Chapter 7

The story of arithmetic progressions

Our main goal in the chapter is to establish the famous theorem of Szemerédi on arithmetic progressions, which says that if $\delta > 0$ and $k \in \mathbb{N}$, then there is an N such that for every $n \geq N$ every subset $A \subset [n]$ with $|A| > \delta n$ contains a k-term arithmetic progression. We prove it in Section 7.3 in Theorem 7.3.2 as an easy corollary of the density Hales–Jewett theorem (Theorem 7.3.1), which says that if $\delta > 0$ and $k \in \mathbb{N}$, then for every $n \geq N$ every subset $A \subset [k]^n$ with $|A| > \delta k^n$ contains a line. In the proof we follow closely the article of Dodos, Kanellopoulos and Tyros.

We explain the notions used. If $d, k > 0$ and a are integers, the *k-term arithmetic progression* (in $\mathbb{Z}$, starting at a and with common difference d) is the k-element set

$$\mathrm{AP}_k(a,d) := \{a, a+d, a+2d, \ldots, a+(k-1)d\} .$$

A *line* ℓ (often one speaks of a *combinatorial line*) in the set $[k]^n$ of words $u = a_1 a_2 \ldots a_n$, $a_i \in \{1, 2, \ldots, k\}$, is the k-element subset given by a word $v \in ([k] \cup \{-1\})^n \backslash [k]^n$:

$$\ell = \{v(i) \mid i \in [k]\}$$

where $v(i)$ arises from v by replacing each occurrence of -1 with i. For $A \subset [k]^n$, we write $\mathcal{L}(A)$ for the set of all lines contained in A. More generally, an *m-subspace* ($m \in \mathbb{N}$) V of $[k]^n$ is the k^m-element subset given by a word $v \in ([k] \cup \{-1, -2, \ldots, -m\})^n$ in which each of the m negative letters occurs at least once:

$$V = \{v(i_1, i_2, \ldots, i_m) \mid i_1, i_2, \ldots, i_m \in [k]\}$$

where $v(i_1, i_2, \ldots, i_m)$ arises from v by replacing for $j \in [m]$ each occurrence of $-j$ with i_j.

In Section 7.2 we give simple proofs for the first nontrivial cases of both theorems. The case $k = 2$ of the density Hales–Jewett theorem follows from the classical result of Sperner (Theorem 7.2.2) that the maximum size of a family of subsets of $[n]$, no two of them in inclusion, is $\binom{n}{\lfloor n/2 \rfloor}$. The case $k = 3$ of Szemerédi's theorem was proved by Roth more than 20 years before the general case. We give two proofs for Roth's theorem (Theorem 7.2.3), a combinatorial one, devised by Szemerédi in his attack on the general case, and an analytic one close to the original argument of Roth.

The initial Section 7.1 presents the original and easier "coloristic" versions of both theorems, as opposed to the stronger density versions in Sections 7.2 and 7.3 (the density version trivially implies the coloristic one). For $r \in \mathbb{N}$, a *coloring*, or an *r-coloring*, of a set X is a mapping $\chi : X \to [r]$ (it is often useful to replace $[r]$ with another r-element set). A *monochromatic set* $Y \subset X$ is one on which χ restricts to a constant mapping. One may call a set monochromatic even in situations when it is not directly a subset of X but it generates a monochromatic subset of X. For example, one may have a coloring $\chi : X = \binom{Z}{k} \to [r]$ of all k-element subsets of a set Z, and call $Y \subset Z$ monochromatic meaning that $\binom{Y}{k} \subset X$ is monochromatic. Van der Waerden's theorem (Theorem 7.1.1) is the coloristic version of Szemerédi's theorem: if $k, r \in \mathbb{N}$, then for every $n \geq N$ every coloring $\chi : [n] \to [r]$ contains a monochromatic k-term arithmetic progression. The original Hales–Jewett theorem (Theorem 7.1.3), preceding by almost 30 years the density version, says that if $k, r \in \mathbb{N}$, then for every $n \geq N$ every coloring $\chi : [k]^n \to [r]$ contains a monochromatic line. We prove van der Waerden's theorem from scratch, prove the Hales–Jewett theorem and deduce the former from the latter (which is trivial). We conclude the section with the theorem of Graham and Rotschild (Theorem 7.1.8), whose special case $p = 1$ and $r = 2$ is used in Section 7.3: if $k, m, r \in \mathbb{N}$ and $p \in \mathbb{N}_0$, then for every $n \geq N$ in every r-coloring of p-subspaces of $[k]^n$ there is a monochromatic m-subspace of $[k]^n$.

In Section 7.1 we also derive both upper and lower bounds on van der Waerden's numbers, defined as the first $n = n(k, r)$ for which the theorem holds. In Theorem 7.1.4 we deduce by a more sophisticated proof of the Hales–Jewett theorem, due to Shelah, an upper bound that is much better than the one provided by the original simple proof. As for lower bounds, Corollary 7.1.7 gives the almost best known lower bound, obtained by applying so called Lovász' local lemma in Proposition 7.1.6.

7.1 van der Waerden's theorem and the Hales–Jewett theorem

We start with van der Waerden's theorem which says that no coloring of any sufficiently long interval of integers can avoid monochromatic arithmetic progressions. It follows easily from the Hales–Jewett theorem: the mapping

$$f : [k]^n \to [kn], \ a_1 a_2 \ldots a_n \mapsto a_1 + a_2 + \cdots + a_n \ ,$$

sends every line to a k-term arithmetic progression, and so for every large n any r-coloring χ of $[kn]$ contains a monochromatic k-term AP because the pullback

r-coloring $\chi \circ f$ of $[k]^n$ contains, by the HJ theorem, a monochromatic line. Note that non-injectivity of f is not an issue here, but for reduction in the density versions one needs injective f. So it would suffice to prove just the HJ theorem, but we give a proof from scratch for the vdW theorem as well. After all, it took people over 35 years to go from arithmetic progressions to lines.

Theorem 7.1.1 (van der Waerden, 1927) *Let $k, r \in \mathbb{N}$. Then for every $n \geq n(k,r)$ every coloring $\chi : [n] \to [r]$ contains a monochromatic k-term arithmetic progression: there is an $\mathrm{AP}_k(a,d) \subset [n]$ such that*

$$\chi(a) = \chi(a+d) = \cdots = \chi(a+(k-1)d)\,.$$

For the proof we introduce some notions. We call a set $A \subset \mathbb{Z}$ an *(i,k)-fan* if it is a union of i k-term arithmetic progressions and a point f, the *focus of A*, extending each progression:

$$A = \bigcup_{j=1}^{i} A_j \cup \{f\} = \bigcup_{j=1}^{i} \mathrm{AP}_k(a_j, d_j) \cup \{f\},\ f = a_1 + kd_1 = \cdots = a_i + kd_i\,.$$

Each A_j focuses to f to create a $(k+1)$-term AP. For a coloring $\chi :\ A \to [r]$ we call A a *proper (i,k)-fan* (with respect to χ) if each A_j is monochromatic and the i colors $\chi(A_j)$ are distinct. Then $i \leq r$, and in the case $i = r$ the focus f, regardless of its color, forms with some A_j a monochromatic $(k+1)$-term AP. A coloring $\chi :\ \mathbb{Z} \to [r]$ is *D-periodic*, $D \in \mathbb{N}$, if $\chi(x) = \chi(x+D)$ for any $x \in \mathbb{Z}$. We make use of the following construction that increases the parameter i.

Lemma 7.1.2 *Suppose that $\chi :\ \mathbb{Z} \to [r]$ is a D-periodic coloring and*

$$A = \bigcup_{j=1}^{i} A_j \cup \{f\} = \bigcup_{j=1}^{i} \mathrm{AP}_k(a_j, d_j) \cup \{f\} \subset \mathbb{Z}$$

is a proper (i,k)-fan with focus f. Then

$$B = \bigcup_{j=1}^{i+1} B_j \cup \{f + kD\} = \bigcup_{j=1}^{i} \mathrm{AP}_k(a_j, d_j + D) \cup \mathrm{AP}_k(f, D) \cup \{f + kD\}$$

is an $(i+1,k)$-fan with focus $f + kD$, in which all B_j are monochromatic and $\chi(B_j) = \chi(A_j)$ for $1 \leq j \leq i$. Hence A contains a $(k+1)$-term monochromatic AP or B is a proper $(i+1,k)$-fan. We call B the D-stretch *of A.*

Proof. It is easy to check that B is an $(i+1,k)$-fan with focus $f + kD$. By the assumption on A and the coloring, $\chi(a_j + t(d_j + D)) = \chi(a_j + td_j) = \chi(a_j)$ for every $j \in [i]$ and $t \in [0, k-1]$ and so $\chi(B_j) = \chi(A_j)$. Similarly, $B_{i+1} = \mathrm{AP}_k(f, D)$ is monochromatic, with color $\chi(f)$. If $\chi(B_{i+1}) = \chi(A_j)$ for some $j \in [i]$, then f and A_j form a monochromatic $(k+1)$-term AP, else B is a proper $(i+1,k)$-fan as its $i+1$ APs have distinct colors. $\square$

Proof of Theorem 7.1.1. By induction on k. For $k = 2$ it trivially holds with $n(2,r) = r+1$. We assume that the theorem holds for $k \geq 2$ and every r, and show that it holds for $k+1$ and every r. For each fixed r we prove by induction on i, $1 \leq i \leq r$, that for large $n = n(k,i,r)$ any r-coloring of $[n]$ either contains a monochromatic $(k+1)$-term AP or a proper (i,k)-fan. For $i = r$ this gives, as we noticed above, presence of a monochromatic $(k+1)$-term AP in any r-coloring of $[n]$, so $n(k+1,r) = n(k,r,r)$ works.

For $i = 1$ we set $n = n(k,1,r) = 2n(k,r)$. By induction on k any r-coloring of $[n]$ contains a monochromatic $A = \mathrm{AP}_k(a,d) \subset [n/2]$, and then $A \cup \{a+kd\} \subset [n]$ is a proper $(1,k)$-fan with focus $a+kd$. We assume that $n(k,i,r)$ exists, for an i with $1 \leq i < r$, and show that $n(k,i+1,r)$ exists. We set

$$n = n(k,i+1,r) = 2n(k,r^N) \cdot N \quad \text{where} \quad N = n(k,i,r) .$$

Let $\chi : [n] \to [r]$ be a coloring. We consider the coloring $\psi : [n(k,r^N)] \to [r]^N$,

$$\psi(a) = (\chi((a-1)N+1), \chi((a-1)N+2), \ldots, \chi((a-1)N+N)) .$$

Thus $\psi(a) = \psi(b)$ means that $\chi((a-1)N+j) = \chi((b-1)N+j)$ for every $j \in [N]$. By the inductive assumption on k, in the coloring ψ there is a monochromatic $\mathrm{AP}_k(a,d) \subset [n(k,r^N)]$. We set $D = dN$. By the inductive assumption on i, in the coloring χ the interval $[(a-1)N+1, (a-1)N+N] \subset [n]$ contains a monochromatic $(k+1)$-term AP or it contains a proper (i,k)-fan A. In the former case we are done and in the latter case we consider the D-stretch B of A. Clearly, $B \subset [n]$. Since ψ is constant on the $\mathrm{AP}_k(a,d)$, we have $\chi(x) = \chi(x+D) = \cdots = \chi(x+(k-1)D)$ for any $x \in [(a-1)N+1, (a-1)N+N]$, and for the purposes of A and B the coloring χ behaves as D-periodic. It follows by Lemma 7.1.2 that $[n]$ contains a monochromatic $(k+1)$-term AP or B is a proper $(i+1,k)$-fan. □

We define $W(k,r)$, the *van der Waerden number*, to be the minimum n such that every r-coloring of $[n]$ contains a k-term AP. Similarly, the *Hales–Jewett number* $HJ(k,r)$ is the minimum n such that every r-coloring of $[k]^n$ has a monochromatic line. We noted above that

$$W(k,r) \leq k \cdot HJ(k,r) .$$

More generally, we define $HJ(k,m,r)$ to be the minimum n such that every r-coloring of $[k]^n$ contains a monochromatic m-subspace; $HJ(k,r) = HJ(k,1,r)$. We prove the Hales–Jewett theorem.

Theorem 7.1.3 (Hales and Jewett, 1963) *Let $k, r \in \mathbb{N}$. Then for every $n \geq N$ every coloring $\chi : [k]^n \to [r]$ contains a monochromatic line.*

Proof. We prove the recurrent inequalities

1. $HJ(k,m+1,r) \leq HJ(k,1,r) + HJ(k,m,r^{k^{HJ(k,1,r)}})$ and
2. $HJ(k+1,1,r+1) \leq HJ(k,1+HJ(k+1,1,r),r+1)$.

Together with the obvious $HJ(1,m,r) = m$ they prove by induction the existence of $HJ(k,m,r)$, and hence $HJ(k,r)$, for every $k,m,r \in \mathbb{N}$.

1. Let $M = HJ(k,1,r)$, $N = HJ(k,m,r^{k^M})$ and χ be an r-coloring of $[k]^{M+N}$. It induces an r^{k^M}-coloring χ' of $[k]^N$: we color each $v \in [k]^N$ by the list of colors $(\chi(uv) \mid u \in [k]^M)$. By the choice of N, there is a χ'-monochromatic m-subspace $V \subset [k]^N$. Its color $\chi'(V)$ gives an r-coloring χ'' of $[k]^M$: $\chi''(u) := \chi(uV)$. By the choice of M, there is a χ''-monochromatic line $\ell \subset [k]^M$. It follows that $\ell V \subset [k]^{M+N}$ is a χ-monochromatic $(m+1)$-subspace.

2. Let $M = HJ(k+1,1,r)$, $N = HJ(k,1+M,r+1)$ and $\chi : [k+1]^N \to [r+1]$ be a coloring. By the choice of N, there is a χ-monochromatic $(1+M)$-subspace $V \subset [k]^N$ with color $c \in [r+1]$; suppose V is given by the word $v \in ([k] \cup \{-1,-2,\ldots,-1-M\})^N$ (using each negative letter). If also $\chi(v(i_1,i_2,\ldots,i_{M+1})) = c$ for $i_1 = k+1$ and some $i_2,\ldots,i_{M+1} \in [k+1]$, we define the word $w \in ([k] \cup \{-1\})^N$ by replacing in v for $j \in [M+1]$ each occurrence of $-j$ with -1 if $i_j = k+1$, and with i_j if $i_j < k+1$ (the rest of v is unchanged). Then w defines a χ-monochromatic line $\ell \subset [k+1]^N$ (ℓ extends by one word a subline of V). In the complementary case we have an r-coloring $\chi' : [k+1]^M \to [r+1]\backslash\{c\}$: $\chi'(a_1a_2\ldots a_M) := \chi(v(k+1,a_1,\ldots,a_M))$. By the choice of M, there is a χ'-monochromatic line $\ell \subset [k+1]^M$. If ℓ is given by $u \in ([k+1] \cup \{-1\})^M$, we define the word $w \in ([k+1] \cup \{-1\})^N$ by replacing in v each occurrence of -1 with $k+1$ and for $j \in [2, M+1]$ each occurrence of $-j$ with u_{j-1} (the rest of v is unchanged). Then w defines a χ-monochromatic line $\ell' \subset [k+1]^N$. □

The two previous proofs do provide upper bounds on $W(k,r)$ and $HJ(k,r)$, but ones that are extremely large. To appreciate how large they are, we introduce the sequence of functions $f_j : \mathbb{N} \to \mathbb{N}$, $j = 1,2,\ldots$, where $f_1(n) = 2n$ and $f_{j+1}(n) = f_j(f_j(\ldots(f_j(1))\ldots))$ with n applications of $f_j(\cdot)$. So $f_2(n) = 2^n$ and $f_3(n)$ is so called tower function

$$f_3(n) = 2^{2^{\cdot^{\cdot^{2}}}} \quad \text{(with } n \text{ twos)}\ .$$

Each $f_{j+1}(n)$ grows to infinity with n much, much faster than $f_j(n)$. The *Ackerman function* $A : \mathbb{N} \to \mathbb{N}$ is the diagonal function $A(n) := f_n(n)$; $A(n)$ grows much, much faster than any $f_j(n)$. Thus $A(1) = 2$, $A(2) = 4$, $A(3) = 16$ but the next value $A(4) = f_3(f_3(f_3(f_3(1)))) = f_3(65536)$, the tower of twos with height 65536, is beyond imagination. When one works out the double recursions in the proofs of Theorems 7.1.1 and 7.1.3, one gets upper bounds on $W(k,r)$ and $HJ(k,r)$ expressible by $A(\cdot)$ but which cannot be upperbounded by any $f_j(\cdot)$. A problem was to improve upon these enormous bounds and to get an upper bound on $W(k,r)$ and $HJ(k,r)$ in terms of some function $f_j(\cdot)$ for fixed j. It was resolved by Shelah.

Theorem 7.1.4 (Shelah, 1988) *Let $k,r \in \mathbb{N}$. The Hales–Jewett numbers satisfy the bound*

$$HJ(k,r) \le f_4(r+2k)\ ,$$

where $f_4(\cdot)$ is the iterated tower function. Hence $W(k,r) \le kf_4(r+2k)$.

The proof uses the next lemma, sometimes called Shelah's pigeonhole. For a word $a = a_1a_2\dots a_n$ of length n and an index $i \in [n]$ we denote by $a\backslash i$ the word of length $n-1$ obtained from a by deleting the i-th term a_i.

Lemma 7.1.5 *For every $n, r \in \mathbb{N}$ there exists an $m \in \mathbb{N}$ such that if*

$$\chi_l : [m]^{2n-1} \to [r],\ l = 1, 2, \dots, n\,,$$

are n colorings, then there exists a word $a \in [m]^{2n}$ such that for every $l = 1, 2, \dots, n$ we have

$$a_{2l-1} < a_{2l} \quad and \quad \chi_l(a\backslash(2l-1)) = \chi_l(a\backslash 2l)\,.$$

Moreover, if $f(n,r)$ is the minimum such m then $f(1,r) = r+1$ and

$$f(n+1,r) \le 1 + r^{f(n,r)^{2n}}\,.$$

Hence $f(n,r) \le f_3(r+4n)$ for every $r, n \in \mathbb{N}$.

Proof. The values $f(1,r) = r+1$ and $f(n,1) = 2$ are clear. We proceed by induction on n. Let $r \ge 2$, $M = f(n,r)$, $N = 1 + r^{M^{2n}}$ and $\chi_l : [N]^{2n+1} \to [r]$, $l \in [n+1]$, be given $n+1$ colorings. We define the coloring

$$\chi' : [N] \to [r]^{([M]^{2n})},\ \chi'(a) = (\chi_{n+1}(xa) \mid x \in [M]^{2n})\,.$$

By the pigeonhole principle and definition of N, there exist two integers $0 < a_{2n+1} < a_{2n+2}$ such that $\chi'(a_{2n+1}) = \chi'(a_{2n+2})$. We consider the colorings $\chi_l'' : [M]^{2n-1} \to [r]$, $l \in [n]$, given by

$$\chi_l''(x_1x_2\dots x_{2n-1}) = \chi_l(x_1x_2\dots x_{2n-1}a_{2n+1}a_{2n+2})\,.$$

By induction, there exists an $a \in [M]^{2n}$ such that $a_{2l-1} < a_{2l}$ and

$$\chi_l(aa_{2n+1}a_{2n+2}\backslash(2l-1)) = \chi_l(aa_{2n+1}a_{2n+2}\backslash 2l)$$

for every $l \in [n]$. But also, by the definition of χ' and the property of a_{2n+1} and a_{2n+2}, $a_{2n+1} < a_{2n+2}$ and

$$\chi_{n+1}(aa_{2n+1}a_{2n+2}\backslash(2n+1)) = \chi_{n+1}(aa_{2n+1}a_{2n+2}\backslash(2n+2))\,.$$

Thus we have proved the recurrent inequality for $f(n+1,r)$. The bound with $f_3(\cdot)$ follows by induction on n: $r+1 \le 2^r \le f_3(r) \le f_3(r+4)$, $f_3(r+4n)^{2n} \le (2n)^{f_3(r+4n)} \le 2^{nf_3(r+4n)} \le f_3(r+4n+2)$ and $1 + r^{f_3(r+4n+2)} \le 2^{rf_3(r+4n+2)} \le f_3(r+4n+4)$. □

Proof of Theorem 7.1.4. Clearly, $HJ(1,r) = 1$. We prove by induction on k that

$$HJ(k+1,r) \le HJ(k,r) \cdot f\left(HJ(k,r), r^{(k+1)^{HJ(k,r)}}\right)\,,$$

where $f(\cdot,\cdot)$ is the function of Lemma 7.1.5. Let $N = HJ(k,r)$ and $m = f(N, r^{(k+1)^N})$. The bound with $f_4(\cdot)$ then follows by induction on k: it holds for $k=1$ and, by Lemma 7.1.5,

$$\begin{aligned} HJ(k+1,r) &\leq Nf_3(r^{(k+1)^N}+4N) \leq f_3(f_3(N)) \leq f_3(f_3(f_4(r+2k))) \\ &= f_4(r+2(k+1)) . \end{aligned}$$

We prove the recurrent inequality for $HJ(k+1,r)$. Let $\chi: [k+1]^{Nm} \to [r]$ be a given coloring. Instead of all words we consider particular words $H_a(g) \in [k+1]^{Nm}$, and will seek a monochromatic line only among them. These are given by a *schema* $a \in [m]^{2N}$, satisfying $a_{2l-1} \leq a_{2l}$ for $l \in [N]$, and by a $g \in [k+1]^N$, as

$$H_a(g)_i = \begin{cases} k+1 & \dots \quad (l-1)m < i \leq (l-1)m + a_{2l-1}, \\ g_l & \dots \quad (l-1)m + a_{2l-1} < i \leq (l-1)m + a_{2l}, \\ k & \dots \quad (l-1)m + a_{2l} < i \leq lm, \ l = 1,2,\dots,N \end{cases}$$

— we split $[Nm]$ into intervals $[Nm] = I_1 I_2 \dots I_N$, $|I_l| = m$, split each interval I_l into the beginning with length a_{2l-1}, the middle with length $a_{2l} - a_{2l-1}$ (the middle may be empty) and the remaining end, and define the word $H_a(g)$ so that at the beginnings of all intervals it is constantly $k+1$, at the ends constantly k, and in the middle of I_l it is constantly g_l. For a fixed schema a such that never $a_{2l-1} = a_{2l}$, the mapping $g \mapsto H_a(g)$ sends a line in $[k+1]^N$ to a line in $[k+1]^{Nm}$. We define the colorings $\chi_l : [m]^{2N-1} \to [r]^{([k+1]^N)}$, $l = 1,2,\dots,N$, by

$$\chi_l(a) = (\chi(H_{a'}(g)) \mid g \in [k+1]^N)$$

where the schema $a' \in [m]^{2N}$ arises from $a \in [m]^{2N-1}$ by repeating the $(2l-1)$-term a_{2l-1} of a on the $2l$-th place and then appending the rest of a; so $a = a'\backslash(2l-1) = a'\backslash 2l$. (If a' is not a schema, we define the values $\chi_l(a)$ arbitrarily.) By the choice of m and Lemma 7.1.5, there is a word $b \in [m]^{2N}$ such that

$$b_{2l-1} < b_{2l} \ \text{ and } \ \chi_l(b\backslash(2l-1)) = \chi_l(b\backslash 2l), \ l = 1,2,\dots,N .$$

Consider the words $H_b(g)$ for g running in $[k+1]^N$. By the choice of N, there is a line $\ell \subset [k]^N$ such that the set $H_b(\ell) \subset [k+1]^{Nm}$, an incomplete line, is monochromatic to χ. But the word extending it to a complete line in $[k+1]^{Nm}$ has the same color. This follows from the crucial fact that, by the property of b and χ_l and the definition of $H_a(g)$, we have $\chi(H_b(g)) = \chi(H_b(g'))$ whenever $g, g' \in [k+1]^N$ differ just in a single l-th term by $g_l = k, g'_l = k+1$. To show it, we modify b to define a schema b', respective b'', by replacing b_{2l} by b_{2l-1}, respective b_{2l-1} by b_{2l}; then $H_b(g) = H_{b'}(g)$ and $H_b(g') = H_{b''}(g)$. We have, indeed,

$$\begin{aligned} \chi(H_b(g)) &= \chi(H_{b'}(g)) = \chi_l(b\backslash 2l)(g) \\ &= \chi_l(b\backslash(2l-1))(g) = \chi(H_{b''}(g)) = \chi(H_b(g')) . \end{aligned}$$

Thus we get a χ-monochromatic line in $[k+1]^{Nm}$. □

A more recent breakthrough, discussed in the final Notes, brought $W(k,r)$ in the scope of the quadruple exponential function $f_2(f_2(f_2(f_2(\cdot))))$. This is still very far from the best known lower bounds on $W(k,r)$, which are only single exponential, $f_2(\cdot)$. Perhaps mankind will never know the function $W(k,r)$ exactly, except for a few initial values.

We turn to the lower bounds on $W(k,r)$, and obtain two which we derive more generally for set systems. The first shows the power of simple counting arguments. The second and better bound is obtained by an interesting counting tool of the probabilistic method, Lovász' local lemma (Theorem B.1.3 in Appendix B.1).

Proposition 7.1.6 *Let $r \in \mathbb{N}$, $X_1, X_2, \ldots, X_t$ be nonempty finite sets, $X = \bigcup_{i=1}^{t} X_i$, $m = \min_{1\le i\le t} |X_i|$ and $d = \max_{1\le i\le t} \deg(i)$ where $\deg(i) = |\{j \in [t] \mid X_i \cap X_j \neq \emptyset\}|$. Each of the following two conditions is sufficient for existence of a coloring $\chi : X \to [r]$ in which no set X_i is monochromatic.*

1. $\sum_{i=1}^{t} r^{1-|X_i|} < 1$.
2. $e(d+1)r^{1-m} < 1$ *where* $e = 2.71828\ldots$ *is Euler's number.*

Proof. 1. Let $n = |X|$. The number of r-colorings of X with monochromatic X_i is $r^{1+n-|X_i|}$, and the number of all colorings is r^n. Since the size of a union of sets is at most the sum of sizes of the sets, $\sum_{i=1}^{t} r^{1+n-|X_i|} < r^n$ implies that there is a coloring with no X_i monochromatic. Division by r^n gives the stated condition.

2. Consider the uniform probability space on the set of all r^n colorings $\chi : X \to [r]$, and the events E_i, $1 \le i \le t$, of colorings that make X_i monochromatic. By Proposition B.1.1, each E_i is mutually independent of the events $\{E_j \mid X_i \cap X_j = \emptyset\}$. Also, $\Pr(E_i) = r^{1-|X_i|} \le r^{1-m}$. Hence, with d as defined, the condition ensures that the hypothesis of Lovász' local lemma (Theorem B.1.3) is satisfied, and by the lemma the complement of the t events E_i has positive probability. There is a coloring making no set X_i monochromatic. □

What does this give for arithmetic progressions and $W(k,r)$?

Corollary 7.1.7 *Let $W(k,r)$ be as defined above, 1+ the maximum $n \in \mathbb{N}$ for which there is an r-coloring of $[n]$ with no monochromatic k-term AP.*
1. Part 1 of the previous proposition gives that

$$W(k,r) > \sqrt{2(k-1)r^{k-1}} \gg \sqrt{kr^{k-1}}\,.$$

2. Part 2 gives that

$$W(k,r) > \frac{r^{k-1}}{ek} \gg r^{k-1}/k\,.$$

Proof. Let $k \ge 2$, $X = [n]$ and the sets X_i be the k-term APs in X; $|X_i| = k$ for each i.

1. There are $n-(1+(k-1)d)+1$ APs in $[n]$ with difference d. Thus the total number t of k-term APs in $[n]$ equals $\sum_{d=1}^{D}(n-(k-1)d)$ where $D=\lfloor\frac{n-1}{k-1}\rfloor$. Since $D\le\frac{n-1}{k-1}$ and $D+1>\frac{n-1}{k-1}$,

$$t=Dn-\frac{(k-1)D(D+1)}{2}=D(n-(D+1)(k-1)/2)<\frac{n^2}{2(k-1)}\,.$$

By part 1 of Proposition 7.1.6, an r-coloring of $[n]$ with no monochromatic k-term AP exists if $tr^{1-k}<1$, that is if $t<r^{k-1}$. We see that this inequality holds if $n\le\sqrt{2(k-1)r^{k-1}}$.

2. Now $m=k$. We claim that each k-term AP in $[n]$ intersects less than kn other k-term APs. Thus $d+1\le kn$ and the condition in Part 2 becomes $ekn<r^{k-1}$, which holds if $n<r^{k-1}/ek$.

To prove the claim, we first show that for $k,n\in\mathbb{N}$, $k\ge 2$, and fixed $x\in[n]$, always $c(k)\le n-1$ where $c(k)$ is the number of k-term APs containing x. By $e(k)$, resp. $b(k)$, we denote the number of k-terms APs in $[n]$ that end, resp. begin, in x. So $c(2)=e(2)+b(2)=(x-1)+(n-x)=n-1$, $c(3)=e(3)+b(3)+\min(e(2),b(2))\le\frac{1}{2}((x-1)+(n-x))+\frac{n-1}{2}=n-1$, $c(4)\le e(3)+e(4)+b(3)+b(4)\le(\frac{1}{2}+\frac{1}{3})(n-1)=\frac{5}{6}(n-1)$ and $c(5)\le e(4)+e(5)+b(4)+b(5)+\min(e(3),b(3))\le(\frac{1}{3}+\frac{1}{4})(n-1)+\frac{n-1}{4}=\frac{5}{6}(n-1)$. One can check that if $k\ge 6$ then $s(k):=\sum_{i=\lceil(k-1)/2\rceil}^{k-1}\frac{1}{i}\le 1$. Thus for $k\ge 6$ we also have $c(k)\le\sum_{i=\lceil(k-1)/2\rceil}^{k-1}(e(i)+b(i))\le s(k)((x-1)+(n-x))\le n-1$. Hence a fixed k-term AP in $[n]$ intersects at most $k(n-1-1)=kn-2k<kn$ other k-term APs. □

We use a particular case of the next theorem in Section 7.3 in the proof of the density Hales–Jewett theorem.

Theorem 7.1.8 (Graham and Rothschild, 1971) *Suppose that $k,m,r\in\mathbb{N}$ and $p\in\mathbb{N}_0$. Then for every $n\ge N$ in every r-coloring χ of p-subspaces of $[k]^n$ there is a monochromatic m-subspace $V\subset[k]^n$ — for every two p-subspaces $A,B\subset[k]^n$ with $A,B\subset V$ we have $\chi(A)=\chi(B)$.*

Proof.

□

7.2 Sperner's theorem and Roth's theorem

We prove the density Hales–Jewett theorem for two-element alphabet.

Proposition 7.2.1 (the density HJT for $k=2$) *Let $\delta\in(0,1]$. Then for every $n\ge N$ every set $A\subset[2]^n$ with $|A|>\delta 2^n$ contains a line.*