Discrete Mathematics Lecture Notes

Oliver Roche-Newton

February 26, 2026

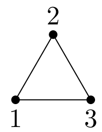
Abstract

Lecture notes for the course on Discrete Mathematics at JKU, Summer Semester 2026.

Contents

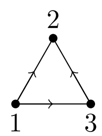
1 Recap of some basic graph theory

An *undirected simple graph* is a pair $G = (V, E)$. V is the set of *vertices* of G and E is the set of *edges*. The elements of E are subsets of V of cardinality 2.



In order to simplify notation, we will write uv as a shorthand for an edge $\{u, v\}$ of a graph. In the above example, $V = \{1, 2, 3\}$ and $E = \{12, 13, 23\}$. If $uv \in E$ then we say that the vertices u and v are *adjacent*.

A *directed simple graph* is a pair $G = (V, E)$. V is the set of *vertices* of G and E is the set of *edges*. The elements of E are *ordered pairs* of elements of $V \times V$.

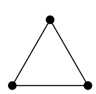


In the above picture, $V = \{1, 2, 3\}$ and $E = \{(1, 2), (1, 3), (3, 2)\}$.

Most of the graphs we consider in this course will be simple graphs (that is, loops and edges with multiplicity are not allowed). Unless otherwise stated, when we talk of a *graph*, we mean a simple undirected graph.

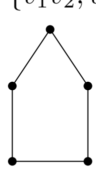
1.1 Some important special graphs

- A graph which contains every possible edge is called a *complete graph*. The complete graph on n vertices is denoted by K_n. So, $K_n = (V, E)$ where $V = \{v_1, v_2, \ldots, v_n\}$ and $E = \{v_i v_j : 1 \leq i < j \leq n\}$. For example, K_3 is a triangle. Note that K_n has $\binom{k}{2}$ edges.

- The notation P_n is used for a graph $G = (V, E)$ with $V = \{v_1, v_2, \ldots, v_n\}$ and $E = \{v_i v_{i+1} : 1 \leq i \leq n - 1\}$. The P here stands for "path".

- The notation C_n is used for a graph $G = (V, E)$ with $V = \{v_1, v_2, \ldots, v_n\}$ and $E = \{v_1 v_2, v_2 v_3, \ldots, v_{n-1} v_n, v_n v_1 :\}$. The C here stands for "cycle".

1.2 More terminology and definitions

- Let $G = (V, E)$ be a graph. A subset $U \subseteq V$ of vertices such that $uv \in E$ for all $u, v \in U$ with $u \neq v$ is called a *clique*. The size of the maximum clique in G is called the *clique number* of G. The clique number is denoted $\omega(G)$.
- A subset $U \subseteq V$ of vertices such that $uv \notin E$ for all $u, v \in U$ is called an *independent set*. The size of the maximum independent set in G is called the *independence number* of G. The independence number is denoted $\alpha(G)$.
- Let $u \in V(G)$. The *neighbourhood* of u, denoted $N(u)$, is the set of all vertices $v \in V(G)$ such that $uv \in E(G)$. So,
$$N(u) := \{v \in V(G) : uv \in E(G)\}.$$
- The *degree* of u is the number of vertices which are adjacent to u. That is, $\deg(u) = |N(u)|$. The *minimum vertex degree* of G is the quantity
$$\delta(G) := \min_{u \in V} \deg(u).$$
- Let $G = (V, E)$ be a graph. The *complement* of G, denoted $\bar{G}$, if the graph with vertex set V and with the property that $e \in E(\bar{G})$ if and only if $e \notin E(G)$.
- A *path* in a graph $G = (V, E)$ is a sequence of distinct vertices $v_1, v_2, \dots, v_k$ such that $v_i v_{i+1} \in E$ for all $1 \leq i \leq k-1$. We can also call this *a path between v_1 and v_k*. The *length* of the path is the number of edges in it (so $k-1$ in this case). We also allow the trivial path v_1, v_1, which has length 0.
- A graph is *connected* if there is a path between every pair of distinct vertices in V.
- The *distance* between two distinct vertices $u, v \in V$, denoted $dist(u, v)$ is the length of a shortest path between u and v.
- The *diameter* of a connected graph , denoted $diam(G)$ is the maximum distance. That is,
$$diam(G) = \max_{u,v \in V} dist(u, v).$$

  A *hypergraph* is a pair $H = (V, E)$ where V is a finite set of *vertices* and the elements of E (called *hyperedges* or simply *edges*) are subsets of V. An r-uniform hypergraph is a hypergraph for which all elements of E have cardinality r. So a 2-uniform hypergraph is simply a graph.

Exercises. Which of the following graphs are connected? For those which are connected, what is their diameter? What are the complements of these graphs? Calculate the clique number and independence number of each graph.

1. $G = (V, E)$ with $V = \{1, 2, 3, 4, 5, 6\}$ and $E = \{12, 13, 23, 34, 56\}$
2. $G = (V, E)$ with $V = \{1, 2, 3, 4, 5, 6\}$ and $E = \{12, 13, 23, 34, 35, 56\}$
3. The disjoint union of a K_3 and a K_2.

1.3 The Handshake Lemma

The following result is very simple, but will also be extremely useful throughout the course.

Theorem 1.1 (The Handshake Lemma)**.** *For any graph $G = (V, E)$,*

$$\sum_{v \in V} deg(v) = 2|E|.$$

Proof.

$$\sum_{v \in V} deg(v) = \sum_{v \in V} \sum_{u \in V : uv \in E} 1 = \sum_{v \in V} \sum_{e \in E} |\{u \in V ; uv = e\}| = \sum_{e \in E} \sum_{v \in V, u \in V : uv = e} 1 = \sum_{e \in E} 2 = 2|E|.$$

□

Exercises. Use the handshake lemma to show that, in any graph, the number of vertices with odd degree is even. Prove that if a graph has exactly two vertices of odd degree, then they are connected by a path.

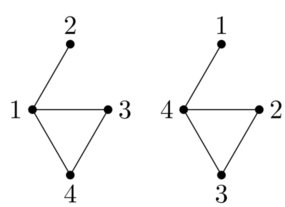
1.4 Isomorphic graphs

Two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are said to be *isomorphic* if there exists a bijection $f : V_1 \to V_2$ with the property that $uv \in E_1$ if and only if $f(u)f(v) \in E_2$.

Example. Let $V_1 = V_2 = \{1, 2, 3, 4\}$, let $E_1 = \{12, 13, 14, 34\}$ and let $E_2 = \{14, 23, 24, 34\}$. Then $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are isomorphic. We define $f : \{1, 2, 3, 4\} \to \{1, 2, 3, 4\}$ by

$$f(1) = 4, f(2) = 1, f(3) = 2, f(4) = 3.$$

We can check that $uv \in E_1 \Leftrightarrow f(u)f(v) \in E_2$. Perhaps it is easier to get an idea of why these two graphs are isomorphic by studying the following simple picture.



These two graphs look the same. The only difference is that the labels for the vertices have changed. One may think of two graphs as being isomorphic if one can be obtained from the other by a relabelling of vertices.

Exercise - Which of the following graphs are isomorphic to one another? If they are isomorphic write down a bijection with the required properties.

1. $G = (V_1, E_1)$ with $V_1 = \{1, 2, 3, 4, 5\}$, $E_1 = \{12, 13, 14, 35, 45\}$
2. $G = (V_2, E_2)$ with $V_2 = \{1, 2, 3, 4, 5\}$, $E_2 = \{13, 14, 25, 35, 45\}$
3. $G = (V_3, E_3)$ with $V_3 = \{1, 2, 3, 4, 5\}$, $E_3 = \{12, 13, 14, 15, 25\}$
4. $G = (V_4, E_4)$ with $V_4 = \{1, 2, 3, 4, 5, 6\}$, $E_4 = \{12, 13, 14, 15, 25\}$
5. $G = (V_5, E_5)$ with $V_5 = \{1, 2, 3, 4, 5, 6\}$, $E_5 = \{12, 14, 24, 34, 46\}$

1.5 A result to warm us up

Theorem 1.2. *Let $G = (V, E)$ be a connected graph with $diam(G) \geq 3$. Then $diam(\bar{G}) \leq 3$.*

Proof. We need to show that any two vertices $a, b \in V$ are connected by a path of length three or less in $\bar{G}$. Since $diam(G) \geq 3$, we know that there exist two vertices is $u, v \in V$ which cannot be connected by a path of length 2 in G. Note that $uv \notin E$, and hence u and v are connected by a path of length 1 in $\bar{G}$ (i.e. $uv \in E(\bar{G})$).

For any $a \in V \setminus \{u, v\}$, the vertex a can be adjacent (in G) to at most one of u and v. Therefore, a is adjacent to at least one of u and v in $\bar{G}$ (that is at least one of $au \in E(\bar{G})$ or $av \in E(\bar{G})$ is true). Since $uv \in E(\bar{G})$, it follows that a is connected to each of u and v by a path of length 2 or less.

It remains to show that $a, b \in V \setminus \{u, v\}$ are connected by a path of length at most 3 in $\bar{G}$. There are four cases to consider, and in each case we can find a path of length 3 or less, as the following picture illustrates.

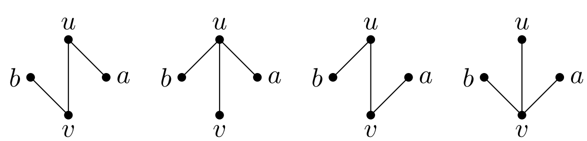


Figure 1: Here we have a visualisation of four possible scenarios for the edges in $\bar{G}$. There may also be more edges than we have drawn here, but that only helps us in the proof.

For example, if $au \in E(\bar{G})$ and $bv \in E(\bar{G})$, then a, u, v, b is a path of length 3 in $\bar{G}$. There are three other cases to check ($au, bu \in E(\bar{G})$, $av.bv \in E(\bar{G})$, and $av, bu \in E(\bar{G})$, and in each case we can find such a path. So, a and b are connected by a path of length 3 in $\bar{G}$.

∎

1.6 More exercises

1. Let $G = (V, E)$ be a disconnected graph. Prove that $diam(\bar{G}) \leq 2$.

2. Let $G = (V, E)$ be a graph with $|V| = n$ and suppose that $|E| > \binom{n-1}{2}$. Prove that G is connected.

3. A graph G is *self complementary* if G is isomorphic to $\bar{G}$. Which of the following graphs are self-complementary: P_5, C_5? Prove that a self-complementary graph on n vertices has exactly $\frac{n(n-1)}{4}$ edges. Prove that there are no self complementary graphs on 3 vertices. Prove that there are no self-complementary graphs on 6 vertices. Are all graphs with $\frac{n(n-1)}{4}$ edges self-complementary?

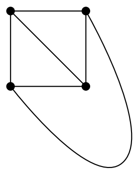
4. Is it possible for a self-complementary graph with 100 vertices to have exactly one vertex with degree 50.

2 Planarity

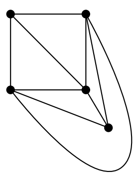
2.1 Euler's formula and applications

A graph G is *planar* if it can be drawn in the plane so that no two edges cross each other.

Examples - K_4 is planar, since we can draw it like this:



The graph $K_5 \setminus \{e\}$, that is the graph K_5 with one edge removed, is planar. It can be drawn like this:

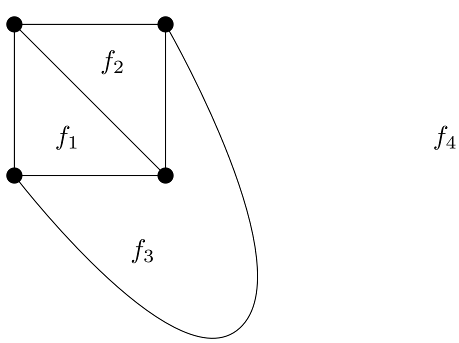


Is K_5 planar? It seems difficult to make a planar drawing of K_5. Perhaps it is impossible? More generally, we can ask the following extremal question about planar graphs: what is the maximum possible number of edges for a planar graph on n vertices?

Exercise - Show that $K_{2,t}$ is planar for any t.

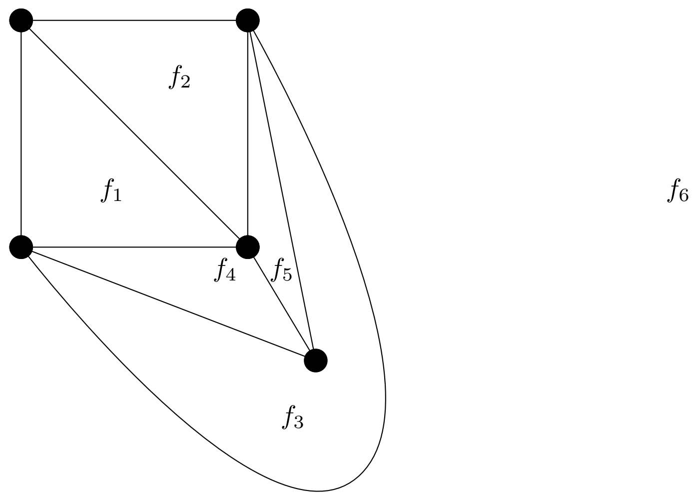
A planar representation of a graph partitions the plane into 2-dimensional, 1-dimensional and 0-dimensional sets. These are called *faces, edges* and *vertices*. Edges and vertices are the same thing as usual. A *face* is a maximal section of the plane in which any two points can be connected by a curve that does not meet any of the vertices or edges in the drawing. A face can be bounded or unbounded.

Examples - Look again at our drawing of K_4. There are 4 faces, labelled f_1, f_2, f_3 and f_4.



So, we have 4 faces, 6 edges and 4 vertices

Look again at our drawing of $K_5 \setminus \{e\}$. There are 6 faces, 9 edges and 5 vertices.



For planar graphs, Euler's formula tells us about the relation between the number of vertices, edges and faces.

Theorem 2.1 (Euler's Formula)**.** *Let G be a connected planar graph with n vertices and m edges. Then, for any planar drawing of the graph, the number of faces f satisfies the formula*

$$n - m + f = 2.$$

This is an important result which can be used to answer the question of how many edges a planar graph can have. We can also use it to prove that graphs with too many edges must contain many crossing pairs of edges.

An interesting class of planar graphs are trees. A graph G is called a *tree* if it is connected and does not contain any cycles. We will not study trees in great detail in this course, but will need the following fact in order to prove Theorem 2.1.

Lemma 2.2. *If $G = (V, E)$ is a tree, then*

$$|E| = |V| - 1. \tag{1}$$

Proof. The proof is by induction on $|V|$. For the base case when $|V| = 1$, we must have $|E| = 0$, and the required identity (1) holds.

Now let $G = (V, E)$ be a tree with $|V| \geq 2$. Since G is a tree, there must be some vertex $v \in V$ with $\deg_G(v) = 1$. Consider the graph $G' = G[V \setminus \{v\}]$. This graph is still a tree and has one edge and one vertex less than G. By the induction hypothesis

$$|E| - 1 = |E(G')| = |V(G')| - 1 = |V| - 2$$

and (1) follows.

$\square$

We can now prove Euler's Formula.

Proof of Euler's Formula. The proof is by induction on m. If $m = 0$ then $G = K_1$. For any planar drawing of the graph, we have $f = 1$ and $n = 1$. So, $n - m + f = 1 - 0 + 1 = 2$, as required.

Now, let $m \geq 1$ and assume the statement is true for all connected planar graphs with n vertices and less than m edges, and let G be a connected planar graph with n vertices and m edges.

Case 1 - There is a sequence of edges forming a cycle; that is $v_1v_2, v_2v_3, \ldots, v_{i-1}v_i, v_iv_1 \in E$. Consider the graph G' obtained from G by removing a single edge in this cycle, say v_1v_2, and note that G' remains planar and connected. The graph G' has n vertices and $m-1$ edges. Also, the drawing of G' has $f - 1$ faces, since two faces from G merge in G'. By the induction hypothesis

$$n - (m-1) + (f-1) = 2,$$

which implies that $n - m + f = 2$.

Case 2 - There are no cycles in G (i.e. G is a tree). Then, since there are no enclosed regions in a planar drawing of G, any drawing of G has one face. So $f = 1$. Furthermore, by Lemma 2.2, any tree on n vertices has exactly $m = n - 1$ edges. Therefore,

$$n - m + f = n - (n-1) + 1 = 2,$$

as required.

□

Corollary 2.3. *If $G = (V, E)$ is a connected planar graph and $|V| \geq 3$ then $|E| \leq 3|V| - 6$.*

Proof. Consider a drawing of the graph G. For each face f, we can trace continuously around the boundary and encounter a sequence of vertices and edges $v_1, e_1, v_2, e_2, \ldots, v_d, e_d, v_1$. The value d is called the *degree* of the face, and we use the notation $\deg(f) = d$ (note that some vertices and edges may appear twice in the aforementioned sequence, for example if there is only one face). If we repeat this process for every face in the drawing, we encounter each edge exactly twice. That is,

$$2|E| = \sum_f \deg(f).$$

Let f_k denote the number of faces with degree k. Then, since G is a graph with at least 3 vertices, $f_1 = f_2 = 0$. Therefore,

$$2|E| = \sum_f \deg(f) = \sum_{k\geq 1} \sum_{f:\deg(f)=k} \deg(f) = \sum_{k\geq 1} k f_k = \sum_{k\geq 3} k f_k \geq 3 \sum_{k\geq 3} f_k = 3 \sum_k f_k = 3f.$$

Then, by Euler's formula,

$$2|E| \geq 3(2 + |E| - |V|),$$

and thus

$$|E| \leq 3|V| - 6.$$

□

It is not difficult to remove the condition that in Corollary 2.3 that the graph is connected.

Corollary 2.4. *If $G = (V, E)$ is a planar graph with $|V| \geq 3$ then $|E| \leq 3|V| - 6$.*

Proof. The planar graph G is a disjoint union of one or more planar connected graphs. We can always form a planar connected graph from G by adding edges to the graph. This can be proved formally by induction on the number of connected components. So, we have $G \subset H$ where H is a planar connected graph with the same vertex set as G. It then follows from Corollary 2.3 that

$$|E(G)| \leq |E(H)| \leq 3|V| - 6.$$

□

Corollary 2.5. *K_5 is not planar.*

Proof. K_5 has 5 vertices and $\binom{5}{2} = 10$ edges. Suppose K_5 is planar. Then, by Corollary 2.4, $10 = |E| \leq 3|V| - 6 = 9$. This is a contradiction. □

Exercise - Give an example of a planar graph $G = (V, E)$ with $|V| \geq 3$ such that $|E| = 3|V| - 6$. This shows that Corollary 2.4 cannot be improved.

The goal of the next two exercises is to show that the converse of Corollary 2.3 is not true. That is, we will give an example of a graph $G = (V, E)$ with $|E| \leq 3|V| - 6$ which is not planar. Namely, we will prove that $K_{3,3}$ is not planar. Can you find any other examples which show that the converse of Corollary 2.3 fails?

Exercise - Suppose that $G = (V, E)$ is a connected planar graph with no triangles and $|V| \geq 3$. Prove that $|E| \leq 2|V| - 4$. (Hint: follow the argument of Corollary 2.3, but use the triangle-free hypothesis to obtain some information about f_3.)

Exercise - Use the previous exercise to show that $K_{3,3}$ is not planar. On the other hand, show that $K_{3,3} \setminus \{e\}$ is planar.

Corollary 2.6. *Every planar graph has a vertex of degree less than or equal to* 5.

Proof. The result holds trivially if the graph has less than 3 vertices. Now let $G = (V, E)$ be a planar graph with $|V| \geq 3$ and suppose for a contradiction that every vertex of G has degree at least 6. Then, by the Handshake Lemma,

$$2|E| = \sum_{v \in V} \deg(v) \geq 6|V|.$$

This tells us that $|E| \geq 3|V|$, which contradicts Corollary 2.4. □

2.2 Graph Colouring

A *vertex colouring* of a graph G is an assignment of colours to the vertices of G such that no two adjacent vertices are assigned the same colour. To put it another way, given a graph $G = (V, E)$, a vertex colouring is a partition of V into subsets $V_1, V_2, \ldots, V_k \subset V$ such that $V_1 \sqcup V_2 \cdots \sqcup V_k = V$ and for all i, the set $\{uv \in E : u, v \in V_i\}$ is empty.

We can also view a vertex colouring by k colours as a function $c : V \to \{1, 2, \dots, k\}$ such that

$$uv \in E \Rightarrow c(u) \neq c(v).$$

A graph is *k-colourable* if there exists vertex colouring by k colours. The sets V_i are called *colour classes.*

This is really another language for a concept we have seen earlier in the course; multipartite graphs. A bipartite graph is by definition 2-colourable.

The *chromatic number* of a graph G, denoted $\chi(G)$, smallest possible value of k such that G is k-colourable.

Example - Any graph G on n vertices is n-colourable, since we can assign all of the n vertices a different colour and then there will be no edges between vertices of the same colour. On the other hand the graph K_n has $\chi(K_n) = n$. Indeed, if there were to exist an $(n-1)$-colouring of K_n, then one of the colour classes V_i must contain at least two vertices. But then, any two distinct edges in this colour class are adjacent, which is a contradiction. By the same reasoning $\chi(G) \geq \omega(G)$.

We are now ready to consider a famous result in graph theory.

The motivation for this problem is the following. Suppose we are given a map showing the countries in some territory. To make the map easier to understand/read, the countries will be given different colours. We want to ensure that two countries sharing a border are given different colours, so that we can distinguish them easily when looking at the map. What is the smallest number of colours that we can use in order colour the map?

We can treat this as a graph theoretical problem, although we need to slightly simplify the map problem; all of the countries in our map must consists of a single connected piece (so we rule out countries consisting of several islands, or the existence of exclaves, for example). Inside each country, draw a vertex. If two countries share a border, we draw an edge between them. We can draw this as a planar graph (try this!). Call this graph G. The question then becomes, what is the smallest number of colours needed to colour the graph. In other words, what is the value $\chi(G)$?

If we try some examples, it seems that we can always do this with four colours. The following theorem is a little weaker.

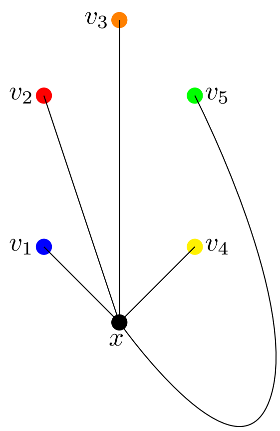
Theorem 2.7 (Five Colour Theorem)**.** *Let G be a planar graph. Then $\chi(G) \leq 5$.*

Proof. Let $G = (V, E)$ be a planar graph and $|V| = n$. The proof is by induction on n. For $n \leq 5$, the statement is trivial. Let G be a planar graph with $n > 5$. Let x be a vertex of degree less than or equal to 5 in G. Such a vertex x exists by Corollary 2.6. Consider the graph $G' = G[V \setminus \{x\}]$. By the induction hypothesis, G' has a 5-colouring $c' : G' \to \{1, 2, 3, 4, 5\}$. Suppose for a contradiction that there is no 5-colouring of G.

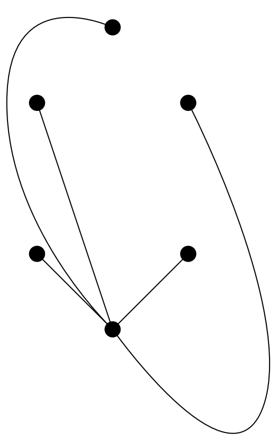
Suppose that $deg_G(x) < 5$. Then, there are at most four elements in the set $\{c'(v) : v \in N_G(x)\}$, and there is some value $a \in \{1, 2, 3, 4, 5\} \setminus \{c'(v) : v \in N_G(x)\}$. Therefore, we can construct a 5-colouring $c : G \to \{1, 2, 3, 4, 5\}$ by taking $c(x) = a$ and $c(v) = c'(v)$ for all $v \in V \setminus \{x\}$.

So, $deg_G(x) = 5$. Similarly, if two elements $u, v \in N_G(x)$ have $c'(u) = c'(v)$, then there is some $a \in \{1, 2, 3, 4, 5\} \setminus \{c'(v) : v \in N_G(x)\}$ and we can construct a 5-colouring of G. So, we can assume that the 5-elements of $N_G(x)$ have different colours, otherwise we are done.

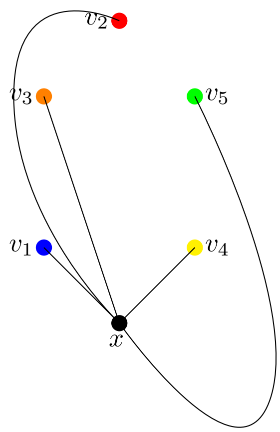
Now, take any planar drawing of G. We label the 5 vertices of $N_G(x)$ as v_1, v_2, v_3, v_4, v_5 in cyclic order as follows:



Note that the vertices have been labelled in clockwise order according to the direction of the edge as it leaves x. For example, if the drawing of the graph was modified to look like this



then it would be labelled as



Without loss of generality, we may assume that

$$c'(v_1) = 1,\ c'(v_2) = 2,\ c'(v_3) = 3, c'(v_4) = 4, c'(v_5) = 5.$$

We make the following claim.

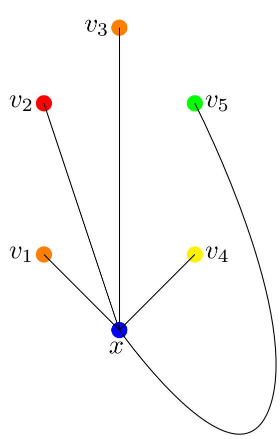
Claim - There exists a path from v_1 to v_3 consisting only of vertices which have colour 1 or 3.

Proof of Claim. Suppose for a contradiction that there is no such path. Now, consider the induced subgraph of G' consisting of all vertices v such that $c'(v) = 1$ or $c'(v) = 3$.

Let V' denote the connected component containing v_1. Then we can define a 5-colouring on G, $c : V \to \{1, 2, 3, 4, 5\}$, by reversing the colours inside the component V', and assigning $c(x) = 1$. That is,

$$\begin{aligned} c(v) &= c'(v), && v \notin V' \\ c(v) &= 3, && v \in V', c'(v) = 1 \\ c(v) &= 1, && v \in V', c'(v) = 3 \\ c(x) &= 1. \end{aligned}$$

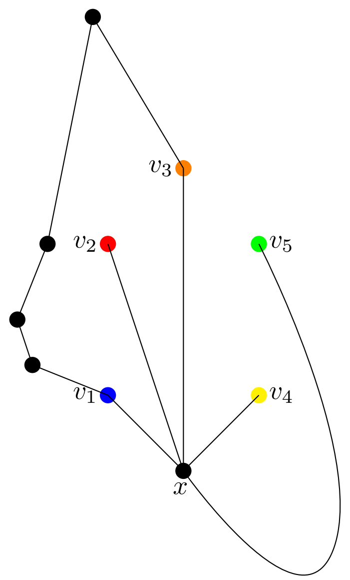
We get the following picture (refering to our original picture):



Note c is still a 5-colouring. There are no edges $uv \in E$ such that $c(u) = c(v) = 1$ or $c(u) = c(v) = 3$. This contradicts our assumption that there is no 5-colouring of G, and so the claim is proven.

□

We have the following picture.



Similarly, we may assume that there is some path from v_2 to v_4 consisting only of vertices with colour 2 or 4. However, as the picture shows, this contradicts the hypothesis that G is planar. Indeed, the path from v_2 to v_4 must, for topological reasons, at some point cross the path from v_1 to v_4. Since these paths do not share any common vertices, this must be an actual crossing. Therefore, our assumption was false, and it must be the case that G is 5-colourable. □

In fact, it really can always be done with as few as four colours, as the following result show.

Theorem 2.8 (Four Colour Theorem)**.** *Let G be a planar graph. Then $\chi(G) \leq 4$.*

The proof is far beyond the range of this course. It involves the use of heavy computation, and is understood fully by very few mathematicians!

3 Ramsey Theory

3.1 Ramsey's Theorem

Roughly speaking, the principle of Ramsey Theory is as follows. If we take a large complete graph and colour the edges using r colours, then no matter how we choose the colouring we will find certain monochromatic configurations, provided that n is sufficiently large (with respect to r).

Formally, an edge colouring of a graph is defined similarly to the vertex colourings we considered in the previous chapter. For a graph $G = (V, E)$, an *edge colouring by r colours* is simply a function

$$c : E \to \{1, \dots, r\}.$$

The main aim of this chapter is to prove the following result.

Theorem 3.1 (Symmetric version of Ramsey's Theorem for Graphs)**.** *Let r and s be positive integers. Then there is a positive integer $n = n(r, s)$ with the following property. For any colouring of the edges of K_n by r colours (i.e. for any function $c : E(K_n) \to \{1, \dots, r\}$) there exists a subset $V' \subset V$ with $|V| = s$ and such that all of the edges between vertices in V' are the same colour (i.e. there exists some $k \in \{1, \dots, r\}$ such that $c(v_iv_j) = k$ for all $v_i, v_j \in V'$).*

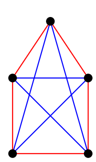
The proof will follow later.

Apart from proving this theorem, we would also like to find lower and upper bounds on the critical value of n.

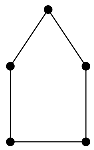
3.2 Two colours

We will largely focus on the case of two colours, i.e. the case $r = 2$.

Example - Let $s = 3$ and $r = 2$ in the statement of Theorem 3.1. This result tells us that there is some n such that every colouring of K_n by 2 colours contains a monochromatic triangle. Here is an example of a colouring of K_5 which does not give us a monochromatic triangle.



For two colours, it is perhaps easier to rephrase the statement slightly. One can view a 2-colouring of K_n, that is a map $c : E(K_n) \to \{1, 2\}$, as simply a graph G on n vertices. Those edges $e \in K_n$ such that $c(e) = 1$ are kept in G, and those e such that $c(e) = 2$ are deleted. The question of whether or not our colouring gives us a monochromatic triangle then changes to the following question: does our graph contain a clique or an independent set of size 3? Using this interpretation, our previous picture becomes C_5.



C_5 is an example of a graph which does not contain any cliques or independent sets of size 3. Are there any graphs on 6 vertices which do not contain cliques or independent sets of size 3? The answer is no.

Theorem 3.2. *Let $G = (V, E)$ be a graph with $|V| = 6$. Then* $\max\{\alpha(G), \omega(G)\} \geq 3$.

Proof. Let $v \in V$ be arbitrary. Suppose that $\deg(v) \geq 3$. Then, let $v_1, v_2, v_3 \in N(v)$. If $v_iv_j \in E$ for some $i, j \in \{1, 2, 3\}$ then $\{v, v_i, v_j\}$ is a clique of size 3. So, $\omega(G) \geq 3$. If not then $v_iv_j \notin E$ for all $i, j \in \{1, 2, 3\}$, which means that $\{v_1, v_2, v_3\}$ is an independent set of size 3. Hence $\alpha(G) \geq 3$.

On the other hand, suppose that $\deg(v) < 3$. Then there are three vertices $v_1, v_2, v_3 \notin N(v) \cup \{v\}$. If all three of these vertices are connected then we have a clique of size 3 and so $\omega(G) \geq 3$. If not, then we have $v_iv_j \notin E(G)$ for some $i, j \in \{1, 2, 3\}$. Therefore $\{v, v_i, v_j\}$ is an independent set of size 3 and $\alpha(G) \geq 3$.

□

We introduce some notation to describe the critical value n in Theorem 3.1 for the case $r = 2$. At the same time, we generalise slightly to consider the possibility that the cliques and independent sets we are searching for may have different sizes.

The minimum n such that every graph with n vertices contains either a clique of size s or an independent set of size t is the Ramsey number $R(s, t)$. We say that $R(s, s)$ is the symmetric Ramsey number.

Using this notation, we can restate Theorem 3.1 for the case r = 2 as follows:

Theorem 3.3 (Ramsey's Theorem for graphs with 2 colours)**.** *Let s be a positive integer. Then the finite integer $R(s, s)$ exists.*

Example - Theorem 3.2 tells us that $R(3, 3) \leq 6$. On the other hand, the fact that C_5 does not contain a clique or an independent set of size 3 tells us that $R(3, 3) > 5$. Therefore $R(3, 3) = 6$.

Exercises - Show that $R(2, k) = k$ for all $k \in \mathbb{N}$. Show that $R(1, k) = 1$ for all $k \in \mathbb{N}$.

Exercise - Prove that $R(s, t) = R(t, s)$.

In general, it is very difficult to find exact values of $R(s, t)$. For some small values of s and t the answer is known, but even the question of determining the precise value of $R(5, 5)$ is an open problem that is thought to be very difficult. Let us look at some of the small s and t for which complete answers are known.

Theorem 3.4. $R(3, 4) = 9$

Proof. First of all we will show that $R(3,4) \leq 9$. That is, we will show that any graph $G = (V, E)$ and with 9 vertices must contain either a clique of size 3 or an independent set of size 4. Suppose for a contradiction that our graph G does not contain a clique of size 3 or an independent set of size 4.

Case 1. Suppose that there is some $v \in V$ such that $|N(v)| \geq 4$. Then let $v_1, v_2, v_3, v_4 \in N(v)$. If none of these four vertices are adjacent, then they form an independent set of size 4, as required. Otherwise, two vertices $v_i, v_j \in N(v)$ are adjacent, in which case $\{v, v_i, v_j\}$ is a clique of size 3.

Case 2 Suppose that there is some $v \in V$ such that $|N(v)| \leq 2$. Then we have six vertices $v_1, \ldots, v_6 \notin N(v) \cup \{v\}$. Since $R(3,3) = 6$, it must be the case that three of these 6 vertices, say $v_i v_j, v_k$, form a clique or an independent set. If $\{v_i, v_j, v_k\}$ is a clique, then we have a clique of size 3 in G and we are done. If not, then the set $\{v, v_i, v_j,\}$ is an independent set of size 4 in G and we are done.

Case 3 Suppose that $|N(v)| = 3$ for all $v \in V$. But then, by the Handshake Lemma

$$2|E| = \sum_{v \in V} \deg(v) = 27.$$

This is a contradiction, since $|E|$ is an integer.

It remains to prove that $R(3,4) > 8$. This is left as an exercise.

□

The next result shows that we can use our knowledge of certain small Ramsey numbers to obtain upper bounds for larger Ramsey numbers.

Theorem 3.5. *Let $s, t \geq 2$ be integers and suppose that $R(s-1, t)$ and $R(s, t-1)$ exist. Then $R(s,t)$ exists and, in particular,*

$$R(s,t) \leq R(s-1,t) + R(s,t-1).$$

Proof. Let $G = (V, E)$ be a graph with $|V| = R(s-1,t) + R(s,t-1)$. We will show that G contains a clique of size s or an independent set of size t. Let $v \in V$ be arbitrary. Consider the two induced subgraphs $G_1 = G[N(v)]$ and $G_2 = G[V \setminus (N(v) \cup \{v\})]$.

Note that $|V(G_1)| + |V(G_2)| + 1 = |V| = R(s-1,t) + R(s,t-1)$. Therefore, at least one of the inequalities

$$|V(G_1)| \geq R(s-1,t) \quad \text{or} \quad |V(G_2)| \geq R(s,t-1)$$

must hold. In the first case, we know that G_1 contains either a clique C of size $s-1$ or an independent set I of size of size t. If G_1 contains a clique of size $s-1$ then $C \cup \{v\}$ is a clique of size s in G. Otherwise, the set I is an independent set of size t in G.

On the other hand, if $|V(G_2)| \geq R(s,t-1)$ then G_2 contains a clique C of size s or an independent set I of size $t-1$. If we are in the first case then C is a clique of size s in G. Otherwise, the set $I \cup \{v\}$ is an independent set of size s in G.

□

Example - It follows from Theorems 3.5 and 3.4 that $R(4,4) \leq 18$. Indeed,

$$R(4,4) \leq R(3,4) + R(4,3) = 2R(3,4) = 18.$$

We can be even more precise.

Theorem 3.6. $R(4,4) = 18$.

Proof. To follow. $\square$

As well as giving us the exact value for certain Ramsey numbers, Theorem 3.5 actually allows us to prove Ramsey's Theorem for $r = 2$ (i.e. Theorem 3.1 for $r = 2$, which was restated above as Theorem 3.3). In fact, it can be be used to prove that the asymmetric Ramsey number $R(s,t)$ exists, and even give us an upper bound for it.

In the next proof, we use the following combinatorial identity.

Exercise - Prove that

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}. \tag{2}$$

holds for all $k, n \in \mathbb{N}$ such that $n \geq k$.

Theorem 3.7. *For all integers $s, t \geq 2$, $R(s,t)$ exists. Moreover,*

$$R(s,t) \leq \binom{s+t-2}{s-1}.$$

Proof. The proof is by induction on $s + t$. The base case when $s = t = 2$ is trivial since

$$R(2,2) = 2 = \binom{2}{1} = \binom{s+t-2}{s-1}.$$

Now let $s, t \geq 2$ be integers not both equal to 2.

Case 1 - Suppose that one of s or t is equal to 2. Without loss of generality we may assume that $s = 2$ (if $t = 2$ then a symmetric argument works in the same way). Then,

$$R(s,t) = R(2,t) = t = \binom{t}{1} = \binom{s+t-2}{s-1}.$$

Case 2 Suppose that $s, t \geq 3$. By the induction hypothesis, $R(s-1,t)$ and $R(s,t-1)$ exists and

$$R(s,t-1) \leq \binom{s+t-3}{s-1}, \quad R(s-1,t) \leq \binom{s+t-3}{s-2}.$$

Combining this with Theorem 3.5 and (2) yields

$$R(s,t) \leq R(s-1,t) + R(s,t-1) \leq \binom{s+t-3}{s-2} + \binom{s+t-3}{s-1} = \binom{s+t-2}{s-1}.$$

$\square$

We are particularly interested in the quantity $R(s, s)$, and this theorem tells us that

$$R(s, s) \leq \binom{2s-2}{s-1}.$$

To get a better idea for how large this number is, we can use some well known facts about the central binomial coefficient, which tell us that $\binom{2s-2}{s-1}$ is asymptotically equal to

$$\frac{4^{s-1}}{\sqrt{\pi(s-1)}}$$

How close might this be to optimal? What lower bounds do we have for $R(s, s)$?

$$R(s, s) \geq 2^{\frac{s-2}{2}}. \tag{3}$$

So, we see that there is a large gap between the upper and lower bounds for $R(s, s)$, and this situation is quite typical in Ramsey Theory.

We will prove (3) in the next chapter, using probabilistic methods. We will also use probabilistic techniques to prove upper bounds for $R(4, t)$.

4 An introduction to the probabilistic method

The Probabilistic Method is a fundamental tool in discrete mathematics, and is responsible for a wealth of seminal results. The basic idea of the method is the following. Suppose that we would like to show the existence of a certain configuration, but we do not know how to do so explicitly. We can try to "construct" such a configuration randomly, by defining an appropriate probability space of structures, and showing that a random element of the space has the desired properties with positive probability.

We will learn more about what this really means by studying several simple examples in this section. As we move through these examples, we will recap some of the basic probability theory that will be used repeatedly throughout this course.

4.1 Lower bounds for Ramsey numbers

We start with a simple and classical example in Ramsey Theory. The *Ramsey number* $R(k,l)$ is the smallest integer n such that any graph on n vertices is guaranteed to contain either a clique of size k or an independent set of size l. Ramsey's Theorem tells us that $R(k,l)$ exists for any $k,l \in \mathbb{Z}$, but determining the correct values of $R(k,l)$ is a major (and very difficult) open problem in combinatorics.

We first consider the *diagonal* Ramsey numbers $R(k,k)$. To lower bound $R(k,k)$ we would like to construct a large graph with no cliques or independent sets of size k. To do this constructively appears to be very difficult, but a simple probabilistic argument gives the following starting point.

Theorem 4.1. *For all $k \geq 3$, $R(k,k) \geq \lfloor 2^{k/2} \rfloor$.*

Proof. We will show that,

$$\binom{n}{k} 2^{1-\binom{k}{2}} < 1 \Rightarrow R(k,k) > n. \tag{4}$$

This will imply the result by taking $n = \lfloor 2^{k/2} \rfloor$. Indeed, for all $k \geq 3$

$$\binom{n}{k} 2^{1-\binom{k}{2}} < \frac{n^k}{k!} 2^{1-\binom{k}{2}} \leq \frac{2^{k^2/2}}{k!} 2^{1-\binom{k}{2}} = \frac{2^{\frac{k}{2}+1}}{k!} < 1.$$

It remains to prove (4). Construct a random graph $G = (V,E)$ on n vertices as follows; for any two distinct vertices $x,y \in V$, $xy \in E$ with probability $1/2$. For any fixed set $R \subset V$ of size k, let X_R be the event that R forms a clique or an independent set in G. Observe that

$$\mathbb{P}[X_R] = 2^{1-\binom{k}{2}}$$

for any such R. By the union bound, we have

$$\mathbb{P}\left[\bigcup_{R \subset V : |R| = k} X_R\right] \leq \sum_{R \subset V : |R| = k} P[X_R] = \binom{n}{k} 2^{1-\binom{k}{2}} < 1.$$

Therefore, with positive probability, the event

$$\bigcup_{R\subset V:|R|=k} X_R$$

does not occur. That is, for every R set of k vertices, the event X_R does not occur, and thus there are no cliques or independent sets of size k in G.

Since this property occurs with positive probability, there exist such a graph G on n vertices, as required.

$\square$

Here we have used the *Union Bound* (also known as Boole's inequality); for a countable set of events $A_1, A_2, \ldots$ in a probability space

$$\mathbb{P}\left(\bigcup_i A_i\right) \leq \sum_i \mathbb{P}(A_i).$$

If the events $A_1, A_2 \ldots$ are pairwise disjoint then

$$\mathbb{P}\left(\bigcup_i A_i\right) = \sum_i \mathbb{P}(A_i).$$

4.2 Small dominating sets

Our next motivating application of the probabilistic method is slightly more delicate. We will make use of an essential tool that will reappear frequently in this course; linearity of expectation.

Recall the definition of expectation: let X be a random variable with a finite number of finite possible values $x_1, \ldots, x_k$, occurring with probabilities $p_1, \ldots, p_k$ respectively. The *expected value* of X is defined as

$$\mathbb{E}(X) = \sum_{i=1}^{k} x_i p_i.$$

We will lean heavily on the following elementary property of expectation, which we call *linearity of expectation*. Let $X_1, \ldots, X_n$ be random variables and $c_1, \ldots, c_n \in \mathbb{R}$. Then

$$\mathbb{E}(c_1 X_1 + \cdots + c_n X_n) = \sum_{i=1}^{n} c_i \mathbb{E}(X_i).$$

This property will be exploited heavily in section 2 of these notes.

A *dominating set* of an undirected graph $G = (V, E)$ is a subset $V' \subset V$ such that, for each $v \in V$, at least one of the following holds:

- $v \in V'$,
- there is some $v' \in V'$ such that $vv' \in E$.

Given a graph $G = (V, E)$, we may be interested to find a small dominating set. One would expect that this may become easier if the number of edges in the graph is somewhat large. The following result supports this intuition.

Theorem 4.2. *Let $G = (V, E)$ be a graph with minimum vertex degree $\delta \geq 0$. Then G has a dominating set $V' \subset V$ with*

$$|V'| \leq |V| \frac{1 + \ln(\delta + 1)}{\delta + 1}.$$

Proof. Let $p \in [0, 1]$, to be specified later. Construct a random set $X \subset V$, where each element $v \in V$ belongs to X with probability p. This random process is carried out independently for each vertex. Let $Y =: Y_X$ be the set

$$Y_X := \{v \in V \setminus X : N_G(v) \cap X = \emptyset\}.$$

Note that $\mathbb{E}(|X|) = |V|p$.

To calculate the expected size of Y, we need to compute the probability that a given vertex $v \in V$ belongs to Y. This is the same as the probability that both v and its neighbours are not in X. Since $|N_G(v)| \geq \delta$, we have

$$\mathbb{P}(v \in Y) \leq (1 - p)^{1+\delta}.$$

The random variable $|Y|$ can be written as a sum of random variables via

$$|Y| = \sum_{v \in V} Y(v).$$

Here we use the $Y(\cdot)$ for the indicator function of Y, so $Y(v) = 1$ if $v \in Y$ and $Y(v) = 0$ otherwise.

By linearity of expectation

$$\mathbb{E}(|Y|) = \sum_{v \in V} \mathbb{E}(Y(v)) \leq |V|(1 - p)^{1+\delta}.$$

Therefore,

$$\mathbb{E}(|X \cup Y|) = \mathbb{E}(|X| + |Y|) \leq |V|(p + (1 - p)^{1+\delta}),$$

and in particular, there exists at least one set X such that $U := X \cup Y_X$ has cardinality

$$|U| \leq |V|(p + (1 - p)^{1+\delta}).$$

It will be convenient to use the bound $1 - x \leq e^{-x}$, which is valid for all $x \in [0, 1]$, and is fairly tight for small values. It follows that

$$|U| \leq |V|(p + e^{-p(1+\delta)}). \tag{5}$$

The set U is a dominating set for G. Indeed, every vertex of V is either in U, or has a neighbour in U.

The inequality (5) holds for all $p \in [0,1]$. The last task is to make a choice of p which minimises the right hand side of (5). To do this, let

$$f(p) = p + e^{-p(1+\delta)},$$

and differentiate to find a minimum. So,

$$f'(p) = 1 - (1+\delta)e^{-p(1+\delta)}$$

and solving $f'(p) = 0$ gives

$$p = \frac{\ln(\delta+1)}{\delta+1}. \tag{6}$$

This gives a minimum value of $f(p)$, and so we choose p as in (6). Note that $p \in [0,1]$. We finally conclude from (5) that there is a dominating set U with size

$$|U| \leq |V| \left(\frac{\ln(\delta+1)}{\delta+1} + \frac{1}{\delta+1} \right) = |V| \left(\frac{\ln(\delta+1)+1}{\delta+1} \right).$$

□

As well as the aforementioned linearity of expectation, there are some other key ideas in this proof which will reappear throughout this course.

- The alteration principle: the randomly chosen set X did not immediately give us what we need, but it got us close. A small alteration to the set, in this case, adjoining Y_X, gave us a set with the required properties. Many more examples of the alteration principle will be explored in section 3.
- The probability p can be treated as a variable, and then specified later in the proof in an optimal way.
- We used the estimate $1 - x \leq e^{-x}$ in order to make the calculations more manageable. The actual minimum was a little tricky to deal with, but by giving up a small quantitative loss, we can bypass this problem. Finding the right trade-off between accuracy and practicality is a key part of the art of the probabilistic method.

4.3 Set systems

Next, we will see our first application of the probabilistic method to yield a fully optimal result. A family

$$\mathcal{F} = \{(A_i, B_i) : 1 \leq i \leq h\}$$

of pairs of sets is a (k,l)-*system* if

$$|A_i| = k \ \ \forall i, \quad |B_j| = l \ \ \forall j, \quad A_i \cap B_i = \emptyset \ \ \forall i, \quad A_i \cap B_j \neq \emptyset \ \ \forall i \neq j.$$

The following result is due to Bollobás.

Theorem 4.3. *If $\mathcal{F} = \{(A_i, B_i) : 1 \leq i \leq h\}$ is a (k,l)-system then*

$$h \leq \binom{k+l}{k}.$$

Proof. Let

$$C = \bigcup_{i=1}^{h} (A_i \cup B_i).$$

Randomly order the elements of C. That is, from all the $|C|!$ bijections $\sigma : C \to \{1, \dots, |C|\}$, choose one at random.

Let X_i be the event that

$$(a, b) \in A_i \times B_i \Rightarrow \sigma(a) < \sigma(b). \tag{7}$$

Note that $\mathbb{P}(X_i) = 1/\binom{k+l}{k}$. Indeed, from each of the $(k + l)!$ orderings of $A_i \cup B_i$, there are $k!l!$ such orderings which satisfy (7), and so

$$\mathbb{P}(X_i) = \frac{k!l!}{(k + l)!} = \frac{1}{\binom{k+l}{k}}.$$

Also, the events X_i are disjoint. Indeed, suppose that both X_i and X_j occur, and assume wlog that

$$\max \sigma(A_i) \le \max \sigma(A_j).$$

Then for all $a \in A_i$ and $b \in B_j$,

$$\sigma(a) \le \max \sigma(A_j) < \sigma(b),$$

but this contradicts the assumption that $A_i \cap B_j \neq \emptyset$.

Therefore,

$$1 \ge \mathbb{P}\left(\bigcup_{i=1}^{h} X_i\right) = \sum_{i=1}^{h} \mathbb{P}(X_i) = \frac{h}{\binom{k+l}{k}}.$$

□

To see that this is indeed optimal, one may take $X = \{1, \dots, k + l\}$ and construct the (k, l)-system

$$\mathcal{F} = \{(A, X \setminus A) : A \subset X, |A| = k\}.$$

4.4 Sum-free sets

A set $A \subset \mathbb{Z}$ is *sum-free* if there are no solutions to

$$a + b = c, \quad a, b, c \in A.$$

The following result shows that every set of integers contain a somewhat large sum-free subset.

Theorem 4.4. *Every set $A \subset \mathbb{Z}^*$ contains a sum-free subset $A' \subset A$ such that $|A'| > |A|/3$.*

Proof. Let $p = 3k+2$ be a sufficiently large prime (there are infinitely many such primes by Dirichlet's Theorem, and so we can always find one bigger than a certain threshold).

Embed A as a subset of $\mathbb{F}_p$ in the obvious way. That is, we map $\{1, \dots, p\}$ to $\mathbb{F}_p$ via the function $\pi(a) = a \mod p$. Write $\mathcal{A} = \pi(A)$.

Let $C \subset \mathbb{F}_p$ be the middle third of $\mathbb{F}_p$, so

$$C = \{k+1, k+2, \dots, 2k+1\}.$$

Note that C is sum-free as a subset of $\mathbb{F}_p$, since $C + C = \{2k+2, 2k+3, \dots, p-1, 0, 1, \dots, k\}$. Let $x \in \mathbb{F}_p^*$ be chosen uniformly at random. Then, by linearity of expectation,

$$\mathbb{E}(|x\mathcal{A} \cap C|) = \mathbb{E}\left(\sum_{a \in \mathcal{A}} C(xa)\right) = \sum_{a \in \mathcal{A}} \mathbb{E}(C(xa)) = \sum_{a \in \mathcal{A}} \frac{k+1}{3k+1} > |A|/3. \tag{8}$$

In the previous line, $C(x)$ denotes the indicator function for the set C. That is,

$$C(x) := \begin{cases} 1 & \text{if } x \in C \\ 0 & \text{otherwise.} \end{cases}$$

In particular, it follows from (8) that there exists some $x \in \mathbb{F}_p^*$ such that $|x\mathcal{A} \cap C| > |A|/3$.

Now let $\mathcal{A}' \subset \mathcal{A}$ be the set $\mathcal{A}' = \mathcal{A} \cap x^{-1}C$. Note that $|\mathcal{A}'| > |A|/3$. Also, the corresponding set $A' := \pi^{-1}(\mathcal{A}') \subset A$ is a sum-free set of size $> |A|/3$. Indeed, suppose that we have a solution to $a + b = c$ with $a, b, c \in A'$. Then

$$x\pi(a) + x\pi(b) = x\pi(c)$$

and $x\pi(a), x\pi(b), x\pi(c) \in C$. This contradicts the fact that C is sum-free in $\mathbb{F}_p$.

□

5 Linearity of expectation

One of the tools used that will be used frequently in this course is *linearity of expectation.* We have already seen this once, in the proof of Theorem 4.2. We recall the statement below.

Proposition 5.1 (Linearity of Expectation)**.** *Let $X_1, \dots, X_n$ be random variables and $c_1, \dots, c_n \in \mathbb{R}$. Then*

$$\mathbb{E}(c_1 X_1 + \cdots + c_n X_n) = \sum_{i=1}^{n} c_i \mathbb{E}(X_i).$$

Note that we make no assumption on the dependence of the random variables. This makes this a powerful and versatile tool. In many cases, the random variables X_i will be characteristic functions for certain events.

5.1 Finding large bipartite subgraps

Recall that a graph $G = (V, E)$ is *bipartite* if we can write V as a disjoint union $V = V_1 \sqcup V_2$ such that

$$uv \in E \Rightarrow u \in V_1, v \in V_2 \text{ or } u \in V_2, v \in V_1.$$

Bipartite graphs are often convenient descriptions for mathematical or physical phenomena, and so it may be useful to know if a certain graph contains a bipartite subgraph. Let us see one simple example of an application of Proposition 5.1 in this direction.

Theorem 5.2. *Let $G = (V, E)$ be a graph with n vertices and e edges. Then G contains a bipartite subgraph with at least $e/2$ edges.*

Proof. Let $T \subset V$ be a random subset of V, where each vertex $v \in V$ belongs to T with probability 1/2. Let $B = V \setminus T$. An edge $xy \in E$, is said to be *crossing* if one of the vertices is in T and the other is in B. Define the indicator random variable

$$X_{xy} = \begin{cases} 1 & \text{if } xy \text{ is crossing} \\ 0 & \text{otherwise.} \end{cases}$$

The expected value of X_{xy} is equal to the probability that xy is crossing, which is 1/2.

Let X denote the number of crossing edges in G, and note that $X = \sum_{xy \in E} X_{xy}$. Therefore, by linearity of expectation

$$\mathbb{E}(X) = \sum_{xy \in E} \mathbb{E}(X_{xy}) = e/2.$$

In particular, there exists a random choice of vertices T for which $X \geq e/2$. We then use this partition to define a bipartite subgraph, as required.

□

A small refinement of the choice of probability space gives a small improvement.

Theorem 5.3. *Let $G = (V, E)$ be a graph with $2n$ vertices and e edges. Then G contains a bipartite subgraph with at least $\frac{en}{2n-1}$ edges.*

Proof. Choose T uniformly at random from all subsets of V with cardinality n.

Define the indicator random variable

$$X_{xy} = \begin{cases} 1 & \text{if } xy \text{ is crossing} \\ 0 & \text{otherwise.} \end{cases}$$

The expected value of X_{xy} is equal to the probability that xy is crossing, but this probability is now $\frac{n}{2n-1}$. Indeed, the probability that $x \in T$ and $y \in B$ is equal to

$$\frac{\binom{2n-2}{n-1}}{\binom{2n}{n}} = \frac{n}{2(2n-1)}.$$

By symmetry, the probability that $x \in B$ and $y \in T$ is the same, and so the total probability of xy crossing is $\frac{n}{2n-1}$. Let X denote the number of crossing edges in G, and note that $X = \sum_{xy \in E} X_{xy}$. Therefore, by linearity of expectation

$$\mathbb{E}(X) = \sum_{xy \in E} \mathbb{E}(X_{xy}) = \frac{en}{2n-1}.$$

Thus there exist sets V and B which gives a partition for a bipartite graph with at least $\frac{en}{2n-1}$ edges, as required.

$\square$

Exercise - Prove the following: If $G = (V, E)$ is a graph with $2n+1$ vertices and e edges, then G contains a bipartite subgraph with at least $\frac{e(n+1)}{2n+1}$ edges.

Exercise - Show that the conclusion in the previous exercise and Theorem 5.3 cannot be improved.

5.2 A Ramsey type problem

Theorem 5.4. *Let $a, n \in \mathbb{N}$. There is a graph on n vertices with at most*

$$\binom{n}{a} 2^{1-\binom{a}{2}}$$

cliques and independent sets of size a.

Proof. Construct a random graph $G = (V, E)$ on n vertices. Each potential edge uv has a probability $1/2$ of belonging to E.

Let the random variable X denote the number of cliques and independent sets of size a in G. Also, for each set $U \subseteq V$, define the random variable

$$X_U = \begin{cases} 1 & \text{if } U \text{ is forms a clique or independent set} \\ 0 & \text{otherwise.} \end{cases}$$

Then we have

$$X = \sum_{U \subseteq V : |U| = a} X_U,$$

and hence by linearity of expectation

$$\mathbb{E}(X) = \sum_{U \subset V : |U| = a} \mathbb{E}(X_U) = \sum_{U \subset V : |U| = a} \mathbb{P}(U \text{ is a clique or independent set})$$
$$= \sum_{U \subset V : |U| = a} 2^{1-\binom{a}{2}} = \binom{n}{a} 2^{1-\binom{a}{2}}.$$

There exists a graph which has at most the expected number of cliques and independent sets of size a, and the proof is complete. □

Note that we can use this result to deduce Theorem 4.1. Try and do this yourself!

5.3 The crossing number

Recall that we proved earlier in Proposition 2.1 that a sufficiently dense graph cannot be planar. In particular, it follows from Proposition 2.4 that any planar graph $G = (V, E)$ satisfies

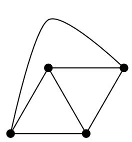
$$|E| \leq 3|V|. \tag{9}$$

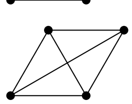
What happens after this threshold? Are we guaranteed that a graph with many more than this number of edges is in some sense far from planar?

To tackle this question more precisely, we need the notion of the *crossing number* of a graph. A *drawing* of G is a representation of G in the plane $\mathbb{R}^2$, where each vertex is represented by a point, and each edge $e = uv$ is represented with a curve connecting u and v. Two edges uv and $u'v'$ with $\{u, v\} \cap \{u', v'\} = \emptyset$ cross in a drawing if the edges in that drawing intersect.

The *crossing number of a drawing* of a graph G is the number of crossing pairs of edges in that drawing. The crossing number of a graph G, denoted $cross(G)$, is the minimum crossing number over all drawings.

There is an important distinction between the crossing number of a drawing and the crossing number of a graph itself. Consider the example K_4 mentioned above. Here are two drawings of it.





The first drawing has crossing number 0 and the second has crossing number 1. However, the crossing number of the abstract graph K_4 is the minimum over all drawings. The first drawing above shows that K_4 has crossing number 0.

So, according to this definition, a planar graph G has $cross(G) = 0$, and non-planar graph must have crossing number at least 1. But don't we expect the crossing number to increase significantly if there are many more than $3|V|$ edges? The following theorem, known as the Crossing Number Lemma, supports this intuition.

Theorem 5.5. *Let $G = (V, E)$ be a graph with $|E| \geq 4|V|$. Then $cross(G) \geq \frac{|E|^3}{64|V|^2}$.*

Proof. First we establish that the weaker bound

$$cross(G) \geq |E| - 3|V| \tag{10}$$

holds for any graph $G = (V, E)$. Indeed, any graph $G = (V, E)$ can be transformed into a planar graph by removing $cross(G)$ edges. Then by (9),

$$|E| - cross(G) \leq 3|V|.$$

Now, let $0 < p \leq 1$, the precise value of p will be specified later. Let V' be a random subset of V, where each element of V belongs to V' with probability p. Let $G' = (V', E')$ be the *induced subgraph* spanned by V'. That is

$$E' := \{uv \in E : u, v \in V'\}.$$

By (10) and linearity of expectation,

$$\mathbb{E}(cross(G')) \geq \mathbb{E}(|E'| - 3|V'|) = \mathbb{E}(|E'|) - 3\mathbb{E}(|V'|) = p^2|E| - 3p|V|. \tag{11}$$

Now consider a drawing of G with the minimal number $cross(G)$ of crossings. Each crossing in this drawing survives the random process if and only if all four of the relevant vertices are in V'. Therefore,

$$\mathbb{E}(cross(G')) \leq p^4 cross(G).$$

Combining this with (11) gives

$$cross(G) \geq \frac{|E|}{p^2} - 3\frac{|V|}{p^3}.$$

The proof is complete if we set

$$p = \frac{4|V|}{|E|}.$$

Note that this is where the bound $|E| \geq 4|V|$ has been used.

$\square$

We will now use the Crossing Number Lemma to prove the Szemerédi-Trotter Theorem, which has been a hugely influential result in combinatorial geometry.

Theorem 5.6 (Szemerédi-Trotter Theorem)**.** *Let P be a set of m points in $\mathbb{R}^2$ and let L be a set of n lines in $\mathbb{R}^2$. Then*

$$I(P, L) \leq 4m^{2/3}n^{2/3} + 4m + n.$$

Proof. We can naturally turn the drawing of the sets P and L into a drawing of a graph $G = (V, E)$. The vertices of G are the points of P, and $p_1p_2 \in E$ if and only if p_1 and p_2 belong to a line $l \in L$ and they are neighbours on that line (i.e there is no $p \in P \cap l$ lying in between p_1 and p_2).

Note that a line $l \in L$ with $|l \cap P| = k$ contributes $k - 1$ edges to the graph G. Therefore

$$|E| = \sum_{l \in L}(|l \cap P| - 1) = \left(\sum_{l \in L}|l \cap P|\right) - n = I(P, L) - n. \tag{12}$$

On the other hand, we can bound $|E|$ via the Crossing Number Lemma. Since any two lines cross in at most one point, the number of crossings in this drawing of G is at most n^2. So, $cross(G) \leq n^2$. By the Crossing Number Lemma, either

$$|E| < 4|V| = 4m,$$

or

$$|E| \leq 4|V|^{2/3} cross^{1/3}(G) \leq 4m^{2/3}n^{2/3}.$$

So, certainly

$$|E| \leq 4m^{2/3}n^{2/3} + 4m$$

and finally by (12)

$$I(P, L) \leq 4m^{2/3}n^{2/3} + 4m + n.$$

□

5.4 Set systems

Given a ground set $[n]$, how large can a family $\mathcal{F}$ of subsets of $[n]$ be such that it has the property that for any two distinct sets $A, B \in \mathcal{F}$, both $A \not\subseteq B$ and $B \not\subseteq A$ hold?

Consider the case when $\mathcal{F}$ is the set of all subsets of $[n]$ of cardinality k. Then $\mathcal{F}$ has the desired property. Indeed, no two finite sets of the same cardinality can be subsets of one another, unless they are identical. So, we can have such a family with

$$|\mathcal{F}| \geq \binom{n}{k}$$

for any k. This value is maximised when $k = \lfloor n/2 \rfloor$.

The following result shows that this simple construction is optimal.

Theorem 5.7. *Let $\mathcal{F}$ be a family of subsets of $[n]$ with the property that for any two distinct sets $A, B \in \mathcal{F}$, both $A \not\subseteq B$ and $B \not\subseteq A$ hold. Then*

$$|\mathcal{F}| \leq \binom{n}{\lfloor n/2 \rfloor}.$$

Proof. Let σ be a random permutation of $[n]$. Consider the random variable

$$X = |\{i : \{\sigma(1), \sigma(2), \ldots, \sigma(i)\} \in \mathcal{F}\}|,$$

and note that $X \leq 1$. Indeed, if two distinct $i, j \in [n]$ are both in the set above, supposing without loss of generality that $\sigma(i) < \sigma(j)$, we have both

$$\{\sigma(1), \ldots, \sigma(i)\}, \{\sigma(1), \ldots, \sigma(j)\} \in \mathcal{F}.$$

But this is a contradiction, since the first set is a subset of the second.

On the other hand, let X_i be the characteristic function of the event that $\{\sigma(1), \ldots, \sigma(i)\} \in \mathcal{F}$. Then $X = \sum_i X_i$, and so by linearity of expectation

$$1 \geq \mathbb{E}(X) = \sum_i \mathbb{E}(X_i). \tag{13}$$

Now, decompose $\mathcal{F} = \mathcal{F}_0 \cup \mathcal{F}_1 \cup \cdots \cup \mathcal{F}_n$, where $\mathcal{F}_i$ is the set of all elements of $\mathcal{F}$ with cardinality i. The expectation of X_i is the probability that $\{\sigma(1), \ldots, \sigma(i)\} \in \mathcal{F}$, which is the same as the probability that $\{\sigma(1), \ldots, \sigma(i)\} \in \mathcal{F}_i$. Therefore,

$$\mathbb{E}(X_i) = \frac{|\mathcal{F}_i|}{\binom{n}{i}} \geq \frac{|\mathcal{F}_i|}{\binom{n}{\lfloor n/2 \rfloor}}.$$

We conclude from (13) that

$$1 \geq \frac{1}{\binom{n}{\lfloor n/2 \rfloor}} \sum_i |\mathcal{F}_i| = \frac{1}{\binom{n}{\lfloor n/2 \rfloor}} |\mathcal{F}|,$$

as required. □

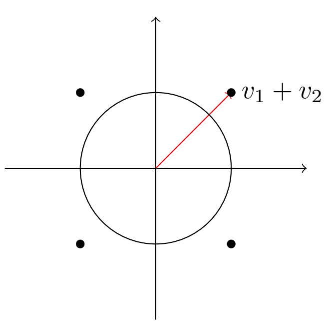
5.5 Balancing vectors

Let v_1 and v_2 be arbitrary unit vectors in $\mathbb{R}^2$. Does there always exist $\epsilon_1, \epsilon_2 \in \{-1, 1\}$ such that

$$\|\epsilon_1 v_1 + \epsilon_2 v_2\|$$

is small? So roughly speaking, can we balance these vectors so that, taking positive and negative combinations of them and going for a walk, we do not stray too far from the original position.

Consider the following simple example: $v_1 = (1, 0)$ and $v_2 = (0, 1)$. There are four choices for the pair (ϵ_1, ϵ_2), and they are all as bad as each other:



The four points here are the four possible vectors $\epsilon_1 v_1 + \epsilon_2 v_2$. We can quickly calculate that all of these vector sums have length $\sqrt{2}$.

The next result shows that this is the absolute worst case.

Theorem 5.8. *For any $v_1, v_2 \in \mathbb{R}^2$ such that $\|v_1\| = \|v_2\| = 1$, there exists $\epsilon_1, \epsilon_2 \in \{-1, 1\}$ such that*

$$\|\epsilon_1 v_1 + \epsilon_2 v_2\| \leq \sqrt{2}.$$

Proof. Choose ϵ_1 and ϵ_2 uniformly at random from $\{-1, 1\}$. Write $v_1 = (x_1, y_1)$ and $v_2 = (x_2, y_2)$. Consider the random variable

$$\begin{aligned}
X &= \|\epsilon_1 v_1 + \epsilon_2 v_2\|^2 \\
&= \|(\epsilon_1 x_1, \epsilon_1 y_1) + (\epsilon_2 x_2, \epsilon_2 y_2)\|^2 \\
&= \|(\epsilon_1 x_1 + \epsilon_2 x_2, \epsilon_1 y_1 + \epsilon_2 y_2)\|^2 \\
&= (\epsilon_1 x_1 + \epsilon_2 x_2)^2 + (\epsilon_1 y_1 + \epsilon_2 y_2)^2 \\
&= x_1^2 + x_2^2 + 2\epsilon_1\epsilon_2 x_1 x_2 + y_1^2 + y_2^2 + 2\epsilon_1\epsilon_2 y_1 y_2 \\
&= 2 + 2\epsilon_1\epsilon_2(x_1 x_2 + y_1 y_2).
\end{aligned}$$

Therefore

$$\mathbb{E}[X] = 2 + 2(x_1 x_2 + y_1 y_2)\mathbb{E}(\epsilon_1\epsilon_2).$$

It remains to calculate the expected value of $\epsilon_1\epsilon_2$. This product has value either 1 or -1, each with probability 1/2. Then by definition of expectation

$$\mathbb{E}(\epsilon_1\epsilon_2) = \frac{1}{2}1 + \frac{1}{2}(-1) = 0.$$

So, $\mathbb{E}(X) = 2$, and thus there exist ϵ_1, ϵ_2 such that

$$\|\epsilon_1 v_1 + \epsilon_2 v_2\|^2 \leq 2,$$

which concludes the proof. $\square$

6 Alterations Method

In the previous two chapters, we have proven the existence of structures with certain desired properties by constructing probability spaces and showing that the desired structure exists in this space with positive probability. It may be the case that the random structure does not have the desired property, but is somewhat close to having it. In this chapter, we look at such cases, and see how we can make some small alterations to the randomly generated sets to get the properties we want.

6.1 Ramsey numbers

We begin by returning to the problem of lower bounding Ramsey numbers, and see how the alteration method can give better results than those given by the simple approach in Chapter 1.

Theorem 6.1. *For any integer n,*

$$R(k,k) > n - \binom{n}{k} 2^{1-\binom{k}{2}}$$

Proof. Recalling the definition of $R(k,k)$, our task is to find a graph on $n - \binom{n}{k}2^{1-\binom{k}{2}}$ vertices which does not contain any cliques or independent sets of size k.

Construct a random graph $G = (V, E)$ on n vertices. Each pair of distinct vertices from V form an edge with probability $1/2$. For any $R \subset V$ of size k, let X_R be the indicator random variable for the event that R forms a clique or independent set. Set

$$X = \sum_{R \subset V : |R| = k} X_R.$$

So X is a random variable counting the number of large cliques and independent sets. Then by linearity of expectation

$$\mathbb{E}(X) = \sum_R \mathbb{E}(X_R) = \binom{n}{k} 2^{1-\binom{k}{2}}.$$

Therefore, there exists a graph on n vertices with at most

$$\binom{n}{k} 2^{1-\binom{k}{2}}$$

subsets of size k which are either a clique or independent set. Let G be such a graph.

Now, for each clique or independent set in G, remove one vertex from G and take the induced subgraph. At the end of this deletion process we have a graph with at least

$$n - \binom{n}{k} 2^{1-\binom{k}{2}}$$

vertices, and with no cliques or independent sets of cardinality k. $\square$

How does this compare with Theorem 4.1? If we take

$$n = k2^{k/2}/e$$

and use our favourite binomial coefficient estimate $\binom{a}{b} \leq \left(\frac{ea}{b}\right)^b$, it follows that

$$R(k,k) \geq \left(\frac{k}{e} - 2\right) 2^{k/2}.$$

So, as k gets large, we win a factor of k/e compared to Theorem 4.1.

We can also get a better estimate for $R(4,k)$ than we obtained in chapter 1. This will be explored in an upcoming exercise sheet.

6.2 Graph Theory

Dense graphs contain large cliques. This is an important fact in graph theory, most famously present in Turán's Theorem. The following result gives a slightly weaker form of Turán's Theorem, but stated in the complimentary form (i.e. sparse graphs have large independent sets).

Theorem 6.2. *Let $G = (V, E)$ be a graph with $|E| \geq n/2$. Then $\alpha(G) \geq \frac{n^2}{4|E|}$.*

Proof. Let $S \subset V$ be a random subset defined by

$$\mathbb{P}(v \in S) = p.$$

The value of $0 \leq p \leq 1$ will be chosen later. Consider the random variables $X = |S|$ and Y the number of edges in the induced subgraph $G[S]$ (that is, the graph with vertex set S and edge set to be the edges of E contained in S). For each $e = \{i, j\} \in E$, let Y_e be the indicator random variable for the event that $i, j \in S$. Then

$$\mathbb{E}[Y_e] = p^2.$$

By linearity of expectation

$$\mathbb{E}[Y] = |E|p^2.$$

Also, $\mathbb{E}[X] = pn$. Therefore, by linearity of expectation again,

$$\mathbb{E}[X - Y] = \mathbb{E}[X] - \mathbb{E}[Y] = pn - p^2|E|.$$

Therefore, there exists a choice of S such that the induced subgraph $G[S] = (S, E')$ has $|S| - |E'| \geq pn - p^2|E|$. We can then find an independent set inside S by removing one vertex for each edge. The resulting independent set has size at least $pn - p^2|E|$.

It remains to choose p. We balance the two terms by taking $p = n/(2|E|)$. This is where we use the assumption that $|E| \geq n/2$. We conclude that

$$\alpha(G) \geq pn - p^2|E| = \frac{n^2}{4|E|}.$$

□

6.2.1 The Erdős-Rogers Problem

Given a graph $G = (V, E)$ which does not contain any cliques of size 4, is it always true that G contains a large induced subgraph which does not contain any cliques of size 3 (aka triangles). The next argument answers with a partial "yes". (By a partial yes, I mean that the answer is yes for a suitable meaning of "large"!)

Theorem 6.3. *Let $G = (V, E)$ be a graph on n vertices such that $\omega(G) \leq 3$. Then G contains an induced subgraph H with at least $\frac{1}{2}\sqrt{n}$ vertices such that $\omega(H) \leq 2$.*

Proof. We will use Theorem 6.2.

Case 1 Suppose first of all that $|E| \leq \frac{1}{2}n^{3/2}$. Then Theorem 6.2 tells us that $\alpha(G) \geq \frac{n^{1/2}}{2}$, and this large independent set certainly has no triangles.

Note that this application of Theorem 6.2 also works if $|E| < n/2$. Indeed, if this is the case we can construct a supergraph $G' = (V, E')$ with $E \subset E'$. Keep adding edges to E' arbitrarily until $|E'| = \lceil n/2 \rceil$. Then apply the theorem to G' to conclude that $\alpha(G') \geq n^2/(4|E'|)$. Since G is a subgraph of G' with the same vertex set, $\alpha(G) \geq \alpha(G')$. Thus

$$\alpha(G) \geq \frac{n^2}{4\lceil n/2 \rceil} \geq \frac{n^2}{2(n+1)} \geq \frac{n^{1/2}}{2}.$$

The last inequality holds for all $n \geq 3$. We do not need to worry about smaller n, as then the theorem is true for trivial reasons.

Case 2 So, we may assume that $|E| \geq \frac{1}{2}n^{3/2}$. By the Handshake Lemma

$$\sum_{v \in V} d(v) = 2|E| \geq n^{3/2}$$

and so so there is some $v \in V$ such that $d(v) \geq n^{1/2}$. However, since G has no cliques of size 4, the neighbourhood of a vertex v in G has no triangles. this completes the proof.

□

Note that there was some space for optimising the constant $\frac{1}{2}$ in the previous statement, but we have skipped this. One can in fact get the bound $\sqrt{n}$ without using probability at all.

On the other hand, a beautiful proof of Wolfovitz established the existence of a graph G with no cliques of size 4, but with the property that every subset of size larger than $n^{1/2}(\log n)^{120}$ contains a triangle. So Theorem 6.3 is close to optimal. The proof of Wolfovitz uses probabilistic arguments very elegantly, although it uses more advanced tools than those we have so far encountered. I hope to prove this theorem towards the end of the course.

6.3 Combinatorial Geometry

Here is geometric question with a similar flavour to the Erdős-Rogers problem. Let P be a point set in the plane with no four points from P contained in a single line. Does there always exists a large subset $P' \subset P$ such that P' does not contain any three points on a line?

The answer is again a partial "yes". We first prove the following lemma. A *collinear triple* is a set of three distinct points which belong to the same line.

Lemma 6.4. *Let $P \subset \mathbb{R}^2$ be a set of points such that no 4 elements of P lie on a line. Then P contains at most $\binom{|P|}{2}$ collinear triples.*

Proof. Fix two distinct points $p, q \in P$. There are $\binom{|P|}{2}$ choices for the pair, and there is a unique line ℓ passing through both points. Since P contains no 4 points on a line, there is at most one other point on ℓ. Therefore, each pair of points belongs to at most one collinear triple.

$\square$

Theorem 6.5. *Let $P \subset \mathbb{R}^2$ be a point set with $|P| = n$ and with no 4 points on a line. Then there exists a subset $P' \subset P$ such that $|P'| \geq \frac{\sqrt{n}}{2}$ and P' does not contain any collinear triples.*

Proof. Construct a random subset $Q \subset P$, where each element $x \in P$ belongs to Q with probability p, independently for each such x. The value of $0 \leq p \leq 1$ will be chosen later. Observe that

$$\mathbb{E}[|Q|] = pn.$$

Let Y be the random variable giving the number of collinear triples in Q. For each collinear triple $t = \{x_1, x_2, x_3\}$ in P, let Y_t denote the indicator random variable for the event that $t \subset Q$. Then $\mathbb{E}[Y_t] = p^3$ and by linearity of expectation

$$\mathbb{E}[Y] = \sum_t \mathbb{E}[Y_t] \leq \binom{n}{2} p^3,$$

where the latter inequality uses Lemma 6.4.

By another application of linearity of expectation

$$\mathbb{E}[|Q| - Y] \geq pn - \binom{n}{2} p^3.$$

Choose $p = \frac{1}{\sqrt{n}}$ to approximately maximise the quantity on the right hand side (you can get a slightly better constant by differentiating, but I prefer to keep the calculations as clean as possible and lose something in the constant). This gives

$$\mathbb{E}[|Q| - Y] \geq \frac{n^{1/2}}{2}.$$

In particular, there exists a subset Q such that the size of Q minus the number of collinear triples in Q is at least $\frac{n^{1/2}}{2}$.

We can prune Q to get a collinear triple free subset P'. For each collinear triple in Q, delete one of the elements of the triple from Q. Repeat until no collinear triples remain. The resulting set P' has size at least $\frac{n^{1/2}}{2}$.

$\square$

How close is Theorem 6.5 to being optimal? In fact, this question is wide-open. Theorem 6.5 tells us that we can always find a collinear triple-free set of size $\Omega(n^{1/2})$. Can we do better? A small improvement was given by Füredi, using some non-trivial graph theory. Still, it is not known whether or not there is guaranteed to be a collinear triple free subset with size $\Omega(n^{1/2+c})$ for some $c > 0$.

From the other side, Balogh and Solymosi showed that, for all $c > 0$ and for all sufficiently large n there exists a set P of n points in $\mathbb{R}^2$ with no collinear quadruples, but with the property that for every $P' \subset P$ such that $|P'| \geq n^{5/6+c}$, P' contains a collinear triple. So, roughly speaking the correct exponent is somewhere in between $1/2$ and $5/6$. The techniques used by Balogh and Solymosi rely on some cutting edge graph theory which is beyond the scope of this course.

7 Second and Higher Moments

7.1 Chebyshev's Inequality

After the expectation, the second most important statistic for a random variable X is its variance. Recall the definition:

$$Var[X] = \mathbb{E}[(X - \mathbb{E}[X])^2].$$

By linearity of expectation, this can also be written as

$$Var[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2 + \mathbb{E}[X]^2 - 2X\mathbb{E}[X]] = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

This gives a measure of how far X strays from its expectation. Following standard notation, let μ denote expectation and σ^2 denote variance.

We would like a result which tells us that, if X has small variance then X is almost always close to its expectation. Our basic tool of this type if Chebyshev's Inequality.

Theorem 7.1 (Chebyshev's Inequality)**.** *Let X be a random variable taking finitely many possible outcomes and set $\mu = \mathbb{E}[X]$. Then for any positive λ*

$$\mathbb{P}[|X - \mu| \geq \lambda\sigma] \leq \frac{1}{\lambda^2}.$$

Proof.

$$\begin{aligned}\sigma^2 = Var[X] = \mathbb{E}[(X-\mu)^2] = \sum_{x \geq 0} x^2 \mathbb{P}[|X - \mu| = x] &\geq \sum_{x : x \geq \lambda\sigma} \lambda^2\sigma^2 \mathbb{P}[|X - \mu| = x] \\ &= \lambda^2\sigma^2\mathbb{P}[|X - \mu| \geq \lambda\sigma].\end{aligned}$$

□

This inequality can be very useful if we can calculate the variance of a random variable. Unfortunately, this calculation is not always so straightforward to carry out, and is much trickier to calculate than the expectation.

Again, we often consider random variables which can be decomposed as a sum of smaller random variables. Suppose that

$$X = X_1 + \cdots + X_n.$$

It can be calculated (exercise) that

$$Var[X] = \sum_i Var[X_i] + \sum_{1 \leq i < j \leq n} 2 \cdot Cov[X_i, X_j], \tag{14}$$

where $Cov[X_i, X_j]$ denotes the *covariance* of X_i and X_j:

$$Cov[X_i, X_j] := \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j].$$

If the random variables X_i and X_j are independent then $Cov[X_i, X_j] = 0$ and calculations become a little easier.

Suppose for instance that $X = X_1 + \cdots + X_n$ and the X_i are indicator random variables (this is a scenario we have seen repeatedly in the previous two chapters). That is, $X_i = 1$ if a certain event A_i occurs, and $X_i = 0$ if it does not. Let $\mathbb{P}[A_i] = p_i$ and suppose that the events $A_1, \ldots, A_n$ are independent of each other (which has also often been the case in the previous chapter). Then

$$Var[X] = \sum_i Var[X_i] = \sum_i (p_i - p_i^2) = \sum_i p_i(1 - p_i) \leq \sum_i p_i = \mathbb{E}[X].$$

7.2 Bounding the middle binomial coefficient

We begin with an elegant application of Chebyshev's Inqeuality. What is the size of the middle (and therefore largest) binomial coefficient? Let n be a natural number and consider the quantity

$$\binom{2n}{n}.$$

This is an important binomial coefficient as it forms part of the definition of the Catalan numbers which form part of the answer to various fundamental counting problems.

Let's focus here on lower bounding $\binom{2n}{n}$. Using our favourite bounds for binomial coefficients estimate

$$\binom{a}{b} \geq \left(\frac{a}{b}\right)^b$$

gives

$$\binom{2n}{n} \geq 2^n.$$

We can do better than this using Chebyshev's Inequality, almost improving the exponent n to $2n$.

Proposition 7.2. *For all $n \geq 1$,*

$$\binom{2n}{n} \geq \frac{2^{2n}}{4\sqrt{n} + 2}.$$

Proof. Consider the random variable

$$X = X_1 \cdots + X_{2n},$$

where the random variables X_i are independent and take values 0 or 1, each with probability 1/2. Note that

$$\mathbb{E}[X] = n$$

and

$$Var[X] = \sum_i Var[X_i] = \frac{n}{2}.$$

Also, given a (possibly negative) integer k,

$$\mathbb{P}[X = n + k] = \binom{2n}{n+k} 2^{-2n} \leq \binom{2n}{n} 2^{-2n}. \tag{15}$$

The inequality above uses the face that $\binom{2n}{n}$ is the largest binomial coefficient.

By Chebyshev's Inequality,

$$\mathbb{P}[|X - n| \geq \sqrt{n}] \leq \frac{1}{2}.$$

Therefore we can write down the complimentary inequality and apply (15), as follows:

$$\frac{1}{2} \leq \mathbb{P}[|X - n| \leq \sqrt{n}] = \sum_{k:|k| \leq \sqrt{n}} \mathbb{P}[X = n + k] \leq (2\sqrt{n} + 1) \binom{2n}{n} 2^{-2n}.$$

A rearrangement gives

$$\binom{2n}{n} \geq \frac{2^{2n}}{4\sqrt{n} + 2}.$$

□

This bound is quite close to giving the actual value of the middle binomial coefficient. The only inaccuracy is in the constant in front of the $\sqrt{n}$. A simple inductive argument gives the upper bound

$$\binom{2n}{n} < \frac{2^{2n}}{\sqrt{2n + 1}}.$$

The actual asymptotic behaviour is

$$\binom{2n}{n} \sim \frac{2^{2n}}{\sqrt{\pi n}}.$$

7.3 Does a random graph contain a triangle?

Consider the random graph $G(n, p)$. This is a graph with n vertices, and for each $u, v \in V(G)$ with $u \neq v$, $uv \in E(G)$ with probability p.

Do we expect $G(n, p)$ to contain a triangle? Of course, this question depends on the value of p. If p is close to 1 then we expect that $G(n, p)$ is a rather dense graph and should contain many triangles. But as p gets close to zero, the number of triangles should shrink drastically.

It turns out that there is a fairly sharp threshold for p and this question. Below this threshold, $G(n, p)$ almost certainly contains no triangles, and above the threshold it almost certainly contains at least one triangle.

Let us first deal with the easier case of what happens below the threshold. We will need another basic tool for estimating the probability that a random variable differs significantly from its expectation - Markov's Inequality.

Theorem 7.3 (Markov's Inequality)**.** *Let X be a non-negative random variable. Then for any real number $\lambda > 0$*

$$\mathbb{P}[X \geq \lambda] \leq \frac{\mathbb{E}[X]}{\lambda}.$$

Proof. First note that

$$X \geq \lambda Y, \tag{16}$$

where Y is the indicator random variable, taking value 1 if $X \geq \lambda$ and taking value 0 otherwise. Take expectations of both sides of inequality (16), giving

$$\mathbb{E}[X] \geq \lambda \mathbb{E}[Y] = \lambda \mathbb{P}[X \geq \lambda],$$

as required. $\square$

Write $G = G(n,p)$. Let X be the random variable counting the number of triangles in G. Using our now standard techniques we can quickly calculate that

$$\mathbb{E}[X] = \binom{n}{3} p^3.$$

Now let $p = \frac{1}{n \cdot f(n)}$ where f is any function which goes to infinity (think of f as a function which goes to infinity very slowly, say $f(n) = \log\log\log\log n$). Then

$$\mathbb{E}[X] \leq n^3 p^3 = \frac{1}{f(n)^3}$$

and

$$\lim_{n\to\infty} \mathbb{E}[X] = 0.$$

Apply Markov's Inequality with $\lambda = 1$, yielding

$$\mathbb{P}[X \geq 1] \leq E[X] \to 0.$$

Since X is integer valued, it follows that

$$\lim_{n\to\infty} \mathbb{P}[X = 0] = 1.$$

We summarise what we have just argued in the following statement:

Theorem 7.4. *Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that*

$$\lim_{n\to\infty} f(n) = \infty.$$

Let $p = \frac{1}{nf(n)}$ and let E be the event that $G(n,p)$ contains a triangle. Then

$$\lim_{n\to\infty} \mathbb{P}[E] = 0.$$

So, when p passes below the *threshold* $\frac{1}{n}$, the random graph $G(n,p)$ almost certainly does not have any triangles.

In fact, when p is slightly larger than this threshold, $G(n,p)$ will almost certainly contain a triangle. We will prove this next, using Chebyshev's Inequality. Actually, the application of Chebyshev's Inequality will be contained in a small technical lemma which lays the foundation for the forthcoming Theorem 7.6 and possible generalisations.

Lemma 7.5. *Consider a sequence $X_1, X_2, \dots$ of non-negative random variables with positive expectation such that*

$$\lim_{n\to\infty} \frac{Var[X_n]}{\mathbb{E}[X_n]^2} = 0. \tag{17}$$

Then

$$\lim_{n\to\infty} \mathbb{P}[X_n > 0] = 1.$$

Proof. Apply Chebyshev's Inequality with

$$\lambda(n) = \frac{\mathbb{E}[X_n]}{\sqrt{Var[X_n]}}.$$

Note that our assumption that $\mathbb{E}[X_n]$ is strictly positive means that $\lambda > 0$ as is required in Chebyshev's Inequality. It follows that

$$\mathbb{P}[|X_n - \mathbb{E}[X_n]| \geq \mathbb{E}[X_n]] \leq \frac{1}{\lambda^2} = \frac{Var[X_n]}{\mathbb{E}[X_n]^2}. \tag{18}$$

Note that the event that $X_n \leq 0$ is a subset of the event that $|X_n - \mathbb{E}[X_n]| \geq \mathbb{E}[X_n]$. So (18) implies that

$$\mathbb{P}[X_n \leq 0] \leq \frac{Var[X_n]}{\mathbb{E}[X_n]^2}.$$

Taking limits of both sides and applying the hypothesis (17), we conclude that

$$\lim_{n\to\infty} \mathbb{P}[X_n < 0] = 0.$$

as required. □

Theorem 7.6. *Let $f(n)$ be any function such that $\lim_{n\to\infty} f(n) = \infty$. Let $p = \frac{f(n)}{n}$ and let E be the event that $G(n,p)$ contains a triangle. Then*

$$\lim_{n\to\infty} \mathbb{P}[E] = 1.$$

Proof. As in the earlier part of this section, let X be the random variable counting the number of triangles in $G(n,p)$. We have already calculated that

$$\mathbb{E}[X] = \binom{n}{3} p^3.$$

In order to make use of Chebyshev's Inequality, we need to calculate the variance of X.

This calculation is easier if we break X down into smaller components. Identify the vertex set of $G(n,p)$ with the interval $[n]$. Write

$$X = \sum_{T\subset[n]:|T|=3} X_T,$$

where X_T is the indicator random variable for the event that the three elements of T form a triangle in $G(n,p)$. Recall that we established a formula (14) for calculating the variance in such an instance:

$$Var[X] = \sum_{T\subset[n]:|T|=3} Var[X_T] + \sum_{T\neq T'} 2\cdot Cov[X_T, X_{T'}]. \tag{19}$$

In the latter sum, we sum over all unordered pairs of distinct subsets of size 3.

We begin with the first sum in (19). For each T, we have the (rather careless) bound

$$Var(X_T) \leq \mathbb{E}[X_T^2] = p^3. \tag{20}$$

For the covariance terms, note that $Cov[X_T, X_{T'}] = 0$ unless the triples T and T' contain two elements in common. This is because the events X_T and $X_{T'}$ are independent of each other unless there is such an intersection.

So, fix two such triples $T = \{t_1, t_2, t_3\}$ and $T' = \{t_1, t_2, t_4\}$. Then

$$Cov[X_T, X_{T'}] = \mathbb{E}[X_T X_{T'}] - \mathbb{E}[X_T]\mathbb{E}[X_{T'}] \leq \mathbb{E}[X_T X_{T'}] = p^5.$$

There are less than $n\binom{n}{3}$ such pairs of triples T, T', and so

$$\sum_{T \neq T'} 2 \cdot Cov[X_T, X_{T'}] \leq 2\binom{n}{3} \cdot n \cdot p^5 \leq n^4 p^5. \tag{21}$$

Combining (21) with (20), we have

$$Var(X) \leq n^3 p^3 + n^4 p^5.$$

We can combine this with our knowledge of $\mathbb{E}[X]$, and look to apply Lemma 7.5. We have

$$\frac{Var[T]}{(\mathbb{E}[X])^2} \leq \frac{n^3 p^3 + n^4 p^5}{(\binom{n}{3} p^3)^2} \leq C\left(\frac{1}{n^3 p^3} + \frac{1}{n^2 p}\right) = C\left(\frac{1}{f(n)^3} + \frac{1}{n f(n)}\right).$$

Therefore,

$$\lim_{n \to \infty} \frac{Var[T]}{(\mathbb{E}[X])^2} = 0.$$

Finally, it follows from Lemma 7.5 that $\lim_{n\to\infty} \mathbb{P}[X > 0] = 1$. Since the random variable X takes only integer values, it follows that $\lim_{n\to\infty} \mathbb{P}[X \geq 1] = 1$, and so the probability that $G(n,p)$ contains a triangle tends to 1.

□

7.4 Exponential Moments - Chernoff Bounds

Another powerful tool for estimating the probability that a random variable differs significantly from its expectation are Chernoff bounds. We can think of such inequalities as a significant quantitative strengthening of Chebyshev's inequality, which is applicable in a more restricted setting. The power of Chernoff type bounds is that they give bounds which shrink *exponentially* as we get further away from the expectation. This can be particularly useful if we are dealing with many random variables simultaneously.

In this subsection, we consider one special case of a Chernoff bound, and an application to *combinatorial discrepancy.*

Theorem 7.7. *Let $X_1, \ldots, X_n$ be independent random variables, each taking value $+1$ and -1 with probability $1/2$. Let $X = X_1 + \cdots + X_n$ and denote $\sigma^2 = Var[X] = n$. Then, for any real $t \geq 0$,*

$$\mathbb{P}[|X| \geq t] < 2e^{\frac{-t^2}{4\sigma^2}}. \tag{22}$$

The statement above can be generalised somewhat, to other random variables of the form $X = X_1 + \cdots + X_n$. However, an important and unavoidable restriction is that the X_i are pairwise independent.

Proof. We will prove that

$$\mathbb{P}[X \geq t] < e^{\frac{-t^2}{4n}}. \tag{23}$$

A similar argument gives the same bound for $\mathbb{P}[X \leq -t]$, and then (22) follows.

We may assume that

$$t \leq n, \tag{24}$$

since the inequality (23) is trivial otherwise.

Consider the auxillary random variable $Y = e^{uX}$, where $0 < u \leq 1$ is a real number which depends on t and will be specified later. The condition (24) will be used later to ensure that $0 < u \leq 1$ is satisfied. By Markov's Inequality

$$\mathbb{P}[X \geq t] = \mathbb{P}[Y \geq e^{ut}] \leq \frac{\mathbb{E}[Y]}{e^{ut}}. \tag{25}$$

The rest of the proof is concerned with calculating $\mathbb{E}[Y]$. Note first that

$$\mathbb{E}[Y] = \mathbb{E}\left[e^{u\sum_i X_i}\right] = \mathbb{E}\left[\prod_i e^{uX_i}\right].$$

Since the random variables X_i are pairwise independent, so are the random variables e^{uX_i}. We can therefore expand the product on the right hand side of the above identity to obtain

$$\mathbb{E}[Y] = \prod_i \mathbb{E}[e^{uX_i}] \tag{26}$$

We can use the Taylor expansion for e^{uX_i}, and the fact that $|uX_i| \leq 1$, to deduce that

$$e^{uX_i} < 1 + uX_i + u^2X_i^2. \tag{27}$$

Indeed,

$$\begin{aligned}
e^{uX_i} &= 1 + uX_i + \frac{(uX_i)^2}{2!} + \frac{(uX_i)^3}{3!} + \frac{(uX_i)^4}{4!} \cdots \\
&\leq 1 + uX_i + \frac{(uX_i)^2}{2!} + \frac{|uX_i|^3}{3!} + \frac{|uX_i|^4}{4!} \cdots \\
&\leq 1 + uX_i + (uX_i)^2\left(\frac{1}{2!} + \frac{1}{3!} + \ldots\right) \\
&< 1 + uX_i + (uX_i)^2\left(\frac{1}{2} + \frac{1}{2^2} + \ldots\right) \\
&= 1 + uX_i + (uX_i)^2.
\end{aligned}$$

(In the last inequality above we have been rather careless, and we could have used the fact that

$$\sum_{n=1}^{\infty} \frac{1}{n!} = e - 1$$

to give a more precise estimate.) Taking expectations of both sides of (27) then gives

$$\mathbb{E}[e^{uX_i}] < 1 + u\mathbb{E}[X_i] + u^2\mathbb{E}[X_i^2] = 1 + u^2 \leq e^{u^2}.$$

Inserting this into (26) gives

$$\mathbb{E}[Y] < e^{nu^2},$$

and then combining this with (25) yields

$$\mathbb{P}[X \geq t] < e^{nu^2 - ut}.$$

This quantity is minimised by setting $u = t/2n$ (and so by (24) we have $u \leq 1$), and the conclusion is that

$$\mathbb{P}[X \geq t] < e^{-\frac{t^2}{4n}},$$

as claimed in (23). □

7.4.1 Combinatorial discrepancy

Let X be an arbitrary set of cardinality n. We may think of $X = \{1, \ldots, n\}$. Let $\mathcal{F}$ be a family of subsets of X. Colour the elements of X by two colours (blue and red), with the aim of making every set in $\mathcal{F}$ have approximately the same number of blue and red elements. To what extent is this possible? The notion of *combinatorial discrepancy* quantifies this idea.

We can identify the two colours with $+1$ and -1 respectively, so the colouring is a function $\chi : X \to \{-1, 1\}$. Then the quantity

$$\chi(S) := \sum_{i \in S} \chi(i)$$

measures how balanced a set S is under this colouring. For a fixed colouring χ, define

$$disc(\mathcal{F}, \chi) := \max_{S \in \mathcal{F}} |\chi(S)|.$$

If this is small then the χ does a good job of balancing the sets in $\mathcal{F}$. Finally, the *discrepancy* of the family $\mathcal{F}$ is

$$disc(\mathcal{F}) := \min_{\chi} disc(\mathcal{F}, \chi).$$

Assume that n is even for simplicity. Consider the case when $\mathcal{F} = 2^X$, the set of all subsets of X. Then any colouring must have at least $n/2$ elements of X taking the same colour. Say we have at least $n/2$ blue elements, and let $B \subset X$ be the set of all blue elements under the colouring χ. Then

$$disc(\mathcal{F}, \chi) \geq |\chi(B)| \geq n/2.$$

So, $disc(\mathcal{F}) \geq n/2$. On the other hand, any colouring which has $n/2$ red and $n/2$ blue elements gives $disc(\mathcal{F}, \chi) = n/2$, which implies that $disc(\mathcal{F}) \leq n/2$. Combining the two inequalities, it follows that $disc(\mathcal{F}) = n/2$.

The next result shows that we can do much better for smaller families $\mathcal{F}$, getting discrepancy close to $\sqrt{n}$.

Theorem 7.8. *Let $\mathcal{F}$ be a family of subsets of $X = \{1, \dots, n\}$ with $|\mathcal{F}| = m$. Then*

$$disc(\mathcal{F}) \leq \sqrt{4n \ln(2m)}.$$

Proof. First let's clarify what we need to prove. We are given an arbitrary family $\mathcal{F}$ of m subsets of X, and we want to show that there exists a 2-colouring $\chi : X \to \{-1, +1\}$ such that

$$|\chi(S)| \leq \sqrt{4n \ln(2m)}, \quad \forall S \in \mathcal{F}.$$

Let χ be a random 2-colouring of X. That is, for each $i \in X$, $\chi(i)$ takes value $+1$ or -1, each with probability $1/2$.

Fix a set $S \in \mathcal{F}$ and consider the random variable

$$\chi(S) := \sum_{i \in S} \chi(i)$$

This is exactly the same as the random variable X in the statement of Theorem 7.8, with role of n there played by $|S|$ here. Hence

$$\mathbb{P}[|\chi(S)| \geq t] < 2e^{\frac{-t^2}{4|S|}} \leq 2e^{\frac{-t^2}{4n}}.$$

We would like this to be less than $1/m$, so that we can use the union bound to conclude that there is a positive probability that $|\chi(S)| < t$ for all S. For this reason, we set

$$t := \sqrt{4n \ln(2m)}$$

and so

$$\mathbb{P}[|\chi(S)| \geq t] < 2e^{-\frac{4n \ln(2m)}{4n}} = \frac{1}{m}.$$

Now we can indeed use the union bound to conclude that

$$\mathbb{P}[\exists\, S \in \mathcal{F} \text{ such that } |\chi(S)| \geq \sqrt{4n \ln(2m)}] \leq \sum_{S \in \mathcal{F}} \mathbb{P}[|\chi(S)| \geq \sqrt{4n \ln(2m)}] < 1.$$

It thus follows that, with positive probability, our randomly chosen colouring χ satisfies

$$|\chi(S)| \leq \sqrt{4n \ln(2m)}, \ \forall S \in \mathcal{F},$$

as required.

□

References

[1] N. Alon and J. Spencer 'The Probabilistic Method' *John Wiley & Sons* (1992).

[2] T. Tao, V. Vu 'Additive combinatorics' *Cambridge University Press* (2006).