Topics in Combinatorics

Dorottya Sziráki
Gergő Nemes

C E U

Preface

These notes are based on the lectures of the course titled "Topics in Combinatorics", which were given by Ervin Győri in the winter trimester of 2013 as part of Central European University's mathematics Ph.D. program. The aim of these notes is to present the material of this course in a clear and easily understandable manner, and to help students in the preparation for the exam. We would like to thank Mohamed Khaled for his help during the preparation of these notes. We welcome any comments, suggestions and corrections.

Dorottya Sziráki & Gergő Nemes

Contents

Chapter 1

Colouring of graphs

We recall that a (simple and undirected) *graph* is an ordered pair $G = (V, E)$ comprising a set $V = V(G)$ of *vertices* together with a set $E = E(G)$ of *edges*, which are 2-element subsets of V. For an edge $\{x, y\}$, we shall use the somewhat shorter notation xy. The *order* of a graph is $|V|$ (the number of vertices). A graph's *size* is $|E|$, the number of edges. The *degree* of a vertex v, denoted by $d(v)$, is the number of edges that connect to it.

1.1 The chromatic number

Definition 1.1.1. *Let G be a graph. The map $c : V(G) \to \mathbb{Z}^+$ is called a **(proper) colouring of** G if $c(x) \neq c(y)$ for all $xy \in E(G)$. The positive integer*

$$\chi(G) \overset{\text{def}}{=} \min\{k \mid G \text{ has a colouring with colours } 1, 2, \ldots, k\}$$

*is called the **chromatic number of** G.*

A *planar graph* is a graph that can be drawn in such a way that no edges cross each other. The origin of graph colouring was the following conjecture (now a theorem):

Conjecture 1.1.1. *Every planar graph G is 4-colourable, i.e., $\chi(G) \leq 4$.*

This statement is equivalent to the *Four Colour Map Theorem.*

Definition 1.1.2. *A graph G is called k-**critical** if $\chi(G) = k$ but $\chi(G - \{e\}) = k - 1$ for all $e \in E(G)$.*

A *path* in a graph is a sequence of edges which connect a sequence of vertices. A path may be infinite, but a finite path always has a first vertex, called its *start vertex*, and a last vertex, called its *end vertex*. A *cycle* is a path such that the start vertex and end vertex are the same. The choice of the start vertex in a cycle is arbitrary.

Theorem 1.1.2. *The graph G is 2-colourable (**bipartite**) if and only if it does not contain any odd cycle.*

Proof. $\boxed{\Rightarrow}$ Suppose that G is 2-colourable with colouring $c : V(G) \to \{1, 2\}$. Let $\mathcal{C} \stackrel{\text{def}}{=} v_0v_1v_2 \dots v_kv_0$ be a cycle in the graph. Without loss of generality, we may assume that $c(v_0) = 1$. Since c is a colouring, we must have $c(v_1) = 2$. In general, we must have $c(v_{2i}) = 1$ and $c(v_{2i+1}) = 2$ $(i = 0, 1, 2, \dots, \lfloor \frac{k-1}{2} \rfloor)$. Hence, k must be odd and $\mathcal{C}$ is an even cycle.

$\boxed{\Leftarrow}$ Without loss of generality, we may assume that G is connected. Suppose that G contains no odd cycles. Denote by $d(x, y)$ the length of the shortest path between two vertices x and y (the length of a path is the number of edges which are traversed). Let v_0 be an arbitrary vertex of G. Define

$$V_1 \stackrel{\text{def}}{=} \{u \in V(G) \mid d(u, v_0) \text{ is even}\},$$

$$V_2 \stackrel{\text{def}}{=} \{u \in V(G) \mid d(u, v_0) \text{ is odd}\}.$$

Define $c : V(G) \to \{1, 2\}$ as

$$c(u) \stackrel{\text{def}}{=} \begin{cases} 1, & \text{if } u \in V_1; \\ 2, & \text{if } u \in V_2. \end{cases}$$

We show that c is a colouring of G. Suppose to the contrary that $xy \in E(G)$ and $c(x) = c(y)$. Let P be the shortest path from x to v_0 and Q be the shortest path from y to v_0. Let w be the first common vertex of P and Q. Clearly, P and Q defines two paths from w to v_0. Since P and Q were shortest paths, these two paths from w to v_0 are shortest paths from w to v_0. In particular, they are of the same length. Consider now the following parts of P and Q: the path P_1 from x to w and the path Q_1 from y to w (see Figure 1.1).

Since $c(x) = c(y)$, the paths P and Q are of the same length. Thus, P_1 and Q_1 must have the same length (since the remaining parts from w to v_0 had the same length). Then $\mathcal{C} \stackrel{\text{def}}{=} (P_1 - \{w\})Q_1^{-1}x$ is an odd cycle, which is a contradiction. ■

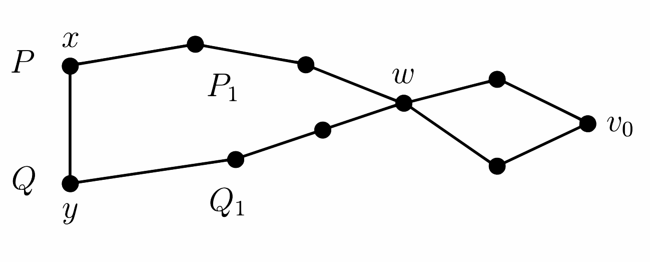

Figure 1.1

Corollary 1.1.3. *The* 3*-critical graphs are the odd cycles.*

Proof. First we show that odd cycles are 3-critical. Clearly, odd cycles have chromatic number three. Removing any edge from a cycle results in a tree (which is 2-colourable). Therefore, odd cycles are 3-critical. By the previous theorem, a graph with no odd cycles is 2-colourable, therefore any graph with chromatic number three contains an odd cycle. In general, however, we know that a k-critical graph G cannot contain another k-critical graph G' as a proper subgraph. (Otherwise, by removing an edge from G which is not in G', the remaining graph would still have chromatic number k, since it contains G'. Hence, G could not be k-critical.) Thus, only odd cycles are 3-critical. ■

Theorem 1.1.4. *Let $\Delta(G)$ be the maximum degree in G. Then $\chi(G) \leq \Delta(G)+1$.*

Proof. By greedy algorithm. We consider the vertices of the graph in sequence and assigns each vertex its first available colour. In the worst case all the $\leq \Delta(G)$ neighbours of a vertex v are coloured with different colours. ■

This upper bound is sharp, as the following examples show.

Example 1.1.1. *The complete graph K_n can be trivially coloured by n colours, giving each vertex a different colour. By the pigeonhole principle, any colouring of K_n with fewer colours would have two vertices with the same colour. But, in K_n, any two vertices are joined by an edge; in particular those vertices of the same colour would be joined by an edge. Thus, the chromatic number of K_n is n. Since $\Delta(K_n) = n - 1$, we have $\chi(K_n) = \Delta(K_n) + 1$.*

Example 1.1.2. *For an odd cycle $\mathcal{C}$ we have $\chi(\mathcal{C}) = 3$. Clearly, any vertex of a cycle has degree 2. Thus $\Delta(\mathcal{C}) = 2$, and $\chi(\mathcal{C}) = \Delta(\mathcal{C}) + 1$.*

Theorem 1.1.5 (Brooks)**.** *For a connected, simple graph G, we have $\chi(G) \leq \Delta(G)$; unless G is isomorphic to a complete graph or an odd cycle.* ■

Remark. *If G' is a subgraph of G then $\chi(G') \leq \chi(G)$.*

1.2 Triangle-free graphs with large chromatic numbers

We shall give constructions for sequence of graphs $G_3, G_4, G_5, \ldots$ such that each G_k is triangle-free and $\chi(G_k) = k$.

(i) Zykov's construction. Let G_3 be the 5-cycle, which has chromatic number 3 and is triangle-free. Suppose that we already have $G_3, G_4, \ldots, G_k$ ($k \geq 3$) such that $\chi(G_i) = i$ and G_i contains no triangles ($i = 1, 2, \ldots, k$). We define G_{k+1} using G_k. Let n_k be the order of G_k.

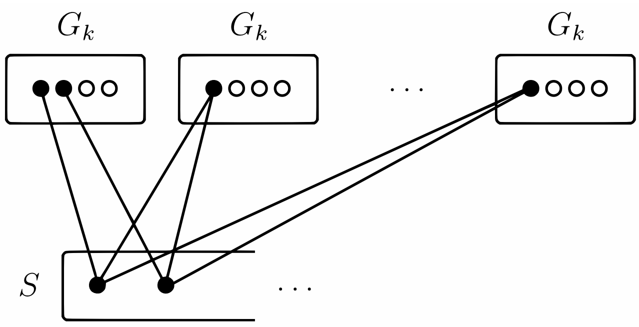

Figure 1.2. *Zykov's construction of G_{k+1}.*

Consider k copies of G_k. From each copy, we choose a vertex and connect them with a new vertex. Thus, this newly added vertex has k adjacent vertices, one from each copy of G_k. Since there are n_k^k possible choices for the vertices, we have that many newly added vertices (denote by S the set they form). This is the graph G_{k+1} (see Figure 1.2). Hence, G_{k+1} has $k \cdot n_k + n_k^k$ vertices.

First, we claim that G_{k+1} contains no triangles. By the induction hypothesis none of the G_k's contains triangles. The copies of the G_k's are connected via the vertices from S. Hence, if G_{k+1} contains a triangle, then at least one of its vertices should come from S. But there are no edges between the vertices of S, nor between the copies of the G_k's, so there is no way to have a triangle.

Second, we show that $\chi(G_{k+1}) = k+1$. Each copy of the G_k's can be coloured by $\chi(G_k) = k$ colours. Since there are no edges between the vertices of S, we can introduce a new colour for them. Thus, we have a colouring with $k+1$ colours: $\chi(G_{k+1}) \leq k+1$. Now, from each copy of G_k, choose a vertex such that all the k vertices have different colours (it is possible, since we have k different colours in each copy of G_k). The vertex from S that corresponds to those k vertices must have a new colour. Therefore, $\chi(G_{k+1}) = k+1$.

The order n_k of the graphs G_k grows quite rapidly as $k \to +\infty$.[1] The first few values are given in Table 1.1. ♣

k	$\lvert V(G_k)\rvert = n_k$
3	5
4	$3 \cdot 5 + 5^3 = 140$
5	$4 \cdot 140 + 140^4 = 384160560$
6	$5 \cdot 384160560 + 384160560^5 \approx 8.366886525 \cdot 10^{42}$

Table 1.1. *The order of the Zykov graphs G_k grows extremely fast.*

(ii) Mycielski's construction. Let G_3 be the 5-cycle, which has chromatic number 3 and is triangle-free. Suppose that we already have $G_3, G_4, \ldots, G_k$ $(k \geq 3)$ such that $\chi(G_i) = i$ and G_i contains no triangles $(i = 1, 2, \ldots, k)$. We define G_{k+1} using G_k.

Let $v_1, v_2, \ldots, v_{n_k}$ be the vertices of G_k. We add the new vertices $u_1, u_2, \ldots, u_{n_k}$ and w. We connect each vertex u_i by an edge to w. In addition, for each edge v_iv_j of G_k, we add two new edges, u_iv_j and v_iu_j. This is the graph G_{k+1} (see Figure 1.3). Hence, G_{k+1} has $2n_k + 1$ vertices.

First, we claim that G_{k+1} contains no triangles. By the induction hypothesis, none of the G_k's contains triangles. Hence, if G_{k+1} contains a triangle, then at least one of its vertices should come from the newly added vertices. Since there are no edges between the u_i's, nor between w and the v_i's, such a triangle must contain a vertex u_i and two vertices from the v_j's. The vertex u_i is adjacent to the

[1]Actually, I proved that $n_k \sim \alpha^{(k-1)!}$ as $k \to +\infty$, where $\alpha = \sqrt{5}\prod_{m=3}^{\infty}\left(1 + mn_m^{1-m}\right)^{1/m!} = 2.27870461723763\ldots$.

vertices that the v_i are adjacent to, but this would imply that G_k has a triangle, where the vertices that u_i forms a triangle with, are vertices of this triangle along with v_i. This would contradict that G_k is triangle-free.

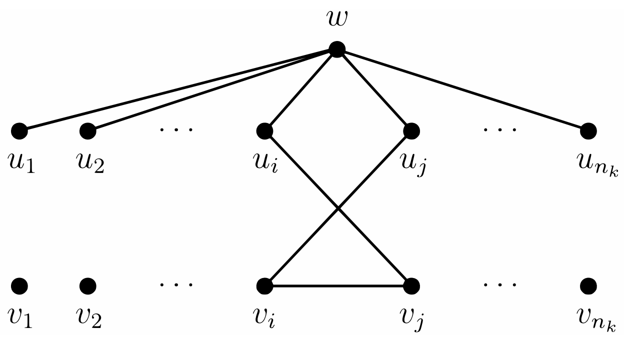

Figure 1.3. *Mycielski's construction of* G_{k+1}.

Second, we show that $\chi(G_{k+1}) = k+1$. Suppose, to the contrary, that $\chi(G_{k+1}) \leq k$, and consider a colouring c of G_{k+1} with k colours. Without loss of generality, we may assume that $c(w) = k$. This forces the u_i's not to have the colour k. If for some $1 \leq i \leq n_k$, $c(v_i) = k$, then we can recolour that v_i to the colour of the corresponding u_i. Hence, we get a proper colouring of G_k with $k-1$ colours. Indeed, if for some j, $v_iv_j \in E(G_k)$, and v_j would have the same colour as v_i; then its adjacent vertex u_i (by construction $u_iv_j \in E(G_{k+1})$) would have the same colour. This would be a contradiction, since we started with a proper colouring. Therefore, $\chi(G_k) \leq k-1$, which is a contradiction.

k	$\lvert V(G_k)\rvert = n_k$	k	$\lvert V(G_k)\rvert = n_k$
3	5	8	$2 \cdot 95 + 1 = 191$
4	$2 \cdot 5 + 1 = 11$	9	$2 \cdot 191 + 1 = 383$
5	$2 \cdot 11 + 1 = 23$	10	$2 \cdot 383 + 1 = 767$
6	$2 \cdot 23 + 1 = 47$	11	$2 \cdot 767 + 1 = 1535$
7	$2 \cdot 47 + 1 = 95$	12	$2 \cdot 1535 + 1 = 3071$

Table 1.2. *The order of the Mycielski graphs* G_k.

The order n_k of the graphs G_k grows much slower than in the previous construction. Indeed, it is easy to show that

$$n_k = 3 \cdot 2^{k-2} - 1 \quad \text{for} \quad k \geq 3.$$

The first few values are given in Table 1.2. ♣

Exercise 1.2.1. *Show that the graph G_k in Mycielski's construction is k-critical. (Hint: use the lemma below.)*

Lemma 1.2.1. *If G is k-critical, then for all $x \in V(G)$, G has a k-colouring such that x is the only one vertex of colour k.* ■

Exercise 1.2.2. *Let G_1 be a single vertex, G_2 be two vertices connected by an edge. Modify Zykov's construction in such a way that instead of considering k copies of G_k, take one copy of each G_i $(1 \leq i \leq k)$. The rest of the construction is the same as in the original one. Prove that G_{k+1} is $k+1$-critical.*

Theorem 1.2.2. *Suppose that G is triangle free, with order at most* 10*. Then $\chi(G) \leq 3$.* ■

The graph G_4 in Mycielski's construction shows that this bound is sharp.

Theorem 1.2.3. *Suppose that G is triangle free, with order at most* 22*. Then $\chi(G) \leq 4$.* ■

The graph G_5 in Mycielski's construction shows that this bound is sharp.

(iii) The Erdős–Hajnal Shift-graph. Here we construct a graph G which is triangle-free and $\chi(G) = k+1$. Unlike the previous constructions, this construction does not use recursion.

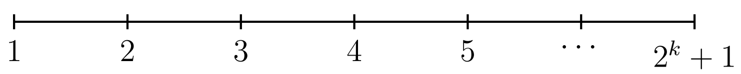

Figure 1.4

The vertices of the graph are (non-degenerate) closed intervals with integer endpoints (cf. Figure 1.4):

$$V(G) = \left\{ [i, j] \mid 1 \leq i < j \leq 2^k + 1 \right\}.$$

Now, we define the edges. Let $[i, j]$ and $[\ell, m]$ be adjacent if either $j = \ell$ or $m = i$. A typical situation is shown in Figure 1.5.

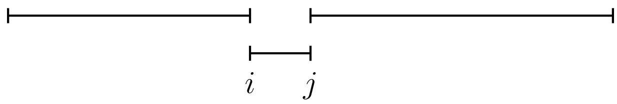

Figure 1.5. *The vertex $[i, j]$ with two adjacent vertices.*

From the definition, it is clear that G does not contain triangles.

Now, we show that $\chi(G) = k + 1$. We divide the interval $\left[1, 2^k + 1\right]$ into two parts: $\left[1, 2^{k-1}\right]$ and $\left[2^{k-1} + 1, 2^k + 1\right]$. They contain 2^{k-1} and $2^{k-1} + 1$ integer points, respectively. By an induction argument, those intervals (vertices) that belong strictly to one of the two halves, can be coloured by k colours. Consider now those intervals that start at the first and end at the second half. By the definition of the edges, these kind of intervals (vertices) are not adjacent. Hence, we can use the colour $k + 1$ for them. Thus, we have showed that $\chi(G) \le k + 1$. Suppose, to the contrary, that $\chi(G) = k$. Denote by $f(i)$ the set of colours of those intervals that begin with i. We claim that $f(i) \ne f(j)$ if $i \ne j$. Indeed, without loss of generality, we may assume that $i < j$ and consider the colour s of $[i, j]$. By definition, $s \in f(i)$. But $s \notin f(j)$, otherwise there would exist an interval $[j, \ell]$ with colour s. This is impossible, since by definition, $[i, j]$ and $[j, \ell]$ are adjacent. Therefore, $f(1), f(2), \ldots, f(2^k + 1)$ represent $2^k + 1$ different subsets of the k-element set of colours. This is a contradiction, since a set with k elements has only 2^k different subsets. ♣

1.3 Cycle-free graphs with large chromatic numbers

Let us denote by $\mathcal{C}_k$ the cycle consisting of k vertices.

Theorem 1.3.1 (Tutte)**.** *For any $k \ge 3$, there exists a graph G_k such that $\chi(G_k) = k$ and G_k does not contain $\mathcal{C}_3$, $\mathcal{C}_4$ or $\mathcal{C}_5$ as proper subgraphs.*

Intuitively, in such a graph, the radius 2 neighbourhood of any vertex looks like a tree.

Proof. Let G_3 be the 7-cycle, which has chromatic number 3 and does not contain $\mathcal{C}_3$, $\mathcal{C}_4$ or $\mathcal{C}_5$. Suppose that we already have $G_3, G_4, \ldots, G_k$ ($k \geq 3$) such that $\chi(G_i) = i$ and G_i does not contain $\mathcal{C}_3$, $\mathcal{C}_4$ or $\mathcal{C}_5$ ($i = 1, 2, \ldots, k$). We define G_{k+1} using G_k. Let n_k be the order of G_k.

First, we take $m \stackrel{\text{def}}{=} (n_k - 1)k + 1$ vertices. Call the set of these vertices as S. For each n_k-tuple from S we add a copy of G_k, and match the vertices of the n_k-tuple with the n_k vertices of the corresponding G_k. (Each vertex from the tuple should have exactly one pair from the vertices of G_k.) This is the graph G_{k+1}, which has $m + n_k\binom{m}{n_k}$ vertices (see Figure 1.6).

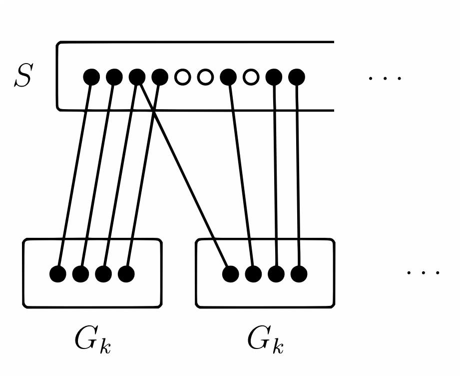

Figure 1.6. *The construction of the graph G_{k+1}.*

First, we prove that $\chi(G_{k+1}) \geq k + 1$. Suppose, to the contrary, that c is a colouring of G_{k+1} with k colours. By the choice of $m = |S|$, there exist n_k vertices in S with the same colour, for example with colour k. Take that copy of G_k which corresponds to these n_k vertices. Since c is a proper colouring, we see that the vertices of this copy of G_k avoid the colour k. Hence, we have a colouring of G_k with $k - 1$ colours. This is a contradiction, since by the induction hypothesis, $\chi(G_k) = k$.

Second, we prove that $\chi(G_{k+1}) \leq k + 1$. Using the induction hypothesis and the fact that there are no edges between the copies of the G_k's, we can colour each G_k with k colours. Since there are no edges between the vertices in S, we can colour them with colour $k + 1$. In this way, we have got a colouring of G_{k+1} with $k + 1$ colours.

Finally, we show that there is no $\mathcal{C}_3$, $\mathcal{C}_4$ or $\mathcal{C}_5$ in G_{k+1}. By the induction hypothesis, G_k does not contain $\mathcal{C}_3$, $\mathcal{C}_4$ or $\mathcal{C}_5$. Moreover, there are no edges between the copies of the G_k's, nor between the vertices in S. This shows immediately that

G_{k+1} cannot have a triangle ($\mathcal{C}_3$). If G_{k+1} would contain a $\mathcal{C}_4$ (or $\mathcal{C}_5$), then exactly two vertices from the cycle would come from S (this is the only possible way). Since a vertex from S is connected to a copy of G_k by only one edge, the other 2 (resp. 3) vertices of the cycle should come from two copies of the G_k's (see Figure 1.7). (It is clear that they can not come from more than two copies, because of the short length of the cycle.)

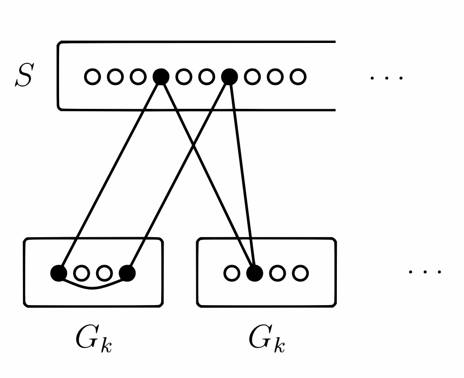

Figure 1.7. *A possible $\mathcal{C}_5$ in G_{k+1}.*

Clearly, one of the G_k's must contain only one vertex from $\mathcal{C}_4$ (or $\mathcal{C}_5$). But then, this vertex is connected to two different vertices from S, which is impossible. Therefore, there is no $\mathcal{C}_4$ or $\mathcal{C}_5$ in G_{k+1}. ∎

Definition 1.3.1. *An **independent set** is a set of vertices in a graph, no two of which are adjacent. A **maximum independent set** is a largest independent set for a given graph G and its size is denoted $\alpha(G)$.*

Theorem 1.3.2 (Erdős)**.** *For every $g \geq 3$ and $k \geq 3$, there is a graph $G_{g,k}$ such that $\chi(G_{g,k}) \geq k$ and $\operatorname{girth}(G_{g,k}) \geq g$. Here $\operatorname{girth}(G_{g,k})$ denotes the length of the shortest cycle in the graph $G_{g,k}$.*

Proof. By probabilistic methods. Let $n \geq 5$ be an integer large enough so that $2g\sqrt{n} \leq \frac{n}{2}$ and $\frac{n^{1/2g}}{6\log n+2} \geq k$ and set $p = n^{-1+1/2g}$. Form a random graph on n vertices by choosing each possible edge to occur independently with probability p. Let G be the resulting graph, and let X be the number of cycles of G with length less than g. The number of cycles of length i in the complete graph on n vertices is $\frac{(i-1)!}{2}\binom{n}{i} = \frac{n!}{2i(n-i)!}$ and each of them is present in G with probability p^i. Hence,

$$\begin{aligned}\mathsf{E}(X) &= \sum_{i=3}^{g-1} \frac{n!}{2i(n-i)!} p^i = \sum_{i=3}^{g-1} \frac{n(n-1)\cdots(n-i+1)}{2i} p^i \le \sum_{i=3}^{g-1} \frac{n^i}{2i} p^i \le \sum_{i=3}^{g-1} n^i p^i \\ &= \sum_{i=3}^{g-1} \left(n^{1/2g}\right)^i \le \sum_{i=3}^{g-1} \left(n^{1/2g}\right)^g \le g\sqrt{n}.\end{aligned}$$

So, by Markov's inequality, we have

$$\mathsf{P}\left(X \ge \frac{n}{2}\right) \overset{2g\sqrt{n}\le n/2}{\le} \mathsf{P}\left(X \ge 2g\sqrt{n}\right) \overset{\substack{\text{Markov's}\\ \text{inequality}}}{\le} \frac{1}{2g\sqrt{n}}\mathsf{E}(X) \le \frac{1}{2}.$$

Thus, $X \le \frac{n}{2}$ with probability $\ge \frac{1}{2}$. By Taylor's formula, we have

$$e^{-p} = 1 - p + \frac{e^{-\xi}}{2}p^2 > 1 - p,$$

for some $0 < p < \xi < 1$. Set $s = \left\lceil \frac{3}{p}\log n \right\rceil$. Then we find

$$\begin{aligned}\mathsf{P}(\alpha(G) \ge s) &\le \binom{n}{s}(1-p)^{\binom{s}{2}} = \frac{n(n-1)\cdots(n-s+1)}{s!}(1-p)^{\frac{s(s-1)}{2}} \\ &\le \frac{n^s}{s!}(1-p)^{\frac{s(s-1)}{2}} \le n^s(1-p)^{\frac{s(s-1)}{2}} \le n^s e^{-p\frac{s(s-1)}{2}} = \left(ne^{-p(s-1)/2}\right)^s \\ &= \left(e^{\log n - p(s-1)/2}\right)^s \le \left(e^{\log n - 3/2\log n}\right)^s = n^{-s/2} \le n^{-1/2} < \frac{1}{2}.\end{aligned}$$

Since $\mathsf{P}(X \le \frac{n}{2}) \ge \frac{1}{2}$ and $\mathsf{P}(\alpha(G) < s) > \frac{1}{2}$, there is a specific G with n vertices for which $X \le \frac{n}{2}$ and $\alpha(G) < s$. Form the graph H from this G by removing from G a vertex from each cycle of length less than g. Then H has no cycles of length less than g, $|V(H)| \ge \frac{n}{2}$, and $\alpha(H) \le \alpha(G) \le s$, so we have

$$\begin{aligned}\chi(H) &\ge \frac{|V(H)|}{\alpha(H)} \ge \frac{n}{2s} \ge \frac{n}{(6/p)\log n + 2} \ge \frac{n}{6n^{1-1/2g}\log n + 2} \\ &= \frac{n^{1/2g}}{6\log n + 2n^{1/2g-1}} \ge \frac{n^{1/2g}}{6\log n + 2} \ge k.\end{aligned}$$

Hence, $G_{g,k} \overset{\text{def}}{=} H$ satisfies the theorem. ∎

1.4 The Art Gallery Theorem

The floor plan of an art gallery modeled as an n-gon (a simple polygon with n vertices). How many watchmen needed to see the whole room? Each watchman is stationed at a fixed point, has 360° vision, and cannot see through the walls. More precisely, let us denote by P the interior of the n-gon. A point $p \in P$ is *visible* from the point $q \in P$, if and only if the line segment pq is completely contained in P (see Figure 1.8).

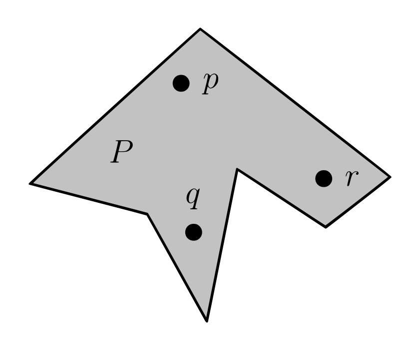

Figure 1.8. *The p is visible from q and r. But q and r are not visible to each other.*

Denote by $w(P)$ the minimum number of points in P such that any point of P is visible from at least one of those points.

Theorem 1.4.1 (The Art Gallery Theorem)**.** *We have $w(P) \leq \left\lfloor \frac{n}{3} \right\rfloor$, i.e., $\left\lfloor \frac{n}{3} \right\rfloor$ watchmen are sufficient to control the interior of an n-gon. Moreover, this bound is sharp as it is shown in Figure 1.9.*

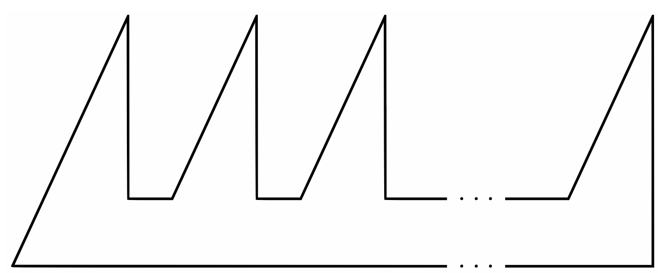

Figure 1.9. *A necessity construction.*

Proof. It is known that every n-gon can be triangulated with pairwise non-intersecting diagonals, and every such triangulation has exactly $n-2$ triangles. The triangulation of our n-gon leads to a (planar) graph G (see Figure 1.10).

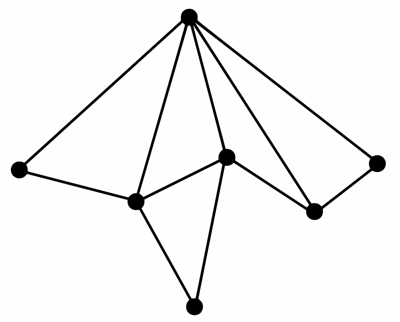

Figure 1.10. *Triangulation of a 7-gon.*

We claim that $\chi(G) = 3$. Since G contains triangles, $\chi(G) \geq 3$. Suppose that $n \geq 4$. (If $n = 3$, this is obvious.) First, we show that there exist at least two non-adjacent vertices with degree 2. We prove this by induction on n. For a triangulated 4-gon, the statement is naturally true. Suppose that the statement is true for every k-gon with $k < n$. Now, consider an n-gon and choose a diagonal from its triangulation with endpoints x and y. This diagonal divides the n-gon into two smaller polygons, both of which have the diagonal xy as a boundary edge. Since in these smaller polygons, x and y are adjacent, by the induction hypothesis, there is an other vertex in each of them which has degree 2.[2] Now, we prove that $\chi(G) = 3$. Suppose that the statement is true for every triangulation graph of a k-gon with $k < n$. Consider a triangulation graph G of an n-gon. Choose a vertex x with degree 2, and delete it. The remaining graph, by the induction hypothesis, is 3-colourable. Let us colour it by 3 colours. We now put back the vertex x and colour it with the colour that is not used by its two adjacent vertices. Hence, $\chi(G) = 3$.

To finish the proof, observe that the least frequent colour appears at most $\lfloor \frac{n}{3} \rfloor$-times. Place the watchmen at these colour positions. A triangle has all 3 colours, so it is seen by at least one watchman. ■

[2]Note that those two vertices cannot be adjacent.

1.5 Perfect graphs

Definition 1.5.1. *A subgraph G' of a graph G is said to be* ***induced*** *if, for any pair of vertices x and y of G', xy is an edge of G' if and only if xy is an edge of G. In other words, G' is an induced subgraph of G if it has exactly the edges that appear in G over the same vertex set. Specially, G is an induced subgraph of itself.*

Definition 1.5.2. *A* ***clique*** *in a graph is a set of pairwise adjacent vertices. The* ***clique number*** *$\omega(G)$ of a graph G is the size of a largest clique in G.*

Clearly, for any graph G, we have $\chi(G) \geq \omega(G)$.

Definition 1.5.3. *A graph G is called* ***perfect****, if $\chi(G') = \omega(G')$ for all induced subgraph G' of G.*

Example 1.5.1. *A path (as a graph) is always perfect. All the even cycles and the* ***tree****s (graphs that do not contain any cycle) are perfect too. The odd cycles with more than 3 vertices are not perfect graphs. The full graphs and the bipartite graphs are perfect.*

Definition 1.5.4. *Consider finitely many closed intervals on the real line. Define a graph G as follows: let $V(G)$ be the set of those intervals. We draw an edge between two vertices (intervals) if they are not disjoint. The resulting graph is called an* ***interval graph****.*

Exercise 1.5.1. *Show that the interval graphs are perfect.*

Definition 1.5.5. *A* ***partial order*** *is a binary relation "$<$" over a set S which is irreflexive and transitive, i.e., for all a, b and c in S, we have that*

(i) $a \not< a$ (irreflexivity);

(ii) if $a < b$ and $b < c$ then $a < c$ (transitivity).

A set with a partial order is called a ***partially ordered set*** *(or simply a* ***poset****). For a, b elements of a partially ordered set $P = (S, <)$, if $a < b$ or $b < a$, then a and b are* ***comparable****. Otherwise they are* ***incomparable****. A subset of a poset in which every two distinct elements are comparable is called a* ***chain****. A subset of a poset in which no two distinct elements are comparable is called an* ***antichain****.*

Definition 1.5.6. *Consider a poset $P = (S, <)$. Define a graph G as follows: let $V(G)$ be the set S. We draw an edge between two vertices (elements of S) if they are comparable. The resulting graph is called a* ***comparability graph****.*

Exercise 1.5.2. *Show that the comparability graphs are perfect. (Hint: the length of the longest chain in a poset is the clique number of the corresponding comparability graph.)*

Definition 1.5.7. *Consider a poset $P = (S, <)$. Define a graph G as follows: let $V(G)$ be the set S. We draw an edge between two vertices (elements of S) if they are not comparable. The resulting graph is called a* ***co-comparability graph****.*

Theorem 1.5.1 (Dilworth's Theorem)**.** *Every co-comparability graph is perfect, i.e., the maximum number of elements in any antichain equals the minimum number of chains whose union is the set.*

Proof. $\boxed{\max \leq \min}$ Take an antichain that consists of r elements. These r elements correspond to r different chains.

$\boxed{\max \geq \min}$ We prove by induction on $|S|$. For $|S| = 1$, the statement is obvious. Let k be the the maximum number of elements in any antichain, and let C be a maximal chain in P. If $C = P$, then $k = 1$ and we are done. So assume that $C \neq P$. Because C can contain at most one element of any maximal antichain, the width of a maximal antichain in $P \setminus C$ can be either k or $k - 1$, and both possibilities can occur.

If it is $k - 1$, then by the induction hypothesis, $P \setminus C$ is the union of $k - 1$ chains, whence P is a union of k chains (i.e., the $k-1$ chains whose union is $P \setminus C$ and C itself).

Suppose that the width of a maximal antichain in $P \setminus C$ is still k. Let $\mathcal{A} = \{a_1, a_2, \ldots, a_k\}$ be a maximal antichain in $P \setminus C$. As $|\mathcal{A}| = k$, it is also a maximal antichain in P. Set

$$L = \{x \in P \mid x \leq a_i \text{ for some } i\},$$

$$U = \{x \in P \mid a_j \leq x \text{ for some } j\}.$$

Every element of P should be comparable with some element of $\mathcal{A}$, otherwise we could increase the length of the maximal antichain. Therefore $P = L \cup U$. Clearly, $\mathcal{A} \subseteq L \cap U$. If $x \notin \mathcal{A}$, then $x \notin L \cap U$, otherwise some elements of $\mathcal{A}$ would be comparable to each other. Therefore $\mathcal{A} = L \cap U$ (see Figure 1.11).

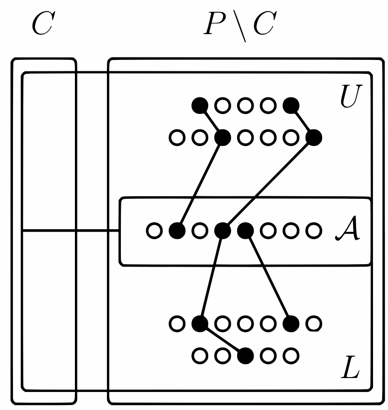

Figure 1.11

Moreover, the maximality of C insures that the largest element of C does not belong to L (remember $\mathcal{A} \subseteq P \setminus C$), so $|L| < |S|$. Similarly, the smallest element of C does not belong to U, whence $|U| < |S|$ also. Therefore, by the induction hypothesis, L is a union of k chains: $L = D_1 \cup D_2 \cup \cdots \cup D_k$, and similarly $U = E_1 \cup E_2 \cup \cdots \cup E_k$ as a union of chains. By renumbering, if necessary, we may assume that $a_i \in D_i \cap E_i$ for $1 \leq i \leq k$, so that $C_i = D_i \cup E_i$ is a chain. Thus $P = L \cup U = C_1 \cup C_2 \cup \cdots \cup C_k$ is a union of k chains. ■

Definition 1.5.8. *The* ***complement*** *of a graph G is a graph $\overline{G}$ on the same vertices such that two vertices of $\overline{G}$ are adjacent if and only if they are not adjacent in G.*

Note that an independent set in a graph is a clique in the complement graph and vice versa. In particular, $\alpha(G) = \omega(\overline{G})$.

Conjecture 1.5.2 (Berge's Weak Perfect Graph Conjecture)**.** *The graph G is perfect if and only if $\overline{G}$ is perfect.*

This conjecture was proved by Lovász in 1972. It is a consequence of the following theorem. We remark that this is not the original way in which he proved the conjecture.

Theorem 1.5.3 (Lovász)**.** *The graph G is perfect if and only if $\alpha(G_0)\omega(G_0) \geq n_0 \stackrel{\text{def}}{=} |V(G_0)|$ for every induced subgraph G_0 of G.*

Proof. $\boxed{\Rightarrow}$ It is true for every graph G_0 that $\alpha(G_0)\chi(G_0) \geq n_0$. If G is perfect then $\chi(G_0) = \omega(G_0)$ for every induced subgraph G_0 of G.

$\boxed{\Leftarrow}$ Suppose, to the contrary, that there exists a non-perfect graph G satisfying the condition, and choose such a graph with $|V(G)|$ minimal. So $\chi(G) > \omega(G)$, while $\chi(G_0) = \omega(G_0)$ for each induced subgraph $G_0 \neq G$ of G.

Claim. Let $U \in V(G)$ be an independent set (here G is our counterexample). Delete from G the vertices contained by U and all the edges between them, and denote the resulting graph by $G - U$. Then we have $\omega(G - U) = \omega(G)$.

Proof. It can be seen that either $\omega(G-U) = \omega(G)$ or $\omega(G-U) = \omega(G) - 1$ must hold. Suppose, to the contrary, that the latter one holds. Since $G - U$ is a proper induced subgraph of G, it is perfect. We can colour it with $\omega(G) - 1$ colours. Doing so, we can colour U with the $\omega(G)$th colour and we see that $\chi(G) \leq \omega(G)$. This is a contradiction, since $\chi(G) > \omega(G)$. ■

Let $\alpha \stackrel{\text{def}}{=} \alpha(G)$ and $\omega \stackrel{\text{def}}{=} \omega(G)$. Take an independent set $A_0 = \{u_1, u_2, \ldots, u_\alpha\}$ of size α in G. According to the claim, we have

$$\boxed{\omega(G - A_0) = \omega.}$$

Also, since $\{u_i\}$ $(i = 1, 2, \ldots, \alpha)$ is an independent set, we have $\omega(G - u_i) = \omega$. Since $G - u_i$ is a proper induced subgraph of G, it is perfect. Hence, it has an ω colouring. Let

$$\{A_{i_1}, A_{i_2}, \ldots, A_{i_\omega}\}$$

be the set of different colour classes of the vertices of $G - u_i$. Since A_{i_j} $(j = 1, 2, \ldots, \omega)$ is a colour class, it is independent. Hence, by the claim, we have

$$\boxed{\omega(G - A_{i_j}) = \omega.}$$

From the two statements in boxes, we conclude that $G - A_0$ contains a clique K^0 of size ω, and $G - A_{i_j}$ contains a clique K^{i_j} of size ω.

Claim. For all $i = 1, 2, \ldots, \alpha$ and $j = 1, 2, \ldots, \omega$, we have

(i) $|V(K^0) \cap A_{i_j}| = 1$;

(ii) $|V(K^{i_j}) \cap A_0| = 1$;

(iii) $|V(K^{i_j}) \cap A_{r_s}| = 1$ for $(i,j) \neq (r,s)$;

(iv) $|V(K^{i_j}) \cap A_{i_j}| = 0$.

Proof. $\boxed{(i)}$ Since K^0 is a clique in $G - A_0 \leq G - u_i$ and A_{i_j} is an independent set in $G - u_i$, we infer that $|V(K^0) \cap A_{i_j}| \leq 1$ for any $j = 1, 2, \ldots, \omega$. Since the ω different sets $A_{i_1}, A_{i_2}, \ldots, A_{i_\omega}$ form a partition of the vertices of $G - u_i$, and the size of K^0 is ω, we must have $|V(K^0) \cap A_{i_j}| = 1$ for all $j = 1, 2, \ldots, \omega$.

$\boxed{(ii)}$ Since K^{i_j} is a clique in $G - A_{i_j} \leq G$ and A_0 is an independent set in G, we infer that $|V(K^{i_j}) \cap A_0| \leq 1$. Similarly, $|V(K^{i_j}) \cap A_{i_s}| \leq 1$ for all $s = 1, 2, \ldots, \omega$. Using this, and (iv), we find

$$\begin{aligned}\left|V\left(K^{i_j}\right) \cap \left(V\left(G\right) - u_i\right)\right| &= \left|V\left(K^{i_j}\right) \cap \dot{\bigcup_{1\leq s\leq\omega}} A_{i_s}\right| = \left|\dot{\bigcup_{1\leq s\leq\omega}} V\left(K^{i_j}\right) \cap A_{i_s}\right| \\ &= \dot{\bigcup_{1\leq s\leq\omega}} \left|V\left(K^{i_j}\right) \cap A_{i_s}\right| \leq \omega - 1.\end{aligned}$$

Since $|V(K^{i_j}) \cap V(G)| = |V(K^{i_j})| = \omega$, we obtain that $u_i \in V(K^{i_j})$, whence $|V(K^{i_j}) \cap A_0| = 1$. Moreover, $|V(K^{i_j}) \cap A_{i_s}| = 1$ for all $s = 1, 2, \ldots, \omega$, $s \neq j$.

$\boxed{(iii)}$ In the previous part of the proof, we showed that $|V(K^{i_j}) \cap A_{i_s}| = 1$ for all $s = 1, 2, \ldots, \omega$, $s \neq j$. Therefore, it is enough to prove that $|V(K^{i_j}) \cap A_{r_s}| = 1$ for $r \neq i$. Since K^{i_j} is a clique in $G - A_{i_j} \leq G$ and A_{r_s} is an independent set in G, we infer that $|V(K^{i_j}) \cap A_{r_s}| \leq 1$. From the previous part of the proof, $V(K^{i_j}) \cap A_0 = \{u_i\}$, whence $|V(K^{i_j}) \cap (V(G) - u_r)| = |V(K^{i_j}) \cap V(G)| = \omega$ for $r \neq i$. Since

$$\begin{aligned}\left|V\left(K^{i_j}\right) \cap \left(V\left(G\right) - u_r\right)\right| &= \left|V\left(K^{i_j}\right) \cap \dot{\bigcup_{1\leq s\leq\omega}} A_{r_s}\right| = \left|\dot{\bigcup_{1\leq s\leq\omega}} V\left(K^{i_j}\right) \cap A_{r_s}\right| \\ &= \dot{\bigcup_{1\leq s\leq\omega}} \left|V\left(K^{i_j}\right) \cap A_{r_s}\right| \leq \omega,\end{aligned}$$

we must have equality, i.e., $|V(K^{i_j}) \cap A_{r_s}| = 1$ for $r \neq i$.

$\boxed{(iv)}$ Since K^{i_j} is a clique in $G - A_{i_j}$ and A_{i_j} is the set of vertices in G that are not in $G - A_{i_j}$, it follows that $|V(K^{i_j}) \cap A_{i_j}| = 0$. ■

Let $v_1, v_2, \ldots, v_n$ be the vertices of G. Take the characteristic vectors of the sets $A_0, A_{1_1}, A_{1_2}, \ldots, A_{\alpha_\omega}$. The characteristic vector of a set of vertices is an n-dimensional vector whose kth coordinate is 1 if the set contains v_k and 0 otherwise. Form a matrix $\mathcal{A}$ whose ℓth row is the characteristic vector of A_ℓ as a row vector ($\ell = 0, 1_1, 1_2, \ldots, \alpha_\omega$). Hence, $\mathcal{A}$ is an $\alpha\omega + 1 \times n$ matrix. Similarly, take the characteristic vectors of the sets $V(K^0), V(K^{1_1}), V(K^{1_2}), \ldots, V(K^{\alpha_\omega})$, and form a matrix $\mathcal{B}$ whose ℓth column is the characteristic vector of $V(K^\ell)$ as a column vector ($\ell = 0, 1_1, 1_2, \ldots, \alpha_\omega$). Thus, $\mathcal{B}$ is an $n \times \alpha\omega + 1$ matrix.

It is easy to see that the product matrix $\mathcal{AB}$ counts the elements in the intersections of the sets $A_0, A_{1_1}, A_{1_2}, \ldots, A_{\alpha_\omega}$ with the sets $V(K^0), V(K^{1_1}), V(K^{1_2}), \ldots,$ $V(K^{\alpha_\omega})$. More precisely, the (s,t)th element of $\mathcal{AB}$ is $|V(K^t) \cap A_s|$ for $s, t = 0, 1_1, 1_2, \ldots, \alpha_\omega$. Therefore, by the second claim, $\mathcal{AB}$ is an $\alpha\omega + 1 \times \alpha\omega + 1$ matrix with zeros in the diagonal and 1 elsewhere. We have $\operatorname{rank}(\mathcal{AB}) = \alpha\omega + 1$ (the columns are linearly independent), and $\operatorname{rank}(\mathcal{A}) \le n$ (the number of columns is always an upper bound). Therefore,

$$\alpha\omega + 1 = \operatorname{rank}(\mathcal{AB}) \le \operatorname{rank}(\mathcal{A}) \le n,$$

whence, $\alpha\omega < n$. This is a contradiction, since G satisfies the assumption of the theorem. ∎

Proof of the Weak Perfect Graph Conjecture. Let G be a graph. First, we show that the complement operation is a bijection between the induced subgraphs of G and the induced subgraphs of $\overline{G}$. Since the complement operation is idempotent (i.e., $\overline{\overline{G_0}} = G_0$), it is enough to show that for any induced subgraph G_0 of G, the $\overline{G_0}$ is an induced subgraph of $\overline{G}$.

Suppose that G_0 is induced by $S \subseteq V(G)$. Clearly, $V(\overline{G_0}) = V(G_0) = S$ and

$$uv \in E\left(\overline{G_0}\right) \overset{\substack{\text{by the definition}\\ \text{of the complement}}}{\iff} uv \notin E\left(G_0\right) \overset{\substack{\text{by the definition}\\ \text{of the induced subgraph}}}{\iff} uv \notin E\left(G\right)$$
$$\overset{\substack{\text{by the definition}\\ \text{of the complement}}}{\iff} uv \in E\left(\overline{G}\right),$$

whence $\overline{G_0}$ is an induced subgraph of $\overline{G}$.

Now, by Lovász's Theorem, a graph G is perfect if and only if

$$\alpha\left(G_0\right)\omega\left(G_0\right) \ge \left|V\left(G_0\right)\right| \Leftrightarrow \omega\left(\overline{G_0}\right)\alpha\left(\overline{G_0}\right) \ge \left|V\left(G_0\right)\right| \Leftrightarrow \omega\left(\overline{G_0}\right)\alpha\left(\overline{G_0}\right) \ge \left|V\left(\overline{G_0}\right)\right|,$$

for every induced subgraph G_0 of G. By the previous argument, as G_0 runs through the induced subgraphs of $\overline{G}$, the $\overline{G_0}$ runs through the induced subgraphs of $\overline{G}$. Hence, the above is equivalent to the fact that $\overline{G}$ is perfect. ■

Definition 1.5.9. *A* ***complete bipartite graph****, $G = (V_1 + V_2, E)$, is a bipartite graph such that for any two vertices, $v_1 \in V_1$ and $v_2 \in V_2$, we have $v_1v_2 \in E$.*

The following theorem is an other example of a combinatorial statement that can be proved by linear algebraic methods.

Theorem 1.5.4 (Graham–Pollak)**.** *Let $G_1, G_2, \ldots, G_m$ be subgraphs of the complete graph K_n. Suppose that*

(i) each G_i is a complete bipartite graph $(1 \leq i \leq m)$;

(ii) $\cup_{i=1}^m E(G_i) = E(K_n)$;

(iii) $E(G_i) \cap E(G_j) = \emptyset$ whenever $i \neq j$.

Then $m \geq n - 1$.

Proof. Suppose that $E(K_n) = E(G_1)\dot{\cup}E(G_2)\dot{\cup}\cdots\dot{\cup}E(G_m)$. Let us denote by $x_1, x_2, \ldots, x_n$ the vertices of K_n. Consider the complete bipartite graph G_i and denote by A_i and B_i the the two independent sets of its vertices. If A_i contains precisely the vertices $x_{i_1}, x_{i_2}, \ldots, x_{i_\ell}$ then assign the polynomial $P_{A_i}(x_{i_1}, x_{i_2}, \ldots, x_{i_\ell}) = x_{i_1} + x_{i_2} + \cdots + x_{i_\ell}$ to A_i. Similarly, we can assign a polynomial Q_{B_i} to B_i. If we expand the product $P_{A_i}Q_{B_i}$, we get a formal sum of the edges of G_i (since G_i is a complete bipartite graph), e.g., the term x_rx_s corresponds to the edge x_rx_s. Hence, by the assumptions (ii) and (iii), it follows that

$$\sum_{i=1}^{m} P_{A_i}Q_{B_i} = \sum_{1 \leq i < j \leq n} x_ix_j = \frac{1}{2}\left[\left(\sum_{i=1}^{n} x_i\right)^2 - \sum_{i=1}^{n} x_i^2\right].$$

Assume, to the contrary, that $m < n - 1$. Consider the following system of $m + 1 < n$ equations:

$$P_{A_1} = 0,\ P_{A_2} = 0,\ \ldots\ , P_{A_m} = 0,\ \sum_{i=1}^{n} x_i = 0.$$

Since the number of variables ($= n$) is strictly larger than the number of equations, the system must have a non-trivial solution $\alpha_1, \alpha_2, \ldots, \alpha_n$, say. It follows that

$$0 = \sum_{i=1}^{m} P_{A_i}(\alpha_1, \alpha_2, \ldots, \alpha_n) \, Q_{B_i}(\alpha_1, \alpha_2, \ldots, \alpha_n) = \frac{1}{2}\left[\left(\sum_{i=1}^{n} \alpha_i\right)^2 - \sum_{i=1}^{n} \alpha_i^2\right]$$
$$= -\frac{1}{2}\sum_{i=1}^{n} \alpha_i^2 < 0,$$

which is a contradiction. ∎

Remark. *It is not hard to show that the bound $m \geq n-1$ is sharp.*

1.6 Chordal graphs

Definition 1.6.1. *A graph G is* ***chordal*** *if and only if $\mathcal{C}_k \not\leq G$ as an induced subgraph for any $k \geq 4$. In other words a graph G is chordal if each of its cycles of four or more vertices has a* ***chord****, which is an edge joining two vertices that are not adjacent in the cycle.*

Theorem 1.6.1. *The interval graphs are chordal.*

Proof. Suppose that we have a cycle $v_1v_2v_3\ldots v_\ell v_1$ with $\ell > 3$. Denote by $I_i = [a_i, b_i]$ the interval that corresponds to the vertex v_i in our interval graph G ($1 \leq i \leq n$). After a possible renumbering, we can assume that $b_1 \leq b_i$ for all $2 \leq i \leq n$. Take the leftmost right endpoint: b_1. Consider the neighbours v_2, v_ℓ of v_1. By the assumption, $b_1 \leq b_2$ and $b_1 \leq b_\ell$. This, together with the facts $I_1 \cap I_2 \neq \emptyset$ and $I_1 \cap I_\ell \neq \emptyset$ implies $b_1 \in I_2, I_\ell$. Thus, $I_2 \cap I_\ell \neq \emptyset$ which means that v_2v_ℓ is a chord of our cycle. ∎

Definition 1.6.2. *Consider two graphs G_1 and G_2. Suppose that we have a clique of size t in both of them consisting of the set of vertices $V_1 \subseteq V(G_1)$ and $V_2 \subseteq V(G_2)$. Form a new graph G whose vertex set is $V(G) = {}^{V(G_1) \cup V(G_2)}/_{\sim}$, where $\sim$ means that we identify each vertex from V_1 with a vertex from V_2 (this can be done in $t!$ ways). The set of edges of G is $E(G) = {}^{E(G_1) \cup E(G_2)}/_{\approx}$, where $\approx$ means that we identify each edge $v_1u_1 \in E(G_1)$ with $v_2u_2 \in E(G_2)$ if and only if* *we identified v_1 with v_2 and u_1 with u_2 by $\sim$. This G is the **pasting** of G_1 and G_2 **along a clique** of size t (see Figure 1.12).*

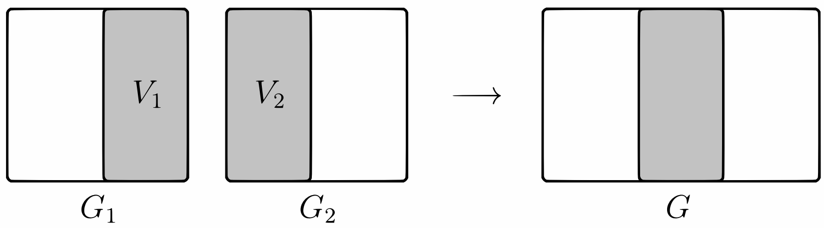

Figure 1.12. *Pasting along a clique.*

Theorem 1.6.2. *A graph is chordal if and only if it can be obtained by means of a sequence of pastings along cliques starting with complete graphs.*

Proof. $\boxed{\Leftarrow}$ By induction on the number of complete graphs we pasted. If no pasting required, i.e., the graph is a single complete graphs, than it is obviously chordal. Now, suppose that we have a graph G_1 which was obtained by means of a sequence of pastings along cliques starting with k complete graphs ($k \geq 1$). We paste a complete graph G_2 and G_1 together along a clique S. We need to prove that the resulting graph G is chordal. Suppose that $\mathcal{C}_\ell \in G$ for $\ell \geq 4$.

If $\mathcal{C}_\ell$ lies entirely in G_1 or G_2, than it has a chord. Indeed, by the induction hypothesis, G_1 is chordal and by assumption, G_2 is complete.

Assume that $\mathcal{C}_\ell$ contains vertices $v_1 \in V(G_1)$ and $v_2 \in V(G_2)$ that are not in S. Clearly, $v_1v_2 \notin E(G)$. Therefore, $\mathcal{C}_\ell$ must have two vertices in S, x and y, say (see Figure 1.13). Since S is a clique, $xy \in E(G)$, which is a chord of $\mathcal{C}_\ell$.

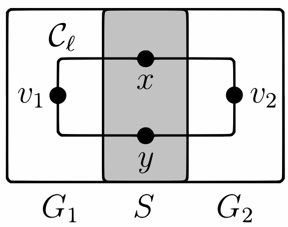

Figure 1.13

$\boxed{\Rightarrow}$ Suppose that G is chordal. We need to prove that it can be obtained by means of a sequence of pastings along cliques starting with complete graphs. We proceed by induction on $|V(G)|$. We can assume that G is connected. Indeed, if $G_1, G_2, \ldots, G_k$ are the connected components of G, then we can obtain G from them by a sequence of pastings along empty cliques. Hence, it is enough to show that each G_i has the required property ($1 \leq i \leq k$). If $|V(G)| = 1$ or 2, the statement is obvious.

Suppose that the statement is true for chordal graphs with vertices less than n. Assume that $|V(G)| = n$. If $G = K_n$, the full graph on n vertices, we are done. Otherwise, there exist $x, y \in V(G)$ such that $xy \notin V(G)$. Take a minimal $S \in V(G)$ that separates x and y, i.e., x and y are in distinct components of $G - S$. Not that S can not be empty, since G is connected.

Let V_1 denote the set of vertices of the component of x in $G - S$. Similarly, denote by V_2 the set of vertices of the component of y in $G - S$. Of course, $G - S$ may have other components beside these.

Claim. The set S forms a clique.

Proof. If $|S| = 1$, we are done. Suppose that $|S| > 1$. Let u and v be two different vertices in S. We need to show that $uv \in E(G)$. There exists $x_1 \in V_1$ such that $x_1u \in E(G)$. Otherwise, $S - u$ would be a set that separates x and y, and $|S - u| < |S|$ which is a contradiction since S was minimal. Similarly, $x_2 \in V_1$ and $y_1, y_2 \in V_2$ such that $x_2v \in E(G)$, $y_1u \in E(G)$ and $y_2v \in E(G)$. Since in V_1 (resp. V_2) are connected in $G - S$, there exists a path between x_1 and x_2, and there exists a path between y_1 and y_2. In this way we obtain a path from u to v in V_1 and another path from u to v in V_2. Choose x_1, x_2 and the path between them so that the resulting path from u to v in V_1 is the shortest possible. Similarly for y_1, y_2 an the path in V_2 (see Figure 1.14).

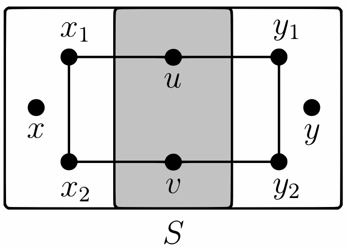

Figure 1.14

So we have obtained a cycle $\mathcal{C}$ in G that contains u and v. Since this cycle also has to contain at least one point in V_1 and one point V_2, its length is at least 4.

Since G is chordal, $\mathcal{C}$ must have a chord. If the chord is uv, we are done. Otherwise there are three different cases.

(i) The chord connects a vertex from V_1 to a vertex from V_2. This is not possible, since V_1 and V_2 are the set of vertices of disjoint components in $G - S$.

(ii) The chord connects two vertices in V_1 (resp. V_2). This is not possible, since in this way we would have a shorter path in V_1 (resp. V_2) from u to v.

(iii) The chord connects u (resp. v) with a vertex in V_1. This is not possible, because then, we would have been able to choose a shorter path in V_1 from u to v. Similarly, we can not have a chord that connects u (resp. v) with a vertex in V_2.

Therefore, the chord must be uv, so $uv \in E(G)$. ■

Thus, we have seen that S is a clique. Now, let $V_2' = V_2$, the set of vertices of the component of y in $G - S$. Let $V_1' = V(G - S) - V_2'$, the set of vertices of all the other components in $G - S$. Finally, let G_1 and G_2 be the graphs induced by $V_1' \cup S$ and $V_2' \cup S$, respectively. Then G is the pasting of G_1 and G_2 along the clique S. Since $x \notin V_2' \cup S = V(G_2)$ and $y \notin V_1' \cup S = V(G_1)$, we have $|V(G_1)| \leq |V(G)| - 1 = n - 1$ and $|V(G_2)| \leq |V(G)| - 1 = n - 1$. Since both G_1 and G_2 are induced subgraphs of the chordal graph G, they are chordal. By the induction hypothesis, G_1 and G_2 were obtained by means of a sequence of pastings along cliques starting with complete graphs. Since G is the pasting of G_1 and G_2 along the clique S, we deduce that G is obtained by a sequence of pastings along cliques starting with complete graphs. ■

Corollary 1.6.3. *Chordal graphs are perfect.*

Proof. Since the full graphs are perfect, it is enough to show that pasting preserves perfectness. For this, it is enough that $\chi(G_i) = \omega(G_i)$ for $i = 1, 2$, and G is obtained by pasting G_1 and G_2 along a clique, implies $\chi(G) = \omega(G)$. This is enough, because

(i) an induced subgraph of a graph G obtained by pasting G_1 and G_2 along a clique S, is the pasting of an induced subgraph of G_1 with an induced subgraph of G_2 along a clique $S' \subseteq S$;

(ii) an induced subgraph of a full graph is a full graph.

Therefore, suppose that G obtained by pasting G_1 and G_2 along a clique S. Let $\chi_i \overset{\text{def}}{=} \chi(G_i)$ and $\omega_i \overset{\text{def}}{=} \omega(G_i)$ for $i = 1, 2$.

Claim. We have

(i) $\chi(G) = \max\{\chi_1, \chi_2\}$;

(ii) $\omega(G) = \max\{\omega_1, \omega_2\}$.

Proof. $\boxed{(i)}$ Colour G_1 with colours $1, 2, \ldots, \chi_1$. Call this colouring c_1. Similarly, colour G_2 with colours $1, 2, \ldots, \chi_2$, and call this colouring c_2. Taking a permutation of $\{1, 2, \ldots, \chi_2\}$ if necessary, we may assume that the vertices in S have the same colour in both these colourings. Then $c \overset{\text{def}}{=} c_1 \cup c_2 : V(G) \to \{1, 2, \ldots, \max\{\chi_1, \chi_2\}\}$ is a proper colouring. Thus, $\chi(G) \leq \max\{\chi_1, \chi_2\}$.

On the other hand, $\chi(G) \geq \chi_i$ for $i = 1, 2$, since G_1 is isomorphic to a subgraph of G induced by $V(G) - V(G_2)$ and G_2 is isomorphic to a subgraph of G induced by $V(G) - V(G_1)$. Hence, $\chi(G) \geq \max\{\chi_1, \chi_2\}$.

$\boxed{(ii)}$ By taking a maximal clique in G, and noting that there are no edges between $V(G_1) - S$ and $V(G_2) - S$, we see that this clique must be entirely in $V(G_1)$ or $V(G_2)$. Hence, its size is at most $\max\{\omega_1, \omega_2\}$, i.e., $\omega(G) \leq \max\{\omega_1, \omega_2\}$.

On the other hand, $\omega(G) \geq \omega_i$ for $i = 1, 2$, since G_1 is isomorphic to a subgraph of G induced by $V(G) - V(G_2)$ and G_2 is isomorphic to a subgraph of G induced by $V(G) - V(G_1)$. ■

Thus, if we also assume that $\chi_i = \omega_i$ for $i = 1, 2$, we have, by the claim, that $\chi(G) = \max\{\chi_1, \chi_2\} = \max\{\omega_1, \omega_2\} = \omega(G)$. ■

Definition 1.6.3. *Suppose that T is a tree, and $T_i \leq T$ is a subtree ($i = 1, 2, \ldots, n$). We construct a graph G in the following way. The vertices of G are the trees $T_1, T_2, \ldots, T_n$. We connect to vertices by an edge if they share a common vertex as subtrees. The resulting G is called the* ***intersection graph of the subtrees*** *$T_1, T_2, \ldots, T_n$* ***of the tree*** *T.*

Exercise 1.6.1 (Gavril's Theorem)**.** *A graph G is chordal if and only if G can be obtained as an in intersection graph of subtrees $T_1, T_2, \ldots, T_n$ of a tree T. (Hint: use the lemma below.)*

Lemma 1.6.4 (Subtrees of a tree have the one-dimensional Helly-property). *Suppose that T is a tree, and $T_i \leq T$ is a subtree ($i = 1, 2, \ldots, n$), such that $V(T_i) \cap V(T_j) \neq \emptyset$ for all $i \neq j$. Then $\cap_{i=1}^{n} V(T_i) \neq \emptyset$.* ■

Conjecture 1.6.5 (Berge's Strong Perfect Graph Conjecture). *The graph G is perfect if and only if G does not contain any $\mathcal{C}_{2k+1}$ or $\overline{\mathcal{C}_{2k+1}}$ as an induced subgraph for $k \geq 2$.*

A proof by Chudnovsky, Robertson, Seymour, and Thomas was announced in 2002 and published by them in 2006.

1.7 Plane and planar graphs

Recall that a planar graph is a graph that can be embedded in the plane, i.e., it can be drawn on the plane in such a way that its edges intersect only at their endpoints. Such a drawing is called a *plane graph*.

Consider the graphs K_5 and $K_{3,3}$ (the full bipartite graph on 6 vertices, where each vertex class has 3 vertices). Are these graphs planar? Figure 1.15 suggests they are not.

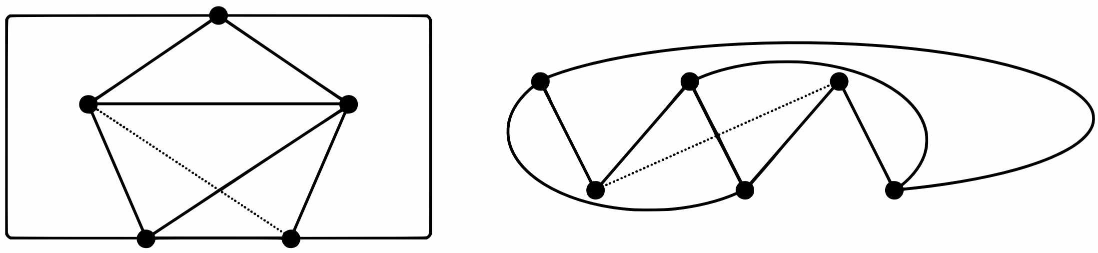

Figure 1.15. *The graphs K_5 and $K_{3,3}$ are not planar.*

To give a rigorous proof, we need the following theorem.

Theorem 1.7.1 (Euler's formula). *Let G be a connected plane graph with n vertices, e edges and f faces (including the external or unbounded face). Then $e + 2 = n + f$.*

Proof. We proceed by induction on e. Since the graph is connected, $e \geq n - 1$. If $e = n - 1$, then G is a tree. In this case, $e = n - 1$, $n = n$, $f = 1$, i.e., $e + 2 = n + f$ holds. Suppose that $e \geq n$. In this case we have at least one cycle in G. Delete one edge of the cycle. The resulting graph has $e-1$ edges, n vertices and $f-1$ faces. By the induction hypothesis, $(e-1)+2 = n+(f-1)$, i.e., $e+2 = n+f$. ■

Theorem 1.7.2. *Let G be a connected planar graph with $n \geq 3$ vertices and e edges. Then $e \leq 3n-6$.*

Proof. Take a plane representation of G. Let f be the number of faces of the graph G. Every face is bounded by at least 3 edges (this is true for the unbounded face as well). Hence, since every edge is a boundary of two faces, we have $2e \geq 3f$. By Euler's formula, $f = e-n+2$, which gives $2e \geq 3f \geq 3e-3n+6$. Solving for e yields $e \leq 3n-6$. ■

For K_5, we have $e = 10$ and $n = 5$. Thus, $e \leq 3n-6$ does not hold, i.e., K_5 can not be planar. However, for $K_{3,3}$ the inequality certainly holds.

Theorem 1.7.3. *Let G be a connected, triangle-free planar graph with $n \geq 3$ vertices and e edges. Then $e \leq 2n-4$.*

Proof. Take a plane representation of G. Let f be the number of faces of the graph G. Since G is triangle-free, every face is bounded by at least 4 edges (this is true for the unbounded face as well). Hence, since every edge is a boundary of two faces, we have $2e \geq 4f$. By Euler's formula, $f = e-n+2$, which gives $2e \geq 4f \geq 4e-4n+8$. Solving for e yields $e \leq 2n-4$. ■

For the triangle-free graph $K_{3,3}$, we have $e = 9$ and $n = 6$. Thus, $e \leq 2n-4$ does not hold, i.e., $K_{3,3}$ can not be planar.

Definition 1.7.1. *A* ***subdivision of a graph*** *G is a graph resulting from the* ***subdivision of edges*** *in G. The subdivision of some edge uv yields a graph containing one new vertex w, and with an edge set replacing uv by two new edges, uw and wv.*

Definition 1.7.2. *A graph H is a* ***topological subgraph*** *of the graph G if a subdivision of H is a isomorphic to a subdivision of a subgraph of G. Equivalently, H is a topological subgraph of G if H can be obtained from G by deleting edges, deleting vertices, subdividing edges, and dissolving degree 2 vertices (which means deleting the vertex and making its two neighbours adjacent).*

Theorem 1.7.4 (Kuratowski)**.** *A graph G is planar if and only if it does not contain K_5 or $K_{3,3}$ as topological subgraphs.* ■

Exercise 1.7.1. *Prove, in two different ways, that the so-called **Petersen graph** (Figure 1.16) is not planar.*

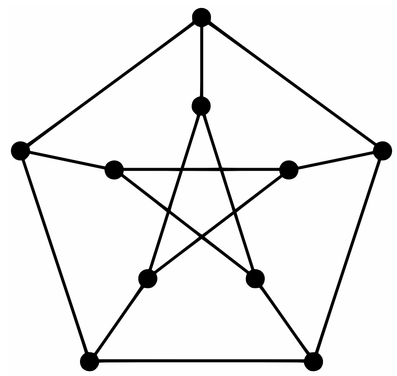

Figure 1.16. *The Petersen graph.*

Definition 1.7.3. *A **contraction of a graph** G is a graph resulting from the **contraction of edges** in G. The edge contraction operation occurs relative to a particular edge in the graph. The edge e is removed and its two adjacent vertices, u and v, are merged into a new vertex w, where the edges adjacent to w each correspond to an edge adjacent to either u or v.*

Definition 1.7.4. *A graph H is a **minor** of the graph G if H is a contraction of a subgraph of G. Equivalently, H is a topological subgraph of G if H can be obtained from G by deleting edges, deleting vertices, and contracting some edges.*

Theorem 1.7.5 (Wagner)**.** *A graph G is planar if and only if it does not contain K_5 or $K_{3,3}$ as minors.* ■

Theorem 1.7.6. *Every planar graph G is 5-colourable, i.e., $\chi(G) \leq 5$.*

Proof. Let G be a planar graph and take a plane representation of it. We proceed by induction on the number of vertices. It is obvious the theorem is true for a graph with only one vertex.

First, we show that G has a vertex with degree at most 5. Suppose, to the contrary, that every vertex of G has degree 6 or more. By Theorem 1.7.2, we have

$$\sum_{v \in V(G)} d(v) = 2e \leq 2(3n-6) = 6n - 12.$$

However, if every vertex has degree greater than 5 as we supposed, $\sum_{v \in V(G)} d(v) \geq 6n$, which is a contradiction. Therefore, G has at least one vertex with at most 5 edges, which we will call x. Remove that vertex x from G to create another graph, G'. This G' is still a planar graph. Therefore, by the induction hypothesis, G' is 5-colourable. We colour it with 5 colours.

If all five colours were not connected to x, then we can give x a missing colour and thus, we obtain an 5-colouring of G.

If all five colours were connected to x, we examine the five vertices x was adjacent to, and call them y_1, y_2, y_3, y_4 and y_5 (in order around x). We colour y_i with colour i $(1 \leq i \leq 5)$.

We now consider the subgraph $G_{1,3}$ of G' induced by the vertices coloured with colour 1 and 3. If there is no walk between y_1 and y_3 in $G_{1,3}$, then we can simply switch the colours 1 and 3 in the component of $G_{1,3}$ containing y_1. Thus, x is no longer adjacent to a vertex of colour 1, so we can colour it with colour 1.

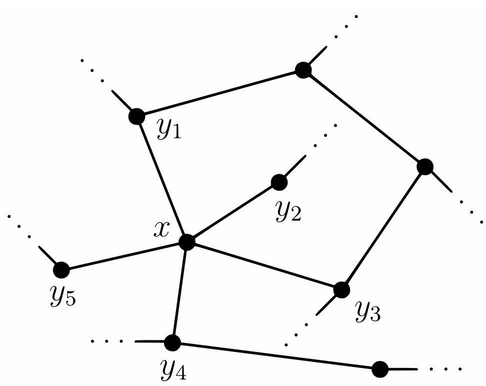

Figure 1.17

Suppose that there is a walk between y_1 and y_3 in $G_{1,3}$. Consider the subgraph $G_{2,4}$ of G' induced by the vertices coloured with colour 2 and 4. However, since G is planar and there is a cycle in G that consists of the walk from y_1 to y_3, x, and the edges xy_1 and xy_3, clearly y_2 cannot be connected to y_4 within G' (the cycle encloses either y_2 or both y_4 and y_5, see Figure 1.17). Thus, we can switch the colours 2 and 4 in the component of $G_{2,4}$ containing y_2. Thus, x is no longer adjacent to a vertex of colour 2, so we can colour it with colour 2. ∎

Definition 1.7.5. *Let G be a graph. For every $v \in V(G)$, take a finite subset $L(v)$ of the positive integers. The $L(v)$ is the **colour list** which corresponds to v. A map $c : V(G) \to \mathbb{Z}^+$, $c(v) \in L(v)$, is called a (**proper**) **list colouring of** G if $c(x) \neq c(y)$ for all $xy \in E(G)$. The positive integer*

$$\chi_\ell(G) \stackrel{\text{def}}{=} \min \{k \mid G \textit{ has a list colouring for every set of lists } L(v),\\ v \in V(G),\ |L(v)| = k\}$$

*is called the **list chromatic number of** G.*

Clearly, $\chi_\ell(G) \geq \chi(G)$, and by the greedy algorithm $\chi_\ell(G) \leq \Delta(G) + 1$.

Theorem 1.7.7. *For all $k \geq 3$, there exists a graph G such that $\chi(G) = 2$ and $\chi_\ell(G) \geq k$.*

Proof. Let $k = 3$ and consider the graph $K_{3,3}$ with the list colouring shown in Figure 1.18. It is easy to check that this choice of colour lists forces χ_ℓ to be at least 3.

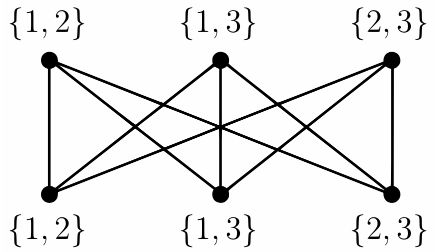

Figure 1.18

The proof for higher k's is left as an exercise. ∎

Exercise 1.7.2. *Prove Theorem 1.7.7 for $k \geq 4$.*

Theorem 1.7.8 (Thomassen)**.** *If G is planar then $\chi_\ell(G) \leq 5$.*

The trick is to find a suitable extension of the theorem. The theorem follows from the following stronger lemma.

Lemma 1.7.9. *Let G be a planar graph which consists of a cycle $\mathcal{C}$: $v_1v_2\dots v_kv_1$, and vertices and edges inside $\mathcal{C}$ such that each bounded face is bounded by a triangle. Suppose that v_1 and v_2 are coloured with colours 1 and 2, respectively, and that $L(v)$ is a list of 3 colours if $v \in V(\mathcal{C}) - \{v_1, v_2\}$ and 5 colours if $v \in V(G - \mathcal{C})$. Then the colouring of v_1 and v_2 can be extended to a list colouring of G.*

Proof. By induction on the number of vertices of G. If $|V(G)| = 3$, then $G = \mathcal{C}$ and there is nothing to prove. We proceed to the induction step.

If $\mathcal{C}$ has a chord v_iv_j, where $2 \le i \le j-2 \le p-1$ ($v_{p+1} = v_1$), then we apply the induction hypothesis to the cycle $v_1v_2\dots v_iv_jv_{j+1}\dots v_1$ and its interior and then to the cycle $v_jv_iv_{i+1}\dots v_{j-1}v_j$ and its interior.

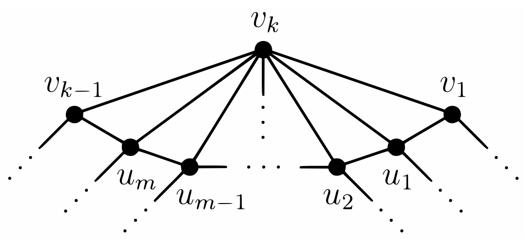

Figure 1.19

Assume that $\mathcal{C}$ has no chord. Let $v_1, u_1, u_2, \dots, u_m, v_{k-1}$ be the neighbours of v_k in that clockwise order around v_k (see Figure 1.19). As the interior of $\mathcal{C}$ is triangulated, G contains the path P: $v_1u_1u_2\dots u_mv_{k-1}$. As $\mathcal{C}$ is chordless, $P \cup (\mathcal{C} - v_k)$ is a cycle $\mathcal{C}'$. Let a, b be two distinct colours in $L(v_k) \setminus \{1\}$ (it is not necessary to have $1 \in L(v_k)$). Now delete from $L(u_i)$ the colours a, b if it contains them, otherwise delete two arbitrary colours ($1 \le i \le m$). The other lists in G remain unchanged. This yields a new colour list L'. Then we apply the induction hypothesis to $\mathcal{C}'$ and its interior and the new list L'. We complete the colouring by assigning a or b to v_k such that v_k and v_{k-1} get distinct colours (v_1 has colour $1 \ne a, b$). ■

Definition 1.7.6. *Let G be a graph. The map $c : E(G) \to \mathbb{Z}^+$ is called an* ***(proper) edge colouring of*** *G if $c(e) \ne c(f)$ for all $e, f \in E(G)$ that share a common vertex. The positive integer*

$$\chi'(G) \stackrel{\text{def}}{=} \min\{k \mid G \text{ has an edge colouring with colours } 1, 2, \dots, k\}$$

is called the ***edge chromatic number*** *(or* ***chromatic index****)* ***of*** *G.*

Clearly, $\Delta(G) \leq \chi'(G)$. Vizing's Theorem states that $\chi'(G) \leq \Delta(G) + 1$. Similarly to χ_ℓ, we can define the *list chromatic index* χ'_ℓ. Clearly, $\Delta(G) \leq \chi'_\ell(G)$.

Conjecture 1.7.10. *For any graph G, we have $\chi'_\ell(G) \leq \Delta(G) + 1$, moreover, $\chi'_\ell(G) = \chi'(G)$.*

Chapter 2

Ramsey theory

2.1 The pigeonhole principle

The *pigeonhole principle* states that, if n objects (pigeons) are put into m pigeonholes, then there exists a pigeonhole that contains at least $\lceil n/m \rceil$ objects, and similarly, there exists a pigeonhole that contains at most $\lfloor n/m \rfloor$ objects.

Proposition 2.1.1. *Let $S \subseteq \{1, 2, \ldots, 2n\}$ such that $|S| = n+1$. Then there exist x and y in S that are coprime.*

Proof. We prove that there exist x and y in S which are consecutive. This is enough, since consecutive numbers are coprime (their difference is 1, so no number greater than 1 can divide both of them). Let $S_1 = \{1, 2\}, S_2 = \{3, 4\}, \ldots, S_i = \{2i-1, 2i\}, \ldots, S_n = \{2n-1, 2n\}$. These are n different pigeonholes. Since the set S is of size $n+1$, by the pigeonhole principle, there must be an S_i that contains two different elements of S $(1 \leq i \leq n)$. This implies that there are two elements x and y in S such that $x = 2i - 1$, and $y = 2i$, so x and y are consecutive. ∎

Proposition 2.1.2. *Let $S \subseteq \{1, 2, \ldots, 2n\}$ such that $|S| = n+1$. Then there exist x and y in S such that $x \mid y$.*

Proof. Let $S_1 = \left\{2^k \mid k \geq 0, 2^k \leq 2n\right\}$. Similarly, let

$$\begin{aligned}
S_2 &= \left\{3 \cdot 2^k \mid k \geq 0, 3 \cdot 2^k \leq 2n\right\}, \\
S_3 &= \left\{5 \cdot 2^k \mid k \geq 0, 5 \cdot 2^k \leq 2n\right\}, \\
S_4 &= \left\{7 \cdot 2^k \mid k \geq 0, 7 \cdot 2^k \leq 2n\right\}, \\
&\vdots \\
S_n &= \{2n-1\}.
\end{aligned}$$

Clearly, we have $\{1, 2, \ldots, 2n\} = \dot{\bigcup}_{i=1}^{n} S_n$. Since we choose $n+1$ elements from this set, by the pigeonhole principle, there must be two from the same S_i for some $1 \leq i \leq n$. But for any $x, y \in S_i$, either $x \mid y$ or $y \mid x$. ■

Proposition 2.1.3 (Erdős–Szekeres)**.** *For $n \geq 0$, let $x_1, x_2, \ldots, x_{n^2+1}$ be a finite sequence of real numbers, such that $x_i \neq x_j$ for $i \neq j$. Then either there exist $i_1 < i_2 < \cdots < i_{n+1}$ such that $x_{i_1} < x_{i_2} < \cdots < x_{i_{n+1}}$, or there exist $j_1 < j_2 < \cdots < j_{n+1}$ such that $x_{j_1} > x_{j_2} > \cdots > x_{j_{n+1}}$.*

Proof. Denote by $f(i)$ the length of the longest monotone increasing subsequence starting with x_i. If $f(i) \geq n+1$ for some i, then we are done (we have the first case). So we may assume that $1 \leq f(i) \leq n$. Since $1 \leq i \leq n^2+1$, by the pigeonhole principle, there exist $j_1 < j_2 < \cdots < j_{n+1}$ such that $f(j_1) = f(j_2) = \cdots = f(j_{n+1})$. Clearly, if $x_{j_1} < x_{j_2}$, then $f(j_1) > f(j_2)$. So we must have $x_{j_1} > x_{j_2}$. Similarly, $x_{j_2} > x_{j_3}$ and so on. Hence, we have $x_{j_1} > x_{j_2} > \cdots > x_{j_{n+1}}$ as required. ■

Exercise 2.1.1. *Prove that the Erdős–Szekeres theorem is sharp, i.e., it does not hold in general for n^2 different real numbers.*

2.2 Ramsey's Theorem

Theorem 2.2.1 (Ramsey's Theorem, Special Case 1)**.** *For every $k, \ell \geq 2$ there exists a positive integer $R(k, \ell)$, such that if we 2-colour the edges of the complete graph on $n \geq R(k, \ell)$ vertices with colours red and blue, then there exists either a red K_k or a blue K_ℓ.*

Note that here colouring does not mean proper colouring, we just simply assign colour 1 or colour 2 into each edge of the graph.

Proof. We shall show that

$$R(k,\ell) = \binom{k+\ell-2}{k-1} \text{ is a good choice.}$$

It does not mean that this is the best possible. We prove by induction on $k+\ell$. If $k,\ell = 2$, by the above choice, $R(2,2) = 2$. It is obvious that K_2 has the required property.

Now consider K_n with $n \geq \binom{k+\ell-2}{k-1}$. Colour the edges with colours red and blue. Choose a vertex v from the graph, and partition the remaining vertices into two sets R and B, such that for every vertex u:

(i) $u \in R$ if uv is red;

(ii) $u \in B$ if uv is blue.

We claim that either

$$|R| \geq \binom{k+\ell-3}{k-2} \quad \text{or} \quad |B| \geq \binom{k+\ell-3}{k-1}.$$

Otherwise we would have

$$n \geq \binom{k+\ell-2}{k-1} = \binom{k+\ell-3}{k-2} + \binom{k+\ell-3}{k-1} \geq |R| + 1 + |B| + 1 = n+1,$$

which is a contradiction. Suppose that $|R| \geq \binom{k+\ell-3}{k-2}$ holds. By the induction hypothesis, the complete subgraph of K_n, induced by R has a blue K_ℓ or a red K_{k-1}. In the former case, we are done. In the latter case, the complete subgraph of K_n, induced by the red K_{k-1} and v, is a red K_k by the definition of R. The case when $|B| \geq \binom{k+\ell-3}{k-1}$, is analogous. ■

Theorem 2.2.2 (Ramsey's Theorem, Special Case 2). *For every $k \geq 2$ and $r_1, r_2, \ldots, r_k \geq 2$ there exists a positive integer $R_k(r_1, r_2, \ldots, r_k)$ such that if we k-colour the edges of K_n, with $n \geq R_k(r_1, r_2, \ldots, r_k)$, then there exists either a K_{r_1} in colour 1, or a K_{r_2} in colour 2,…, or a K_{r_k} in colour k.*

This number $R_k(r_1, r_2, \ldots, r_k)$ is called the *Ramsey number* for $r_1, r_2, \ldots, r_k$.

Proof. The proof is again by induction, this time on the number of colours k. The case $k = 2$ is exactly the previous theorem. Now let $k > 2$. We shall show that

$$R_k(r_1, r_2, \ldots, r_k) \leq R_{k-1}(R_2(r_1, r_2), \ldots, r_k).$$

Note that the right-hand side contains only Ramsey numbers for $k-1$ colours and 2 colours, and therefore, by the inductive hypothesis, exists. So proving this will prove the theorem. Call the number on the right-hand side as t. Consider the complete graph K_t and colour its edges with k colours. Now "go colour-blind" and pretend that r_1 and r_2 are the same colour. Thus the graph is now $k-1$-coloured.

By the induction hypothesis, it contains either

(i) a complete graph K_{r_i} in colour i for some $3 \leq i \leq k$, or

(ii) a complete graph $K_{R_2(r_1,r_2)}$ coloured in the "blurred colour".

In case (i), we are done. In case (ii), we recover our sight again and see from the definition of $R_2(r_1, r_2)$ we must have either a complete graph K_{r_1} in colour 1, or a complete graph K_{r_2} in colour 2. In either case the proof is complete. ∎

Definition 2.2.1. *An r-**uniform hypergraph** $\mathscr{H}$ is an ordered pair $(S, \mathcal{E})$ where S is a finite set (the set of **vertices**), and $\mathcal{E} = \mathcal{E}(\mathscr{H}) \subseteq \binom{S}{r}$ (the set of **hyperedges**). Here $\binom{S}{r}$ is the set of all r-element subsets of S, so $\left|\binom{S}{r}\right| = \binom{|S|}{r}$. The $\mathscr{H}$ is a **complete** r-**uniform hypergraph**, denoted by K_n^r, if it is an r-uniform hypergraph and $\mathcal{E}(\mathscr{H}) = \binom{S}{r}$, $n = |S|$.*

Theorem 2.2.3 (Ramsey's Theorem). *For every $r \geq 2$, $k \geq 2$ and $n_1, n_2, \ldots, n_k \geq r$ there exists a positive integer $R_k^r(n_1, n_2, \ldots, n_k)$ such that if we k-colour the hyperedges of K_n^r, with $n \geq R_k^r(n_1, n_2, \ldots, n_k)$, then there exists either a $K_{n_1}^r$ in colour 1, or a $K_{n_2}^r$ in colour 2,..., or a $K_{n_k}^r$ in colour k.*

Proof. Let $k \geq 2$ be fixed. We prove the theorem by double induction:

(i) induction on r;

(ii) induction on $n_1 + n_2 + \cdots + n_k$.

If $r = 2$ and $n_1, n_2, \dots, n_k \geq 2$ are arbitrary, then the conclusion is true; this is Theorem 2.2.2.

Now, let $r > 2$, and assume that the conclusion of the theorem is true for $r-1$ with any $m_1, m_2, \dots, m_k \geq r-1$.

Let $n_1, n_2, \dots, n_k \geq r$ be arbitrary. If $n_1 = n_2 = \cdots = n_k = r$, then $R_k^r(r, r, \dots, r) = r$ satisfies the conditions. Indeed, suppose we have K_n^r for $r \leq n$, and we k-colour the hyperedges. Then there is an $i \in \{1, 2, \dots, k\}$ such that there is a hyperedge $E \in \binom{[n]}{r}$ of colour i. Then E gives us a K_r^r of colour i. Assume now that the theorem is true for fixed k, r and any $m_1, m_2, \dots, m_k \geq r$ such that $m_1 + m_2 + \cdots + m_k < n_1 + n_2 + \cdots + n_k$. We will show that

$$R_k^r(n_1, n_2, \dots, n_k) \leq \underbrace{R_k^{r-1}\left(R_k^r(n_1 - 1, n_2, \dots, n_k), \dots, R_k^r(n_1, n_2, \dots, n_k - 1)\right)}_{R} + 1.$$

This is enough, because the right-hand side only contains Ramsey numbers for $(r-1)$-uniform hypergraphs, and r-uniform hypergraphs with $m_1 + m_2 + \cdots + m_k < n_1 + n_2 + \cdots + n_k$, so by the induction hypothesis, it exists.

To show that $R_k^r(n_1, n_2, \dots, n_k) \leq R+1$, take an K_n^r for some $n \geq R+1$, and k-colour it. We can assume that the set of vertices of K_n^r is $[n] \stackrel{\text{def}}{=} \{1, 2, \dots, n\}$. Remove the last vertex $v = n$ from the set of vertices of K_n^r and take the $(r-1)$-uniform hypergraph on the remaining $n-1$ vertices (which is the set $[n-1]$), K_{n-1}^{r-1}. For a hyperedge $E \in \binom{[n-1]}{r-1}$ of K_{n-1}^{r-1}, let the colour of E be the colour of the hyperedge $E \cup \{v\} \in \binom{[n]}{r}$ in the original colouring of K_n^r.

Now, because $n \geq R+1$, we have $n-1 \geq R$, so by the definition of R, with this colouring of K_{n-1}^{r-1} with k colours, we have for some $i \in \{1, 2, \dots, k\}$, a complete $(r-1)$-uniform subgraph of K_{n-1}^{r-1} on $R_k^r(n_1, n_2, \dots, n_i - 1, \dots, n_k)$ edges, which is of colour i, i.e., a $K_{R_k^r(n_1, n_2, \dots, n_i - 1, \dots, n_k)}^{r-1}$ of colour i. In other words, we have a subset A of $[n-1]$ that is of size $R_k^r(n_1, n_2, \dots, n_i - 1, \dots, n_k)$ and all $(r-1)$-tuples of A are of colour i in the new colouring of K_{n-1}^{r-1}.

Without loss of generality, we can assume that $i = 1$, so $|A| = R_k^r(n_1 - 1, n_2, \dots, n_k)$. By the definition of this Ramsey number, if we consider the original colouring of K_n^r, then either there is a $K_{n_j}^r$ of colour j in A for some $j \in \{2, 3, \dots, k\}$, or there is a $K_{n_1 - 1}^r$ of colour 1 in A. In the first case, we have found a $K_{n_j}^r$ of colour j in K_n^r, for some $j \in \{2, 3, \dots, k\}$, so we are done. In the second case, we have a $K_{n_1 - 1}^r$ of colour 1 in A, i.e., we have a set B of vertices from A, of size $n_1 - 1$, such that the colour of any element of $\binom{B}{r}$ is 1. Consider the set of vertices $B \cup \{v\}$. The size of this set is n_1. Let $E \in \binom{B \cup \{v\}}{r}$ be a hyperedge, and determine the colour of E in the original colouring of K_n^r.

If $E \subseteq B$, then this colour is 1, as we have just seen. If $v \in E$, then $E - \{v\} \in \binom{B}{r-1} \subseteq \binom{A}{r-1}$. Now recall that in the new colouring of K_{n-1}^{r-1}, all $(r-1)$-tuples of A were of colour 1. This means, by the definition of the new colouring of K_{n-1}^{r-1}, that for all $(r-1)$-tuples $E' \subseteq A$, the colour of $E' \cup \{v\}$ in K_n^r was 1. Specifically, for $E' = E - \{v\}$, the colour of $E' \cup \{v\} = E$ in K_n^r was 1.

Thus, we have seen that for all $E \in \binom{B \cup \{v\}}{r}$, the colour of E in K_n^r is 1, and so, $B \cup \{v\}$ gives us a $K_{n_1}^r$ of colour 1 in K_n^r. ■

Theorem 2.2.4 (Infinite Ramsey Theorem)**.** *If we 2-colour the edges of the infinite complete graph $K_{\mathbb{N}}$ with colours red and blue, then there exists either an infinite red complete subgraph or an infinite blue complete subgraph.*

Proof. Identify the vertices of $K_{\mathbb{N}}$ with the set $\mathbb{N}$. Choose $v_1 \in \mathbb{N}$. There are infinitely many edges from v_1, so we can find an infinite set $B_1 \subset \mathbb{N} - \{v_1\}$ such that all edges from v_1 to B_1 have the same colour, c_1.

Now choose $v_2 \in B_1$. There are infinitely many edges from v_2 to vertices in $B_1 - \{v_2\}$, so we can find an infinite set $B_2 \subset B_1 - \{v_2\}$ such that all edges from v_2 to B_2 have the same colour, c_2.

By induction, we obtain a sequence $v_1, v_2, v_3 \ldots$ of distinct elements of $\mathbb{N}$, and a sequence $c_1, c_2, c_3, \ldots$ of colours such that the edge $v_i v_j$ $(i < j)$ has colour c_i. Since we have two colours, we must have $c_{i_1} = c_{i_2} = c_{i_3} = \cdots$ for some infinite subsequence. Then the set of vertices $\{v_{i_1}, v_{i_2}, v_{i_3}, \ldots\}$ induces an infinite monochromatic complete subgraph. ■

2.3 Applications

Theorem 2.3.1 (Schur's Theorem)**.** *For every positive integer k, there exists a positive integer $n = n(k)$, such that for every k-colouring of the integers $\{1, 2, \ldots, n\}$, one of the colour classes contains (not necessarily different) integers i, j and ℓ such that $i + j = \ell$.*

Proof. Let $n = R_k(3, 3, \ldots, 3)$ and k-colour the integers $1, 2, \ldots, n$. Now consider the complete graph K_n with vertices labeled with the integers $1, 2, \ldots, n$. Colour the edges of K_n as follows: an edge ab is coloured with the colour of the vertex corresponding to the integer $|a - b|$.

Now from the definition of $R_k(3, 3, \ldots, 3)$ and Theorem 2.2.2, K_n will definitely contain a triangle in colour red, say. Let $a < b < c$ be the vertices of the triangle. Since the edges of the triangle are red, the vertices $b - a$, $c - a$ and $c - b$ are also red in the original colouring of the integers. It only remains to take $i = c - b$, $j = b - a$ and $\ell = c - a$ to complete the proof. ■

Exercise 2.3.1. *Prove the theorem when i, j and ℓ are different. (Hint: choose $n = R_k(4, 4, \ldots, 4)$.)*

Theorem 2.3.2 (Erdős–Szekeres)**.** *For every $k \geq 3$ there is an $n = n(k)$ such that if n points are placed in the plane, with no three on a line, then there exist k points in convex position, i.e., they constitute a convex k-gon.*

To prove the theorem, we need two lemmas.

Lemma 2.3.3 (Klein)**.** *Any set of five points in the plane, with no three on a line, has a subset of four points that form the vertices of a convex quadrilateral.*

Proof. If the convex hull of the five points is a quadrilateral or a pentagon, we are done.

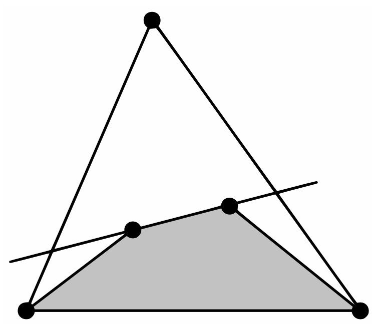

Figure 2.1

If the convex hull is a triangle, the two points left inside the triangle define a line that splits the triangle so that two of the triangle's points are on one side of the line. These two points plus the two interior points form a convex quadrilateral (see Figure 2.1). ■

The statement of the lemma is sometimes called as the "Happy End Problem". The problem was so-named by Erdős when two investigators who first worked on the problem, Eszter Klein and György Szekeres, became engaged and subsequently married.

Lemma 2.3.4. *Suppose that we have k points in the plane, with no three on a line. If any 4 points form the vertices of a convex quadrilateral, then these k points constitute a convex k-gon.*

Proof. Suppose that the set of k points S satisfies the requirements but they do not constitute a convex k-gon. Then their convex hull is an ℓ-gon for some $\ell \leq k-1$, so at least one of the points $s \in S$ is in the interior of this convex hull (see Figure 2.2).

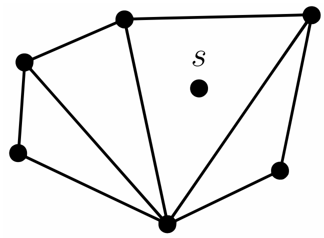

Figure 2.2

Triangulate the convex hull. Then s is in the interior of one of the triangles, otherwise there would be three points in a line. Then the vertices of this triangle are 3 different points from S, and together with s, they form the vertices of a triangle. This contradicts the assumption that any 4 points from S form the vertices of a convex quadrilateral. ■

Proof of Theorem 2.3.2. We show that $n(k) = R_2^4(k, 5)$ is enough. Suppose S is a set of $n \geq n(k)$ points in the plane, with no three on a line. Consider the complete 4-uniform hypergraph K_n^4 whose vertices are the points in S. Colour the hyperedges of K_n^4 as follows. Let $E \in \binom{S}{4}$ be a hyperedge.

(i) If the convex hull of E is a quadrilateral, colour E with colour red;

(ii) if the convex hull of E is a triangle, colour E with colour blue.

By Ramsey's Theorem, and the definition of $n(k)$, either there is

(i) an $S' \in \binom{S}{k}$ such that each 4-tuple in S' is of colour red and therefore convex;

(ii) or there is an $S'' \in \binom{S}{5}$ such that each 4-tuple in S'' is of colour blue and therefore concave.

By Lemma 2.3.3, the case (ii) is impossible. So we must have case (i), i.e., there exists a set $S' \subseteq S$ of k points such that each 4-element subset of S' forms the vertices of a quadrilateral. Lemma 2.3.4 then implies that S' constitutes a convex k-gon. ∎

2.4 Lower bounds for diagonal Ramsey numbers

In the proof of Theorem 2.2.1, we showed that

$$R(k, \ell) \leq \binom{k+\ell-2}{k-1}.$$

For the *diagonal Ramsey numbers* $R_2(k) \stackrel{\text{def}}{=} R(k,k)$, this gives

$$R_2(k) \leq \binom{2k-2}{k-1} < 2^{2k-2}.$$

This 2^{2k-2} is not a rough estimate for the binomial coefficient, since, using Stirling's formula, $\binom{2k-2}{k-1} \sim \frac{2^{2k-2}}{\sqrt{\pi(k-1)}}$ as $k \to +\infty$. We would like to obtain a non-trivial lower bound for $R_2(k)$.

Theorem 2.4.1 (Erdős)**.** *For $k \geq 3$, we have $R_2(k) > 2^{k/2}$.*

Proof. By probabilistic method. In order to show that $R_2(k) > n$, it is sufficient to show that there exists a colouring of the edges of K_n that contains no monochromatic K_k. Consider an edge colouring of K_n in which colours are assigned randomly. Let each edge be coloured independently, and such that for all $e \in E$:

$$\mathsf{P}(e \text{ is coloured red}) = \mathsf{P}(e \text{ is coloured blue}) = \frac{1}{2}.$$

There are $\binom{n}{k}$ copies of K_k in K_n. Let A_i be the event that the ith K_k is monochromatic ($1 \leq i \leq \binom{n}{k}$). Then

$$\mathsf{P}(A_i) = 2 \cdot \left(\frac{1}{2}\right)^{\binom{k}{2}} = 2^{1-\binom{k}{2}},$$

where the leading 2 is because there are two colours from which to choose. Take $n = \lfloor 2^{k/2} \rfloor$, then

$$\begin{aligned}\mathsf{P}(\text{there exists a monochromatic } K_k) = \mathsf{P}(\cup_i A_i) \le \sum_i \mathsf{P}(A_i) = \binom{n}{k} 2^{1-\binom{k}{2}} \\ \le \frac{n^k}{k!} 2^{1-\frac{k(k-1)}{2}} \le \frac{\left(2^{\frac{k}{2}}\right)^k}{k!} 2^{1-\frac{k^2}{2}+\frac{k}{2}} = \frac{2^{1+\frac{k}{2}}}{k!}.\end{aligned}$$

For $k = 3, 4$, direct computation shows that $\frac{2^{1+\frac{k}{2}}}{k!} < 1$. For $k > 4$, we can use the inequality $k! \ge \left(\frac{k}{e}\right)^k$, so $\frac{2^{1+\frac{k}{2}}}{k!} \le 2\left(\frac{\sqrt{2}e}{k}\right)^k < 2\left(\frac{4}{k}\right)^k < 1$. Therefore,

$$\mathsf{P}(\text{there does not exist a monochromatic } K_k) > 0.$$

Hence, there exists a colouring with no monochromatic K_k. ∎

Although Erdős' proof shows that there is a colouring of K_n with no monochromatic K_k if $n = \lfloor 2^{k/2} \rfloor$, it does not give us a concrete colouring. The following theorem is just a polynomial lower bound for $R_2(k)$, nevertheless, it is constructive.

Theorem 2.4.2 (Nagy)**.** *For $k \ge 4$, we have $R_2(k) > \binom{k-1}{3}$.*

Proof. Take $k-1$ points in the plane and consider all the $\binom{k-1}{3}$ possible triples. Define the graph G as follows. Let $V(G)$ be the set of all triples and let G be the complete graph on $V(G)$. Colour an edge xy with colour red if $|x \cap y| = 1$, and with colour blue if $|x \cap y| = 0$ or 2.

Claim. There is no red K_k.

Proof. Take the characteristic vectors of the vertices in $\mathbb{R}^{k-1}$. Suppose, to the contrary, that we have a red K_k. Then there exist characteristic vectors $v_1, v_2, \ldots, v_k$ such that $\langle v_i, v_j \rangle = 1$ for $i \ne j$, and $\langle v_i, v_i \rangle = 3$ for all $i = 1, 2, \ldots, k$. Here we used the usual scalar product in $\mathbb{R}^{k-1}$. Suppose that

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k = 0,$$

for some real numbers $\lambda_1, \lambda_2, \ldots, \lambda_k$. Multiplying both sides with v_i yields

$$\lambda_1 + \lambda_2 + \cdots + 3\lambda_i + \cdots + \lambda_k = 0.$$

Summing up this for all $i = 1, 2, \ldots, k$ and dividing by 3, we obtain

$$\lambda_1 + \lambda_2 + \cdots + \lambda_k = 0,$$

so by subtraction, $2\lambda_i = 0$, i.e., $\lambda_i = 0$ for all $i = 1, 2, \ldots, k$. Therefore, $v_1, v_2, \ldots, v_k$ are k different linearly independent vectors in the $(k-1)$-dimensional space $\mathbb{R}^{k-1}$, which is obviously a contradiction. ■

Claim. There is no blue K_k.

Proof. Take the characteristic vectors of the vertices in $\mathbb{Z}_2^{k-1}$. Suppose, to the contrary, that we have a blue K_k. Then there exist characteristic vectors $u_1, u_2, \ldots, u_k$ such that $\langle u_i, u_j \rangle = 0$ or $2 = 0$ for $i \neq j$, and $\langle u_i, u_i \rangle = 3 = 1$ for all $i = 1, 2, \ldots, k$. Suppose that

$$\mu_1 u_1 + \mu_2 u_2 + \cdots + \mu_k u_k = 0,$$

for some real numbers $\mu_1, \mu_2, \ldots, \mu_k$. Multiplying both sides with u_i yields $\mu_i = 0$ for all $i = 1, 2, \ldots, k$. Therefore, $u_1, u_2, \ldots, u_k$ are k different linearly independent vectors in the $(k-1)$-dimensional space $\mathbb{Z}_2^{k-1}$, which is a contradiction. ■

Hence, there exists a colouring with no monochromatic K_k in $K_{\binom{k-1}{3}}$, therefore, $R_2(k) > \binom{k-1}{3}$. ■

2.5 Van der Waerden's Theorem

Theorem 2.5.1 (Van der Waerden's Theorem)**.** *For any given $k, \ell \geq 2$, there is a positive integer $W(k, \ell)$ such that for any k-colouring of the integers $1, 2, \ldots, W(k, \ell)$, there exists a monochromatic arithmetic progression of length ℓ.*

The infinite version of the theorem is not true. Consider for example, the colouring of the positive integers with colours red and blue. For every $j \geq 0$, colour the numbers in $\mathbb{N} \cap [2^{2j}, 2^{2j+1})$ with colour red and the numbers in $\mathbb{N} \cap [2^{2j+1}, 2^{2j+2})$ with colour blue. Clearly, there is no monochromatic infinite arithmetic progression.

We shall deduce Van der Waerden's Theorem from the following stronger result we do not prove.

Theorem 2.5.2 (Hales–Jewett Theorem)**.** *Consider the n-dimensional cube of edge length ℓ:*

$$Q^n(\ell) \stackrel{\text{def}}{=} \{(x_1, x_2, \ldots, x_n) \mid x_i \in \mathbb{N}, 0 \leq x_i \leq \ell - 1\}.$$

It contains ℓ^n points. We say that the ℓ different points

$$(x_{1_1}, x_{1_2}, \ldots, x_{1_n}), (x_{2_1}, x_{2_2}, \ldots, x_{2_n}), \ldots, (x_{\ell_1}, x_{\ell_2}, \ldots, x_{\ell_n})$$

*form a **diagonal**, if for each $i \in \{1, 2, \ldots, n\}$ we have*

(i) $x_{1_i} = x_{2_i} = \cdots = x_{\ell_i}$ *or*

(ii) $x_{1_i} = 0, x_{2_i} = 1, \ldots, x_{\ell_i} = \ell - 1$ *or*

(iii) $x_{1_i} = \ell - 1, x_{2_i} = \ell - 2, \ldots, x_{\ell_i} = 0.$

Then for any given $k, \ell \geq 2$, there is a positive integer $HJ(k, \ell)$ such that for any k-colouring of the points of $Q^n(\ell)$ ($n \geq HJ(k, \ell)$), there exists a monochromatic diagonal (of length ℓ). ∎

In other words, the higher-dimensional, multi-player, ℓ-in-a-row generalization of game of tic-tac-toe cannot end in a draw, no matter how large ℓ is, no matter how many people k are playing, and no matter which player plays each turn, provided only that it is played on a board of sufficiently high dimension $HJ(k, \ell)$.

It is known that

$$HJ(k, \ell) \geq 2^{2^{\cdot^{\cdot^{\cdot 2^k}}}},$$

where the power-tower contains ℓ copies of 2.

Proposition 2.5.3. *The Hales–Jewett Theorem implies Van der Waerden's Theorem, more precisely, we have*

$$W(k, \ell) \leq \ell^{HJ(k,\ell)}.$$

Proof. Consider the $HJ \stackrel{\text{def}}{=} HJ(k, \ell)$-dimensional cube of edge length ℓ. Define the bijective map $\Phi : Q^{HJ}(\ell) \to \{0, 1, 2, \ldots, \ell^{HJ} - 1\}$ by the formula

$$\Phi((x_1, x_2, \ldots, x_{HJ})) = x_1 + x_2\ell + \cdots + x_{HJ}\ell^{HJ-1}.$$

So we write the integers $0, 1, 2, \ldots, \ell^{HJ} - 1$ in base-ℓ. By Φ, a k-colouring of $Q^{HJ}(\ell)$ corresponds to a k-colouring of the integers $0, 1, 2, \ldots, \ell^{HJ} - 1$, and vice-versa. By the Hales–Jewett Theorem, for any k-colouring of the points of $Q^{HJ}(\ell)$, there exists a monochromatic diagonal. Take such a diagonal

$$(x_{1_1}, x_{1_2}, \ldots, x_{1_{HJ}}), (x_{2_1}, x_{2_2}, \ldots, x_{2_{HJ}}), \ldots, (x_{\ell_1}, x_{\ell_2}, \ldots, x_{\ell_{HJ}}).$$

By the definition of a diagonal, Φ maps the ℓ (different) elements of this diagonal to a monochromatic arithmetic progression of length ℓ. So for any k-colouring of the integers $0, 1, 2, \ldots, \ell^{HJ} - 1$, there exists a monochromatic arithmetic progression of length ℓ, i.e., $W(k, \ell) \leq \ell^{HJ}$ $(= |\{0, 1, 2, \ldots, \ell^{HJ} - 1\}|)$. ■

Theorem 2.5.4. *For $k \geq 2$, we have $W(2, k) > \sqrt{k-1}2^{(k-1)/2}$.*

Proof. By probabilistic method. In order to show that $W(2, k) > n$, it is sufficient to show that there exists a colouring of the integers $1, 2 \ldots, n$ that contains no monochromatic arithmetic progression of length k. Consider a colouring of the integers $1, 2 \ldots, n$ in which colours are assigned randomly. Let each number be coloured independently, and such that for all $i \in \{1, 2, \ldots, n\}$:

$$\mathsf{P}(i \text{ is coloured red}) = \mathsf{P}(i \text{ is coloured blue}) = \frac{1}{2}.$$

Let $N(n, k)$ be the number of arithmetic progression of length k from the set $\{1, 2, \ldots, n\}$. An arithmetic progression is completely determined by its first term a and its difference d. To obtain an arithmetic progression of length k from the set $\{1, 2, \ldots, n\}$, we can choose a at most $n-1$ different ways, and we can choose d at most $\left\lfloor \frac{n-1}{k-1} \right\rfloor$ different ways. In general, not all such pairs give an arithmetic progression of length k from the set $\{1, 2, \ldots, n\}$ (some terms may be strictly larger than n). So $N(n, k) \leq (n-1) \cdot \left\lfloor \frac{n-1}{k-1} \right\rfloor$. Let A_i be the event that the ith arithmetic progression of length k is monochromatic $(1 \leq i \leq N(n, k))$. Then

$$\mathsf{P}(A_i) = 2 \cdot \left(\frac{1}{2}\right)^k = 2^{1-k},$$

where the leading 2 is because there are two colours from which to choose. Take $n = \left\lfloor \sqrt{k-1}2^{(k-1)/2} \right\rfloor$, then

$$\begin{aligned}
&\mathsf{P}(\text{there exists a monochromatic arithmetic progression of length } k) \\
&= \mathsf{P}\left(\cup_i A_i\right) \leq \sum_i \mathsf{P}\left(A_i\right) \leq (n-1) \cdot \left\lfloor \frac{n-1}{k-1} \right\rfloor \cdot 2^{1-k} \\
&< \frac{n^2}{k-1} 2^{1-k} \leq \frac{\left(\sqrt{k-1}2^{(k-1)/2}\right)^2}{k-1} 2^{1-k} = 1.
\end{aligned}$$

Therefore, the probability that there is no monochromatic arithmetic progression of length k is larger than 0. Hence, there exists a colouring with no monochromatic arithmetic progression of length k. ∎

Exercise 2.5.1. *Prove that* $W(2,3) = 9$.

Chapter 3

Set systems and extremal graph theory

3.1 Extremal set families

Theorem 3.1.1 (Fisher's inequality). *Let $A_1, A_2, \ldots, A_m$ be distinct subsets of a set S such that $|A_i \cap A_j| = k$ for some fixed $1 \leq k \leq |S|$ and every $i \neq j$. Then $m \leq |S|$.*

Proof. Let $v_1, v_2, \ldots, v_m$ be the characteristic vectors of $A_1, A_2, \ldots, A_m$. Let $n \stackrel{\text{def}}{=} |S|$. Since $\dim v_i = n$ for all $1 \leq i \leq m$, it is enough to show that these vectors are linearly independent. Assume to the contrary, that

$$\sum_{i=1}^{m} \lambda_i v_i = 0,$$

with not all coefficients being zero. We have $\langle v_i, v_j \rangle = k$ for $i \neq j$, and $\langle v_i, v_i \rangle = |A_i|$. Therefore,

$$\begin{aligned} 0 = \left\langle \sum_{i=1}^{m} \lambda_i v_i, \sum_{i=1}^{m} \lambda_i v_i \right\rangle &= \sum_{i=1}^{m} \lambda_i^2 \langle v_i, v_i \rangle + \sum_{1 \leq i \neq j \leq m} \lambda_i \lambda_j \langle v_i, v_j \rangle \\ &= \sum_{i=1}^{m} \lambda_i^2 |A_i| + \sum_{1 \leq i \neq j \leq m} \lambda_i \lambda_j k = \sum_{i=1}^{m} \lambda_i^2 (|A_i| - k) + k \left(\sum_{i=1}^{m} \lambda_i \right)^2 . \end{aligned}$$

Clearly, $|A_i| \geq k$ for all $1 \leq i \leq m$ and $|A_i| = k$ for at most one i, since otherwise the intersection condition would not be satisfied. But then the right-hand side is greater than 0 (because the last sum can vanish only if at least two of the coefficients λ_i are non-zero), a contradiction. ■

Let S be a set with n elements. In what follows, $\mathscr{F}$ denotes a family of subsets of S, i.e., $\mathscr{F} \subseteq 2^S \stackrel{\text{def}}{=} \mathcal{P}(S)$.

Proposition 3.1.2. *Suppose that for all $X, Y \in \mathscr{F}$, $X \cap Y \neq \emptyset$. Then $|\mathscr{F}| \leq 2^{n-1}$.*

Proof. For any $X \in 2^S$ consider the pair $(X, \overline{X})$, where $\overline{X}$ is the complement of X. There are 2^{n-1} such pairs and, by assumption, $\mathscr{F}$ can contain at most one set from each pair. ■

Proposition 3.1.3. *Suppose that for all $X, Y \in \mathscr{F}$, $X \cup Y \neq S$. Then $|\mathscr{F}| \leq 2^{n-1}$.*

Proof. Consider $\overline{\mathscr{F}} \stackrel{\text{def}}{=} \left\{ \overline{X} \mid X \in \mathscr{F} \right\}$. Clearly, $|\overline{\mathscr{F}}| = |\mathscr{F}|$, and by Proposition 3.1.2, $|\overline{\mathscr{F}}| \leq 2^{n-1}$. ■

Proposition 3.1.4. *Suppose that for all $X, Y \in \mathscr{F}$, $X \cap Y \neq \emptyset$ and $X \cup Y \neq S$. Then $|\mathscr{F}| \leq 2^{n-2}$.*

The proof is based on the following lemma, whose proof is left as an exercise.

Lemma 3.1.5. *Let $\mathscr{U}$ be an **up-set family**, i.e., for all $X \in \mathscr{U}$, $X \subseteq X'$ implies $X' \in \mathscr{U}$. Similarly, let $\mathscr{D}$ be an **down-set family**, i.e., for all $X \in \mathscr{D}$, $X' \subseteq X$ implies $X' \in \mathscr{D}$. Suppose that $\mathscr{U}, \mathscr{D} \subseteq 2^S$, $|S| = n$. Then*

$$|\mathscr{U} \cap \mathscr{D}| \leq \frac{|\mathscr{U}||\mathscr{D}|}{2^n}.$$

■

Exercise 3.1.1. *Prove the lemma above. (Hint: by induction on n.)*

Proof of Proposition 3.1.4. Define

$$\widetilde{\mathscr{F}} \stackrel{\text{def}}{=} \{X \mid \text{there exists } X_0 \subseteq X, X_0 \in \mathscr{F}\};$$

$$\widehat{\mathscr{F}} \stackrel{\text{def}}{=} \{X \mid \text{there exists } X_0 \supseteq X, X_0 \in \mathscr{F}\}.$$

By Proposition 3.1.2, $|\widetilde{\mathscr{F}}| \leq 2^{n-1}$. Similarly, by Proposition 3.1.3, $|\widehat{\mathscr{F}}| \leq 2^{n-1}$. Clearly, $\widetilde{\mathscr{F}}$ is an up-set family and $\widehat{\mathscr{F}}$ is a down-set family. Using the lemma above, we obtain

$$|\mathscr{F}| \leq |\widetilde{\mathscr{F}} \cap \widehat{\mathscr{F}}| \leq \frac{|\widetilde{\mathscr{F}}||\widehat{\mathscr{F}}|}{2^n} \leq \frac{2^{n-1}2^{n-1}}{2^n} = 2^{n-2},$$

which completes the proof. ■

Theorem 3.1.6 (Sperner's Theorem)**.** *Suppose that for all $X, Y \in \mathscr{F}$, $X \not\subseteq Y$. Then $|\mathscr{F}| \leq \binom{n}{\lfloor n/2 \rfloor}$. This bound is sharp as the family of all $\lfloor n/2 \rfloor$-element subsets of S shows.*

Note that $\mathscr{F}$ is precisely an antichain.

Proof. This proof is due to Lubell, the original one was more complicated. Let $s_1, s_2, \ldots, s_n$ be a fixed ordering of the elements of S. We will say that a permutation π of the elements of S begins with $X \in \mathscr{F}$ if $X = \left\{s_{\pi(1)}, s_{\pi(2)}, \ldots, s_{\pi(|X|)}\right\}$. The number of permutations beginning with X must be $|X|!\,(n-|X|)!$. Also, no permutation can begin with two different sets in $\mathscr{F}$, since one of these sets would contain the other; therefore permutations beginning with different sets in $\mathscr{F}$ are distinct. Thus, the number of permutations π begins by some $X \in \mathscr{F}$ is

$$\sum_{X \in \mathscr{F}} |X|!\,(n-|X|)!,$$

which cannot be larger than the total number of permutations, $n!$, therefore,

$$\sum_{X \in \mathscr{F}} |X|!\,(n-|X|)! \leq n!.$$

Dividing both sides by $n!$, we have

$$\sum_{X \in \mathscr{F}} \frac{1}{\binom{n}{|X|}} = \sum_{X \in \mathscr{F}} \frac{|X|!\,(n-|X|)!}{n!} \leq 1.$$

Since $\binom{n}{|X|} \leq \binom{n}{\lfloor n/2 \rfloor}$, we obtain

$$|\mathscr{F}| \frac{1}{\binom{n}{\lfloor n/2 \rfloor}} = \sum_{X \in \mathscr{F}} \frac{1}{\binom{n}{\lfloor n/2 \rfloor}} \leq \sum_{X \in \mathscr{F}} \frac{1}{\binom{n}{|X|}} \leq 1,$$

which is equivalent to the statement. ■

Theorem 3.1.7 (Erdős–Ko–Rado Theorem)**.** *Suppose that for all* $X, Y \in \mathscr{F}$, $X \not\subseteq Y$, $X \cap Y \neq \emptyset$ *and* $|X|, |Y| \leq k \leq \frac{n}{2}$. *Then* $|\mathscr{F}| \leq \binom{n-1}{k-1}$. *This bound is sharp: fix an element of* S *and consider all the* k*-element subsets of* S *that contain it.*

Proof. The original proof of 1961 used induction on n. In 1972, Katona gave the following short proof using double counting. We may assume that every element of $\mathscr{F}$ has size k. Arrange the elements of S in any cyclic order, and consider the sets from $\mathscr{F}$ that form intervals of length k within this cyclic order. However, it is not possible for all of the intervals of the cyclic order to belong to $\mathscr{F}$, because some pairs of them are disjoint. We claim that at most k of the intervals for a single cyclic order may belong to $\mathscr{F}$. To see this, note that if $(x_1, x_2, \ldots, x_k)$ is one of these intervals in $\mathscr{F}$, then every other interval of the same cyclic order that belongs to $\mathscr{F}$ separates x_i and x_{i+1} for some i (that is, it contains precisely one of these two elements). The two intervals that separate these elements are either disjoint or one contains the other, so at most one of them can belong to $\mathscr{F}$. Thus, the number of intervals in $\mathscr{F}$ is one plus the number of separated pairs, which is at most k.

Based on this observation, we may count the number of pairs (X, π), where $X \in \mathscr{F}$ and π is a cyclic order for which X is an interval, in two ways. First, for each set X one may generate π by choosing one of $k!$ permutations of X and $(n-k)!$ permutations of the remaining elements, showing that the number of pairs is $|\mathscr{F}|k!(n-k)!$. And second, there are $(n-1)!$ cyclic orders, each of which has at most k intervals of $\mathscr{F}$, so the number of pairs is at most $k(n-1)!$. Combining these two counts gives the inequality

$$|\mathscr{F}|\, k!\, (n-k)! \leq k\, (n-1)!$$

and dividing both sides by $k!(n-k)!$ gives the result

$$|\mathscr{F}| \leq \frac{k\,(n-1)!}{k!\,(n-k)!} = \frac{(n-1)!}{(k-1)!\,(n-k)!} = \binom{n-1}{k-1}.$$

■

3.2 Extremal graph theory

Let H be a graph and n be a positive integer. Define

$$\operatorname{ex}(n,H) \overset{\text{def}}{=} \max_{|V(G)|=n} \left\{ e(G) \mid H \not\subseteq G \right\},$$

where $e(G) = |E(G)|$. So we are looking for a graph G on n vertices with maximal number of edges such that G does not contain the graph H as a subgraph.

Theorem 3.2.1 (Erdős)**.** *We have*

$$\operatorname{ex}(n,\mathcal{C}_4) = \frac{1}{2}n^{3/2} + o(n^{3/2}),$$

as $n \to +\infty$.

Proof. $\boxed{\leq}$ A $\mathcal{C}_4$ can be viewed as two paths of length 2 with the same start and endpoint. Let G be a $\mathcal{C}_4$-free graph. Denote by $\#P_3$ the number of paths of length 2 in G (such a path consists of 3 vertices). Since the graph does not contain any $\mathcal{C}_4$, a start and an endpoint uniquely determine a path of length 2, so $\#P_3 \leq \binom{n}{2}$. On the other hand, if $x \in V(G)$ is the middle point of the path, then we can choose its start and endpoint in $\binom{d(x)}{2}$ ways. Therefore,

$$\binom{n}{2} \geq \#P_3 = \sum_{x\in V(G)} \binom{d\left(x\right)}{2}.$$

Hence, we have

$$\sum_{x\in V(G)} \frac{d^2\left(x\right)}{2} - \sum_{x\in V(G)} \frac{d\left(x\right)}{2} \leq \frac{n\left(n-1\right)}{2},$$
$$\sum_{x\in V(G)} d^2\left(x\right) - \underbrace{\sum_{x\in V(G)} d\left(x\right)}_{2e\overset{\text{def}}{=}2e(G)} \leq n\left(n-1\right).$$

Using the inequality between the arithmetic and quadratic mean, we obtain

$$\sum_{x\in V(G)} d^2\left(x\right) = n\left(\sqrt{\frac{1}{n}\sum_{x\in V(G)} d^2\left(x\right)}\right)^2 \geq n\left(\frac{1}{n}\sum_{x\in V(G)} d\left(x\right)\right)^2 = n\left(\frac{2e}{n}\right)^2 = \frac{4e^2}{n}$$

So by the above,

$$\frac{4e^2}{n} - 2e \leq n\left(n-1\right).$$

Solving for e, we deduce the required result:

$$e \le \frac{n + n\sqrt{4n-3}}{4} = \frac{n^{3/2}}{2}\sqrt{1-\frac{3}{4n}} + \frac{n}{4} = \frac{n^{3/2}}{2}\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right) + \frac{n}{4}$$
$$= \frac{n^{3/2}}{2} + \mathcal{O}\left(\sqrt{n}\right) + \frac{n}{4} = \frac{n^{3/2}}{2} + \mathcal{O}\left(n\right) = \frac{n^{3/2}}{2} + o\left(n^{3/2}\right),$$

as $n \to +\infty$.

$\boxed{\ge}$ We prove the bound for n's of the form p^2+p+1, where $p>2$ is a prime. Let $p>2$ be a fixed prime. Consider the 3-tuples formed by elements of $\mathbb{F}_p$. We have p^3-1 non-zero ($\neq (0,0,0)$) such 3-tuples. Define the equivalence relation $\sim$ on them as follows

$$(x,y,z) \sim (x',y',z') \Leftrightarrow \text{ there is an } a \in \mathbb{F}_p\backslash\{0\} \text{ such that } (x',y',z') = (ax,ay,az).$$

Each equivalence class contains $p-1$ elements, and the number of equivalence classes is

$$\frac{p^3-1}{p-1} = p^2+p+1.$$

We call the equivalence classes points. So we have p^2+p+1 different points.

Exactly the same way, we define the lines $[a,b,c]$. Again, there are p^2+p+1 different lines. We say that the line $[a,b,c]$ contains the point (x,y,z) if and only if $ax+by+cz=0$ in $\mathbb{F}_p$. It is easy to see that this concept is well-defined.

We shall count the number of points in a line. Consider a line $[a,b,c]$. We may assume that $a\neq 0$. Suppose that $ax+by+cz=0$. Any fixed $(y,z)\neq(0,0)$ gives a unique solution x, so we have p^2-1 solutions. But some of these points (x,y,z) are equivalent. Thus, there are

$$\frac{p^2-1}{p-1} = p+1$$

points in a line.

We show that every two different lines intersects at one point. Let $[a,b,c]$ and $[a',b',c']$ two different lines. We may assume that $(b',c')\neq(\lambda b,\lambda c)$ for any $\lambda \in \mathbb{F}_p \setminus \{0\}$. Suppose that

$$ax+by+cz=0 \quad \text{and} \quad a'x+b'y+c'z=0.$$

The $x\neq 0$ can be chosen in $p-1$ different ways. For any x the above gives a system of equations for y and z with unique solutions. Hence, we have $p-1$ solutions, which clearly yield equivalent points. Thus, every two different lines intersects at one point.

Note that there is a duality between points and lines. Consider a graph G on $n = p^2 + p + 1$ points. We view each vertex as a point (x, y, z) and, at the same time, as a line $[x, y, z]$. We draw an edge between two vertices if and only if one contains the other, as a line contains a point. This graph G does not contain any $\mathcal{C}_4$, otherwise there would be two different lines intersecting in more than one point. Each of the n lines contains $p + 1$ points, maybe one of them is itself as a point. So each vertex has degree $p + 1$ or p. Since $2e = \sum_{x \in V(G)} d(x)$, we obtain

$$np \leq 2e \leq n\,(p+1) \quad \text{or} \quad \frac{p}{\sqrt{n}} \leq \frac{e}{n^{3/2}/2} \leq \frac{p+1}{\sqrt{n}} = \frac{p}{\sqrt{n}} + \frac{1}{\sqrt{n}}.$$

Therefore,

$$e = \frac{n^{3/2}}{2}\left(\frac{p}{\sqrt{n}} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right)$$

as $n \to +\infty$ (or equivalently $p \to +\infty$). But

$$\frac{p}{\sqrt{n}} = \frac{p}{\sqrt{p^2 + p + 1}} = \frac{1}{\sqrt{1 + \frac{1}{p} + \frac{1}{p^2}}} = 1 + \mathcal{O}\left(\frac{1}{p}\right) = 1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$$

as $n \to +\infty$. Thus,

$$\begin{aligned} e &= \frac{n^{3/2}}{2}\left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right) = \frac{n^{3/2}}{2}\left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right) \\ &= \frac{n^{3/2}}{2} + \mathcal{O}\,(n) = \frac{n^{3/2}}{2} + o\left(n^{3/2}\right) \end{aligned}$$

as $n \to +\infty$. ∎

Definition 3.2.1. *For every $r \geq 2$, define the **Turán number** as*

$$t_r(n) \stackrel{\text{def}}{=} \operatorname{ex}(n, K_{r+1}).$$

*The **Turán graph** $T_r(n)$ is a graph formed by partitioning a set of n vertices into r subsets, with sizes as equal as possible, and connecting two vertices by an edge whenever they belong to different subsets. The graph will have $(n \bmod r)$ subsets of size $\lceil n/r \rceil$, and $r - (n \bmod r)$ subsets of size $\lfloor n/r \rfloor$. That is, it is a **complete***

*r-**partite graph** (i.e., a graph with chromatic number r and maximum possible number of edges)*

$$K_{\lceil n/r\rceil,\lceil n/r\rceil,\ldots,\lfloor n/r\rfloor,\lfloor n/r\rfloor}.$$

Each vertex has degree either $n-\lceil n/r\rceil$ or $n-\lfloor n/r\rfloor$.

Theorem 3.2.2 (Turán's Theorem)**.** *For any fixed $r\geq 2$, the graph $T_r(n)$ is the only graph G that does not contain K_{r+1} as a subgraph and for which $e(G)=t_r(n)$.*

First, we prove Turán's Theorem in the case $r=2$.

Proof. First, we prove that for any K_3-free graph G on n vertices, we have $e(G)\leq\lfloor n^2/4\rfloor$. We proceed by induction on $V(G)=n$. For $n=2,3$ the statement is obviously true. Suppose that the statement is true for $n\geq 2$ and we proceed to $n+2$ (this is enough, because the statement is true for both $n=2$ and $n=3$). Let G be a K_3-free graph with $n+2$ vertices. Choose two vertices x and y such that $xy\in E(G)$. Since G is triangle-free, x and y do not have common neighbours, so there are at most n edges that have x or y as an endpoint and which are different from xy. Hence, by the induction hypothesis,

$$e(G)\leq 1+n+e(G-xy)\leq 1+n+\left\lfloor\frac{n^2}{4}\right\rfloor=\left\lfloor\frac{n^2+4n+4}{4}\right\rfloor=\left\lfloor\frac{(n+2)^2}{4}\right\rfloor.$$

The Turán graph $K_{\lfloor n/2\rfloor,\lceil n/2\rceil}$ is triangle-free and has $\lfloor n^2/4\rfloor$ many edges. The induction above shows that it is the unique such graph. ■

Exercise 3.2.1. *Let G be a graph consisting of two triangles sharing a common edge. Prove that*

$$\operatorname{ex}(n,G)=\left\lfloor\frac{n^2}{4}\right\rfloor.$$

Exercise 3.2.2. *Let G be a graph consisting of two triangles sharing a common vertex. Prove that*

$$\operatorname{ex}(n,G)=\left\lfloor\frac{n^2}{4}\right\rfloor+1.$$

Turán's Theorem is the consequence of the following lemma.

Lemma 3.2.3. *Let G be a K_r-free graph. Then there exists a graph G' such that*

(i) $V(G')=V(G)$;

(ii) $d_{G'}(x) \geq d_G(x)$ *for all* $x \in V$*;*

(iii) G' *is* $(r-1)$*-partite, i.e.,* $\chi(G') = r - 1$.

Proof. The method of the proof is called *the method of symmetrization.* It is due to Zykov. We proceed by induction on r. For $r = 2$, the statement is obvious. Suppose that the statement is true for $r-1$ and we proceed to r. Let G be a K_r-free graph. Let $x \in V(G)$ be a vertex of maximum degree. Let N be the set of neighbours of x, and M be the set of non-neighbours of x. So $V(G) = \{x\} \dot{\cup} N \dot{\cup} M$. Let G_0 be the graph induced by N. Since G is K_r-free, we must have that G_0 is K_{r-1}-free. By the induction hypothesis, there is an $(r-2)$-partite graph G_0' on the vertices N, such that $d_{G_0'}(v) \geq d_{G_0}(v)$ for all $v \in N$.

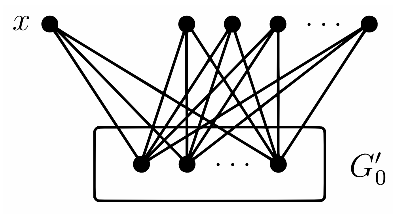

Figure 3.1. *The graph* G'.

We form a new graph $G' = (V, E')$. There are three type of vertices: the x itself and the sets N and M. On the vertex set N we form the graph G_0'. Then we introduce an edge between every vertex in $V(G_0')$ and every vertex in $V \setminus N$ (this is x and M). This is the graph G' (see Figure 3.1), and it satisfies the requirements of the lemma. ■

Remark. *It is not hard to see that if there are edges between the vertices in* M *then the new graph* G' *has strictly larger number of edges. Hence, if* $e(G) = t_{r-1}(n)$*, then there can not be edges between the vertices in* M*, otherwise we would have* $e(G') > e(G)$*. Also, since* $e(G)$ *is maximal, we must have* $G' = G$.

Proof of Turán's Theorem. Consider any a K_r-free graph G on n vertices. Form the graph G' whose existence is stated in the previous lemma. We have $e(G) \leq e(G') \leq t_{r-1}(n)$ (G' is $(r-1)$-partite, so it does not contain a K_r). If $e(G')$ is not a complete $(r-1)$-partite graph then we can add new edges to it without creating a K_r, until we get a complete $(r-1)$-partite graph.

We claim that the number of edges in a complete $(r-1)$-partite graph is maximized when the size of the parts differs by at most one. If G' is a complete $(r-1)$-partite graph with parts X and Y and $|X| > |Y|+1$, then we can increase the number of edges in G' by moving a vertex from part X to part Y. By moving a vertex from part X to part Y, the graph loses $|Y|$ edges, but gains $|X|-1$ edges. Thus, it gains at least $|X|-1-|Y| \geq 1$ edge.

It follows that $e(G') \leq e(T_{r-1}(n))$, and therefore, $e(G) \leq e(T_{r-1}(n))$ for all K_r-free graph G on n vertices. Consequently, $e(T_{r-1}(n)) = t_{r-1}(n)$.

It remains to prove that $T_{r-1}(n)$ is the unique such graph. Suppose that G is an extremum. Form the graph G' whose existence is stated in the previous lemma. We have $t_{r-1}(n) = e(G) \leq e(G') \leq t_{r-1}(n)$, i.e., $e(G') = t_{r-1}(n)$. From the previous argument, it follows that G' must be the Turán graph $T_{r-1}(n)$. According to our remark, $G = G'$, i.e., $G = T_{r-1}(n)$. ■

Definition 3.2.2. *Let G be an arbitrary graph with vertices $v_1, v_2, \ldots, v_n$. We construct a new graph $\widetilde{G}$, by replacing each v_i with an independent set of vertices V_i $(1 \leq i \leq n)$. We draw an edge between each vertex in V_i and each vertex in V_j if and only if $v_iv_j \in E(G)$. This new graph $\widetilde{G}$ is called a **blow-up** of the original graph G.*

For example, $T_r(n)$ is a blow-up of K_r.

Exercise 3.2.3. *Let G be a blow-up of a graph H. Prove that there is a maximal sized bipartite subgraph G_0 of G which is a blow-up of a bipartite subgraph H_0 of H.*

Exercise 3.2.4. *With the notations of the previous exercise, find a G and H such that G_0 is not a blow-up of a maximal bipartite subgraph H_0 of H.*

Conjecture 3.2.4 (Erdős). *Let G be a triangle-free graph. Then there exists a subgraph G_0 of G such that*

(i) G_0 *is bipartite;*

(ii) $e(G_0) \geq e(G) - \frac{n^2}{25}$,

and an appropriate blow-up of a $\mathcal{C}_5$ shows that this is sharp.

Exercise 3.2.5. *Prove that Erdős's conjecture is true if $e(G) \geq \frac{n^2}{5}$.*

Theorem 3.2.5. *For every graph G, there is a subgraph G_0 of G such that*

(i) G_0 *is bipartite;*

(ii) $e(G_0) \geq \frac{1}{2}e(G)$.

Proof. Divide the vertices of the graph G into two arbitrary non-empty classes A and B. Consider a vertex v in A (or B). If v has more neighbours in A (or B) than in B (or B), then put it into the class B (or A). Do this with every vertex of G. Finally, every vertex in A will have at least as many neighbours in B as in A; and every vertex in B will have at least as many neighbours in A as in B. Therefore, if we delete the edges between every two vertices in A and every two vertices in B, we get a subgraph G_0 which is bipartite and has at least half as many edges as the original graph G. ■

Theorem 3.2.6 (Erdős–Stone–Simonovits)**.** *Let H be a fixed graph such that $\chi(H) = r$. Then*

$$\mathrm{ex}(n, H) = t_{r-1}(n) + o(n^2),$$

as $n \to +\infty$.

The formula in the theorem is indeed an asymptotic one since

$$t_{r-1}(n) \sim \frac{r-2}{2r-2}n^2$$

as $n \to +\infty$. We can extend $t_r(n)$ to $r = 1$ as $t_1(n) = 0$. In this case we have the weaker result $\mathrm{ex}(n, H) = o(n^2)$ for any bipartite graph H.

We shall give a sketch of the proof using Szemerédi's Regularity Lemma.

Definition 3.2.3. *Let G be a graph, $X, Y \subseteq V(G)$ such that $X \cap Y = \emptyset$ and $|X| = |Y|$. Let $\varepsilon > 0$ be fixed. The pair (X, Y) is called an ε-**regular pair** if for any $X_0 \subseteq X$, $|X_0| \geq \varepsilon|X|$ and for any $Y_0 \subseteq Y$, $|Y_0| \geq \varepsilon|Y|$, we have*

$$d(X, Y) - \varepsilon \leq d(X_0, Y_0) \leq d(X, Y) + \varepsilon.$$

Here

$$d(X, Y) \stackrel{\text{def}}{=} \frac{\textit{number of edges between } X \textit{ and } Y}{|X||Y|}$$

and similarly for $d(X_0, Y_0)$.

Theorem 3.2.7 (Szemerédi's Regularity Lemma)**.** *For every $\varepsilon > 0$ and $M \in \mathbb{Z}^+$, there are positive integers $N_0(\varepsilon, M)$ and $N(\varepsilon, M)$ such that for any graph G for which $|V(G)| > N_0(\varepsilon, M) \gg N(\varepsilon, M)$ ($\gg$ means much bigger), there is a p such that G has an ε-regular partition of $N(\varepsilon, M) \geq p > M$ classes, i.e., we have a partition of the vertices into sets $S_1, S_2, \ldots, S_p$ with equal size and to a small set with size at most $\varepsilon|V(G)|$ such that all pairs (S_i, S_j) are ε-regular with the exception of $\varepsilon\binom{p}{2}$ pairs.* ■

Proposition 3.2.8. *Let (X, Y) be an ε-regular pair. Then the number of vertices x belonging to X such that $d(x, Y) < d(X, Y) - \varepsilon$ is less than $\varepsilon|X|$.* ■

Sketch of the proof of the Erdős–Stone–Simonovits Theorem. The theorem states that with fixed H, $\chi(H) = r$, for every $\varepsilon > 0$ there is an $n_0 \in \mathbb{N}$ such that for $n \geq n_0$, we have

$$|\mathrm{ex}\,(n, H) - t_{r-1}\,(n)| \leq \varepsilon n^2.$$

Suppose, to the contrary, that there is an $\varepsilon > 0$ and a sequence of graphs $G_1, G_2, \ldots$ with $e(G_1) < e(G_2) < \cdots$ such that H is not a subgraph of any G_i and

$$e\,(G_i) \geq t_{r-1}\,(n_i) + \varepsilon n_i^2,$$

where $n_i = |V(G_i)|$. Choose $G \stackrel{\text{def}}{=} G_i$ such that G_i has an ε-regular partition with p classes. Let $S_1, S_2, \ldots, S_p$ be the classes. Form a (reduced) graph R on p vertices $v_1, v_2, \ldots, v_p$ as follows. Let $v_iv_j \in E(R)$ if and only if (S_i, S_j) is ε-regular and $d(S_i, S_j) \geq \varepsilon$.

In R, we have more edges, consequently, it contains a K_r. Dividing the classes S_i into t equal parts and using Proposition 3.2.8, it can be shown that in G there is a blow-up of K_r: $K_{t,t,\ldots,t}$, a complete r-partite graph. If $t = |V(H)|$, then $K_{t,t,\ldots,t}$ contains H. Therefore G contains H, which is a contradiction. ■

Theorem 3.2.9. *For every $r \geq 3$, there is a $c(r) > 0$ such that if $e(G) \geq c(r)n$, then G contains a topological K_r as a subgraph.*

The theorem was proved with $c(r) = 2^{\binom{r}{2}}$ by Mader. Komlós and Szemerédi improved the constant to $c(r) = 256r^2$. The theorem is a consequence of the following stronger assertion with $H = K_r$.

Lemma 3.2.10. *Let H be a connected graph with m edges and r vertices. If the graph G has at least $2^m|V(G)|$ edges, then a subdivision of H is contained in G.*

Proof. We proceed by induction on m. Since H is connected, $m = r-1$ is the first case (so H is a tree). Suppose that G is a graph on n vertices with $e(G) \geq 2^{r-1}n$. Then its average degree is

$$\overline{d}\,(G) \overset{\text{def}}{=} \frac{1}{n} \sum_{x \in V(G)} d\,(x) = \frac{2}{n} e\,(G) \geq 2^r.$$

We claim that there is a G_0 subgraph of G such that

$$\delta\,(G_0) \overset{\text{def}}{=} \min_{x \in V(G_0)} d\,(x) \geq \frac{\overline{d}\,(G)}{2} \geq 2^{r-1}.$$

Indeed, suppose that G contains a vertex v such that $d\,(v) < \overline{d}\,(G)\,/2$. If we delete this v form G, the average degree of the resulting graph is

$$\overline{d}\,(G-v) = \frac{1}{n-1} \sum_{x \in V(G-v)} d\,(x) = \frac{n}{n-1}\overline{d}\,(G) - \frac{d\,(v)}{n-1} > \frac{n-1/2}{n-1}\overline{d}\,(G) > \overline{d}\,(G)\,.$$

Therefore, the average degree in the resulting graph did not decrease. We repeat this procedure until the resulting subgraph of G satisfies $\delta \geq \frac{\overline{d}(G)}{2}$.

Hence, there is a G_0 subgraph of G such that $\delta\,(G_0) \geq 2^{r-1} \geq r-1$. Using this bound we can build up the tree H in G_0 using a greedy algorithm.

Now, suppose that the statement in true for $m-1 \geq r-1$ and we proceed to m. Thus, H is a connected graph with $m \geq r$ edges. Since it is not a tree, there is an $e \in E(H)$ such that the graph $H - e$ is still connected. Choose an edge e like that and delete it from the graph H. For an $A \subseteq V(G)$, denote by $G\,[A]$ the subgraph of G which is induced by A. Take an $S \subseteq V(G)$ such that

(i) the $G\,[S]$ is connected;

(ii) if S is contracted to one vertex then the average degree of the resulting graph is still at least 2^{m+1} (note that this is the lower bound for $\overline{d}(G)$ too);

(iii) the S is maximal sized with the properties (i) and (ii).

Denote by $N(S)$ the neighbours of S, and consider the induced subgraph $G\,[N(S)]$. We claim that

$$d_{G[N(S)]}(x) \geq 2^m$$

for each $x \in V(G[N(S)])$. Indeed, if $d_{G[N(S)]}(x) < 2^m$ for some $x \in V(G[N(S)])$, then add x to S. We loose at most 2^m edges, but just one vertex after contracting $S \cup \{x\}$ to a vertex. Hence, the average degree of the resulting graph did not decrease. Thus, the set $S \cup \{x\}$ satisfies the properties (i) and (ii), and it is bigger than S. This is a contradiction. Using our claim, we obtain

$$e(G[N(S)]) = \frac{1}{2} \sum_{x \in V(G[N(S)])} d_{G[N(S)]}(x) \geq \frac{1}{2} \sum_{x \in V(G[N(S)])} 2^m = 2^{m-1} |N(S)| .$$

By the induction hypothesis, $G[N(S)]$ contains a subdivision of $H - e$. Find the endpoints of e in H in the subdivision of $H - e$. These are in $N(S)$, so they can be connected to vertices in S. If they can be connected to the same vertex in S we get a subdivision of H in G. If they can only be connected to two different vertices, then those vertices can be connected by a path in S (since S is connected). In this way we get a subdivision of H in G. ■

For the first few K_r, we can compute the exact value of the extremum.

Proposition 3.2.11. *We have* $\mathrm{ex}(n, \text{topological } K_3) = n - 1$ *for* $n \geq 3$.

Proof. A topological K_3 is a cycle. The cycle-free graph on n vertices that has maximal number of edges is a connected tree, which has $n - 1$ edges. ■

Proposition 3.2.12. *We have* $\mathrm{ex}(n, \text{topological } K_4) = 2n - 3$ *for* $n \geq 4$.

Proof. $\boxed{\geq}$ For $n = 4$ take K_4 and delete an edge from it. For $n \geq 5$ consider $K_{2,n-2}$ and draw an edge between the two vertices in the first class.

$\boxed{\leq}$ The proof follows from the following lemma:

Lemma 3.2.13. *If $d(x) \geq 3$ for all $x \in V(G)$ then G contains a topological K_4 as a subgraph.* ■

Suppose that G is a graph on n vertices such that $e(G) \geq 2n - 2$. If $d(x) \geq 3$ for all $x \in V(G)$ then $e = \sum_{x \in V(G)} d(x) \geq 3n > 2n - 3$ and by the previous lemma, G contains a topological K_4 as a subgraph. Suppose that there is an $x \in V(G)$ such that $d(x) < 3$ and delete this vertex. Then $|V(G - x)| = n - 1$ and $e(G - x) \geq e(G) - 2 \geq 2n - 4 = 2(n - 1) - 2$. Using an induction argument on the number of vertices, we have that $G - x$ contains a topological K_4, and hence so does G.

To prove the lemma above, we prove the following one that can be proved by induction.

Lemma 3.2.14. *If $d(x) \geq 3$ for all but one $x \in V(G)$ then G contains a topological K_4 as a subgraph.*

Proof. We proceed by induction on $n = |V(G)|$. For $n = 4$, the statement is true. Suppose that it is true for $n - 1 \geq 4$ and we proceed to $n > 4$. Let x be a vertex with minimal degree. If $d(x) = 0, 1$ or 4, we delete x and we can use the induction hypothesis.

Consider the case $d(x) = 2$. Denote by y and z the neighbours of x. There are two main cases:

(i) The y and z are not adjacent, i.e., $yz \notin E(G)$. In this case, delete x from the graph and add the edge yz to it. For this new graph we can use the induction hypothesis: it contains a topological K_4. If this topological K_4 does not contain the edge yz, then the original graph contained this topological K_4 too. If it contains the edge yz, delete this edge and draw back the vertex x with the edges xy and xz. Clearly, we get a topological K_4 in G.

(ii) The y and z are adjacent, i.e., $yz \in E(G)$. There are two cases:

   (a) The y and z do not have a common neighbour other than x. Denote by $v_1, v_2, \ldots, v_s$ the neighbours of z. Delete x and z from the graph and connect $v_1, v_2, \ldots, v_s$ to y. For this new graph we can use the induction hypothesis: it contains a topological K_4. If this topological K_4 does not contain any of the edges yv_i, then the original graph contained this topological K_4 too. If it contains any of the edges yv_i, delete those edges and draw back the vertices x and z with the edges xy, xz, yz and the zv_i's. Clearly, we get a topological K_4 in G by considering the path $yxzv_i$ instead of yv_i.

   (b) The y and z have a common neighbour $w \neq x$. If $d(y) \geq 4$ or $d(z) \geq 4$, we can delete x and use the induction hypothesis. If both of them have degree 3, delete x, y and z. In the remaining graph, w is the only vertex that could have degree less than 3. Hence, we can use the induction hypothesis for this graph.

Consider the case $d(x) = 3$. Denote by y, z and v the neighbours of x. There are two main cases:

(i) All the neighbours of x are adjacent to each other, i.e., $yz, yv, zv \in E(G)$. In this case, x, y, z, v form a K_4, so we are done.

(ii) There exist two neighbours of x that are not adjacent to each other. We may assume that $yz \notin E(G)$. In this case, delete x from the graph and add the edge yz to it. For this new graph we can use the induction hypothesis: it contains a topological K_4. If this topological K_4 does not contain the edge yz, then the original graph contained this topological K_4 too. If it contains the edge yz, delete this edge and draw back the vertex x with the edges xy and xz. Clearly, we get a topological K_4 in G.

■

The proof of the proposition is complete. ■

Exercise 3.2.6. *Show that*

$$\mathrm{ex}(n, \text{a cycle with a chord}) = 2n - 4.$$

(Hint: first, consider the case $\delta(G) \geq 3$, then use induction.)

Exercise 3.2.7. *A theta graph is the union of three internally disjoint (simple) paths that have the same two distinct end vertices. Show that*

$$\mathrm{ex}(n, \text{theta graph}) = \left\lfloor \frac{3}{2}(n-1) \right\rfloor.$$

Theorem 3.2.15 (Erdős–Gallai)**.** *Denote by P_k a path on k vertices. We have*

$$\mathrm{ex}(n, P_k) \leq \frac{k-2}{2} n,$$

and equality holds if $k - 1 \mid n$.

Proof. By induction on k, then by induction on n. The case $k = 2$ is obvious. We proceed to $k > 2$. Suppose that

$$e(G) > \frac{k-2}{2} n.$$

Since $\frac{k-2}{2} n > \frac{k-3}{2} n$, by the induction hypothesis, G contains a path P_{k-1}. Denote by $v_1, v_2, \ldots, v_{k-1}$ the vertices of the path in the order as they follow each other. If we can extend it to a P_k, then we are done. Otherwise, consider the neighbours of v_1 and v_{k-1}. Since we can not extend the path to a longer one, the neighbours of them are among the vertices $v_2, v_3, \ldots, v_{k-2}$. There are two cases:

(i) There is an $2 \leq i \leq k-1$ such that $v_1 v_i$ and $v_{k-1} v_{i-1}$ are edges in the graph. In this case, we have the cycle $v_1 v_i v_{i+1} v_{i+2} \dots v_{k-1} v_{i-1} v_{i-2} \dots v_1$ (see Figure 3.2). If any of the vertices v_i has a neighbour different from the ones in the cycle, we obtain a path P_k and we are done. If not, these $k-1$ vertices form a component of the graph G. Consider the graph $G' = G - \{v_1, v_2, \dots, v_{k-1}\}$. We have

$$|V(G')| = n - (k-1) \quad \text{and}$$
$$e(G') \geq e(G) - \binom{k-1}{2} > \frac{k-2}{2} n - \binom{k-1}{2} = \frac{k-2}{2}(n-(k-1)).$$

If the theorem holds for graphs with vertices at most $k-1$, then we can proceed by induction on n and conclude that G' contains a path P_k (and so does G). Clearly, a graph on at most $k-1$ vertices does not contain a P_k and satisfies

$$e \leq \binom{|V|}{2} = \frac{(|V|-1)|V|}{2} \leq \frac{(k-1)-1}{2}|V| = \frac{k-2}{2}|V|.$$

Therefore, in this case we are done.

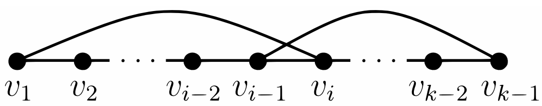

Figure 3.2

(ii) In this case we assume that if v_i is adjacent to v_1 then v_{i-1} can not be adjacent to v_{k-1} $(2 \leq i \leq k-1)$. Therefore, v_{k-1} can have at most $k-2-d(v_1)$ neighbours, i.e.,

$$d(v_1) + d(v_{k-1}) \leq k-2.$$

Consider the graph $G' = G - \{v_1, v_{k-1}\}$. We have

$$|V(G')| = n-2 \quad \text{and}$$
$$e(G') \geq e(G) - (d(v_1) + d(v_{k-1})) > \frac{k-2}{2} n - (k-2) = \frac{k-2}{2}(n-2).$$

As in the previous case, we can proceed by induction on n and conclude that G' contains a path P_k (and so does G).

To prove the last statement, consider $\frac{n}{k-1}$ distinct copies of K_{k-1}. ■

Definition 3.2.4. *A* ***Hamiltonian path*** *is a path in a graph that visits each vertex exactly once. A* ***Hamiltonian cycle*** *is a Hamiltonian path that is a cycle.*

Theorem 3.2.16 (Dirac)**.** *Let G be a graph. If $d(x) \geq \frac{n}{2}$ for all $x \in V(G)$, then G has a Hamiltonian cycle.*

Proof. Take a maximal counterexample G, i.e., $d(x) \geq \frac{n}{2}$ for all $x \in V(G)$, G does not contain a Hamiltonian cycle, and adding any $xy \notin E(G)$ edge to G, the resulting graph contains a Hamiltonian cycle. It is easy to see that G contains a Hamiltonian path $v_1v_2\dots v_n$ ($|V(G)| = n$). If there is an $2 \leq i \leq n$ such that v_1v_i and v_nv_{i-1} are edges in the graph, then we have the Hamiltonian cycle $v_1v_iv_{i+1}v_{i+2}\dots v_nv_{i-1}v_{i-2}\dots v_1$. Therefore, v_n can have at most $n-1-d(v_1)$ neighbours, i.e.,

$$d\left(v_1\right) + d\left(v_n\right) \leq n - 1 < n.$$

On the other hand, by assumption, $d\left(v_1\right) + d\left(v_n\right) \geq \frac{n}{2} + \frac{n}{2} = n$. This is a contradiction. ■

3.3 Algebraic methods

Let G be an r-regular graph with girth g. First, suppose that g is odd. Choose an arbitrary vertex x of G. Let us call the set $\{x\}$ as the zeroth level. This vertex has precisely r different neighbours, call the set of these neighbours as the first level. There are $r = r(r-1)^0$ edges between the zeroth and the first level, and this is the number of vertices in the first level too. If $g > 3$, then there are no edges between the vertices of the first level. Consider the neighbours of the vertices of the first level. These vertices form the second level. If $g > 3$, then there are $r(r-1) = r(r-1)^1$ edges between the first and the second level, and this is the number of vertices in the second level too. If $g > 5$, then there are no edges between the vertices of the second level. An example is shown in Figure 3.3. We can continue this procedure further. The $\frac{g-1}{2}$th level is the first level that might contain edges between the vertices of the level. Hence,

$$|V\left(G\right)| = n \geq 1 + \sum_{i=0}^{\frac{g-1}{2}-1} r\left(r-1\right)^i = 1 + r\frac{\left(r-1\right)^{\frac{g-1}{2}} - 1}{r-2}.$$

When g is even, we start with an edge xy and we carry out the above procedure both with x and y. It follows that

$$|V(G)| = n \geq 2\sum_{i=0}^{\frac{g}{2}-1}(r-1)^i = 2\frac{(r-1)^{\frac{g}{2}}-1}{r-2}.$$

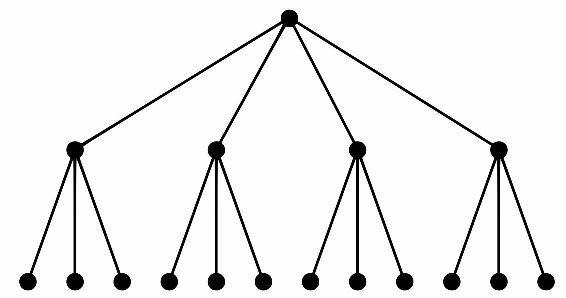

Figure 3.3

Let us consider now the case $g = 5$. By the above argument, if G is an r-regular graph with girth 5, we must have

$$|V(G)| = n \geq 1 + r\frac{(r-1)^2-1}{r-2} = 1 + r\frac{r(r-2)}{r-2} = r^2+1.$$

The question arises naturally whether there are r-regular graphs on r^2+1 vertices with girth 5. For $r = 2$ and 3 the 5-cycle $\mathcal{C}_5$ and the Petersen graph provide positive answer. In general, we have the following theorem:

Theorem 3.3.1 (Hofmann–Singleton)**.** *If there exists an r-regular graph of r^2+1 vertices with girth 5 then r must be $2, 3, 7$ or 57.*

Remark. *We saw that for $r = 2, 3$ such graphs exist. Hofmann and Singleton constructed a graph for $r = 7$. The existence of such a graph for $r = 57$ is not known.*

Proof. Let $v_1, v_2, \ldots, v_n$ be the vertices of the r-regular graph G ($n = r^2+1$). Let A be the *adjacency matrix* of G, i.e., a 0-1 matrix whose ijth element is 1 if $v_iv_j \in E(G)$ and 0 otherwise. Consider the matrix $A^2 = \{a_{ij}\}_{i,j=0}^{n}$. It is easy to see that a_{ij} is the number of walks from v_i to v_j of length 2. Thus, the a_{ii} is precisely the number of neighbours of v_i, i.e., $a_{ii} = r$. If $i \neq j$, a_{ij} is precisely the number of common neighbours of v_i and v_j. If v_i and v_j have more than 1 common neighbours, then they form a $\mathcal{C}_4$ which is impossible since the girth number of G is 5. Therefore, $a_{ij} \leq 1$ for $i \neq j$. If $v_i v_j \in E(G)$, then $a_{ij} = 0$ otherwise we would have a triangle in the graph. If $v_i v_j \notin E(G)$, then $a_{ij} = 1$. Indeed, starting with $x = v_i$, the vertices can be divide into three levels as in the argument of the first paragraph of this section (this is because $g = 5$ and $n = r^2 + 1$). Then v_j must be in the second level. They have a common neighbour in the first level (see Figure 3.3), therefore $a_{ij} = 1$. Therefore, we have the matrix equation

$$A + A^2 - (r-1)\,I = J,$$

where I is the $n \times n$ identity matrix and J is the $n \times n$ matrix whose each entry is 1.

Since A is a symmetric matrix ($A = A^\top$), all its eigenvalues are real and some of the eigenvectors constitute an orthogonal basis. Note that r is an eigenvalue with eigenvector $\mathrm{e} \stackrel{\text{def}}{=} (1, 1, \ldots, 1)^\top$. Let $\lambda \neq r$ be an other eigenvalue of A with eigenvector $v \neq 0$. By the above matrix equation, we have

$$\lambda v + \lambda^2 v - (r-1)\,v = Jv.$$

The right-hand side is $Jv = \langle \mathrm{e}, v \rangle = 0$ by the orthogonality. Therefore, we find

$$\lambda + \lambda^2 - (r-1) = 0.$$

This quadratic equation has the solutions

$$\lambda_{1,2} = \frac{-1 \pm \sqrt{4r-3}}{2}.$$

Thus, A has three different eigenvalues: r, λ_1 and λ_2. The multiplicity of r is 1, and denote by m_1 and m_2 the multiplicities of λ_1 and λ_2. Since $1 + m_1 + m_2 = n$, we obtain

$$m_1 + m_2 = n - 1 = r^2.$$

On the other hand,

$$0 = \operatorname{tr}(A) = r + m_1\lambda_1 + m_2\lambda_2.$$

Substituting the values of λ_1 and λ_2 yields

$$2r - m_1 + m_1\sqrt{4r-3} - m_2 - m_2\sqrt{4r-3} = 0.$$

If $m_1 = m_2$, then $r = 2$. If $m_1 \neq m_2$, we must have $4r - 3 = s^2$ for some positive integer s. Hence,

$$2r - m_1 + m_1 s - m_2 - m_2 s = 0,$$

or

$$32r - 16(m_1 + m_2) + 16(m_1 - m_2)s = 0.$$

Since $m_1 + m_2 = r^2$ and $r = \frac{s^2+3}{4}$, we obtain

$$8\left(s^2 + 3\right) - \left(s^2 + 3\right)^2 + 16(m_1 - m_2)s = 0,$$

or

$$-s^4 + 2s^2 + 16(m_1 - m_2)s + 15 = 0.$$

The right-hand side is divisible by s, so has to be the left-hand side, i.e., $s \mid 15$. Therefore, $s = 1, 3, 5$ or 15 which give the values $r = 1, 3, 7$ and 57. ■

Consider $\mathbb{R}^d$ with the usual metric $d(\,\cdot\,,\,\cdot\,)$. Let $c > 0$ be a constant. Let $S \subset \mathbb{R}^d$ be a set of n distinct points:

$$S = \{p_1, p_2, \ldots, p_n\},$$

such that $d(p_i, p_j) = c$ for all $i \neq j$. Then it is known that $n \leq d + 1$ and equality holds if S is the standard d-dimensional simplex.

A similar problem is the following. Let $c_1, c_2 > 0$ be constants. Let $S \subset \mathbb{R}^d$ be a set of n distinct points:

$$S = \{p_1, p_2, \ldots, p_n\},$$

such that $d(p_i, p_j) = c_1$ or c_2 for all $i \neq j$. Such a set is called a *two-distance set*. The question is, how large n can be? Consider the set of $(d+1)$-dimensional vectors whose two coordinates are 1 and all the others are 0. There are $\binom{d+1}{2} = \frac{d^2}{2} + \frac{d}{2}$ such vectors. They are contained in the hyperplane

$$\sum_{i=1}^{d+1} x_i = 2.$$

Hence, by translation, they are points in $\mathbb{R}^d$. This shows that

$$\frac{d^2}{2} + \frac{d}{2} \leq n = |S|$$

is possible. The precise value of the maximum is not known. Nevertheless, we shall show that it is $\frac{d^2}{2} + \mathcal{O}(d)$ as $d \to +\infty$. Suppose that $S = \{p_1, p_2, \ldots, p_n\}$ satisfies the conditions. Consider the polynomials

$$P_i\left(x_1, x_2, \ldots, x_d\right) = \left(\|\mathbf{x} - p_i\|^2 - c_1^2\right)\left(\|\mathbf{x} - p_i\|^2 - c_2^2\right) \quad i = 1, 2, \ldots, n$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$. Let $\mathcal{P}$ be the vector space, generated by the polynomials $P_1, P_2, \ldots, P_n$. Suppose that

$$\sum_{i=1}^{n} \lambda_i P_i = 0$$

holds for some real constants $\lambda_1, \lambda_2, \ldots, \lambda_n$. Substituting $\mathbf{x} = p_j$ shows that $\lambda_j = 0$, therefore, our polynomials are linearly independent. Consequently $\dim \mathcal{P} = n$. The polynomials can contain the following terms:

$$\text{constant term;}$$

$$x_1, x_2, \ldots, x_d;$$

$$x_i x_j \;\text{ for }\; i, j = 1, 2, \ldots, d;$$

$$\left(\sum_{i=1}^{d} x_i^2\right)^2;$$

$$\left(\sum_{i=1}^{d} x_i^2\right) x_j \;\text{ for }\; i = 1, 2, \ldots, d.$$

These polynomials are linearly independent in the space of d-variable polynomials and $\mathcal{P}$ is contained in the subspace generated by these polynomials. The dimension of the space that they generate is

$$1 + d + \left(\binom{d}{2} + d\right) + 1 + d = \frac{d^2}{2} + \frac{5}{2}d + 2,$$

whence, $|S| = n = \dim \mathcal{P} \leq \frac{d^2}{2} + \frac{5}{2}d + 2$.

About the lecturer

Ervin Győri was born in 1954. He earned his Ph.D. degree in mathematics in 1980. He has been the doctor of the Hungarian Academy of Sciences since 1994. He has been working as a research fellow at the Alfréd Rényi Institute of Mathematics since 1977. He is an internationally recognized, leading researcher of discrete mathematics. He has written more than 60 papers and edited five conference proceedings. His main fields of interest are extremal graph theory, connectivity of graphs, cycles in graphs and minimax theorems in combinatorics.