1 Relationship between Ramsey theory and extremal theory

Consider the following theorem, which falls within the framework of Ramsey theory.

**Theorem 1** (Van der Waerden, 1927). For any \( k \geq 2, \ell \geq 3 \), there is \( n \) such that any \( k \)-coloring of \( [n] \) contains a monochromatic arithmetic progression of length \( \ell \): \( \{a, a+b, a+2b, \ldots, a+(\ell-1)b\} \).

This is a classical theorem which predates even Ramsey’s theorem about graphs. We are not going to present the proof here. Note, however, that in order to prove such a statement, it would be enough to show that for a sufficiently large \( [n] \), any subset of at least \( n/k \) elements contains an arithmetic progression of length \( \ell \). This is indeed what Szemerédi proved, much later and using much more involved techniques.

**Theorem 2** (Szemerédi). For any \( \delta > 0 \) and \( \ell \geq 3 \), there is \( n_0 \) such that for any \( n \geq n_0 \) and any set \( S \subseteq [n], |S| \geq \delta n \), \( S \) contains an arithmetic progression of length \( \ell \).

It can be seen that this implies Van der Waerden’s theorem, since we can set \( \delta = 1/k \) and for any \( k \)-coloring of \( [n] \), one color class contains at least \( \delta n \) elements. Szemerédi’s theorem is an extremal type of statement - stating, that any object of sufficient size must contain a certain structure.

2 Bipartite graphs

**Definition 1.** A graph \( G \) is called bipartite, if the vertices can be partitioned into \( V_1 \) and \( V_2 \), so that there are no edges inside \( V_1 \) and no edges inside \( V_2 \).

Equivalently, \( G \) is bipartite if its vertices can be colored with 2 colors so that the endpoints of every edge get two different colors. (The 2 colors correspond to \( V_1 \) and \( V_2 \).) Thus, bipartite graphs are called equivalently 2-colorable.

We also have the following characterization, which is useful to know.

**Lemma 1.** \( G \) is bipartite, if and only if it does not contain any cycle of odd length.

*Proof.* Suppose \( G \) has an odd cycle. Then obviously it cannot be bipartite, because no odd cycle is 2-colorable.

Conversely, suppose \( G \) has no odd cycle. Then we can color the vertices greedily by 2 colors, always choosing a different color for a neighbor of some vertex which has been colored already. Any additional edges are consistent with our coloring, otherwise they would close a cycle of odd length with the edges we considered already.

The easiest extremal question is about the maximum possible number of edges in a bipartite graph on \( n \) vertices.
Lemma 2. A bipartite graph on \( n \) vertices can have at most \( \frac{1}{4}n^2 \) edges.

Proof. Suppose the bipartition is \((V_1, V_2)\) and \( |V_1| = k \), \( |V_2| = n - k \). The number of edges between \( V_1 \) and \( V_2 \) can be at most \( k(n - k) \), which is maximized for \( k = n/2 \). \(\square\)

3 Graphs without a triangle

Let us consider Ramsey’s theorem for graphs, which guarantees the existence of a monochromatic triangle for an arbitrary coloring of the edges. An analogous extremal question is, what is the largest number of edges in a graph that does not have any triangle? We remark that this is not the right way to prove Ramsey’s theorem - even for triangles, it is not true that for any 2-coloring of a large complete graph, the larger color class must contain a triangle.

Exercise: what is a counterexample?

The question how many edges are necessary to force a graph to contain a triangle is very old and it was resolved by the following theorem.

Theorem 3 (Mantel, 1907). For any graph \( G \) with \( n \) vertices and more than \( \frac{1}{4}n^2 \) edges, \( G \) contains a triangle.

Proof. Assume that \( G \) has \( n \) vertices, \( m \) edges and no triangle. Let \( d_x \) denote the degree of \( x \in V \). Whenever \((x, y) \in E\), we know that \( x \) and \( y \) cannot share a neighbor (which would form a triangle), and therefore \( d_x + d_y \leq n \). Summing up over all edges, we get

\[
mn \geq \sum_{(x, y) \in E} (d_x + d_y) = \sum_{x \in V} d_x^2.
\]

On the other hand, applying Cauchy-Schwartz to the vectors \((d_1, d_2, \ldots, d_n)\) and \((1, 1, \ldots, 1)\), we obtain

\[
n \sum_{x \in V} d_x^2 \geq \left( \sum_{x \in V} d_x \right)^2 = (2m)^2.
\]

Combining these two inequalities, we conclude that \( m \leq \frac{1}{4}n^2 \). \(\square\)

We remark that the analysis above can be tight only if for every edge, any other vertex is connected to exactly one of the two endpoints. This defines a partition \( V_1 \cup V_2 \) such that we have all edges between \( V_1 \) and \( V_2 \), i.e. a complete bipartite graph. When \( |V_1| = |V_2| \), this is the unique extremal graph without a triangle, containing \( \frac{1}{4}n^2 \) edges.

4 Graphs without a clique \( K_{t+1} \)

More generally, it is interesting to ask how many edges \( G \) can have if \( G \) does not contain any clique \( K_{t+1} \). An example of a graph without \( K_{t+1} \) can be constructed by taking \( t \) disjoint sets of vertices, \( V = V_1 \cup \ldots \cup V_t \), and inserting all edges between vertices in different sets. Now, obviously there is no \( K_{t+1} \), since any set of \( t + 1 \) vertices has two vertices in the same set \( V_i \). The number of edges in such a graph is maximized, when the sets \( V_i \) are as evenly sized as possible, i.e. \( |V_i| - |V_j| \in \{-1, 0, +1\} \) for all \( i, j \). We call such a graph on \( n \) vertices the Turán graph \( T_{n,t} \). Turán proved in 1941 that this is indeed the graph without \( K_{t+1} \) containing the maximum number of edges. Note that the number of edges in \( T_{n,t} \) is \( \frac{1}{2}(1 - \frac{1}{t})n^2 \), assuming for simplicity that \( n \) is divisible by \( t \).
Theorem 4 (Turán, 1941). Among all $K_{t+1}$-free graphs on $n$ vertices, $T_{n,t}$ has the most edges.

Proof. Let $G$ be a graph without $K_{t+1}$ and $v_m$ a vertex of maximum degree $d_m$. Let $S$ be the set of neighbors of $v_m$, $|S| = d_m$, and $T = V \setminus S$. Note that by assumption, $S$ has no clique of size $t$.

We modify the graph into $G'$ as follows: we keep the graph inside $S$, we include all possible edges between $S$ and $T$, and we remove all edges inside $T$. For each vertex, the degree can only increase: for vertices in $S$, this is obvious, and for vertices in $T$, the new degrees are at least $d_m$, i.e. at least as large as any degree in $G$. Thus the total number of edges can only increase.

By induction, we can prove that $G[S]$ can be also modified into a union of $t-1$ disjoint independent sets with all edges between them. Therefore, the best possible graph has the structure of a Turán graph.

To prove that the Turán graph is the unique extremal graph, we note that if $G$ had any edges inside $T$, then we strictly gain by modifying the graph into $G'$.

We present another proof of Turán’s theorem, which is probabilistic. Here, we only prove the quantitative part, that $\frac{1}{2}(1 - \frac{1}{t})n^2$ is the maximum number of edges in a graph without $K_{t+1}$.

Proof. Let’s consider a probability distribution on the vertices, $p_1, \ldots, p_n$ such that $\sum_{i=1}^n p_i = 1$. We start with $p_i = 1/n$ for all vertices. Suppose we sample two vertices $v_1, v_2$ independently according to this distribution - what is the probability that $\{v_1, v_2\} \in E$? We can write this probability as

$$\Pr[\{v_1, v_2\} \in E] = \sum_{i,j: \{i,j\} \in E} p_ip_j.$$  

At the beginning, this is equal to $\frac{2}{n^2} |E|$.

Now we modify the distribution in order to make $\Pr[\{v_1, v_2\} \in E]$ as large as possible. We claim that the probability distribution that maximizes this probability is uniform on some maximal clique. We proceed as follows: If there are two non-adjacent vertices $i, j$ such that $p_i, p_j > 0$, let $s_i = \sum_{k: \{i,k\} \in E} p_k$ and $s_j = \sum_{k: \{j,k\} \in E} p_k$. If $s_i \geq s_j$, we set the probability of vertex $i$ to $p_i + p_j$ and the probability of vertex $j$ to 0 (and conversely if $s_i < s_j$). It can be verified that this increases $\Pr[\{v_1, v_2\} \in E]$ by $p_j(s_i - s_j)$ or $p_i(s_j - s_i)$, respectively.

Eventually, we reach a situation where there are no two non-adjacent vertices of positive probability, i.e. the distribution is on a clique $Q$. Then, $\Pr[\{v_1, v_2\} \in E] = \Pr[v_1 \neq v_2] = 1 - \sum_{i \in Q} p_i^2$.

By Cauchy-Schwartz, this is maximized when $p_i$ is uniform on $Q$, i.e.

$$\Pr[\{v_1, v_2\} \in E] \leq 1 - \frac{1}{|Q|} \leq 1 - \frac{1}{t}$$

assuming that there is no clique larger than $t$. Recall that the probability we started with was $\frac{2}{n^2} |E|$ and we never decreased it in the process. Therefore,

$$|E| \leq \left(1 - \frac{1}{t}\right) \frac{n^2}{2}.$$  

}$\square$