

PRIMITIVE RECURSIVE BOUNDS FOR VAN DER WAERDEN NUMBERS

SAHARON SHELAH

0. INTRODUCTION

We shall give primitive recursive upper bounds for the van der Waerden numbers, the Hales-Jewett numbers, Graham-Rothschild and the affine Ramsey numbers.

In 1927 van der Waerden published a proof of the following result.

Van der Waerden's Theorem [15]. *For all positive integers n and c there exists an integer N such that if the set of integers $\{1, 2, \dots, N\}$ is c -colored, then there exists a monochromatic n -term arithmetic progression.*

Let $W(n, c)$ be the least such integer. The known proofs of van der Waerden's theorem proceed by a double induction on n and c and yield extremely large upper bounds for $W(n, c)$. For example, the following upper bound $U(n, c)$ is obtained in [8]. Suppose that for some n $U(n, r)$ has been defined for all r . For each $c \in \mathbb{N}$ define a sequence by

$$a_0 = 1, \quad a_{i+1} = 2U(n, c^{a_i}), \quad i \geq 0.$$

Then $U(n+1, c) = a_c$. In order to get some idea of the rate of growth of $U(n, c)$ (and of the functions which will be introduced later in this paper), we shall define the classes of the Grzegorzczk hierarchy. For each $n \in \mathbb{N}$, define a function E_n by

$$\begin{aligned} E_0(x, y) &= x + y, \\ E_1(x) &= x^2 + 2, \\ E_{n+2}(0) &= 2, \\ E_{n+2}(x+1) &= E_{n+1}(E_{n+2}(x)). \end{aligned}$$

Then, for example,

$$2^{2^x} \leq E_2(x) \leq 3^{3^x}$$

Presented at a meeting of the British Mathematics Colloquium; received by the editors February 10, 1988.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 05A99; Secondary 15A03.

Key words and phrases. van der Waerden theorem, Hales-Jewett theorem, Graham-Rothschild theorem, Ramsey theory, primitive recursive functions.

We thank S. Thomas for telling us the problem and rewriting the proofs.

and

$$2^{2^{2^{\dots^2}}} \leq E_3(x) \leq 3^{3^{3^{\dots^3}}},$$

where both towers have height $2x + 1$. Define the class of functions \mathcal{E}^n , $n \in \mathbb{N}$, as follows. \mathcal{E}^0 is the class whose initial functions are the zero function, the successor function, and the projection functions and is closed under composition and limited recursion. \mathcal{E}^{n+1} is defined similarly except that the function E_n is added to the list of initial functions. \mathcal{E}^n is called the n th Grzegorzcyk class. (To say that a class \mathcal{E} is closed under limited recursion means that if $g, h, j \in \mathcal{E}$ and if f is defined by primitive recursion using g, h and satisfies

$$f(x_1, \dots, x_k) \leq j(x_1, \dots, x_k)$$

then $f \in \mathcal{E}$.) The most important properties of the Grzegorzcyk classes are summarized below. (A clear treatment of this material may be found in [11].)

Fact 1. If $n \geq 2$ and $f \in \mathcal{E}^n$, then there is an integer m such that

$$f(x_1, \dots, x_k) \leq E_{n-1}^m(\max\{x_1, \dots, x_k\}),$$

where E_{n-1}^m is the m th iterate of E_{n-1} .

Fact 2. If $n \geq 2$ and $f \in \mathcal{E}^n$ is a unary function, then f is eventually majorized by E_n (i.e., there exists an integer t such that $f(x) < E_n(x)$ for all $x > t$).

Fact 3. If $g, h \in \mathcal{E}^n$ and f is defined by primitive recursion using g and h , then $f \in \mathcal{E}^{n+1}$.

Fact 4. $\bigcup_{n \in \mathbb{N}} \mathcal{E}^n$ is the class of primitive recursive functions.

It is easily checked that the function $V(x) = U(x, 2)$ eventually majorizes E_n for every $n \in \mathbb{N}$. Hence V is not primitive recursive. In fact, whenever a function F is defined by double recursion, then F is essentially a variant of the Ackermann function $E_\omega(x) = E_{x+1}(x)$. (Actually E_ω is a slight modification of the function which Ackermann [1] constructed.) Hence to obtain a substantially improved upper bound for $W(n, c)$ it is necessary to find a proof of van der Waerden's theorem in which the use of a double induction is avoided.

The known lower bounds for $W(n, c)$ are of a completely different order of magnitude. For example, in [8] the Lovász local lemma is used to show that

$$W(n, 2) \geq \left(\frac{2^n}{2en} \right) (1 + o(1)).$$

This improves an earlier result of Erdős and Rado [3]. In [2], Berlekamp gave a constructive proof that if p is a prime, then

$$W(p + 1, 2) \geq p2^p.$$

Given the enormous discrepancy between the known upper and lower bounds for $W(n, c)$, the problem of finding the true order of magnitude became a

central question of Ramsey theory. This problem is discussed in numerous papers; for example, see [6] and [8].

In 1936, Erdős and Turán had a seminal idea of how to obtain a more reasonable upper bound; namely, prove a much stronger theorem. So they proposed the following conjecture:

If A is a set of positive integers with positive upper density, that is, satisfying

$$(*) \quad \limsup_N \frac{|A \cap [1, N]|}{N} > 0$$

then A contains arbitrarily long arithmetic progressions.

This conjecture was a very fruitful one. Roth [12] used analytic number theory to prove that $(*)$ implies that A contains a 3-term arithmetic progression. In 1974, Szemerédi [14] settled the general conjecture affirmatively with a highly complicated combinatorial argument. Unfortunately Szemerédi's proof made use of van der Waerden's theorem and so did not give a better upper bound for $W(n, c)$. A few years later, Furstenberg and Weiss gave a proof of Szemerédi's theorem using topological dynamics; and Furstenberg and Katznelson proved some generalizations. (For an account of this work see [4].) Unfortunately it seemed that the analytic methods of Furstenberg were such that no upper bound for $W(n, c)$ could be derived from them.

Then, in the seventies, Solovay had an intriguing idea: use logic to show that $W(n, c)$ really did grow at essentially the same rate as the Ackermann function. The idea was to translate the problem into logic as follows. Work in a weak (but natural) version of first order Peano arithmetic, in which the principle of double induction is not provable, and show that van der Waerden's theorem cannot be proved in this system. It would follow that $W(n, c)$ is not primitive recursive and that the known upper bounds were essentially correct. (Solovay had already had a striking success with the opposite approach. In [10], he and Ketonen showed that the function associated with the Paris-Harrington variant of Ramsey's theorem grew faster than every provably recursive function. This gave a new proof that the Paris-Harrington theorem was independent of first order Peano arithmetic.) This proposal generated a lot of excitement, and also some controversy. While many logicians hoped that Solovay was right, some (most notably Kriesel and MacIntyre) felt that present methods for proving independence results from weak systems of arithmetic were simply not powerful enough for his proposal to stand any chance of success. On the other hand, Graham did not even believe that $W(n, c)$ was not primitive recursive, and he conjectured that

$$W(n, 2) \leq G(n) = 2^{2^{\dots^2}}, \quad \text{a tower of height } n.$$

Some evidence in favor of Solovay's position came from recent work of Girard, who made a proof-theoretic analysis of the Furstenberg-Weiss proof of Szemerédi's theorem and showed that it yielded an upper bound for $W(n, c)$

which was essentially the Ackermann function. Considering how different the various proofs were, the fact that they give the same upper bound is very striking. Certainly the present author has to admit that he thought that Solovay's conjecture that $W(n, c)$ was not primitive recursive was more convincing than the opposing conjecture of Graham. However, in this paper, we shall see that Graham was correct: $W(n, c)$ is a primitive recursive function. In fact, we shall show that $W(n, c) \in \mathcal{E}^5$. The question of whether $W(n, 2) \leq G(n)$, or even whether $W(n, 2) \in \mathcal{E}^4$, remains open.

Since the sixties, many extensions of van der Waerden's theorem have been proved. The first of these was the Hales-Jewett theorem [9], which now forms the cornerstone of Ramsey theory. In order to state this result, we need to introduce some notation. The k -cube over n elements ${}^k n$ is defined by

$${}^k n = \{\eta = (\eta(0), \dots, \eta(k-1)) \mid \eta(i) \in \{0, 1, \dots, n-1\}\}.$$

Let $1 \leq l \leq k$. We shall define what we mean by an l -parameter subset of ${}^k n$. Let

$$\{0, 1, \dots, k-1\} = A_0 \cup A_1 \cup \dots \cup A_l$$

be a partition with $A_i \neq \emptyset$ for $0 \leq i \leq l-1$. Let $f: A_l \rightarrow \{0, 1, \dots, n-1\}$ be any function. We define a map $\hat{f}: {}^l n \rightarrow {}^k n$ by

$$\hat{f}(y_0, \dots, y_{l-1}) = (x_0, \dots, x_{k-1})$$

where

$$\begin{aligned} x_i &= f(i) & \text{if } i \in A_l, \\ x_i &= y_j & \text{if } i \in A_j, \ 0 \leq j \leq l-1. \end{aligned}$$

An l -parameter subset of ${}^k n$ is a set that is the range of \hat{f} for some choice of A_0, A_1, \dots, A_l, f . A 1-parameter subset is called a line.

Hales-Jewett Theorem [9]. *For all n, c there exists an integer $HJ(n, c)$ such that whenever $k \geq HJ(n, c)$ and ${}^k n$ is c -colored, then there exists a monochromatic line.*

As all known proofs of this result proceeded by a double induction on n and c , it was not known whether the function $HJ(n, c)$ was primitive recursive. In §1, we shall present a new proof which shows that $HJ(n, c)$ is indeed primitive recursive; in fact, $HJ(n, c) \in \mathcal{E}^5$. This easily implies that $W(n, c) \in \mathcal{E}^5$.

In 1971, Graham and Rothschild proved a far-reaching generalization of the Hales-Jewett theorem.

Graham-Rothschild Theorem [7]. *For all n, t, l, c there exists an integer $GR(n, t, l, c)$ such that whenever $k \geq GR(n, t, l, c)$ and d is a c -coloring of the l -parameter subsets of ${}^k n$, then there exists a t -parameter subset of ${}^k n$, all of whose l -parameter subsets have the same color.*

A year later, Graham, Leeb, and Rothschild went on to prove the following theorem, which had been conjectured by Rota.

Affine Ramsey Theorem [5]. For all q, t, l, c there exists an integer $n = N(q, t, l, c)$ such that if the l -subspaces of $AG(n, q)$ are c -colored, then there exists a t -subspace, all of whose l -subspaces have the same color.

Spencer [13] later found a beautiful and transparent proof of this result. The Hales-Jewett theorem is used in the proofs of both of the above theorems as the basis step of a double induction. In §§2 and 3, we will present new proofs which avoid the use of double induction, and hence show that the above functions are also primitive recursive. More precisely, both functions lie in \mathcal{E}^6 .

A number of interesting questions remain open, the most notable being whether there is a density version of the Hales-Jewett theorem.

Question 1. Does there exist for each $\varepsilon > 0$ and $n \in \mathbb{N}$ an integer $k = k(n, \varepsilon)$ such that if $A \subseteq {}^k n$ with $|A| \geq \varepsilon n^k$, then A contains a line?

I do not even know the answer to the following question.

Question 2. Does there exist for each $\varepsilon > 0$ and $n, c \in \mathbb{N}$ an integer $k = k(n, c, \varepsilon)$ with the following property? If d is a c -coloring of ${}^k n$, then there exist pairwise disjoint lines

$$\begin{aligned} L_1 &= \{\underline{x}_i^1 \mid 0 \leq i \leq n-1\}, \\ &\vdots \\ L_N &= \{\underline{x}_i^N \mid 0 \leq i \leq n-1\} \end{aligned}$$

such that $d(\underline{x}_0^r) = d(\underline{x}_1^r)$ for all $1 \leq r \leq N$ and $|\bigcup_{r=1}^N L_r|/n^k \geq 1 - \varepsilon$.

1. HALES-JEWETT AND VAN DER WAERDEN NUMBERS

In this section, we will give primitive recursive bounds for the Hales-Jewett number $HJ(n, c)$ and hence also for the van der Waerden number $W(n, c)$. We begin by explaining the main idea of the proof.

We shall argue by induction on n that the number $HJ(n, c)$ exists. However, when trying in the induction step to show that $H(n+1, c)$ exists, we are only allowed to use the existence of $HJ(n, c)$. If we also used the existence of $HJ(n, c')$ for very large values of c' , then our argument would become a double induction and this would lead to Ackermann-like bounds. Let k be some (as yet unspecified) integer. We shall attempt to prove that $k \geq HJ(n+1, c)$. So let d be a c -coloring of ${}^k(n+1)$. We shall try to reduce the problem to one concerning c -colorings of ${}^k n$. Let $l = HJ(n, c)$. A natural approach is to try to find an l -parameter subset of ${}^k(n+1)$, say X , in which two letters cannot be distinguished by d . In more detail, we first partition the set $\{0, 1, \dots, k-1\}$ into l pairwise disjoint nonempty subsets A_0, \dots, A_{l-1} . Each A_i will eventually contribute one degree of freedom to X , but one in which the letters $n-1$ and n cannot be distinguished by d . For each $i < l$, we will choose a nonempty subset B_i of A_i and a function $\eta_i \in {}^{A_i \setminus B_i}(n+1)$. Then X will be the l -parameter

subset consisting of those $\eta \in {}^k(n+1)$ such that

1.1.1. for $s \in A_i \setminus B_i$, $\eta(s) = \eta_i(s)$;

1.1.2. if $s_1, s_2 \in B_i$, then $\eta(s_1) = \eta(s_2)$.

To say that $n-1$ and n cannot be distinguished by d means the following.

1.1.3. Suppose that $\eta, \tau \in X$ and that there exists $i < l$ such that

(a) $\eta(s) = \tau(s)$ for $s \in k \setminus B_i$;

(b) $\eta(s) = n-1$ and $\tau(s) = n$ for $s \in B_i$.

Then $d(\eta) = d(\tau)$.

But now it is easy to see that $k \geq HJ(n+1, c)$. First, by repeated applications of 1.1.3, we find that the following holds.

1.1.4. Suppose that $\eta, \tau \in X$ and that for each $s < k$

(a) $\eta(s) < n-1$ if and only if $\tau(s) < n-1$;

(b) if $\eta(s) < n-1$, then $\eta(s) = \tau(s)$.

Then $d(\eta) = d(\tau)$.

Now consider the canonical embedding $\pi: {}^l(n+1) \rightarrow X \subseteq {}^k(n+1)$. Let $d_1 = d \circ \pi$ be the induced c -coloring of ${}^l(n+1)$. If $d_2 = d_1 \upharpoonright {}^l n$, then there is certainly a d_2 -monochromatic line of ${}^l n$, say $L_0 = \{\tau_0, \dots, \tau_{n-1}\}$. But then 1.1.4 implies that $\pi[L_0] = \{\pi(\tau_0), \dots, \pi(\tau_{n-1})\}$ can be extended to a d -monochromatic line L of ${}^k(n+1)$.

The sets A_i, B_i and the functions $\eta_i \in {}^{A_i \setminus B_i}(n+1)$ will be chosen in a simple way. Namely, k will have the form ml for some integer m , and we will let $A_i = [mi, m(i+1))$. Then for some $a_i < b_i < m$, we will let $B_i = [mi + a_i, mi + b_i)$ and

$$\begin{aligned} \eta_i(mi+r) &= n \quad \text{if } r < a_i, \\ &= n-1 \quad \text{if } b_i \leq r < m. \end{aligned}$$

Consider our requirement 1.1.3. We want to replace the i th block of η

$$\begin{array}{c|c|c} n & n-1 & n-1 \\ \hline mi+a_i & mi+b_i & \end{array}$$

by the i th block of τ

$$\begin{array}{c|c|c} n & n & n-1 \\ \hline mi+a_i & mi+b_i & \end{array}$$

without changing the color. This can be regarded as saying that a_i and b_i should be very similar with respect to some information which is stored by $\{a_j, b_j | j \neq i\}$. (Imagine that we obtain τ from η by replacing $mi + a_i$ by $mi + b_i$.) When we try to express this more precisely, we find that we have translated the Hales-Jewett theorem into the problem which is dealt with in Lemma 1.3.

Definition 1.2. For $l, c \in \mathbb{N}$, $f(l, c)$ is the least integer k satisfying the following property.

(*) If $\{d_j | j < l\}$ are c -colorings of ^{2l-1}k , then there exist $a_j < b_j < k$ for $j < l$ such that

$$\begin{aligned} d_j(a_0, b_0, \dots, a_{j-1}, b_{j-1}, a_j, a_{j+1}, b_{j+1}, \dots, a_{l-1}, b_{l-1}) \\ = d_j(a_0, b_0, \dots, a_{j-1}, b_{j-1}, b_j, a_{j+1}, b_{j+1}, \dots, a_{l-1}, b_{l-1}) \end{aligned}$$

for all $j < l$.

Lemma 1.3. (a) $f(1, c) = c + 1$.

(b) For $l \geq 1$, $f(l + 1, c) \leq c^{f(l, c)^{2l}} + 1$.

Proof. (a) Trivial.

(b) Let $m = f(l, c)$, $k = c^{m^{2l}} + 1$, and let $\{d_j | j < l + 1\}$ be c -colorings of ^{2l+1}k . Define a function

$$d^1: k \rightarrow \{g | g: ^{2l}m \rightarrow c\}$$

as follows. For each $a < k$,

$$d^1(a)(a_0, b_0, \dots, a_{l-1}, b_{l-1}) = d_l(a_0, b_0, \dots, a_{l-1}, b_{l-1}, a).$$

Then there exist $a_l < b_l < k$ such that $d^1(a_l) = d^1(b_l)$. Now define c -colorings $\{d_j^2 | j < l\}$ of ^{2l-1}m by

$$\begin{aligned} d_j^2(a_0, b_0, \dots, a_{j-1}, b_{j-1}, a_j, a_{j+1}, b_{j+1}, \dots, a_{l-1}, b_{l-1}) \\ = d_j(a_0, b_0, \dots, a_{j-1}, b_{j-1}, a_j, a_{j+1}, b_{j+1}, \dots, a_{l-1}, b_{l-1}, a_l, b_l). \end{aligned}$$

By the definition of m , there exist $a_j < b_j < m$ for $j < l$ such that (*) holds with respect to $\{d_j^2 | j < l\}$. But now the integers $a_j < b_j$ for $j < l + 1$ satisfy (*) with respect to $\{d_j | j < l + 1\}$. \square

Lemma 1.4. (a) $HJ(1, c) = 1$.

(b) For $n \geq 1$,

$$HJ(n + 1, c) \leq HJ(n, c) \times f(HJ(n, c), c^{(n+1)^{HJ(n, c)}}).$$

Proof. (a) Trivial.

(b) Let $l = HJ(n, c)$, $m = f(l, c^{(n+1)^l})$, and $k = lm$. Let d be a c -coloring of $^k(n + 1)$. Define colorings $\{d_j^1 | j < l\}$ of ^{2l-1}m as follows. If $j < l$, then

$$d_j^1(a_0, b_0, \dots, a_{j-1}, b_{j-1}, a_j, a_{j+1}, b_{j+1}, \dots, a_{l-1}, b_{l-1})$$

is a function with domain $^l(n + 1)$ and range included in the range of d such that each $\eta \in ^l(n + 1)$ is sent to $d(\nu_\eta)$, where $\nu_\eta \in ^k(n + 1)$ is defined as follows.

(i) Suppose that $i < l$, $i \neq j$, $r < m$, and $a_i < b_i$. Then

$$\begin{aligned} \nu_\eta(mi+r) &= n \quad \text{if } r < a_i, \\ &= n-1 \quad \text{if } b_i \leq r < m, \\ &= \eta(i) \quad \text{if } a_i \leq r < b_i. \end{aligned}$$

(ii) If $i < l$, $i \neq j$, $r < m$, and $a_i \geq b_i$, then $\nu_\eta(mi+r) = 0$. (This case will be of no importance in what follows.)

(iii) Suppose that $i = j$ and $r < m$. Then

$$\begin{aligned} \nu_\eta(mj+r) &= n \quad \text{if } r < a_j, \\ &= n-1 \quad \text{if } r \geq a_j. \end{aligned}$$

Clearly $|\text{rang } d_j^1| \leq c^{(n+1)^l}$. So there exist $a_j < b_j < m$ for $j < l$ satisfying 1.2(*) with respect to $\{d_j^1 \mid j < l\}$.

For each $i < l$, let $A_i = [mi, m(i+1))$, $B_i = [mi + a_i, mi + b_i)$ and define $\eta_i \in {}^{A_i \setminus B_i}(n+1)$ by

$$\begin{aligned} \eta_i(mi+r) &= n \quad \text{if } r < a_i, \\ &= n-1 \quad \text{if } b_i \leq r < m. \end{aligned}$$

Let X be the l -parameter subset of ${}^k(n+1)$ consisting of those η which satisfy conditions 1.1.1 and 1.1.2. Suppose that $\eta, \tau \in X$ and that there exists $i < l$ such that

- (a) $\eta(s) = \tau(s)$ for $s \in k \setminus B_i$,
- (b) $\eta(s) = n-1$ and $\tau(s) = n$ for $s \in B_i$.

Using the fact that

$$\begin{aligned} d_i^1(a_0, b_0, \dots, a_{i-1}, b_{i-1}, a_i, a_{i+1}, b_{i+1}, \dots, a_{l-1}, b_{l-1}) \\ = d_i^1(a_0, b_0, \dots, a_{i-1}, b_{i-1}, b_i, a_{i+1}, b_{i+1}, \dots, a_{l-1}, b_{l-1}), \end{aligned}$$

we find that $d(\eta) = d(\tau)$. Hence condition 1.1.3 is also satisfied, and this implies that $k \geq HJ(n+1, c)$. \square

Theorem 1.5. *The function $HJ(n, c)$ is primitive recursive. In fact, $HJ(n, c) \in \mathcal{E}^5$.*

Proof. This is an immediate consequence of Lemmas 1.3 and 1.4. \square

Theorem 1.6. *The function $W(n, c)$ is primitive recursive. In fact, $W(n, c) \in \mathcal{E}^5$.*

Proof. Let $k = HJ(n, c)$. Define a function $\Psi: {}^k n \rightarrow (n-1)k$ by $\Psi(\eta) = \sum_{i < k} \eta(i)$. Then Ψ maps each line of ${}^k n$ onto an n -term arithmetic progression. Hence $W(n, c) \leq (n-1)HJ(n, c)$. \square

We end this section with some remarks on extensions of the Hales-Jewett theorem.

Definition 1.7. $HJ(n, l, c)$ is the least natural number k such that if ${}^k n$ is c -colored, then there exists a monochromatic l -parameter subset of ${}^k n$.

It is well known that the existence of $HJ(n, l, c)$ follows easily from that of $HJ(n, c)$ (for example, see [8]). Alternatively, essentially the same argument as that in Lemma 1.4 yields.

Lemma 1.8. (a) $HJ(1, l, c) = l$.

(b) $HJ(n + 1, l, c) \leq HJ(n, l, c) \times f(HJ(n, l, c), c^{(n+1)^{HJ(n, l, c)}})$. \square

The proof shows that the monochromatic l -parameter subset can always be chosen so that its degrees of freedom $A_0, \dots, A_{l-1} \subset k$ are pairwise disjoint convex hulls, i.e., whenever $i < j < l$ and $r \in A_i, s \in A_j$, then $r < s$.

The $*$ -version of the Hales-Jewett theorem, due to Voigt [16], can also be proved using the method of Lemma 1.4.

In the appendix to this paper, we shall show that $HJ(2, l, c)$ is just a double exponential function.

2. GRAHAM-ROTHSCHILD NUMBERS

In this section, we shall present a new proof of the Graham-Rothschild theorem, which shows that the function $GR(n, t, l, c)$ is primitive recursive.

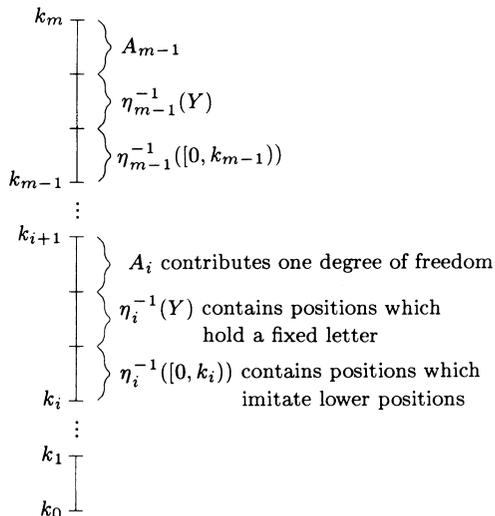
Definition 2.1. Let $RAM(t, l, c)$ denote the least natural number m such that $m \rightarrow (t)_c^l$ (i.e., whenever the set of l -subsets of m is c -colored, then there exists a t -subset of m , all of whose l -subsets have the same color).

Theorem 2.2. Let $n, t, l, c \in \mathbb{N}$ and let $m = RAM(t, l, c)$. Define k_i for $i \leq m$ by

$$k_0 = 0, \quad k_{i+1} = k_i + HJ(n + k_i, c^{(n+l)^{k_i+m-i-1}}).$$

Then $GR(n, t, l, c) \leq k_m$.

Picture of proof.



Proof. Let $k = k_m$. For notational simplicity, Y will be an alphabet on n letters which is disjoint from $k = \{0, 1, \dots, k-1\}$. Let $\{\lambda_\nu | \nu < l\}$ be a set of letters which is disjoint from $Y \cup \{0, 1, \dots, k-1\}$. We shall regard an l -parameter subset Φ of ${}^k Y$ as a function from k into $Y \cup \{\lambda_\nu | \nu < l\}$ such that

- (i) each λ_ν is included in the range of Φ ;
- (ii) $\min \Phi^{-1}(\lambda_i) < \min \Phi^{-1}(\lambda_j)$ for $i < j < l$.

Let d be a c -coloring of the l -parameter subsets of ${}^k Y$.

We shall define by downwards induction on $i < m$ a nonempty subset A_i of $[k_i, k_{i+1})$ and a function η_i from $B_i = [k_i, k_{i+1}) \setminus A_i$ into $Y \cup [0, k_i)$, R_i will denote the set of l -parameter subsets Φ of ${}^k Y$ which satisfy the following conditions.

2.2.1. If $i \leq j < m$, $s \in B_j$, and $\eta_j(s) \in Y$, then $\Phi(s) = \eta_j(s)$.

2.2.2. If $i \leq j < m$, $s \in B_j$, and $\eta_j(s) \in [0, k_j)$, then $\Phi(s) = \Phi(\eta_j(s))$.

2.2.3. If $i \leq j < m$, then $\Phi \upharpoonright A_j$ is a constant function.

A_i and η_i will be chosen so that the following condition holds.

(*) _{i} . Suppose that $\Phi, \Psi \in R_i$ satisfy

2.2.4. $\Phi \upharpoonright A_j = \Psi \upharpoonright A_j$ for $i < j < m$;

2.2.5. $\Phi \upharpoonright [0, k_i) = \Psi \upharpoonright [0, k_i)$;

2.2.6. if $\Phi \upharpoonright A_i$ is constantly λ_ν for some $\nu < l$, then $\min \Phi^{-1}(\lambda_\nu) < k_i$; and similarly for $\Psi \upharpoonright A_i$.

Then $d(\Phi) = d(\Psi)$.

Suppose, for the moment, that the above induction has been accomplished. We shall show how to complete the proof of the theorem. Let S be the m -parameter subset of ${}^k Y$ consisting of those $\eta \in {}^k Y$ which satisfy

- (a) for all $i < m$, if $s \in B_i$ and $\eta_i(s) \in Y$, then $\eta(s) = \eta_i(s)$;
- (b) for all $i < m$, if $s \in B_i$ and $\eta_i(s) \in [0, k_i)$, then $\eta(s) = \eta(\eta_i(s))$;
- (c) for all $i < m$, $\eta \upharpoonright A_i$ is a constant function.

Let Φ, Ψ be l -parameter subsets of S , considered as functions from k to $Y \cup \{\lambda_\nu | \nu < l\}$. Suppose that

$$(+) \quad \min \Phi^{-1}(\lambda_\nu) = \min \Psi^{-1}(\lambda_\nu) \quad \text{for all } \nu < l.$$

We claim that $d(\Phi) = d(\Psi)$. To see this, let $s = \min\{r | \Phi(r) \neq \Psi(r)\}$ and let $s \in [k_i, k_{i+1})$. By (a) and (b), we must have that $s \in A_i$. Define Φ' by

$$\begin{aligned} \Phi'(r) &= \Psi(r) \quad \text{if } r \in A_i, \\ &= \Phi'(\eta_j(r)) \quad \text{if } i \leq j < m, r \in B_j \text{ and } \eta_j(r) \in [0, k_j), \\ &= \Phi(r) \quad \text{otherwise.} \end{aligned}$$

By (+) and $(*)_i$, $d(\Phi) = d(\Phi')$. Continuing in this fashion, we can transform Φ stepwise into Ψ without changing the color. Thus the claim is established.

Hence the color of an l -parameter subset Φ of S is determined by the sequence $\langle i_0, \dots, i_{l-1} \rangle$ where $\min \Phi^{-1}(\lambda_\nu) \in A_{i_\nu}$ for $\nu < l$. Let d^* be the coloring of the l -subsets of $\{0, 1, \dots, m-1\}$ defined by $d^*(\{i_0, \dots, i_{l-1}\}) = d(\Phi)$. Since $m \rightarrow (t)_c^l$, there exists a t -subset $U \subseteq \{0, 1, \dots, m-1\}$ such that $d^*(\{i_0, \dots, i_{l-1}\})$ is constant for all $i_0 < \dots < i_{l-1} \in U$. Finally choose an element $y_0 \in Y$ and let

$$T = \{\eta \in S \mid \eta(s) = y_0 \text{ for } s \in A_i, i \notin U\}.$$

Then T is a t -parameter subset of ${}^k Y$, all of whose l -parameter subsets have the same color.

Hence to complete the proof of the theorem, we need only check that the inductive construction of A_i and η_i can be carried out. Suppose that A_j and η_j have been defined for all $i < j < m$. Define a coloring d_1 of ${}^{[k_i, k_{i+1})}(Y \cup [0, k_i))$ as follows.

(1) For each $\eta \in {}^{[k_i, k_{i+1})}(Y \cup [0, k_i))$, $d_1(\eta)$ is a function with domain

$$\mathcal{D}_i = \{\Psi \upharpoonright [0, k_i) \cup [k_{i+1}, k_m) \mid \Psi \in R_j \text{ for } i < j < m\}.$$

Thus $|\mathcal{D}_i| \leq (n+l)^{k_i+m-i-1}$.

(2) For each $\Psi \in \mathcal{D}_i$ and $\eta \in {}^{[k_i, k_{i+1})}(Y \cup [0, k_i))$, define

$$\eta(\Psi) \in {}^{[0, k)}(Y \cup \{x_r : r < l\})$$

by

$$\begin{aligned} \eta(\Psi)(s) &= \eta(s) && \text{if } s \in [k_i, k_{i+1}) \text{ and } \eta(s) \in Y, \\ &= \eta(\Psi)(r) && \text{if either (a) } s \in B_j \text{ for some } i < j < m \text{ and } \eta_j(s) = r \in \\ & && [0, k_j), \text{ or (b) } s \in [k_i, k_{i+1}) \text{ and } \eta(s) = r \in [0, k_i), \\ &= \Psi(s) && \text{otherwise.} \end{aligned}$$

Then $d_1(\eta) = \langle d(\eta(\Psi)) \mid \Psi \in \mathcal{D}_i \rangle$. Of course, $d(\eta(\Psi))$ is not always defined, since there may exist $\nu < l$ such that $\lambda_\nu \notin \text{rang } \eta(\Psi)$. In this case, let $d(\eta(\Psi)) = c_0$ for some fixed $c_0 \in \text{rang } d$.

Clearly

$$|\text{rang } d_1| \leq c^{(n+l)^{k_i+m-i-1}}.$$

So by the definition of k_{i+1} , ${}^{[k_i, k_{i+1})}(Y \cup [0, k_i))$ contains a d_1 -monochromatic line L , say determined by

$$\emptyset \neq A_i \subseteq [k_i, k_{k+1}) \quad \text{and} \quad \eta_i \in {}^{[k_i, k_{i+1}) \setminus A_i}(Y \cup [0, k_i)).$$

It is now clear that with this choice of A_i and η_i , condition $(*)_i$ holds. This completes the proof of the theorem. \square

3. The affine Ramsey theorem

In this section, we will give a proof of the affine Ramsey theorem, which yields primitive recursive upper bounds. The proof is heavily influenced by that of Spencer [13]. Throughout F will be an arbitrary but fixed finite field. If V is an affine space over F , then $\left[\begin{smallmatrix} V \\ l \end{smallmatrix} \right]$ will denote the set of l -dimensional subspaces of V . The affine space $\{(x_0, \dots, x_{k-1}) \mid x_i \in F\}$ will be denoted by F^k .

Definition 3.1. Let V be an affine space over F and let χ be a coloring of $\left[\begin{smallmatrix} V \\ l \end{smallmatrix} \right]$. Let $B \in \left[\begin{smallmatrix} V \\ u+1 \end{smallmatrix} \right]$ and let $p: B \rightarrow F^u$ be a surjective projection. We say that B is special with respect to χ, p if whenever $S_1, S_2 \in \left[\begin{smallmatrix} B \\ l \end{smallmatrix} \right]$ satisfy $p(S_1) = p(S_2)$, then $\chi(S_1) = \chi(S_2)$.

Notice that the condition on S_1, S_2 is nontrivial only when $\dim p(S_1) = l$. For if $\dim p(S_1) = l - 1$, then $S_1 = p^{-1}(p(S_1)) = p^{-1}(p(S_2)) = S_2$. The next result is a slight generalization of Lemma 1 [13].

Lemma 3.2. Let $m, l \in \mathbb{N}$ and let $\bar{c} = \langle c_i \mid i \leq l \rangle$ be a finite sequence of natural numbers. Let

$$k = m + HJ \left(|F|^{m+1}, \prod_{i=0}^l c_i^{\nu_i} \right),$$

where $\nu_i = \left| \left[\begin{smallmatrix} F^m \\ i \end{smallmatrix} \right] \right|$ for $i \leq l$. Let $p: F^k \rightarrow F^m$ be the natural projection obtained by taking the first m coordinates. Then if χ_i is a c_i -coloring of $\left[\begin{smallmatrix} F^k \\ i \end{smallmatrix} \right]$ for $i \leq l$, there exists $B \in \left[\begin{smallmatrix} F^k \\ m+1 \end{smallmatrix} \right]$ which is simultaneously special with respect to χ_i, p for each $i \leq l$. \square

Definition 3.3. If X is a set, then $\binom{X}{i}$ denotes the set of i -subsets of X . Let $t, l \in \mathbb{N}$ and let $\bar{c} = \langle c_i \mid i \leq l \rangle$ be a finite sequence of natural numbers. We write $m \rightarrow (t)_{\bar{c}}^l$ if whenever $d_i: \binom{m}{i} \rightarrow c_i$ for $i \leq l$, then there exists $T \in \binom{m}{t}$ such that $d_i \upharpoonright \binom{T}{i}$ is constant for all $i \leq l$.

Definition 3.4. Let $t, l \in \mathbb{N}$ and let $\bar{c} = \langle c_i \mid i \leq l \rangle$ be a finite sequence of natural numbers. $A_F(t, l, \bar{c})$ is the least k such that whenever $d_i: \left[\begin{smallmatrix} F^k \\ i \end{smallmatrix} \right] \rightarrow c_i$ for $i \leq l$, then there exists $S \in \left[\begin{smallmatrix} F^k \\ t \end{smallmatrix} \right]$ such that $d_i \upharpoonright \left[\begin{smallmatrix} S \\ i \end{smallmatrix} \right]$ is constant for all $i \leq l$.

Definition 3.5. Let

$$\Phi(m, l, \bar{c}) = m + HJ \left(|F|^{m+1}, \prod_{i=0}^l c_i^{\nu_i} \right)$$

be the function given by Lemma 3.2.

Theorem 3.6. Let $t, l \in \mathbb{N}$ and let $\bar{c} = \langle c_i \mid i \leq l \rangle$ be a finite sequence of natural numbers. Suppose that $m \rightarrow (t)_{\bar{c}}^l$. For each $i \leq l$, define c_i^j by downward

induction on $j \leq m$ by

$$\begin{aligned} c_i^m &= c_i, \\ c_i^j &= c_i^{j+1} \times c_{i+1}^{j+1} \quad \text{if } i < l, \\ &= c_i \quad \text{if } i = l. \end{aligned}$$

Now define k_j for $j \leq m$ by

$$k_0 = 0, \quad k_{j+1} = \Phi(k_j, l, \langle c_i^{j+1} | i \leq l \rangle).$$

Then $A_F(t, l, \bar{c}) \leq k_m$.

Proof. For each $j \leq m$, let $V_j = F^{k_j}$ and let $p_j: V_{j+1} \rightarrow V_j$ be the natural projection. Let d_i be a c_i -coloring of $[V_i^m]$ for $i \leq l$. Define by downwards induction subspaces $V_j^+ \leq V_{j+1}$ for $j < m$ and colorings $\langle d_i^j | i \leq l \rangle$ for $j \leq m$ such that the following conditions hold.

- (a) $\dim V_j^+ = k_j + 1$.
- (b) For each $i \leq l$, d_i^j is a c_i^j -coloring of $[V_i^j]$.
- (c) For each $i \leq l$, $d_i^m = d_i$.
- (d) If $j < m$, $i \leq l$, $L_1, L_2 \in [V_i^+]$ and $p_j(L_1) = p_j(L_2)$, then $d_i^{j+1}(L_1) = d_i^{j+1}(L_2)$. (Since $k_{j+1} = \Phi(k_j, l, \langle c_i^{j+1} | i \leq l \rangle)$, a suitable subspace V_j^+ can always be found.)
- (e) If $j < m$ and $L \in [V_j^+]$, then $d_i^j(L) = d_i^{j+1}(L')$, where $L' \in [V_i^+]$ satisfies $p_j(L') = L$. (This is well defined by (d).)
- (f) If $j < m$, $i < l$, and $L \in [V_j^+]$, then $d_i^j(L) = \langle d_i^{j+1}(L_1), d_{i+1}^{j+1}(L_2) \rangle$, where $L_1 \in [V_i^+]$ and $L_2 \in [V_{i+1}^+]$ satisfy $p_j(L_1) = L = p_j(L_2)$. (Again, this is well defined by (d).)

Now choose j -dimensional subspaces $S_j \leq V_j$ for $j \leq m$ such that

$$\begin{aligned} S_0 &= V_0, \\ S_{j+1} &= \text{some } S \in \left[\begin{array}{c} V_j^+ \\ j+1 \end{array} \right] \text{ such that } p_j(S) = S_j. \end{aligned}$$

Let $0 \leq i \leq l$ and let $L \in [S_i^m]$. Consider the sequence of spaces $L_j \leq V_j$, $0 \leq j \leq m$, defined by

$$L_m = L, \quad L_{j-1} = p_{j-1}(L_j).$$

Let $j_1 < \dots < j_i$ be the numbers j such that $\dim L_j < \dim L_{j+1}$. Then (e) and (f) imply that $d_i(L)$ depends only on $\{j_1, \dots, j_i\}$. Let $\chi_i: \binom{m}{i} \rightarrow c_i$ be the induced coloring. Since $m \rightarrow (t)_{\bar{c}}^l$ there exists $U \in \binom{m}{i}$ such that $\chi_i \upharpoonright \binom{U}{i}$ is constant for $0 \leq i \leq l$. Finally choose a sequence of subspaces $T_j \leq V_j$ for $j \leq m$ such that $T_0 = V_0$, and

- (i) if $j \in U$, then $T_{j+1} \in \left[\begin{array}{c} S_{j+1} \\ \dim(T_j)+1 \end{array} \right]$ satisfies $p_j(T_{j+1}) = T_j$,

(ii) if $j \notin U$, then $T_{j+1} \in [{}_{\dim(T_j)}^{S_{j+1}}]$ satisfies $p_j(T_{j+1}) = T_j$.

Then $T_m \in [{}^V_m]$ and $d_i \upharpoonright [{}^T_m]$ is constant for each $0 \leq i \leq l$. \square

Exactly the same proof works for the parameter systems, introduced by Spencer in [13]. This means that all of the variants of the Graham-Rothschild theorem proved in [7] now have primitive recursive upper bounds.

4. APPENDIX

Proposition 4.1. $HJ(2, l, c) \leq c^{2^{2^l - 1}}$

Proof. Fix a natural number c . Define integers r_l by

$$r_1 = c + 1, \quad r_{l+1} = c \prod_{j=1}^l \binom{r_j}{2} + 1.$$

It is enough to show that $HJ(2, l, c) \leq \sum_{j=1}^l r_j$. To accomplish this, we shall show that if $k = \sum_{j=1}^l r_j$ and $d: {}^k 2 \rightarrow c$, then there exist $a_i < b_i$ for $1 \leq i \leq l$ such that the following conditions hold.

(1) $\sum_{j=1}^{i-1} r_j \leq a_i < b_i < \sum_{j=1}^i r_j$.

(2) d is constant on the l -parameter set consisting of those $\eta \in {}^k 2$ such that for each $1 \leq i \leq l$,

$$\begin{aligned} \eta(s) &= 0 && \text{if } \sum_{j=1}^{i-1} r_j \leq s < a_i, \\ &= 1 && \text{if } b_i \leq s < \sum_{j=1}^i r_j, \\ &= \eta(a_i) && \text{if } a_i \leq s < b_i. \end{aligned}$$

Suppose that this holds for some $l \geq 1$, and let $m = \sum_{j=1}^{l+1} r_j$. For each $k \leq t < m$, define an embedding $\pi_t: {}^k 2 \rightarrow {}^m 2$ by

$$\begin{aligned} \pi_t(\eta)(s) &= \eta(s) && \text{if } s < k, \\ &= 0 && \text{if } k \leq s < t, \\ &= 1 && \text{if } t \leq s < m. \end{aligned}$$

Let $d_t = d \circ \pi_t$. By the induction hypothesis, there exists $P_t = \langle a'_i, b'_i \mid 1 \leq i \leq l \rangle$ satisfying (1) and (2) with respect to d_t . Let d_t be the constant color c_t on the l -parameter set defined by P_t . Then there exist $t_0 < t_1$ such that $\langle P_{t_0}, c_{t_0} \rangle = \langle P_{t_1}, c_{t_1} \rangle$. Hence we can take

$$\begin{aligned} \langle a_j, b_j \rangle &= \langle a_j^{t_0}, b_j^{t_0} \rangle && \text{if } 1 \leq j \leq l, \\ &= \langle t_0, t_1 \rangle && \text{if } j = l + 1. \quad \square \end{aligned}$$

¹ Probably noticed by anybody who looks at the problem.

Remark 4.2. Hence $HJ(3, 1, c)$ can be bounded similarly, as (see §1) $HJ(3, 1, c) \leq HJ(2, c + 1, c)$.

REFERENCES

1. W. Ackermann, *Zum Hilbertschen Aufbau der reellen Zahlen*, Math. Ann. **99** (1928), 118–133.
2. E. R. Berlekamp, *A construction for partitions which avoid long arithmetic progressions*, Canad. Math. Bull. **11** (1968), 409–414.
3. P. Erdős and R. Rado, *Combinatorial theorems on classification of subsets of a given set*, Proc. London Math. Soc. **2** (1952), 417–439.
4. H. Furstenberg, *Recurrence in ergodic theory and combinatorial number theory*, Princeton Univ. Press, Princeton, N.J., 1982.
5. R. L. Graham, K. Leeb, and B. L. Rothschild, *Ramsey's theorem for a class of categories*, Adv. in Math. **8** (1972), 417–433.
6. R. L. Graham and V. Rödl, *Numbers in Ramsey theory*, Surveys in Combinatorics 1987, London Math. Soc. Lecture Notes **123** (1987), 111–153.
7. R. L. Graham and B. L. Rothschild, *Ramsey's theorem for n -parameter sets*, Trans. Amer. Math. Soc. **159** (1971), 257–292.
8. R. L. Graham, B. L. Rothschild, and J. H. Spencer, *Ramsey theory*, Wiley-Interscience Ser. in Discrete Math., New York, 1980.
9. A. W. Hales and R. I. Jewett, *Regularity and positional games*, Trans. Amer. Math. Soc. **106** (1963), 222–229.
10. J. Ketonen and R. Solovay, *Rapidly growing Ramsey functions*, Ann. of Math. **113** (1981), 267–314.
11. H. E. Rose, *Subrecursion: functions and hierarchies*, Oxford Logic Guides 9, Oxford Univ. Press, Oxford, 1984.
12. K. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 104–109.
13. J. H. Spencer, *Ramsey's theorem for spaces*, Trans. Amer. Math. Soc. **249** (1979), 363–371.
14. E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245.
15. B. L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw Arch. Wisk. **15** (1927), 212–216.
16. B. Voigt, *The partition problem for finite abelian groups*, J. Comb. Theory Ser. A. **28** (1980), 257–271.

INSTITUTE OF MATHEMATICS, THE HEBREW UNIVERSITY, JERUSALEM, ISRAEL

DEPARTMENT OF MATHEMATICS, RUTGERS UNIVERSITY, NEW BRUNSWICK, NEW JERSEY 08903

AT&T BELL LABORATORIES, 600 MOUNTAIN AVENUE, MURRAY HILL, JERSEY 07974