ACTA ARITHMETICA
LXI.3 (1992)

$B_2[\infty]$-sequences of square numbers

by

Javier Cilleruelo and Antonio Córdoba (Madrid)

Let us denote by $r(n)$ the number of representations of the integer n as a sum of two squares. It is an outstanding arithmetical function whose study is a familiar topic in Number Theory. This function has a rather "irregular" behavior because the value $r(n)$ depends upon the decomposition of n in prime factors.

On the other hand, it is well known that its average satisfies $\sum_{n\le x} r(n) \sim \pi x$ as x tends to infinity. A particularly interesting expression of this "irregularity" is given by the fact that the order of magnitude of $\sum_{n\le x} r^2(n)$ is strictly greater than $\sum_{n\le x} r(n)$. More precisely,

$$\frac{\sum_{n\le x} r^2(n)}{\sum_{n\le x} r(n)} \gg \log x\,.$$

It is then natural to consider the following class of sequences.

Definition. We say that a sequence of positive integers $\{n_k\}$ *has the property* $B_2[\infty]$, or that it is a $B_2[\infty]$-*sequence*, if

$$\limsup_{x\to\infty} \frac{\sum_{n\le x} r^2(n)}{\sum_{n\le x} r(n)} < +\infty\,,$$

where $r(n) = \#\{n = n_k + n_j, n_k \le n_j\}$.

This definition is also a natural extension of the concept of $B_2[g]$-sequences, i.e. those for which $r(n) \le g$ for every n. We refer to [5] for details and applications.

As we have seen in our previous discussion, the whole sequence of squares is not a $B_2[\infty]$-sequence. It seems that J. E. Littlewood was the first to observe that this fact is "atypical" among sequences with growth similar to the square numbers. In his Ph.D. thesis, A. O. L. Atkin [1] confirmed Littlewood's prediction proving the existence of a $B_2[\infty]$-sequence very close to the squares ($n_k = k^2 + O(\log k)$). He also observed that no sequence $\{n_k\}$ such that $n_k = k^2 + o(\sqrt{\log k})$ has the property $B_2[\infty]$.

Later on, P. Erdős and A. Rényi [4], using probabilistic methods, were able to show that, in a certain sense, almost every sequence close to the squares ($n_k = k^2(1 + o(1))$) is $B_2[\infty]$. In both [1] and [4], the sequences considered keep the growth of the squares but their terms are not square numbers. In this paper we shall proceed in the opposite direction and we construct subsequences of the squares with the required property $B_2[\infty]$ and with a moderate rate of growth.

This paper is closely related to [3], where the existence of a B_2-sequence $\{a_k^2\}$ such that $a_k \ll k^2$ is proved.

THEOREM. *There exists an infinite sequence*, $a_1^2 < a_2^2 < \ldots < a_k^2 < \ldots$, *of square numbers satisfying property* $B_2[\infty]$ *and such that* $a_k^2 \ll k^2 \log^2 k$.

Proof. We shall construct explicitly a family of such sequences. Let

$$I_j^\nu = \{2n \mid 2^j(1+\nu/j) \le 2n < 2^j(1+(\nu+1)/j)\}, \qquad j = 1, \ldots,\ 0 \le \nu \le j-1.$$

With a choice of ν_j, for each j, consider the sets

$$I^{(\nu)} = \bigcup_j I_j^{\nu_j}\,, \qquad A^{(\nu)} = \{a_k^2 \mid a_k \in I^{(\nu)}\}\,.$$

We claim that $A^{(\nu)}$ satisfies the conditions required by the Theorem. (To simplify notation we shall consider the case $\nu_j = 0$. The proof for general values of ν_j, $j = 1, 2, \ldots$, is identical.)

Clearly we have $a_{[2^{j-1}/j]} \le 2^j + 2^j/j \le 2^{j+1}$, therefore, given k, there exists j such that $2^{j-1}/j \le k < 2^j/(j+1)$. Then $a_k \le a_{[2^j/(j+1)]} < 2^{j+2} \ll k \log k$.

Let us consider

$$r(n) = \#\Big\{n = a_k^2 + a_l^2 \,\Big|\, a_k \le a_l\,,\ a_k, a_l \in I = \bigcup_j I_j^0\Big\}\,.$$

We have the following inequality:

$$\text{(I)} \qquad \sum_{n \le x} r(n) = \sum_{\substack{a_k^2 + a_l^2 \le x \\ a_k \le a_l}} 1 \ge \frac{1}{2}\Big(\sum_{a_k^2 \le x/2} 1\Big)^2 \gg x/\log^2 x\,.$$

Therefore to prove the Theorem it is enough to show the following estimate:

$$\text{(II)} \qquad \sum_{n \le x} r^2(n) \ll x/\log^2 x\,.$$

To see this, let us define, for every j, the function

$$r_j(n) = \#\{n = a_k^2 + a_l^2 \mid a_k \le a_l\,,\ a_k, a_l \in I_j^0\}\,.$$

Then

$$\text{(III)} \qquad \sum_{n\le x} r^2(n) \ll \Big[\sum_{j\le[\log_4 x]} \Big(\sum_n r_j^2(n)\Big)^{1/2}\Big]^2 .$$

This is a consequence of the Littlewood–Paley theory of Fourier series, together with the natural interpretation of the sums $\sum_{n\le x} r^2(n)$ as L^4-norms of the corresponding trigonometric polynomials:

$$\begin{aligned}
\sum_{n\le x} r^2(n) &\simeq \Big\| \sum_{j\le(\log_2 x)/2} \sum_{n\in I_j^0} e^{2\pi i n^2 x}\Big\|_4^4 \\
&\simeq \Big\| \Big(\sum_{j\le(\log_2 x)/2} \Big| \sum_{n\in I_j^0} e^{2\pi i n^2 x}\Big|^2\Big)^{1/2}\Big\|_4^4 \\
&\ll \Big(\sum_{j\le(\log_2 x)/2} \Big\| \sum_{n\in I_j^0} e^{2\pi i n^2 x}\Big\|_4^2\Big)^2 \\
&\simeq \Big(\sum_{j\le(\log_2 x)/2} \Big(\sum_n r_j^2(n)\Big)^{1/2}\Big)^2 .
\end{aligned}$$

We use the standard notation $A(z) \simeq B(z)$ with the following meaning: there exist constants $0 < C_1 \le C_2 < \infty$ such that $C_1 A(z) \le B(z) \le C_2 A(z)$ for every z.

We claim that

$$\text{(IV)} \qquad \sum_n r_j^2(n) \ll (2^j/j)^2 .$$

Then we obtain

$$\text{(V)} \qquad \sum_{n\le x} r^2(n) \ll \Big(\sum_{j\le(\log_2 x)/2} 2^j/j\Big)^2 \ll x/\log^2 x .$$

To finish the proof we have to show estimate (IV). We shall do this in two steps: first we state a lemma.

LEMMA. *If $r_j(n) \ge 2$, then $r_j^2(n) \ll F(n)$, where $F(n)$ is the number of integer solutions (a_1, b_1, a_2, b_2) of the following system of inequalities*:

$$0 < |a_1/b_1| < 1/j , \qquad 0 < |a_2/b_2 - 1| < 1/j .$$

Assuming this lemma we can complete the proof of (IV). With j fixed we have

$$\sum_n r_j^2(n) \le \sum_{r_j(n)=1} r_j(n) + \sum_{r_j(n)\ge 2} r_j^2(n) .$$

It is clear that the first term satisfies the estimate $\sum r_j(n) \ll (2^j/j)^2$.

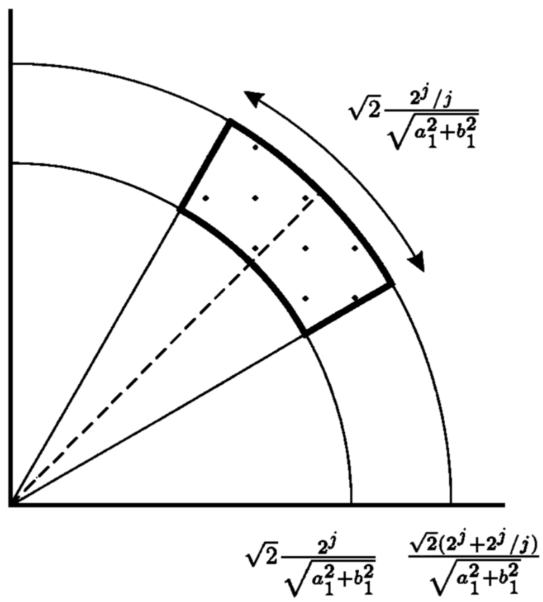
By the Lemma, the second term is bounded by the number of solutions (a_1, b_1, a_2, b_2) of the following inequalities:

$$2(2^j)^2 \le (a_1^2 + b_1^2)(a_2^2 + b_2^2) \le 2(2^j + 2^j/j)^2 ,$$
$$0 < |a_1/b_1| < 1/j , \qquad 0 < |a_2/b_2 - 1| < 1/j .$$

Keeping a_1, b_1 fixed we shall estimate the number of pairs (a_2, b_2) satisfying

$$\sqrt{2}\frac{2^j}{\sqrt{a_1^2 + b_1^2}} \le \sqrt{a_2^2 + b_2^2} \le \frac{\sqrt{2}(2^j + 2^j/j)}{\sqrt{a_1^2 + b_1^2}} , \qquad 0 < \left|\frac{a_2}{b_2} - 1\right| < \frac{1}{j} .$$

One can identify each solution (a_2, b_2) with a lattice point in the region described in the figure.



The condition $0 < |a_2/b_2 - 1| < 1/j$ yields $b_2 > j$ and, therefore,

$$\sqrt{2}\frac{2^j/j}{\sqrt{a_1^2 + b_1^2}} \gg 1 ,$$

which allows the number of lattice points inside that region to be estimated by its area. Thus

$$\sum_{r_j(n)\ge 2} r_j^2(n) \ll \sum_{\substack{a_1, b_1 \\ 0<|a_1/b_1|<1/j}} \frac{4^j}{(a_1^2 + b_1^2) j^2} \ll \sum_{b_1 < 2^j} \frac{4^j}{b_1 j^3} \ll \frac{4^j}{j^2} ,$$

which easily yields inequality (IV).

Proof of the Lemma. First we observe (see [3] for details) that if we have two different decompositions of $n \equiv 0 \pmod 4$ as a sum of two squares, $n = a^2 + b^2 = c^2 + d^2$, then there exist integers a_1, b_1, a_2, b_2 such that

$$n = (a_1^2 + b_1^2)(a_2^2 + b_2^2)\,,$$
$$\frac{a_1}{b_1} = \tan\left(\frac{\arctan(a/b) + \arctan(c/d)}{2}\right),$$
$$\frac{a_2}{b_2} = \tan\left(\frac{\arctan(a/b) - \arctan(c/d)}{2}\right).$$

Let $n = a_r^2 + b_r^2 = a_s^2 + b_s^2$ be two different representations of n as the sum of two squares satisfying

$$2^j \le a_r \le b_r \le 2^j + 2^j/j\,,$$
$$2^j \le a_s \le b_s \le 2^j + 2^j/j\,,$$
$$a_r \equiv b_r \equiv a_s \equiv b_s \equiv 0 \pmod 2\,.$$

By our previous argument there exist a_1, b_1, a_2, b_2 such that

$$n = (a_1^2 + b_1^2)(a_2^2 + b_2^2)\,,$$
$$\frac{a_1}{b_1} = \tan\left(\frac{\arctan(a_r/b_r) + \arctan(a_s/b_s)}{2}\right),$$
$$\frac{a_2}{b_2} = \tan\left(\frac{\arctan(a_r/b_r) - \arctan(a_s/b_s)}{2}\right).$$

The conditions imposed on a_r, b_r, a_s, b_s easily yield that

$$|a_2/b_2| < 1/j\,, \qquad |a_1/b_1 - 1| < 1/j\,.$$

On the other hand, both quantities are strictly greater than 0 because otherwise we could not have started with two different representations of n as a sum of two squares.

Each pair of different representations

$$\{n = a_r^2 + b_r^2 = a_s^2 + b_s^2,\ a_s \equiv b_s \equiv a_r \equiv b_r \equiv 0 \,(\mathrm{mod}\, 2)\}$$

produces different angles and, consequently, different integers a_1, b_1, a_2, b_2. Therefore $\binom{r_j(n)}{2}$ is bounded above by the number of solutions of the problem

$$n = (a_1^2 + b_1^2)(a_2^2 + b_2^2)\,,$$
$$0 < |a_1/b_1| < 1/j\,, \qquad 0 < |a_2/b_2 - 1| < 1/j\,.$$

The Lemma follows by observing that if $r_j(n) \ge 2$ then $r_j^2(n) \le 4\binom{r_j(n)}{2}$.

Remarks. It is not difficult to see that no subsequence of the squares $\{n_k^2\}$ such that $n_k^2 = o(k^2 \log k)$ can have the property $B_2[\infty]$.

Furthermore, one can build sequences of square numbers which grow very fast and which, nevertheless, are not $B_2[\infty]$-sequences.

In the proof of the Theorem we have used some of the methods introduced in [2] and [3] and we have obtained the following estimate:

$$\Big\|\sum_{N\le n\le N+N/\log N} e^{2\pi i n^2 x}\Big\|_4 \simeq \frac{N^{1/2}}{(\log N)^{1/2}}\,,$$

uniformly in N, which improves the results contained in [2]. More precisely, we have obtained

$$\Big\|\sum_{N\le n\le N+N/M} e^{2\pi i n^2 x}\Big\|_4 = 2^{1/4}\frac{N^{1/2}}{M^{1/2}}\left(1+O\left(\left(\frac{\log N}{M}\right)^{1/4}\right)\right)$$

for every $M\ge 1$.

References

[1] A. O. L. Atkin, *On pseudo-squares*, Proc. London Math. Soc. (3) 14A (1965), 22–27.

[2] J. Cilleruelo and A. Córdoba, *Trigonometric polynomials and lattice points*, Proc. Amer. Math. Soc., to appear.

[3] J. Cilleruelo, *B_2-sequences whose terms are squares*, Acta Arith. 55 (1990), 261–265.

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

[5] H. Halberstam and K. F. Roth, *Sequences*, Clarendon Press, Oxford 1966.

DEPARTAMENTO DE MATEMÁTICAS
UNIVERSIDAD AUTÓNOMA DE MADRID
28049 MADRID, ESPAÑA

Received on 8.10.1990
and in revised form on 5.7.1991 (2087)