

Combinatorial Number Theory

LECTURE NOTES

Contents

1	Ramsey's Theorem	3
1.1	Ramsey's Theorem for graphs	3
1.2	The compactness principle for colorings	5
1.3	Ramsey's Theorem for 2-sets	6
1.4	Schur's Theorem	7
1.5	Ramsey's Theorem for k -sets	10
1.6	Ramsey's Theorem for hypergraphs	11
1.7	Erdős-Szekeres' Theorem on convex polygons	12
1.8	Erdős-Szekeres' Theorem on monotone paths	15

Chapter 1

Ramsey's Theorem

1.1. Ramsey's Theorem for graphs

Definition 1. A graph $G = (V, E)$ is a set V of points, called *vertices*, and a set E of distinct pairs of vertices, called *edges*.

Definition 2. A subgraph $G' = (V', E')$ of a graph $G = (V, E)$ is a graph such that $V' \subseteq V$ and $E' \subseteq E$.

Figure 1.1 below depicts a graph G with four vertices $V = \{V_1, V_2, V_3, V_4\}$ and four edges $E = \{e_1, e_2, e_3, e_4\}$, where $e_1 = \{V_1, V_2\}$, $e_2 = \{V_2, V_3\}$, $e_3 = \{V_3, V_4\}$, and $e_4 = \{V_2, V_4\}$. Note that edges are *unordered* pairs of vertices, meaning that $\{V_1, V_2\}$ and $\{V_2, V_1\}$ refer to the same edge. Next to it is a graph $G' = (V', E')$ with $V' = V = \{V_1, V_2, V_3, V_4\}$ and $E' = \{e_1, e_3\}$. Since $V' \subseteq V$ and $E' \subseteq E$, we deduce that G' is a subgraph of G .

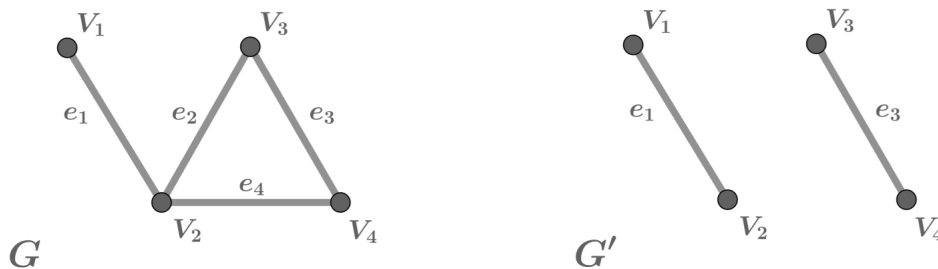


Figure 1.1: A graph G and one of its subgraphs G' .

Definition 3. Given $n \in \mathbb{N}$, a *complete graph on n vertices*, denoted by K_n , is a graph with n vertices and the property that every pair of distinct vertices is connected by an edge.

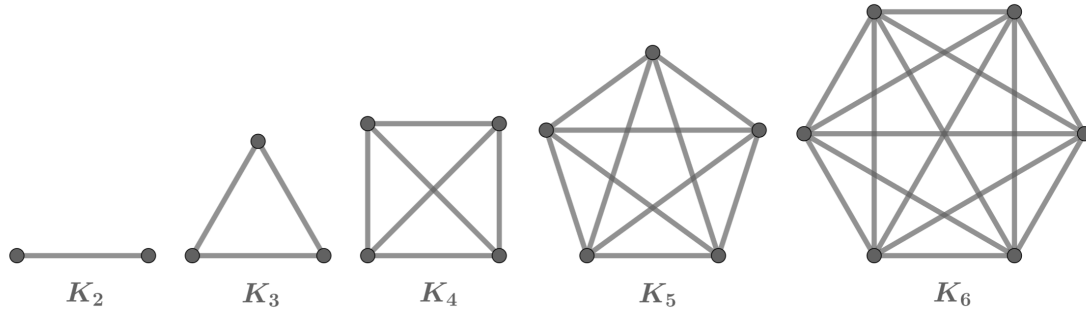


Figure 1.2: A depiction of K_n for $n = 2, 3, 4, 5$, and 6.

Definition 4. An *edge-coloring* of a graph $G = (V, E)$ is an assignment of a color to each edge of the graph. A graph that has been edge-colored is called *monochromatic* if all of its edges are the same color.

An edge-coloring of a graph can also be viewed as a function where the domain is the set of edges of the graph and the codomain is the set of colors. For example, suppose one has a graph with edges $E = \{e_1, e_2, e_3\}$ and a set of colors $C = \{\text{red}, \text{blue}\}$. A valid coloring of this graph can be seen as a function $\chi: E \rightarrow C$, where, for instance, $\chi(e_1) = \text{red}$, $\chi(e_2) = \text{blue}$, and $\chi(e_3) = \text{red}$.

Ramsey's Theorem for graphs. For any $n, m \in \mathbb{N}$ there exists $R = R(n, m) \in \mathbb{N}$ such that any edge-coloring of K_R with at most m colors contains a monochromatic copy of K_n as a subgraph.

Let us illustrate the content of Ramsey's Theorem for graphs by looking at an example. If the edge-coloring consists only of two colors, say **red** and **blue**, and we assume $n = 3$, then Ramsey's Theorem asserts that there exists a number $R(3, 2)$ such that any edge-coloring of a complete graph on $R(3, 2)$ vertices admits a monochromatic triangle. Note that $R(3, 2)$ cannot equal 5, because Figure 1.3 below shows a 2-coloring of K_5 containing no monochromatic triangle. However, taking

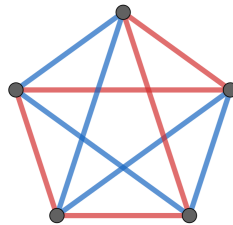


Figure 1.3: An edge-coloring of K_5 containing no monochromatic copy of K_3 .

Proof. The implication (ii) \implies (i) is immediate, so it only remains to prove (i) \implies (ii). We can view a coloring of Y that uses no more than m colors as a function $\chi: Y \rightarrow \{1, \dots, m\}$ simply by associating a number from 1 to m with each color. This means the space of all possible colorings of Y can be identified with the product space $\{1, \dots, m\}^Y$. Note that the finite set $\{1, \dots, m\}$, endowed with the discrete topology, is a compact Hausdorff space. By Tychonoff's theorem, $\{1, \dots, m\}^Y$ endowed with the product topology is therefore also a compact Hausdorff space.

For any finite non-empty set $Z \subseteq Y$ let \mathcal{C}_Z be the set of all colorings in $\{1, \dots, m\}^Y$ for which there is monochromatic $F \in \mathcal{F}$ with $F \subseteq Z$. Then \mathcal{C}_Z is an open set in the product topology on $\{1, \dots, m\}^Y$. Moreover, in light of statement (i), we have

$$\bigcup_{\substack{Z \subseteq Y \\ 0 < |Z| < \infty}} \mathcal{C}_Z = \{1, \dots, m\}^Y.$$

By compactness, the above cover admits a finite subcover, or in other words, there exist finite non-empty sets $Z_1, \dots, Z_\ell \subseteq Y$ such that $\mathcal{C}_{Z_1} \cup \dots \cup \mathcal{C}_{Z_\ell} = \{1, \dots, m\}^Y$. Taking $Z = Z_1 \cup \dots \cup Z_\ell$ and noting that $\mathcal{C}_{Z_1} \cup \dots \cup \mathcal{C}_{Z_\ell} \subseteq \mathcal{C}_{Z_1 \cup \dots \cup Z_\ell}$, it follows that $\mathcal{C}_Z = \{1, \dots, m\}^Y$, completing the proof. \square

1.3. Ramsey's Theorem for 2-sets

Definition 5. A *2-set* is a set consisting of exactly two elements. Given a set X , a *2-subset* of X is any subset of X that is a 2-set. We will use $X^{(2)}$ to denote the set of all 2-subsets of X .

We have already seen examples of 2-subsets in the previous section. Indeed, the set of edges E of a graph $G = (V, E)$ consists of 2-subsets of the set of vertices V . In other words, $E \subseteq V^{(2)}$. Note that a graph $G = (V, E)$ is a complete graph if and only if $E = V^{(2)}$.

Definition 6. Let X be a set. A *coloring of $X^{(2)}$* is an assignment of a color to each 2-subset of X . We call $X^{(2)}$ *monochromatic* if all elements in $X^{(2)}$ have the same color.

The following can be viewed as an “infinitary” version of Ramsey's Theorem for graphs.

Ramsey's Theorem for 2-sets. *Let X be an infinite set. Then for any finite coloring of $X^{(2)}$ there exists an infinite subset $Y \subseteq X$ such that $Y^{(2)}$ is monochromatic.*

Proof. Fix an arbitrary element $x_1 \in X$ and note that any 2-set of the form $\{x_1, x\}$ for $x \in X \setminus \{x_1\}$ has a certain color. Since the number of colors is finite but the set $X \setminus \{x_1\}$ is infinite, there exists an infinite subset $X_1 \subseteq X \setminus \{x_1\}$ such that all 2-sets of the form $\{x_1, x\}$ for $x \in X_1$ have the same color. Now fix an arbitrary element $x_2 \in X_1$

and let us repeat the same procedure. Any 2-set of the form $\{x_2, x\}$ for $x \in X_1 \setminus \{x_2\}$ has a certain color. For the same reason as before, since the number of colors is finite but the set $X_1 \setminus \{x_2\}$ is infinite, there exists an infinite subset $X_2 \subseteq X_1 \setminus \{x_2\}$ such all 2-sets of the form $\{x_2, x\}$ for $x \in X_2$ have the same color. Continuing this procedure produces an infinite sequence of distinct elements x_1, x_2, x_3, \dots and a nested family of infinite sets $X \supseteq X_1 \supseteq X_2 \supseteq X_3 \supseteq \dots$ such that for all $i \in \mathbb{N}$ we have $x_{i+1} \in X_i$ and the set $\{\{x_i, x\} : x \in X_i\}$ is monochromatic.

Let c_i denote the color of elements in the set $\{\{x_i, x\} : x \in X_i\}$. Then c_1, c_2, c_3, \dots is an infinite sequence of colors. Since there are only finitely many different colors, one color must appear infinitely often in this sequence. In other words, there exists a color c and an infinite sequence $i_1 < i_2 < i_3 < \dots \in \mathbb{N}$ such that $c_{i_k} = c$ for all $k \in \mathbb{N}$.

To finish the proof, define $Y = \{x_{i_k} : k \in \mathbb{N}\}$ and observe that any 2-subset of Y is of the form $\{x_{i_k}, x_{i_\ell}\}$ for $k < \ell \in \mathbb{N}$. Since $x_{i_\ell} \in X_{i_{\ell-1}}$ and $X_{i_{\ell-1}} \subseteq X_{i_k}$, the 2-set $\{x_{i_k}, x_{i_\ell}\}$ has the color c . Hence all 2-subsets of Y have the color c , which proves that $Y^{(2)}$ is monochromatic. \square

Proposition 7. *Ramsey's Theorem for 2-sets implies Ramsey's Theorem for graphs.*

Proof. Fix $n, m \in \mathbb{N}$. Let $\mathbb{N}^{(2)}$ denote the set of all 2-element subsets of \mathbb{N} , and define

$$\mathcal{F} = \{F^{(2)} : F \subseteq \mathbb{N}, |F| = n\}.$$

By Ramsey's Theorem for 2-sets, any coloring of $\mathbb{N}^{(2)}$ with at most m colors admits a monochromatic set of the form $F^{(2)}$ for some $F \subseteq \mathbb{N}$ with $|F| = n$.

By the Compactness Theorem for finite colorings, there exists a finite set $Z \subseteq \mathbb{N}^{(2)}$ such that any coloring of Z with at most m colors already contains a monochromatic $F^{(2)}$ with $F \subseteq \mathbb{N}$ and $|F| = n$. Enlarging Z if necessary, we may assume $Z = H^{(2)}$ for some finite $H \subseteq \mathbb{N}$.

In other words, for every m -coloring of $H^{(2)}$, there exists a subset $F \subseteq H$ of size n such that $F^{(2)}$ is monochromatic. Define $R = R(n, m) := |H|$. Identifying $H^{(2)}$ with the edge set of the complete graph K_R , the subset $F^{(2)}$ corresponds to a copy of K_n inside K_R . Thus we have shown that any edge-coloring of K_R with at most m colors contains a monochromatic copy of K_n . This completes the proof. \square

1.4. Schur's Theorem

Fermat's Last Theorem states that for $m \geq 3$ the equation

$$x^m + y^m = z^m \tag{1.4.1}$$

has no positive integer solutions $x, y, z \in \mathbb{N}$. For centuries, this remained one of the biggest open problems in mathematics, and one whose intriguing nature captivated many mathematicians. Among them was also Issai Schur, who investigated a

natural, localized version of Fermat's Last Theorem. More precisely, he wondered whether for any $m \geq 2$ the congruence equation

$$x^m + y^m \equiv z^m \pmod{p} \quad (1.4.2)$$

possesses non-trivial solutions for all but finitely many primes p . Note that any non-trivial solution to Fermat's equation $x^m + y^m = z^m$ also offers a non-trivial solution to Schur's equation $x^m + y^m \equiv z^m \pmod{p}$ for all primes p satisfying $p > z^m$, but not the other way around. In order to address (1.4.2), Schur proved a theorem that is often regarded as the earliest result in Ramsey Theory:

Schur's Theorem ([Sch17]). *For any $m \in \mathbb{N}$ there exists $S = S(m) \in \mathbb{N}$ such that if the set $\{1, \dots, S\}$ is colored using at most m colors then there exist monochromatic $x, y, z \in \{1, \dots, S\}$ with $x + y = z$.*

Proof. Take $S = R(3, m)$, where $R(3, m)$ is the Ramsey number for $(3, m)$. Let K_S denote the complete graph on S vertices and denote the vertices of K_S by V_1, V_2, \dots, V_S . Any coloring of the set $\{1, \dots, S\}$ induces an edge-coloring on K_S by assigning to each edge $\{V_i, V_j\}$ the color of the number $|i - j| \in \{1, \dots, S\}$. According to Ramsey's Theorem for graphs, K_S contains a monochromatic triangle. Let V_a, V_b , and V_c , for $a < b < c$, be the vertices of this monochromatic triangle. By setting

$$x = b - a, \quad y = c - b, \quad \text{and} \quad z = c - a,$$

it is then easy to check that x, y, z have the same color and satisfy $x + y = z$. \square

The smallest possible positive integer $S(m)$ for which the conclusion of Schur's Theorem holds is referred to as the *Schur number* for m . The known Schur numbers to date are:

m	Schur Number
2	5
3	14
4	45
5	161
6	unknown
7	unknown
\vdots	

Here is an example from Schur's original paper [Sch17] of a 3-coloring of $\{1, \dots, 13\}$ admitting no monochromatic solution to the equation $x + y = z$:

color 1: {2, 3, 11, 12}

color 2: {5, 6, 8, 9}

color 3: {1, 4, 7, 10, 13}

More examples along these lines can be found here: <https://oeis.org/A030126>.

The proof that the Schur number for 5-colorings equals 161 took up 2 petabytes of space. Even though every 5-coloring of $\{1, \dots, 161\}$ admits a monochromatic solution to $x + y = z$, there are 2447113088 many 5-colorings of $\{1, \dots, 160\}$ admitting no monochromatic solution to $x + y = z$.

With the help of the above theorem, Schur was able to show that, contrary to Fermat's equation (1.4.1), its "local" counterpart (1.4.2) does possess non-trivial solutions.

Theorem 8. *Let $m \in \mathbb{N}$. There exists $F = F(m)$ such that for all prime numbers $p > F$ there exist $x, y, z \in \{1, \dots, p-1\}$ with $x^m + y^m \equiv z^m \pmod{p}$.*

For the proof of Theorem 8, we will need the following basic fact from algebra, the proof of which is left to the interested reader.

Lemma 9. *Let $(K, +, \cdot)$ be a field and $f(x) \in K[x]$ a polynomial of degree $\deg(f) = m$ with coefficients in K . Then the number of roots of $f(x)$ is at most m .*

Let us now see the proof of Theorem 8.

Proof of Theorem 8. Take $F = S(m)$, where $S(m)$ is as guaranteed by Schur's Theorem. Let p be any prime number bigger than F . The set $\mathbb{F}_p = \{0, 1, \dots, p-1\}$ of congruence classes modulo p naturally forms a field $(\mathbb{F}_p, +, \cdot)$ under the modular arithmetic operations $+$ and \cdot . Let $\mathbb{F}_p^\times = \mathbb{F}_p \setminus \{0\}$ and consider the set

$$C := \{x^m : x \in \mathbb{F}_p^\times\}.$$

Note that C is a subgroup of the multiplicative group $(\mathbb{F}_p^\times, \cdot)$. This means that \mathbb{F}_p^\times can be covered by cosets of C . More precisely, there exist coset representatives $g_1, g_2, \dots, g_r \in \mathbb{F}_p^\times$ such that

$$\mathbb{F}_p^\times = g_1 C \cup g_2 C \cup \dots \cup g_r C. \quad (1.4.3)$$

Any element of \mathbb{F}_p^\times is a root of the polynomial $x^m - a$ for some $a \in C$, or equivalently,

$$\mathbb{F}_p^\times = \bigcup_{a \in C} \{x : x^m - a \equiv 0 \pmod{p}\}.$$

But since each such polynomial has at most m roots due to Lemma 9, we have

$$|\{x : x^m - a \equiv 0 \pmod{p}\}| \leq m.$$

It follows that

$$|\mathbb{F}_p^\times| = \left| \bigcup_{a \in C} \{x : x^m - a \equiv 0 \pmod{p}\} \right| \leq \sum_{a \in C} |\{x : x^m - a \equiv 0 \pmod{p}\}| \leq m \cdot |C|.$$

So C can have at most m cosets, or in other words, $r \leq m$. Since $p > F$, the set $\{1, \dots, F\}$ is a subset of $\mathbb{F}_p^\times = \{1, \dots, p-1\}$ and hence (1.4.3) yields a partition of the set $\{1, \dots, F\}$ involving r disjoint cells. We can think of this partition as a coloring of $\{1, \dots, F\}$ using r colors. Since $F = S(m)$ and $r \leq m$, it follows from Schur's Theorem that there exist monochromatic $\tilde{x}, \tilde{y}, \tilde{z} \in \{1, \dots, F\}$ for which $\tilde{x} + \tilde{y} = \tilde{z}$. Since $\tilde{x}, \tilde{y}, \tilde{z}$ have

the same color, they all belong to the same coset. In other words, there exists a coset representative $g_i \in \{g_1, \dots, g_r\}$ such that $\tilde{x}, \tilde{y}, \tilde{z} \in g_i C$. Take any $x, y, z \in \mathbb{F}_p^\times$ for which

$$\tilde{x} \equiv g_i x^m \pmod{p}, \quad \tilde{y} \equiv g_i y^m \pmod{p}, \quad \text{and} \quad \tilde{z} \equiv g_i z^m \pmod{p},$$

which is possible because $\tilde{x}, \tilde{y}, \tilde{z} \in g_i C$. Then we have

$$g_i x^m + g_i y^m \equiv g_i z^m \pmod{p},$$

from which it follows that

$$x^m + y^m \equiv z^m \pmod{p},$$

because $g_i \not\equiv 0 \pmod{p}$. □

1.5. Ramsey's Theorem for k -sets

Definition 10. A k -set is a set consisting of exactly k elements. Given a set X , a k -subset of X is any subset of X that is a k -set. We will use $X^{(k)}$ to denote the set of all k -subsets of X .

We have already seen Ramsey's Theorem for 2-sets. Here is Ramsey's result in full generality.

Ramsey's Theorem for k -sets ([Ram30]). *Let X be an infinite set and $k \geq 2$. Then for any finite coloring of $X^{(k)}$ there exists an infinite subset $Y \subseteq X$ such that $Y^{(k)}$ is monochromatic.*

Proof. Let us use a proof by induction on k . The base case of the induction, when $k = 2$, follows from Ramsey's Theorem for 2-sets established in Section 1.3. To prove the inductive step, assume $k \geq 3$ and Ramsey's Theorem has already been proven for $(k - 1)$ -sets. Let $Y_0 = X$ and fix an arbitrary element $y_1 \in Y_0$. Note that any k -set of the form $\{y_1, x_2, \dots, x_k\}$ for $\{x_2, \dots, x_k\} \in (Y_0 \setminus \{y_1\})^{(k-1)}$ has a certain color, which induces a finite coloring on $(Y_0 \setminus \{y_1\})^{(k-1)}$. Applying Ramsey's Theorem for $(k - 1)$ -sets, we can find an infinite subset $Y_1 \subseteq Y_0 \setminus \{y_1\}$ such that all k -sets of the form $\{y_1, x_2, \dots, x_k\}$ for $\{x_2, \dots, x_k\} \in Y_1^{(k-1)}$ are monochromatic. Next, fix an arbitrary element $y_2 \in Y_1$ and repeat the same procedure. The given coloring of k -sets of the form $\{y_2, x_2, \dots, x_k\}$ for $\{x_2, \dots, x_k\} \in (Y_1 \setminus \{y_2\})^{(k-1)}$ induces a finite coloring of $(Y_1 \setminus \{y_2\})^{(k-1)}$. Applying Ramsey's Theorem for $(k - 1)$ -sets once more yields an infinite subset $Y_2 \subseteq Y_1 \setminus \{y_2\}$ such that all k -sets of the form $\{y_2, x_2, \dots, x_k\}$ for $\{x_2, \dots, x_k\} \in Y_2^{(k-1)}$ are monochromatic. Continuing this procedure produces an infinite sequence of distinct elements y_1, y_2, y_3, \dots and a nested family of infinite sets $X = Y_0 \supseteq Y_1 \supseteq Y_2 \supseteq Y_3 \supseteq \dots$ such that for all $i \in \mathbb{N}$ the set $\{\{y_i, x_2, \dots, x_k\} : \{x_2, \dots, x_k\} \in Y_i^{(k-1)}\}$ is monochromatic. Moreover, we have $y_{i+1} \in Y_i$ for all $i \in \mathbb{N}$.

Let c_i denote the color of elements in the set $\{y_i, x_2, \dots, x_k\} : \{x_2, \dots, x_k\} \in Y_i^{(k-1)}\}$. Since the sequence c_1, c_2, c_3, \dots is infinite but the number of colors is finite, one color must appear infinitely often in this sequence. In other words, there exists a color c and an infinite subsequence $c_{i_1}, c_{i_2}, c_{i_3}, \dots \in \mathbb{N}$ such that $c_{i_\ell} = c$ for all $\ell \in \mathbb{N}$. To finish the proof, define $Y = \{y_{i_k} : k \in \mathbb{N}\}$ and observe that any k -subset of Y is of the form $\{y_{i_{\ell_1}}, \dots, y_{i_{\ell_k}}\}$ for $\ell_1 < \dots < \ell_k \in \mathbb{N}$. Since $\{y_{i_{\ell_2}}, \dots, y_{i_{\ell_k}}\} \in Y_{i_{\ell_1}}$ because $\ell_1 < \ell_2 < \dots < \ell_k$, the k -set $\{y_{i_{\ell_1}}, \dots, y_{i_{\ell_k}}\}$ has the color c . Hence all k -subsets of Y have the color c , which proves that $Y^{(k)}$ is monochromatic. \square

1.6. Ramsey's Theorem for hypergraphs

A hypergraph is a generalization of a graph in which an edge can join multiple vertices at once.

Definition 11. Let $k \in \mathbb{N}$. A k -uniform hypergraph is a pair $G = (V, E)$ where V is a set of points, called *vertices*, and $E \subseteq V^{(k)}$ is a set of k -subsets of V , called *hyperedges*.

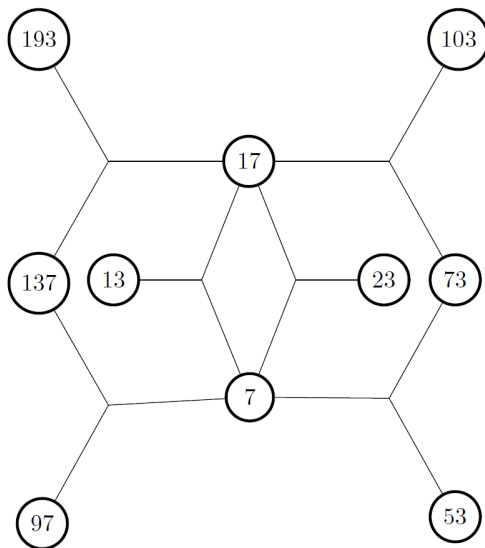


Figure 1.5: Here is an example of a 3-uniform hypergraph with vertices $V = \{7, 13, 17, 23, 53, 73, 97, 103, 137, 193\}$, where three vertices are connected by a hyperedge if and only if their squares form a 3-term arithmetic progression. For example, $\{7, 13, 17\}$ is an edge, because $7^2, 13^2, 17^2$ are in an arithmetic progression.

Given $k, n \in \mathbb{N}$ with $k \leq n$, a *complete k -uniform hypergraph on n vertices* is a k -uniform hypergraph $G = (V, E)$ where the set of vertices has cardinality n and where every set of k distinct vertices in V is connected by an edge. In other words, $G = (V, E)$ is a complete k -uniform hypergraph on n vertices if $|V| = n$ and $E = V^{(k)}$.

Ramsey's Theorem for hypergraphs. For any $n, m, k \in \mathbb{N}$ there exists a number $R = R_k(n, m) \in \mathbb{N}$ such that any edge-coloring of a complete k -uniform hypergraph on R vertices with at most m colors admits a monochromatic copy of a complete k -uniform hypergraph on n vertices.

Proof. Let $n, m, k \in \mathbb{N}$ be given. It follows from Ramsey's Theorem for k -sets that for any m -coloring of $\mathbb{N}^{(k)}$ there exists a set $S \subseteq \mathbb{N}$ with $|S| = n$ such that $S^{(k)}$ is monochromatic. If we now apply the Compactness Theorem for finite colorings to this statement (with $Y = \mathbb{N}^{(k)}$ and $\mathcal{F} = \{S^{(k)} : S \subseteq \mathbb{N}, |S| = n\}$), it follows that there exists some integer $R = R_k(n, m)$ such that for any m -coloring of $\{1, \dots, R\}^{(k)}$ there exists a set $S \subseteq \{1, \dots, R\}$ with $|S| = n$ such that $S^{(k)}$ is monochromatic. Note that $\{1, \dots, R\}^{(k)}$ can be identified with a complete k -uniform hypergraph on R vertices, and $S^{(k)}$ with a complete k -uniform hypergraph on n vertices. This finishes the proof. \square

1.7. Erdős-Szekeres' Theorem on convex polygons

Definition 12. A non-empty set $C \subseteq \mathbb{R}^2$ is called *convex* if for any $\vec{x}, \vec{y} \in C$ and $\lambda \in [0, 1]$ one has $\lambda\vec{x} + (1 - \lambda)\vec{y} \in C$.

The point $\lambda\vec{x} + (1 - \lambda)\vec{y}$ is usually referred to as a *convex combination* of the points \vec{x} and \vec{y} . Also observe that the set $\{\lambda\vec{x} + (1 - \lambda)\vec{y} : \lambda \in [0, 1]\}$ is just an algebraic description for the line segment joining the points \vec{x} and \vec{y} .



Figure 1.6: A convex polygon (left) and a non-convex polygon (right).

Definition 13. The *convex hull* of a non-empty set $K \subseteq \mathbb{R}^2$ is the smallest convex set that contains K .

Since the intersection of convex sets is again a convex set, it follows that the convex hull of K equals the intersection of all convex sets that contain K . The convex hull can also be described algebraically as the set of all finite convex combinations

of elements in the set. More precisely, if K is a subset of \mathbb{R}^2 and we use $\text{conv}(K)$ to denote its convex hull, then

$$\text{conv}(K) = \{w_1\vec{z}_1 + \dots + w_\ell\vec{z}_\ell : \ell \in \mathbb{N}, \vec{z}_1, \dots, \vec{z}_\ell \in K, w_1, \dots, w_\ell \in [0, 1], w_1 + \dots + w_\ell = 1\}. \quad (1.7.1)$$

Mind that the convex hull of K should not be confused with the *closed convex hull* of K , which is defined as the smallest closed convex set that contains K , and is usually denoted by $\overline{\text{conv}}(K)$ instead of $\text{conv}(K)$.

Definition 14. A non-empty set of points $K \subseteq \mathbb{R}^2$ is said to be in *convex position* if no point $\vec{x} \in K$ belongs to the convex hull of $K \setminus \{\vec{x}\}$.

For example, a finite set $K \subseteq \mathbb{R}^2$ is in convex position if and only if its elements are the corners of a convex polygon.

Definition 15. A set $K \subseteq \mathbb{R}^2$ is called *discrete* if it has no accumulation points.

Erdős-Szekeres' Theorem on points in convex position. *Let K be an infinite discrete set of points in \mathbb{R}^2 . Then either there is an infinite subset of K whose points lie on a straight line or there is an infinite subset of K whose points are in convex position.*

For the proof of Erdős-Szekeres' Theorem on points in convex position we will need the following classical result from convex geometry.

Carathéodory's theorem. *Let $K \subseteq \mathbb{R}^2$ with $|K| \geq 4$ be given. Then K is in convex position if and only if any four distinct points from K form a convex quadrilateral.*

Proof. Clearly, if K is in convex position then any quadrilateral formed using points from K is convex. To prove the converse, we will show that if K is not in convex position then there exist four points in K such that one of these points lies within the triangle spanned by the others.

Suppose K is not in convex position. Then there exists a point $\vec{x} \in K$ lying in the convex hull of $K' = K \setminus \{\vec{x}\}$. In light of (1.7.1), this means that we can write \vec{x} as

$$\vec{x} = w_1\vec{z}_1 + \dots + w_\ell\vec{z}_\ell, \quad (1.7.2)$$

where $\vec{z}_1, \dots, \vec{z}_\ell \in K'$ and $w_1, \dots, w_\ell \in [0, 1]$ with $w_1 + \dots + w_\ell = 1$. Note that we can assume without loss of generality that $\vec{z}_1, \dots, \vec{z}_\ell$ are in convex position. Indeed, if for example \vec{z}_ℓ belongs to the convex hull of $\vec{z}_1, \dots, \vec{z}_{\ell-1}$ then we can express \vec{z}_ℓ as a convex combination of $\vec{z}_1, \dots, \vec{z}_{\ell-1}$ and substitute this representation in (1.7.2), allowing us to represent \vec{x} as a convex combination of $\vec{z}_1, \dots, \vec{z}_{\ell-1}$ instead of $\vec{z}_1, \dots, \vec{z}_\ell$. Thus, invoking induction on ℓ , we may assume that $\vec{z}_1, \dots, \vec{z}_\ell$ are in convex position. This implies that $\vec{z}_1, \dots, \vec{z}_\ell$ form the corners of a convex polygon. Since \vec{x} lies inside this polygon and since convex polygons decompose into triangles (as illustrated in Figure 1.7), there exists $i < j < k \in \{1, \dots, \ell\}$ such that \vec{x} lies in the triangle spanned by $\vec{z}_i, \vec{z}_j, \vec{z}_k$, finishing the proof. \square

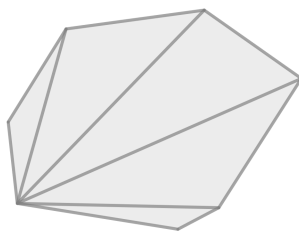


Figure 1.7: A convex polygon divided into triangles.

Proof of Erdős-Szekeres' Theorem on points in convex position. Let $K \subseteq \mathbb{R}^2$ be infinite. We begin by coloring $K^{(3)}$ by assigning the color red to $\{\vec{x}, \vec{y}, \vec{z}\} \in K^{(3)}$ if the points $\vec{x}, \vec{y}, \vec{z}$ are collinear and the color blue otherwise. According to Ramsey's Theorem for k -sets, there exists an infinite set $L \subseteq K$ such that all 3-sets in $L^{(3)}$ have the same color. If this color is red, then any three distinct points in L are collinear. This can only happen if all the points in L lie on a straight line, in which case we are done.

It remains to deal with the case when all elements in $L^{(3)}$ are blue, i.e., when no three points in L are collinear. In this situation, we need to apply Ramsey's Theorem one more time. Note that L is a discrete set. This implies that for any three points $\vec{x}, \vec{y}, \vec{z} \in L$ the triangle $\Delta\vec{x}\vec{y}\vec{z}$ contains only finitely many points from L . Color all elements in $L^{(3)}$ by assigning the color red to the 3-set $\{\vec{x}, \vec{y}, \vec{z}\} \in L^{(3)}$ if the triangle $\Delta\vec{x}\vec{y}\vec{z}$ contains an even number of points from L , and the color blue otherwise. By Ramsey's Theorem for k -sets there exists an infinite set $C \subseteq L$ such that $C^{(3)}$ is monochromatic. We claim that C is in convex position. Indeed, if C were not in convex position then, in view of Carathéodory's theorem, there exist four points $\vec{w}, \vec{x}, \vec{y}, \vec{z} \in C$ such that \vec{w} lies inside the triangle $\Delta_0 = \Delta\vec{x}\vec{y}\vec{z}$. Note that Δ_0 splits into three smaller triangles, $\Delta_1 = \Delta\vec{w}\vec{y}\vec{z}$, $\Delta_2 = \Delta\vec{w}\vec{x}\vec{z}$, and $\Delta_3 = \Delta\vec{w}\vec{x}\vec{y}$, as seen in Figure 1.8. For $i = 0, 1, 2, 3$ let $\#\Delta_i$ denote the number of points from L inside the

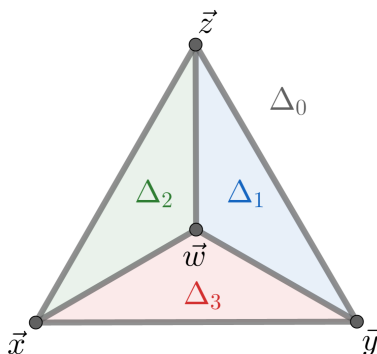


Figure 1.8

triangle Δ_i . Since no three points from L are collinear, there are no points on the boundary of any of these triangles aside from their corners. This means that the

number of points from L inside Δ_0 equals the combined number of points inside the three smaller triangles plus the point \vec{w} , or in other words,

$$\#\Delta_0 = \#\Delta_1 + \#\Delta_2 + \#\Delta_3 + 1. \quad (1.7.3)$$

Recall that $C^{(3)}$ is monochromatic. If all elements in $C^{(3)}$ are red then the quantities $\#\Delta_0$, $\#\Delta_1$, $\#\Delta_2$, and $\#\Delta_3$ are even numbers. This would imply that the left hand side of (1.7.3) is an even number whereas the right hand side is an odd number, a contradiction. Similarly, if all elements in $C^{(3)}$ are blue then $\#\Delta_0$, $\#\Delta_1$, $\#\Delta_2$, $\#\Delta_3$ are odd numbers, implying that the left hand side of (1.7.3) is odd whereas the right hand side is even. Either way, we have obtained a contradiction, which means that C is in convex position. \square

The following is a big open conjecture at the interface of convex geometry and Ramsey theory, posed by Erdős and Szekeres in 1960.

Conjecture (Erdős-Szekeres convex polygon problem). *Let $n \geq 3$. Any set of $2^{n-2} + 1$ points in the plane, no three of which are collinear, contains a subset of n points in convex position.*

1.8. Erdős-Szekeres' Theorem on monotone paths

Erdős-Szekeres' Theorem on monotone paths. *Fix $n, m \in \mathbb{N}$. Any sequence of distinct real numbers of length at least $nm + 1$ admits either a monotonically increasing subsequence of length $n + 1$ or a monotonically decreasing subsequence of length $m + 1$.*

Proof. Let $x_1, x_2, \dots, x_{nm+1}$ be a sequence of distinct real numbers. Label each element x_i in the sequence with the pair (a_i, b_i) , where a_i is the length of the longest monotonically increasing subsequence ending with x_i and b_i is the length of the longest monotonically decreasing subsequence ending with x_i . Note that any two elements in the sequence are labeled with a different pair: if $i < j$ and $x_i < x_j$ then $a_i < a_j$, and on the other hand if $x_i > x_j$ then $b_i < b_j$. If $a_i \leq n$ and $b_i \leq m$ for all i then there are only nm possible labels, contradicting the fact that there are $nm + 1$ elements in the sequence each with a unique label. It follows that either $a_i > n$ or $b_i > m$ for some i , yielding either an increasing sequence of length at least $n + 1$ or a decreasing sequence of length at least $m + 1$. \square

Bibliography

- [Ram30] F. P. RAMSEY, On a Problem of Formal Logic, *Proceedings of the London Mathematical Society* **s2-30** (1930), 264–286. <http://dx.doi.org/10.1112/plms/s2-30.1.264>.
- [Sch17] I. SCHUR, Über die Kongruenz $x^m + y^m \equiv z^m \pmod{p}$, *Jahresbericht der Deutschen Mathematiker-Vereinigung* **25** (1917), 114–116. Available at <http://eudml.org/doc/145475>.