Hilbert cubes in multiplicatively defined sets

Christian Elsholtz,
TU Graz
Austria

joint work with R. Dietmann
1) Israel Journal of Mathematics, 2012, Volume 192 (1), 59-66
2) Hilbert cubes II, 2012 arxiv, and further work in progress

with A. Dujella 3) Sumsets being squares, Acta Math Hungarica

26 August 2013,
Budapest (Fourier workshop)

Abstract

Problem in number theory?

Need some further tool!

Try combinatorics!

Complexity of sequences

$$5 + \{0, 2\} + \{0, 6\} + \{0, 96\} = \{5, 7, 11, 13, 101, 103, 107, 109\}$$

$$11 + \{0, 2\} + \{0, 6\} + \{0, 90\} + \{0, 15630\}$$
$$= \{11, 15641, 101, 15731, 17, 15647, 107, 15737,$$
$$13, 15643, 103, 15733, 19, 15649, 109, 15739\}$$

A Hilbert cube, or affine cube, is an iterated sumset

$$H(a_0; a_1, \ldots, a_d) := a_0 + \{0, a_1\} + \{0, a_2\} + \cdots + \{0, a_d\}$$

all 2^d subsetsums of d base elements, with an affine shift a_0.
Gowers norm:

$$\|f\|_{U^d(g)} = \mathbf{E}_{x,h_1,\ldots,h_d \in G} \prod_{\omega_1,\ldots,\omega_d \in \{0,1\}} J^{\omega_1+\cdots+\omega_d} f\left(x + h_1\omega_1 + \cdots + h_d\omega_d\right)$$

Central question:

Suppose we study the additive structure of a set S, then it is natural to ask: what is the largest dimension d such that there is a d-dimensional Hilbert cube with $H \subset S \subset [1, N]$?

Selection of known results:

Theorem (Hilbert, 1892)

For every r, d, there exists a least number $h(d, r)$ so that for every colouring

$$\chi : [1, h(d, r)] \to [1, r]$$

there exists an affine d-dimensional monochromatic cube.

(Early Ramsey type result).

Theorem (Szemerédi, 1969, density version)

For each d, there exists a constant c_d so that (for large $N \geq N_d$): if $S \subset [1, N]$ and $|S| \geq c_d N^{1-\frac{1}{2^d}}$, then S contains an affine d-dimensional Hilbert cube.

Corollary

If S has positive lower density, i.e. $\liminf \frac{S(N)}{N} \geq c > 0$, then $d \geq \log \log N$.

Theorem (Hegyvári)

Let $d(S,N) = \max\{k : S \cap [1,N]$ *contains a* k*-dimensional cube*$\}$. *There exists an infinite sequence* S *of positive integers with positive lower density and*

$$d(S,N) \ll \sqrt{\log N \log\log N}.$$

Theorem (Conlon, Fox, Sudakov)

For any $0 < \delta < 1$, *there exists* $c > 0$ *such that with high probability a random subset of* $[1,N]$, *where each element is chosen independently with probability* δ, *does not contain Hilbert cubes of dimension* $c\sqrt{\log N}$. *(Best possible, by another result of Hegyvári).*

Further results by Gunderson & Rödl, Hegyvári, Sandor et al.

If S is the set of primes. (For $a_i \neq a_j$).

$$5 + \{0, 2\} + \{0, 6\} + \{0, 96\} = \{5, 7, 11, 13, 101, 103, 107, 109\}$$

For distinct integers a_i:

Theorem (Hegyvári and Sárközy)

$$d(primes, N) = O(\log N).$$

Theorem (Wood, CE, independently)

$$d(primes, N) = O(\frac{\log N}{\log \log N})$$

Corollary (Wood)

Suppose C_n is a Σ_3^2 circuit which tests whether $X_1 \ldots X_n$ are the binary digits of a prime number. Then the number of AND gates which are used as inputs to the output OR gate must be at least:

$$2^{n-(9/2+o(1))\frac{n \log \log n}{\log n}}.$$

Squares

$$a_0 + \{0, a_1\} + \{0, a_2\} + \cdots + \{0, a_d\}.$$

Let S be the set of squares. $1 + \{0, 15\} + \{0, 48\} = \{1, 16, 49, 64\}$.
JL Nicolas (1977), J. Lagrange (1981):

$$A = \{-15863902, 17798783, 21126338, 49064546, 82221218, 447422978\}$$

$$A +' A = \{1934881, 5262436, 33200644, 38925121, 66357316, 66863329,$$

$$70190884, 100020001, 103347556, 131285764, 431559076,$$

$$465221761, 468549316, 496487524, 529644196\}$$

$A + B \subset$ Squares . If $|A| = 2$ what can we say on $|B|$???
Finite/infinite, upper/lower bounds ???

$A + B \subset$ Squares, and $|A| = 2$

$a_1 + b_i = x^2, a_2 + b_i = y^2$. $a_2 - a_1 = y^2 - x^2 = (y + x)(y - x)$.
$|B|$ may be arbitrarily large! An integer $n = \prod_{i=1}^{s} p_i$ has 2^s divisors. Each divisor induces a solution $a_2 - a_1 = (y - x)(y + x)$.
If A fixed, $|B|$ is bounded, but not uniformly.
Example: $a_2 - a_1 = 3 \times 5 \times 7 \times 11 = 1155$.

y	x	$y - x$	$y + x$
34	1	33	35
38	17	21	55
46	31	15	77
58	47	11	105
86	79	7	165
118	113	5	231
194	191	3	385
578	577	1	1155

$|A| = 2$

If only $|A| = 2$ is fixed, but the set A may vary, $A, B \subset [1, N]$, then

$$|B| \sim \exp\left((\log 2 + o(1))\frac{\log N}{\log\log N}\right)$$

is possible. (Finite, but may tend to infinity with varying A and N).

$|A| = 3$

If $|A| = 3$, question already asked by Euler. $|A| = 3$ and $A, B \subset [1, N]$ and B can be arbitrarily large,

$$|B| \geq (\log N)^{5/7}$$

Alon, Angel, Benjamini, Lubetzky (Israel JM 2012).
(a certain elliptic curve, here of rank 5, has many rational points, applications to expanders)

Theorem (Dujella, CE)

There exists A and $B \subset [1, N]$ and $A + B \subset$ squares, with $|A| = 3$ and

$$|B| \geq c(\log N)^{15/17}.$$

Moreover, if A, B and $A + B$ are all squares:

$$|A| = 3, |B| \geq c(\log N)^{9/11}$$

$|A| = 3$, continued

$$y^2 = x^3 + x^2 - 614100573770591167151980664421796984 0x + 585743317734880315858628578592963147780809517115906318 8.$$

$$a_1 = 0, a_2 = 28678999^2, a_3 = 43105370^2.$$

Conjecture (CE, upper bounds)

$$|A| = 3, |B| \leq \textit{ something better than } \exp\left(\frac{c \log N}{\log \log N}\right)$$

$|A| \geq 4$

The Bombieri-Lang conjecture (completely hopeless, very general) implies:
If $|A| = k \geq 4$, $|B| \leq C_k$ is **uniformly** bounded.
(Argument above but with curves of genus $g > 1$ have very few rational points).
(Details, see Solymosi; Cilleruelo and Granville).

Brown, Erdős and Freedman (1990) asked about the maximal dimension of Hilbert cube.

Theorem (Hegyvári and Sárközy, 1999)

$d(\textit{squares}, N) = O((\log N)^{1/3})$.

Theorem (Dietmann, C.E.)

$d(\textit{squares}, N) = O(\exp((\log\log N)^{1/2}\log\log\log N)$.
$d(\textit{squares}, N) = O((\log\log N)^2)$.
$d(\textit{squares}, N) = O(\log\log N)$.

Theorem (Dietmann, C.E.)

$d(k\text{-}th\ \textit{powers}\ (k \geq 3), N) = O(\log\log N)$.

Conjecture (Solymosi, special case of Bombieri-Lang)

$d(\textit{squares}, N) = O(1)$.

$|A| \geq 4$, Heuristics

Caporaso, Harris, Mazur proved: If the Bombieri-Lang conjecture is true, there exists an integer k such that for any polynomial $f \in \mathbb{Z}[x]$ of degree six, which does not have repeated roots, there are no more than k rational numbers m, for which $f(m)$ is a square. If $x^2 \in H(a_0; a_3, a_4, \ldots, a_d)$ then $x^2 + a_1, x^2 + a_2, x^2 + a_1 + a_2$ are all squares, and so is $f(x) = (x^2 + a_1)(x^2 + a_2)(x^2 + a_1 + a_2)$, which is only possible for at most k values of x, hence $d \leq k$.

Theorem (Folklore ?, Hegyvári and Sárközy)

There is no infinite Hilbert cube in the squares.

Proof 1: By Siegel-Baker: For any given a_1, a_2, there are finitely many integer points on the elliptic curve $y^2 = f(x)$, but ineffective, no uniform bound.

Proof 2: Infinite Hilbert cubes have fixed difference infinitely often.
$a_2 - a_1 = y^2 - x^2$.
The sequence of differences of the squares tends to infinity, hence no fixed difference can occur infinitely often.

Proof 3: $a_2 - a_1$ has finitely many divisors only.

Hilbert cubes in the squares
Sketch of the Hegyvari-Sárközy approach

- local-global method:
  Instead of studying the problem for integers, they study it modulo primes. The squares modulo primes are in half of the residue classes modulo p.
- If $A, B \subset \mathbb{Z}/p\mathbb{Z}$, and $A + B \subset \{x^2 : x \in \mathbb{Z}/p\mathbb{Z}\}$, then $\min(|A|, |B|) \leq \sqrt{p}$. (Gauß-sum, character sum)
- If $\{a_1, \ldots, a_d\} \subset \mathbb{Z}/p\mathbb{Z}$ are in t distinct residue classes, their sumset $\{0, a_1\} + \cdots + \{0, a_d\}$ is in at least $\frac{t^2}{2} \leq \sqrt{p}$ classes modulo p.
  Example $A = \{1, 2, 3, 4, ..., d\} \Rightarrow \Sigma A = \{1, 2, 3, \ldots, \frac{d(d+1)}{2}\}$.
  So, $\{a_1, \ldots, a_d\}$ is modulo p in $O(p^{1/4})$ residue classes.
- A restriction of this type, modulo many primes, leads (by a sieve method) to $d = O((\log N)^{1/3})$.

What should we do?

- Do not use curves of higher genus, too complicated.
- Subsetsum mod p argument seems (here) very inefficient.
- analyze subsetsums for sets with forbidden "global" structure.
- Would like to use argument of the type:
  1) If $C + D$ is in a set without ... then $\min(|C|, |D|) \leq$ ...
  2) If the sumset of $A = \{a_1, \ldots, a_d\}$ is in a set without ..., then $\Sigma A \geq$ ... (hope for 2^d), use better combinatorics here.

Sets without arithmetic progressions

Roth, Bourgain, Gowers, Tao, Bombieri, Green, Sanders, Granville, Pintz, Zannier et al. worked on density of sets without an arithmetic progression, or on squares in arithmetic progresson etc. Why should we study this?

- k-th powers and squares do not have long arithmetic progressions!
  $x^k + y^k = 2z^k$
- Arithmetic progressions are the bad case for the subsetsums modulo primes.

Lemma (Gyarmati)

For all $k > 1$, there is a N_k such that for all $N > N_k$: If $A, B \subset \{1, \dots, N\}$ with $A + B \subset S_k := \{x^k : x \in \mathbb{N}\}$ for some $k \geq 2$. Then: $\min\{|A|, |B|\} \leq 4k \log N$.

This is proved using Gauß sums.

Lemma (Darmon-Merel, compare with Fermat-Wiles)

For $k \geq 3$: $x^k + y^k = 2z^k$ does not have nontrivial solutions that is the set of k-th powers does not have arithmetic progressions of length 3.

All of the 2^d subset-sums must be distinct:

$$a_0 + a_{i_1} + \cdots + a_{i_r} + a_{j_1} + \cdots + a_{j_s} = a_0 + a_{i_1} + \cdots + a_{i_r} + a_{k_1} + \cdots + a_{k_t}.$$

$$a_0, a_0 + a_{j_1} + \cdots + a_{j_s}, a_0 + a_{j_1} + \cdots + a_{j_s} + a_{k_1} + \cdots + a_{k_t}$$

is an arithmetic progression. $A = a_0 + \{0, a_1\} + \cdots + \{0, a_{\frac{d}{2}}\}$ and $B = \{0, a_{\frac{d}{2}+1}\} + \cdots + \{0, a_d\}$, $2^{d/2} \leq 4k \log N$ and so $d = O(\log \log N + \log k)$.

Lemma (Fermat, Euler)

There are no four squares in arithmetic progression.

Lemma (Growth lemma I)

$|A| \geq h$ and let $S \subset \mathbb{N}$ be a set without progression of length k. If

$$a_0 + \{0, a_1\} + \{0, a_2\} + \cdots + \{0, a_d\} \subseteq S,$$

then $|h\hat{}A| \geq \frac{(|A|-h)^h}{(k-2)^h\, h!\, h!}$. Here $h\hat{}C$ is the disjoint sumset

$$\{c_1 + c_2 + \cdots + c_h : c_i \in C, \text{ but } c_i \neq c_j \text{ for } i \neq j\}.$$

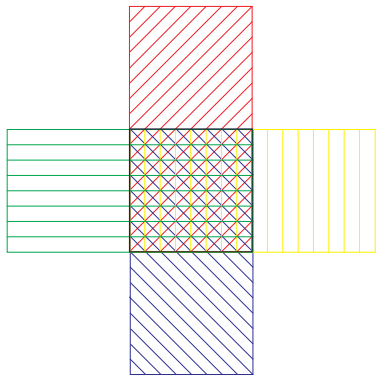
$A_1 = \{a_1, a_2, \ldots, a_{\frac{d}{2}}\}, A_2 = \{a_{\frac{d}{2}+1}, \ldots, a_d\}$.

For $i = 1$ or $i = 2$: $\frac{(|A_i|-h)^h}{(k-2)^h\, h!\, h!} \leq |h\hat{}A_i| \leq 8 \log N$.

$$|A_i| \leq (8\, 2^h\, h!\, h! \log N)^{1/h} + h \ll h^2 (\log N)^{1/h} \ll (\log \log N)^2,$$

$$(h = \log \log N).$$

Alon observed: proof contains a number theoretic version of a result of Erdős-Rado on sun flowers (Δ systems).
Jukna: "A sunflower with k petals and a core X is a collection of sets $S_1, \ldots, S_k$ such that $S_i \cap S_j = X$ for all $i \neq j$.
(The sets $S_i \backslash X$ are petals, and we require that none of them is empty...)"



Lemma (Erdős-Rado, 1960, Sunflower Lemma)

Let $\mathcal{F}$ be family of sets S_i with $|S_i| = s$. If $|\mathcal{F}| > s!(k-1)^s$, then $\mathcal{F}$ contains a sunflower with k petals.

Conjecture (Erdős-Rado, \$1000)

For every k there exists a C_k such that: If $|\mathcal{F}| > C_k^s$, then $\mathcal{F}$ contains a sunflower with k petals.

Sets without arithmetic progression

Lemma (Growth Lemma II, inspired by T. Schoen)

Let $k \geq 3$ be a positive integer, and let S denote a set of integers without an arithmetic progression of length k. Moreover, let c be a real number such that

$$1 < c < \frac{k}{k-1}.$$

Let $H = a_0 + \{0, a_1\} + \cdots + \{0, a_d\} \subset S$. Then $|H| \geq 2c^{d-1}$.

Theorem (Dietmann, CE)

Hilbert cube $H \subset S_k$, S_k no k-progression.

$$d \leq \frac{2(k-2)}{(k-1)\log c} \log N.$$

Hilbert cubes in the squares

By Fermat no 4-progression
By Gyarmati $A_1 + A_2 \subset$ squares, $\min |A_i| = O(\log N)$.
$A_1 = \{a_1, a_2, \ldots, a_{\frac{d}{2}}\}, A_2 = \{a_{\frac{d}{2}+1}, \ldots, a_d\}$.
By growth lemma II, for $i = 1$ or $i = 2$:

$$2c^{d-1} \leq |h\hat{}A_i| \leq 8 \log N.$$

Hence:

$$d = O(\log \log N).$$

Note that the growth lemmas I/II use "no 4-progression" (Fermat), i.e. an important global property. This is the key to improvement over the more local approach (modulo primes) by Hegyvári and Sárközy.

For progression-free sets almost best possible.

A lower bound of $\log N$.
Let $a_0 = 0$, and let the multi-set A consist of $k-2$ copies of $a_{i+1} = k^i : i = 0, \ldots, s = \lfloor \frac{\log N}{\log k} \rfloor$.
Let $H = H(0; a_1, \ldots, a_1, \ldots, a_s, \ldots, a_s)$, (each element $k-2$-fold).
All $n \in H$ can be written as $n = \sum_{i=0}^{s} b_i k^i$, with $b_i \in \{0, 1, \ldots, k-2\}$.
H does not contain any arithmetic progression of length k.
Digit restrictions, compare Behrend's example for progression-free sets.

Sketch proof of Growth lemma II, preparation

Lemma (Preparation)

$0 < \alpha < 1$, *and* $h \in \mathbb{Z}$, $B \subset \mathbb{N}$ *non-empty set. If* $|B \cap (B+h)| > (1-\alpha)|B|$, *then* B *contains an arithmetic progression of length* $\lfloor \frac{1}{\alpha} \rfloor + 1$ *and gap* h.

Proof.

Shift Operator $f : \mathbb{Z} \to \mathbb{Z}$: $f(b) = b + h$ and iterated shifts. For fixed $b \in \mathbb{Z}$, $r(b)$ is the least integer $r \geq 0$ with

$$\{b, f(b), \ldots, f^{r-1}(b)\} \subset B, \text{ and } f^r(b) = b + rh \notin B.$$

From $|\{b \in B : b + h \in B\}| > (1-\alpha)|B|$ it follows, that for every $r \geq 0$ there are at most $\alpha|B|$ elements $b \in B$ with this value $r = r(b)$. For $k \in \mathbb{N}$, the number of $b \in B$ with $r(b) \leq k$ is at most $k\alpha|B|$. If $k\alpha|B| < |B|$, then there exists $b \in B$ with $r(b) \geq k+1 \geq \lfloor \frac{1}{\alpha} \rfloor + 1$. Then B contains an arithmetic progression $\{b, b+h, \ldots, b+(r-1)h\}$ of length $r \geq \lfloor \frac{1}{\alpha} \rfloor + 1$. □

Sketch proof of Growth lemma II, part 2

Growth lemma.

$$H_i = a_0 + \{0, a_1\} + \cdots + \{0, a_i\}.$$

If $|H| < 2c^{d-1}$, then there exists $i \in \{1, \ldots, d-1\}$ with

$$\frac{|H_{i+1}|}{|H_i|} < c.$$

For this i one has $|(H_i + a_{i+1}) \cap H_i| > (2-c)|H_i|$. By the preparation lemma and by assumption $1 < c < \frac{k}{k-1}$:
H_i contains an arithmetic progression of length

$$\lfloor \frac{1}{c-1} \rfloor + 1 \geq k$$

contradicting the assumption that S does not contain such progressions! □

This topic involves

additive combinatorics (sumsets, progressions),
Ramsey theory,
complexity theory,
sieve methods,
character sums,
group theory (not in this talk),
arithmetic geometry (Fermat type and uniform bounds on number of solutions),
elliptic curves (Siegel-Baker).
sun flower lemma
sum set growth

Thank you for your attention