Partition calculus

1 Introduction

Partition calculus is the area of set theory that deals with Ramsey theory; it is devoted to Ramsey’s theorem and its infinite and infinitary generalizations. This means both strengthenings of Ramsey’s theorem for sets of natural numbers (like the Carlson-Simpson or the Galvin-Prikry theorems characterizing the completely Ramsey sets in terms of the Baire property) and for larger cardinalities (like the Erdős-Rado theorem), as well as variations in which the homogeneous sets are required to possess additional structure (like the Baumgartner-Hajnal theorem).

Ramsey theory is a vast area and by necessity we won’t be able to cover (even summarily) all of it. There are many excellent references, depending on your particular interests. Here are but a few:

- Stevo Todorčević, High-dimensional Ramsey theory and Banach space geometry, in Ramsey methods in Analysis, Spiros Argyros, Stevo Todorčević, Birkhäuser (2005), 121–257.

I taught a course on Ramsey theory at Caltech a couple of years ago, and expect to post notes from it at some point. Here we will concentrate on infinitary combinatorics, but I will briefly mention a few finitary results.

In the early 1950s, Rado introduced the partition calculus notation, several variants of which we will be using through these lectures.

Definition 1 Let $X$ be a set and $\tau$ be a cardinal. Then $[X]^\tau = \{Y \subseteq X: |Y| = \tau\}$. Let $\rho$ be a cardinal. A function $f : [X]^\tau \to \rho$ is referred to as a $\rho$-partition or $\rho$-coloring of $[X]^\tau$. Sometimes one writes the partition “explicitly” as $[X]^\tau = \bigcup_{\beta<\rho} A_\beta$, where $A_\beta = f^{-1}(\{\beta\})$. 

1
Let $\kappa, \tau, \rho$ be cardinals, and let $(\mu_i : i \in \rho)$ be a sequence of cardinals. The **arrow notation** is defined as follows:

$$\kappa \rightarrow (\mu_i)_{i<\rho}^\tau$$

iff for any set $X$ of size $\kappa$ and any coloring $f : [X]^\tau \rightarrow \rho$ there is an $i < \rho$ and a set $H \in [X]^{\mu_i}$ such that $f''[H]^\tau = \{i\}$. Here, $g''Y$ denotes the pointwise image of the set $Y$ under the function $g$. We say that $H$ is **monochromatic**, or **homogeneous** or $i$-**homogeneous** for the coloring $f$, and also say that $f$ **admits** an $i$-homogeneous subset of $\kappa$ of the prescribed size $\mu_i$.

One reads $\rightarrow$ as **arrows**. The negation of the relation above is denoted by replacing $\rightarrow$ with $\not\rightarrow$.

Although we are mostly interested in the version above, we will also consider the following:

**Definition 2** An **order type** is an equivalence class of the collection of partial orders under the relation of order isomorphism; we denote the order type of a poset $(X, \prec)$ by $\text{ot}(X, \prec)$.

When talking about order types of partially ordered sets, by $\eta = \eta_0$ we mean the order type of $\mathbb{Q}$ and by $\lambda = \lambda_0$ we mean the order type of $\mathbb{R}$. The order type of an ordinal $\alpha$ is denoted also by $\alpha$, so the notation is consistent with the standard use for well-ordered sets.

If $X$ is a partially ordered set and $\Phi$ is an order type, then $[X]^\Phi$ denotes

$$\{Y \subseteq X : \text{ot}(Y) = \Phi\}.$$  

There is risk of ambiguity here, so when it matters we use $[Y]^{|\alpha|}$ for the set of subsets of $Y$ of size $\aleph_\alpha$ while $[Y]^{<\alpha}$ is used for the collection of subsets of (the partially ordered set) $Y$ of **order type** $\omega_\alpha$.

If $\Phi, \Psi$ are order types, $\Phi \leq \Psi$ means that whenever $X \in \Phi$ and $Y \in \Psi$, there is an order preserving injection (an **embedding**) of $X$ into $Y$.

If $\kappa, \tau$ denote order types, $\rho$ is a cardinal, and $(\mu_i : i \in \rho)$ is a sequence of order types, then

$$\kappa \rightarrow (\mu_i)_{i<\rho}^\tau$$

iff for any poset $(X, \prec)$ of order type $\kappa$ and any coloring $f : [X]^\tau \rightarrow \rho$ there is an $i < \rho$ and a set $H \in [X]^{\mu_i}$ such that $f''[H]^\tau = \{i\}$. As before, such an $H$ is called homogeneous for $f$.

The notation is very flexible; if all the $\mu_i$ are equal to $\mu$, one simply writes $\kappa \rightarrow (\mu)^\tau_\rho$ and if $\rho = 2$ then one simply says $\kappa \rightarrow (\mu)^\tau$. For example, $\omega \rightarrow (\omega)^2$, $\omega \rightarrow (\omega)_2^2$ and...
\[ \omega \rightarrow (\omega, \omega)^2 \] have the same meaning, and are true whether \( \omega \) denotes size or order type.

However daunting it appears at first sight, this notation is a very useful shorthand, and it compresses a large amount of information. We assume choice throughout unless otherwise stated; the results vary wildly in the absence of choice.

Note that if \( \kappa \rightarrow (\mu_i)_{i<\rho}^\tau \) then \( \kappa' \rightarrow (\mu'_i)_{i<\rho'}^\tau \) for any \( \kappa' \geq \kappa, \rho' \leq \rho \) and \( (\mu'_i : i < \rho') \) such that \( \mu'_i \leq \mu_i \) for all \( i < \rho' \).

## 2 The classical Ramsey’s theorem

Ramsey theory is to a large extent the study of the “arrow” relation, in a sense, a series of (vast) generalizations of the pigeonhole (or Schubfach-, or Dirichlet’s) principle.

The result that gave birth to the field is Frank Ramsey’s theorem from 1928, published in 1930, and obtained in order to show a decidability result.

**Theorem 3 (Ramsey)**  
For any \( n, m < \omega \), one has \( \omega \rightarrow (\omega)^n_m \).

**Proof:** The result is clear if \( n \leq 1 \), by the pigeonhole principle.

We proceed by induction on \( n \). Assume the result for \( n \) and let’s show it for \( n + 1 \). Let \( f : [\omega]^{n+1} \rightarrow m \) be given. For \( a \in \omega \) let \( f_a : [\omega \setminus (a+1)]^n \rightarrow m \) be the function \( f_a(x) = f(\{a\} \cup x) \). By induction define a sequence \( (a_i, k_i, A_i : i < \omega) \), as follows:

- \( a_0 = 0 \). By induction, there is a number \( k_0 \in m \) and an infinite set \( A_0 \) that is \( k_0 \)-homogeneous for \( f_{a_0} \).
- Inductively, let \( a_{i+1} = \min(A_i) \). There is then a \( k_{i+1} \in m \) and an infinite \( A_{i+1} \subseteq A_i \setminus (a_{i+1} + 1) \) that is \( k_{i+1} \)-homogeneous for \( f_{a_{i+1}} \).

Now consider the function \( g : \omega \rightarrow m \) given by \( g(i) = k_i \). Let \( A \) be infinite and homogeneous for \( g \). Then the set \( \{a_i : i \in A\} \) is infinite and homogeneous for \( f \). \( \square \)

This proof has the advantage that it is easily finitizable. By this I mean, from the proof one can easily extract bounds that allow us to prove the following:

**Theorem 4**  
For all \( n, m, k < \omega \) there is an \( N < \omega \) such that \( N \rightarrow (n)_m^k \).

**Proof:** Argue by induction on \( k \). We denote by \( r(k, m, n) \) the least \( N \leq \omega \) as in the statement, and our task is to show that in fact \( r(k, m, n) < \omega \). (More generally, we
let $r(k; n_0, \ldots, n_m)$ be the smallest number $N \leq \omega$ such that $N \to (n_0, \ldots, n_m)^k$, so $r(k, m, n) = r(k; n, \ldots, n)$.)

To begin with, $r(1, m, n) = m(n - 1) + 1$ by the pigeonhole principle.

Given $k > 1$, let $M = r(1, m, n)$, and define $r_i$ for $1 \leq i \leq M$ by $r_1 = 1$, $r_{j+1} = r(k - 1, m, r_j)$ for $1 \leq j < M$. The proof of Theorem 3 above easily shows that $r(k, m, n) \leq r_M$. □

One can see that, in fact, for all $m < \omega$ there is a constant $c_m$ such that for all $k, n < \omega$,

$$r(k, m, n) \leq 2^{2^{c_m n}}$$

where the tower of 2s has height $k - 1$.

The case $k = 2$ of Ramsey’s theorem has received special attention, the statement $6 \to (3)^2$ being particularly well-known. Write $r(a, b)$ for $r(2; a, b)$ for $a, b < \omega$. This is the smallest number $n$ of vertices required to guarantee that whenever each edge of the complete graph $K_n$ is colored either blue or red, there is either going to be a blue copy of $K_a$ or a red copy of $K_b$.

Theorem 5 (Erdős-Szekeres) For all $k, l < \omega$, $r(k + 1, l + 1) \leq \binom{k + l}{k}$.

PROOF: The proof is by induction on $k + l$. Note that $r(1, n) = r(n, 1) = 1$. Now argue that $r(k + 1, l + 1) \leq r(k, l + 1) + r(k + 1, l)$, from which the statement follows immediately.

To see the claimed inequality, consider a graph on $r(k, l + 1) + r(k + 1, l)$ many vertices, a coloring of its edges in colors blue and red, and let $a$ be one of the vertices. Let $A$ be the set of vertices that are joined to $a$ by a blue edge and let $B$ be the set of vertices joined to $a$ by a red edge, so $|A| + |B| = r(k, l + 1) + r(k + 1, l) - 1$. It follows that either $|A| \geq r(k, l + 1)$ or $|B| \geq r(k + 1, l)$.

Without loss, assume that $|A| \geq r(k, l + 1)$. It follows that either there is a red copy of $K_{l+1}$ with vertices from $A$, and we are done, or else, there is a blue copy of $K_k$ with vertices from $A$. In this case, all these vertices are joined to $a$ by blue edges, so, by adding $a$ to them, we obtain a blue copy of $K_{k+1}$. □

In fact, $r(k + 1, l + 1) = o\left(\binom{k + l}{k}\right)$, this was shown by Rödl. Very recently (in a paper to appear in the Annals of Mathematics), David Conlon has shown that for all $\varepsilon \in (0, 1]$ there is a constant $C_\varepsilon$ such that whenever $k \geq l \geq \varepsilon k$, then

$$r(k + 1, l + 1) \leq k^{-C_\varepsilon \frac{\ln k}{\ln \ln k}} \binom{k + l}{k}.$$
In joint work by David Conlon, Jacob Fox, and Benny Sudakov, the known bounds on $r(3, m, n)$ have also been recently improved.

### 2.1 The happy end problem

The Erdős-Szekeres Theorem 5 above is responsible for popularizing Ramsey’s theorem. Erdős and Szekeres reproved Ramsey’s result in 1934 when studying a question in elementary geometry, sometimes referred to as the “happy end problem.” (Another proof of Ramsey’s result had been found in 1933 by Skolem.)

**Definition 6** For all natural numbers $n \geq 3$ let $N(n)$ denote the least $N$, if it exists, such that given any $N$ points in general position in the plane (i.e., no four of them lying on the same circle, no three of them being collinear, no two of the lines they determine being parallel), there are $n$ that form the vertices of a convex $n$-gon.

For example, $N(3) = 3$ and $N(4) = 5$, as observed by Esther Klein, who also asked the general problem of showing that $N(n) < \omega$ for all $n$; $N(5) = 9$ is due to Makai, but the first published proof is due to J.D. Kalbfleisch, J.G. Kalbfleisch, and R. Stanton, in 1970.

To see that $N(4) = 5$, notice that 4 points do not suffice by considering the three vertices of a triangle and a fourth point inside. Now, given 5 points, consider their convex hull. We are done if it is not a triangle. If it is a triangle, there are two points inside, and the line $\ell$ they determine does not cross one of the sides of the triangle. The quadrilateral with vertices these two points and the two vertices of the segment not crossed by $\ell$ is convex.

**Theorem 7 (Erdős-Szekeres)** For all $n \geq 3$, $N(n)$ exists.
Shortly after their paper was published, Szekeres married Klein, which gave the name to the problem.

**Proof:** In fact, \( N(n) \leq r(3, 2, n) \). This is an argument of Michael Tarsy. As an undergraduate student at the Technion University, Tarsy was taking a combinatorics course under Lewin, and missed the lecture where the usual proof of the existence of \( N(n) \) was presented so, when asked to in an exam, he produced the following instead:

Let \( m = r(3, n, n) \) and let \( x_1, \ldots, x_m \) be \( m \) points in general position in the plane. For \( 1 \leq i < j < k \leq m \) color \( \{x_i, x_j, x_k\} \) blue if one encounters them in this order when going clockwise around the triangle they determine, and color them red otherwise.

One easily checks that four points determine a convex quadrilateral iff the set they form is homogeneous under the coloring just described, i.e., iff all its triples receive the same color. An easy induction shows that a set of \( n \) points determines a convex \( n \)-gon iff any 4 of them determine a convex quadrilateral. The result follows. \( \square \)

The best known upper bound is \( N(n) \leq \left( \frac{2n - 5}{n - 3} \right) + 2 \), shown by Tóth and Valtr in 1998. Shortly before, Bárány and Valtr showed that for any \( k \) there is a positive constant \( c_k \) such that any sufficiently large finite set \( X \) of points in the plane contains \( k \) subsets \( Y_1, \ldots, Y_k \), each of size at least \( c_k |X| \), such that every set \( \{y_1, \ldots, y_k\} \) with \( y_i \in Y_i \) is in convex position.

It is conjectured that \( N(n) = 2^{n-2} + 1 \) for all \( n \geq 3 \). That the right hand side is a lower bound is also due to Erdős and Szekeres.

### 2.2 A decidability result

Ramsey proved his result to obtain an instance of the *Entscheidungsproblem*, a decidability theorem. Fix a first order language with no function symbols, and consider the formulas in the language, of the form

\[
\varphi \equiv \exists x_1 \ldots \exists x_k \forall y_1 \ldots \forall y_m \phi,
\]

where \( \phi \) is quantifier free. These are the so-called Bernays-Schönfinkel formulas.

Ramsey proved that there is an algorithm that decides whether these formulas have (finite) models. In contrast, it is well known (from results of Church and Trakhtenbrot) that the problem is undecidable for general first order formulas, and even for those of the form

\[
\exists x_1 \ldots \exists x_k \forall y_1 \ldots \forall y_m \exists z_1 \ldots \exists z_n \phi
\]

with \( \phi \) quantifier free.
Definition 8 Let \((X, <)\) be a linearly ordered set. Let \(k < \omega\). Two sequences \((x_i : i < k)\) and \((y_i : i < k)\) of elements of \(X\) have the same ordering, 

\[
(x_i : i < k) \sim (y_i : i < k),
\]

iff for all \(i, j < k\), \(x_i < x_j\) iff \(y_i < y_j\).

Equivalently, if \(\sigma = \{x_i : i < k\}\) and \(\tau = \{y_i : i < k\}\), then \((x_i : i < k) \sim (y_i : i < k)\) iff \((\sigma, x_0, \ldots, x_{k-1}, <) \cong (\tau, y_0, \ldots, y_{k-1}, <)\).

The somewhat awkward definition allows for the possibility that two or more of the \(x_i\) (and consequently, of the \(y_i\)) are the same.

Definition 9 A \(k\)-ary relation \(R\) on a linearly ordered set \(X\) is \textit{canonical on} \(X\) iff whenever \((x_1, \ldots, x_k) \sim (y_1, \ldots, y_k)\), then 

\[
(x_1, \ldots, x_k) \in R \text{ iff } (y_1, \ldots, y_k) \in R.
\]

For example, the canonical unary relations on a set \(X\) are “true” (i.e., \(X\)) and “false” (i.e., \(\emptyset\)). The canonical binary relations are “false,” “\(<\),” “\(\leq\),” “\(\neq\),” “\(\geq\),” “\(>\),” and “true.”

Lemma 10 For all \(n_1, \ldots, n_k, t < \omega\) there is \(m < \omega\) such that whenever \([m]^t\) is \(n_i\)-colored for all \(i\) with \(1 \leq i \leq k\), then there is an \(H \in [m]^t\) such that \([H]^i\) is monochromatic for all such \(i\).

**Proof:** Let \(m_0 = t\) and for \(1 \leq i \leq k\) let \(m_i = r(i, n_i, m_{i-1})\). Then we can take \(m = m_k\). \(\square\)

Theorem 11 For all \(s_1, \ldots, s_k, t < \omega\) there is \(m < \omega\) such that whenever \(m \geq m\) (whether finite or infinite), and \(\mathcal{R}_i\) is a set of \(s_i\) many \(i\)-ary relations on \(m\) for all \(i\) with \(1 \leq i \leq k\), then there is \(S \in [m]^t\) on which all the relations in all the \(\mathcal{R}_i\) are canonical.

**Proof:** This is a consequence of Lemma 10. Consider a cardinal \(m\) and sets \(\mathcal{R}_i\) of \(i\)-ary relations on \(m\) for \(1 \leq i \leq k\), as in the statement of the theorem.

For \(1 \leq i \leq k\) define an equivalence relation on \([m]^t\) by saying that whenever \(x, y \in [m]^t\), if \((x_j : j < i)\) and \((y_j : j < i)\) are the increasing enumerations of \(x\) and \(y\), respectively, then \(x \sim y\) iff the following holds:
For all $l$ with $i \leq l \leq k$, all surjections $f : l \to i$, and all $R \in \mathcal{R}_l$,
\[ R(x_{f(0)}, \ldots, x_{f(l-1)}) \iff R(y_{f(0)}, \ldots, y_{f(l-1)}) \, . \]

For example, the equivalence class of a singleton $\{x\}$ is completely determined by the tuple of truth values of the relations $R(x, \ldots, x)$ for $R \in \bigcup_l \mathcal{R}_l$.

Clearly, there is an absolute (and easily computable) bound $n_i < \omega$ on the number of equivalence classes of these relations; “absolute” means here that $n_i$ only depends on $i$ and $(s_1, \ldots, s_k)$, but is independent of $m$ and of the specific relations in $\bigcup_l \mathcal{R}_l$.

Now use Lemma 10 with the $n_i$ just described and $t$ to find the required $m$, noting that if $S$ is monochromatic for the partition of $[m]^i$ into equivalence classes for all $i$ with $1 \leq i \leq k$, then all the relations in $\bigcup_l \mathcal{R}_l$ are canonical on $S$. □

**Theorem 12 (Ramsey)** Consider a language with only relation symbols, and let $T$ be a finite theory all of whose axioms are universal, i.e., of the form
\[ \forall x_1 \ldots \forall x_m \phi, \]
where $\phi$ is quantifier free. There is a finite number $m$ that only depends on the number $t$ of variables, and the number and arity of the relations used in $T$, such that for any $m \geq m$ (whether finite or infinite), there is a model of $T$ of size $m$ iff there is a canonical model of $T$ with universe $t$. Here, we say that a model is canonical iff all the relations in its language are canonical.

**Proof:** Letting $s_i$ be the number of $i$-ary relations mentioned in $T$, for $1 \leq i \leq k$ (for an appropriate $k < \omega$), we can take $m$ as in Theorem 11.

First, if there is a model $M$ of $T$ with universe $m \geq m$, there is certainly a canonical substructure of $M$ of size $t$, since the language of $T$ only has relation symbols. The universality of the axioms guarantees that this substructure is indeed a model of $T$.

Second, if there is a canonical model $M$ of $T$ with universe $t$, then there is one with universe any larger cardinal, obtained simply by extending each relation in the same canonical fashion. (For example, if a binary relation symbol $R$ is interpreted in $M$ by $\geq$, then we also let $\geq$ be its interpretation in any larger model.) That this extended structure is indeed a model of $T$ follows easily from canonicity. □

It is now relatively easy to extend the above result to give the decidability of the question of whether a Bernays-Schönfinkel sentence in a language with no function symbols is consistent. Essentially, one first replaces such a formula with a finite disjunction $\Phi$ of universal formulas as above (in a possibly larger language).

Now, for any such formula $\phi$, take $T = \{ \phi \}$ and $m, t$ as in the statement of Theorem 12, and note that whether or not $\phi$ has a model of size $\geq m$ reduces to whether it has
a canonical model with universe $t$. There are only finitely many such models, so one can easily determine whether this is the case. If not, then one only has to examine the finitely many structures in the language of $\phi$ of size less than $m$, to see whether at least one of them models $\phi$.

Finally, whether the original formula is consistent or not is equivalent to whether our search was successful for at least one of the formulas $\phi$ in the disjunction $\Phi$.

### 2.3 Tree arguments

It is convenient to examine other proofs of Ramsey’s Theorem 3, as they may illustrate different features than the argument above. Here is a proof that generalizes to the context of large cardinals; moreover, the technique it uses (trees and end-homogeneous sets) is widely applied in partition calculus. For the purpose of proving $\omega \rightarrow (\omega)^m_n$, the argument is certainly more involved than the one we have given; on the other hand, it provides us with a fairly concrete (constructive) description of how to find homogeneous sets for given colorings:

**Proof:** The proof is by induction on $m$. For $m = 1$ the result is obvious.

The key to the inductive step is a mixture of König’s lemma and the construction of what now are called end-homogeneous sets. We first present the case $m = 2$.

**Definition 13** A tree is a poset $(T, <)$ such that the $<$-predecessors of any element of the tree are well-ordered by $<$.

To any tree $T$ we can associate a rank function as usual,

$$\|x\|_T = \sup\{\|y\|_T + 1 : y < x\}.$$ 

For each $\alpha$, the set of $x \in T$ with $\|x\|_T = \alpha$ is called the $\alpha$-th level of the tree. The first $\alpha$ such that the $\alpha$-th level of $T$ is empty is called the height of $T$. We are primarily interested in $\omega$-trees, i.e., trees of height $\omega$.

**Definition 14** A branch through a tree $T$ is a maximal linearly ordered subset of $T$.

**Lemma 15** (König) Let $T$ be an $\omega$-tree such that all its levels are finite (we say that $T$ is infinite and finite branching). Then there is an infinite branch through $T$.

**Proof:** Define a branch $x = (x_n : n < \omega)$ by induction, so each $x_n$ is in the $n$-th level of $T$ and $\{s \in T : x_n < s\}$ is infinite. □
Given $f : [\omega]^2 \to n$ define a tree $(T, <)$ as follows: The nodes of the tree are the elements of $\omega$. We define $<$ by induction; in what follows, $<$ refers always to the ordering of $T$ rather than the usual $<$ relation among integers. Denote by $T_k$ the $k$-th level of $T$. We define $T$ by inductively defining the levels $T_k$. Let $T_0 = \{0\}$; the potential successors of $0$ are the elements of $\omega \setminus \{0\}$. Given $T_k$ and $a \in T_k$, let $A_a$ be the set of potential successors of $a$. Divide $A_a$ into the $n$ sets

$$A_{a,i} = \{b \in A_a : f(\{a, b\}) = i\}$$

for $i \in n$. The immediate successors of $a$ (all of which belong to $T_{k+1}$) are the least elements of these $n$ sets (there may be less than $n$ many of them, if one of the $A_{a,i}$ is empty). The potential successors of $j_i = \min A_{a,i}$ are the elements of $A_{a,i} \setminus \{j_i\}$, if any. This completes the construction of the tree. Notice that if $A_a$ is infinite, at least one of the $A_{a,i} \setminus \{j_i\}$ is infinite as well. It follows that the tree $T$ is an $\omega$-tree and that all its levels are finite; in fact, $|T_{s+1}| \leq n|T_s|$ for each $s \in \omega$.

By König’s lemma, there is an infinite branch $b = (b_l : l \in \omega)$ through $T$. By construction, this branch is end-homogeneous, i.e., it has the property that for any $i$ and any (larger) numbers $j, k$,

$$f(\{b_i, b_j\}) = f(\{b_i, b_k\}).$$

If this value is $l \in n$ say that $b_i$ is an $l$-node. Now apply the case $m = 1$ of the theorem: For some $l$, the set $H$ of $l$-nodes $b_i$ is infinite, and we have ensured it is homogeneous for $f$. This completes the proof of Ramsey’s theorem for $m = 2$.

Assume Ramsey’s theorem for $m$, and let $f : [\omega]^{m+1} \to n$. By a further trivial induction, we may also assume that $n = 2$. We define an $\omega$-tree $(T, <)$ all of whose levels are finite, and explain how to extract from any branch through $T$ an infinite homogeneous set for $f$.

The first $m$ levels of $T$ are simply given by $T_k = \{k\}$. To choose the successors of $m - 1$, divide $\omega \setminus m$ into two classes,

$$A_0 = \{j : m+1 \in j + 1 \text{ and } f(\{0, 1, \ldots, m - 1, j\}) = 0\}$$

and

$$A_1 = \{j : m+1 \in j + 1 \text{ and } f(\{0, 1, \ldots, m - 1, j\}) = 1\},$$

and define $T_m$ as the set of least elements of each class (so $T_m$ has size at most 2; it may be a singleton if one of the $A_i$ is empty). Say $j_i = \min A_i$. Then the elements of $\{k : j_i < k\}$ are to be chosen among the remaining elements of $A_i$ (if any).

Suppose we have already defined $T_s$, that $j \in T_s$, and that we have selected a subset $B_j$ of $\{k : j \in k\}$ of potential successors of $j$. Let $C_j = \{k \in T : k \leq j\}$, so $|C_j| = s + 1$. Define an equivalence relation on $B_j$ by

$$i \sim k \text{ iff } f(x \cup \{i\}) = f(x \cup \{j\}) \text{ for all } x \in [C_j]^m.$$
The immediate successors of \( j \) are the least elements of the \( \sim \)-equivalence classes, and the successors of any such element are to be chosen among the remaining elements of its class.

It follows that each node has finitely many immediate successors (since for each \( j \), the equivalence relation \( \sim \) on \( B_j \) admits only finitely many classes). If \( B_j \) is infinite, at least one of the \( \sim \)-classes is infinite as well, and it follows that \( T \) is an \( \omega \)-tree. By König’s lemma, it has an infinite branch \( r \). Again, \( r \) is end-homogeneous: By construction, \( r \) has the property that given any \( x \in [r]^m \), \( f(x \cup \{j\}) \) is constant for all \( j \in r \setminus (\max(x) + 1) \).

Say that \( X \) is of type 0 if this value is 0 and of type 1 otherwise. This gives a coloring \( g : [r]^m \to 2 \). By induction, there is an infinite \( g \)-homogeneous set \( H \subseteq r \). But then \( H \) is also \( f \)-homogeneous, and we are done. \( \square \)

### 2.4 Compactness

As it is well-known, an easy compactness argument shows that Ramsey’s theorem implies Theorem 4. The disadvantage of using compactness is that no discernible finite bounds seem to follow from such an argument:

**Proof:** We want to show that for all natural numbers \( n, k, l \) there is an \( N < \omega \) such that \( N \rightarrow (n)^k \). Suppose otherwise and fix counterexamples \( n, k, l \), so for all \( N \) there is a function \( f : [N]^k \to l \) that does not admit homogeneous sets of size \( n \). Given two counterexamples \( f, g \), say that \( f < g \) iff \( f \subseteq g \), i.e., there must be integers \( N_1 < N_2 \) such that \( \text{dom}(f) = [N_1]^k \), \( \text{dom}(g) = [N_2]^k \), and \( g \rceil [N_1]^k = f \).

This gives a tree structure to the set of all such functions. Notice that this tree is infinite, since there is a counterexample with domain \([N]^k\) for all \( N \); it is also finite branching, since there are only finitely many counterexamples with domain \([N]^k\) for any given \( N \) because, in fact, there are only finitely many functions \( f : [N]^k \to l \).

By König’s lemma there is a branch through this tree. The functions along this branch cohere, and we can then define a function \( F : [\omega]^k \to l \) by taking their union. But, by construction, \( F \) does not admit a homogeneous set of size \( n \), which of course contradicts Ramsey’s theorem that claims that there must in fact be an infinite homogeneous set. \( \square \)

This is a typical example of a compactness argument, called this way because it can be recast as a corollary of the compactness of the Tychonoff product \( \prod_n X_n \) of certain finite (discrete) sets \( X_n \). It is also a consequence of the compactness of first-order logic. There is in fact a very general compactness result that applies even in situations where no tree argument as above is available.
Theorem 16 (Rado’s selection principle) Let $A_i$ be a finite nonempty set for all $i \in I$, and let $A = \prod_i A_i$. For $V \in [I]^{<\omega}$ set $B_V = \prod_{i \in V} A_i$ and let $f_V \in B_V$ be given. Then there exists $f \in A$ such that for all $W \in [I]^{<\omega}$ there is $V \in [I]^{<\omega}$ such that $V \supseteq W$ and

$$f \upharpoonright W = f_V \upharpoonright W.$$ 

**Proof:** For $W \in [I]^{<\omega}$ let

$$T_W = \{g \in B_W : \exists V \supseteq W \text{ finite such that } g = f_V \upharpoonright W\}$$

and

$$A_W = \{f \in A : f \upharpoonright W \in T_W\}.$$

We need to show that $\bigcap_{W \in [I]^{<\omega}} A_W \neq \emptyset$.

This is an immediate consequence of Tychonoff’s theorem: Note that each $A_W$ is a closed subset of the Tychonoff product $A$ of the discrete sets $A_i$, because each $T_W$ is finite, being a subset of the finite set $B_W$, and we can write

$$A_W = \bigcup_{g \in T_W} \{f \in A : f \upharpoonright W = g\}$$

as a finite union of closed sets. Let

$$F = \{A_W : W \in [I]^{<\omega}\}.$$

We claim that $F$ has the finite intersection property. If this is the case, it follows from the compactness of $A$ that $\bigcap F \neq \emptyset$ and we are done.

To prove the claim, let $n < \omega$ and let $W_k$ for $k < n$ be finite subsets of $I$. Then $V = \bigcup_k W_k$ is finite, and $f_V$ has an extension $f \in A$. By definition,

$$f \upharpoonright W_k = f_V \upharpoonright W_k \in T_{W_k}$$

for all $k < n$, so $f \in A_{W_k}$ for all $k < n$, and we are done.

Actually, we can quickly provide a self-contained proof of the remainder of the argument: Since $F$ has the finite intersection property, it can be extended to an ultrafilter $\mathcal{U}$ over $A$. For $i \in I$ and $x \in A_i$ let

$$A_{i,x} = \{f \in A : f(i) = x\},$$

so $A = \bigcup_{x \in A_i} A_{i,x}$ for any fixed $i \in I$. This is a finite union, so it follows that for each $i$ there is some $x_i$ such that $A_{i,x_i} \in \mathcal{U}$. Let $f \in A$ be the function $f(i) = x_i$. Then $f \in \bigcap F$ as required: Let $W \in [I]^{<\omega}$. Then

$$\bigcap_{i \in W} A_{i,f(i)} \cap A_W \in \mathcal{U},$$

12
since $W$ is finite, all the $A_{i,f(i)}$ are in $U$, and $A_W \in F$. In particular, this intersection is nonempty, and we can fix one of its elements, say $g$. Then $g \upharpoonright W = f \upharpoonright W$ and $g \upharpoonright W \in T_W$. Hence, $f \in A_W$, and we are done. □

One can also prove Rado’s selection principle as a consequence of the compactness of first order logic; for example, it is easy to give a proof using the ultrapower construction as in the proof of 2 implies 6 of Theorem 8 in lecture II.11.

Theorem 4 is an easy consequence of Theorem 16: Given $n, k, m < \omega$ we want to show that for some $N < \omega$ we have $N \to (n)^k_m$. Assume otherwise. It follows that for each finite $W \subset \omega$ there is some function $g_W : [W]^k \to m$ without a homogeneous set of size $n$. Let $I = [\omega]^k$ and for each $i \in I$ let $A_i = m$. For each $W \subset \omega$ finite let $f_W \in \prod_{i \in [W]^k} m$ be $g_W$ as above. For an arbitrary $V \in [I]^<\omega$ pick some $W \subset \omega$ finite such that $V \subseteq [W]^k$ and let $f_V \in \prod_{i \in V} m$ be $f_W \upharpoonright V$. Rado’s selection principle now gives us an $m$-coloring of $[\omega]^k$ without homogeneous sets of size $n$, as any such set would be contained in some $[W]^n$, $W \subset \omega$ finite, such that $f \upharpoonright [W]^k = g_W$, and we have a contradiction.

Another well-known consequence of Rado’s principle is the de Bruijn-Erdős theorem on the chromatic number of infinite graphs:

**Definition 17** If $G$ is a graph, $\chi(G)$, the chromatic number of $G$, is the smallest cardinal $\kappa$ such that there is a map $f : V \to \kappa$, where $V$ is the set of vertices of $G$, such that whenever $f(a) = f(b)$, then there is no edge in $G$ between $a$ and $b$.

**Theorem 18 (de Bruijn-Erdős)** Let $G$ be a graph. If $n < \omega$ and every finite subgraph $G$ of $G$ has chromatic number at most $n$, $\chi(G) \leq n$, then the same holds of $G$, $\chi(G) \leq n$.

**Proof:** This is also an easy consequence of Theorem 16. Now take $I$ to be the set of vertices of $G$ and each $A_i$ to be $n$, and take $f_W$ for $W \in [I]^<\omega$ to be a coloring witnessing that the chromatic number of the subgraph of $G$ spanned by $W$ is at most $n$. □

Unlike the finite version of Ramsey’s theorem, the de Bruijn-Erdős result is not a consequence of König’s lemma, since $G$ may well be uncountable. It follows immediately from Theorem 18 and the well-known 4-color theorem that every planar graph is 4-colorable.

A well-known open problem in geometric graph theory, the **Hadwiger-Nelson problem**, asks for the minimum number of colors required to color the plane such that no two points at distance one from each other have the same color. Abusing notation a bit, let’s call this number $\chi(\mathbb{R}^2)$. It is known that $4 \leq \chi(\mathbb{R}^2) \leq 7$; the drawing below (taken from the wikipedia entry on the problem) shows both a coloring of the plane with 7 colors witnessing $\chi(\mathbb{R}^2) \leq 7$, and a finite graph showing that $4 \leq \chi(\mathbb{R}^2)$. 

13
Whatever the true value of $\chi(\mathbb{R}^2)$, Theorem 18 shows that any lower bound can be verified by examining some finite configuration. A recent remark of Shelah and Soifer suggests that the value of $\chi(\mathbb{R}^2)$ may depend on whether the axiom of choice is adopted or not; Falconer has shown, for example, that $\chi(\mathbb{R}^2) \geq 5$ if we require the monochromatic sets to be Lebesgue measurable. In particular, this must be the case in Solovay’s model, or under determinacy, even if every finite subset of the plane has chromatic number at most 4.

Truszczynski-Tuza and Rav, among others, have studied consequences of Rado’s principle in algebra and combinatorics.

2.5 Ultrafilters

There is yet another proof of Ramsey’s theorem that is worth discussing. It explains that infinite homogeneous sets can be found because one of the colors is “large” in a precise sense. For this, we need a construction closely related to the product of ultrafilters discussed in lecture II.11.

It is convenient to think of ultrafilters (over $\omega$ in this discussion, but the argument holds in general) as generalized quantifiers: Let $\mathcal{U}$ be an ultrafilter over $\omega$. Given a formula $\varphi(x)$, we interpret $\mathcal{U}n \varphi(n)$ as saying that “almost all” $n$ have property $\varphi$,

$$\{ n \in \omega : \varphi(n) \} \in \mathcal{U}.$$ 

Clearly, we have:

- $\neg \mathcal{U}n \varphi(n) \Leftrightarrow \mathcal{U}n \neg \varphi(n)$.
- $(\mathcal{U}n \varphi(n) \land \mathcal{U}n \psi(n)) \Leftrightarrow \mathcal{U}n (\varphi(n) \land \psi(n))$.
- Similarly with any of the standard binary connectives in place of $\land$. 

14
Definition 19 Given a nonprincipal ultrafilter $U$ over $\omega$ and $0 < k \in \omega$, define $U^k$ as the set of all $X \in [\omega]^k$ such that

$$Un_0 \ldots Un_{k-1} (\{n_0, \ldots, n_{k-1}\} \in X).$$

Note that $UnUm (n \neq m)$, since $U$ is nonprincipal. It then follows that for any $X \in [\omega]^k$, $X \in U^k$ iff $[\omega]^k \setminus X \notin U^k$. Similarly, $U^k$ is closed under finite intersections and supersets, so it is an ultrafilter. It is nonprincipal, since $Un (n \notin A)$ for any finite $A \subset \omega$, from which one easily gets that any $X \in U^k$ is infinite.

We now fix a nonprincipal $U$, and use $U^k$ to give a proof that $\omega \to (\omega)^k_m$. In fact, given $f : [\omega]^k \to m$, we will obtain an infinite homogeneous set $M$ such that $[M]^k$ is contained in a color in $U^k$. Without loss, we may assume that $m = 2$.

- Note that there is a color $P = P_0 \in U^k$. More precisely, $P = f^{-1}(\{i\})$ for some $i \in 2$.

Given $s \in [\omega]^l$, $l < k$, let

$$P_s = \{t \in [\omega]^{k-l} : \max(s) < \min(t) \text{ and } s \cup t \in P\}.$$

We then have:

- $\{s \in [\omega]^l : P_s \in U^{k-l}\} \in U^l$ whenever $0 \leq l < k$,

where “$P_s \in U^0$” means “$s \in P$.”

Let $T = \{s \in [\omega]^{\leq k} : P_s \in U^{k-|n|}\}$. Then:

- $\emptyset \in T$.

- If $s \in T \cap [\omega]^{<k}$ then $T_s = \{n : s \cup \{n\} \in T\} \in U$.

Hence we can build an increasing sequence $m_0 < m_1 < \ldots$ such that for all $j$ and all $s \in [\{m_i : i < j\}]^{<k}$, we have $T_s \supseteq \{m_i : i \geq j\}$. But then, letting $M = \{m_j : j \in \omega\}$, we have that $[M]^k \subseteq P$, and we are done.

(Notice the similarities between this and the tree argument.)

Bibliography

Here are some additional references (besides those mentioned at the beginning) consulted while preparing this note:


*Typeset using LaTeX2WP.*