PROBABILITY THEORY

Lagrange's Theorem and Thin Subsequences of Squares

*Paul Erdös and Melvyn B. Nathanson**

ABSTRACT

Probabilistic methods are used to prove that for every $\varepsilon > 0$ there exists a sequence A_ε of squares such that every positive integer is the sum of at most four squares in A_ε and $A_\varepsilon(x) = O(x^{3/8+\varepsilon})$.

Key words and phrases: Sums of squares, additive bases, probabilistic methods in additive number theory.

The set A of positive integers is a *basis of order h* if every positive integer is the sum of at most h elements of A. Lagrange proved in 1770 that the set of squares is a basis of order 4. Let $A(x)$ denote the number of elements of

* The research of Melvyn B. Nathanson was supported in part by the National Science Foundation under Grant No. MCS78-07908.

ISBN 0-12-274460-8

the set A not exceeding x. The number of choices with repetitions of at most h elements in A not exceeding x is the binomial coefficient $(A(x) + h, h)$. If A is a basis of order h, then for x sufficiently large and $h \geq 2$

$$x < \binom{A(x) + h}{h} < A(x)^h$$

and so $A(x) > x^{1/h}$. In particular if $h = 4$, then $A(x) > x^{1/4}$. But the sequence of squares $A = \{n^2\}_{n=1}^{\infty}$ satisfies $A(x) \sim x^{1/2}$. It is a natural problem [7] to look for "thin" subsequences of the squares that are still bases of order 4. We shall prove (Theorem 1) that for every $\varepsilon > 0$ there exists a set A_ε of squares such that A_ε is a basis of order 4 and $A_\varepsilon(x) = O(x^{3/8+\varepsilon})$. We conjecture that for every $\varepsilon > 0$ there is a sequence A^* of squares such that A^* is a basis of order 4 and $A^*(x) = O(x^{1/4+\varepsilon})$.

Choi *et al.* [3] have improved Theorem 1 in the following finite case: For every $N > 1$ there is a finite set A of squares such that $|A| < (4/\log 2)N^{1/3} \log N$ and every nonnegative integer $n \leq N$ is the sum of four squares in A.

The proof of Theorem 1 uses the probabilistic method of Erdös and Rényi [4]. (The Halberstam-Roth book [6] contains an excellent exposition of this method.) Consider the following general situation. Let $F_j = F_j(x_1, x_2, \ldots, x_{h(j)})$ be a function in $h(j) \leq h$ variables, and let $\mathscr{F} = \{F_j\}_{j \in J}$. Let $A = \{a_n\}_{n=1}^{\infty}$ be a strictly increasing sequence of positive integers. Let $\mathscr{F}(A)$ denote the set consisting of all numbers of the form $F_j(a_{n_1}, a_{n_2}, \ldots, a_{n_{h(j)}})$, where $F_j \in \mathscr{F}$ and $a_{n_i} \in A$ for $i = 1, 2, \ldots, h(j)$. Let $s \in \mathscr{F}(A)$ and

$$s = F_j(a_{n_1}, a_{n_2}, \ldots, a_{n_{h(j)}}) = F_k(a_{m_1}, a_{m_2}, \ldots, a_{m_{h(k)}})$$

be two representations of s. These representations are *disjoint* if

$$\{a_{n_1}, a_{n_2}, \ldots, a_{n_{h(j)}}\} \cap \{a_{m_1}, a_{m_2}, \ldots, a_{m_{h(k)}}\} = \varnothing.$$

In Lemma 1 we apply probabilistic methods to show that if $S \subseteq \mathscr{F}(A)$ and each $s \in S$ has sufficiently many pairwise disjoint representations, then there is a "thin" subsequence A^* of A such that $S \subseteq \mathscr{F}(A^*)$. We also use this Lemma to obtain a best possible result for sums of three squares (Theorem 2) and to obtain a "thin" version of Chen's result on Goldbach's problem (Theorem 3).

LEMMA 1. *Let $A = \{a_n\}_{n=1}^{\infty}$ be a strictly increasing sequence of positive integers such that*

$$a_n \geq c_1 n^{\alpha} \tag{1}$$

for constants $c_1 > 0$, $\alpha \geq 1$, and all $n \geq 1$. Let $F_j = F_j(x_1, x_2, \ldots, x_{h(j)})$ be a function in $h(j) \leq h$ variables, and let $\mathscr{F} = \{F_j\}_{j \in J}$. Suppose there exist con-

stants $c_2 > 0$ *and* $\beta > 0$ *such that, if* $F_j \in \mathscr{F}$ *and* $F_j(x_1, x_2, \ldots, x_{h(j)}) = s$, *then*

$$x_i \le c_2 s^{\beta} \tag{2}$$

for $i = 1, 2, \ldots, h(j)$. *Let* $\mathscr{F}(A)$ *be the set consisting of all numbers of the form* $F_j(a_{n_1}, a_{n_2}, \ldots, a_{n_{h(j)}})$ *with* $F_j \in \mathscr{F}$, $a_{n_i} \in A$. *For* $s \in \mathscr{F}(A)$, *let* $R(s)$ *denote the maximum number of pairwise disjoint representations of* s *in the form* $s = F_j(a_{n_1}, a_{n_2}, \ldots, a_{n_{h(j)}})$. *Let* $S \subseteq \mathscr{F}(A)$. *Suppose there exist constants* $c_3 > 0$, $\gamma > 0$, *and* γ' *such that*

$$R(s) \ge c_3 s^{\gamma} / \log^{\gamma'} s \tag{3}$$

for all $s \in S$, $s > 1$. *Then for every* $\varepsilon > 0$ *there exist a constant* $c = c(\varepsilon) > 0$ *and a subsequence* A^* *of* A *such that* $S \subseteq \mathscr{F}(A^*)$ *and*

$$A^*(x) \le c x^{(1/\alpha - \gamma/\beta h + \varepsilon)}. \tag{4}$$

Proof. By the method of Erdös and Rényi [4, 6], every sequence of real numbers $p(1)$, $p(2), \ldots$ satisfying $0 \le p(n) \le 1$ determines a probability measure μ on the space Ω of all strictly increasing sequences of positive integers. The measure μ has the property that, if $B^{(n)}$ denotes the set of all sequences containing n, then $B^{(n)}$ is measurable and $\mu(B^{(n)}) = p(n)$. Moreover, the events $B^{(1)}, B^{(2)}, \ldots$ are independent. Let $0 < \varepsilon < \gamma/\beta h$. Then $\delta = (\alpha\gamma/\beta h) - \alpha\varepsilon > 0$. We consider the measure μ on Ω determined by the sequence of probabilities

$$p(n) = 1/n^{\delta} = 1/n^{(\alpha\gamma/\beta h) - \alpha\varepsilon}. \tag{5}$$

Each sequence $U = \{u(k)\}_{k=1}^{\infty} \in \Omega$ determines the subsequence $A^U = \{a_{u(k)}\}_{k=1}^{\infty}$ of A. This establishes a one-to-one correspondence between subsequences of A and sequences U in Ω. The probability that $a_n \in A^U$ is the same as the probability that $n \in U$, which is precisely $p(n) = n^{-\delta}$.

Let $s = F_j(a_{n_1}, a_{n_2}, \ldots, a_{n_{h(j)}}) \in S$. Inequalities (1) and (2) imply that

$$c_1 n_i^{\alpha} \le a_{n_i} \le c_2 s^{\beta}$$

for $i = 1, 2, \ldots, h(j)$, and so

$$n_i \le (c_2 s^{\beta}/c_1)^{1/\alpha} = c_4 s^{\beta/\alpha}.$$

The integers $n_1, n_2, \ldots, n_{h(j)}$ are not necessarily distinct. Let $m_1, m_2, \ldots, m_t$ be pairwise distinct integers such that $\{n_1, n_2, \ldots, n_{h(j)}\} = \{m_1, m_2, \ldots, m_t\}$. The probability that a subsequence $A^U = \{a_{u(k)}\}_{k=1}^{\infty}$ of A contains each of the numbers $a_{n_1}, a_{n_2}, \ldots, a_{n_{h(j)}}$ is the same probability that the sequence

$U = \{u(k)\}_{k=1}^{\infty} \in \Omega$ contains each of the numbers $n_i \in \{n_1, n_2, \ldots, n_{h(j)}\} = \{m_1, m_2, \ldots, m_t\}$. This probability is

$$\begin{aligned} p(m_1)p(m_2)\cdots p(m_t) &= \frac{1}{(m_1 m_2 \cdots m_t)^{\delta}} \\ &\geq \frac{1}{(n_1 n_2 \cdots n_{h(j)})^{\delta}} \\ &\geq \frac{1}{(c_4 s^{\beta/\alpha})^{\delta h(j)}} \\ &\geq \frac{c_5}{s^{\beta\delta h/\alpha}} \\ &= \frac{c_5}{s^{\gamma - \beta h\varepsilon}}. \end{aligned}$$

Therefore, the probability that the subsequence A^U does not contain at least one of the numbers $a_{n_1}, a_{n_2}, \ldots, a_{n_{h(j)}}$ is at most

$$1 - \frac{c_5}{s^{\gamma - \beta h\varepsilon}}.$$

There are $R(s)$ disjoint representations of $s \in S$. By (3), the probability that A^U does not contain at least one term from each of these $R(s)$ representations of s is at most

$$(1 - c_5/s^{\gamma - \beta h\varepsilon})^{R(s)} \leq (1 - (c_5/s^{\gamma - \beta h\varepsilon})^{c_3 s^{\gamma}/\log^{\gamma'} s}.$$

The corresponding series of probabilities converges:

$$\sum_{s=2}^{\infty} (1 - c_5/s^{\gamma - \beta h\varepsilon})^{c_3 s^{\gamma}/\log^{\gamma'} s} < \infty.$$

The Borel–Cantelli lemma implies that for almost all sequences $U \in \Omega$, the subsequence A^U of A represents all but finitely many $s \in S$. Adjoining a finite set to A^U, we obtain a subsequence A^* of A such that $S \subseteq F(A^*)$. The law of large numbers implies that for almost all $U \in \Omega$,

$$U(x) \sim c_6 x^{1-\delta} = c_6 x^{1 - \alpha\gamma/\beta h + \alpha\varepsilon}.$$

Since $a_n \geq c_1 n^{\alpha}$ by (1), it follows that

$$A^U(x) \leq U((x/c_1)^{1/\alpha}) \leq c x^{1/\alpha - \gamma/\beta h + \varepsilon}.$$

This completes the proof of Lemma 1.

LEMMA 2. *Let* $S = \{n \geq 1 \mid n \not\equiv 0 \ (mod\ 4)\}$. *Let* $R(s)$ *denote the maximum number of pairwise disjoint representations of* s *as the sum of at most four* *squares. Then for every $\varepsilon > 0$ there is a constant $c = c(\varepsilon) > 0$ such that*

$$R(s) > cs^{1/2-\varepsilon}$$

for all $s \in S$.

Proof. Let $r_k(s)$ denote the number of representations of s as the sum of at most k squares. It is well known that $r_2(s) \leq c_1 s^{\varepsilon}$ for every $\varepsilon > 0$. This implies that $r_3(s) \leq c_2 s^{1/2+\varepsilon}$, since if $s = a^2 + b^2 + c^2$, there are at most $s^{1/2}$ choices for a and, for each a, we have $r_2(s - a^2) \leq c_1 s^{\varepsilon}$ choices of b and c.

Let $s = a_1^2 + a_2^2 + a_3^2 + a_4^2$. The number of representations of s as a sum of at most four squares that include the number a_i is $r_3(s - a_i^2)$. It follows that the number of representations of s that include at least one of the numbers a_1, a_2, a_3, a_4 is at most

$$\sum_{i=1}^{4} r_3(s - a_i^2) \leq c_3 s^{1/2+\varepsilon}.$$

There are $R(s)$ disjoint representations of s as the sum of four squares, and so there are at most

$$c_3 s^{1/2+\varepsilon} R(s)$$

representations of s as a sum of four squares. But Jacobi's theorem on the number of representations of an integer as the sum of four squares implies that each $s \in S$ has at least $c_4 s$ such representations. Therefore,

$$c_4 s \leq c_3 s^{1/2+\varepsilon} R(s).$$

This completes the proof of Lemma 2.

THEOREM 1. *For every $\varepsilon > 0$ there exists a sequence A_ε of squares such that every positive integer is the sum of at most four squares in A_ε and $A_\varepsilon(x) \leq cx^{3/8+\varepsilon}$ for some $c = c(\varepsilon) > 0$.*

Proof. Let $A = \{n^2\}_{n=1}^{\infty}$. Let $F_j = F_j(x_1, \ldots, x_j) = x_1 + \cdots + x_j$, let $J = \{1, 2, 3, 4\}$, and let $\mathscr{F} = \{F_j\}_{j \in J}$. Lagrange's theorem asserts that $\mathscr{F}(A) = \{1, 2, 3, \ldots\}$. Let $S = \{s \geq 1 \mid s \not\equiv 0 \pmod 4\}$. We apply Lemma 1 with $\alpha = 2$, $\beta = 1$, $h = 4$, and, by Lemma 2, with $\gamma = \frac{1}{2} - \varepsilon$. Then there is a sequence A^* of squares such that each $s \in S$ is a sum of four squares in A^* and

$$A^*(x) \leq cx^{1/2-[1/2-\varepsilon]/4+\varepsilon} = cx^{3/8+5\varepsilon/4}.$$

Let $A_\varepsilon = \{2^k a \mid a \in A^*, k \geq 0\}$. Let $n \geq 1$. Then $n = 4^k s$ for some $s \in S$. There exist $j \in J$ and $a_1, \ldots, a_j \in A^*$ such that $s = \sum_{i=1}^{j} a_i^2$. Then $2^k a_i \in A_\varepsilon$ and $\sum_{i=1}^{j} (2^k a_i)^2 = 4^k \sum_{i=1}^{j} a_i^2 = 4^k s = n$. Therefore, each $n \geq 1$ is a sum of at most four squares in A_ε. Moreover, if $2^k a \leq x$, then $k \leq \log x/\log 2$ and so

$$A_\varepsilon(x) \leq \left(1 + \frac{\log x}{\log 2}\right) A^*(x) \leq \left(1 + \frac{\log x}{\log 2}\right) cx^{3/8 + 5\varepsilon/4}$$
$$\leq cx^{3/8 + 2\varepsilon}.$$

Replacing ε by $\varepsilon/2$ completes the proof of Theorem 1.

THEOREM 2. *For every $\varepsilon > 0$ there exists a sequence B_ε of squares such that every positive integer $n \neq 4^k(8m + 7)$ is the sum of at most three squares in B_ε and*

$$B_\varepsilon(x) \leq cx^{1/3+\varepsilon} \qquad \text{for some} \quad c = c(\varepsilon) > 0.$$

Proof. Let $A = \{n^2\}_{n=1}^{\infty}$. Let $F_j = x_1 + \cdots + x_j$ for $j \in J = \{1, 2, 3\}$, and let $\mathscr{F} = \{F_j\}_{j \in J}$. Gauss showed that $\mathscr{F}(A)$ consists of all positive integers not of the form $4^k(8m + 7)$. Let $S = \{s \geq 1 \mid s \not\equiv 0, 4, 7 \pmod 8\}$. Then $S \subseteq \mathscr{F}(A)$. Siegel [8] and Bateman [1] showed that for every $\varepsilon > 0$ and $s \in S$ there are at least $c_1 s^{1/2-\varepsilon}$ representations of s as a sum of three squares. The argument used to prove Lemma 2 shows that if $s \in S$, then s has at least $c_2 s^{1/2-\varepsilon}$ pairwise disjoint representations as a sum of three squares. We apply Lemma 1 with $\alpha = 2$, $\beta = 1$, $h = 3$, and $\gamma = \frac{1}{2} - \varepsilon$. This yields a subsequence $A^* \subseteq A$ such that $S \subseteq \mathscr{F}(A^*)$ and

$$A^*(x) \leq cx^{1/2 - [1/2 - \varepsilon]/3 + \varepsilon} = cx^{1/3 + 4\varepsilon/3}.$$

If $n \in \mathscr{F}(A)$, then $n = 4^k s$ for some $k \geq 0$ and $s \in S$. Let $B_\varepsilon = \{2^k a \mid k \geq 0, a \in A^*\}$. Then $\mathscr{F}(B_\varepsilon) = \mathscr{F}(A) = \{a^2 + b^2 + c^2 \mid a, b, c, \geq 0\}$ and $B_\varepsilon(x) \leq c'(\log x)A^*(x) \leq cx^{(1/3)+2\varepsilon}$. This completes the proof of Theorem 2.

THEOREM 3. *Let C consist of all numbers of the form p or pq, where p, q are odd primes. Then for every $\varepsilon > 0$ there is a set $C_\varepsilon \subseteq C$ such that every sufficiently large even integer is the sum of two elements of C_ε and*

$$C_\varepsilon(x) \leq cx^{1/2+\varepsilon}.$$

Proof. Chen [2, 5] proved that every even number $n \geq n_0$ has at least $c_1 n/\log^2 n$ representations as the sum of two elements of C. These representations are pairwise disjoint. Apply Lemma 1 with $\alpha = 1$, $\beta = 1$, $h = 2$, and $\gamma = 1$. This yields a sequence $C_\varepsilon \subseteq C$ such that every even number $n \geq n_0$ is the sum of two elements of C_ε and $C_\varepsilon(x) \leq cx^{1/2+\varepsilon}$. This completes the proof of Theorem 3.

References

[1] P. T. Bateman, On the representations of a number as the sum of three squares. *Trans. Amer. Math. Soc.* **71** 70–101, (1951).

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

[3] S. L. G. Choi, P. Erdös, and M. B. Nathanson, Lagrange's theorem with $N^{1/3}$ squares. *Proc. Amer. Math. Soc. 79* 203–205 (1980).

[4] P. Erdös and A. Rényi, Additive properties of random sequences of positive integers. *Acta Arith.* **6** 83–110, (1960).

[5] H. Halberstam and H. -E. Richert, "Sieve Methods". Academic Press, New York, 1974.

[6] H. Halberstam and K. F. Roth, "Sequences", Vol. I. Oxford Univ. Press (Clarendon), London and New York, 1966.

[7] E. Härtter and J. Zöllner, Darstellungen natürlichen Zahlen als Summe und als Differenz von Quadraten. *K. Norske Vidensk. Selk. Skr.* no. 1, 1–8, (1977).

[8] C. L. Siegel, Uber die Klassenzahl quadratischer Zahlkörper. *Acta. Arith.* **1** 83–86, (1935).

Paul Erdös
Mathematical Institute of
the Hungarian Academy of
Sciences
Budapest V., Realtanoda
Hungary

Melvyn B. Nathanson
Department of Mathematics
Southern Illinois University
Carbondale, Illinois

1980 Mathematics Subject
Classification. Primary 10J05.
Secondary 10L05, 10K99.