DIMACS Technical Report 98-34
July 1998

On the dimension of the Hilbert-cubes

by

Norbert Hegyvári[1]

[1]Research partially supported by Hungarian National Foundation for Scientific Research, Grant No. T025617 and by DIMACS (Center for Discrete Mathematics and Theoretical Computer Science) NSF-STC-91-19999.

DIMACS is a partnership of Rutgers University, Princeton University, AT&T Labs-Research, Bell Labs, Bellcore and NEC Research Institute.

DIMACS is an NSF Science and Technology Center, funded under contract STC–91–19999; and also receives support from the New Jersey Commission on Science and Technology.

ABSTRACT

Let A be a sequence of positive integers with positive density. Then $A \cap \{1, 2, \ldots, n\}$ contains a Hilbert (or combinatorial) cube of dimension $c \log \log n$. We prove that this bound can not be replaced by $c' \sqrt{\log n \log \log n}$.

1. Introduction

In [1] D. Hilbert showed (using different terminology) that for any $k \geq 1$, if $\mathbf{N}$ is finitely colored than there exists in one color infinitely many transletes of a k-cube. We call $H \subset \mathbf{N}$ a k-cube if there exist $a > 0$ and $x_1, x_2, \ldots, x_k$

$$H = H(a, x_1, \ldots, x_k) = \left\{ a + \sum_{i=1}^{k} \varepsilon_i a_i : \quad \varepsilon_i = 0 \text{ or } 1 \right\}.$$

This result was essentially the first Ramsey-type theorem. The density version of the Hilbert result is the following:

Theorem A: Let $k \geq 1$ be a integer and assume $A \subset [1, n]$ and

$$|A| \geq n^{1-2^{-k}},$$

where $\varepsilon \geq 2$. Then A contains a k-cube. (see [2],[3]).

This result implies the following

Corollary: Let A be an infinite sequence of integers with

$$d(A) = \lim_{n\to\infty} \inf A(n)/n > 0,$$

where $A(n) = \sum_{a_i \leq n} 1$ Then there exists a β so that for every n, $A \cap [1, n]$ contains a k-cube where $k > \beta \log\log n$.

Proof of the Corollary: Indeed there exists a $c > 0$ such that for every n, $A(n) > cn$. Now if $\alpha n^{1-2^{-k}}$ or equivalently if $k \gg \log\log n$ then by Theorem A, A contains a k-cube as we stated.

Definition: Let A be an infinite sequence of integers. Let

$$H(n) = \max\{k : A \cap [1, n] \text{ contains a k-cube}\}$$

So by the corollary we have if $\underline{d}(A)>0$ then $H(n) \gg \log\log n$.

The aim of this note to prove

Theorem: There exists an infinite sequence A of positive integers with $\underline{d}(A)>0$ and

$$H(n) \leq c\sqrt{\log n \log\log n}$$

where $c = 6 \cdot (\log(5/4))^{-1/2}$.

The proof of the theorem is not constructive. A Sárközy and the author investigated the dimension of the Hilbert-cube some "conventional" sequence as well. For instance we prove if Q is the sequence of squares then $H_Q < c' \sqrt[3]{\log n}$ and for the sequence of primes P we get $c_1 \cdot \log\log n < H_P(n) < c_2 \cdot \log n$ (see [6]). Finally we mention a related question of P. Erdős which solved by E. Straus (see [5]): is it true if A is an infinite sequence of integer, A has positive upper density then A contains an infinite Hilbert-cube. The answer is negative.

For sequences A, B of integers let $A + B = \{a + b : a \in A, b \in B\}$ and $D(A) = \{a - a' : a > a' \quad a, a' \in A\}$

2. The proof of the theorem

For the proof of the theorem we need some lemmas.

Lemma 1: Let $B = \{b_1 < b_2 < \ldots < b_k\}$ be a sequence of integers. Then

$$\binom{k+1}{2} + 1 \leq |\Sigma\{A\}| \leq \ 2^k.$$

where $\Sigma\{B\} = \left\{x : x = \sum_{i=1}^{k} \epsilon_i \ b_i : \ \epsilon_i = 0 \ \text{ or } \ 1, \quad b_i \in B\right\}$.

The proof is easy or see [4].

Lemma 2: We have

$$T = |\{A : A \subseteq \ [1, n]; |A| = k \text{ and} |\Sigma(A)| < \ k^3\}| < \ (kn)^{5 \log_2 k} \cdot 4^{k^2}$$

Proof of Lemma 2: Let U be the event $U = \{A : A \subseteq \ [1, n]; |A| = k \text{ and} |\Sigma(A)| < \ k^3\}$. Clearly

$$Pr(U) = \frac{T}{\binom{n}{k}} \tag{1}$$

where $Pr(U)$ is the probability of the event U.

Now we are going to give an upper bound for $Pr(U)$. Let $A = \{a_1 < \ldots < a_k\} \subset [1, n]$ and let B the event that there are at least $[4 \log_2 k]$ indeces j for which

$$\Sigma\{a_1, a_2, \ldots, a_{j-1}\} \cap \Sigma\{a_1, a_2, \ldots, a_j\} = \{\emptyset\}. \tag{2}$$

In this case

$$|\Sigma\{a_1, a_2, \ldots, a_j\}| = 2|\Sigma\{a_1, a_2, \ldots, a_{j-1}\}| \tag{3}$$

using the fact

$$\Sigma\{a_1, a_2, \ldots, a_j\} = \{a_j\} + \Sigma\{a_1, a_2, \ldots, a_j\}.$$

So by (3) we have

$$|\Sigma\{a_1, a_2, \ldots, a_k\}| \geq 2^{[4 \log_2 k]} > k^3.$$

This yields $U \subset \overline{B}$, so it is enough to find an upper bound for $Pr(\overline{B})$. $\overline{B}$ holds if there are at most $[4\log_2 k]$ indeces j for which (1) holds. If j is an index for which (1) does not hold then
$$a_j \in D(\Sigma\{a_1, \dots, a_{j-1}\}. \tag{4}$$

Let C_j be the event fo which (4) holds. Then by Lemma 1

$$Pr(C_j) \leq \frac{|\Sigma\{a_1, \cdots, a_{j-1}\}|}{n-(j-1)} \leq \frac{(2^j)^2}{n-(j-1)}.$$

Let $R = [4\log_2 k]$. Thus

$$Pr(\overline{B}) \leq \sum_{1\leq t_1,\dots,t_s\leq k,\ s\leq R} \frac{(4^k)^{k-s}(n-t_1)\dots(n-t_s)}{n(n-1)\cdots(n-k+1)}$$

$$\leq \sum_{s\leq R} \binom{k}{s} \frac{4^{k^2} n^R}{n\cdots(n-k+1)} \leq R\frac{(kn)^R 4^{k^2}}{n\cdots(n-k+1)}.$$

This yields
$$T \leq \binom{n}{k} Pr(\overline{B}) \leq (kn)^{R+1} 4^{k^2} \leq (kn)^{5\log_2 k} 4^{k^2}$$

which proves Lemma 2.

Now we turn to the proof of the Theorem.

Proof of the Theorem: Let X be a random sequence of integers with $Pr(x \in X) = \frac{1}{25}$. Clearly with probability 1 we have $\underline{d}(X) > 0$.

Let H_n be the event $H_X(n) > c\sqrt{\log n \log\log n}$ where $c = 6 \cdot (\log 5/4)^{-1/2}$. We are going to show
$$Pr(H_n) < \frac{1}{n^2} \tag{5}$$

We have
$$Pr(H_n) \leq \sum_{\substack{1\leq a\leq n \\ 1\leq x_1,\dots,x_k\leq n}} \left(\frac{1}{25}\right)^{|\sum(x_1,\dots,x_k)|} =$$

$$\sum_{\substack{1\leq a\leq n \\ 1\leq x_1,\dots,x_k\leq n, \\ |\sum(x_1,\dots,x_k)|<k^3}} \left(\frac{1}{25}\right)^{|\sum(x_1,\dots,x_k)|} + \sum_{\substack{1\leq a\leq n \\ 1\leq x_1,\dots,x_k\leq n, \\ |\sum(x_1,\dots,x_k)|\geq k^3}} \left(\frac{1}{25}\right)^{|\sum(x_1,\dots,x_k)|} \tag{6}$$

By Lemma 1 and 2 we have

$$\sum_{\substack{1\le a\le n\\ 1\le x_1,\dots,x_k\le n,\\ |\sum(x_1,\dots,x_k)|<k^3}} \left(\frac{1}{25}\right)^{|\sum(x_1,\dots,x_k)|} \le \sum_{1\le a\le n} (kn)^{5\log_2 k}\cdot 4^{k^2}\cdot\left(\frac{1}{25}\right)^{k^2/2} = n\cdot (kn)^{5\log_2 K}\cdot\left(\frac{4}{5}\right)^{k^2}$$

which is less than $1/2n^2$ if $k\ge 6\cdot\log(5/4)^{-1/2}\sqrt{\log n\log\log n}$. Furthermore

$$\sum_{\substack{1\le a\le n\\ 1\le x_1,\dots,x_k\le n,\\ |\sum(x_1,\dots,x_k)|\ge k^3}} \left(\frac{1}{25}\right)^{|\sum(x_1,\dots,x_k)|} \le n\cdot\binom{n}{k}\cdot\left(\frac{1}{25}\right)^{k^3} < \frac{1}{2n^2}$$

holds if $k>3\sqrt{\log_2 n}$.

By (5) we have $\sum_{n=1}^{\infty} Pr(H_n)<\infty$ so that by the Borel-Cantelli lemma with probability 1, at most finite number of events H_n occur.

This completes the theorem.

Remarks: 1. We split the sum in (6) into two parts according the value of $|\sum(x_1,\dots,x_k)|$. We mention here for the sets $A_d=\{d,2d,\dots,kd\}$, $d=1,2,\dots,\left[\frac{n}{k}\right]$ we have

$$|\sum(A_d)|=\binom{k+1}{2}\sim k^2.$$

So we have to count these sets in the first sum which yields that our method works only if $k>c\sqrt{\log n}$.

2. There are many pairs A and A' for which $\sum(A)\cap\sum(A')$ contains a fix "big" sets, that is why we use only the trivial inequality $Pr(\cup A)\le\sum Pr(A)$.

3. Concluding remarks

In this section we are going to show that for *random* sequence our bound appart from the factor $\sqrt{\log\log n}$ is the best possible.

We prove

Proposition: Let A be a random sequence of positive integers with $Pr(a\in A)=p$. Then with probability 1, we have

$$H_A(n)>c_p\sqrt{\log n}.$$

Proof of the Proposition: Let $0 < p < 1$ be a fixed real number and let A be a random sequence of integers with $Pr(a \in A) = p$ and let $k_n = max_{a,k}\{k : a+1, \dots a+k \in A\}$. By a theorem of Erdös and Rényi [7], with probability 1 $k_n = c(p)\log n$. But let us observe that if $a, a+1, \dots . a+k \in A$ then $H(a, 1, 2, \dots, [\sqrt{2k} - 1]) \subset A$. (Indeed $H(a, 1, 2, \dots, [\sqrt{2k} - 1]) \supseteq [a, a+1, \dots, k]$.) This yields that with probability 1 we have $H_A(n) > c_p\sqrt{\log n}$, where $c_p = \sqrt{2c(p)}$.

References

[1] D. Hilbert: Über die Irreducibilität rationaler Functionen mit ganzzäligen Koefficienten, J. Reine Angew. Math. 110 (1892), 104–109.

[2] P. Erdős, A. Sárközy, V.T. Sós: On a conjecture of Roth and related problems I, Algorithms and Combinatorics 8, (Ed. R.L. Graham, B. Körte, Lovász), 47–59.

[3] R. L. Graham, B. Rothschild, J. Spencer: Ramsey Theory, Wiley Interscience Publications, 1980.

[4] N. Hegyvári: On representation problems in the additive number theory, Acta Math. Hung. 72 (1-2) 1996, 35–44.

[5] N. Hindman: Ultrafilters and combinatorial number theory, In: M. Nathanson (ed.) Number Theory Carbonadle. Lecture Notes in Math. 751 119–184.

[6] N. Hegyvári, A. Sárközy: On the siye of a combinatorial cube of certain sets (preprint).

[7] P. Erdős and A. Rényi: On a new law of large numbers, Journ. Analyse Mathématique 22 (1970) 103–111.

Norbert Hegyvári
Department of Mathematics,
ELTE TFK, Eötvös University,
Budapest, Markó u. 29.
Hungary, H-1055
email: NORB@ludens.elte.hu