Informal notes on van der Waerden's Theorem and Ramsey's Theorem

Tibor Szabó
Winter 2011/12 — Discrete Mathematics II

On van der Waerden's Theorem. Szemerédi's Theorem is hard to prove, we rather consider the following weakening about monochromatic arithmetic progressions in two-colored integers.

Is it unavoidable to have a monochromatic (*m.c.*) arithmetic progression of length 3 (a 3-*AP*) if we two-color the integers?
YES, of course.
Roth's Theorem says that the larger of the two color-classes, the one whose density is at least 50% , will contain a 3-AP. However, Roth's Theorem is too big of a gun to shoot down such a simple statement.

Let $c : [N] \to \{\texttt{red}, \texttt{blue}\}$ be an arbitrary two-coloring of the first N integers with no m.c. 3-AP. (We do not specify N now, only at the very end. We work under the assumption that it is "large enough".) In any *block* of five consecutive integers $y, y+1, y+2, y+3, y+4$ we find a triple of integers forming a 3-AP, such that the color of the first two is the same, while the third one (of course) has the opposite color. (If $c(y) = c(y+1)$ then $y, y+1, y+2$ is such a triple, if $c(y) = c(y+2)$ then $y, y+2, y+4$ is such a triple, otherwise $y+1, y+2, y+3$ is such a triple.) Let us consider the first $5 \cdot (2^5 + 1)$ integers as the union of $2^5 + 1 = 33$ disjoint blocks of fives. Each of these blocks can have one of the 2^5 coloring patterns on it. By the Pigeonhole Principle (PP), two of these 33 blocks have identical coloring pattern. In these two blocks we have two 3-APs: a_1, a_1+d, a_1+2d in the first block and a_2, a_2+d, a_2+2d in the second one, so that $c(a_1) = c(a_1+d) = c(a_2) = c(a_2+d)$ (by symmetry we can assume that this color is `red`) , and $c(a_1+2d) = c(a_2+2d)$ is of the opposite color, that is, `blue`. But then what is the color of $z = a_1 + 2d + 2d'$? (Here we denote by $d' = a_2 - a_1$.) If $c(z)$ is `blue` then a_1+2d, a_2+2d, z is a `blue` 3-AP (with difference d'). If $c(z)$ is `red` then a_1, a_2+d, z is a `red` 3-AP (with difference $d'+d$). So we proved that there is a m.c. 3-AP if $N = 5 \cdot (2^5 + 1 + 2^5) = 325$.

Proposition. For any two-coloring of [325] there is a m.c. 3-AP.
Remark. BTW The tight answer is: $N = 9$ integers are enough.

Can we find a m.c. 4-AP if N is even larger enough? YES, of course, Szemerédi's Theorem for arithmetic progressions of length 4 implies that the larger of the two color classes contains a 4-AP. The proof of this would be several level harder than the one of Roth's Theorem. Can we apply somehow the previous idea to prove just the coloring statement? How large should N be for such proof

to work?
Let $c : [N] \to \{\texttt{red}\ , \texttt{blue}\}$ be a coloring with no m.c. 4-AP. How large blocks should we consider to use the previous idea? Well, we know that within 325 consecutive integers there are three forming a m.c. 3-AP. The extension of this 3-AP to a 4-AP is within the next 162 integers. So any block of 487 consecutive integers contains a 4-AP $a_1, a_1+d, a_1+2d, a_1+3d$, such that $c(a_1) = c(a_1+d) = c(a_1+2d)$ and (of course) $c(a_1 + 3d)$ is of the opposite color. IF (and it's a big IF) we were able to find *a* 3*-AP of blocks* having the same coloring pattern, we would be DONE. (Indeed: then we would have three 4-APs $a_1, a_1 + d, a_1 + 2d, a_1 + 3d$, $a_2, a_2+d, a_2+2d, a_2+3d$, $a_3, a_3+d, a_3+2d, a_3+3d$, such that the starting integers a_1, a_2, a_3 form a 3-AP (say with difference d'), furthermore $c(a_1) = c(a_1 + d) = c(a_1 + 2d) = c(a_2) = c(a_2 + d) = c(a_2 + 2d) = c(a_3) = c(a_3 + d) = c(a_3 + 2d)$, say `red`, and $c(a_1 + 3d) = c(a_2 + 3d) = c(a_3 + 3d)$ is the opposite color `blue`.

Then the integer $z = a_1 + 3d + 3d'$ will be the fourth member of a m.c. AP (which either starts at a_1 and has difference $d+d'$ (if its color is `red`) or at a_1+3d and its difference is d' (if its color is `blue`)).)
So how do we find this 3-AP of blocks having the same coloring pattern?? It was so easy in the previous proof, when we just needed to find two (one can say, a 2-AP of) blocks with the same coloring pattern: we just used the Pigeonhole Principle (PP). It turns out that we must do the same here except the PP-use is in a bit more complex setting. Consider each block (of length 487) as one entity and each of its 2^{487} possible coloring patterns as one possible color of this "entity" and try to find a m.c. 3-AP in this setup. Hence, it seems that to find a m.c. 4-AP we need to extend first the above Proposition to arbitrary number of *colors*.

Theorem. For any r there is a number $W = W(r, 3)$, such that no matter how we color the first W integers with r colors, there will be a m.c. 3-AP.

Remark. By the above, we can then use this Theorem to find a m.c. 4-AP in any two-coloring of the first $487 \cdot W(2^{487}, 3)$ integers. (This is an admittedly weak bound, it would be enough to two-color 35 integers. But while this proof generalizes to arbitrary number of colors and length of APs, the 35 bound is ad hoc.)

Proof of Thm. Induction on r. For the base case we can take $r = 2$ which is just our Proposition, which shows that $W(2, 3) \leq 325$. We prove first the statement for $r = 3$ to see better the pattern. Let us have a 3-coloring $c : [N] \to \{\texttt{red}, \texttt{blue}, \texttt{yellow}\}$ with no m.c. 3-AP. In any block of 4 consecutive integers we find two identically colored, so in any block of 7 integers we find a 3-AP $a_1, a_1 + d, a_1 + 2d$, such that $c(a_1) = c(a_1 + d)$ and (of course) $c(a_1 + 2d)$ is different from the color of the other two. Taking 3^7+1 consecutive disjoint blocks of 7 integers, we find two that have identical coloring pattern. Hence there are two arithmetic progressions $a_1, a_1 + d, a_1 + 2d$ and $a_2, a_2 + d, a_2 + 2d$, such that

$c(a_1) = c(a_1 + d) = c(a_2) = c(a_2 + d)$, say is `red`, and $c(a_1 + 2d) = c(a_1 + 2d)$ is NOT `red`, say is `blue`. Then the integer $z = a_1 + 2d + 2d'$ (where $a_2 = a_1 + d'$) does not have color `red` (because of the 3-AP $a_1, a_2 + d, z$) and it does not have color `blue` (because of the 3-AP $a_1 + 2d, a_2 + 2d, z$). So z is colored `yellow`.

Hence in any block of $7 \cdot (2 \cdot 3^7 + 1)$ integers we find a_1, d, d' such that $c(a_1) = c(a_1 + d + d')$ is one color, $c(a_1 + 2d) = c(a_1 + 2d + d')$ is another color, and (of course) $c(a_1 + 2d + 2d')$ is the third color.

Let's find two blocks of $7 \cdot (2 \cdot 3^7 + 1)$ integers with identical color pattern. These surely exists if we take $3^{7\cdot(2\cdot3^7+1)} + 1$ blocks. Let the distance of these two identically colored blocks be d''.

In the first block we find a_1, d, d' such that $c(a_1) = c(a_1 + d + d')$, say of color `red`, $c(a_1 + 2d) = c(a_1 + 2d + d')$ is of another color, say `blue`, and $c(a_1 + 2d + 2d')$ is of the third color (in our setup it is assumed to be `yellow`). Since the second block has identical color pattern we also have that $c(a_1 + d'') = c(a_1 + d + d' + d'')$ is `red`, $c(a_1 + 2d + d'') = c(a_1 + 2d + d' + d'')$ is `blue`, and $c(a_1 + 2d + 2d' + d'')$ is `yellow`.

Now depending on the color of the integer $y = a_1 + 2d + 2d' + 2d''$ we have a m.c. 3-AP (the possibilities: in color `red` $a_1, a_1 + d + d' + d'', y$, in color `blue` $a_1 + 2d, a_1 + 2d + d' + d'', y$, and in color `yellow` $a_1 + 2d + 2d', a_1 + 2d + 2d' + d'', y$.) And we are done for $r = 3$ colors. I am sure I made a mistake somewhere with the numbers, but if not then clearly $W(3, 3) \leq (2 \cdot 3^{7\cdot(2\cdot3^7+1)} + 1) \cdot (7 \cdot (2 \cdot 3^7 + 1))$.

Now we just need to iterate this idea and we get a bound on $W(r, 3)$, which then we can plug into the formula $487 \cdot W(2^{487}, 3)$ to get an upper bound to guarantee a m.c. 4-AP in two-colored sequences.

Hmmm..... The bound is a bit wild.

The general theorem can be formulated as follows. Let

$$W(r, k) = \min\{N \in \mathbb{N} : \text{ for any } r\text{-coloring } c : [N] \to [r] \text{ there is a m.c. } k\text{-AP}\}.$$

Van der Waerden's Theorem For any integers $r, k \geq 1$ $W(r, k) < \infty$.

Proof: Analogous to the above idea. A *family of crossing k-APs* is a family of k-APs with starting elements $a^{(1)}, \ldots, a^{(\ell)}$, and differences $d_1, \ldots, d_\ell$, respectively such that the $(k + 1)st$ element of each of these APs is the same integer: $a^{(1)} + kd_1 = \ldots = a^{(\ell)} + kd_\ell$

Let $L(r, k, \ell)$ be the smallest positive integer N such that for any r-coloring $c : [N] \to [r]$ there is a m.c. $(k + 1)$-AP or a family of ℓ crossing m.c. k-APs in ℓ distinct colors.

We show by induction on k that $L(r, k, \ell) < \infty$ for every $k \geq 1$ and $r \geq \ell \geq 1$. Then we are done, since $W(r, k + 1) \leq 2L(r, k, r)$.

Base case: $L(r, 1, \ell) = \ell$ for all $\ell \leq r$.

Let $k \geq 2$: Induction on ℓ.

For $\ell = 1$, $L(r, k, 1) \leq W(r, k) < 2L(r, k - 1, r) < \infty$.

For $\ell \geq 2$, $L(r, k, \ell) < 2W(r^{2L(r,k,\ell-1)}, k)2L(r, k, \ell - 1) < \infty$ □

Remark The auxiliary function L used in this proof is a bit different from the auxiliary function in the proof on the transparencies. It is still the same proof.

The bound following from this proof is enoooormous. For a long time there was no primitive recursive upper bound known, until Shelah gave a proof for that. The best known bound today is due to Gowers (who got the Fields-medal partly for his work on this problem (or rather on the stronger Szemerédi's Theorem) and stands at a five times iterated exponential:

$$W(r,k) \leq 2^{2^{r^{2^{2^{k+9}}}}}.$$

On Ramsey's Theorem. Recall from Discrete Math I: Ramsey's Theorem in party of 6, Definition of *Ramsey number.* Given a coloring $c : E(K_N) \to \{\texttt{red}, \texttt{blue}\}$ we say that a subgraph $K \cong K_k$ is a ”$\texttt{red}$ K_k” if c is constant $\texttt{red}$ on $E(K)$. The definition of ”$\texttt{blue}$ K_l” is analogous.

$$R(k,l) = \min\{N : \ \forall c : E(K_N) \to \{\texttt{blue,red}\} \exists \ \texttt{red}K_k \text{ or } \texttt{blue}K_l \ \}$$

Examples: $R(k,2) = R(2,k) = k$, HW: $R(4,3) = 9$, $R(5,3) = 14$, $R(4,4) = 18$

$R(4,5)$ is known after a huge computer search. $R(5,5)$ is not known (today's computers are not fast enough to handle it). For $R(6,6)$ probably no computer will ever be fast enough.

Definition of *Paley-graph* P_p, for primes $p \equiv 1 \pmod 4$ (this congruence assumption is needed to have that -1 is a quadratic residue modulo p, and that's only to make the definition of an edge symmetric):

Vertex set $V(P_p) = \mathbb{F}_p$ (field of p elements).

Edge set $E(P_p) = \{xy : x - y \in Q_p\}$, where $Q_p = \{z^2 : z \in \mathbb{F}_p\}$ is the set of *quadratic residues* modulo p.

In a Paley-graph every vertex has $(p-1)/2$ neighbors, since $|Q_p| = (p-1)/2$.

P_5 is the 5-cycle. It does not contain a K_3 and no $\bar{K}_3$. This example, together with the ”party of 6”-proposition proves that $R(3,3) = 6$.

HW: P_{17} does not contain a K_4 and no $\bar{K}_4$.

An alternative definition of Ramsey numbers is

$$R(k,l) = \min\{N : \ \forall G \text{ with } |V(G)| = n, \, \omega(G) \geq k \text{ or } \alpha(G) \geq l\}$$

For a *Paley-coloring* just color a pair xy with $\texttt{red}$ if $x - y \in Q_p$, otherwise $\texttt{blue}$.

Theorem. $R(k,l) \leq R(k,l-1) + R(k-1,l)$

Proof: Take $N = R(k, l-1) + R(k-1, l)$ and an arbitrary `red`/`blue` coloring of $E(K_N)$. Pick an arbitrary vertex $x \in V$.
Case 1: x has at least $R(k-1, l)$ `red` neighbors
Case 2: x has at least $R(k, l-1)$ `blue` neighbors
First of all: one of these cases does happen. (Otherwise there are at most $R(k, l-1) - 1 + R(k-1, l) - 1 = N - 2$ neighbors of x, a contradiction.)
In Case 1: if there is a `red` K_{k-1} among the red neighbors of x, then together with x they form a `red` K_k, done. Otherwise there is a `blue` K_l among the red neighbors of x and we are also done.
Case 2 is analogous: if there is a `blue` K_{l-1} among the `blue` neighbors of x, then together with x they form a `blue` K_l, done. Otherwise there is a `red` K_k among the `blue` neighbors of x and we are also done. □

Corollary For all $k, l \geq 1$, $R(k, l) \leq \binom{k+l-2}{k-1}$. In particular, $R(k, l)$ exists.
Proof. Induction on $k + l$, using the similar identity for binomial coefficients.

Corollary $R(k, k) \leq 4^k$.

The finiteness of $R(k, l)$ is called *Ramsey's Theorem.* The proof above with the estimate is due to Erdős and Szekeres.

How about a lower bound?

A set of vertices of a graph G is called a *homogenous set* of G if it is a clique or an independent set. A graph with no homogenous set of order k is called a *k-Ramsey graph.*

How good Ramsey-graphs are the Paley-graph? It is not known. Numerical data suggests that the largest clique (and hence also independent set) might be much smaller than the square root of the number of vertices. For example, for $p = 6997$ the clique number is only 17. (Shearer) However, provably it is only known that the largest clique and independent set is *at most* of the order $\sqrt{p}$. This gives $R(k, k) = \Omega(k^2)$.

But for this we have the more trivial Turán-coloring: Partition $(k-1)^2$ vertices into parts of size $k-1$ and color each edge within parts by `red` and edges between parts with `blue`. The largest m.c. clique has size $k-1$, proving $R(k, k) \geq (k-1)^2 + 1$. Pretty weak considering that the upper bound is exponential.

Is there something better?

Theorem (Erdős) $R(k, k) \geq \sqrt{2}^k$
Proof. By the probabilistic method (see in about three weeks). Only proves existence; no "construction". □

Nobody is able to *construct* explicitly k-Ramsey-graphs with 1.000001^k vertices. One needs a quite unexpected idea even to construct something better than k^2 vertices. (We will come back to this question later in the semester when we discuss the Linear Algebra method.)

\$1000 dollar question: Determine $\lim_{k\to\infty} \sqrt[k]{R(k, k)}$. (Currently it is not even known that this limit exists. The existence of the limit alone would bring you

\$500.)

Application: (HW)

Theorem. Color the integers with r colors. Prove that there are three numbers of the same color, such that one is equal to the sum of the other two.

Remark. Notice the similarity of this homework exercise to van der Waerden's Theorem. This exercise proves the existence of a monochromatic solution to the equation $x + y = z$. Van der Waerden's Theorem proves the existence of a monochromatic solution to the equation $x + y = 2z$.

Another application.

Proposition (Eszter Klein) Among 5 points in the plane in general position (i.e. no three on a line) there are always at least 4 in convex position.

Happy Ending Problem (Klein) Let $M(n)$ be the smallest number such that among any set of $M(n)$ points in the plane there are at least n in convex position. Is $M(n)$ finite?

$M(3) = 3$, $M(4) = 5$.

(a) there are two things that can happen to four points in general position: they are either in convex position or not.

(b) n points are in convex position iff every four element subset is in convex position. (Proof of "if" statement: take convex hull, if there is point inside, it is also contained in some triangle of an arbitrary triangulation of the convex hull: these are four points in non-convex position.)

(c) among any 5 points there are four which are not in non-convex position

Ramsey framework:

(a) provides a natural two-coloring of the 4-subsets of the point set (`red`: "convex 4-gon", `blue`: "non-convex 4-set")

(b) says that we really want is a m.c. subset in color `red` of LARGE size (i.e., of size n)

(c) says we CANNOT have a m.c. subset of size 5 in color `blue`.

We need Ramsey's Theorem in a situation when we color 4-sets instead of edges.

Definitions Graph: $G = (V, E)$ on vertex set V with edge set $E \subseteq \binom{V}{2}$

Hypergraph: on vertex set V with edge set $E \subseteq 2^V$

s-uniform hypergraph: if edge set $E \subseteq \binom{V}{s}$

Complete s-uniform hypergraph $K_k^{(s)}$ on k vertices is defined by $E(K_k^{(s)}) = \binom{[k]}{s}$.

Example. graph: 2-uniform hypergraph

Definition *s-uniform Ramsey number* $R^{(s)}(k, l)$ is the smallest integer N such that for any 2-coloring $c : \binom{[N]}{s} \to \{\texttt{red}, \texttt{blue}\}$ there exists a subset $I_r \subseteq [N]$ such that $c(J) = \texttt{red}$ for every $J \in \binom{I_r}{s}$ or there exists a subset $I_b \subseteq [N]$ such that $c(J) = \texttt{blue}$ for every $J \in \binom{I_b}{s}$

Theorem $R^{(s)}(k, l)$ is finite for every $s, k, l \geq 1$

Proof. Induction on s: $R^{(1)}(k,l) = k + l - 1$.
Let $s \geq 2$. Induction on $k + l$.
Base cases: For $k \geq l$, $l < s$, we have $R^{(s)}(k,l) = l$,
for $k \leq l$, $k < s$, we have $R^{(s)}(k,l) = k$,
for $k \geq s$, we have $R^{(s)}(k,s) = k$,
and for $l \geq s$, we have $R^{(s)}(s,l) = l$.

Let $c : \binom{[N]}{s} \to \{\texttt{red}, \texttt{blue}\}$ be a two-coloring of the s-sets. Pick an arbitrary vertex, say $N \in [N]$. Canonical projection of c on the $(s-1)$-sets of $[N-1]$: $c^* : \binom{[N-1]}{s-1} \to \{\texttt{red}, \texttt{blue}\}$ defined by $c^*(A) := c(A \cup \{N\})$.
By induction (used for $s - 1$), there is a "large" subset $J_r \subset [N-1]$ such that every $(s-1)$-subset of J_r is `red` or there is a "large" subset $J_b \subset [N-1]$ such that every $(s-1)$-subset of J_b is `blue`.
How large should "large" be?

In $|J_r|$ it would be enough to have $R^{(s)}(k-1,l)$ vertices. This would guarantee that either there is a m.c. l-subset in `blue` or an m.c $(k-1)$-subset in `red` , which together with x would form an m.c. k-subset in `red`. (Remember that we are within J_r!)
The argument for $|J_b|$ is analogous.

So if $N - 1 = R^{(s-1)}(R^{(s)}(k-1,l), R^{(s)}(k,l-1))$, then we are guaranteed to have the appropriate size J_r or the appropriate size J_b. Hence $R^{(s-1)}(R^{(s)}(k-1,l), R^{(s)}(k,l-1)) + 1$ is an upper bound on $R^{(s)}(k,l)$ and its finiteness is implied by the finiteness of the functions involved in this upper bound; these are all finite by induction. □

Corollary $M(n) \leq R^{(4)}(n,5) < \infty$

Remark The best known bounds for $M(n)$ are pretty far from each other:

$$2^{n-2} + 1 \leq M(n) \lessapprox \frac{4^n}{\sqrt{n}}.$$

The lower bound is conjectured to be tight by Erdős and Szekeres. It is proven to be tight for $n = 3, 4, 5, 6$.