BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 14, Number 2, April 1986

IP_r-SETS, SZEMERÉDI'S THEOREM, AND RAMSEY THEORY

BY H. FURSTENBERG AND Y. KATZNELSON

We set forth here combinatorial theorems relating to Szemerédi's theorem on arithmetic progressions in sets of integers of positive density. The method of proof (which we shall not even outline here) rests heavily on ergodic theory, and is developed in detail in [**3**]. Here we shall focus on the combinatorial ramifications. The common thread in our results is that a subset of a group which is "substantial" in some quantitative sense must contain certain kinds of patterns or configurations. In describing these configurations a special role is played by what we call IP_r-sets. (The letters IP refer to idempotence of operators that occur in the ergodic-theoretic development of this subject.)

DEFINITION 1. Let H be an abelian group. An IP_r-set in H is a set of $2^r - 1$ elements $\{h_n\} \subset H$ indexed by subsets $\alpha \subset \{1, 2, \ldots, r\}$, $\alpha \neq \emptyset$ and satisfying: if $\alpha \cap \beta = \emptyset$, then $h_{\alpha\cup\beta} = h_\alpha + h_\beta$.

The configurations whose existence is obtained in our discussion come about as follows. We suppose $\{h_\alpha^{(1)}\}, \{h_\alpha^{(2)}\}, \ldots, \{h_\alpha^{(k)}\}$ are IP_r-sets in a group. For a "substantial" set of the type to be described, we will find an element x and an index $\alpha \subset \{1, 2, \ldots, r\}$ such that all the elements

$$x + h_\alpha^{(1)}, x + h_\alpha^{(2)}, \ldots, x + h_\alpha^{(k)}$$

are inside the given set. If, for example, $H = \mathbf{Z}$ and $h_\alpha^{(q)} = qh_\alpha$, where $\{h_\alpha\}$ is a given IP_r-set, then the above sequence has the form $x+h_\alpha$, $x+2h_\alpha, \ldots, x+kh_\alpha$, i.e., forms an arithmetic progression.

Another possibility is to take $H = \mathbf{R}^d$ = d-dimensional Euclidean space and to let $v_1, v_2, \ldots, v_k$ be any k vectors in $\mathbf{R}^d$. If $\{\lambda_\alpha\}$ is an IP_r-set in $\mathbf{R}$ then $\{\lambda_\alpha v_1\}, \{\lambda_\alpha v_2\}, \ldots, \{\lambda_\alpha v_k\}$ is a family of IP_r-sets in $\mathbf{R}^d$. The sequence $x + \lambda_\alpha v_1, x + \lambda_\alpha v_2, \ldots, x + \lambda_\alpha v_k$ forms a configuration similar to the arbitrary preassigned finite configuration $\{v_1, \ldots, v_k\}$. Thus, a special case of our results will ensure the existence of the vertices of an equilateral triangle or a square in certain planar sets.

An example of a "substantial" subset of a group is a subset of positive density. We will generally be interested in finite sets and so a "substantial" subset can be described as follows. In any countable abelian group H there exists a sequence J_n of finite subsets such that for any element $h \in H$ the symmetric difference of J_n and $J_n + h$ tends to zero relative to the size of J_n ("Følner" sets). A substantial subset is a subset of a Følner set containing a proportion bounded away from zero of elements in the Følner set. We now formulate this precisely.

Received by the editors September 20, 1985.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 05B30.

We use the following notation. If S is a finite set, $|S|$ denotes its cardinality. For two sets S_1, S_2, $S_1 \,\Delta\, S_2$ denotes the symmetric difference, α always denotes a subset of $\{1, 2, \ldots, r\}$, and when α is a singleton $\alpha = \{p\}$ we write h_p for h_α.

THEOREM 1. *Let $k \in \mathbf{N}$ and let $\delta > 0$ be given. There exists $\varepsilon(k, \delta) > 0$ and $r(k, \delta) < \infty$ such that if H is an abelian group, $\{h_\alpha^{(1)}, h_\alpha^{(2)}, \ldots, h_\alpha^{(k)}\}$ any k IP_r-sets in H for $r \geq r(k, \delta)$, $J \subset H$ a finite subset satisfying*

$$(1) \qquad |(J + h_p^{(i)}) \,\Delta\, J| < \varepsilon(k, \delta)|J| \qquad i = 1, \ldots, k; \; p = 1, \ldots, r,$$

and $S \subset J$ any subset satisfying $|S| > \delta|J|$, then for some $\alpha \subset \{1, 2, \ldots, r\}$ and for some $x \in J$, we will have

$$x + h_\alpha^{(1)} \in S, x + h_\alpha^{(2)} \in S, \ldots, x + h_\alpha^{(k)} \in S.$$

Let us reformulate this for $H = \mathbf{Z}$. Here the Følner set J will be a long interval $\{1, 2, \ldots, N\}$.

THEOREM 2. *Let $\delta > 0$ and k be given. There exists $N(k, \delta)$ and $R(r, \delta)$ such that if $r \geq R(k, \delta)$ and $u_1, u_2, \ldots, u_r$ are any r integers and $N > N(k, \delta) \max|u_j|$, then whenever A is a subset of $\{1, 2, \ldots, N\}$ with more than δN elements, A contains an arithmetic progression of length k and common difference of the form $u_\alpha = u_{i_1} + u_{i_2} + \cdots + u_{i_l}$ for some $1 \leq i < i_2 < \cdots < i_l \leq r$.*

This result adds to the Szemerédi theorem the fact that we can specify to some extent the form of the common difference d of the arithmetic progressions $a, a + d, \ldots, a + ld$. Results of this type were known for van der Waerden's theorem. (See [**2** and **4**].) For example, it follows from the Hales-Jewett theorem [**4**] that if $\mathbf{Z}$ is partitioned into finitely many sets, one of these contains arbitrarily long arithmetic progressions whose common difference is a number with only 0's and 1's in its expansion to the base q. The numbers n of this form for $n < q^r$ form a simple example of an IP_r-set.

Next suppose that H is a finite group. In that case we may take $J = H$ and condition (1) in Theorem 1 is automatically satisfied. We then have

THEOREM 3. *Let $r(k, \delta)$ be as in Theorem 1. If H is a finite group and S is a subset with $|S| > \delta|H|$, and $\{h_\alpha^{(1)}, h_\alpha^{(2)}, \ldots, h_\alpha^{(k)}\}$ are any k IP_r-sets in H for $r > r(k, \delta)$, then we can find $x \in H$ and $\alpha \subset \{1, 2, \ldots, r\}$ with*

$$(2) \qquad x + h_\alpha^{(1)} \in S, \quad x + h_\alpha^{(2)} \in S, \quad \ldots, \quad x + h_\alpha^{(k)} \in S.$$

COROLLARY. *Let G and H be abelian groups, let δ be given and let $r > r(|G|, \delta)$. If (θ_α), $\alpha \subset \{1, 2, \ldots, r\}$, is an IP_r-set in the group of homomorphisms $\mathrm{Hom}(G, H)$, and S is a subset of H with $|S| > \delta|H|$, then for some $\alpha \subset \{1, 2, \ldots, r\}$ a coset of $\theta_\alpha(G)$ will be contained in S.*

As an application of this corollary we take F_q to be a finite field and suppose that G is a vector space over F_q of dimension d and that H is a vector space over F_q of dimension D. Let $r > r(|G|, \delta) = r(q^d, \delta)$. If $D \geq rd$, then it is easy to construct an IP_r-set of homomorphisms of $G \to H$ which are all isomorphisms. Then $\theta_\alpha(G)$ is a subspace of dimension d of H. These considerations lead to the following.

THEOREM 4. *With the notation of Theorem* 1, *if* V *is a vector space over* F_q *of dimension* $\geq (r(q^d, \delta) + 1)d$ *and* S *is any subset of* V *with* $|S| > \delta|V|$, *then* S *contains a* d-*dimensional affine plane.*

This extends the theorem of Brown and Buhler [**1**].

The following geometric result is easy to deduce from Theorem 1. I^d denotes the unit cube in $\mathbf{R}^d$.

THEOREM 5. *For* k *a natural number and* $\delta > 0$ *there exist two numbers,* $r(k, \delta) < \infty$ *and* $\lambda(k, \delta) > 0$ *such that for any* d, *if* $S \subset I^d$ *is a measurable set of measure* $> \delta$, *and* $F \subset I^d$ *is any set of* k *points and* Λ *is any* IP_r-*set of reals with* $|\lambda| < \lambda(k, \delta)$ *for* $\lambda \in \Lambda$, *then there exists a* $\lambda \in \Lambda$ *and an* $x \in I^d$ *with* $x + \lambda F \subset S$.

It is noteworthy that any set Λ of reals has the property described in the theorem. Replacing I^d by a larger cube, we see that if $S \subset \mathbf{R}^d$ has positive upper density in the sense that it intersects arbitrarily large cubes in a proportion $> \delta$ of the cube, then the property of the theorem will be true for arbitrary IP_r-sets of reals. For example we have

COROLLARY. *A subset of the plane of positive upper density contains the vertices of an equilateral triangle* (*of preassigned orientation*) *whose side has integral length. More generally, if* $A \subset \mathbf{R}^d$ *has positive upper density and* $F \subset \mathbf{R}^d$ *is a finite set, we can find* $x \in \mathbf{R}^d$ *with* $x + nF \subset A$, *where* n *is an integer* > 0.

Theorem 1 gives us additional information regarding arithmetic progressions occurring in infinite subsets of $\mathbf{Z}$ of positive density, as well as the configuration occurring in subsets of positive density in other groups.

DEFINITION 2. Let H be an abelian group. An IP_r^*-set in H is a subset that intersects every IP_r-set in H.

Note that the smaller r is the more stringent the requirement to be an IP_r^*-set.

THEOREM 6. *Let* H *be an abelian group and let* G *be an abelian group of endomorphisms of* H. *For every* k *and every* $\delta > 0$ *there is a number* $r(k, \delta)$ *such that, if* S *is a subset of* H *of upper density* $> \delta$, $F \subset H$ *with* $|F| = k$, *and* $r > r(k, \delta)$, *then* $\{g \in G$: *some translate of* $g(F) \subset S\}$ *is an* IP_r^*-*set.*

In particular, if $H = \mathbf{Z}$ and $G = \mathbf{Z}$ acts on itself by multiplication and $F = \{1, 2, \ldots, k\}$, then the set of g in question consists of the common difference in length k arithmetic progressions occurring in a subset $S \subset \mathbf{Z}$ of positive upper density. That this set is large now follows from the next theorem which we formulate for $\mathbf{Z}$.

THEOREM 7. *If* $R \subset \mathbf{Z}$ *is an* IP_r^*-*set then*

$$\liminf_{N \to \infty} \frac{(R \cap \{1, 2, \ldots, N\})}{N} \geq 2^{-r+1}.$$

Thus the density of the set of difference for length k *arithmetic progressions occurring in subsets of* $\mathbf{Z}$ *of density* $> \delta$ *is bounded away from zero by a number depending on* k *and* δ.

This is of course trivial for $k = 2$.

REFERENCES

1. T. C. Brown and J. P. Buhler, *A density version of a geometric Ramsey theorem*, J. Combin. Theory Ser. A **32** (1982), 20–34.

2. H. Furstenberg, *Recurrence in ergodic theory and combinatorial number theory*, Princeton Univ. Press, Princeton, N.J., 1981.

3. H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for IP-systems and combinatorial theory* J. d' Analyse Math. **45** (1985), 117–168.

4. R. L. Graham, B. Rothschild and J. Spencer, *Ramsey theory*, John Wiley, 1980.

INSTITUTE OF MATHEMATICS, HEBREW UNIVERSITY, JERUSALEM, ISRAEL