GREGORY A. FREIMAN

Structure theory of set addition

Astérisque, tome 258 (1999), p. 1-33

<http://www.numdam.org/item?id=AST_1999__258__1_0>
STRUCTURE THEORY OF SET ADDITION

by

Gregory A. Freiman

Abstract. — We review fundamental results in the so-called structure theory of set addition as well as their applications to other fields.

1. ‘Structure theory of set addition’(1) is a shorthand for a direction in the study of sets which extracts structures from sets for which some properties of their sums (or products in a non-abelian case) are known.

Here is an indication of what is meant by “structure”. The first stage is to build an equivalence relation on sets. Then, by taking well chosen representatives of an equivalence class we are able to reveal its properties and thereby describe its structure (see, for example, the Definition and Theorem in §6).

2. This review is written in the following way. In §§3–8 we explain the main ideas. In §§9–12 we make some historical remarks. Then in §§13–19 we present several concrete problems in additive and combinatorial number theory, showing how new results may be obtained with the help of the described new approach. Further then in §§20–27 we try to show a diversity of fields where the ideas of “Structure Theory” may be applied. Finally in §§28–35 we discuss methods and problems. In the bibliography we include references to a wider spectrum of subjects which may be treated from the point of view of Structure Theory.

3. This approach to additive problems was originally given the name “Inverse problems of additive number theory”. A series of nine papers under this heading was published in 1955–1964 (see [85], [86], [87], [88], [89], [90], [91], [92] and [98]).

4. I quote from my lecture in the Fourth All-Union Mathematical Congress, Leningrad, 3-12 July 1961 (see [84]):

1991 Mathematics Subject Classification. — 11 02, 11Z05.

Key words and phrases. — Structure theory of set addition, inverse problems of additive number theory, small doubling property, isomorphism of subsets.

(1)This paper is based on my review lecture given at the conference on Structure theory of set addition held at CIRM (Centre International des Rencontres Scientifiques), Luminy, Marseille, on 10 June 1993.
"The term inverse problems of additive number theory appeared in 1955 in two of my papers [85](2) and [86]. In [85] the following problem was studied. Let
\[ a_1, a_2, \ldots, a_r, \ldots \]  
be an unbounded, monotonically increasing sequence of positive numbers. To have an asymptotic formula
\[ \log q(u) \sim A u^\alpha, \quad \text{where } A > 0, 0 < \alpha < 1 \]
it is necessary and sufficient that
\[ n(u) \sim B(A, \alpha) u^{\alpha/(1-\alpha)} \]
where \( n(u) \) is the number of terms of a sequence (1) not exceeding \( u \), and \( q(u) \) is the number of solutions of the inequality
\[ a_1n_1 + a_2n_2 + \cdots \leq u. \]

In [86] the case \( \log q(u) = A u^\alpha + O(u^{\alpha_1}), \quad \text{where } 0 < \alpha_1 < \alpha, \)
was studied and an estimate of the error term in the asymptotic formula for \( n(u) \) was obtained.

One can easily see that if \( q(u) \) is known then (1) is determined in a unique way (see [85]). In 'direct' problems we study \( q(u) \) when the sequence (1) is given; a particular case is the classical problem on the representation of positive integers as sums of an unlimited number of positive integers.

Thus a direct problem in additive number theory is a problem in which, given summands and some conditions, we discover something about the set of sums. An inverse problem in additive number theory is a problem in which, using some knowledge of the set of sums, we learn something about the set of summands.

Several cases of inverse problems were studied earlier; see [14] and [67].

Paul Erdős, in 1942, found an asymptotic formula for \( n(u) \) when
\[ \log p(u) \sim a\sqrt{u} \]
where \( p(u) \) is the number of solutions of an equation
\[ a_1n_1 + a_2n_2 + \cdots = u \]
where \( \{a_i\} \) is some sequence of positive integers (see [67]).

In the same paper another inverse problem was studied; if \( q(u) \sim C' u^{2\alpha} \), where \( q(u) \) is the number of solutions of an inequality
\[ a_i + a_j \leq u, \]

(2) The reference numbers given accord with the bibliography of this paper and not the original text.
then

\[ n(u) \sim C_1 u^\alpha. \]

In 1960 V. Tashbaev [252] studied the problem of estimating the error term for this inverse problem.

We will now explain how problems on the distribution of prime numbers are connected with inverse problems. If we define

\[ q(u) = \lfloor e^u \rfloor \]

then \( a_i = \log p_i \), where \( p_i \) denotes the \( i^{th} \) prime number. Thus the problem of the distribution of prime numbers may be treated as an inverse problem of additive number theory of the type described above. The study of inverse problems for different \( q(u) \) close to \( \lfloor e^u \rfloor \), and also of direct problems when \( n(u) \) is close to \( e^u / u \), may give some insight into the problem of the distribution of primes, in a way similar to that in which the behaviour of a function in the vicinity of a point may help to find its value at that point (see A. Beurling [14] and B. M. Bredichin [30], [31], [32] and [33])."

The results of Diamond (see [57], [58], [59], [60] and [61]) should of course be mentioned.

The treatment of prime distribution problems as inverse additive problems have not developed up to now. I still consider this approach very hopeful.

5. We pass on now to the study of additive problems with a fixed number of summands. The majority of papers mentioned in §3 treat the addition of two equal sets. The study of this particular case is usually sufficient to develop ideas, methods and results as well as their use in applications.

Let us start with \( K \subseteq \mathbb{Z} \) with \( |K| = k \). Define

\[ 2K = K + K = \{ x \mid x = a_i + a_j, \quad a_i, a_j \in K \}. \]

We may ask the question what is the minimal cardinality of \( 2K \)? Evidently,

\[ |2K| \geq 2k - 1. \quad (2) \]

Suppose now that \( K \) is such that \( |2K| \) is minimal i.e. \( |2K| = 2k - 1 \). What can be said about such a \( K \)? It is clear that,

\[ |2K| = 2k - 1, \quad (3) \]

only if \( K \) is an arithmetic progression.

Suppose now that \( |K + K| \) is not much greater than this minimal value. In that case we have the following result [87], describing the structure of \( K \).

**Theorem 1.** — Let \( K \) be a finite set, \( K \subseteq \mathbb{Z} \). If

\[ |K + K| \leq 2k - 1 + b, \quad 0 \leq b \leq k - 3 \]

then \( K \) is contained in an arithmetic progression of length \( k + b \).
Further, suppose that we know that
\[ |2K| < Ck, \]
where \( C \) is any given positive number, we may ask what then is the structure of \( K \)?

6. The theorem answering this question (we will quote it as a main theorem) was proved in a previously mentioned series of papers, expositions of it were given in [81] and [82], and an improved version of a proof was presented in [105]. We are citing here the result of Y. Bilu [16], where he studies a case when \( C \) in (4) is a slowly growing function of \( k \).

**Definition.** — Let \( A \) and \( B \) be groups, and let \( K \subseteq A \) and \( L \subseteq B \). The map \( \phi: K \rightarrow L \) is called an \( \mathbb{F}_s \)-homomorphism, if for any \( x_1, \ldots, x_s \) and \( y_1, \ldots, y_s \) in \( K \) we have
\[ x_1 + \cdots + x_s = y_1 + \cdots + y_s \Rightarrow \phi(x_1) + \cdots + \phi(x_s) = \phi(y_1) + \cdots + \phi(y_s). \]
The \( \mathbb{F}_s \)-homomorphism \( \phi \) is an \( \mathbb{F}_s \)-isomorphism if it is invertible and the inverse \( \phi^{-1} \) is also an \( \mathbb{F}_s \)-homomorphism.

Let \( P \subseteq \mathbb{Z}^n \) be given by
\[ P = \{0, \ldots, b_1 - 1\} \times \cdots \times \{0, \ldots, b_n - 1\}. \]
We have \( |P|=b_1 \ldots b_n \). In this paper we will call \( P \) an \( n \)-dimensional parallelepiped.

**Theorem 2.** — Let \( K \subseteq \mathbb{Z} \) and suppose that
\[ |K + K| < \sigma k \]
where
\[ k = |K| \geq k_0(\sigma) = \frac{[\sigma][\sigma + 1]}{2(\sigma + 1) - \sigma} + 1, \]
then there exists an \( n \)-dimensional parallelepiped, \( P \), such that \( n \leq [\sigma - 1] \) and \( |P| < ck \), where \( c \) depends only on \( \sigma \) and \( s \) and there also exists a map \( \phi: P \rightarrow \mathbb{Z} \) which is such that \( P \rightarrow \phi(P) \) is an \( \mathbb{F}_s \)-isomorphism while \( K \subseteq \phi(P) \).

Let us now return to §1. The equivalence relation that we talked about there, is now seen to be \( \mathbb{F}_s \)-isomorphism. A representative of an equivalence class is an \( n \)-dimensional parallelepiped, \( P \). We now understand that \( K \), a subset of the one-dimensional space \( \mathbb{R} \), has, in fact, a multidimensional structure, being a dense subset of an \( n \)-dimensional set \( P \) (i.e. \( \phi^{-1}(K) \subseteq P \)). Consider the numbers
\[ a = \phi((0, \ldots, 0)), \ a_1 = \phi((1, 0, \ldots, 0)) - a, \ldots, \ a_n = \phi((0, 0, \ldots, 1)) - a. \]
Then,
\[ \phi(P) = \{a + a_1x_1 + a_2x_2 + \cdots + a_nx_n, \text{ with } 0 \leq x_i \leq b_i - 1\}. \]
Imre Rusza has called such a set \( \phi(P) \) a generalized arithmetic progression of rank \( n \). He gave a new and shorter proof, based on new ideas, of the main theorem together with an important generalization; in this the summands \( A \) and \( B \) may be different, although however the condition \( |A| = |B| \) is required (see [233]). His generalization to the case of subsets of abelian groups is to be found in [238].
7. We can now describe an “algorithm” for solving an inverse additive problem, by the following steps.

(i) Choose some (usually numerical) characteristic of the set under study.
(ii) Find an extremal value of this characteristic within the framework of the problem that we are studying.
(iii) Study the structure of the set when its characteristic is equal to its extremal value.
(iv) Study the structure of a set when its characteristic is near to its extremal value.
(v) (vi),... continue, taking larger and larger neighbourhoods for the characteristic.

From estimates obtained by Yuri Bilu it follows that in (5) we can take, for \( \sigma \), the following very slowly growing function of \( k \),
\[
\sigma = c \log \log \log \log k.
\]

It will be very important to study the cases
\[
\sigma = (\log k)^\epsilon \tag{6}
\]
and
\[
\sigma = k^\epsilon, \ \epsilon > 0, \tag{7}
\]
even if \( \epsilon \) is a very small number.

Here to simplify this extremely difficult problem a little, it is better to take \( |rK| \) as a characteristic value, where \( r \) is a fixed, positive, but rather large, integer. So our condition is now
\[
|rK| < k^{1+\epsilon}
\]
which is much stronger than (5); \( rK \) contains \( k^r \) sums, but no more than \( k^{1+\epsilon} \) of them are different.

8. I have here added a playful description of the comparative difficulty of the problems discussed, which should not be taken too literally. To prove (2) took one minute. Condition (3) was studied in three minutes. The proof of the theorem of §5 together with the description of \( K \) under the condition \( |2K| = 3k - 3 \) took one month. Proof of the main theorem took five years. I will be very happy if we will see results for (6) in the next thirty years but I am not certain that for (7) we will have satisfactory results even in the next hundred years.

9. L. Schnirelman [242] was one of the first who passed from studying fixed sets to studying general additive properties. Schnirelman introduced the notion of the density of a sequence.

**Definition.** — Let \( A = (a_1,a_2,\ldots,a_n,\ldots) \) be an increasing sequence of positive integers and further let,
\[
A(x) = |\{y \in A \mid 0 < y \leq x\}|,
\]
and
\[
d(A) = \inf_{x \in \mathbb{N}} A(x)/x.
\]
The number \( d(A) \) is called the Schnirelman density of the sequence \( A \) (see step (i) of §7).
10. Define

\[ A + B = \{a + b \mid a \in A, \ b \in B\} \]

and denote

\[ \alpha = d(A), \ \beta = d(B), \ \gamma = d(A + B). \]

Schnirelman proved that

\[ \gamma \geq \alpha + \beta - \alpha \beta. \]

L. Schnirelman and E. Landau conjectured in 1932 and Mann [178] has proved in 1942 that

\[ \gamma \geq \alpha + \beta. \]  \hspace{1cm} (8)

11. The famous \( \alpha + \beta \) theorem of Mann cannot be improved. Take a sequence

\[ A = \{0, 1, \ldots, r, l + 1, l + 2, \ldots, l + r, 2l + 1, 2l + 2, \ldots, 2l + r, \ldots\} \]

It is clear that if \( r \leq l \) then,

\[ \alpha = d(A) = r/l. \]

However if \( 2r < l \) then

\[ \gamma = d(2A) = 2r/l = 2\alpha. \]

But for \( A = B \) we always have from (8) that \( \gamma \geq 2\alpha. \) So step 2 of §7 is now completed.

Thus Mann has entirely solved the problem of increase of the density under summation of sequences. Its solution took ten years. Khinchine [151] writes in his book:

"The problem has become ‘fashionable’. Scientific societies proposed a prize for its solution. My friends from England wrote me in 1935 that half of English mathematicians tried to solve it, putting aside all other obligations"

When Mann had solved the problem, the interest in these subjects disappeared. But what about proving the inequality \( \gamma \geq 3\alpha? \) Or, equivalently, what are the sequences \( A \) for which \( \gamma < 3\alpha? \) These questions were not asked.

12. However, Schnirelman density is not a good characteristic. Take \( A = \{2, 3, 4, \ldots\}. \)

For this sequence we have \( A(1) = 0 \) and \( d(A) = 0. \) We feel, however, that the value 1 would be more appropriate for a density. So we arrive at a notion of an asymptotic density:

\[ d(A) = \lim \inf_{x \to \infty} A(x)/x. \]

In 1953 Martin Kneser [153] proved an analog of the \( \alpha + \beta \) theorem for asymptotic densities. He described the structure of \( A \) and \( B \) in the case when

\[ d(A) + d(B) < d(A + B). \]

Recently Yuri Bilu analysed the case when

\[ d(A + A) \leq \sigma d(A), \]

where \( \sigma \in [2, 5/2]. \)

To prove his theorem Kneser had to consider, for some positive integer \( g, \) sets of residues \( A \) and \( B \) modulo \( g \) for which

\[ |A + B| = |A| + |B| - 1. \]
Cauchy [38] and Davenport [50] have proved that if \( A \subseteq \mathbb{Z}_p \) and \( B \subseteq \mathbb{Z}_p \), where \( p \) is a prime, then
\[
|A + B| \geq \min(p, |A| + |B| - 1).
\]
This inequality is analogous to (8).

Vosper [257] proved that if \( A, B \subseteq \mathbb{Z}_p \), \( |A| + |B| - 1 < p - 2 \) and \( \min(|A|, |B|) > 2 \) then from \( |A + B| = |A| + |B| - 1 \) it follows that \( A \) and \( B \) are arithmetic progressions in \( \mathbb{Z}_p \) with the same difference.

Theorems of Kneser, Cauchy-Davenport and Vosper were amongst the first results giving solutions of inverse additive problems.

13. We may ask, are there any applications of the ideas and results described in §§4–8? For an answer to this question we turn now to the extremal combinatorial problems of Paul Erdős.

We begin with the problem raised by Erdős and Freud [68]. Fix some positive integer, \( \ell \). Denote by \( A \) a set of \( x \) natural numbers, \( \{a_1, a_2, \ldots, a_x\} \), with \( 1 \leq a_1 < a_2 < \cdots < a_x \leq \ell \). Take the set, \( A_0 = \{3, 6, 9, \ldots, 3 \left\lfloor \frac{\ell}{3} \right\rfloor \} \). For each subset \( B \subseteq A_0 \) the sum of elements in \( B \), the subset sum, is divisible by 3 and thus not equal to any power of 2. In this case \( |A_0| = \left\lfloor \frac{\ell}{3} \right\rfloor \).

However if we take \( |A| > \left\lfloor \frac{\ell}{3} \right\rfloor \) then for sufficiently large \( \ell \) there exist \( B \subseteq A \) and \( s \in \mathbb{N} \) such that \( \sum_{a_i \in B} a_i = 2^s \). This was proved in [70]. E. Lipkin [167] proved that, for sufficiently large \( \ell \), a set of maximal cardinality, none of whose subset sums is equal to a power of two, must be exactly the set \( A_0 \).

The desired result was achieved with the help of analytical methods. However, there was a difficulty — how to apply them to prove a result which is valid for some integer, say, \( \left\lfloor \frac{\ell}{3} \right\rfloor + 1 \), but is not valid for an integer which is one less. To cope with this, some conditions were formulated, so that when satisfied an analytical treatment could be used. The case where these conditions were not fulfilled was treated as an inverse additive problem. The structure of such sets was thus determined and it then became possible to finish the proof. (For more details, see §28.)

One might think that the problem of representing powers of two by subset sums is rather special, even artificial and therefore not that interesting. But, Paul Erdős knows how to ask questions. Ideas developed in order to solve the problem explained here, have turned out to be sufficient to solve a wide range of problems in Integer Programming, see §23 and [41]–[44].

14. In the framework of the problem of the previous section we may ask the following questions.

1) Let \( |A| > \left\lfloor \frac{\ell}{3} \right\rfloor \). What is the minimal cardinality \( |B| \) of \( B \subseteq A \), whose subset sum is equal to some power of 2?

2) What is the minimal number of summands required in the representation of a power of 2, if equal summands are allowed?

These questions were asked and answered in a paper of M. Nathanson and A. Sárközy [201]. The sufficient number of summands required was estimated to be at most 30360 and 3503, respectively. Using the Theorem of §5 it appeared to be possible to improve these estimates to 8 and 6, respectively (see [104]). We will here briefly
explain the main ideas. If we apply the Theorem of §5 to some set $A \subseteq [1, \ell]$, then under doubling the number of elements is multiplied, roughly, by 3 and the length of the segment where the sum $2A$ is situated is multiplied by 2. So, the density is multiplied, roughly, by $\frac{3}{2}$. After the doubling is repeated twice, the density of $4A$ will be $\geq \frac{1}{3} \cdot \frac{3}{2} \cdot \frac{3}{2} = \frac{3}{4}$. One more doubling (or more accurately summing $4A + 2A$) will give a long interval, in $8A$ (or even in $6A$), containing then some power of 2.

Noga Alon gave a simple example showing that 4 summands in the case of different and 3 summands in a case of possibly repeating summands are not, in general, sufficient. Recently, Vsevolod Lev [160] found the exact number of summands, in a case of possibly repeating ones. He showed that four summands are sufficient.

The following questions are of interest.

1) For given $|A|$ and $s$, find, $f(|A|, \ell, s)$, the minimum over all sets $A \subseteq [1, \ell]$ of order $|A|$, of the maximal length arithmetic progression contained in $sA$.

2) For given $|A|$ and $L$, find, $f(|A|, \ell, L)$, the maximum over all sets $A \subseteq [1, \ell]$ of order $|A|$, of the minimum number of summands, $s$, such that $sA$ contains an arithmetic progression of length $L$.

15. Denote by $s^A$ the set of integers which can be written as a sum of $s$ pairwise distinct elements from $A$. The set $A$ is called admissible if, and only if, $s \neq t$ implies that $s^A$ and $t^A$ have no element in common.

E.G. Straus [247] showed that the set $\{N - k + 1, N - k + 2, \ldots, N\}$ is admissible if, and only if, $k \leq 2\sqrt{N + \frac{1}{4}} - 1$. He proved that for any admissible set $A \subseteq [1, N]$ we have $|A| \leq (4/\sqrt{3} + o(1))\sqrt{N}$. The constant involved was slightly reduced by P. Erdős, J-L. Nicolas and A. Sárközy (cf. [75]). In the paper of J-M. Deshouillers and G. Freiman [52] (see also [51]) Erdős’ conjecture was proved, at least when $N$ is sufficiently large.

**Theorem 3.** — There exists an integer $N_0$ such that for any integer $N \geq N_0$ and any admissible subset $A \subseteq [1, N]$ we have,

$$|A| \leq 2\sqrt{N + \frac{1}{4}} - 1.$$ 

The proof was obtained with the help of methods of the type quoted in §5.

16. Let $A \subseteq [1, n]$. If $A \cap (A + A) = \emptyset$, the set $A$ is called sum-free. P. Erdős and P.J. Cameron conjectured that for the number $I_n$ of sum-free sets we have,

$$I_n = O(2^{n/2}). \quad (9)$$

The typical example of sum-free set $A \subseteq [1, n]$ is the set $\{1, 3, 5, \ldots\}$ of odd numbers. We can show that $\left\lfloor \frac{n+1}{2} \right\rfloor$ is the maximal cardinality of a sum-free set.

In G. Freiman [101] and the paper of J-M. Deshouillers, G. Freiman, V. Sos and M. Temkin [54], the problem of structure of sum-free sets was raised and studied. It was solved in the case of large cardinality of $A$, namely, when $|A| > 0.4\ell - c$, where $c$ is some positive constant. An example of such a structure is one in which all the elements of $A$ are congruent to 2 or 3 modulo 5.
The structure of \( A \) having been found, the estimate (9) for this class of \( A \), now follows immediately. An open question is to describe the structure of \( A \) for smaller cardinalities.

17. In the paper of G. Freiman, L. Low and J. Pitman [106], the following conjecture of Erdős and Heilbronn [73] is proved for sufficiently large primes. For \( A \subset \mathbb{Z}_p \), where \( p \) is a prime, \(|A| = k < p/50 \) and \( k > 60 \), we have

\[
|A + A| \geq 2k - 3.
\]

Also, the structure of \( A \) was described in the case when \(|A + A| < 2.06k - 3\). The conjecture of Erdős and Heilbronn was proved independently by J.A. Dias da Silva and Y.O. Hamidoune, see [246].

18. In the paper of A. Yudin [261], an example of large sets of integers, \( A \), was constructed for which

\[
|A + A| < |A - A|^c
\]

where \( c = 0.756 \). The previous example [113] gave only \( c = 0.89 \). In [113] the estimate \( c \geq 0.75 \) was proved. The result of A. Yudin puts the important additive characteristic,

\[
\liminf \frac{\log |A + A|}{\log |A - A|} = \alpha,
\]

in a very narrow interval, \( 0.75 \leq \alpha \leq 0.756 \), and allows one to begin to study the structure of sets with values of \( c \) which are close to \( \alpha \). Possibly the example of Yudin is not far from an extremal structure (look at §7).

19. In the paper of E. Lipkin [169], the Diderich conjecture [62] was studied. We now describe the conjecture. Let \( G \) be a finite Abelian group, \( A \subset G \) with \( 0 \notin A \). Let \( A^* \) denote the set of subset sums of the set \( A \). G.T. Diderich called the minimal number \( n \) such that, if \(|A| \geq n\) then \( A^* = G \), the critical number, \( c(G) \) of the group \( G \).

Let \( G \) be an Abelian group of odd order \(|G| = ph\) where \( p \) is the least prime divisor of \(|G|\) and \( h \) is a composite integer. Diderich conjectured, and E. Lipkin proved for \( G = \mathbb{Z}_q \) when \( q \) is sufficiently large, that

\[
c(G) = p + h - 2.
\]

20. In §§21–27 we will give a few examples of problems in different fields which may be looked at and treated as Structure Theory problems. These examples will be chosen from Additive Number Theory (§21), Combinatorial Number Theory (§22), Integer Programming (§23), Probability Theory (§24), Coding Theory (§25), Group Theory (§26) and Mathematical Statistics (§27). Our aim is not so much to enumerate these problems as to show how ideas and methods of Structure Theory may influence their solution and to show their interdependence. Not many examples are chosen and they do not cover the whole stock of related problems.

21. Additive Number Theory. We now present a paper (see [109]) of G. Freiman, H. Halberstam and I.Z. Ruzsa. This paper confronts the problem of how to show that, starting from some set of integers \( A \), the set \( rA \) contains an arithmetic progression of integers of length, \( L \), and difference, \( d \).
One obvious set of sufficient conditions is as follows. Firstly, that the set \((r - 1)A\) contains an arithmetic progression of length \(\ell\) and difference \(d\). Further that in some arithmetic progression of integers of length \(L + 2\ell\) and difference \(d\), we have that every part of it which forms an arithmetic progression of length \(\ell\) contains a number from \(A\).

These conditions are very simple and satisfactory but, how may one find such an arithmetic progression of length \(\ell\), even if \(\ell\) is much smaller than \(L\)? It is supplied by results of the paper mentioned! The final result is given below.

**Theorem 4.** — Let \(B\) be an infinite set of integers such that \(\Delta_B(N) \equiv \frac{B(N)}{N} > (\log N)^{-\alpha}\) for every integer \(N > N_0\), where \(\alpha\) is some fixed number in the interval \((0, \frac{1}{3})\), and \(N_0 = N_0(\alpha)\). Suppose further that \(B\) has the following "local" property.

Corresponding to each \(N > 12N_0\) there exists an integer \(M\) with \(N_0 \leq M < \frac{1}{12}N\), such that every arithmetic progression modulo \(q\) in \([1, N]\) of length \(\lfloor \frac{1}{2}A(M) \rfloor\) contains an element of \(B_N := B \cap [1, N]\), where \(2 \leq q \leq M\) and

\[
A(M) = e^{\frac{1}{2} C_0 \log M} (\log M)^{1-3\alpha}.
\]

Then \(B\) is an asymptotic basis of order 4.

The first version of this paper was built on methods of [82] and [105], but later changed to methods of [233], proposed by I. Rusza in his proof of the main theorem. The results of [109] were improved by Bourgain [21].

22. **Combinatorial Number Theory.** See examples given in §§13–19.

23. **Integer Programming.** Let us discuss problems connected with one linear equation,

\[
a_1x_1 + a_2x_2 + \cdots + a_mx_m = b. \tag{10}
\]

Suppose that the coefficients in (10) are positive integers, with \(a_1 < a_2 < \cdots < a_m < \ell\), and we wish to find a solution in the Boolean case with \(x_i \in \{0, 1\}\). Remember that we are dealing here with problems which we would not be encountering in Number Theory. We have to find an algorithm with the help of which a computer has to be able, in a reasonable time, to answer the question, whether or not there exists a solution and then, to find it. And a most important point must be borne in mind, namely that the algorithm has to achieve this task for *any* choice of coefficients in a given range. The number of unknowns in (10) is equal to \(m\), and each unknown may take two values, so the number of possibilities to check, if we decided to do it, is \(2^m\). Existing methods (branch and bound, partial enumeration, etc.) try to diminish this number but progress has been slow. If the coefficients \(a_j \in [1, \ell]\) and \(\ell = 10^{12}\), say, then \(m\) has to be not bigger than about 100 or 200 for the equation to be solved by today’s computers. The dynamic programming approach gives times of \(O(\ell^m)\). If, for example, \(m = 10^6\) the time is of order \(10^{24}\) verifications, too long to see results in our lifetime.

A different approach to the problem was outlined in [96]. We began to study the structure of the set of values of a linear form, using Analytic Number Theory. This structure appeared to be rather simple, it is in essence, the union of several arithmetic
progressions with the same difference. To characterize an arithmetic progression we have to know its difference $d$, its first member and its length.

The time required to answer a question of solubility of an equation is $O(m)$ and in our example it is of order $10^6$ verifications, a matter of seconds. The main idea is explained in §28. For detailed exposition and literature see a review of Mark Chaimovich [42] and a paper [43].

24. Probability Theory. Estimates for concentration functions and local limit theorems — these are two domains where today there exist applications of the Structure Theory approach to Probability Theory.

Let $\xi_1, \ldots, \xi_n$ be a sequence of independent identically distributed random variables taking values in $\mathbb{Z}$. Further, let $s_n = \sum_{j=1}^{n} \xi_j$. Define

$$Q\xi(\ell) = Q(\ell) = \sup_{x} P(x \leq \xi < x + \ell),$$

the concentration function of the random variable $\xi$, and let $Q_{s_n}(\ell) = Q_n(\ell)$ be the concentration function of $s_n$.

The paper of J-M. Deshouillers, G. Freiman, A. Yudin [55], gives a new estimate for $Q_n(1)$. Previous results, see for example G. Kesten [150], give an estimate of the type

$$Q_n(1) \leq \frac{c}{n^{2/3}}$$

where $c$ is independent of $n$. In this estimate the exponent $1/3$ cannot in general be replaced by a larger number. Indeed, let us fix some integer valued random variable with variance $\sigma^2$. Then by the local limit theorem we have

$$P\{s_n = N\} = \frac{1}{\sigma \sqrt{2\pi n}} \left( \exp \left( - \frac{(\mu n - N)^2}{2n\sigma^2} \right) + o(1) \right).$$

From here we see that the estimate (11) cannot, in general, be improved. If we want to improve (11) we have to impose additional conditions and this is what is done in [55].

**Theorem 5.** — Let $\sigma \in \left( \frac{\log 4}{\log 3}, 2 \right)$, $\epsilon > 0$, $A \geq 1$ and $a > 0$ be given real numbers. Let $n$ be a positive integer and let $\{X_1, \ldots, X_n\}$ be a set of independent identically distributed integral random variables such that

$$\max_{q \geq 2} \max_{s(\text{mod } q)} \sum_{\ell \equiv s(\text{mod } q)} P\{X_1 = \ell\} \leq 1 - \epsilon,$$

$$\forall L \geq A : Q(X_1; L) \leq 1 - aL^{-\sigma}.$$  

Then we have

$$Q(S_n; 1) \leq cn^{-1/\sigma},$$

where $c$ depends at most on $\sigma, \epsilon, A, a$ and $Q(X_1; 1)$.

We have here two conditions, one excludes the case when the support is a part of some class mod $q$, $q \geq 2$ and the second asks for the tail to be 'heavy'. Conditions of both types are necessary to get results of the form of the Theorem above. In the first
version of a paper [55] the condition of type 1 was formulated for a series of random variables as follows. For any \( q \in \mathbb{Z}, q \geq 2 \)

\[
\max_r \sum_{k \equiv r \pmod{q}} p_k < 1 - 10 \sqrt{\frac{\ln n}{n}}.
\]

Let us also stress that the result of Esseen, cited in [55], gives a condition from which the concentration may be estimated from below. All these results give us the possibility to begin to study the distribution of a given random variable \( \xi \), if we know something about the value of \( Q_n(1) \), for example if we know that

\[
Q_n(1) \asymp \frac{1}{n^\theta},
\]

where \( \frac{\ln 4}{\ln 3} < \theta \leq 2 \). We can ask the same question for series. In this case we have to describe distributions where numbers \( a_i \) and numbers \( p_i \) may depend on \( n \).

25. Coding Theory. This section and §35 were written jointly with A. Yudin. The connection between coding theory and structure theory was shown by Zemor (see [262] and [263]) and Cohen & Zemor (see [265], [266], [46] and [47]). We will now try to explain that the main problems of coding theory are, in fact, inverse additive problems.

Let \( A = \{a_1, \ldots, a_k\} \) be a word in an alphabet of 2 symbols, say, \( a_i \in \{0,1\} \). Let \( A_n \) be the set of all words in this alphabet of length \( n \), so that we have \( |A_n| = 2^n \). The distance, \( g(x,y) \), between two words \( x = \{x_1, x_2, \ldots, x_n\} \) and \( y = \{y_1, y_2, \ldots, y_n\} \) is defined to be

\[
g(x,y) = \left| \{ i \mid x_i \neq y_i, \quad i = 1, \ldots, n \} \right|,
\]

that is, the number of positions in which the symbols in the words \( x \) and \( y \) differ. It is not difficult to check that \( g(x,y) \) satisfies all the axioms for a distance function. The question is how to ensure the correction of possible errors during transmission of information?

Consider some subset, \( U \), of the set of all words \( A_n \). Such a subset is called a code. A portion of information has assigned to it some word from \( U \) which is then transmitted through the channel. If during the transmission only a small number of mistakes occurred then we are still not far from the code word which was transmitted and thus we can then restore it. Let us put this question in a more precise formulation. We let the word transmitted be \( x = \{x_1, \ldots, x_n\} \) and the word received be \( \tilde{x} = \{\tilde{x}_1, \ldots, \tilde{x}_n\} \). If during the transmission of a word through a channel no more than \( t \) mistakes take place, it means that

\[
g(x,\tilde{x}) \leq t
\]

and so it is necessary that \( \tilde{x} \) be closer to \( x \) than to any other word in the code. That is, for any \( y \in U \) with \( y \neq x \), we have to ensure that

\[
g(y, \tilde{x}) > t.
\]

By the triangle inequality

\[
g(x,y) \leq g(x,\tilde{x}) + g(\tilde{x},y),
\]
and when
\[ g(x, y) > 2t, \]  
we can obtain (13) from the inequality (12).

If there exists \( y \) such that \( g(x, y) = 2t \), then we can find \( \tilde{x} \) for which (12) and (13) become equalities and then \( g(\tilde{x}, y) = t \). Thus, the condition (15) is necessary and sufficient for code correcting \( t \) mistakes. We have a set, \( A_n \), and a subset \( U \), but to speak about inverse additive problem is still premature, since an algebraic operation is missing. So we will consider \( A_n \) as a vector space over the field \( \mathbb{Z}_2 \). In this field \(-1 = 1\) and for each \( n \)-dimensional vector \( x \in A_n \) the equality \(-x = x\) holds. The distance \( g(x, y) \) is equal to the number of 1s in the vector \( x - y = x + y \), i.e. to the distance of the element \( x + y \) from 0. The condition (15) may now be written as
\[ g(x + y, 0) > 2t. \]

Thus, a code, correcting \( t \) mistakes, is a \( U \subset A_n \) such that \( \forall z \in 2U \) we have \( g(z, 0) > 2t \). We have now come to a well known situation, namely, we have a group \( A_n \), a subset \( U \) and a condition on \( 2U \).

In §12 the first results about sums of sets in a group were mentioned. The doubling of sets in groups was studied in the works of Kemperman [146], [147], [148], Freiman [83], Olson [207], [208], [209], Brailovsky & Freiman [27], [29], Brailovsky [22–25] and Hamidoune [124–137]. If \( n \) is a minimal number such that for \( A \subset G \) we have \( nA = G \), \( A \) is called a basis of \( G \) of order \( n \). This theme is reviewed in [9] and [140].

What are the main aims which we are trying to achieve in coding? Atoms of information are transmitted by words of code. Thus, if the quality of a code is fixed, i.e. the number of mistakes to be corrected is fixed, then the code will be the better, the greater the cardinality of the code \( U \). And conversely, if the number \( |U| \) is given, how do we choose the best code?

We shall reiterate the formulation of the two problems mentioned above. Let \( U \subset A_n = \mathbb{Z}_2^n \) for some fixed \( n \in \mathbb{N} \). Assume that for all \( z \in 2U \)
\[ g(z, 0) \geq d, \]  
where \( d \in \mathbb{N} \).

**Problem I.** Let \( d \) be fixed. What is the maximum value of \( |U| \) for which (16) is valid?

**Problem II.** Let \( |U| \) be fixed. What is the maximum value of \( d \) for which (16) is valid for some \( U \) of order \( |U| \).

We have formulated two inverse additive problems which are the major problems of coding theory but are, in essence, not yet solved satisfactorily. In a paper of Gerard Cohen and Gilles Zemor [47] other inverse additive problems are presented and their connection with coding theory is explained.

**26. Group Theory.** Results in group theory are reflected in the reviews of M. Herzog [140] and Y. Berkovich [9] and the bibliography to this review. We try now to find an example where our approach gives some progress on a theme which was investigated earlier in group theory.

For a set
\[ \{a_1, a_2, a_3\} \]  
(17)

SOCIÉTÉ MATHÉMATIQUE DE FRANCE 1999
of elements of a group \( G \), we build all the products,

\[
a_1a_2a_3, \ a_1a_3a_2, \ a_2a_1a_3, \ a_2a_3a_1, \ a_3a_1a_2, \ a_3a_2a_1. \tag{18}
\]

If at least one product in (18) is equal to another one, the set (17) is called \emph{rewritable}. If every 3-element set in \( G \) is rewritable, then \( G \) is called a \emph{rewritable group}, that is \( G \in Q_3 \), where by \( Q_3 \) we denote the class of rewritable groups. If every product in (18) is equal to some other product, then the set (17) is called \emph{totally rewritable}. If every set (17) in \( G \) is totally rewritable, then \( G \) is said to be a \emph{totally rewritable group}, written \( (G \in P_3) \).

The problem is to describe all groups in the classes \( P_n \) and \( Q_n \). See Kaplansky [145], Blyth & Robinson [19], Freiman & Schein [117] and [118], Longobardy & Maj [170], [171] and [172].

The main tool to use in this study is a notion of ‘permutational isomorphism’, a realization of the equivalence relation we talked about in §1. This notion is somewhat different from the one introduced in §6, but it is suited very well to the study of this particular problem.

A \emph{permutational isomorphism} of \( A \) onto \( B \) (where \( A \subset S \) and \( B \subset R \), while \( S \) and \( T \) are two sets with binary operations) is a pair of bijections \( \varphi: A \rightarrow B \) and \( \psi: A^{[3]} \rightarrow B^{[3]} \) such that for all pairwise distinct elements \( a_1, a_2, a_3 \in A \) we have

\[
\psi(a_1a_2a_3) = \varphi(a_1)\varphi(a_2)\varphi(a_3).
\]

Here \( A^{[3]} \) is the set of all products of triples of distinct elements.

To begin our approach we have only to pay attention to the fact that amongst the six products in (18) there are no more than five distinct ones, if the set (17) is rewritable. Thus, we take as a numerical characteristic, \( r \), the maximal number of different products for all sets (17) in a group \( G \). We thus obtain the classes of groups \( P(3,r) \) for \( 1 \leq r \leq 6 \) (see Freiman & Schein [117]). In [117] all classes of isomorphic triples, 19 classes in all, were obtained and then used to study the classes \( P(3,r) \).

Similarly one can define the classes of groups \( P(4,r) \) of which there are 24. It turns out that \( P_3 = P(3,2) \) (see [117]). In [118] the class \( P(3,3) \) was described. G. Freiman, D. Robinson and B. Schein [115] partially described the class \( P(3,4) \). The next step is the study of \( P(3,5) = Q_3 \).

27. \textbf{Mathematical statistics.} Let \( F = \{f_i\}_{i=0}^n \) be a set of continuous functions on \([a,b]\), and let \( F^* = \{f_if_j\}_{i,j=0}^n \). In the paper of B. Granovsky and Eli Passow [120] conditions were determined for the set \( F^* \) to consist of exactly \( 2n + 1 \) distinct functions. The additional requirement is that \( F^* \) has to be a Chebyshev system on \([a,b]\).

A set \( \{u_i\}_{i=0}^n \) of continuous functions on \([a,b]\) is said to be a \emph{Chebyshev system} on \([a,b]\) if every nontrivial “polynomial” \( \sum_{i=0}^n g_iu_i(x) \) has at most \( n \) zeros on \([a,b]\). The number \( n + 1 \) is called the degree of the Chebyshev system. In [120] necessary and sufficient conditions were given on the set \( \{f_i\}_{i=0}^n \) so that the set \( \{f_if_j\}_{i,j=0}^n \) is a Chebyshev system of minimal degree \((2n + 1)\). These results have applications to the field of experimental design. See also I. Efrat [66], Kiefer & Wolfowitz [152] and E. Passow [213].
It is clear how this problem can be formulated as a problem of small doubling of a set of real numbers. Given \( n + 1 \) functions pick some fixed argument \( x_0 \). Consider the \( n + 1 \) numbers \( \{ f_i(x_0) \}_{i=0}^{n} \). Leaving for further investigation the case when they are not all distinct, or some of them are not positive, we have the set \( D \), of logarithms of these numbers, \( D = \{ \log f_i(x_0) \}_{i=0}^{n} \) subject to the condition \( |2D| = 2n + 1 \). So \( D \) is a set with small doubling and it is very simple to show, not only for integers but also for real numbers, that \( D \) is an arithmetic progression. I. Efrat [66] has used the results of Theorem 1 and described all Chebyshev systems with \( |F^*| < 3n \).

28. In this section we want to point out the unity of approach and similarity of methods when different problems are treated from the point of view of Structure Theory.

In Combinatorial additive problems we mainly study finite sets of integers. In many of such problems the theorems of §§5 and 6 about a structure of sets of integers with small doubling may be applied directly. In §§13–19 such results were given. These theorems may also be applied to sets in other algebraic systems, such as \( \mathbb{Z}_p \), see [88], \( \mathbb{T}^1 \), see [197], \( \mathbb{R}^k \), see [82], page 94, and to functional spaces, see [66]. The sets in \( \mathbb{Z} \) may be infinite, see [91] and [82]. The structure of sets with a small product in a nonabelian torsion-free group, see [26], is described with the help of methods developed to prove Mann’s theorem.

To solve inverse problems of additive number theory, analytical methods are used. They reveal some unity and similarity when applied to the study of different problems, see §30. Problems in number theory of the evaluation of measure and of the determination of the structure of sets with large trigonometric sum, see [260], [100], [13], and in probability theory, of sets with large characteristic function, see [197] and [55], are often studied by similar methods.

A tool of investigation which can be used in many situations, may be called “multiple use of structural argument”. To ensure the existence in Integer Programming, of a solution of an equation (10), see [96], we assume a condition on \( A(q) \equiv \{ x \in A \mid q|x \} \), namely

\[
|A(q)| < |A| - |A|^\delta, \tag{19}
\]

where \( \delta < 1 \) is independent \( q \). In analytical number theory it is usual to place such a uniformity condition on the distribution of residues. When it does not hold, the case is not studied. However let us now consider the case when (19) is not valid. Then there exists \( q > 1 \) such that

\[
|A(q)| \geq |A| - |A|^\delta.
\]

This is a very strong condition to impose on the structure of \( A \) and so we can continue our analysis and describe the structure in full. In papers [56] and [197], where problems in probability were studied, a condition of the type (19) is present. This observation opens up the possibility to obtain new results, stronger than those in [56] and [197].

The very notion of a set with small doubling, when brought to group theory, resulted in the appearance of new problems.
The notion of isomorphism which was introduced in the course of proving the main theorem (§6) became a useful tool. In group theory, it provided the possibility of building an equivalence relation on finite sets, describing its equivalence classes and then studying the property of a group in connection with the existence or nonexistence of some classes in this group. In rewritable groups, see §26, a version of isomorphism was given suited to the purpose. In [53] a notion of isomorphism for random variables was introduced, which gave the possibility of describing the behavior of a one-dimensional random variable with the help of a multidimensional one.

29. First results about the structure of sets with small doubling were obtained with the help of elementary methods. Afterwards, the analytic methods were introduced. In fact, there exists an exact dividing point. If $|K + K| \leq 3k - 3$, then the elementary approach very quickly gave a full description of $K$. For larger values the elementary methods did not give results in spite of big efforts.

Very little has been done to get elementary results in the multidimensional case. In [82] the case on the plane of $|K + K| < \frac{10}{3}k - 5$ is studied and I. Stanchescu studied the case $|K + K| < (4 - \varepsilon)k$. I don’t know the range of the doubling coefficient $C_n$ in an inequality $|K + K| < C_n k$, where $K \subset \mathbb{Z}^n$ for which elementary results may be obtained.

To obtain here a clear picture is very desirable and not very difficult. Then it can be used to make the results of the main theorem more precise. Results for doubling coefficients $\frac{10}{3}$ and $4 - \varepsilon$ show that the structure of $A$ after it becomes multidimensional may be described more accurately with the help of elementary methods.

Many interesting problems arise from a study of $K$ when two, or more, numerical characteristics are given. A long list of invariants is given in [82], page 41.

30. In direct problems of additive number theory one is usually studying an integral which yields the number of representations of a number expressed as a sum of terms of a certain type. Further, a transform of this integral yields an asymptotic formula for the number of representations. Characteristic of the analytic method in Structure Theory is the fact that an integral with a known value serves as a starting point.

**Examples**

(i) (See Roth [224].) Sets $A$ without arithmetic progressions of length three. We have

$$\sum_{x \in A} \sum_{y \in A} \sum_{z \in A} \int_0^1 e^{2\pi i \alpha(x+y-2z)} d\alpha = |A| = \int_0^1 S^2 S_1 d\alpha,$$

where

$$S = \sum_{x \in A} e^{2\pi i ax}, \quad S_1 = \sum_{x \in A} e^{-4\pi i ax}.$$

(ii) A set $K$ with small doubling (see Freiman [82]). Here

$$\sum_{x \in K} \sum_{y \in K} \sum_{z \in 2K} \int_0^1 e^{2\pi i \alpha(x+y-z)} d\alpha = \int_0^1 S^2 S_1 d\alpha = |K|^2,$$
where
\[ S = \sum_{x \in K} e^{2\pi i \alpha x}, \quad S_1 = \sum_{x \in 2K} e^{-2\pi i \alpha x}. \]

(iii) Sum-free sets. We have
\[ \sum \sum \sum \int_0^1 e^{2\pi i \alpha (x+y-z)} \, d\alpha = \int_0^1 S_2 S \, d\alpha = 0, \]

where
\[ S = \sum_{x \in A} e^{2\pi i \alpha x}, \quad A \subset [1,l], \quad l \in \mathbb{N}. \]

The next step is to obtain a large trigonometric sum for a certain value (sometimes, for several values) of the argument. Consider an example from Freiman [82], page 48. Let \( K \) be a set of residues modulo a prime \( p \). Then
\[ I = \sum \sum \sum_{x_1, x_2, x_3 \in K} e^{2\pi i \frac{a}{p} (x_1 + x_2 - x_3)} = \sum_{a=0}^{p-1} S_2 S_1 = k^2 p, \]

where
\[ S = \sum_{x \in K} e^{2\pi i \frac{a}{p} x}, \quad S_1 = \sum_{x \in 2K} e^{-2\pi i \frac{a}{p} x}. \]

Let \( T = |K + K| \) and assume that \( |S| < \frac{3}{5} k \) for every \( a \not\equiv 0(p) \) then
\[ |I| \leq k^2 T + \sum_{a=1}^{p-1} |S|^2 |S_1| \leq k^2 T + 3\frac{k}{5} \left( \sum_{a=0}^{p-1} |S|^2 \sum_{a=0}^{p-1} |S_1|^2 \right)^{1/2} \]
\[ = k^2 T + 3k \frac{k}{5} \sqrt{kp \cdot Tp}. \]

In the example just considered the conditions \( T < \frac{12}{5} k \) and \( k < \frac{p}{35} \) were assumed, from which it follows that \( |I| < k^2 p \), a contradiction. We have therefore proved that there exists \( a' \not\equiv 0(\text{mod}p) \) such that
\[ |S(a')| = \left| \sum_{j=0}^{k-1} e^{2\pi i \frac{a' j}{p}} \right| > \frac{3}{5} k. \]

The presence of a large trigonometric sum makes it possible to obtain data about the set \( A \) which can then be processed using elementary techniques.

31. In the first papers on sets with small doubling information about only one large trigonometric sum was used. In the proof of the main theorem we have used several, but finite number of large sums. The next step was to begin to study a set of all ‘large’ trigonometric sums. It was first done in 1973 in probability theory field, in the proof of local limit theorems (see D. Moskvin, G. Freiman & A. Yudin [197]). In this case we were dealing with the characteristic function of a lattice random variable,
\[ f(\alpha) = \sum_{k \in \mathbb{Z}} p_k e^{2\pi i \alpha k} \]
studying the measure and structure of the sets $E$, where the characteristic function is large.

The reasoning is, in short, as follows. We use the fact that, if for some $\alpha_1$ and $\alpha_2$ we have $|f(\alpha_1)| \geq 1-u$ and $|f(\alpha_2)| \geq 1-u$ then $|f(\alpha_1 + \alpha_2)| \geq 1-4u$. We take the set

$$E = \left\{ \alpha \mid |f(\alpha)| > 1 - \frac{\sqrt{\log n}}{n}, \ n \in \mathbb{N} \right\}$$

and begin to double, obtaining sets $2E$, $2^2E$, $2^3E$, .... If the measure is growing steadily we will cover the set $[0,1)$ very quickly, thus obtaining a contradiction. If at some stage we meet a set with small doubling, we will get a structure. For some $q \in \mathbb{N}$, the arguments $\frac{\alpha}{q}$, with $0 \leq p < q$, will be included in this structure which will lead to the conclusion that almost all the probability measure is concentrated in an arithmetical progression modulo $q$, which gives a contradiction.

32. We are naturally led to a study of sets with a large measure of large trigonometric sums.

Let $k$ be a positive integer and $u < k$ a positive real. For a set

$$K = \{a_1 < a_2 < \cdots < a_k\}, \quad a_j \in \mathbb{Z}, \quad 1 \leq j \leq k$$

let

$$S_K(\alpha) = \sum_{j=1}^{k} e^{2\pi i a_j}, \quad s_K(\alpha) = |S_K(\alpha)|,$$

$$E_{K,u} = \{\alpha \in [0,1), \text{ for which } s_K(\alpha) \geq k - u\}$$

and

$$\mu_K(u) = \mu(E_{K,u})$$

when $\mu$ is the Lebesgue measure on $[0,1]$.

**Problem.** — Find the set $K$ which maximizes $\mu_K(u)$ and find its maximal value.

We denote by $\mu_{\text{max}}(k,u)$ the supremum of $\mu_K(u)$ over all sets $K$ of size $k$. The first results on this problem were obtained by Freiman (see [95], page 144) and Yudin (see [260], page 163). I sketched an approach for solving the problem in [100]. In [13] A. Besser carried out and extended this plan very widely. He showed that up to the second order

$$\mu_{\text{max}}(k,u) = 2\beta \approx \frac{2\sqrt{6}}{\pi} \left(1 + \frac{5u}{8k}\right)$$

and $K_{ex}$ may be described, in the main case, as the union of an arithmetic progression of length $k_0 = k - \frac{5}{12}u$, symmetric around zero, and, for any non-zero integer $n$, an arithmetic progression of length

$$\frac{1}{2}k_n = \frac{u}{(\pi n)^2} \left(1 - \frac{(-1)^n}{2}\right)$$

centered around $\frac{n}{\beta}$.
We will try to explain from where the structure of $K_{ex}$ comes. If $a$ is small the term $e^{2\pi i a_j}$ has a value close to 1 if $a_j$ is small. That is why we take an arithmetic progression with difference 1 centered around 0. We have, for $\alpha > 0$,

$$s_k(\alpha) = \frac{\sin(\pi \alpha k)}{\sin(\pi \alpha)} \approx \frac{\pi \alpha k - \pi^3 \alpha^3 k^3 / 6}{\pi \alpha} = k - \frac{\pi^2 \alpha^2}{6} k^3.$$  

As $\alpha$ increases, $s_k(\alpha)$ decreases and reaches $k - u$ for $\alpha$ determined by

$$k - \frac{\pi^2 \alpha^2}{6} k^3 = k - u$$

that is,

$$\alpha^2 \approx \frac{6}{\pi^2} \frac{u}{k^3}$$

and thus

$$\alpha_0 \approx \frac{\sqrt{6}}{\pi} \frac{1}{k} \left( \frac{u}{k} \right)^{\frac{1}{2}}.$$  

Consider the trigonometric sum at this point $\alpha_0$. Our set is positioned on the segment $[-\frac{k}{2}, \frac{k}{2}]$. If we add another number, $\frac{k}{2} + 1$, to the arithmetic progression, the term $e^{2\pi i a_0(\frac{k}{2} + 1)}$ will be added to the trigonometric sum. If we add $[\frac{1}{\alpha_0}]$, then $e^{2\pi i a_0[1/\alpha_0]}$ will be closer to unity, it will lie in a smaller neighborhood of the $x$ axis and will influence the increasing of $S(\alpha)$ more critically. This consideration explains the appearance of segments near to the points $\frac{n}{\alpha_0}$.

33. An analysis of the remarkable results of A. Besser does not reveal an easy future. The set $K_{ex}$ is of a rather complex two-dimensional structure which becomes more complex as $n$ increases and will, in all likelihood, become multi-dimensional. The structure of $K_{ex}$ has only been found for very small values of $u$, $u < \frac{k}{32000}$ and an increase is only gained with some effort. Thus, further progress in the problem under consideration would be of great interest, but reaching it is very difficult.

The sets $K_{ex}$ found by Besser have small density for small $u$'s. But in many open problems the situation is different. For example, in the problem of sum-free sets, the density of the set to be considered is close to 0.4. When attempting to strengthen the theorem on the structure of $K$, with small doubling, outside the bounds $|2K| = 3k - 3$, we should begin by considering sets whose densities are close to 0.5. So, we state the problem on measure of large trigonometric sums as follows. Let $K$ be a set of integers in $[0,1]$ with $|K| = k$. Define

$$E_{K,m} = \{ \alpha \in [0,1), \text{ for which } s_K(\alpha) \geq m \}$$

and let $\mu_K(m) = \mu(K,m)$ denote the measure of $E_{K,m}$. Also we set

$$\mu_k(m,l) = \max_{K \subset [0,l]} \mu(K,m).$$

Then if $l = k - 1$, the problems is a trivial one. As $l$ increases, it becomes more complex. After the quantities $\mu_k(m,l)$ have been found, one should proceed to describe the structure of those $K$'s for which $\mu(K,m)$ does not differ greatly from $\mu_k(m,l)$.
34. In the problem on sum-free sets, the following integral was being considered, 
\[ \int_0^1 |S|^2 S \, d\alpha = 0, \]
and it follows from this that 
\[ \int_0^1 |S|^2 \Re(S) \, d\alpha = 0. \]
In a neighbourhood of zero the integrand is of order \( k^3 \) and its contribution to the integral is of order \( k^2 \). Since the integral over the whole interval equals zero, the measure of the set of \( \alpha \)'s where \( |\Re(S)| \) has order \( k \) and is negative, should be large.

We come to the following general problem. Let \( K \subset \mathbb{Z} \) with \( |K| = k \) and set 
\[ E_{K,-m} = \{ \alpha \in [0, 1), \text{ for which } |\Re(S)| \leq -m, \quad 0 < m < k \}. \]
Let \( \mu(K, -m) \) be the measure of \( E_{K,-m} \) and 
\[ \mu_k(-m,l) = \max_{K \subset [0,l]} \mu(K, -m). \]
The usual questions may be asked once again about the quantities \( \mu_k(-m,l) \) and about the structure of the set \( K \) for which the measure \( \mu(K, -m) \) is close to the maximal value. At the next, deeper stage of study, one may investigate combining two or more numerical characteristics. The first step here should be the study of trigonometric sums when some conditions are imposed not only on \( |S| \) but also on \( \arg S \).

35. Let \( G \) be an abelian group whose operation will be denoted by \( + \), and \( \hat{G} \) be the group dual to \( G \), that is the group of characters of \( G \). Let \( A \) be a subset of \( G \) and define a map 
\[ f_\chi: A \mapsto \sum_{a \in A} \chi(a), \quad \text{for } \chi \in \hat{G}, \]
that is, to the set \( A \) we correspond a function of a character \( \chi \in \hat{G} \).

As is shown in [82], from the fact that \( |2A| \leq C|A| \) in the case \( G = \mathbb{Z} \) it follows that the set on which \( |f(\chi)| \) is rather large has a large measure. With the help of methods from harmonic analysis we can describe the structure of the set \( A \).

It is important to stress that to the set \( A \) with small doubling from \( G \) corresponds a set 
\[ \hat{A}_\alpha = \left\{ \chi \in \hat{G} \text{ such that } |\sum_{a \in A} \chi(a)| > \alpha|A| \right\}, \quad \text{for } \alpha \in \mathbb{R}^+, \]
which also has small doubling. From the fact that \( \hat{G} = G \) we may, it seems, suppose that from \( B \subset \hat{G} \) and \( |2B| \leq C|B| \) it will follow that \( \hat{B} \subset G \) and \( |2\hat{B}| \leq C|\hat{B}| \). Note that the constants in different places of this section may differ. For a given additive problem it is possible to find the equivalent problem on the dual group and vice versa, and then to study the version which is preferable.

The following observations are also important. Suppose that \( A_1 \subset G \) and \( A_2 \subset G \) are sets which are structurally 'near' to each other. A natural question to ask is whether \( \hat{A}_1 \) and \( \hat{A}_2 \) are also 'near' to each other and what kind of topology is
induced by the correspondence $A \mapsto \hat{A}$. Again, from $\hat{G} = G$ it follows that these topologies induce one other. It would be very interesting to determine what kind of neighbourhoods they define and to what extent these topologies are `metrisable', because metric characteristics of these topologies will be of great interest during the study of problems of addition of sets.

The analytic tool in the case $G = \mathbb{Z}$ was the equality

$$\int_0^1 S^2S_1 \, d\alpha = |A|^2,$$

where

$$S = \sum_{x \in A} e^{2\pi i x} \quad \text{and} \quad S_1 = \sum_{x \in 2A} e^{-2\pi i x}.$$

In the case of a finite abelian group, $A$, we can write the parallel expression

$$\sum_\chi \left( \sum_{a \in A} \chi(a) \right)^2 \sum_{a \in 2A} \overline{\chi}(a) = |A|^2.$$

Generalization to the nonabelian case should also be studied.

36. I am greatly indebted to Dr. Ruth Lawrence and Mr. Harry Lawrence for their invaluable help in producing this manuscript.

References


[13] Besser A., *Sets of integers with large trigonometric sums*, this volume,


[16] Bilu Y., *Structure of sets with small sumset*, this volume,


[34] Brodsky S., *On groups generated by a pair of elements with small third or fourth power*, this volume,


Cameron P.J., Erdos P., On the number of sets of integers with various properties, in “Number Theory”, Banff, Alberta 1988 conference proceedings, de Gruyter Berlin (1990), 61–79.


Chaimovich M., Solving value–independent knapsack problem with the use of methods of additive number theory, Congressus Numerantium 72 (1990), 115–123.


Chaimovich, M., New structural approach to integer programming: a survey, this volume,

Chaimovich M., New algorithm for Dense Subset-Sum Problem, this volume,


Cohen G.D., Zemor G., Subset sums and coding theory, this volume


Deshouillers J-M., Freiman G.A., On an additive problem of Erdős and Straus II, this volume,

Deshouillers J-M., Freiman G.A., Moran W., On series of discrete random variables 1: Real trinomial distribution with fixed probabilities, this volume,

Deshouillers J-M., Freiman G.A., Sos V., Temkin M., On the structure of sum-free sets 2, this volume,

Deshouillers J-M., Freiman G.A., Yudin A., On Bounds for the Concentration Function, 1 this volume,


SOCIÉTÉ MATHÉMATIQUE DE FRANCE 1999


[106] Freiman G.A., Low L., Pitman J., *Sumsets with distinct summands and the conjecture of Erdős-Heilbronn on sums of residues*, this volume,
[114] Freiman G., *Structure theory of set addition*, this volume,
[136] Hamidoune Y.O., Subsets with a small product in groups, this volume.
[140] Herzog M., New results on subset multiplication in groups, this volume.
[158] Lev V.F., On the structure of sets of integers with small doubling property (|A + A| < \frac{5}{4}|A| − 5), unpublished manuscript.
[162] Lev V.F., On small subsets in abelian groups, this volume.
[163] Lev V., The structure of multisets with small number of subset sums, this volume.
[165] Levy M.P., Théorie d’addition des variables aléatoires.
[169] Lipkin E., Subset sums of sets of residues, this volume,
[172] Longobardi P., Maj M., On the derived length of groups with some permutational properties, manuscript.


[177] Manning W.A., *Groups in which a large number of operators may correspond to their inverses*, Trans. Amer. Math. Soc. 7 (1906), 233–240.


Nathanson M., Tenenbaum G., *Inverse theorems and the number of sums and products*, this volume.


G.A. Freiman, School of Mathematical Sciences, Raymond and Beverly Sackler, Faculty of Exact Sciences, Tel Aviv University, 69978 Tel Aviv, Israel • E-mail: grisha@math.tau.ac.il