Astérisque

IMRE Z. RUZSA

Probabilistic constructions in additive number theory

Astérisque, tome 147-148 (1987), p. 173-182

<http://www.numdam.org/item?id=AST_1987__147-148__173_0>

Société Mathématique de France
Astérisque 147-148 (1987), p.173-182.

PROBABILISTIC CONSTRUCTIONS IN ADDITIVE NUMBER THEORY

by
Imre Z. Ruzsa

1. Introduction

Sometimes it is easier to find a lot of examples of a certain phenomenon than one, by showing that, in a suitable probability, almost all objects are good. Examples abound in combinatorics (see Erdős-Spencer [1] as well as in number theory (see Halberstam-Roth [4]). Now we use a random construction to produce a wide class of "additively very effective" sets of positive integers. The novelty in the proof is that we use probability to study the Fourier transform of our random set, in contrast to previous works where generally probability was applied directly to the representation. While this is not a panacea, sometimes it can yield superior results. I have applied this approach to construct a thin essential component (Ruzsa [7]).

I should like to use this occasion to point out that the distinction between "probabilistic" and "deterministic" constructions is a rather unclear one. After having proved, by probability or other methods, the existence of a sequence with certain properties, it is generally not too difficult to pinpoint a concrete example, say the first in some lexicographical ordering, thus making our proof formally deterministic. Speed may seem a distinctive feature (and some insist it is): the above procedure of selecting the "first" works in exponential time, while most "real" constructions work much faster. On the other hand, I am rather inclined to accept a definition like "the remainder of $2^n!$ mod 3^n+1", which is very slow to compute, as deterministic. This - perhaps more philosophical than mathematical - problem may be worth clarifying.

2. The main result

We use capital letters to denote a set of integers, and we use the same letter to denote its counting function; that is, if A is

a set of integers, we write

$$A(x) = |A \cap [1, x]| .$$

Observe that even if A has negative elements, they are not counted into $A(x)$. The Shnirelmann density of A is defined by

$$\sigma(A) = \inf A(n)/n,$$

n runs over the integers. The asymptotic density is

$$d(A) = \lim_{x \to \infty} A(x)/x$$

if it exists. The upper limit of $A(x)/x$ is called the upper density of A and is denoted by $\overline{d}(A)$. As usual, we write $A+B$ $(A-B)$ to denote the set of all numbers $a+b$ $(a-b)$, $a \in A$, $b \in B$.

THEOREM. Let g be a positive-valued function defined on $[2, \infty)$, such that $g(x)/\log x$ is increasing while $g(x)/x$ tends monotonically to 0. There exists a set H of positive integers with the properties that

$$H(x) \sim \sum_{n \leqslant x} \frac{\log n}{g(n)} \leqslant \frac{x}{2g(x)} \log^2 (x+1) , \tag{2.1}$$

and that if A is another set of integers and

$$A(x_k)/g(x_k) \to \infty \tag{2.2}$$

for a sequence (x_k), then their sum $S=A+H$ and difference $D=H-A$ satisfy

$$\frac{S(2x_k)-S(x_k)}{x_k} \to 1 , \quad D(x_k)/x_k \to 1 . \tag{2.3}$$

In particular, if $A(x)/g(x) \to \infty$, then both S and D have asymptotic density one.

Comments. 1) The flexibility built in the theorem (the possibility to choose $g(x)$ and (x_k)) will be useful for the applications, cf. the next section.

2) In (2.3), we cannot assert $S(x_k)/x_k \to 1$. It can

namely happen that almost all elements of A between 1 and x_k are near to x_k, and then the majority of sums in $A+H$ will be greater than x_k. Similarly, instead of $H-A$ we cannot consider $A-H$, since if almost all elements of A in $[1, x_k]$ are small, then so are the differences in $A-H$.

3) To get a positive portion of the numbers up to x into $A+H$ we must have $A(x)H(x)>cx$; thus if this has to happen whenever $A(x)/g(x)\to\infty$, $H(x)$ must not be $o(x/g(x))$. (2.1) is at most a $\log^2 x$ factor higher. If $g(x)$ is near to x, this is necessary as we shall see in the next section. For $g(x)=\sqrt{x}$, (2.1) yields $H(x)=O(\sqrt{x}\log x)$, and I cannot decide whether this can be improved. If $g(x)$ is as small as $\log x$, then (2.1) does not produce any nontrivial bound. I can obtain better bounds than (2.1) for slowly increasing $g(x)$ (up to $g(x)=x^{1/4-\varepsilon}$) by different methods; this will be discussed in another paper.

4) The second inequality in (2.1) is easy to see. If $g(x)/x$ is decreasing, then $g(n)/n\geq g(x)/x$ for $n\leq x$, hence

$$\sum_{n\leq x}\frac{\log n}{g(n)}\leq\frac{x}{g(x)}\sum_{n\leq x}\frac{\log n}{n}\leq\frac{x}{g(x)}\,\frac{1}{2}\log^2(x+1)\ .$$

3. Applications

We show how a solution to a couple of older problems can be obtained as a corollary to our theorem.

Following Erdős and Sárközy [2, 3], we call a set H sum-intersective, if $H\cap(A+A)\neq\emptyset$ whenever $\overline{d}(A)>0$. This is an analogon to the much investigated concept of a difference-intersective set (see Ruzsa [8] for further reference), but it is much less understood. In particular, no "natural" set is known to have this property. From Theorem 7 of Erdős and Sárközy [3] it follows that if H is sum-intersective, then $H(x)\neq o(\log^2 x)$, and I can prove the slightly stronger

$$H(x)/\log^2 x\to\infty\ . \tag{3.1}$$

In [2] they gave a probabilistic construction of a sum-intersective H with $H(x)=O(x^{2/3+\varepsilon})$. In [7] I gave a deterministic construction with $H(x)=O(x^{1/2+\varepsilon})$. Here we show that the lower estimate (3.1) actually gives the correct order of magnitude.

COROLLARY 1. Let w be a positive function defined on $[1, \infty)$

and tending to infinity. There exists a sum-intersective set H with

$$H(x) = O(w(x) \log^2 x) . \tag{3.2}$$

Proof. Apply the Theorem with some function $g(x)=o(x)$, $g(x)\geqslant x/w(x)$ (we can have $g(x)=x/w(x)$ if $w(x)$ is regular enough). We obtain a set H that clearly satisfies (3.2). We have to show that it is sum-intersective.

Since A has a positive upper density, there is a sequence (x_k) such that

$$A(x_k)/x_k > c > 0 . \tag{3.3}$$

Then also $A(x_k)/g(x_k)\to\infty$, thus by the Theorem, for $D=H-A$ we have

$$D(x_k)/x_k \to 1 . \tag{3.4}$$

By (3.3) and (3.4), $A(x_k)+D(x_k)>x_k$ for large k, and then A and D cannot be disjoint. This means that there are numbers a_1, $a_2\in A$ and $h\in H$ such that $a_1=h-a_2$; this yields $h=a_1+a_2$ as wanted.

For another application, we consider a problem that is discussed in Halberstam-Roth [4], Ch. 1. §5. They ask whether a sequence A of density 0 exists such that $\sigma(A+B)>0$ for every basis B. P. Erdős has shown that the sequence of primes does not have this property.

COROLLARY 2. Let w be a positive-valued function on $[1, \infty)$ satisfying $w(x)=O(x^{\varepsilon})$ for all $\varepsilon>0$. There exists a set H satisfying

$$H(x) = O(x/w(x)) \tag{3.5}$$

such that for every basis B, $H+B$ has asymptotic density one and a positive Shnirelmann-density.

Remark. If B is a basis of order k, then clearly

$$B(n) \geqslant n^{1/k} \tag{3.6}$$

for all n; and this is the only property we shall need, it will be irrelevant whether B is indeed a basis.

Proof. Apply the Theorem with some g satisfying both

$$g(x) > w(x) \log^2 x , \quad g(x) = O(x^{\varepsilon}) .$$

(3.5) follows from (2.1). If B is a basis, say of order k, then (3.6) yields $B(x)/g(x) \to \infty$ and then for $S=H+B$ we have $d(S)=1$.

To obtain $\sigma(S)>0$, observe that it is equivalent to $d(S)>0$ and $1 \in S$. Now if B is a basis, it contains both 0 and 1, and if we include 1 to H, $1 \in S$ is guaranteed.

4. Outline of the proof

We are going to select the positive integers into H independently, with probability

$$p_n = P(n \in H) = \frac{\log n}{g(n)} . \tag{4.1}$$

We use ξ_n to denote the indicator of H, that is, $\xi_n=1$ if $n \in H$, $=0$ otherwise. Thus the ξ_n's are independent random variables and

$$\xi_n = \begin{cases} 1 & \text{with probability } p_n \\ 0 & \text{with probability } 1-p_n \end{cases} . \tag{4.2}$$

Hence its expectation and variance are

$$E(\xi_n) = p_n , \quad D^2(\xi_n) = p_n(1-p_n) . \tag{4.3}$$

By the strong law of large numbers we can immediately conclude that with probability one

$$H(x) = \sum_{n \leqslant x} \xi_n \sim \sum_{n \leqslant x} p_n , \tag{4.4}$$

and hence that (2.1) holds with probability one.

The proof of the additive properties (2.3) will not be so straightforward. We forst estimate some trigonometrical sums

$$\sum_{h \in H, h \leqslant x} a_h \, e(ht) ,$$

and this will be applied in a more or less classical way.

5. A random trigonometrical sum

We shall need an estimate for certain sums

$$(5.1) \qquad r_m(k) = \sum_{h\in H, h\leqslant m} a_h\, e(kh/m) = \sum_{n\leqslant m} a_n \xi_n\, e(kn/m) ,$$

where $e(x)=\exp 2\pi i x$ as usual. We shall be interested only in rational arguments with denominator m, and we need sums that are small for $k\not\equiv 0 \pmod m$. For a fixed k and m, $r_m(k)$ is a random variable, a sum of the independent variables $\zeta_n = a_n e(kn/m)\xi_n$. Its expectation is

$$(5.2) \qquad E(r_m(k)) = \sum_{n\leqslant m} a_n e(kn/m) E(\xi_n) = \sum_{n\leqslant m} a_n p_n e(kn/m) .$$

We cannot hope having small values without a small expectation. This expectation turns into 0 if all the coefficients $a_n p_n$ are equal. Recalling that $p_n = (\log n)/g(n)$, this can be achieved by putting

$$(5.3) \qquad a_n = \frac{g(n)}{\log n} .$$

With this choice of a_n the expectation of $r_m(k)$ will be

$$(5.4) \qquad E(r_m(k)) = \begin{cases} 0 & \text{if } k\not\equiv 0 \pmod m , \\ m & \text{if } k\equiv 0 \pmod m . \end{cases}$$

To this sum we shall apply a version of Bernstein's inequality.

LEMMA 1. *Let $\zeta_1, \dots \zeta_n$ be independent complex-valued random variables. Assume that $|\zeta_j - E\zeta_j| \leqslant K$ for all j and*

$$\sum_{j=1}^{n} D^2(\zeta_j) \leqslant D^2 .$$

Write $M = \Sigma E\zeta_j$. For $0<\lambda\leqslant D/K$ we have

$$(5.5) \qquad P(|\Sigma\zeta_j - M| > \lambda D) < c_1 \exp -c_2\lambda^2 ,$$

with positive absolute constants c_1 and c_2.

Real-valued versions of this inequality are standard (see e.g. Rényi [5]). To obtain the above complex variant one may apply the real inequality separately to the real and imaginary parts of the sum. In this way for c_2 we obtain a rather weak estimate (about $1/100$), but this is quite irrelevant for our goals.

Our aim is to deduce

LEMMA 2. Let $d_{mk}=0$ if $m\nmid k$, $=m$ if $m\mid k$. With suitable absolute constants c, C the inequality

$$P(|r_m(k)-d_{mk}| > c\sqrt{mg(m)}) < Cm^{-3} \tag{5.6}$$

holds for all but a finite number of values of m.

Proof. We use Lemma 1 with $\zeta_j=a_j e(kj/m)\xi_j$. Clearly $|\zeta_j-E\zeta_j|\leqslant a_j$, thus we can choose

$$K = \max_{j\leqslant m} a_j = \max_{j\leqslant m} \frac{g(j)}{\log j} = \frac{g(m)}{\log m} . \tag{5.7}$$

To obtain Cm^{-3} at the right side of (5.6), our choice of λ must be

$$\lambda = c_3\sqrt{\log m} , \quad c_3 = \sqrt{3/c_2} . \tag{5.8}$$

We need an estimate for the variance. Clearly

$$D^2(\zeta_j) = a_j^2 p_j(1-p_j) \leqslant a_j^2 p_j = g(j)/\log j$$

independently of k and m, thus

$$\sum_{j=1}^{m} D^2(\zeta_j) \leqslant \sum_{j=1}^{m} g(j)/\log j \leqslant m \max g(j)/\log j$$

$$= mg(m)/\log m .$$

With this as D^2, our λD will just be the $c_3\sqrt{mg(m)}$ appearing in (5.6). The condition $\lambda\leqslant D/K$ is equivalent to

$$g(m)/m\leqslant c_2/3 ,$$

which is satisfied for large m.

LEMMA 3. There is a set H of positive integers with

$$H(x) \sim \sum_{n\leqslant x} \frac{\log n}{g(n)}$$

and such that the trigonometrical sums

$$r_m(k) = \sum_{h\in H, h\leqslant m} \frac{g(h)}{\log h}\, e(kh/m)$$

satisfy

$$|r_m(k) - d_{mk}| \leqslant c\sqrt{mg(m)}$$

for all k and $m>m_0$.

Proof. We need to check only the values $k=1,\dots m$ because of the periodicity. Since the sum of probabilities of these events is, by Lemma 2,

$$\sum_m \sum_{k=1}^{m} Cm^{-3} = \sum_m Cm^{-2} < \infty ,$$

by the Borel-Cantelli lemma almost all sets described in the previous chapter fulfill the requirements of Lemma 3.

6. A modular additive property

Addition of sets of residue classes modulo m is connected with trigonometrical sums via the following lemma.

LEMMA 4. Let G be a set of (mod m) residue classes. Let a_n, $n\in G$ be arbitrary complex numbers and put

$$r(t) = \sum_{n\in G} a_n\, e(nt/m) . \tag{6.1}$$

If

$$|r(t)| \leqslant \eta |r(0)| \tag{6.2}$$

for all integers $t\not\equiv 0$ (mod m), then for any set X of residue classes we have

$$|X+G| \geqslant \frac{m|X|}{|X|+\eta^2(m-|X|)} . \tag{6.3}$$

For a proof, see Ruzsa [7].

(6.3) immediately yields

$$m-|X+G| \leqslant \eta^2 m/|X| . \tag{6.4}$$

Now let H_m be the set of residues modulo m of the elements of H up to m . In view of Lemma 3, (6.2) holds with some

$$\eta = \eta_m = c'\sqrt{g(m)/m}$$

with any $c'>c$ for sufficiently large m . Hence

$$m - |X+H_m| \prec c'mg(m)/|X| \tag{6.5}$$

for all X .

7. Completion of the proof

First consider S=A+H . We put $m=2x_k$. In view of $A(x_k)/g(x_k)\to\infty$, there is an interval

$$I = [u, u+\varepsilon_k x_k] \subset [1, x_k]$$

of length $\varepsilon_k x_k$ that contains more than $\omega_k g(x_k)$ elements of A , where $\varepsilon_k\to 0$ and $\omega_k\to\infty$. Let $X=A\cap I$.

With a slight abuse of the notation, let H_m be the set of elements of H up to m . By (6.5) we conclude that $X+H_m$ contains at least

$$m - c'mg(m)/|X| \succ m(1-c'/\omega_k)$$

different residues mod m . Thus its cardinality is also at least so much, and it is contained in

$$[u, u+\varepsilon_k x_k+m] ,$$

an interval of length $m+\varepsilon_k x_k$. Therefore it contains all but

$$\varepsilon_k x_k + c'm/\omega_k = x_k(\varepsilon_k+2c'/\omega_k) = o(x_k)$$

elements of that interval. This interval contains $[x_k, 2x_k]$, thus all but $o(x_k)$ numbers between x_k and $2x_k$ are in S , qu. e. d.

To obtain information about D=H-A we consider H_m-X in the same way. Our basic interval will then be $[-u-\varepsilon_k x_k, m-u]$ and this always contains $[0, x_k]$.

Added in proof. Recently I learned that some similar investigations were done by P. Erdős and A. Rényi, "On some applications of probability methods to additive number theoretic problems", in : Contributions to Ergodic Theory and Probability (Proc. Conf., Ohio State University, Columbus, USA 1970), Springer 1970, p. 37-44.
They show (Theorem 3) that for every function $f(n) \to \infty$ there is a set A of density 0 with the property that $d(A+B)=1$ whenever the set B satisfies $B(n) > f(n)$ for all large n. This does not imply and is not implied by my Corollary 2, but it also solves the problem that I quote from Halberstam-Roth's book and which is actually due to A. Stöhr, Gelöste und ungelöste Fragen über Basen der natürlichen Zahlenreihe, J. reine angew. Math. 197 (1957), 216-219. Compared to my analytic method, their elementary probabilistic approach is superior for the treatment of dense random sets (x/log x elements or more), but weaker for thin ones.

References

1. P.Erdős, J.Spencer, Probabilistic methods in combinatorics, Akadémiai Kiadó, Budapest 1974.
2.-3. P.Erdős, A.Sárközy, On differences and sums of integers I-II; I, J.Number Theory 10 (1978), 430-450; II, Bull. Greek Math. Soc. 18 (1977), 204-223.
4. H.Halberstam, K.F.Roth, Sequences I, Clarendon Press, Oxford 1966.
5. A.Rényi, Probability theory, Amsterdam 1970.
6. I.Z.Ruzsa, Sets of sums and differences, Séminaire de Théorie des Nombres, Paris 1982/83, 267-273, Birkhäuser, Boston 1984.
7. I.Z.Ruzsa, Essential components, to appear in the J. London Math. Soc.
8. I.Z.Ruzsa, Difference sets and the Bohr topology I, submitted to the Michigan J. Math.

Imre Z. Ruzsa
Mathematical Institute of the
Hungarian Academy of Sciences
Budapest, Pf. 127
H-1364 Hungary