $M_1(n)$ and $M_2(n)$ are defined likewise, but with the condition $D < d \leqslant m/D$ replaced by $d \leqslant D$ for M_1 and $d > \max(D, m/D)$ for M_2. Reversing the order of summation in M_1 yields

$$M_1(n) = \sum_{d \leqslant D} \chi(d) \sum_{\substack{l=0 \\ l \equiv n \bmod d}}^{n-1} t(l,n). \tag{3.4}$$

We transform the sum $M_2(n)$ similarly by writing $m = dk$ and $l = n - m$. Then

$$M_2(n) = \sum_{k \leqslant D} \sum_{\substack{0 \leqslant l < n - Dk \\ l \equiv n \bmod k}} t(l,n) \chi\left(\frac{n-l}{k}\right) + O\left(\frac{n^{1+\varepsilon}}{D}\right). \tag{3.5}$$

It is convenient to extend the inner sum in (3.5) to $l \leqslant n-1$. Since $t(l,n) \leqslant r(l)$, we can use Lemma 2.11 to see that the extra terms introduce an error that does not exceed

$$\sum_{k \leqslant D} \sum_{\substack{n - Dk < l < n \\ l \equiv n \bmod k}} t(l,n) \ll \sum_{k \leqslant D} \frac{\tau((n;k)) Dk \log n}{\varphi(k)} \ll D^2 (\log n)^2.$$

By (3.1), we may now combine (3.4) and (3.5) as

$$M_1(n) + M_2(n) = M(n) + O(n/\log n)$$

where

$$M(n) = \sum_{d \leqslant D} \sum_{\substack{l=0 \\ l \equiv n \bmod d}}^{n-1} t(l,n) \left(\chi(d) + \chi\left(\frac{n-l}{d}\right)\right). \tag{3.6}$$

The scene has now been prepared to evaluate $M(n)$ asymptotically by invoking the results on smooth numbers and on sums of two squares in arithmetic progressions. This will be the subject of the next subsection. The more arithmetic analysis of the main term is performed in § 3.3, and in § 3.4, we then show that $E(n)$ is an error term. Theorem 4 will follow from the formula

$$R_2(n, \theta) = 4M(n) + 4E(n) + O(n/\log n) \tag{3.7}$$

that itself now is inferred from (3.2).

3.2 The main term: smooth numbers

Our analysis of $M(n)$ is modelled on Hooley [Hoo57] but also borrows from Plaksin [Pla81]. Note that when $x_1^2 + x_2^2 \equiv n \bmod d$, the value of $\chi((n - x_1^2 - x_2^2)/d)$ depends only on x_1, x_2 modulo $2d$. Recalling the definition of $t(l,n)$ and (3.6), this allows us to write

$$M(n) = \sum_{d \leqslant D} \sum_{\substack{a_1, a_2 = 1 \\ a_1^2 + a_2^2 \equiv n \bmod d}}^{2d} \left(\chi(d) + \chi\left(\frac{n - a_1^2 - a_2^2}{d}\right)\right) \sum_{\substack{x_1^2 + x_2^2 \leqslant n \\ x_j \equiv a_j \bmod 2d \\ P(x_1 x_2) \leqslant n^{\theta/2}}} 1.$$

Here the innermost sum extends over integers x_1, x_2, but for the argument to follow it is more convenient to sum only over natural numbers that we shall denote by k_1, k_2. For negative values of x_j we may change a_j to $-a_j$ to see that

$$M(n) = 4 \sum_{d \leqslant D} \sum_{\substack{a_1,a_2=1 \\ a_1^2+a_2^2 \equiv n \bmod d}}^{2d} \left(\chi(d) + \chi\left(\frac{n-a_1^2-a_2^2}{d}\right)\right) \sum_{\substack{k_1^2+k_2^2 \leqslant n \\ k_j \equiv a_j \bmod 2d \\ P(k_1k_2) \leqslant n^{\theta/2}}} 1. \tag{3.8}$$

In (3.8), large common factors between d and a_j are a nuisance in the later analysis, and we proceed by removing such terms. Let

$$L = (\log n)^{15}. \tag{3.9}$$

Since the inner sum over k_1, k_2 in (3.8) is at most n/d^2, we deduce from the second clause in Lemma 2.12 that the contribution of terms in (3.8) with $(a_1; d) > L$ does not exceed $\ll nL^{-1/3}(\log D)^4$. By symmetry in a_1, a_2, we conclude from (3.8) that $M(n) = 4M^*(n) + O(n/\log n)$, where $M^*(n)$ is defined as the sum in (3.8), but with the additional constraints $(a_1; d) \leqslant L$, $(a_2; d) \leqslant L$. This is still unsatisfactory for the use of a mean value estimate in the evaluation of (3.8). The reason for this is that for large values of k_1 the sum over k_2 may be rather short. To prevent this happening, we restrict the summation over k_1 in (3.8) to $k_1 \leqslant \sqrt{n - n/L}$. This will introduce an error not exceeding

$$\begin{aligned} &\sum_{d \leqslant D} \sum_{\substack{a_1,a_2=1 \\ a_1^2+a_2^2 \equiv n \bmod d}}^{2d} \sum_{\substack{\sqrt{n-n/L} < k_1 < \sqrt{n} \\ k_1 \equiv a_1 \bmod 2d}} \frac{\sqrt{n}}{d} \\ &\qquad \ll \sum_{d \leqslant D} d\tau(d) \left(\frac{\sqrt{n}}{dL} + 1\right) \frac{\sqrt{n}}{d} \ll \sqrt{n} D \log D + n(\log D)^2 L^{-1}, \end{aligned}$$

and we may finally conclude that

$$M(n) = 4M_0(n) + O(n/\log n) \tag{3.10}$$

where $M_0(n)$ is defined as the sum in (3.8), but subject to the additional constraints

$$(a_1; d) \leqslant L, \quad (a_2; d) \leqslant L, \quad k_1 \leqslant \sqrt{n - \frac{n}{L}}.$$

We are now ready to apply the variance estimate provided by Lemma 2.2 to extract a main term from $M_0(n)$. We use the notation introduced in Lemma 2.3, and for simplicity we also write

$$\Xi_{d,a}(u) = \varrho\left(\frac{\log(u/(a; 2d))}{\log n^{\theta/2}}\right)\frac{u}{2d} + \varrho'\left(\frac{\log(u/(a; 2d))}{\log n^{\theta/2}}\right)\frac{(\Pi(2d/(a; 2d)) + \gamma - 1)u}{2d \log n^{\theta/2}}. \tag{3.11}$$

Then, by Lemma 2.3, whenever $n^{(\theta+\varepsilon)/2} \leqslant u \leqslant n^{1/2}$, $d \leqslant D$, and $(a; 2d) \leqslant L$, one has

$$\Psi(u, n^{\theta/2}; 2d, a) = \Xi_{d,a}(u) - E(u, n^{\theta/2}; 2d, a) + O\left(\frac{u(\log\log n)^3}{d(\log n)^2}\right). \tag{3.12}$$

In the definition of $M_0(n)$ implicit in (3.8) and (3.10), insert (3.12) for the sum over k_2. This produces

$$M_0(n) = \sum_{d\leqslant D} \sum_{\substack{a_1,a_2=1 \\ a_1^2+a_2^2\equiv n \bmod d \\ (a_j;d)<L}}^{2d} \left(\chi(d) + \chi\left(\frac{n-a_1^2-a_2^2}{d}\right)\right) \sum_{\substack{k_1\leqslant\sqrt{n-n/L} \\ k_1\equiv a_1 \bmod 2d \\ P(k_1)\leqslant n^{\theta/2}}} \Big(\Xi_{d,a_2}(\sqrt{n-k_1^2}) - E(\sqrt{n-k_1^2}, n^{\theta/2}; 2d, a_2) + O\left(\frac{\sqrt{n}(\log\log n)^3}{d(\log n)^2}\right)\Big). \tag{3.13}$$

Using the first statement in Lemma 2.12 we may sum the error term inside the triple sum of (3.13), and the resulting total error does not exceed

$$\sum_{d\leqslant D} \sum_{\substack{a_1,a_2=1 \\ a_1^2+a_2^2\equiv n \bmod d}}^{2d} \frac{n(\log\log n)^3}{d^2(\log n)^2} \ll n\frac{(\log\log n)^6}{\log n}.$$

Next, we consider the terms involving E in (3.13). Their sum does not exceed $O(L\Sigma)$ where

$$\Sigma = \sum_{d\leqslant D} \sum_{a_2=1}^{2d} \tau(d) \max_{1\leqslant a_1\leqslant 2d} \sum_{\substack{k\leqslant\sqrt{n-n/L} \\ k\equiv a_1 \bmod 2d}} |E(\sqrt{n-k^2}, n^{\theta/2}; 2d, a_2)|. \tag{3.14}$$

The simple inequality (2.5) allows us to split the sum over k into ranges where E can be made independent of k. To do so, let $\Delta = D(\log n)^{50}$, $H = (\log n)^{50}\sqrt{1-1/L}$, and sort k into ranges $(h-1)\Delta < k \leqslant h\Delta$. By (2.5) and (3.14) one finds that

$$\begin{aligned}\Sigma &\ll \sum_{h\leqslant H} \sum_{d\leqslant D} \sum_{a=1}^{2d} \tau(d) \max_{1\leqslant a_1\leqslant 2d} \sum_{\substack{(h-1)\Delta<k\leqslant h\Delta \\ k\equiv a_1 \bmod 2d}} \left(|E(\sqrt{n-(h\Delta)^2}, n^{\theta/2}; 2d, a)| + \frac{h\Delta^2}{d\sqrt{n-(h\Delta)^2}} + 1\right) \\ &\ll \sum_{h\leqslant H} \sum_{d\leqslant D} \tau(d)\frac{\Delta}{d} \sum_{a=1}^{2d} |E(\sqrt{n-(h\Delta)^2}, n^{\theta/2}; 2d, a)| + \sum_{h\leqslant H} \sum_{d\leqslant D} \tau(d)\frac{\Delta}{d}\left(\frac{h\Delta^2}{d\sqrt{n-(h\Delta)^2}} + 1\right).\end{aligned}$$

The second term in this last estimate is $O(\Delta\sqrt{n} + H\Delta(\log D)^2) = O(n(\log n)^{-50})$. For the first term, we apply Cauchy's inequality and Lemma 2.2 to find that this contribution to (3.14) is bounded by

$$\begin{aligned}\Delta \sum_{h\leqslant H} &\left(\sum_{d\leqslant D} \frac{\tau(d)^2}{d}\right)^{1/2} \left(\sum_{d\leqslant D} \sum_{a=1}^{2d} |E(\sqrt{n-(h\Delta)^2}, n^{\theta/2}; 2d, a)|^2\right)^{1/2} \\ &\ll \Delta(\log D)^2 \sum_{h\leqslant H} (D\sqrt{n-(h\Delta)^2} + n(\log n)^{-300})^{1/2} \\ &\ll n^{3/4}D^{1/2}(\log D)^2 + n(\log n)^{-30} \ll n(\log n)^{-30}.\end{aligned}$$

By (3.9) and (3.14), this shows that the contribution from terms involving E in (3.13) does not exceed $O(n/\log n)$. The terms involving Ξ contain the main term. In fact, by partial summation and a change of variable,

$$\sum_{\substack{k\leqslant\sqrt{n-n/L}\\ k\equiv a_1 \bmod 2d\\ P(k)\leqslant n^{\theta/2}}} \Xi_{d,a_2}(\sqrt{n-k^2}) = \Xi_{d,a_2}(\sqrt{n/L}) \sum_{\substack{k\leqslant\sqrt{n-n/L}\\ k\equiv a_1 \bmod 2d\\ P(k)\leqslant n^{\theta/2}}} 1 + \int_{\sqrt{n/L}}^{\sqrt{n}} \Xi'_{d,a_2}(u) \sum_{\substack{k\leqslant\sqrt{n-u^2}\\ k\equiv a_1 \bmod 2d\\ P(k)\leqslant n^{\theta/2}}} du.$$

The first summand on the right-hand side is $O(nL^{-1/2}d^{-2})$, whence by Lemma 2.12, the contribution to (3.13), after summation over a_1, a_2, d, is $O(n/\log n)$, and is therefore negligible. For the second summand, we split the integral into two parts, over $[\sqrt{n/L}, \sqrt{n-n/L}]$ and $[\sqrt{n-n/L}, \sqrt{n}]$. The portion over $[\sqrt{n-n/L}, \sqrt{n}]$ will make, after summing over a_1, a_2, d, a contribution not exceeding $n/\log n$. To see this, note that

$$\begin{aligned}\Xi'_{d,a}(u) = \frac{1}{2d}\bigg(\varrho\bigg(\frac{\log(u/(a;2d))}{\log n^{\theta/2}}\bigg) + \frac{\Pi(2d/(2d;a))+\gamma}{\log n^{\theta/2}}\varrho'\bigg(\frac{\log(u/(a;2d))}{\log n^{\theta/2}}\bigg)\bigg)\\ + O\bigg(\frac{1+\Pi(2d/(2d;a))}{d(\log n)^2}\bigg)\end{aligned} \tag{3.15}$$

whence $\Xi'_{d,a}(u) \ll d^{-1}$, and then proceed in much the same way as in the argument leading to (3.10) to see that these terms contribute in total

$$\sum_{d\leqslant D} \sum_{\substack{a_1,a_2=1\\ a_1^2+a_2^2\equiv n \bmod d}}^{2d} \frac{\sqrt{n}}{d} \int_{\sqrt{n-n/L}}^{\sqrt{n}} |\Xi'_{d,a_2}(u)|\, du \ll \sum_{d\leqslant D} \sum_{\substack{a_1,a_2=1\\ a_1^2+a_2^2\equiv n \bmod d}}^{d} \frac{n}{d^2L} \ll n(\log n)^{-1}.$$

We may summarize the previous deliberations by inserting all estimates into (3.13). This shows that $M_0(n)$ equals

$$\sum_{d\leqslant D}\sum_{a_1,a_2}\bigg(\chi(d)+\chi\bigg(\frac{n-a_1^2-a_2^2}{d}\bigg)\bigg)\int_{\sqrt{n/L}}^{\sqrt{n-n/L}} \Xi'_{d,a_2}(u) \sum_{\substack{k\leqslant\sqrt{n-u^2}\\ k\equiv a_1 \bmod 2d\\ P(k)\leqslant n^{\theta/2}}} du + O(n(\log n)^{\varepsilon-1}) \tag{3.16}$$

where the sum over a_1, a_2 is subject to the same constraints as in (3.13). It is now possible to repeat some aspects of the previous argument. The sum over k in the above integral equals $\Psi(\sqrt{n-u^2}, n^{\theta/2}; 2d, a_1)$ and may be replaced by $\Xi_{d,a_1}(\sqrt{n-u^2})$ using (3.12). The bound $\Xi'_{d,a_2}(u) \ll 1/d$ is valid in the range $\sqrt{n/L} < u < \sqrt{n-n/L}$ by (3.15), and as before, we may use (3.12) to see that all errors introduced into (3.16) contribute a total amount not exceeding $O(n(\log n)^{\varepsilon-1})$. Therefore, by (3.16), $M_0(n)$ now equals

$$\sum_{d\leqslant D}\sum_{a_1,a_2}\bigg(\chi(d)+\chi\bigg(\frac{n-a_1^2-a_2^2}{d}\bigg)\bigg)\int_{\sqrt{n/L}}^{\sqrt{n-n/L}} \Xi'_{d,a_2}(u)\Xi_{d,a_1}(\sqrt{n-u^2})\, du + O(n(\log n)^{\varepsilon-1}) \tag{3.17}$$

where the sum over a_1, a_2 is still as in (3.13).

We now evaluate the integral in (3.17). To reduce the notational complexity, let

$$U_1 = \frac{\log(\sqrt{n-u^2}/(2d;a_1))}{\log n^{\theta/2}}, \quad U_2 = \frac{\log(u/(2d;a_2))}{\log n^{\theta/2}}.$$

https://doi.org/10.1112/S0010437X09004254 Published online by Cambridge University Press

Then, by (3.11) and (3.15),

$$\Xi'_{d,a_2}(u)\Xi_{d,a_1}(\sqrt{n-u^2}) = \frac{\sqrt{n-u^2}}{4d^2}\bigg(\varrho(U_1)\varrho(U_2) + \frac{\varrho(U_1)\varrho'(U_2)(\Pi(2d/(2d;a_2))+\gamma)}{\log n^{\theta/2}} + \frac{\varrho'(U_1)\varrho(U_2)(\Pi(2d/(2d;a_1))+\gamma-1)}{\log n^{\theta/2}} + O\bigg(\bigg(\frac{\log\log n}{\log n}\bigg)^2\bigg)\bigg)$$

uniformly for $d \leqslant D$, $\sqrt{n/L} \leqslant u \leqslant \sqrt{n-n/L}$. It is now convenient to write $u = \sqrt{n}v$ with $\sqrt{1/L} \leqslant v \leqslant \sqrt{1-1/L}$, and we note that $U_j = 1/\theta + \eta_j$ with

$$\eta_1 = \frac{\log\sqrt{1-v^2} - \log(2d;a_1)}{\log n^{\theta/2}}, \quad \eta_2 = \frac{\log v - \log(2d;a_2)}{\log n^{\theta/2}}.$$

In the ranges for v as above, and $(2d;a_j) \leqslant L$, we have $\eta_j < 0$ and $|\eta_j| \ll \log L/\log n$. When $\theta \neq 1/2$, then ϱ is twice differentiable in a neighbourhood of $1/\theta$, and hence, by Taylor's theorem,

$$\varrho(U_j) = \varrho\bigg(\frac{1}{\theta}\bigg) + \eta_j\varrho'\bigg(\frac{1}{\theta}\bigg) + O\bigg(\bigg(\frac{\log L}{\log n}\bigg)^2\bigg),$$

$$\varrho'(U_j) = \varrho'\bigg(\frac{1}{\theta}\bigg) + O\bigg(\frac{\log L}{\log n}\bigg).$$

But when $\theta = 1/2$, then we may use the formula $\varrho(u) = 1 - \log u$ $(1 < u < 2)$ and the fact that $\eta_j < 0$ to confirm these expansions also in this exceptional case. We insert this above and find that

$$\Xi'_{d,a_2}(u)\Xi_{d,a_1}(\sqrt{n-u^2}) = \frac{\sqrt{n-u^2}}{4d^2}\bigg(\varrho\bigg(\frac{1}{\theta}\bigg)^2 + \frac{\varrho(1/\theta)\varrho'(1/\theta)}{\log n^{\theta/2}}B + O\bigg(\bigg(\frac{\log\log n}{\log n}\bigg)^2\bigg)\bigg)$$

where

$$B = 2\gamma - 1 + \Pi\bigg(\frac{2d}{(2d;a_1)}\bigg) + \Pi\bigg(\frac{2d}{(2d;a_2)}\bigg) + \log(v\sqrt{1-v^2}) - \log((2d;a_1)(2d;a_2)).$$

We insert this formula into (3.17) and take $v = u/\sqrt{n}$ as the variable of integration. From the error term $O((\log\log n/\log n)^2)$ there arises a contribution to $M_0(n)$ that can be bounded with the aid of Lemma 2.12 as $O(n(\log n)^{\varepsilon-1})$, which is acceptable. For the main terms, we write

$$I_1(n) = \int_{\sqrt{1/L}}^{\sqrt{1-1/L}} \sqrt{1-v^2}\,dv, \quad I_2(n) = \int_{\sqrt{1/L}}^{\sqrt{1-1/L}} \sqrt{1-v^2}\log(v\sqrt{1-v^2})\,dv,$$

$$\mathfrak{s}_1(n) = \sum_{d\leqslant D}\frac{1}{d^2}\sum_{a_1,a_2}\bigg(\chi(d) + \chi\bigg(\frac{n-a_1^2-a_2^2}{d}\bigg)\bigg),$$

$$\mathfrak{s}_2(n) = \sum_{d\leqslant D}\frac{1}{d^2}\sum_{a_1,a_2}\bigg(\chi(d) + \chi\bigg(\frac{n-a_1^2-a_2^2}{d}\bigg)\bigg)\bigg(\Pi\bigg(\frac{2d}{(2d;a_1)}\bigg) - \log(2d;a_1)\bigg)$$

where the sums over a_1, a_2 are still subject to the conditions in (3.13). Observing symmetry in a_1, a_2, we then find that

$$4M_0(n) = n\mathfrak{s}_1(n)I_1(n)\bigg(\varrho\bigg(\frac{1}{\theta}\bigg)^2 + O\bigg(\frac{1}{\log n}\bigg)\bigg) + \varrho\bigg(\frac{1}{\theta}\bigg)\varrho'\bigg(\frac{1}{\theta}\bigg)\bigg(\frac{2n}{\log n^{\theta/2}}\mathfrak{s}_2(n)I_1(n) + \frac{n}{\log n^{\theta/2}}\mathfrak{s}_1(n)I_2(n)\bigg) + O(n(\log n)^{\varepsilon-1}).$$

To see that $I_1(n)$ is essentially independent of n, it suffices to integrate over $[0,1]$ instead to confirm the formula $I_1(n) = \pi/4 + O(L^{-1/2})$. Moreover, we obviously have $I_2(n) \ll \log L$. The previous formula for $M_0(n)$ now simplifies to

$$4M_0(n) = \frac{\pi}{4} n\mathfrak{s}_1(n)\biggl(\varrho\Bigl(\frac{1}{\theta}\Bigr)^2 + O\Bigl(\frac{\log\log n}{\log n}\Bigr)\biggr) + \frac{2\varrho(1/\theta)\varrho'(1/\theta)n}{\log n^{\theta/2}}\mathfrak{s}_2(n)I_1(n) + O(n(\log n)^{\varepsilon-1}). \tag{3.18}$$

3.3 Two singular series

It remains to evaluate the sums $\mathfrak{s}_1$ and $\mathfrak{s}_2$. The treatment of $\mathfrak{s}_1(n)$ is rather straightforward, for $\mathfrak{s}_2(n)$ some extra complications will arise. The summation conditions for a_1, a_2 contain the artificial conditions $(a_j; d) \leqslant L$ that we remove to recover the genuinely multiplicative nature of these sums. If we add in terms with $(a_1; d) > L$ or $(a_2; d) > L$, then by Lemma 2.12, the sum $\mathfrak{s}_1(n)$ will be altered by an amount not exceeding $L^{-1/3}(\log n)^4 \ll (\log n)^{-1}$, and since $\Pi(2d/(2d; a_1)) - \log(2d; a_1) \ll \log d$, the same argument shows that $\mathfrak{s}_2(n)$ is altered by $O(1)$. We write

$$\Gamma(d) = \sum_{\substack{a_1,a_2=1\\ a_1^2+a_2^2\equiv n \bmod d}}^{2d} \biggl(\chi(d) + \chi\Bigl(\frac{n-a_1^2-a_2^2}{d}\Bigr)\biggr). \tag{3.19}$$

Then we have shown

$$\mathfrak{s}_1(n) = \sum_{d\leqslant D} d^{-2}\Gamma(d) + O((\log n)^{-1}) \tag{3.20}$$

and

$$\mathfrak{s}_2(n) = \sum_{d\leqslant D}\frac{1}{d^2}\sum_{\substack{a_1,a_2=1\\ a_1^2+a_2^2\equiv n \bmod d}}^{2d}\biggl(\chi(d)+\chi\Bigl(\frac{n-a_1^2-a_2^2}{d}\Bigr)\biggr)\biggl(\Pi\Bigl(\frac{2d}{(2d;a_1)}\Bigr) - \log(2d;a_1)\biggr) + O(1). \tag{3.21}$$

We shall now complete the sums over d by Dirichlet series techniques. The function $\Gamma(d)$ factors with respect to the decomposition into prime factors. To see this, let $d = 2^\delta d_0$ with odd d_0, and apply the Chinese remainder theorem in (3.19) to confirm the formula

$$\Gamma(d) = \chi(d_0)d_0\lambda_n(d_0)\Gamma(2^\delta) \tag{3.22}$$

where it is convenient to note that $\chi(d_0)d_0\lambda_n(d_0)$ is multiplicative. Moreover, if $4\nmid n$, one has

$$\Gamma(2^\delta) = 0 \quad (\delta \geqslant 1). \tag{3.23}$$

For a proof, consider (3.19) with $d = 2^\delta$ and $\delta \geqslant 2$. Suppose that a_1, a_2 is a pair that meets the summation conditions. Then, in particular, $a_1^2 + a_2^2 \equiv n \bmod 2^\delta$, and since $4 \nmid n$, at least one of a_1, a_2 must be odd, say a_1. Then $a_1 + 2^\delta, a_2$ is another pair that meets the summation condition, and one has

$$\chi\Bigl(\frac{n-a_1^2-a_2^2}{2^\delta}\Bigr) = -\chi\Bigl(\frac{n-(a_1+2^\delta)^2-a_2^2}{2^\delta}\Bigr), \quad \chi(2^\delta) = 0.$$

This confirms (3.23) for $\delta \geqslant 2$, and for $\delta = 1$ it may be checked by hand.

We may combine (3.22) and (3.23) to the simpler formula $\Gamma(d) = \Gamma(1)\chi(d)d\lambda_n(d)$ for all d. Hence, (3.20) now reads

$$\mathfrak{s}_1(n) = \Gamma(1)S(D) + O((\log n)^{-1}), \tag{3.24}$$

where

$$S(D)=\sum_{d\leqslant D} d^{-1}\chi(d)\lambda_n(d), \tag{3.25}$$

and we evaluate this sum through the Dirichlet series

$$F_n(s)=\sum_{d=1}^{\infty}\frac{\chi(d)\lambda_n(d)}{d^s}.$$

By Lemma 2.8 we see that $F_n(s)$ converges absolutely for $\operatorname{Re} s>1$, and may therefore be written as an Euler product. Using (2.22) in conjunction with (2.20) and (2.21), one may calculate the Euler factors explicitly. This yields

$$F_n(s)=(1-2^{-1-s})^{-1}L(s)\zeta(s+1)^{-1}\prod_{\substack{p\mid n\\ p\nmid 2}}E_{n,p}(s) \tag{3.26}$$

where $\zeta(s)$ is the Riemann zeta function, $L(s)$ is the Dirichlet L-function with character χ, and $E_{n,p}(s)$ is defined by $p^\nu\|n$ and

$$E_{n,p}(s)=\frac{1-p^{-(\nu+1)s}}{1-p^{-s}}. \tag{3.27}$$

Hence, (3.26) provides the analytic continuation of $F_n(s)$ to $\operatorname{Re} s>0$.

We now compare the finite sum in (3.24) with its limit by standard analytic techniques, but because we need to take care of the dependence on n, moderate details will be presented. The Dirichlet series with coefficients $\chi(d)\lambda_n(d)d^{-1}$ is $F_n(s+1)$. An effective version of Perron's formula, such as Theorem II.2.2 in Tenenbaum [Ten95], shows that

$$S(D)=\frac{1}{2\pi i}\int_{1/4-in}^{1/4+in}F_n(s+1)\frac{D^s}{s}\,ds+O\biggl(D^{1/4}\sum_{d=1}^{\infty}\frac{\lambda_n(d)}{d^{5/4}(1+n|\log d/D|)}\biggr).$$

Here the error term is readily seen to be bounded by $O(n^{-1/4})$. The integrand is analytic in $\operatorname{Re} s>-1$ except for a simple pole of residue $F_n(1)$ at $s=0$. From (3.27) we find that $E_{n,p}(s)\ll 1+p^{-1/2}$ uniformly in $\operatorname{Re} s\geqslant 1/2$. Since $\prod_{p\mid n}(1+p^{-1/2})\ll n^{\varepsilon}$, we deduce from (3.26) and (3.25) that in $\operatorname{Re} s\geqslant 1/2$ one has

$$F_n(s)\ll n^{\varepsilon}|L(s)|.$$

Hence, we may integrate $F_n(s+1)D^s s^{-1}$ over the rectangle with corners $1/4\pm in$ and $-1/2\pm in$, and a standard estimate like $L(s)\ll 1+|\operatorname{Im} s|^{1/6}$ suffices to estimate the integrals along the horizontal lines and the segment on $\operatorname{Re} s=-1/2$ by $n^{1/6+\varepsilon}D^{-1/2}$. The residue theorem now yields

$$S(D)=F_n(1)+O(n^{\varepsilon-1/12}). \tag{3.28}$$

It is straightforward to recover the formulae (1.1) from $F_n(1)$. Indeed, the classical evaluations $\zeta(2)=\pi^2/6$ and $L(1)=\pi/4$ combined with (3.25), (3.26) and (3.27) yield

$$F_n(1)=\frac{2}{\pi}\prod_{\substack{p^\nu\| n\\ p\neq 2}}\frac{1-p^{-\nu-1}}{1-p^{-1}},$$

and (3.19) gives $\Gamma(1)=4+\chi(n)+2\chi(n-1)+\chi(n-2)$. Thus $\Gamma(1)=4$ when n is odd, and

$\Gamma(1)=6$ when $n\equiv 2 \bmod 4$. Using multiplicativity we now find that whenever $4\nmid n$ one has

$$\pi n\Gamma(1)F_n(1)=8(2+(-1)^n)\sum_{\substack{d|n\\ d\equiv 1 \bmod 2}} d,$$

which is the term in (1.1). In particular, it follows that $S(D)\ll F_n(1)\ll \log\log n$, and that

$$\mathfrak{s}_1(n)=\Gamma(1)F_n(1)+O((\log n)^{-1}).$$

If we now anticipate the bound $\mathfrak{s}_2(n)\ll(\log n)^{\varepsilon}$, we conclude from (3.18) and (3.28) that

$$16M_0(n)=\varrho\left(\frac{1}{\theta}\right)^2\mathfrak{S}(n)n+O(n(\log n)^{\varepsilon-1}). \tag{3.29}$$

It remains to consider $\mathfrak{s}_2(n)$. We mimic the above treatment of $\mathfrak{s}_1(n)$ via a conventional contour integral approach. The presence of terms $\log(2d;a_1)$ and $\Pi(2d/(2d;a_1))$ in (3.21) is a cause for extra complication, but since we only need an upper bound we may economize in part of the argument. We study the sums

$$\Gamma_v(d)=\sum_{\substack{a_1,a_2=1\\ a_1^2+a_2^2\equiv n \bmod d\\ (a_1;2d)=v}}^{2d}\left(\chi(d)+\chi\left(\frac{n-a_1^2-a_2^2}{d}\right)\right)$$

that allow us to rewrite (3.21) in the form

$$\mathfrak{s}_2(n)=\sum_{d\leqslant D}\frac{1}{d^2}\sum_{v|2d}\Gamma_v(d)(\Pi(2d/v)-\log v)+O(1). \tag{3.30}$$

Suppose that $d=2^{\delta}d_0$ with $2\nmid d_0$. When $v|2d$ one has $v=2^{\eta}v_0$ with $v_0|d_0$ and $0\leqslant\eta\leqslant\delta+1$. Any residue class $a_j \bmod 2d$ has exactly one representative $2^{\delta+1}b_j+d_0c_j$ with $1\leqslant b_j\leqslant d_0, 1\leqslant c_j\leqslant 2^{\delta+1}$, and one readily confirms the identity

$$\chi(d_0)\Gamma_v(d)=\sum_{b_1,b_2,c_1,c_2}\left(\chi(2^{\delta})+\chi\left(\frac{n-(d_0c_1)^2-(d_0c_2)^2}{2^{\delta}}\right)\right)$$

where the sum is over all $1\leqslant b_j\leqslant d_0$, $1\leqslant c_j\leqslant 2^{\delta+1}$ subject to the additional constraints $(b_1;d_0)=v_0$, $(c_1;2^{\delta+1})=2^{\eta}$ and

$$(2^{\delta+1}b_1)^2+(2^{\delta+1}b_2)^2\equiv n \bmod d_0,\quad (d_0c_1)^2+(d_0c_2)^2\equiv n \bmod 2^{\delta}.$$

Substitute b_j for $2^{\delta+1}b_j$ and c_j for d_0c_j to see that

$$\Gamma_v(d)=\chi(d_0)\Gamma_{2^{\eta}}(2^{\delta})\Lambda_{v_0}(d_0) \tag{3.31}$$

where $\Lambda_v(d)$ denotes the number of solutions of the congruence $a_1^2+a_2^2\equiv n \bmod d$ with $1\leqslant a_1,a_2\leqslant d$ and $(a_1;d)=v$. Note that $\Lambda_v(d)=\Lambda_v(d;n)$ in the notation of § 2.3; we suppress dependence of n here as long as there is no risk of confusion.

The argument that proved (3.23) is still applicable in the new context, and now shows that

$$\Gamma_{2^{\eta}}(2^{\delta})=0\quad(\delta\geqslant 2, 0\leqslant\eta\leqslant\delta+1), \tag{3.32}$$

but in some cases now $\Gamma_{2^{\eta}}(2)\neq 0$. We sort the terms in (3.30) according to the values of δ and η where $2^{\delta}\|d$, $2^{\eta}\|v$. Then, by (3.31) and (3.32), and recalling the definition of Π, the sum on the right-hand side of (3.30) becomes

$$\sum_{\delta=0}^{1}\sum_{\eta=0}^{\delta+1} 4^{-\delta}\Gamma_{2^\eta}(2^\delta) \sum_{d\leqslant 2^{-\delta}D} \frac{\chi(d)}{d^2} \sum_{v\mid d} \Lambda_v(d)\left(\Pi\left(2^{\delta+1-\eta}\frac{d}{v}\right) - \log 2^\eta v\right)$$
$$= \sum_{\delta=0}^{1}\sum_{\eta=0}^{\delta+1} 4^{-\delta}\Gamma_{2^\eta}(2^\delta)(T(2^{-\delta}D) - c(\delta,\eta)S(2^{-\delta}D))$$

where $S(D)$ is the sum defined by (3.24), where

$$T(D) = \sum_{d\leqslant D} \frac{\chi(d)}{d^2} \sum_{v\mid d} \Lambda_v(d)\left(\Pi\left(\frac{d}{v}\right) - \log v\right), \tag{3.33}$$

and where $c(\delta,\eta) = (\eta-1)\log 2$ if $\eta \leqslant \delta$, and $c(\delta,\delta+1) = (\delta+1)\log 2$. The term involving T is independent of η, and since $\Gamma_1(2)+\Gamma_2(2)+\Gamma_4(2) = \Gamma(2) = 0$ by (3.23), we may combine our results with (3.30) to the simpler formula

$$\mathfrak{s}_2(n) = \Gamma(1)T(D) + O(\log\log n). \tag{3.34}$$

We estimate $T(D)$ by a rough version of Dirichlet's hyperbola method and the analytic technique that was used to evaluate $\mathfrak{s}_1(n)$. Let

$$V = (\log n)^{21}, \tag{3.35}$$

and consider the contribution to $T(D)$ of the subsum defined by $v > V$ in (3.33). Because one has $\prod(d/v) - \log v \ll \log D$ for $v\mid d, v\leqslant D$, we may apply Lemma 2.12 to see that terms with $v > V$ amount to

$$\ll (\log D)\sum_{d\leqslant D}\frac{1}{d^2}\sum_{\substack{v\mid d\\ v>V}} \Lambda_v(d) \ll (\log n)^{-2}$$

in (3.33), and hence,

$$T(D) = T_1(D) - T_2(D) + O((\log n)^{-2}) \tag{3.36}$$

where

$$T_1(D) = \sum_{v\leqslant V}\frac{\chi(v)}{v^2}\sum_{u\leqslant D/v}\Pi(u)\frac{\chi(u)}{u^2}\Lambda_v(uv), \tag{3.37}$$

$$T_2(D) = \sum_{v\leqslant V}\frac{\chi(v)}{v^2}(\log v)\sum_{u\leqslant D/v}\frac{\chi(u)}{u^2}\Lambda_v(uv). \tag{3.38}$$

The sum (3.38) is ready for direct treatment by analytic methods. For odd $v\in\mathbb{N}$, we study the Dirichlet series

$$G(s; v, n) = G(s) = \sum_{u=1}^{\infty}\chi(u)\Lambda_v(uv)u^{-s}. \tag{3.39}$$

The inequalities

$$\Lambda_v(uv) \leqslant uv\lambda_n(uv) \ll uv\tau(u)\tau(v) \tag{3.40}$$

follow from (2.19), (2.25) and Lemma 2.8, and imply that $G(s)$ converges absolutely for $\mathrm{Re}\,(s) > 2$. The quasi-multiplicative property of Λ expressed by (2.27) shows that in the same

https://doi.org/10.1112/S0010437X09004254 Published online by Cambridge University Press

range Re $(s) > 2$ we may write $G(s)$ as a product of Eulerian type. This takes the shape

$$G(s) = \prod_{p^h \| v} \sum_{l=0}^{\infty} \Lambda_{p^h}(p^{h+l})\chi(p^l)p^{-ls} = \prod_p G_p(s), \text{ say.} \tag{3.41}$$

As was the case with $F(s)$ in the discussion of $\mathfrak{s}_1(n)$, it is again possible to compute $G_p(s)$ explicitly, and as we shall see, $G(s)$ contains $L(s-1)$ as a factor. Indeed, for odd primes $p \nmid v$, we may use (2.28) to confirm the formula

$$G_p(s) = 1 + \sum_{l=1}^{\infty} p^{l-1}\Lambda_1(p)\chi(p^l)p^{-ls} = \frac{1 + \chi(p)p^{1-s}((\Lambda_1(p)/p) - 1)}{1 - \chi(p)p^{1-s}}. \tag{3.42}$$

When $p^h \| v$ with $h \geqslant 1$, it is convenient to write $n = p^\nu n_p$ with $p \nmid n_p$. There are three cases. First suppose that $h > \nu$. If ν is odd, then $G_p(s) = 0$ by (2.31), and if ν is even, then again by (2.31),

$$G_p(s) = p^{\nu/2}\left(1 + \left(\frac{n_p}{p}\right)\right)\sum_{l=0}^{\infty} \varphi(p^l)\chi(p^l)p^{-ls} = p^{\nu/2}\left(1 + \left(\frac{n_p}{p}\right)\right)\frac{1 - \chi(p)p^{-s}}{1 - \chi(p)p^{1-s}}. \tag{3.43}$$

Next, suppose that $1 \leqslant h \leqslant \nu/2$. Then, by (2.30),

$$G_p(s) = \sum_{l=0}^{h} \varphi(p^l)p^{[(h+l)/2]}\frac{\chi(p^l)}{p^{ls}} + \sum_{l=h+1}^{\infty} p^{h+l-1}\Lambda_1(p; np^{-2h})\frac{\chi(p^l)}{p^{ls}}.$$

We may rearrange this in the form

$$\begin{aligned}(1 - \chi(p)p^{1-s})G_p(s) = p^{[h/2]} &+ \sum_{l=1}^{h} \frac{\chi(p^l)}{p^{ls}}(\varphi(p^l)p^{[(h+l)/2]} - p\varphi(p^{l-1})p^{[(h+l-1)/2]}) \\ &+ \frac{\chi(p^{h+1})}{p^{(h+1)s}}p^{2h}(\Lambda_1(p; np^{-2h}) - p + 1).\end{aligned} \tag{3.44}$$

Finally, suppose that $\nu/2 < h \leqslant \nu$. Here, when ν is odd, (2.31) gives

$$G_p(s) = \sum_{l=0}^{\nu-h} \varphi(p^l)p^{[(h+l)/2]}\frac{\chi(p^l)}{p^{ls}};$$

this we rewrite as

$$\begin{aligned}&(1 - \chi(p)p^{1-s})G_p(s) \\ &\quad = p^{[h/2]} + \sum_{l=1}^{\nu-h} \frac{\chi(p^l)}{p^{ls}}(\varphi(p^l)p^{[(h+l)/2]} - p\varphi(p^{l-1})p^{[(h+l-1)/2]}) - \frac{\chi(p^{\nu-h+1})}{p^{(\nu-h+1)s}}p\varphi(p^{\nu-h})p^{[\nu/2]}.\end{aligned} \tag{3.45}$$

When $\nu/2 < h \leqslant \nu$ and ν is even, then by (2.31) we find likewise

$$G_p(s) = \sum_{l=0}^{\nu-h} \varphi(p^l)p^{[(h+l)/2]}\frac{\chi(p^l)}{p^{ls}} + \sum_{l=\nu-h+1}^{\infty} \varphi(p^l)p^{\nu/2}\frac{\chi(p^l)}{p^{ls}}\left(1 + \left(\frac{n_p}{p}\right)\right)$$

https://doi.org/10.1112/S0010437X09004254 Published online by Cambridge University Press

and much as in (3.45), this takes the more convenient form

$$(1-\chi(p)p^{1-s})G_p(s) = p^{[h/2]} + \sum_{l=1}^{\nu-h} \frac{\chi(p^l)}{p^{ls}}(\varphi(p^l)p^{[(h+l)/2]} - p\varphi(p^{l-1})p^{[(h+l-1)/2]}) + \frac{\chi(p^{\nu-h+1})}{p^{(\nu-h+1)s}} p^{\nu/2}\bigg(\varphi(p^{\nu-h+1})\bigg(1+\bigg(\frac{n_p}{p}\bigg)\bigg) - p\varphi(p^{\nu-h})\bigg).$$

In the particular case $h=\nu$, this simplifies to

$$(1-\chi(p)p^{1-s})G_p(s) = p^{[\nu/2]} + \frac{\chi(p)}{p^s}p^{\nu/2}\bigg(\bigg(\frac{n_p}{p}\bigg)(p-1)-1\bigg), \tag{3.46}$$

and when $h<\nu$ one gets

$$p^{[h/2]} + \sum_{l=1}^{\nu-h} \frac{\chi(p^l)}{p^{ls}}(\varphi(p^l)p^{[(h+l)/2]} - p\varphi(p^{l-1})p^{[(h+l-1)/2]}) + \bigg(\frac{n_p}{p}\bigg)\frac{\chi(p^{\nu-h+1})}{p^{(\nu-h+1)s}}p^{\nu/2}\varphi(p^{\nu-h+1}). \tag{3.47}$$

By (3.41)–(3.47), we obtain

$$G(s) = L(s-1)J_{n,v}(s)K(s;v,n) \tag{3.48}$$

where once again L is the Dirichlet L-function for χ, and where

$$J_{n,v}(s) := \prod_{p\nmid v}\bigg(1+\chi(p)p^{1-s}\bigg(\frac{\Lambda_1(p;n)}{p}-1\bigg)\bigg),$$

$$K(s;v,n) := \prod_{p\mid v}(1-\chi(p)p^{1-s})G_p(s).$$

By (3.44)–(3.48) it follows that $G(s)$ is regular in $\mathrm{Re}\,(s)>1$, in particular.

Further progress now depends on an upper bound for $|G(s)|$ in $\mathrm{Re}\,(s)\geqslant 3/2$ that we now derive. By (2.29), we have $|p^{-1}\Lambda_1(p,n)-1|\leqslant 3/p$ when $p\nmid n$, and $|p^{-1}\Lambda_1(p,p)-1|\leqslant 1+2/p$. If $\sigma=\mathrm{Re}\,s$ it follows that

$$|J_{n,v}(s)| \leqslant \prod_p\bigg(1+\frac{3}{p^\sigma}\bigg)\prod_{p\mid n}(1+p^{1-\sigma}), \tag{3.49}$$

and in particular, in $\mathrm{Re}\,(s)\geqslant 3/2$, this yields the convenient bound $J_{n,v}(s)\ll\tau(n)$. Similarly, an inspection of (3.43)–(3.47) shows

$$|K(s;v,n)| \ll \tau(v)(v;n)^{1/2} \quad (\mathrm{Re}\,(s)\geqslant 3/2). \tag{3.50}$$

Let $D/V^3\leqslant U\leqslant D$. Then, by (3.41) and Perron's formula [Ten95, Theorem II.2.2],

$$\sum_{u\leqslant U}\frac{\chi(u)}{u^2}\Lambda_v(uv) = \frac{1}{2\pi i}\int_{1/4-in}^{1/4+in} G(s+2;v,n)\frac{U^s}{s}\,ds + O\bigg(U^{1/4}\sum_{u=1}^{\infty}\frac{\Lambda_v(uv)}{u^{9/4}(1+n|\log u/U|)}\bigg).$$

For $v\leqslant V$, one finds from (3.40) and crude estimates that the error term here is $O(n^{-1/4})$. Now integrate $G(s+2;v,n)U^s s^{-1}$ along the rectangle with corners $1/4\pm in$ and $-1/2\pm in$. By (3.48), (3.49) and (3.50), we have

$$|G(s+2)| \ll |L(s+1)|(vn)^\varepsilon(v;n) \ll V^{1+\varepsilon}n^\varepsilon|L(s+1)| \ll n^{2\varepsilon}|L(s+1)|, \tag{3.51}$$

and hence we may argue exactly as in the corresponding treatment of $F_n(s)$ in the evaluation of $S(D)$ to show that the integral over the horizontal parts as well as over the vertical portion on

https://doi.org/10.1112/S0010437X09004254 Published online by Cambridge University Press

Re $(s)=-1/2$ contribute $O(n^{\varepsilon-1/12})$. Thus, by the residue theorem,

$$\sum_{u\leqslant U}\frac{\chi(u)}{u^2}\Lambda_v(uv)=G(2;v,n)+O(n^{\varepsilon-1/12}),$$

uniformly in $v\leqslant V$. By (3.49) and (3.50), $G(2;v,n)\ll(\log\log n)\tau(v)v^{1/2}$, and so, by the previous formula with $U=D/v$, and (3.38),

$$T_2(D)\ll\log\log n. \tag{3.52}$$

The treatment of the sum $T_1(D)$ is similar, but some preparation is required before we can bring in our earlier analysis of the Dirichlet series $G(s)$. Recall the definition of $\Pi(u)$ in (2.6) and then reverse the order of summation in (3.37) to see that

$$T_1(D)=\sum_{p\leqslant D}\frac{\chi(p)\log p}{p^2(p-1)}\sum_{v\leqslant V}\sum_{u\leqslant D/pv}\frac{\chi(v)\chi(u)}{v^2u^2}\Lambda_v(uvp). \tag{3.53}$$

We estimate the contribution of primes $p>V$ to (3.53) with the aid of (3.40) as

$$\ll\sum_{p>V}\frac{\log p}{p^2}\sum_{v\leqslant V}\frac{\tau(v)}{v}\sum_{u\leqslant D}\frac{\tau(u)}{u}\ll V^{-1}(\log V\log D)^2\ll 1.$$

For the remaining primes $p\leqslant V$ in (3.53), we write $v=p^hv_0,u=p^lu_0$ with $p\nmid u_0v_0$. Then, by (2.26),

$$T_1(D)=\sum_{h,l=0}^{\infty}\sum_{p\leqslant V}\frac{\chi(p^{1+h+l})\Lambda_{p^h}(p^{h+l+1})}{p^{2(h+l+1)}(p-1)}\log p\sum_{\substack{v\leqslant Vp^{-h}\\ p\nmid v}}\frac{\chi(v)}{v^2}\sum_{\substack{u\leqslant Dp^{-1-l}v^{-1}\\ p\nmid u}}\frac{\chi(u)}{u^2}\Lambda_v(uv)+O(1). \tag{3.54}$$

Here we bound terms with $p^l>V$ by invoking (3.40) again. Since only terms with $p^l\leqslant D$ make a non-zero contribution, these terms amount to

$$\ll\sum_{h,l\ll\log n}\sum_{\substack{p\leqslant V\\ V<p^l\leqslant D}}\frac{(\log p)(h+1)(l+1)}{p^{h+l+1}(p-1)}\sum_{v\leqslant V}\frac{\tau(v)}{v}\sum_{u\leqslant D}\frac{\tau(u)}{u}\ll V^{-1}(\log n)^4\ll 1.$$

Thus, (3.54) remains valid with the additional contraint $p^l\leqslant V$ on the sum over p. In this situation, the innermost sum over u in (3.54) is of the type

$$\sum_{\substack{u\leqslant U\\ p\nmid u}}\frac{\chi(u)}{u^2}\Lambda_v(uv) \tag{3.55}$$

where $p\nmid v$, and where $U=Dp^{-1-l}v^{-1}\in[DV^{-3},D]$. By (3.41),

$$\sum_{\substack{u=1\\ p\nmid u}}^{\infty}\chi(u)\Lambda_v(uv)u^{-s}=G(s)/G_p(s),$$

and the 'missing' Euler factor $G_p(s)$ is given by (3.42) (because $p\nmid v$). Hence, an inspection of the analysis of $G(s)$ shows that for $G(s)/G_p(s)$ the bound (3.51) still holds, and that we may therefore adapt the argument that followed (3.50) to conclude that the sum (3.55) equals $G(2)/G_p(2)+O(n^{\varepsilon-1/12})$, and is therefore bounded by $O(v^{1/2+\varepsilon}\log\log n)$, uniformly in $v\leqslant V, p\leqslant V, p^l\leqslant V$. Consequently, by (3.54),

$$T_1(D)\ll 1+(\log\log n)\sum_{h,l\geqslant 0}\sum_{\substack{p\leqslant V\\ p\leqslant V^{1/l}}}\frac{\log p}{p^{1+2(h+l+1)}}\Lambda_{p^h}(p^{h+l+1})\ll\log\log n;$$

here (3.40) is again sufficient for the final inequality. When combined with (3.52), (3.36) and (3.34), it follows that $\mathfrak{s}_2(n)\ll\log\log n$, and hence, the proof of (3.29) is complete.

3.4 The error term

Recall that it remains to estimate $E(n)$, defined by (3.3) above. We initiate our treatment by using the inequality $t(l,n)\leqslant r(l)$ in (3.3) to infer the bound

$$|E(n)|\leqslant\sum_{m<n}r(n-m)|H(m)| \tag{3.56}$$

where

$$H(m)=\sum_{\substack{d|m\\ D<d\leqslant m/D}}\chi(d).$$

We explore the fact that numbers m with $H(m)\neq 0$ are rare through the estimate provided by Lemma 2.7. Our path will follow the pattern laid down by Hooley [Hoo57], and with this in mind, we introduce the sets

$$\mathcal{A}_n=\{a\in\mathbb{N}:p|a\Rightarrow p\nmid n\text{ and }p\leqslant Y(n)\}$$
$$\mathcal{B}_n=\{b\leqslant n:p|b\Rightarrow p>Y(n)\text{ or }p|n\}$$

where $Y(n)=\exp(\log n/\log\log n)$, in accordance with the setup in Lemmata 2.4 and 2.6.

Note that $\mathcal{A}_n\cap\mathcal{B}_n=\{1\}$ and that, by unique factorisation, any $m\leqslant n$ has exactly one representation $m=ab$ with $a\in\mathcal{A}_n, b\in\mathcal{B}_n$. We use this in (3.56), and separate terms with $a>n^{1/50}$ from the resulting double sum over a and b. By the definition of $\mathcal{A}_n$, any $a\in\mathcal{A}_n$ with $a>n^{1/50}$ has a divisor $a_1\in\mathcal{A}_n$ with $n^{1/50}<a_1<n^{1/25}$. Therefore, these terms contribute to (3.56) at most

$$\sum_{\substack{a\in\mathcal{A}_n,b\in\mathcal{B}_n\\ ab<n\\ a>n^{1/50}}}r(n-ab)|H(ab)|\leqslant\sum_{m<n}r(n-m)|H(m)|\sum_{\substack{a_1\in\mathcal{A}_n\\ a_1|m\\ n^{1/50}<a_1\leqslant n^{1/25}}}1.$$

Write $n-m=l$ and replace a_1 by a again. Then reverse the order of summation to find that the right-hand side equals

$$\sum_{\substack{n^{1/50}<a\leqslant n^{1/25}\\ a\in\mathcal{A}_n}}\sum_{\substack{l<n\\ l\equiv n\bmod a}}r(l)|H(n-l)|.$$

Now $r(l)\leqslant 4\tau(l)$ and $|H(m)|\leqslant\tau(m)$ by trivial estimates. From Cauchy's inequality we infer that the previous expression does not exceed

$$4\bigg(\sum_{\substack{n^{1/50}<a\leqslant n^{1/25}\\ a\in\mathcal{A}_n}}\sum_{\substack{l<n\\ l\equiv n\bmod a}}\tau(l)^2\bigg)^{1/2}\bigg(\sum_{\substack{n^{1/50}<a\leqslant n^{1/25}\\ a\in\mathcal{A}_n}}\sum_{\substack{m<n\\ m\equiv 0\bmod a}}\tau(m)^2\bigg)^{1/2};$$

here we replaced $n-l$ by m again. In both factors, Shiu's version of the Brun–Titchmarsh theorem [Shi80] bounds the inner sum by $\tau(a)\varphi(a)^{-1}n(\log n)^3$, and we may use Lemma 2.6 to sum over a. It follows that

$$\sum_{\substack{a\in\mathcal{A}_n,b\in\mathcal{B}_n\\ ab<n\\ a>n^{1/50}}} r(n-ab)|H(ab)| \ll n(\log n)^{-2}, \tag{3.57}$$

say. For the complementary part with $a\leqslant n^{1/50}$, we note that whenever $a\in\mathcal{A}_n$, $b\in\mathcal{B}_n$, one has $(a;b)=1$, and hence

$$H(ab)=\sum_{\substack{d|ab\\ D<d<ab/D}}\chi(d)=\sum_{\substack{a_1\in\mathcal{A}_n,b_1\in\mathcal{B}_n\\ a_1|a,b_1|b\\ D<a_1b_1<ab/D}}\chi(a_1)\chi(b_1).$$

It follows that

$$\sum_{\substack{a\in\mathcal{A}_n,b\in\mathcal{B}_n\\ ab<n\\ a\leqslant n^{1/50}}} r(n-ab)|H(ab)| \leqslant \sum_{m<n} r(n-m)\Bigg|\sum_{\substack{a_1\in\mathcal{A}_n,b_1\in\mathcal{B}_n\\ a_1b_1|m,a_1\leqslant n^{1/50}\\ D<a_1b_1<m/D}}\chi(a_1)\chi(b_1)\Bigg|$$

because $a_1\leqslant n^{1/50}$ holds for any $a_1|a$, and we have added non-negative terms on the right in cases where m is not of the form $m=ab$ with $a\leqslant n^{1/50}$. We sum over b_1 first and use the triangle inequality. Then, it follows that the above expression does not exceed

$$\sum_{\substack{b_1\in\mathcal{B}_n\\ b_1\leqslant n/D}}\sum_{\substack{m\leqslant n\\ m\equiv 0 \bmod b_1}} r(n-m)\Bigg|\sum_{\substack{a_1\in\mathcal{A}_n,a_1\leqslant n^{1/50}\\ a_1|(m/b_1)\\ D<a_1b_1<m/D}}\chi(a_1)\Bigg|.$$

In the interest of notational simplicity, we write $a_1=a$, $b_1=b$ and $m=bs$. Then the previous expression takes the form

$$\sum_{\substack{b\in\mathcal{B}_n\\ b\leqslant n/D}}\sum_{s\leqslant n/b} r(n-bs)\Bigg|\sum_{\substack{a\in\mathcal{A},a|s\\ D/b<a<\min(n^{1/50},s/D)}}\chi(a)\Bigg|. \tag{3.58}$$

The interval for a is empty unless $b>Dn^{-1/50}$. Moreover, for a given s, the innermost sum is also empty unless s has a divisor $s_0=s/a$ with $a\in\mathcal{A}_n$ and $D/b<a<\min(n^{1/50},s/D)$. But then $s_0>D$ and $s_0\leqslant n/ab\leqslant n/D$. Hence, if we define

$$\mathbf{S}=\{s\leqslant n : s \text{ has a divisor } s_0\in[D,n/D]\},$$

we may add the conditions $b>Dn^{-1/50}$ and $s\in\mathbf{S}$ in (3.58) without altering the sum. By Cauchy's inequality, it follows that

$$\sum_{\substack{a\in\mathcal{A}_n,b\in\mathcal{B}_n\\ a\leqslant n^{1/50},ab<n}} r(n-ab)|H(ab)| \leqslant (S_1S_2)^{1/2} \tag{3.59}$$

where

$$S_1 = \sum_{\substack{Dn^{-1/50} < b \leqslant n/D \\ b \in \mathcal{B}_n}} \sum_{\substack{s \leqslant n/b \\ s \in \mathbf{S}}} r(n-bs),$$

$$S_2 = \sum_{\substack{b \leqslant n/D \\ b \in \mathcal{B}_n}} \sum_{s \leqslant n/b} r(n-bs) \sum_{\substack{a_1, a_2 \in \mathcal{A}_n \\ D/b < a_1, a_2 < \min(n^{1/50}, s/D) \\ a_1 | s, a_2 | s}} \chi(a_1 a_2).$$

By (3.56), (3.57) and (3.59), we then have

$$E(n) \ll (S_1 S_2)^{1/2} + n(\log n)^{-2}, \tag{3.60}$$

and it remains to bound S_1 and S_2.

We begin the estimation of S_1 by reversing the order of summation. Then

$$S_1 \leqslant \sum_{\substack{s \leqslant n^{51/50} D^{-1} \\ s \in \mathbf{S}}} \sum_{\substack{b \leqslant n/s \\ b \in \mathcal{B}_n}} r(n-bs).$$

An integer k with $k \leqslant n$ is in $\mathcal{B}_n$ if and only if k is not divisible by any of the primes in $\mathcal{A}_n$. Hence, by properties of the Möbius function,

$$\sum_{\substack{b \leqslant n/s \\ b \in \mathcal{B}_n}} r(n-bs) = \sum_{a \in \mathcal{A}_n} \mu(a) \sum_{\substack{k \leqslant n/s \\ k \equiv 0 \bmod a}} r(n-ks),$$

and therefore,

$$S_1 \leqslant \sum_{\substack{s \leqslant n^{51/50} D^{-1} \\ s \in \mathbf{S}}} \sum_{a \in \mathcal{A}_n} \mu(a) \sum_{l \leqslant n/as} r(n-las). \tag{3.61}$$

Here we split the sum over $a \in \mathcal{A}_n$ into two parts. The portion where $a > n^{1/50}$ will be estimated first. Since any $s \in \mathbf{S}$ satisfies $s \geqslant D$, the innermost sum in (3.61) is empty unless $a \leqslant n/D$. Hence, the contribution of terms with $a > n^{1/50}$ to (3.61) does not exceed

$$\sum_{\substack{a \in \mathcal{A}_n \\ n^{1/50} < a \leqslant n/D}} \sum_{D \leqslant s \leqslant n^{51/50} D^{-1}} \sum_{l \leqslant n/as} r(n-las) \leqslant \sum_{\substack{a \in \mathcal{A}_n \\ n^{1/50} < a \leqslant n/D}} \sum_{ls \leqslant n/a} r(n-las),$$

where now the inner sum is a double sum over l and s. Since this sum is symmetric in l and s, the above sum is

$$\ll \sum_{\substack{a \in \mathcal{A}_n \\ n^{1/50} < a \leqslant n/D}} \sum_{l \leqslant n/a} \sum_{l \leqslant s \leqslant n/al} r(n-las).$$

Note that the innermost sum is empty unless $l^2 \leqslant n/a$. On summing over s, we see that the previous expression is bounded by

$$\sum_{\substack{a \in \mathcal{A}_n \\ n^{1/50} < a \leqslant n/D}} \sum_{l \leqslant \sqrt{n/a}} \sum_{\substack{m \leqslant n \\ m \equiv n \bmod al}} r(m).$$

But $al \leqslant \sqrt{an} \leqslant nD^{-1/2}$, and hence, by Lemma 2.11, the above does not exceed

$$\sum_{\substack{a\in\mathcal{A}_n\\ n^{1/50}<a\leqslant n/D}} \sum_{l\leqslant\sqrt{n/a}} \frac{\tau((n,al))n\log n}{\varphi(al)}.$$

By elementary estimates, $\varphi(al) \geqslant \varphi(a)\varphi(l)$ and $\tau((n;al)) \leqslant \tau(a)\tau(l)$, and hence, by routine estimates and Lemma 2.6, the final bound for the portion of (3.61) with $a > n^{1/50}$ is

$$\ll n(\log n) \sum_{\substack{a\in\mathcal{A}\\ n^{1/50}<a\leqslant n/D}} \frac{\tau(a)}{\varphi(a)} \sum_{l\leqslant n} \frac{\tau(l)}{\varphi(l)} \ll \frac{n}{(\log n)^2}.$$

For the portion of (3.61) where $a \leqslant n^{1/50}$ we substitute the asymptotic formula from Lemma 2.11 for the innermost sum over l and find from (3.61) and the preceding discussion that

$$S_1 \ll n \sum_{\substack{s\leqslant n^{51/50}D^{-1}\\ s\in\mathbf{S}}} \Bigg| \sum_{\substack{a\in\mathcal{A}_n\\ a\leqslant n^{1/50}}} \mu(a)\frac{\lambda_n(as)}{as}\Bigg| + n^{1/2+\varepsilon} \sum_{s\leqslant n^{51/50}D^{-1}} \sum_{a\leqslant n^{1/50}} \left(\frac{(as;n)}{as}\right)^{1/4} + \frac{n}{(\log n)^2}.$$

By elementary estimates, the middle summand is $O(n^{1/2+2\varepsilon}\tau(n)^2(n^{26/25}D^{-1})^{3/4})$, and may therefore be absorbed into the third term $n/(\log n)^2$. In the first term, we wish to remove the condition $a \leqslant n^{1/50}$, and instead sum over all $a \in \mathcal{A}_n$. This introduces an error term that we may estimate via $\lambda_n(as) \leqslant \tau(s)\tau(a)$ (from Lemma 2.8) and Lemma 2.6 as

$$\ll n \sum_{s\leqslant n} \frac{\tau(s)}{s} \sum_{\substack{a\in\mathcal{A}_n\\ a>n^{1/50}}} \frac{\tau(a)}{a} \ll \frac{n}{(\log n)^2}.$$

Thus far, we have shown that

$$S_1 \ll n \sum_{\substack{s\leqslant n\\ s\in\mathbf{S}}} \frac{1}{s} \Bigg| \sum_{a\in\mathcal{A}} \frac{\mu(a)\lambda_n(as)}{a} \Bigg| + \frac{n}{(\log n)^2}. \tag{3.62}$$

By multiplicativity, we may write the sum over a in (3.62) as a product of Eulerian type. This takes the shape

$$\sum_{a\in\mathcal{A}_n} \frac{\mu(a)\lambda_n(as)}{a} = \prod_{\substack{p^\sigma\|s\\ p\notin\mathcal{A}_n}} \lambda_n(p^\sigma) \prod_{\substack{p^\sigma\|s\\ p\in\mathcal{A}_n}} (\lambda_n(p^\sigma) - p^{-1}\lambda_n(p^{\sigma+1})).$$

But $p \in \mathcal{A}_n$ implies $p \nmid n$, and in this case, $\lambda_n(p^l) = \lambda_n(p)$ for all $l \geqslant 2$ by (2.22). Hence, the previous identity simplifies to

$$\sum_{a\in\mathcal{A}_n} \frac{\mu(a)\lambda_n(as)}{a} = \lambda_n(s) \prod_{\substack{p\in\mathcal{A}_n\\ p\mid s}} \left(1 - \frac{1}{p}\right) \prod_{\substack{p\in\mathcal{A}_n\\ p\nmid s}} (1 - p^{-1}\lambda_n(p)).$$

But $\lambda_n(p) = 1 - \chi(p)p^{-1}$ whenever $p \nmid n$ (by (2.20) and (2.21)), and hence, all Euler factors are positive. Moreover, we see that

$$1 - \frac{\lambda_n(p)}{p} \leqslant 1 - \frac{1}{p} + \frac{1}{p^2} \leqslant \left(1 - \frac{1}{p}\right)\left(1 + \frac{2}{p^2}\right),$$

https://doi.org/10.1112/S0010437X09004254 Published online by Cambridge University Press

whence, recalling the definition of $\mathcal{A}_n$, we may conclude that

$$0 \leqslant \sum_{a\in\mathcal{A}_n} \frac{\mu(a)\lambda_n(as)}{a} \leqslant \lambda_n(s)\biggl(\prod_p\biggl(1+\frac{2}{p^2}\biggr)\biggr)\prod_{p\in\mathcal{A}_n}\biggl(1-\frac{1}{p}\biggr)$$
$$\ll \lambda_n(s)\prod_{p\mid n}\biggl(1-\frac{1}{p}\biggr)^{-1}\prod_{p\leqslant Y(n)}\biggl(1-\frac{1}{p}\biggr)$$
$$\ll \lambda_n(s)(\log n)^{-1}(\log\log n)^2;$$

here we have used elementary prime number theory for the final estimate. Now insert this bound into (3.62) to confirm the bound

$$S_1 \ll \frac{n(\log\log n)^2}{\log n}\sum_{\substack{s\leqslant n\\ s\in\mathbf{S}}}\frac{\lambda_n(s)}{s}+\frac{n}{(\log n)^2} \ll \frac{n(\log\log n)^3}{\log n}\sum_{\substack{s\leqslant n\\ s\in\mathbf{S}}}\frac{\tau\bigl((n;s)\bigr)}{s}+\frac{n}{(\log n)^2}; \tag{3.63}$$

for the last line we have used Lemma 2.8 and the elementary bound $s/\varphi(s) \ll \log\log s$. Put $d=(n;s)$. Then

$$\sum_{\substack{s\leqslant n\\ s\in\mathbf{S}}}\frac{\tau\bigl((n;s)\bigr)}{s} \leqslant \sum_{d\mid n}\frac{\tau(d)}{d}\sum_{\substack{s\leqslant n/d\\ ds\in\mathbf{S}}}\frac{1}{s}. \tag{3.64}$$

Let $X(n)=\exp((\log n)^{1/250})$, as in Lemma 2.7. First consider the portion of (3.64) where $d>X(n)^{1/2}$. Using the trivial bound $O(\log n)$ for the sum over s, and Lemma 2.5 for the divisor sum, we see that these terms contribute to (3.65), at most,

$$\ll (\log n)\sum_{\substack{d\mid n\\ d>X(n)^{1/2}}}\frac{\tau(d)}{d} \ll (\log n)^{-2}.$$

Now consider the terms with $d\leqslant X(n)^{1/2}$. If $ds\in\mathbf{S}$, then ds has a divisor $s_0\in[D,n/D]$. We may write $s_0=s_1s_2$ with $s_1\mid d$, $s_2\mid s$. Although such a decomposition is not necessarily unique, we may conclude that s has a divisor s_2 with $s_2\leqslant s_0\leqslant n/D$ and $s_2\geqslant s_0/d\geqslant DX(n)^{-1/2}$. In particular, recalling the definitions of D and $X(n)$, we have $s_2\in[\sqrt{n}X(n)^{-1},\sqrt{n}X(n)]$, that is, $s\in\mathcal{S}$, in the notation of Lemma 2.7. Consequently,

$$\sum_{\substack{s\leqslant n/d\\ ds\in\mathbf{S}}}\frac{1}{s} \leqslant \sum_{\substack{s\leqslant n\\ s\in\mathcal{S}}}\frac{1}{s} \ll (\log n)^{37/40},$$

uniformly in $d\leqslant X(n)^{1/2}$. A standard estimate for the outer sum over d in (3.64) suffices to conclude from the above discussion that the sum in (3.64) does not exceed $(\log n)^{37/40+\varepsilon}$. Hence, (3.63) reduces to the final estimate

$$S_1 \ll n(\log n)^{\varepsilon-3/40}. \tag{3.65}$$

It remains to estimate S_2. We reverse the order of summation in the definition implicit in (3.59) and take the resulting formula

$$S_2=\sum_{\substack{b\leqslant n/D\\ b\in\mathcal{B}_n}}\ \sum_{\substack{a_1,a_2\in\mathcal{A}_n\\ D/b<a_j\leqslant\min(n^{1/50},n/Db)}}\chi(a_1a_2)\sum_{\substack{s\leqslant n/b\\ s>Da_j\,(j=1,2)\\ s\equiv 0 \bmod [a_1;a_2]}} r(n-bs) \tag{3.66}$$

https://doi.org/10.1112/S0010437X09004254 Published online by Cambridge University Press

as the starting point. For convenience, we write $a^*=\max(a_1,a_2)$ and $q=[a_1;a_2]b$. Then $Dba^*\leqslant n$ by the conditions of summation for a_1,a_2, and hence, the inner sum in (3.66) equals

$$\sum_{\substack{Dba^*<m\leqslant n\\ m\equiv 0 \bmod q}} r(n-m)=\pi\frac{\lambda_n(q)}{q}(n-Dba^*)+O\biggl(n^{1/2+\varepsilon}\frac{(n;q)^{1/4}}{q^{1/4}}\biggr);$$

here we invoked Lemma 2.11. We insert this into (3.66) and sum the error term to

$$\begin{aligned}\sum_{b\leqslant n/D}\sum_{a_1,a_2\leqslant n^{1/50}} & n^{1/2+\varepsilon}\frac{(n;[a_1;a_2]b)^{1/4}}{([a_1;a_2]b)^{1/4}}\\ &\ll n^{1/2+\varepsilon}\sum_{a_1,a_2\leqslant n^{1/50}}(a_1a_2)^{1/4}\sum_{b\leqslant n/D}\frac{(n;b)^{1/4}}{b^{1/4}}\ll n^{1/2+1/20+\varepsilon}\Bigl(\frac{n}{D}\Bigr)^{3/4}\ll n^{49/50}.\end{aligned}$$

The formula for S_2 in (3.66) now takes the shape

$$S_2=\pi\sum_{\substack{b\leqslant n/D\\ b\in\mathcal{B}_n}}\sum_{\substack{a_1,a_2\in\mathcal{A}_n\\ D/b<a_j\leqslant\min(n^{1/50},n/Db)}}\chi(a_1a_2)(n-Dba^*)\frac{\lambda_n(q)}{q}+O(n^{49/50}).$$

We partially reverse the previous procedures by writing

$$n-Dba^*=\sum_{Dba^*<m\leqslant n}1+O(1).$$

Within the last expression for S_2, the error $O(1)$ will result in a total contribution to S_2 that does not exceed a value very much less than

$$\sum_{b\leqslant n/D}\sum_{a_1,a_2\leqslant n^{1/50}}\frac{\lambda_n(q)}{q}\ll\sum_{b\leqslant n/D}\frac{\tau(b)}{b}\sum_{a_1,a_2\leqslant n^{1/50}}\tau(a_1)\tau(a_2)\ll n^{3/50};$$

here we only used $\lambda_n(q)\ll\tau(q)$ and crude bounds. We now insert this information in the last formula for S_2 and reverse the order of summation. Since $a_j\in\mathcal{A}_n$, $b\in\mathcal{B}_n$ implies $(b;[a_1;a_2])=1$, we find

$$S_2=\pi\sum_{\substack{b\leqslant n/D\\ b\in\mathcal{B}_n}}\frac{\lambda_n(b)}{b}\sum_{D^2<m\leqslant n}\sum_{\substack{a_1,a_2\in\mathcal{A}_n\\ D/b<a_j\leqslant\min(n^{1/50},m/Db)}}\chi(a_1a_2)\frac{\lambda_n([a_1,a_2])}{[a_1;a_2]}+O(n^{49/50}).$$

We wish to remove the conditions $a_j\leqslant n^{1/50}$ from the main term here. By symmetry in a_1,a_2, this will introduce an extra contribution not exceeding

$$\begin{aligned}&\ll n\sum_{b\leqslant n/D}\frac{\tau(b)}{b}\sum_{\substack{a_1\in\mathcal{A}_n\\ a_1>n^{1/50}}}\sum_{a_2\in\mathcal{A}_n}\frac{\tau([a_1;a_2])}{[a_1;a_2]}\\ &\ll n(\log n)^2\sum_{\substack{a_1\in\mathcal{A}_n\\ a_1>n^{1/50}}}\sum_{a_2\in\mathcal{A}_n}\frac{\tau(a_1a_2/(a_1;a_2))(a_1;a_2)}{a_1a_2}\\ &\ll n(\log n)^2\sum_{d\in\mathcal{A}_n}\sum_{\substack{a_1'\in\mathcal{A}_n\\ da_1'>n^{1/50}}}\sum_{a_2'\in\mathcal{A}_n}\frac{\tau(da_1'a_2')}{da_1'a_2'}.\end{aligned}$$

https://doi.org/10.1112/S0010437X09004254 Published online by Cambridge University Press

But $\tau(da_1'a_2') \leqslant \tau(da_1')\tau(a_2')$, and $da_1' = a_1$ has at most $\tau(a_1)$ solutions in d, a_1'. Therefore, the above expression does not exceed

$$\ll n(\log n)^2 \sum_{\substack{a_1\in\mathcal{A}_n\\ a_1>n^{1/50}}} \frac{\tau(a_1)^2}{a_1} \sum_{a_2'\in\mathcal{A}_n} \frac{\tau(a_2')}{a_2'}. \tag{3.67}$$

We may now apply Lemma 2.6 to bound the sum over a_1, and for the sum over a_2', we use the elementary bound

$$\sum_{a\in\mathcal{A}_n} \frac{\tau(a)}{a} \leqslant \prod_{p\leqslant Y(n)} \sum_{l=0}^{\infty} \frac{l+1}{p^l} \ll (\log n)^2.$$

It follows that (3.67) does not exceed $O(n(\log n)^{-2})$, and consequently, the asymptotic formula for S_2 may be rewritten as

$$S_2 = \pi \sum_{\substack{b\leqslant n/D\\ b\in\mathcal{B}_n}} \frac{\lambda_n(b)}{b} \sum_{D^2<m\leqslant n} \sum_{\substack{a_1,a_2\in\mathcal{A}_n\\ D/b<a_1,a_2\leqslant m/Db}} \chi(a_1a_2)\frac{\lambda_n([a_1;a_2])}{[a_1;a_2]} + O\biggl(\frac{n}{\log n}\biggr). \tag{3.68}$$

Here, we sort the inner double sum according to the value of $a=(a_1;a_2)$. Then, on writing $a_j = aa_j'$, our preliminary finding is

$$\sum_{\substack{a_1,a_2\in\mathcal{A}_n\\ D/b<a_1,a_2\leqslant m/Db}} \chi(a_1a_2)\frac{\lambda_n([a_1;a_2])}{[a_1;a_2]} = \sum_{\substack{a\in\mathcal{A}_n\\ a\leqslant m/Db}} \frac{\chi(a)^2}{a} \sum_{\substack{a_1',a_2'\in\mathcal{A}_n\\ D/b<aa_j'\leqslant m/Db\\ (a_1';a_2')=1}} \chi(a_1'a_2')\frac{\lambda_n(a_1'a_2'a)}{a_1'a_2'}.$$

Note that $\chi(a)=0$ when a is even. Hence, the outer sum extends over odd values of $a\in\mathcal{A}_n$ only in which case $\chi(a)^2=1$. But whenever $a\in\mathcal{A}_n$ is odd, it follows from (2.22), (2.23) and (2.24) that $\lambda_n(a)>0$. Hence, for these a, the function $t\mapsto \lambda_n(at)/\lambda_n(a)$ is defined and multiplicative, and we may rewrite the right-hand side above as

$$\sum_{\substack{a\in\mathcal{A}_n\\ a\leqslant m/Db\\ a\equiv 1 \bmod 2}} \frac{\lambda_n(a)}{a} \sum_{\substack{a_1',a_2'\in\mathcal{A}_n\\ D/b<aa_j'\leqslant m/Db\\ (a_1';a_2')=1}} \chi(a_1')\chi(a_2')\frac{\lambda_n(a_1'a)\lambda_n(a_2'a)}{a_1'a_2'\lambda_n(a)^2}.$$

Now pick up the coprimality condition with the indicator $\sum_{d|(a_1';a_2')}\mu(d)$ and pull the sum over d outside. Then, on writing $a_j' = du_j$, the expression in the previous display equals

$$\begin{aligned}
&\sum_{\substack{a\in\mathcal{A}_n\\ a\leqslant m/Db\\ a\equiv 1 \bmod 2}} \frac{\lambda_n(a)}{a} \sum_{\substack{d\in\mathcal{A}_n\\ d\leqslant m/Dba}} \mu(d) \sum_{\substack{u_1,u_2\in\mathcal{A}_n\\ D/b<adu_j\leqslant m/Db}} \chi(d)^2\chi(u_1)\chi(u_2)\frac{\lambda_n(u_1da)\lambda_n(u_2da)}{d^2u_1u_2\lambda_n(a)^2}\\
&\quad= \sum_{\substack{a\in\mathcal{A}_n\\ a\leqslant m/Db\\ a\equiv 1 \bmod 2}} \frac{\lambda_n(a)}{a} \sum_{\substack{d\in\mathcal{A}_n\\ d\leqslant m/Dba\\ d\equiv 1 \bmod 2}} \frac{\mu(d)}{d^2}\biggl(\sum_{\substack{u\in\mathcal{A}_n\\ D/b<adu\leqslant m/Db}} \chi(u)\frac{\lambda_n(uda)}{u\lambda_n(a)}\biggr)^2.
\end{aligned} \tag{3.69}$$

The final expression may be rearranged by introducing $v=ad$ into the outer sum. Note that v is odd if and only if $a\equiv d\equiv 1 \bmod 2$, and that likewise one has $v\in\mathcal{A}_n$ if and only if $a\in\mathcal{A}_n$, $d\in\mathcal{A}_n$.

For such values of v, we also have $\lambda_n(v) > 0$ by Lemma 2.5, and hence (3.69) can be written in the form

$$\sum_{\substack{v\in\mathcal{A}_n\\ v\leqslant m/Db\\ v\equiv 1 \bmod 2}} \frac{\lambda_n(v)^2}{v} \sum_{ad=v} \frac{\mu(d)}{d\lambda_n(a)} \Bigg(\sum_{\substack{u\in\mathcal{A}_n\\ D/b<uv\leqslant m/Db}} \frac{\chi(u)}{u} \frac{\lambda_n(uv)}{\lambda_n(v)} \Bigg)^2. \tag{3.70}$$

For $v \in \mathcal{A}_n$ we have $(v;n) = 1$. Hence, we may use Lemma 2.9 for the convolution in (3.70). On the one hand, it shows that for any fixed v the summands in (3.70) are non-negative, and on the other, it then implies the upper bound

$$\sum_{\substack{v\in\mathcal{A}_n\\ v\leqslant m/Db\\ v\equiv 1 \bmod 2}} \frac{\lambda_n(v)\varphi(v)}{v^2} \Bigg(\sum_{\substack{u\in\mathcal{A}_n\\ D/b<uv\leqslant m/Db}} \frac{\chi(u)}{u} \frac{\lambda_n(uv)}{\lambda_n(v)} \Bigg)^2$$

for the sum in (3.70). Recalling that (3.70) is an expression for the innermost sum in (3.68), we thus deduce from the above discussion the preliminary inequality

$$S_2 \ll \sum_{\substack{b\leqslant n/D\\ b\in\mathcal{B}_n}} \frac{\lambda_n(b)}{b} \sum_{D^2<m\leqslant n} \sum_{\substack{v\in\mathcal{A}_n\\ v\leqslant m/Db\\ v\equiv 1 \bmod 2}} \frac{\lambda_n(v)}{v} \Bigg(\sum_{\substack{u\in\mathcal{A}_n\\ D/b<uv\leqslant m/Db}} \frac{\chi(u)}{u} \frac{\lambda_n(uv)}{\lambda_n(v)} \Bigg)^2 + \frac{n}{\log n}. \tag{3.71}$$

We now consider the sum over u within the square in (3.71). Recall the function $h_a(q)$ introduced in Lemma 2.10. Then, for any $v \in \mathcal{A}_n$, $v \equiv 1 \bmod 2$, Möbius inversion yields

$$\frac{\lambda_n(uv)}{\lambda_n(v)} = \sum_{w|u} h_v(w).$$

In particular, this shows that whenever $U_2 \geqslant U_1 \geqslant 1$, one has

$$\sum_{\substack{U_1<u\leqslant U_2\\ u\in\mathcal{A}_n}} \frac{\chi(u)}{u} \frac{\lambda_n(uv)}{\lambda_n(v)} = \sum_{\substack{U_1<uw\leqslant U_2\\ u\in\mathcal{A}_n, w\in\mathcal{A}_n}} h_v(w) \frac{\chi(w)}{w} \frac{\chi(u)}{u}$$

whence by Lemma 2.10 and the triangle inequality one finds that

$$\Bigg| \sum_{\substack{U_1<u\leqslant U_2\\ u\in\mathcal{A}_n}} \frac{\chi(u)}{u} \frac{\lambda_n(uv)}{\lambda_n(v)} \Bigg| \leqslant \sum_{w\leqslant U_2} \frac{1}{w^2} \Bigg| \sum_{\substack{U_1/w<u\leqslant U_2/w\\ u\in\mathcal{A}_n}} \frac{\chi(u)}{u} \Bigg|.$$

We insert this bound into (3.71) with $U_1 = D/bv, U_2 = m/Dbv$. Then we may write $k = bv$ and recall that the representation of k in this form with $v \in \mathcal{A}_n, b \in \mathcal{B}_n$ is unique. It follows that

$$S_2 \ll \sum_{D^2<m\leqslant n} \sum_{k\leqslant m/D} \frac{\lambda_n(k)}{k} \Bigg(\sum_{w\leqslant m/Dk} \frac{1}{w^2} \Bigg| \sum_{\substack{D/kw<u\leqslant m/Dkw\\ u\in\mathcal{A}_n}} \frac{\chi(u)}{u} \Bigg| \Bigg)^2 + \frac{n}{\log n}. \tag{3.72}$$

The endgame begins with an estimate for the sum over u in (3.72). We take $U_1 = D/kw$ and $U_2 = m/Dkw$, and note that $U_2/U_1 = m/D^{-2} \ll (\log n)^{200}$. Therefore, with this choice of parameters,

$$\sum_{\substack{U_1<u\leqslant U_2\\ u\in\mathcal{A}_n}} \frac{\chi(u)}{u} \ll \sum_{U_1<u\leqslant U_2} \frac{1}{u} \ll 1 + \log \frac{U_2}{U_1} \ll \log\log n. \tag{3.73}$$

https://doi.org/10.1112/S0010437X09004254 Published online by Cambridge University Press

Now let $0 < \alpha < 1/100$ denote a real number to be fixed later, and put

$$K = \sqrt{n} \exp(-(\log n)^{\alpha}).$$

The trivial bound (3.73) suffices to estimate the contribution of terms with $K < k \leqslant m/D$ in (3.72). These in fact will contribute to S_2 at most

$$\ll (\log\log n)^2 \sum_{D^2 < m \leqslant n} \sum_{K < k \leqslant m/D} \frac{\lambda_n(k)}{k} \ll (\log\log n)^3 n \sum_{K < k \leqslant n/D} \frac{\tau((n;k))}{k},$$

and standard estimates give

$$\begin{aligned}\sum_{K < k \leqslant n/D} \frac{\tau((n;k))}{k} &= \sum_{d \mid n} \frac{\tau(d)}{d} \sum_{K/d < l \leqslant n/Dd} \frac{1}{l} \\ &= \sum_{d \mid n} \frac{\tau(d)}{d} \left(\log \frac{n}{KD} + O(1) \right) \ll (\log n)^{\alpha} (\log\log n)^2\end{aligned}$$

so that the total contribution of terms with $K < k \leqslant m/D$ in (3.72) is $O(n(\log n)^{\alpha+\varepsilon})$.

It remains to consider terms with $k \leqslant K$ in (3.72). The contribution of the subsum with the extra condition $w > \log n$ is again estimated by (3.73), and is therefore bounded by

$$\ll (\log\log n)^2 \sum_{D^2 < m \leqslant n} \sum_{k \leqslant K} \frac{\lambda_n(k)}{k} \left(\sum_{w > \log n} \frac{1}{w^2} \right)^2 \ll \left(\frac{\log\log n}{\log n} \right)^2 n \sum_{k \leqslant \sqrt{n}} \frac{\tau(k)}{k} \ll n (\log\log n)^2,$$

so that we are now reduced to the subsum of (3.72) where $k \leqslant K$ and $w \leqslant \log n$. We take $U_1 = D/kw$, $U_2 = m/Dkw$, as before. Then, $U_1 \geqslant D/(K \log n) \geqslant (\log n)^{-100} \exp((\log n)^{\alpha}) \geqslant \exp((\log n)^{\alpha/2})$ when n is large. Moreover, $U_2 \leqslant n$. Under these conditions, one has

$$\sum_{\substack{U_1 < u \leqslant U_2 \\ u \in \mathcal{A}_n}} \frac{\chi(u)}{u} \ll (\log n)^{-5}, \tag{3.74}$$

as we shall see in a moment. Equipped with (3.74), we now see that the portion of (3.72) with $k \leqslant K$ and $w \leqslant \log n$ contributes to S_2 an amount not exceeding

$$\ll (\log n)^{-10} \sum_{D^2 < m \leqslant n} \sum_{k \leqslant m/D} \frac{\lambda_n(k)}{k} \ll n (\log n)^{-8};$$

here it suffices to use Lemma 2.8 ($\lambda_n(k) \ll \tau(k)$) and straightforward bounds. We have now completed the estimation of the right-hand side of (3.72), and may infer the bound $S_2 \ll n(\log n)^{\alpha+\varepsilon}$ from the above discussion. Since $0 < \alpha < 1/100$ was an arbitrary real number, it follows that $S_2 \ll n(\log n)^{\varepsilon}$. Hence, by (3.60) and (3.65), we finally conclude that $E(n) \ll n(\log n)^{\varepsilon - 3/80}$. The proof of Theorem 4 is now complete.

It remains to establish (3.74). Actually, much better bounds are available from Lemma 2.4. To see this, note that

$$\sum_{\substack{U_1 < u \leqslant U_2 \\ u \in \mathcal{A}_n}} \frac{\chi(u)}{u} = \sum_{\substack{U_1 < u \leqslant U_2 \\ P(u) \leqslant Y(n) \\ (u;n)=1}} \frac{\chi(u)}{u} = \sum_{\substack{d \mid n \\ P(d) \leqslant Y(n)}} \mu(d) \frac{\chi(d)}{d} \sum_{\substack{U_1/d < k \leqslant U_2/d \\ P(k) \leqslant Y(n)}} \frac{\chi(k)}{k}$$

whence

$$\Bigl|\sum_{\substack{U_1<u\leqslant U_2\\ u\in\mathcal{A}_n}}\frac{\chi(u)}{u}\Bigr| \leqslant \sum_{\substack{d\mid n\\ d\leqslant U_2}}\frac{1}{d}\Bigl|\sum_{\substack{U_1/d<k\leqslant U_2/d\\ P(k)\leqslant Y(n)}}\frac{\chi(k)}{k}\Bigr|. \tag{3.75}$$

Note that for $U_1 < d \leqslant U_2$, the innermost sum starts at $k=1$. In the ranges for U_1, U_2 for which we claim (3.74), we import from Lemma 2.4 with $X=n, Y=Y(n)$ the bounds

$$\sum_{\substack{U_1/d<k\leqslant U_2/d\\ P(k)\leqslant Y(n)}}\frac{\chi(k)}{k} \ll \min\Bigl(\frac{d}{U_1},1\Bigr)+\exp(-(\log n)^{2/5}),$$

and by (3.75), we then find that

$$\sum_{\substack{U_1<u\leqslant U_2\\ u\in\mathcal{A}_n}}\frac{\chi(u)}{u} \ll \sum_{\substack{d\mid n\\ d\leqslant U_2}}\frac{1}{d}\min\Bigl(\frac{d}{U_1},1\Bigr)+\exp\Bigl(-\frac{1}{2}(\log n)^{2/5}\Bigr).$$

For the remaining divisor sum on the right-hand side here, we use the upper bound

$$\leqslant \sum_{\substack{d\mid n\\ d>U_1}}\frac{1}{d}+U_1^{-1}\sum_{\substack{d\mid n\\ d\leqslant U_1}}1 \leqslant \sum_{\substack{d\mid n\\ d>U_1}}\frac{1}{d}+\sum_{\substack{d\mid n\\ \sqrt{U_1}<d\leqslant U_1}}\frac{1}{d}+U_1^{-1}\sum_{\substack{d\mid n\\ d\leqslant\sqrt{U_1}}}1.$$

The sum on the far right is trivially bounded by $O(U_1^{-1/2})$, and the first two sums may be recombined and bounded by Lemma 2.5 as

$$\sum_{\substack{d\mid n\\ d>\sqrt{U_1}}}\frac{1}{d} \ll \exp(-(\log n)^{\alpha/4}),$$

provided one has $U_1 \geqslant \exp((\log n)^{\alpha/2})$, as we have assumed.

4. A combinatorial interlude

We begin by a detailed description of an application of the inclusion–exclusion principle that will ultimately yield a lower bound for $R(n,\theta)$. For a readier exposition, let $\mathcal{X}(n)$ be the set of all $\mathbf{x}=(x_1,\dots,x_4)\in\mathbb{Z}^4$ satisfying (1.3). Recall the definition of $R_j(n)$ from the introduction, and note that $R(n,\theta)=R_4(n,\theta)$. We wish to relate R_4 with R_2 and R_1 because the latter two can be evaluated asymptotically. By the inclusion–exclusion principle,

$$\begin{aligned} R_4(n,\theta) = R_2(n,\theta) &- \#\{\mathbf{x}\in\mathcal{X}(n) : P(x_1x_2)\leqslant n^{\theta/2}, P(x_3)>n^{\theta/2}\}\\ &- \#\{\mathbf{x}\in\mathcal{X}(n) : P(x_1x_2)\leqslant n^{\theta/2}, P(x_4)>n^{\theta/2}\}\\ &+ \#\{\mathbf{x}\in\mathcal{X}(n) : P(x_1x_2)\leqslant n^{\theta/2}, P(x_3)>n^{\theta/2}, P(x_4)>n^{\theta/2}\}.\end{aligned}$$

The two middle terms in the sum on the right are equal, by symmetry. It follows that

$$R_4(n,\theta)\geqslant R_2(n,\theta)-2U(n) \tag{4.1}$$

where

$$U(n)=\#\{\mathbf{x}\in\mathcal{X}(n) : P(x_1x_2)\leqslant n^{\theta/2}, P(x_3)>n^{\theta/2}\}.$$

https://doi.org/10.1112/S0010437X09004254 Published online by Cambridge University Press

Yet another application of the same idea, this time with respect to x_1, shows that

$$U(n) = \#\{\mathbf{x} \in \mathcal{X}(n) : P(x_2) \leqslant n^{\theta/2}, P(x_3) > n^{\theta/2}\} \\ - \#\{\mathbf{x} \in \mathcal{X}(n) : P(x_1) > n^{\theta/2}, P(x_2) \leqslant n^{\theta/2}, P(x_3) > n^{\theta/2}\}.$$

However, directly from the definition,

$$R_1(n,\theta) - R_2(n,\theta) = \#\{\mathbf{x} \in \mathcal{X}(n) : P(x_1) \leqslant n^{\theta/2}, P(x_2) > n^{\theta/2}\},$$

and hence, by symmetry in $x_1, \ldots, x_4$,

$$U(n) \leqslant R_1(n,\theta) - R_2(n,\theta).$$

From (4.1) we now infer that

$$R(n,\theta) \geqslant 3R_2(n,\theta) - 2R_1(n,\theta). \tag{4.2}$$

Our initial manipulations finished, we now turn to the evaluation of the right-hand side of (4.2) with the aid of Theorem 4. When $1/2 < \theta < 1$, then $\varrho(1/\theta) = 1 + \log\theta$ by the definition of ϱ. Consequently,

$$R(n,\theta) \geqslant (3(\log\theta)^2 + 4\log\theta + 1)\pi^2\mathfrak{S}(n)n + O(n(\log n)^{-\gamma})$$

holds for all $4 \nmid n$. The first factor on the right-hand side is positive for $\theta > e^{-1/3}$. Moreover, as we saw in § 3.3, the singular series (1.2) may be rewritten as an Euler product. If $n = \prod_p p^{e_p}$, then

$$\mathfrak{S}(n) = \frac{1 + (-1)^n}{2^{e_2}} \prod_{p \neq 2} \left(1 + \frac{1}{p}\right)\left(1 - \frac{1}{p^{e_p+1}}\right). \tag{4.3}$$

In particular, it follows that $\mathfrak{S}(n) \gg 1$ for $4 \nmid n$. This proves Theorem 1 for all large n with $4 \nmid n$. If $4|n$, write $n = 4^k m$ with $4 \nmid m$. From (4.3) or (1.1) it follows that either $\mathfrak{S}(n)n = \mathfrak{S}(m)m$ or $\mathfrak{S}(n)n = 3\mathfrak{S}(m)m$. Also, any solution of $x_1^2 + \cdots + x_4^2 = m$ with $P(x_1 \cdots x_4) \leqslant m^{\theta/2}$ can be multiplied by 4^k to produce a solution of $y_1^2 + \cdots + y_4^2 = n$ with $P(y_1 \cdots y_4) \leqslant m^{\theta/2}$. Hence, $R(n,\theta) \gg n\mathfrak{S}(n)$ for all n with $4 \nmid n$ implies $R(n,\theta) \gg \mathfrak{S}(n)n$ for all n. This completes the proof of Theorem 1.

5. Ternary quadratic forms

In this section we shall prove Theorem 3. As indicated in the introduction, Theorem 2 will be a simple consequence. Let $\varepsilon > 0$, let $n \not\equiv 0, 4, 7 \bmod 8$ be sufficiently large (in terms of ε) and choose three different primes $p_1, p_2, p_3 \in \mathbb{Z} \cap [n^{1/148-\varepsilon}, 2n^{1/148-\varepsilon}]$ such that $p_1p_2p_3 \nmid n$ and $p_j \equiv 1 \bmod 4$. We show that the Diophantine equation

$$n = q(\mathbf{x}) := (p_1x_1)^2 + (p_2x_2)^2 + (p_3x_3)^2$$

has a solution $\mathbf{x} \in \mathbb{Z}^3$. Obviously this implies Theorem 3. The number of representations $r(q,n)$ of n by the quadratic form q can be approximated by the Siegel mean (see for example [Blo08] for the definitions)

$$r(\text{gen } q, n) = \frac{2\pi}{p_1p_2p_3} \prod_p r_p(q,n)$$

where

$$r_p(q,n) = \lim_{\nu\to\infty} \frac{1}{p^{2\nu}} \#\{\mathbf{x} \in (\mathbb{Z}/p^\nu\mathbb{Z})^3 \mid q(\mathbf{x}) \equiv n \bmod p^\nu\}$$

https://doi.org/10.1112/S0010437X09004254 Published online by Cambridge University Press

are the local densities. We have $r_p(q, n) = 1 + (\frac{-n}{p})p^{-1}$ if $p \nmid 2p_1p_2p_3n$, and $r_p(q, n) \geqslant 1 - 1/p$ if $p \nmid 2p_1p_2p_3$ [Sie35, Hilfssätze 12, 16]. In order to calculate the local densities at the ramified primes, we observe that for $d \equiv 1$ (modulo 4) and $(h, d) = 1$ one has by quadratic reciprocity

$$G(h, d) = \sum_{b \bmod d} e\left(\frac{hb^2}{d}\right) = \left(\frac{h}{d}\right)\sqrt{d}.$$

For $j = 1, 2, 3$ this gives

$$r_{p_j}(q, n) = \sum_{k=0}^{\infty} \frac{1}{p_j^{3k}} \sum_{\substack{h \bmod p_j^k \\ (h; p_j)=1}} G(hp_1^2, p_j^k)G(hp_2^2, p_j^k)G(hp_3^2, p_j^k)e\left(-\frac{hn}{p_j^k}\right) = 1 - \frac{1}{p_j}.$$

Finally one can check by direct computation that $r_2(q, n) = 3/2$ if $n \equiv 1, 2 \bmod 4$ and $r_2(q, n) = 1$ if $n \equiv 3 \bmod 8$. Using Siegel's lower bound for $L(1, \chi_{-n})$, we find

$$r(\text{gen } q, n) \gg \frac{n^{1/2-\varepsilon}}{p_1p_2p_3}.$$

By [Ome73, (102:10)], the quadratic form q has only one spinor genus per genus. Thus $r(\text{gen}\, q, n) = r(\text{spn}\, q, n)$ (cf. [Blo08] for the notation), and by [Blo08, (2.7)] we have

$$r(q, n) - r(\text{spn}\, q, n) \ll n^{\varepsilon} p_1^{1/4}(p_1p_2p_3)^2(n^{7/16} + (p_1p_2p_3)^{1/2}n^{3/8}) \ll \frac{n^{1/2-2\varepsilon}}{p_1p_2p_3}$$

for our choice of p_1, p_2, p_3. In particular, $r(q, n) > 0$ for sufficiently large n. This completes the proof of Theorem 3. □

In order to deduce Theorem 2, we start with the following (essentially known) observation concerning smooth numbers in short intervals, see e.g. [FL87]. Let m be any positive integer. Choose $a := \lceil \sqrt{m} \rceil$ and $b \in \mathbb{Z} \cap [\sqrt{a^2 - m}, 9 + \sqrt{a^2 - m}]$, and write $k := a^2 - b^2$. Then we have $m - (36 + o(1))m^{1/4} \leqslant k \leqslant m$ and $P(k) \leqslant a + b \leqslant (1 + o(1))\sqrt{m}$. For the proof of Theorem 2 we can assume that $4 \nmid n$, cf. the remark after (4.3). We may choose $x_4 \in \mathbb{Z} \cap [n^{1/2} - 37n^{1/8}, n^{1/2}]$ such that $P(x_4) \leqslant 2n^{1/4}$ and $n - x_4^2 \not\equiv 0, 4, 7 \bmod 8$. Note that $n - x_4^2 \ll n^{5/8}$. By Theorem 3, we can find x_1, x_2, x_3 such that $n - x_4^2 = x_1^2 + x_2^2 + x_3^2$ and $P(x_1x_2x_3) \ll n^{((5/8)\cdot(73/148))+\varepsilon}$, and Theorem 2 follows with $\theta > 365/592$.

Acknowledgements

The authors wish to express their sincere gratitude to an unknown referee who has read the original manuscript with utmost care, and made very useful suggestions towards an improved exposition of the material herein.

References

Bal89 A. Balog, *On additive representation of integers*, Acta Math. Hungar. **54** (1989), 297–301.

Blo04 V. Blomer, *Uniform bounds for Fourier coefficients of theta-series with arithmetic applications*, Acta Arith. **114** (2004), 1–21.

Blo08 V. Blomer, *Ternary quadratic forms, and sums of three squares with restricted variables*, in *The anatomy of integers*, CRM Proceedings and Lecture Notes, vol. 46 eds A. Granville, F. Luca and J.-M. de Koninck (American Mathematical Society, Providence, RI, 2008), 1–17.

BF94 J. Brüdern and E. Fouvry, *Lagrange's four squares theorem with almost prime variables*, J. Reine Angew. Math. **454** (1994), 59–96.

Che73 J. R. Chen, *On the representation of a large even integer as the sum of a prime and the product of at most two primes*, Sci. Sinica **16** (1973), 157–176.

Est45 T. Estermann, *On the sign of the Gaussian sum*, J. Lond. Math. Soc. **20** (1945), 66–67.

FT91 E. Fouvry and G. Tenenbaum, *Entiers sans grand facteur premier en progressions arithmétiques*, Proc. Lond. Math. Soc. (3) **63** (1991), 449–494.

FL87 J. Friedlander and J. C. Lagarias, *On the distribution in short intervals of integers having no large prime factor*, J. Number Theory **25** (1987), 249–273.

HT88 R. R. Hall and G. Tenenbaum, *Divisors*, Cambridge Tracts in Mathematics, vol. 90 (Cambridge University Press, Cambridge, 1988).

Har99 G. Harcos, *On sums of four smooth squares*, J. Number Theory **77** (1999), 145–154.

HW79 G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, fifth edition (Oxford University Press, New York, 1979).

HT03 D. R. Heath-Brown and D. I. Tolev, *Lagrange's four squares theorem with one prime and three almost-prime variables*, J. Reine Angew. Math. **558** (2003), 159–224.

Hoo57 C. Hooley, *On the representation of a number as the sum of two squares and a prime*, Acta Math. **97** (1957), 189–210.

Hoo75 C. Hooley, *On the Barban–Davenport–Halberstam theorem. III*, J. Lond. Math. Soc. (2) **10** (1975), 249–256.

Hoo79 C. Hooley, *On a new technique and its applications to the theory of numbers*, Proc. Lond. Math. Soc. (3) **38** (1979), 115–151.

Hoo98 C. Hooley, *On the Barban–Davenport–Halberstam theorem. X*, Hardy-Ramanujan J. **21** (1998), 9 (electronic).

Kov82 F. B. Kovalchik, *Analogues of the Hardy–Littlewood equation,* in *Integral lattices and finite linear groups*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **116** (1982), 86–95, 163 (In Russian).

Neu06 S. Neumann, *Mean square value theorems for integers without large prime factors*, PhD thesis, Universität Stuttgart (2006).

Ome73 O. T. O'Meara, *Introduction to quadratic forms* (Springer, Berlin, 1973).

Pla81 V. A. Plaksin, *Asymptotic formula for the number of solutions of an equation with primes*, Izv. Akad. Nauk SSSR Ser. Mat. **45** (1981), 321–397, 463 (In Russian).

Pla84 V. A. Plaksin, *Asymptotic formula for the number of representations of a natural number by a pair of quadratic forms, the arguments of one of which are primes*, Izv. Akad. Nauk SSSR Ser. Mat. **48** (1984), 1245–1265 (In Russian).

Shi79 P. Shields, *Some applications of sieve methods in number theory*, PhD thesis, University College, Cardiff (1979).

Shi80 P. Shiu, *A Brun–Titchmarsh theorem for multiplicative functions*, J. Reine Angew. Math. **313** (1980), 161–170.

Sie35 C. L. Siegel, *Über die analytische Theorie quadratischer Formen I*, Ann. of Math. (2) **36** (1935), 527–606.

Smi68 R. A. Smith, *The circle problem in an arithmetic progression*, Canad. Math. Bull. **11** (1968), 175–184.

Ten84 G. Tenenbaum, *Sur la probabilité qu'un entier possède un diviseur dans un intervalle donné*, Compositio Math. **51** (1984), 243–263.

Ten95 G. Tenenbaum, *Introduction to analytic and probabilistic number theory* (Cambridge University Press, Cambridge, 1995).

Tol03 D. Tolev, *Lagrange's four squares theorem with variables of special type*, in *Proc. of the session in analytic number theory and diophantine equations*, Bonner Mathematische Schriften, vol. 360, eds D. R. Heath-Brown and B. Z. Moroz (Universität Bonn, 2003), 17 pp.

Vau98 R. C. Vaughan, *On a variance associated with the distribution of general sequences in arithmetic progressions. I, II*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **356** (1998), 781–791, 793–809.

VW02 R. C. Vaughan and T. D. Wooley, *Waring's problem: a survey*, in *Number theory for the millennium, III,* Urbana, IL, 21–26 May 2000, eds M. A. Bennett, B. C. Berndt, N. Boston, H. G. Diamond, A. J. Hildebrand and W. Philipp (A K Peters, Natick, MA, 2002), 301–340.

Woo02 Wooley T. D., *Slim exceptional sets for sums of four squares*, Proc. Lond. Math. Soc. (3) **85** (2002), 1–21.

V. Blomer vblomer@math.toronto.edu
Department of Mathematics, University of Toronto, 40 St. George Street, Toronto, Ontario, Canada M5S 2E4, USA

J. Brüdern bruedern@mathematik.uni-stuttgart.de
Universität Stuttgart, Institut für Algebra und Zahlentheorie, Pfaffenwaldring 57,
D-70511 Stuttgart, Germany

R. Dietmann dietmarr@mathematik.uni-stuttgart.de
Universität Stuttgart, Institut für Algebra und Zahlentheorie, Pfaffenwaldring 57,
D-70511 Stuttgart, Germany

Current address: Department of Mathematics, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK