Journal of Number Theory **77**, 145–154 (1999)
Article ID jnth.1999.2380, available online at http://www.idealibrary.com on IDEAL®

On Sums of Four Smooth Squares

Gergely Harcos*

Department of Mathematics, University of Illinois at Urbana-Champaign, 1409 West Green Street, Urbana, Illinois 61801
E-mail: harcos@math.uiuc.edu, gharcos@math.princeton.edu

Communicated by A. Hildebrand

Received May 12, 1998

We show that almost every natural number M is the sum of four squares with all their prime factors smaller than $\exp(20(\log M \log\log M)^{1/2})$.

1. INTRODUCTION

The present paper serves as a supplement to the author's previous work [3] which initiated the study of Waring's problem in smooth variables. That work was based on a conjecture of Sárközy stating that given any $\varepsilon > 0$ every sufficiently large natural number N is the sum of four squares with all their prime factors smaller than N^{ε}. Such a result would be the exact counterpart to the theorem of Brüdern and Fouvry [2] that every sufficiently large natural number N congruent to 4 modulo 24 is the sum of four squares with all their prime factors greater than $N^{1/69}$. While Sárközy's conjecture remains open, [3] was able to confirm a strong form of the analogous statement for five squares. The techniques developed there are flexible enough to show that exceptions to the Sárközy conjecture are rare. In fact, writing $P(n)$ for the greatest prime factor of a number n we have

THEOREM. *Almost every natural number M not divisible by* 8 *has at least*

$$M \exp(-(\log M \log\log M)^{1/2}/8)$$

representations of the form

$$n_1^2 + n_2^2 + n_3^2 + n_4^2 = M,$$

* Current address: Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton, NJ 08544.

0022-314X/99 $30.00

where

$$P(n_1 n_2 n_3 n_4) < \exp(20(\log M \log\log M)^{1/2}).$$

More precisely, the number of exceptions up to N is

$$\ll N \exp(-(\log N \log\log N)^{1/2}/16).$$

As any representation of M gives rise to a representation of $4M$ by multiplying each square term by 4 we also have the following

COROLLARY. *Almost every natural number M is the sum of four squares with all their prime factors smaller than* $\exp(20(\log M \log\log M)^{1/2})$. *More precisely, the number of exceptions up to N is* $\ll N \exp(-(\log N \log\log N)^{1/2}/16)$.

2. NOTATIONS AND PRELIMINARY REMARKS

Our exposition mainly follows [3] so that a variant of the Hardy–Littlewood method will be applied which goes back to the work of Balog and Sárközy [1]. A new ingredient is the careful study of the convergence of the singular series which is based on well-known properties of Gauß and Ramanujan sums.

We write $e^x = \exp(x)$ and $e^{2\pi i\alpha} = e(\alpha)$. We define the empty sum to be 0 and the empty product to be 1. We denote the least prime factor of n by $p(n)$, while the greatest prime factor of n is denoted by $P(n)$. We write $\omega(n) = \sum_{p \mid n} 1$ for the number of prime factors of n and $\Omega(n) = \sum_{p^\alpha \parallel n} \alpha$ for the number of prime factors of n with multiplicity. The Vinogradov symbols $\ll$, $\gg$ have their usual meaning, namely that for functions f and g with g taking non-negative values $f \ll g$ and $g \gg f$ means $|f| \leqslant Cg$ where C is a constant. The dependence of the implicit constants in the O, $\ll$ and $\gg$ notations, if any, will be indicated explicitly in the subscript of these symbols. Whenever we use the o-symbol without any comment, we understand $N \to \infty$ to be implicit.

We shall always assume that $1 \leqslant M \leqslant N$ and that N is sufficiently large to fit in all our statements. Define w to be $\exp((\log N^{1/2} \log\log N^{1/2})^{1/2})$, and put $y = w^{27}$ and $z = y^{2/9} = w^6$. Let

$$Q = \frac{N}{z^{1/4}}, \qquad U = \left[4\frac{N}{y}\right] + 1,$$

$$\mathscr{L} = \left\{ l : \frac{9}{10}\frac{N^{1/2}}{y} \leqslant l \leqslant \frac{N^{1/2}}{y} \text{ and } z < p(l) \leqslant P(l) \leqslant y \right\}.$$

We introduce the weights

$$d_n = \sum_{\substack{ml=n \\ m \leqslant y \\ l \in \mathscr{L}}} 1 \qquad (\text{for } 1 \leqslant n \leqslant N^{1/2}),$$

and the generating functions

$$f(\alpha) = \sum_{1 \leqslant n \leqslant N^{1/2}} d_n e(n^2 \alpha), \qquad u(\alpha) = \frac{1}{U} \sum_{n=0}^{U-1} e(n\alpha),$$

$$h(\alpha) = f(\alpha)\, u(\alpha) = \sum_{n=1}^{N+U-1} h_n e(n\alpha)$$

so that

$$h_n = \frac{1}{U} \sum_{n-U<j^2 \leqslant n} d_j.$$

We shall study for $1 \leqslant M \leqslant N$ the integrals

$$J(M) = \int_0^1 f(\alpha)^4\, e(-M\alpha)\, d\alpha = \sum_{n_1^2+n_2^2+n_3^2+n_4^2=M} d_{n_1} d_{n_2} d_{n_3} d_{n_4}$$

by comparing them to

$$I(M) = \int_0^1 h(\alpha)^4\, e(-M\alpha)\, d\alpha = \sum_{n_1+n_2+n_3+n_4=M} h_{n_1} h_{n_2} h_{n_3} h_{n_4}.$$

It will be convenient to work on the unit interval $\mathfrak{U} = [1/Q, 1+1/Q]$ instead of $[0, 1]$. Note that the generating functions $e(\alpha)$, $f(\alpha)$, $u(\alpha)$ and $h(\alpha)$ appearing here are all periodic modulo 1, hence they depend only on the fractional part of α. For any $1 \leqslant a \leqslant q \leqslant z$ and $(a, q) = 1$ introduce the intervals

$$\mathfrak{M}(q, a) = \left\{ \alpha : \left| \alpha - \frac{a}{q} \right| \leqslant \frac{1}{Q} \right\},$$

which are pairwise disjoint and lie in $\mathfrak{U}$. Hence if we define the *major arcs* $\mathfrak{M}$ as the union of these intervals and the *minor arcs* as $\mathfrak{m} = \mathfrak{U} \backslash \mathfrak{M}$, then clearly $\mathfrak{U}$ is the disjoint union of $\mathfrak{M}$ and $\mathfrak{m}$. According to this decomposition we can write $J(M)$ and $I(M)$ as

$$J(M) = J_{\mathfrak{M}}(M) + J_{\mathfrak{m}}(M), \qquad I(M) = I_{\mathfrak{M}}(M) + I_{\mathfrak{m}}(M),$$

respectively.

We put, for every $1 \leqslant M \leqslant N$,

$$S_M(q) = \sum_{\substack{1 \leqslant a \leqslant q \\ (a, q) = 1}} (q^{-1} S(q, a))^4 \, e(-Ma/q),$$

where $S(q, a)$ denotes the Gauß sum

$$S(q, a) = \sum_{r=1}^{q} e(r^2 a/q).$$

Then the singular series $\mathfrak{S}(M)$ and the truncated singular series $\mathfrak{S}(M, P)$ is defined as

$$\mathfrak{S}(M) = \sum_{q=1}^{\infty} S_M(q), \qquad \mathfrak{S}(M, P) = \sum_{1 \leqslant q \leqslant P} S_M(q),$$

respectively.

3. PRELIMINARY LEMMATA

The first two lemmata deal with estimations of the coefficients of f and h. The first result improves upon [3, Lemma 3]. While the previous version would also suffice for the present purposes the author takes the opportunity to include a stronger statement (which was pointed out to him by Imre Z. Ruzsa).

Lemma 1. *We have*

$$\max_{1 \leqslant n \leqslant N^{1/2}} d_n \ll \left(\frac{\log N}{\log \log N} \right)^2.$$

Proof. Decompose each n as $n = n^- n^+$ where

$$n^- = \prod_{\substack{p^\alpha \| n \\ p \leqslant z}} p^\alpha \qquad \text{and} \qquad n^+ = \prod_{\substack{p^\alpha \| n \\ p > z}} p^\alpha.$$

Then in any representation $n = ml$ contributing to d_n we must have $l \mid n^+$, i.e., $n^- \mid m$. Hence $m^* = m/n^-$ is a positive integer dividing n^+. In particular,

$$z < p(n^+) \leqslant p(m^*) \leqslant m^* \leqslant m \leqslant y < z^5$$

which implies that m^* has at most 4 prime factors. These prime factors are among those of n^+, therefore the number of choices for m^* is at most

$$\sum_{r=0}^{4} \binom{\Omega(n^+)}{r} \ll \Omega(n^+)^4 \leqslant \left(\frac{\log n^+}{\log z}\right)^4 \ll \left(\frac{\log N}{\log\log N}\right)^2.$$

This completes the proof of the lemma as m^* determines the representation $n=ml$ uniquely. ▮

In contrast to the previous lemma we deduce a lower bound for the averages of the d_n.

LEMMA 2. *We have*

$$\min_{U \leqslant n \leqslant 3N/4} h_n \geqslant N^{-1/2} w^{-1/27-o(1)}.$$

Proof. This follows from a combination of [3, Lemmata 4 and 5] with the choices $\lambda = 9/10$ and $c_3 = 27$. ▮

We now collect our main estimates for the generating functions f and h. We describe the behavior of these functions on the major arcs and then proceed to give a nontrivial upper estimate for them on the minor arcs. Finally, an appropriate mean value result is given.

LEMMA 3. *If* $\alpha \in \mathfrak{M}$, *e.g.* $\alpha \in \mathfrak{M}(q, a)$ *then*

$$f(\alpha) = q^{-1} S(q, a)\, f(\beta) + O(N^{1/2} y^{-1} z^2), \qquad h(\alpha) = \delta_q h(\beta) + O(N^{1/2} y^{-1} z^2),$$

where $\beta = \alpha - a/q$ *and* $\delta_q = 1$ *or* 0 *according as* $q = 1$ *or* $q > 1$.

Proof. This follows from [3, Lemma 10]. ▮

LEMMA 4. *We have*

$$\sup_{\mathfrak{m}} |f(\alpha)| \leqslant \frac{N^{1/2}}{z^{1/48}} w^{o(1)}.$$

Proof. This is a special case of [3, Lemma 13]. ▮

LEMMA 5. *There exists a positive constant* c *such that*

$$\int_0^1 |f(\alpha)|^4 \, d\alpha \leqslant N (\log N)^c.$$

Proof. This follows from Lemma 1 and Hua's inequality [4, Satz 4] by considering the underlying diophantine equation (cf. [3, Lemma 12]). ▮

We conclude this section by giving an appropriate estimate for $S_M(q)$ which enables to deduce that the singluar series $\mathfrak{S}(M)$ converges quickly enough for our purposes.

LEMMA 6. *We have*

$$|S_M(q)| \leqslant 4q^{-2}(q, M).$$

Proof. It is known [5, Lemma 2.11] that $S_M(q)$ is multiplicative in q, hence it suffices to show that

$$|S_M(2^l)| \leqslant 2^{2-2l}(2^l, M)$$

and

$$|S_M(p^l)| \leqslant p^{-2l}(p^l, M)$$

for any odd prime p.

For the first inequality use [5, Lemma 4.4] to see that

$$S(2^l, a)^4 = \begin{cases} -2^{2l+2}, & \text{if } \ l > 1; \\ 0, & \text{if } \ l = 1; \end{cases}$$

whenever $2 \nmid a$ so that

$$|S_M(2^l)| \leqslant 2^{2-2l} \left| \sum_{\substack{1 \leqslant a \leqslant 2^l \\ (a, 2) = 1}} e(-Ma/2^l) \right|$$

which yields the result.

The proof of the second estimate is similar but now based on the fact that $S(p^l, a)^4 = p^{2l}$ whenever $p \nmid a$ which follows from [5, Lemma 4.4] combined with the basic identity $S(p, a)^4 = p^2$. ▮

LEMMA 7. *We have, for any* $1 \leqslant M \leqslant N$,

$$\mathfrak{S}(M) = \mathfrak{S}(M, P) + O(P^{-1/2} w^{o(1)}).$$

In particular,

$$\mathfrak{S}(M) \ll w^{o(1)}.$$

Proof. By the previous lemma we have

$$\mathfrak{S}(M)-\mathfrak{S}(M,P)=\sum_{q>P} S_M(q) \ll \sum_{q>P} q^{-2}(q,M)$$

$$\leqslant \sum_{d\mid M} d \sum_{\substack{q>P\\ d\mid q}} q^{-2} = \sum_{d\mid M} d^{-1} \sum_{r>P/d} r^{-2}$$

$$\ll \sum_{\substack{d\mid M\\ d\leqslant P}} P^{-1} + \sum_{\substack{d\mid M\\ d>P}} d^{-1} \leqslant P^{-1/2} \sum_{d\mid M} d^{-1/2}.$$

The sum on the right is

$$\sum_{d\mid M} d^{-1/2} = \prod_{p^\alpha \parallel M} \sum_{\beta=0}^{\alpha} p^{-\beta/2} = \prod_{p\mid M} \exp(O(p^{-1/2})) = \exp\{O(\sum_{p\mid M} p^{-1/2})\},$$

hence we are left with estimating $\sum_{p\mid M} p^{-1/2}$. If $p_1<p_2<\cdots$ denotes the series of prime numbers in increasing order then the sum in question is clearly at most $\sum_{k=1}^{\omega(M)} p_k^{-1/2}$, whence the Chebyshev bound $p_k \gg k \log k$ implies that

$$\sum_{p\mid M} p^{-1/2} \ll \sum_{k=1}^{\omega(M)} (k\log k)^{-1/2} \ll \left(\frac{\omega(M)}{\log \omega(M)}\right)^{1/2} \ll \frac{(\log N)^{1/2}}{\log\log N} = o(\log w).$$

The proof is complete. ∎

LEMMA 8. *If M is not divisible by 8 then*

$$\mathfrak{S}(M) \gg 1.$$

Proof. This is a consequence of [5, Lemma 2.15 and Theorem 4.5] on noting that the congruence

$$x_1^2+x_2^2+x_3^2+x_4^2 \equiv M \pmod 8$$

is soluble with $2\nmid x_1$. ∎

4. PROOF OF THE THEOREM

Let $\mathscr{M}$ denote the set of exceptional natural numbers M not divisible by 8 in the Theorem. We have to prove that

$$|\mathscr{M}\cap[1,N]| \ll N\exp(-(\log N \log\log N)^{1/2}/16).$$

Clearly, it suffices to show that, for N large enough,

$$|\mathscr{M}\cap[N/2,N]|\leqslant N\exp(-(\log N\log\log N)^{1/2}/16). \qquad (1)$$

As

$$y=w^{27}<\exp(20(\log M\log\log M^{1/2}) \qquad (N/2\leqslant M\leqslant N),$$

Lemma 1 implies that for any exceptional $M\in\mathscr{M}\cap[N/2,N]$ we have

$$\begin{aligned} J(M)=&\sum_{n_1^2+n_2^2+n_3^2+n_4^2=M} d_{n_1}d_{n_2}d_{n_3}d_{n_4}\\ &\leqslant w^{o(1)}M\exp(-(\log M\log\log M)^{1/2}/8).\end{aligned}$$

As a result,

$$J(M)\leqslant\frac{N}{w^{1/6}} \qquad (M\in\mathscr{M}\cap[N/2,N]). \qquad (2)$$

Hence our aim is to prove that $J(M)$ cannot be small frequently which we do by considering it on average. We shall estimate the integrals

$$J(M)=\int_{\mathfrak{U}} f(\alpha)^4\,e(-N\alpha)\,d\alpha, \qquad I(M)=\int_{\mathfrak{U}} h(\alpha)^4\,e(-N\alpha)\,d\alpha$$

simultaneously. It will be seen that the contributions from the major arcs are almost equal apart from the factor $\mathfrak{G}(M)$ about which we have good information (cf. Lemmata 7 and 8) while the contributions from the minor arcs are small on average.

First, let $\alpha\in\mathfrak{M}(q,a)$. If we put $\beta=\alpha-a/q$ then Lemma 3 combined with the obvious estimates $z^2\leqslant y$, $|f(\alpha)|\leqslant N^{1/2}$, $|u(\alpha)|\leqslant 1$ implies that

$$\begin{aligned} f(\alpha)^4&=(q^{-1}S(q,a))^4\,f(\beta)^4+O(N^2y^{-1}z^2),\\ h(\alpha)^4&=\delta_q h(\beta)^4+O(N^2y^{-1}z^2).\end{aligned}$$

Therefore, using also Lemma 7,

$$\begin{aligned} J_{\mathfrak{M}}(M)&=\sum_{\substack{1\leqslant a\leqslant q\leqslant z\\(a,q)=1}}\int_{\mathfrak{M}(q,a)} f(\alpha)^4\,e(-M\alpha)\,d\alpha\\ &=\mathfrak{G}(M,z)\int_{-1/Q}^{1/Q} f(\beta)^4\,e(-M\beta)\,d\beta+O(Q^{-1}N^2y^{-1}z^4)\\ &=\mathfrak{G}(M)\int_{-1/Q}^{1/Q} f(\beta)^4\,e(-M\beta)\,d\beta+O\left(\frac{N}{z^{1/4}}\,w^{o(1)}\right).\end{aligned}$$

In order to estimate the contribution from the minor arcs on average we observe that

$$J_{\mathfrak{m}}(M)=\int_{\mathfrak{m}} f(\alpha)^4\, e(-M\alpha)\, d\alpha$$

is the Mth Fourier coefficient of the function defined on $\mathfrak{U}$ which agrees with $f(\alpha)^4$ on $\mathfrak{m}$ and vanishes elsewhere. Hence Bessel's inequality combined with Lemmata 4 and 5 yields that

$$\sum_{1\leqslant M\leqslant N} |J_{\mathfrak{m}}(M)|^2 \leqslant \int_{\mathfrak{m}} |f(\alpha)|^8\, d\alpha \leqslant \frac{N^3}{z^{1/12}}\, w^{o(1)}.$$

If we denote by $\mathscr{M}_1$ the set of those M for which $|J_{\mathfrak{m}}(M)| \geqslant N/z^{1/30}$ then, by the previous inequality,

$$|\mathscr{M}_1| \leqslant \frac{N}{z^{1/60}}\, w^{o(1)},$$

and for any $M\notin\mathscr{M}_1$ we have

$$J(M)=\mathfrak{S}(M)\int_{-1/Q}^{1/Q} f(\beta)^4\, e(-M\beta)\, d\beta+O\left(\frac{N}{z^{1/30}}\right).$$

We can deduce, by similar means, that

$$I(M)=\int_{-1/Q}^{1/Q} h(\beta)^4\, e(-M\beta)\, d\beta+O\left(\frac{N}{z^{1/30}}\right)$$

holds for all $1\leqslant M\leqslant N$ with the exception of a set $\mathscr{M}_2$ of cardinality

$$|\mathscr{M}_2| \leqslant \frac{N}{z^{1/60}}\, w^{o(1)}.$$

As in [3, p. 183] we have

$$\left|\int_{-1/Q}^{1/Q} f(\beta)^4\, e(-N\beta)\, d\beta-\int_{-1/Q}^{1/Q} h(\beta)^4\, e(-N\beta)\, d\beta\right| \ll \frac{N}{z^4},$$

hence an appeal to Lemma 7 shows that

$$|J(M)-\mathfrak{S}(M)\, I(M)| \leqslant \frac{N}{z^{1/30}}\, w^{o(1)} \qquad (M\notin\mathscr{M}_3), \tag{3}$$

where $\mathscr{M}_3=\mathscr{M}_1\cup\mathscr{M}_2$.

By means of Lemma 2 it is possible to estimate $I(M)$ for $N/2 \leqslant M \leqslant N$. First of all,

$$I(M)=\int_0^1 h(\alpha)^4\, e(-M\alpha)\, d\alpha = \sum_{n_1+n_2+n_3+n_4=M} h_{n_1}h_{n_2}h_{n_3}h_{n_4}$$

$$\geqslant \Big(\min_{N/12<n<3N/4} h_n\Big)^4 \sum_{\substack{n_1+n_2+n_3+n_4=M\\ N/12<n_1,\,n_2,\,n_3,\,n_4}} 1 \gg (N^{-1/2} w^{-1/27-o(1)})^4\, N^3,$$

whence

$$I(M) \geqslant N w^{-4/27-o(1)} \qquad (N/2 \leqslant M \leqslant N).$$

This inequality yields by (2) and Lemma 8 that

$$|\mathfrak{S}(M)\, I(M) - J(M)| \geqslant \frac{N}{w^{4/27+o(1)}} \qquad (M \in \mathcal{M} \cap [N/2, N]). \tag{4}$$

On noting that $z = w^6$ we see that (3) and (4) together imply that

$$\mathcal{M} \cap [N/2, N] \subseteq \mathcal{M}_3,$$

whence

$$|\mathcal{M} \cap [N/2, N]| \leqslant |\mathcal{M}_3| \leqslant |\mathcal{M}_1| + |\mathcal{M}_2| \leqslant \frac{N}{z^{1/60}}\, w^{o(1)} \leqslant \frac{N}{w^{1/11}}.$$

As

$$w^{1/11} > \exp((\log N \log\log N)^{1/2}/16),$$

we have proved (1).

REFERENCES

1. A. Balog and A. Sárközy, On sums of integers having small prime factors, I, *Studia Sci. Math. Hungar*. **19** (1984), 35–47.
2. J. Brüdern and E. Fouvry, Lagrange's four squares theorem with almost prime variables, *J. Reine Angew. Math*. **454** (1994), 59–96.
3. G. Harcos, Waring's problem with small prime factors, *Acta Arith*. **80** (1997), 165–185.
4. L.-K. Hua, "Additive Primzahltheorie," Teubner, Leipzig, 1959 [In German]; English translation, "Additive Theory of Prime Numbers," Amer. Math. Soc., Providence, 1965.
5. R. C. Vaughan, "The Hardy–Littlewood Method," 2nd ed., Cambridge Tracts in Math., Vol. 125, Cambridge Univ. Press, Cambridge, UK, 1997.