6. On the Distribution of Monochromatic Configurations

P. Frankl, R.L. Graham and V. Rödl

0. Introduction

Much of Ramsey theory is concerned with the study of structure which is preserved under finite partitions, (eg., see [8], [9], [12]). Some of the earliest results in the field were the following.

Schur's Theorem (1916) [17]. *For any partition of the set* $\mathbf{N}$ *of positive integers into finitely many classes, say* $\mathbf{N} = C_1 \cup \ldots \cup C_r$, *some* C_i *must contain a set of the form* $\{x, y, x+y\}$.

Van der Waerden's Theorem (1927) [19]. *For any finite partition of* $\mathbf{N} = C_1 \cup \ldots \cup C_r$, *some* C_i *must contain arbitrarily long arithmetic progressions.*

Ramsey's Theorem (1930) [16]. *For any finite partition of the set* $\binom{\mathbf{N}}{k}$ *of* k*-element subsets of* $\mathbf{N}$, *say* $\binom{\mathbf{N}}{k} = C_1 \cup \ldots \cup C_r$, *some* C_i *must contain the set* $\binom{X}{k}$ *of all the* k*-element sets of some infinite set* $X \subseteq \mathbf{N}$.

It is common in Ramsey theory to call the classes "colors", the partition into r classes an "r-coloring", and the objects belonging to a single class "monochromatic" (cf. [8], [9]).

Each of these results in fact enjoys a "finite" form, which measures the onset of monochromaticity. We abbreviate the interval $\{1, 2, \ldots, N\}$ by $[N]$.

Schur's Theorem (finite form). *For all* $r \in \mathbf{N}$ *there is a least integer* $Sc(r)$ *such that in any* r*-coloring of* $[Sc(r)]$ *there is a monochromatic set of the form* $\{x, y, x+y\}$.

Van der Waerden's Theorem (finite form). *For any* k *and* r *in* $\mathbf{N}$ *there is a least integer* $W(k,r)$ *such that in any* r*-coloring of* $[W(k,r)]$ *there is a monochromatic* k*-term arithmetic progression.*

Ramsey's Theorem (finite form). *For any* k, l *and* r *in* $\mathbf{N}$ *there is a least integer* $R = R(k, l; r)$ *such that in any* r*-coloring of* $\binom{[R]}{k}$ *there is an* l*-element set* $X \subseteq [R]$ *with* $\binom{X}{k}$ *monochromatic.*

The determination of the true orders of growth of the functions $Sc(r)$, $W(k,r)$ and $R(k,l;r)$ are among the most difficult problems in combinatorics. Indeed, each new factor of $\log\log n$ (or even 2!) is usually considered a significant achievement in this quest. A fairly complete survey of this work (as of the time this article is being written) can be found in [10].

In this article we want to focus on somewhat different quantitative aspects of these partition theorems. In one direction, we will ask not *when* the desired monochromatic structure must occur, but rather *how many* monochromatic structures we must have as the size of the set. We are partitioning tends to infinity, and the parameters k, l and r are fixed (cf. Theorems 1,2 and 3).

In another direction we will investigate a measure of the frequency of occurrence of monochromatic structures first suggested for Schur's theorem by Bergelson [1] (cf. Theorems 4 and 5).

We feel that both types of results can contribute to a deeper understanding of these (and other related) fundamental partition results and their various generalizations.

1. Rado's Theorem

In 1930, R. Rado [15] published a far-reaching generalization of Schur's theorem, which dealt with solution sets to systems of homogeneous linear equations over **Z**. (In fact, this striking work formed the basis of Rado's dissertation written under Schur's direction). To describe his results, we first need some terminology.

For an l by k matrix $A = (a_{ij})$ of integers, denote by $\mathcal{L} = \mathcal{L}(A)$ the system of homogeneous linear equations

$$(1.1) \qquad \sum_{j=1}^{k} a_{ij}x_j = 0, \quad 1 \le i \le l.$$

We can abbreviate this by writing

$$(1.2) \qquad A\bar{x} = \bar{0}, \ \bar{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} = (x_1, \ldots, x_k)^t$$

We say that $\mathcal{L}$ is *partition regular* if for any r-coloring of $\mathbf{N}$, there is always a solution to (1.1) with all x_i having the same color.

The matrix A is said to satisfy the *Columns Condition* if it is possible to re-order the column vectors $\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_k$ so that for some choice of indices $1 \leq k_1 < k_2 < \ldots < k_t = k$, if we set

$$A_i := \sum_{j=k_{i-1}+1}^{k_i} \bar{a}_j$$

then

(i) $A_1 = \bar{0}$;

(ii) For $1 < i \leq t$, A_i can be expressed as a rational linear combination of $\bar{a}_j$, $1 \leq j \leq k_{i-1}$.

A classical results of Rado asserts the following.

Rado's Theorem ([15], [9]). *The system $A\bar{x} = \bar{0}$ is partition regular if and only if A satisfies the Columns Condition.*

Let us call a set $\mathcal{H} \subseteq \mathbf{N}$ *large* if for any partition regular system $A\bar{x} = 0$ and finite coloring of $\mathcal{H}$, there is always a monochromatic solution to $A\bar{x} = 0$. It was shown by Deuber [3] (settling a conjecture of Rado) that large sets have the following partition property: If $\mathcal{H}$ is large and $\mathcal{H} = \mathcal{H}_1 \cup \ldots \cup \mathcal{H}_r$ then for some i, $\mathcal{H}_i$ is large. We next introduce some notation due to Deuber [3].

Definition. $D(m, p, c,) := \{(\lambda_1, \ldots, \lambda_m) :$ for some $i < m$, $\lambda_j = 0$ for $j < i$, $\lambda_i = c > 0$ and $|\lambda_k| \leq p$ for $k > i\}$

A set $S \subseteq \mathbf{Z}^+$ is called an *(m, p, c)-set* if

$$S = \{\sum_{i=1}^{m} \lambda_i y_i : (\lambda_1, \lambda_2, \ldots, \lambda_m) \in D(m, p, c)\}$$

for some choice of $y_1, y_2, \ldots, y_m > 0$.

As shown by Deuber, sets of solutions for partition regular systems $A\bar{x} = 0$ correspond to subsets of (m, p, c)-sets in the following way. Let A be an l by k matrix satisfying the Columns Condition, and let $A_1, A_2, \ldots, A_t$ be the column vector sums coming from the definition of the Columns Condition. We can assume without loss of generality that A has rank l. Then there exist $k - l$ linearly independent solutions to $A\bar{x} = \bar{0}$ which (by the Columns Condition) have the following form *:

---

* $\bar{x}^t$ denotes the transpose of $\bar{x}$; we will occasionally omit the t if it is clear from context.

$$\begin{array}{llll}
 & k_1 & k_2-k_1 & k_t-k_{t-1} \\
\bar{w}_1 & =(1,1,\ldots,1, & 0,0,\ldots,0,\ldots\ldots & 0,0,\ldots,0)^t \\
\bar{w}_2 & =(\alpha_{21},\ldots,\alpha_{2k_1}, & 1,1,\ldots,1,\ldots\ldots & 0,0,\ldots,0)^t \\
\vdots & & & \\
\bar{w}_t & =(\alpha_{t1}, & \ldots & \ldots\alpha_{tk_{t-1}},1,1,\ldots,1)^t \\
\bar{w}_{t+1} & =(\alpha_{t+1,1}, & \ldots & \ldots\ldots \quad \alpha_{t+1,k})^t \\
\vdots & & & \\
\bar{w}_{k-l} & =(\alpha_{k-l,1}, & \ldots & \ldots\ldots \quad \alpha_{k-l,k})^t
\end{array}$$

where all the α_{ij} are rational. Multiplying all the entries by a sufficiently large integer c, we obtain linearly independent vectors of the following form:

$$\begin{aligned}
\bar{v}_1 &= (c,c,\ldots,c,0,\ldots,0,\ldots,0,\ldots,0)^t \\
\bar{v}_2 &= (\beta_{21},\ldots,\beta_{2,k_1},c,\ldots,c,0,\ldots,0)^t \\
&\ \vdots \\
\bar{v}_{t+1} &= (\beta_{t+1,1},\ldots,\beta_{t+1,k})^t \\
&\ \vdots \\
\bar{v}_{k-l} &= (\beta_{k-l,1},\ldots,\beta_{k-l,k})^t
\end{aligned} \tag{1.3}$$

where all entries are integers. Set $p = |\max \beta_{ij}|$. Since every solution to $A\bar{x} = \bar{0}$ can be expressed as a linear combination of the vectors $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_{k-l}$, say,

$$\bar{x} = \sum_{i=1}^{k-l} y_i v_i,$$

then in fact each solution of $A\bar{x} = \bar{0}$ is always a subset of some $(k-l, p, c)$-set, and conversely, as claimed.

We are now ready to give the following quantitative version of Rado's Theorem.

Theorem 1. *Let A be an l by k matrix of rank l which satisfies the Columns Condition. Then for any r there exists $c_r(A) > 0$ such that in any r-coloring of $[N]$, $N > N_0$ there are at least $c_r(A)N^{k-l}$ monochromatic solutions to the partition regular system $A\bar{x} = \bar{0}$.*

If we let $\nu_{\mathcal{L}}(N, r)$ denote the minimum possible number of monochromatic solution sets to a system $\mathcal{L}$ whenever $[N]$ is r-colored (so that $\nu_{\mathcal{L}}(N) = \nu_{\mathcal{L}}(N, 1)$), then we have as an immediate consequence:

Corollary 1. *If $\mathcal{L}$ is partition regular then for any r there exists $c_r(\mathcal{L}) > 0$ so that*

$$\liminf_{N\to\infty} \frac{\nu_{\mathcal{L}}(N,r)}{\nu_{\mathcal{L}}(N)} \geq c_r(\mathcal{L})$$

Proof of Theorem 1. The proof will use the following version of Deuber's theorem ([3]).

Theorem. *For every choice of m, p, c and r there exist M, P and C such that for any r-coloring of*

$$J = \{\sum_{i=1}^{M} \lambda_i Y_i : (\lambda_1, \ldots, \lambda_M) \in D(M, P, C)\}$$

there exist pairwise disjoint sets $B_1, B_2, \ldots, B_m \subseteq [m]$ and

$$y_i = \sum_{j \in B_i} \xi_j Y_j, \quad 1 \le |\xi_j| \le P,\ 1 \le i \le m,$$

such that all linear combinations

$$\sum_{i=1}^{m} \lambda_i y_i, \quad (\lambda_1, \lambda_2, \ldots, \lambda_m) \in D(m, p, c)$$

are monochromatic.

Now, given our l by k matrix A of rank l satisfying the Columns Condition, we know by the preceding remarks that the entries of the set of solution vectors of $A\bar{x} = \bar{0}$ all belong to some $(k-l, p, c)$-set. Set $m = k - l$ and let M, P and C be the integers from Deuber's theorem. Choose $N \gg M$ to be very large. Consider all the M-tuples $(Y_1, Y_2, \ldots, Y_M)$ of integers Y_i satisfying

$$(1.4) \qquad 0 < Y_i \le \frac{N}{MC} \text{ and } Y_i \equiv (2P+1)^i \ mod(2P+1)^M$$

for $1 \le i \le M$. There are at least $c_1 N^M$ such M-tuples for some constant $c_1 > 0$ not depending on N. For such an M-tuple $(Y_1, Y_2, \ldots, Y_M)$, consider the (M, P, C)-set

$$J(Y_1, \ldots, Y_M) = \{\sum_{i=1}^{M} \lambda_i Y_i : (\lambda_1, \ldots, \lambda_M) \in D(M, P, C)\}$$

Let $[N] = C_1 \cup \ldots \cup C_r$ be an r-coloring of $[N]$. By Deuber's theorem we can find disjoint subsets $B_1, \ldots, B_{k-l} \subseteq [M]$ and $y_i = \sum_{j \in B_i} \xi_j Y_j$, $1 \le |\xi_j| \le P$, so that all the linear combinations $\sum_{i=1}^{m} \lambda_i y_i$ $(\lambda_1, \lambda_2, \ldots, \lambda_m) \in D(m, p, c)$, have the same color. In particular, $\bar{x} = \sum_{i=1}^{k-l} y_i \bar{u}_i$ (from (1.3)) is a monochromatic solution to the system $A\bar{x} = \bar{0}$. This therefore gives, with multiplicity, at least $c_1 N^M$ monochromatic solutions (one for each choice of $(Y_1, \ldots, Y_M)$). Our proof will be complete if we can show that each of these solutions can occur at most $N^{M-(k-l)}$ times.

To see this, suppose $(x_1, \ldots, x_k)$ is some solution obtained above, i.e., for some choice of $(y_1, \ldots, y_{k-l})$, the x_i are *fixed* linear combinations of the y_i. Then, we must show that the same monochromatic (m,p,c)-set is obtained at most $N^{M-(k-l)}$ times. However, given y_i, its residue modulo $(2P+1)^M$ uniquely determines the λ_j, $1 \le j \le M$ from (1.4). Thus, the possible $Y_1, \ldots, Y_M$ must satisfy $k-l$ linear equations which involve pairwise disjoint sets of unknowns among them. This gives the required bound and the proof is complete. □

2. Van der Waerden's Theorem

A natural question raised in connection with van der Waerden's Theorem by Erdős and Turán over 50 years ago was that of identifying which of the color classes must contain the desired arbitrarily long arithmetic progressions. In particular, they conjectured that the "largest" color class should always have this property. To make this precise, for a set $X \subseteq \mathbf{N}$, define the upper density $\bar{d}(X)$ of X by:

$$\bar{d}(X) := \limsup_{n\to\infty} \frac{|X \cap [n]|}{n}$$

In 1975, Szemerédi finally settled the conjecture of Erdős and Turán, by proving the following celebrated result.

Theorem of Szemerédi [18]. *If $X \subseteq \mathbf{N}$ satisfies $\bar{d}(X) > 0$ then X contains arbitrarily long arithmetic progressions.*

This result of course implies van der Waerden's Theorem, and it was, in fact, hoped that it might lead to improved estimates for $W(k,r)$. This did not happen (yet) though since Szemerédi's proof in fact uses van der Waerden's Theorem. More recently, Furstenberg and Katznelson [6], [7] have given alternate proofs and generalizations of Szemerédi's Theorem using techniques from ergodic theory and topological dynamics (which however, do not shed any light on the true values of $W(k,r)$).

Observe that a k-term arithmetic progression $(a, a+d, a+2d, \ldots, a+(k-l)d)$ can be viewed as a solution $\bar{x} = (x_1, x_2, \ldots, x_k)$ to the system of equations (over $\mathbf{N}$)

$$x_2 - x_1 = x_3 - x_2 = \ldots = x_k - x_{k-1} \neq 0$$

In this section we establish the density analogue to Theorem 1 for the appropriate systems of linear equations.

The system

$$A\bar{x} = \bar{0} \tag{2.1}$$

is said to be *density regular* if for any set $X \subseteq \mathbf{N}$ of positive upper density there is a vector $\bar{x}$ satisfying (2.1) and having all entries belonging to X.

If it happens that (2.1) has the vector $\bar{x} = \bar{1} = (1, 1, \ldots, 1)$ as a solution then, of course, for any $m \in \mathbf{N}$, $\bar{x} = m \cdot \bar{1} = (m, m, \ldots, m)$ is also a solution. In this case, (2.1) is trivially density regular. However, the solution $m \cdot \bar{1}$ is normally not considered to be very interesting. For example, for the density regular system

$$x_1 - 2x_2 + x_3 = 0$$

the solutions (x_1, x_2, x_3) are just the 3-term arithmetic progressions, provided the x_i are distinct.

With these considerations in mind, let us call the system (2.1) *irredundant*, if (2.1) *does not* imply that $x_i = x_j$ for $i \neq j$. Also, let us call a solution $\bar{x} = (x_1, \ldots, x_k)$ to (2.1) *proper* if all the x_i are distinct.

Fact 2.1. *If $A\bar{x} = \bar{0}$ is irredundant then it has a proper solution.*

Proof. For each choice of $i < j$, let $\bar{x}^{(ij)} = (x_1^{(ij)}, x_2^{(ij)}, \ldots, x_k^{(ij)})$ be a solution to (2.1) with $x_i^{(ij)} \neq x_j^{(ij)}$, which exists by hypothesis. Thus, for any integer N, $\bar{x}^* = (x_1^*, x_2^*, \ldots, x_k^*)$ with

$$x_t^* = \sum_{i<j} N^{ki+j} x_t^{(ij)}$$

is also a solution to (2.1) by linearity. However, if $N > \max_{i,j,t}(x_t^{(ij)})$ then all x_t^* are distinct. □

Fact 2.2. *An irredundant system $A\bar{x} = \bar{0}$ has a proper solution in every set X of positive upper density if and only if $A \cdot \bar{1} = 0$.*

Proof. First, since X has positive upper density then by Szemerédi's Theorem, X contains arbitrarily long arithmetic progressions. Suppose $\bar{x}_0 = (b_1, b_2, \ldots, b_k)$ is a proper solution of $A\bar{x}_0 = \bar{0}$, i.e., all the b_k are distinct. Let $B := \max_k b_k$ and let $P = \{c + \lambda d : \lambda \in [B]\}$ be a B-term arithmetic progression in X. If $\bar{1}$ also satisfies $A \cdot \bar{1} = \bar{0}$ then so does the linear combination

$$\bar{x}^* = c \cdot \bar{1} + d\bar{x}_0 = (c + b_1 d, c + b_2 d, \ldots, c + b_n d)$$

which is proper, and furthermore, has all entries in $P \subseteq X$, as desired.

In the other direction, suppose $A\bar{x} = \bar{0}$ has a proper solution in every set of positive upper density. Let $N > \sum_{i,j} |a_{ij}|$ where a_{ij} ranges over all entries of A. Consider the set $Y = \{Ny + 1 : y \in \mathbf{N}^+\}$ with (upper) density $1/N$. Suppose $\bar{x} = (x_1, \ldots, x_n)$ satisfies $A\bar{x} = \bar{0}$ where each $x_k = Ny_k + 1 \in Y$. Thus,

$$0 = \sum_j a_{ij} x_j = \sum_j a_{ij}(Ny_j + 1) = N \sum_j a_{ij} y_j + \sum_j a_{ij}$$

for $1 \leq i \leq m$. By the choice of N, this implies that $\sum_j a_{ij} = 0$ for all i. This is exactly the statement that $A\bar{1} = \bar{0}$, as required. This completes the proof. □

Theorem 2. *Let A be an l by k matrix of rank l so that $A\bar{x} = 0$ is irredundant and $A\bar{1} = \bar{0}$. Then for any $\epsilon > 0$ there is a constant $c_\epsilon = c_\epsilon(A) > 0$ so that if $N > N_0(A, \epsilon)$ and $X \subseteq [N]$ with $|X| > \epsilon N$ then X must contain at least $c_\epsilon N^{k-l}$ proper solutions $\bar{x}$ to $A\bar{x} = \bar{0}$.*

Proof. Let $\epsilon > 0$ be arbitrary (but fixed) and let $X \subseteq [N]$ with $|X| > \epsilon N$ be given, where it will be useful to think of N as being very large. Since A has rank l, the space of all (rational) solutions $\bar{x}$ to $A\bar{x} = \bar{0}$ has dimension $k - l$. Let $\bar{v}_0 = \bar{1}$, $\bar{v}_1, \ldots, \bar{v}_m$ be linearly independent integer solutions to $A\bar{x} = \bar{0}$ where $m := k - l - 1$ and for $\bar{v}_i = (v_{i1}, \ldots, v_{ik})^t$, we can assume without loss of generality, all $v_{ij} \geq 0$ (since if not, then we can repeatedly add $\bar{1}$ to $\bar{v}_i$ until this is true). Define $t := 1 + \max_{ij} v_{ij}$.

For $u \in \mathbf{N}$ and each vector $\bar{y} = (y_1, \ldots, y_m)$ with $y_i \in \mathbf{N}$, define the m-box $B_u(\bar{y})$ to be the set

$$\{(a_1y_1, a_2y_2, \ldots, a_my_m) : 0 \leq a_i < u,\ 1 \leq i \leq m\}$$

Further, define the projection $\pi : B_u(\bar{y}) \to \mathbf{Z}$ by

$$\pi[(a_1y_1, \ldots, a_my_m)] = \sum_{i=1}^{m} a_iy_i$$

By a theorem of Furstenberg and Katznelson ([6], [7]), there is an integer T so that for any $\bar{Y} = (Y_1, \ldots, Y_m)$ with $Y_i \in \mathbf{N}^+$, if $X^* \subseteq B_T(\bar{Y})$ with $|X^* > \frac{\epsilon}{2}|B_T(\bar{Y})| = \frac{\epsilon}{2}T^m$ then there exists a "translated" m-box $\bar{A} + B_t(A_0\bar{Y}) \subseteq X^*$, where $\bar{A} = (A_1Y_1, \ldots, A_mY_m)$ and $A_0, A_1, \ldots, A_m \in \mathbf{N}^+$.

Now, consider the set of all integer vectors $\bar{Y} = (Y_1, Y_2, \ldots, Y_m)$ which satisfy the following constraints:

(i) $0 \leq Y_i < \epsilon^2 N/mT$, $1 \leq i \leq m$;

(ii) $Y_i \equiv T^{i-1} (\text{mod } T^m)$, $1 \leq i \leq m$.

Note that if $\bar{P} = (a_1Y_1, \ldots, a_mY_m) \in B_T(\bar{Y})$, $\bar{P}' = (a_1'Y_1, \ldots, a_m'Y_m) \in B_T(\bar{Y})$ and $\pi(\bar{P}) = \pi(\bar{P}')$ then by (ii)

$$\sum_{i=1}^{m} a_iT^{i-1} \equiv \sum_{i=1}^{m} a_i'T^{i-1} \pmod{T^m}$$

which in turn implies $a_i = a_i'$ for all i, since $0 \leq a_i$, $a_i' < T$. Thus, π is 1-to-1 on $B_T(\bar{Y})$. Also, by (i)

$$0 \leq \pi(\bar{P}) < \epsilon^2 N$$

Let us call an integer $\underline{a} \in [N]$ "good" if

$$B(a) := a + \pi(B_T(\bar{Y})) \subseteq [N]$$

and

$$|X \cap B(a)| > \frac{\epsilon}{2}T^m$$

It is easy to see that for a fixed constant $\delta = \delta(\epsilon) > 0$, the set $A = \{a \in [N] : a \text{ is good } \}$ satisfies

$$|A| > \delta N.$$

By the choice of T, for each $a \in A$, $X \cap B(a)$ contains the translated projection

$$Y_0 + \pi(B_t(A_0\bar{Y}))$$

for some $Y_0, A_0 \in \mathbf{N}$. Furthermore, by the choice of t, this in turn contains all components of the solution

$$\bar{x} = Y_0 \cdot \bar{1} + \sum_{i=1}^{m} A_0 Y_i \bar{v}_i$$

to $A\bar{x} = \bar{0}$. Since there are cN^{m+1} ways to choose the $Y_0, Y_1, \ldots, Y_m$ for a positive constant c (depending on ϵ and A) then the theorem will be proved if we can show that no solution $\bar{x}$ to $A\bar{x} = \bar{0}$ can arise this way in more than a bounded number of ways.

To see this, first note that since the $(k-l)$ by k matrix $V = (v_{ij})$ formed from the (linearly independent) solution vectors $\bar{v}_i$, $0 \le i < k-l$, has rank $k-l$ then we can assume without loss of generality (by relabelling, if necessary) that the $(k-l)$ by $(k-l)$ submatrix $V' = (v_{ij})_{0 \le i,j < k-l}$ is non-singular. Suppose $\bar{x} = (x_1, \ldots, x_k)$ has all its components x_i lying in some set $Y_0 + \pi(B_T(\bar{Y})) \subseteq [N]$ where $\bar{Y} = (Y_1, \ldots, Y_m)$ satisfies (i) and (ii). For each of the $k!$ permutations σ on $[k]$, consider the vector $\bar{x}_\sigma = (x_{\sigma(1)}, \ldots, x_{\sigma(k)})$. If

$$\bar{x}_\sigma = \sum_{i=0}^{m} Y_i \bar{v}_i$$

then by the non-singularity of V', the first $k-l$ coordinates of $\bar{x}_\sigma$ determine all the Y_i. Thus, each such $\bar{x}$ can arise from at most $k!$ choices for the Y_i.

Finally, we observe that almost all of these $c'N^{k-l}$ solutions $\bar{x}$ to $A\bar{x} = \bar{0}$ are proper solutions. This is because, by hypothesis, for $i \ne j$, the space of solutions $\bar{x}$ with $x_i = x_j$ corresponds to a non-trivial dependence between the coefficients Y_i, $1 \le i \le m$, resulting in at most $O(N^{k-l-1})$ such solutions.

This completes the proof of the theorem. □

Let $v^*_{\mathcal{L}}(N;\epsilon)$ denote the minimum possible number of proper solutions to a system $\mathcal{L} = \mathcal{L}(A)$ which can belong to a set $X \subseteq [N]$ having $|X| > \epsilon N$. The following corollary is immediate.

Corollary. *If A is irredundant and $\mathcal{L} = \mathcal{L}(A)$ is density regular (i.e., $A \cdot \bar{1} = \bar{0}$) then for any $\epsilon > 0$ there exists $c^*_\epsilon(\mathcal{L}) > 0$ such that*

$$\liminf_{N \to \infty} \frac{v^*_{\mathcal{L}}(N;\epsilon)}{v_{\mathcal{L}}(N)} \ge c^*_\epsilon(\mathcal{L}),$$

where $v_{\mathcal{L}}(N)$ denotes the total number of solutions $\mathcal{L}$ has in $[N]$.

When the Corollary is applied to the system

$$\mathcal{L}^* : x_2 - x_1 = x_3 - x_2 = \ldots = x_k - x_{k-1}$$

we obtain the desired qualitative form of Szemerédi's Theorem, namely for some $c = c^*_\epsilon(\mathcal{L}^*) > 0$, if $N > N_0$ then

$$v^*_{\mathcal{L}^*}(N;\epsilon) \geq cN^2,$$

which is, of course, up to the value of c, the best one could hope for here.

3. Ramsey's Theorem

It turns out the analogues of the preceding results for Ramsey's Theorem follow rather easily from an averaging argument. We sketch it here for completeness.

Theorem 3. *For all k, l and r in* N *there exists $c = c(k,l,r) > 0$ such that for any r-coloring of $\binom{[N]}{k}$, $N \geq R(k,l;r)$, there are at least $c\binom{N}{l}$ l-sets $Y \subseteq [N]$ for which $\binom{Y}{k}$ is monochromatic.*

Proof. Let $R = R(k,l;r)$, and suppose $\binom{[N]}{k}$ is arbitrarily r-colored with $N \geq R$. Then for any R-set $Z \in \binom{[N]}{R}$ there is always some l-set $Y \in \binom{Z}{l}$ with $\binom{Y}{k}$ monochromatic. Call such an l-set Y "good". Now each good Y can occur in at most $\binom{N-l}{R-l}$ different $Z \in \binom{[N]}{R}$. Since there are $\binom{N}{R}$ different $Z \in \binom{[N]}{R}$ then there must be at least

$$\frac{\binom{N}{R}}{\binom{N-l}{R-l}} > c_R\binom{N}{l}$$

good sets $Y \in \binom{[N]}{l}$, i.e., such that $\binom{Y}{k}$ is monochromatic. □

Note that as in Theorems 1 and 2, a positive proportion of the objects under consideration is guaranteed to be monochromatic. This phenomenon does not always occur, however, as the following example shows. It is known (see [8]) that for each r there is an $F(r)$ so that for any r-coloring of all the subsets of $[F(r)]$ we can always find nonempty disjoint sets $A, B \subseteq [F(r)]$ so that A, B and $A \cup B$ all have the same color. Now, for a fixed N, consider the 2-coloring $\mathcal{X}$ of the subsets of $[N]$ given by:

$$\mathcal{X}(X) = \begin{cases} 0, & \text{if } |X| \leq N/2, \\ 1, & \text{if } |X| > N/2. \end{cases}$$

However, with this coloring there are only $O(2^{3N/2})$ monochromatic triples $\{A, B, A \cup B\}$ while in $[N]$ there are $(1 + o(1))3^N$. Thus, we do not get a positive proportion in this case.

4. An Iterated Density Theorem for the Strong van der Waerden Theorem

The following strengthening of van der Waerden's Theorem was used by Rado [15] in his work on partition regular systems.

Strong van der Waerden Theorem. *In any finite coloring of* $\mathbf{N}$ *there must exist for all* $p \in \mathbf{N}$, *a monochromatic set of the form*

$$\{x, y, x+y, 2x+y, \ldots, (p-1)x+y\}.$$

Note that this set consists of a p-term arithmetic progression together with its common difference. For the special case of $p = 2$, this reduces to the set $\{x, y, x+y\}$ which occurs in Schur's Theorem. However, even in this case it is clear that there is no direct density analogue to this theorem (as Szemerédi's Theorem was for the ordinary van der Waerden Theorem) since, for example, the set of odd integers $\{2k+1 : k \in \mathbf{N}\}$ has (upper) density $1/2$ and yet contains *no* set of the form $x, y, x+y$. Nevertheless, it is possible to prove a result which asserts that in any finite partition of $\mathbf{N}$, there are "many" monochromatic sets of the form $\{x, y, x+y\}$. This was first done by Bergelson, who recently proved the following.

Theorem 1. *In any finite coloring of* $\mathbf{N}$, *we have*

$$\bar{d}\{x : \bar{d}\{y : \{x, y, x+y\} \text{ is monochromatic} > 0\} > 0\} \tag{4.1}$$

Actually, Bergelson proves the following somewhat stronger result which does not, however, guarantee that either set has upper density bounded away from 0 as a function only of the number of colors.

Bergelson's Theorem (1986) [1]. *In any finite coloring of* $\mathbf{N}$, *there is always some color class* C *with* $\bar{d}(C) > 0$ *such that for any* $\epsilon > 0$,

$$\bar{d}\{x : \bar{d}\{y : x, y, x+y \text{ is monochromatic }\} \geq \bar{d}(c)^2 - \epsilon\} > 0$$

In this section we will prove an iterated density version of the strong van der Waerden Theorem, which will imply, in particular, a strengthening of (4.1), both in having explicit functions in the lower bounds, and in the replacement of $\bar{d}$ by $\mathbf{d}$. We first introduce a slightly modified form of the (m, p, c)-sets introduced in Section 1, called $(m, p, c)'$-sets, which will be useful in what follows.

For m, p and c in N, we mean by an $(m,p,c)'$-set, a subset of N which can be formed as follows, for suitable $a_1, a_2, \ldots, a_m \in \mathbf{N}$:

$$< a_1, \ldots, a_m > := \{\lambda_1 a_1 + \ldots + \lambda_{i-1} + c a_i : 0 \leq \lambda_j < p,\ 1 \leq j < i,\ 1 \leq i \leq m\}.$$

We also define

$$[a_1, \ldots, a_m] := \{\lambda_1 a_1 + \ldots + \lambda_m a_m : 0 \leq \lambda_i < p\},$$

where in both cases the values of p and c will be understood from the context if only the left-hand sides are used.

Thus,

$$(4.2) \qquad < a_1, \ldots, a_m > = < a_1, \ldots, a_{m-1} > \cup \{[a_1, \ldots, a_{m-1}] + c a_m\}.$$

Suppose now that N is arbitrarily r-colored. A basic result of Deuber [3] then implies that there exist $M, P, C \in \mathbf{N}$ so that any $(M,P,C)'$-set must always contain a monochromatic $(m,p,c)'$-set.

Our result will deal with $(2,p,1)'$-sets. These are just sets of the form $< x, y > = \{x, y, x+y, 2x+y, \ldots, (p-1)x+y\}$. In what follows, the integers M, P and C will denote the values needed in Deuber's theorem to force monochromatic $(2,p,1)'$-sets.

Theorem 4. **(Iterated Strong van der Waerden Theorem).** *In any r-coloring of* N *and for any p, there is a $\delta = \delta(r,p) > 0$ such that*

$$(4.3) \qquad \mathbf{d}\{x : \bar{d}\{y :< x, y > \text{(is monochromatic)}\} > \delta\} > \delta.$$

Proof. We assume $p > 2$ (the case $p = 2$ is very similar, and is omitted). Define

$$B = B(\delta) := \{x : \bar{d}y :< x, y > \text{ is monochromatic} \leq \delta\}$$

and let $\bar{B} := \mathbf{N} \backslash B$, the complement of B in N. Assume to the contrary that $\mathbf{d}(\bar{B}) = 1 - \bar{d}(B) \leq \delta$. Let $B_1 := \{d \in B : < b > \subseteq B\}$, and let $a_1 \in B_1$ be the least element in B_1. Next, define

$$B_2 = \{b \in B_1 : [a_1] + Cb \subseteq B_1\}$$

Thus, if $b \in B_2$, then

$$[a_1] + Cb \subseteq B_1 \subseteq B \text{ and } < a_1 > \subseteq B$$

so that $< a_1, b > \subseteq B$. Next, select (if possible) $a_2 \in B_2$ so that $< a_1, a_2 >$ contains no monochromatic $(2,p,1)'$-set. Define

$$B_3 := \{b \in B_2 : [a_1, a_2] + Cb \subseteq B_2\}.$$

Thus, for $b \in B_3, < a_1, a_2, b > \subseteq B$. Continuing, we select (if possible) $a_3 \in B_3$ so that $< a_1, a_2, a_3 >$ contains no monochromatic $(2, p, 1)'$-set, and we define

$$B_4 := \{b \in B_3 : [a_1, a_2, a_3] + Cb \subseteq B_3\}, \text{ etc.}$$

In general, after

$$B_j := \{b \in B_{j-1} : [a_1, \ldots, a_{j-1}] + Cb \subseteq B_{j-1}\}$$

is formed, we see that for $b \in B_j$, $< a_1, \ldots, a_{j-1}, b > \subseteq B$. We then select (if possible) $a_j \in B_j$ so that $< a_1, \ldots, a_j >$ contains no monochromatic $(2, p, 1)'$-set, and define $B_{j+1} := \{b \in B_j : [a_1, \ldots, a_j] + Cb \subseteq B_j\}$, etc. (where, of course, throughout this construction $< a_1, \ldots, a_i >$ denotes an (i, P, C)-set). By Deuber's theorem this process must terminate with the formation of B_t, for some $t < M$. In order to guarantee that this is actual cause for termination, we need to know that the various B_j's are nonempty. This fact is implied by the following elementary lemma.

Lemma. *Let $A \subseteq \mathrm{N}$, $D \subseteq \mathrm{N}$, $C \in \mathrm{N}$ with A finite, and define*

$$D' = \{d \in D : Cd + A \subseteq D\}$$

Then

$$1 - \bar{d}(D') \leq C \mid A \mid (1 - \bar{d}(D)) \tag{4.4}$$

The proof of this result is elementary and will be omitted. Thus, by (4.4) and (4.2) (with C in place of c), we have

$$1 - \bar{d}(B_{j+1}) \leq C \mid [a_1, \ldots, a_j] \mid (1 - \bar{d}(B_j)) \leq CP^j(1 - \bar{d}(B_j)). \tag{4.5}$$

Consequently,

$$1 - \bar{d}(B_t) \leq C^t P^{\binom{t+1}{2}}(1 - \bar{d}(B)) \leq \delta C^M P^{\binom{M}{2}} \tag{4.6}$$

so that for δ sufficiently small, all $\bar{d}(B_i)$ are at least $1/2$ (say) for $1 \leq i \leq t$.

Therefore, the set $\{a_1, \ldots, a_{t-1}\}$ has the properties:

(i) $< a_1, \ldots, a_{t-1} >$ contains no monochromatic $(2, p, 1)'$-set;

(ii) For any $b \in B_t$, $< a_1, \ldots, a_{t-1}, b > \subseteq B$ contains some monochromatic $(2, p, 1)'$-set, say $< x(b), y(b) >$.

Thus,

$$x(b) = \lambda_1 a_1 + \ldots + \lambda_{j-1} a_{j-1} + C a_j \tag{4.7}$$

for some $j = j(b) \leq t - 11$. To see this, suppose otherwise, i.e., suppose that

$$x(b) = \lambda_1 a_1 + \ldots + \lambda_{t-1} a_{t-1} + Cb.$$

Since $p > 2$ by hypothesis then $< x(b), y(b) >$ contains the element $2x(b) + y(b) > 2Cb$. However, this is impossible if b is sufficiently large since we have assumed

$$< x(b), y(b) >\subseteq< a_1, \ldots, a_{t-1}, b >$$

and the largest element of $< a_1, \ldots, a_{t-1}, b >$ is less than $(a_1 + \ldots + a_{t_1})p + Cb$.

On the other hand, we must have

$$y(b) = \lambda'_1 a_1 + \ldots + \lambda'_{t-1} a_{t-1} + Cb \tag{4.8}$$

since, if not, say $y(b) = \lambda'_1 a_1 + \ldots + \lambda'_{j-1} a_{j-1} + Ca_j$ for some $j < t$, then no element of $< x(b), y(b) >$ is large enough to use b, so that $< a_1, \ldots, a_{t-1} >$ must have already contained $< x(b), y(b) >$, which is a contradiction.

Now, for each $b \in B_t$, there are fewer than MP^{2M} choices for $j(b)$, $\bar{\lambda}(b) = (\lambda_1, \ldots)$ and $\bar{\lambda}'(b) = (\lambda'_1, \ldots)$. Hence, for some choice of $j_0(b), \lambda_0(b)$ and $\bar{\lambda}'_0(b)$, the set of (large) $b \in B_t$ with $j(b) = j_0(b)$, $\bar{\lambda}(b) = \bar{\lambda}_0(b)$ and $\bar{\lambda}'(b) = \bar{\lambda}'_0(b)$ has upper density at least $\bar{d}(B_t)/MP^{2M}$. Call this set B^*. Also

$$\bar{d}\{y : y = y(b) \text{ for } b \in B^*\} \geq \frac{1}{C}\bar{d}(B^*) \geq \bar{d}(B_t)/CMP^{2M} \tag{4.9}$$

since $y = y(b) = \lambda'_1 a_1 + \ldots + \lambda'_{t-1} a_{t-1} + Cb$. Note that

$$x(b) = \lambda_1 a_1 + \ldots + \lambda_{j0} a_{j0} \in< a_1, \ldots, a_{t-1} >\subseteq B.$$

Therefore, $\bar{d}\{y(b) : b \in B^*\} \leq \delta$ since $< x(b), y(b) >$ is monochromatic. This implies

$$\bar{d}(B_t)/CMP^{2M} \leq \delta,$$

i.e.,

$$\bar{d}(B_t) \leq \delta MP^{2M}C \tag{4.10}$$

However, this contradicts (4.6) if δ is sufficiently small. Hence, the initial assumption that $\mathrm{d}(\bar{B}) \leq \delta$ is untenable, and (4.3) must hold. This completes the proof. □

Note that this proof shows that δ can be chosen to be $(M^2 C^{M+1} P^{2M^2})^{-1}$, for example.

5. An Iterated Density Ramsey Theorem

The obvious density version of the finite form of Ramsey's theorem is clearly false, as can be seen, for example, by considering the complete bipartite graph $K_{n,n}$. This graph has more than half the possible number of edges for a graph with $2n$ vertices but contains no triangle. However, there is an iterated density version (in the spirit of Theorem 4) which is valid. This we now give.

Fix $k \le l$ and r, and suppose $\binom{N}{k}$ is r-colored. For $a_1 < \ldots < a_{l-1}$, define

$$\Gamma(a_1,\ldots,a_{l-1}) := \{x : \binom{a_1,\ldots,a_{l-1},x}{k} \text{is monochromatic}\},$$

and for $0 \le i < l-1$,

$$\Gamma(a_1,\ldots,a_i) := \{x : \bar{d}(\Gamma(a_1,\ldots,a_i,x)) > \delta\} \tag{5.1}$$

where $\delta = \delta(k,l,r) = 2^{-R}$ and $R := (k,l;r)$, the ordinary Ramsey number. For $i = 0$, we denote the expression in (5.1) by Γ.

Theorem 5. *For all $k \le l$ and r,*

$$\mathrm{d}(\Gamma) > \delta. \tag{5.2}$$

Proof. Assume $\mathrm{d}(\Gamma) \le \delta$. Thus

$$\bar{d}(\mathbf{N}\backslash\Gamma) \ge 1 - \delta.$$

Note that

$$x \notin \Gamma(a_1,\ldots,a_i) \Rightarrow \bar{d}(\Gamma(a_1,\ldots,a_i,x)) \le \delta.$$

Define S_i, $i = 1,2,\ldots$, as follows:

$$S_1 = \{s_1\} \text{ where } s_1 \in \mathbf{N}\backslash\Gamma \text{ is arbitrary.}$$

Suppose $S_j = \{s_1, s_2, \ldots, s_j\}$ has been defined. Form $S_{j+1} = S_j \cup \{s_{j+1}\}$ by choosing s_{j+1} (if possible) so that:

(i) $s_{j+1} \in \mathbf{N} - \cup_{u=0}^{l-1} \cup_{1 \le i_1 < \ldots < i_u \le j} \Gamma(S_{i_1},\ldots,S_{i_u}) := C_j$,

(ii) No $Y \in \binom{S_{j+1}}{l}$ has $\binom{Y}{k}$ monochromatic.

Note that since

$$\bar{d}(C_j) \ge 1 - \{\binom{j}{0} + \binom{j}{1} + \ldots + \binom{j}{k-1}\}\delta \ge 1 - 2^j\delta$$

then we never get stuck because of (i). However, by Ramsey's Theorem, we must eventually halt because of condition (ii), say with the formation of S_t, for some $t < R$. By the definition of S_t, for each $c \in C_t$, there is a set $X(c) \in \binom{S_t}{l-1}$ such that $\binom{X(c)\cup c}{k}$ is monochromatic. Thus, there exists a set $X_0 = \{s_{j_1} < \ldots < s_{j_{l-1}}\} \in \binom{S_t}{l-1}$ such that

(5.3)

$$\bar{d}\{c \in C_t : X(c) = X_0\} \ge \bar{d}(C_t)/\binom{t}{l-1} \ge (1 - 2^t\delta)/\binom{t}{l-1} > 2^{-R} = \delta$$

by the choice of δ. However, by construction,

$$s_{j_{l-1}} \notin \Gamma(s_{j_1},\ldots,s_{j_{l-2}})$$
$$\Rightarrow \bar{d}(\Gamma(s_{j_1},\ldots,s_{j_{l-1}})) \le \delta$$

$$\Rightarrow \bar{d}\{u : \binom{X_0 \cup u}{l} \text{ is monochromatic } \} \leq \delta$$

which contradicts (5.3). This proves the theorem. □

6. Concluding Remarks

There are a number of other examples known for which some of the preceding extensions can be proved (cf. [4], [5]). These include several of the canonical partition theorems, in which an arbitrary number of colors can be used (but a wider class of colored objects is allowed); see [2], [11], [13], [14]. However, we have barely scratched the surface for what might be looked at here. For example, if $<< x, y, z >>:= \{x, y, z, x+y, x+z, y+z, x+y+z\}$ then is it true that for any r-coloring of $\mathbf{N}$ there is a $\delta > 0$ such that

$$\mathrm{d}\{x : \bar{d}\{y : \bar{d}\{z : << x, y, z >> \text{ is monochromatic } \} > \delta\} > \delta\} > \delta?$$

We also point out that we have very little idea as to the true values of the various constants ($\delta' s$ and $c' s$) appearing in our theorems. Of course, since these typically depend on the corresponding values of the classical Ramsey numbers $Sc(r)$, $W(k, r)$ and $R(k, l; r)$ which themselves are far from being completely understood, (see [10]) then it is not surprising that our current knowledge here is very incomplete.

References

[1] V.Bergelson, A density statement generalizing Schur's theorem, *J.Combin. Th. (A)* **43** (1986), 338-343.

[2] W.Deuber, R.L.Graham, H.J.Prömel, B.Voigt, A Canonical Partition Theorem for Equivalence Relations on 2^t, *J.Combinatorial Th. (A)*, **34**, (1983), 331-339.

[3] W.Deuber, Partitionen und lineare Gleichungsysteme, *Math. Z.* **133** (1973), 109-123.

[4] P.Frankl, R.L.Graham and V.Rödl, Quantitative theorems for regular systems of equations, (to appear).

[5] P.Frankl, R.L.Graham and V.Rödl, Iterated combinatorial density theorems, (to appear).

[6] H.Furstenberg, Y.Katznelson, An ergodic Szemerédi theorem for commuting transformations, *J.Analyse Math.* **34** (1978), 275-291.

[7] H.Furstenberg, Y.Katznelson and D.Ornstein, The ergodic theoretic proof of Szemerédi's Theorem, *Bull Amer. Math. Soc.* **7** (1982), 527-552.

[8] R.L.Graham, Rudiments of Ramsey Theory, *Regional Conference Series in Mathematics*, no. 45, AMS, 1981.

[9] R.L.Graham, B.L.Rothschild, J.H.Spencer, *Ramsey Theory*, John Wiley & Sons, Inc., 1980.

[10] R.L.Graham and V.Rödl, Numbers in Ramsey theory, *in Surveys in Combinatorica 1987* (C. Whitehead, ed) LMS Lecture Note Series 123, Cambridge, 1987, 111–113.

[11] H.Lefmann, Kanonische Partitionssatze, *Ph.D. Dissertation*, Univ. of Bielefeld (1985).

[12] J. Nešetřil, V.Rödl, *Partition Theory and its Applications*, Surveys in Combinatorics 1979 (B. Bollobás, ed) LMS Lecture Note Series 38, Cambridge, 1979, 96–157.

[13] H.J.Prömel, V.Rödl, An elementary proof of the canonizing version of Gallai-Witt's Theorem, *J.Combinatorial Theory (A)*, **42**, 2 (1986), 144-149.

[14] H.J.Prömel, B.Voigt, Canonical Partition Theorems for Parameter Sets, *J. Combinatorial Theory (A)*, **35** (1985), 309-327.

[15] R.Rado, Studien zur Kombinatorik, *Math. Z.* **36** (1933), 242-280.

[16] F.P.Ramsey, On a problem in formal logic, *Proc. London Math. Soc.* (2) **30** (1930), 264-285.

[17] I.Schur, Über die Kongruenz $x^m + y^m = z^m \pmod p$, *Jber.Deutsch. Math. Verein.* **25** (1916), 114-116.

[18] E.Szemerédi, On Sets of Integers containing no k Elements in Arithmetic Progression, *Acta Arith.* **27** (1975), 199-245.

[19] B.L. van der Waerden, Beweis einer Baudetschen Vermutung, *Nieuw. Arch. Wisk.* **15** (1927), 212-216.

P. Frankl
C.N.R.S.
Paris France

R.L. Graham
AT & T Bell Laboratories
Murray Hill,
New Jersey 07974

V. Rödl
Czech. Technical University
Husova, Czechoslovakia