

# Part II – Graph Theory

Based on lectures by Prof. P. Russell

Notes taken by Bhavik Mehta

Michaelmas 2017

## Contents

<b>0</b>	<b>Introduction</b>	<b>1</b>
0.1	Preliminary . . . . .	1
0.2	Informal definitions . . . . .	2
0.3	Where do such structures arise? . . . . .	2
<b>1</b>	<b>Ramsey Theory</b>	<b>4</b>
1.1	Basic Terminology . . . . .	10
<b>2</b>	<b>Extremal Graph Theory</b>	<b>12</b>
2.1	Forbidden Subgraph Problem . . . . .	12
2.1.1	Triangles . . . . .	12
2.1.2	Complete graphs . . . . .	13
2.1.3	Bipartite graphs . . . . .	14
2.1.4	General graphs . . . . .	16
2.1.5	Proof of Erdős-Stone . . . . .	18
2.2	Hamiltonian graphs . . . . .	19
<b>3</b>	<b>Graph Colouring</b>	<b>21</b>
3.1	Planar Graphs . . . . .	21
3.2	General Graphs . . . . .	23
3.2.1	Lower bounds . . . . .	23

## 0 Introduction

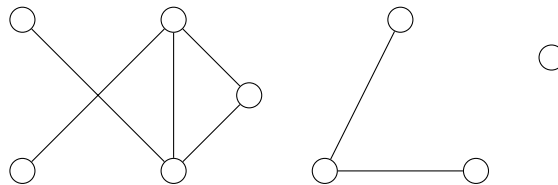
### 0.1 Preliminary

This course has no prerequisites, being the first course in the tripos in this field. However, introductory facts about, say, eigenvalues and limits will be used. As usual, books are not required, but the most relevant is Modern Graph Theory, B. Bollobas.

This course consists of six chapters: corresponding to paragraphs 2 to 7 in the schedules but in a different order. Paragraph 1 is introductory, and hence is split between chapters.

## 0.2 Informal definitions

A *graph* consists of ‘vertices’ with some pairs of vertices joined by ‘edges’.



In this course, we make the following assumptions:

- The number of vertices is finite.
- No ‘multiple edges’: every pair of vertices is connected by at most one edge.
- No loops: no vertex can be joined to itself - edges have to go between different vertices.

## 0.3 Where do such structures arise?

**Example.**

1. Problem: Can we walk across each bridge in this city precisely once, returning to our starting point? (Euler, 1736)

We turn the problem into a graph problem - in particular, a multigraph to allow multiple edges between vertices. The vertices of this graph are the four bits of the city, while the edges are the seven bridges. In terms of this graph, the problem becomes to walk around the edges of this graph precisely once, and return to the starting vertex.

2. Map colouring problem (1850s)

How many colours are required to colour a map such that neighbouring countries get different colours? (picture) Here, the vertices are countries, and the edges are neighbouring pairs. If we can draw a graph with no crossing edges, how many colours are needed to colour vertices with all edges having two different colours?

3. Cosets of finite subgroups of finite groups

Let  $G$  be a finite group,  $H \leq G$ ,  $|G : H| = n$ . By Lagrange,  $\exists a_1, \dots, a_n \in G$  such that the left cosets of  $H$  are  $a_1H, \dots, a_nH$ . Similarly,  $\exists b_1, \dots, b_n \in G$  such that the right cosets of  $H$  are  $Ha_1, \dots, Ha_n$ . Can we do this simultaneously? In particular, are there  $a_1, \dots, a_n \in G$  such that  $a_1H, \dots, a_nH \in G$  are the left cosets, and  $Ha_1, \dots, Ha_n \in G$  are the right cosets?

In the case where  $H \triangleleft G$  this is easy since we have  $aH = Ha \forall a \in G$ . However it is less obvious if  $H \not\triangleleft G$ . Let  $L$  be the set of left cosets, and let  $R$  be the set of right cosets. We can create a graph where the vertices are  $L \cup R$ . Join  $X \in L$  to  $Y \in R$  by an edge if we can find some  $a$  such that  $X = aH$  and  $Y = Ha$ , that is, if  $X$  and  $Y$  have a common element. Ideally, we would like a graph like this (insert bipartite graph here)

Formally, we ask: In this graph, can we find a set of edges meeting each vertex precisely once?

#### 4. Fermat equation mod $p$

Consider the equation  $x^n + y^n = z^n$ . Does this have solutions mod  $p$  for  $p$  a prime? Rule out trivial solutions such as  $x = 0, y = z$ , in particular we search for solutions in  $\mathbb{Z}_p$  where  $xyz \neq 0$ . Let  $G = \mathbb{Z} \setminus \{0\}$  under multiplication, and  $H = \{g^n : g \in G\}$ . We can check that  $H \leq G$ , and that  $|G : H| \leq n$ , by considering the number of  $n$ th roots an element can have. So,  $G$  can be partitioned into  $g_1H, g_2H, \dots, g_mH$  for some  $g_1, g_2, \dots, g_m$  and  $m \leq n$ . Suppose in some  $g_iH$ , we have  $a, b, c$  with  $a + b = c$ . Then  $a = g_ix^n, b = g_iy^n, c = g_iz^n$  for some  $x, y, z \in G$ . Then

$$\begin{aligned} g_ix^n + g_iy^n &= g_iz^n \\ x^n + y^n &= z^n \end{aligned}$$

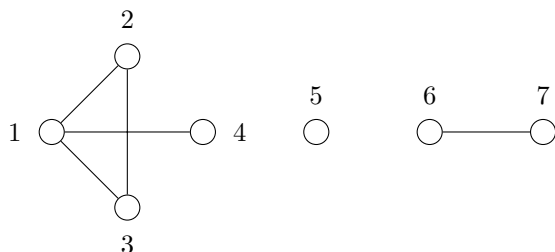
It finally remains to show Schur's Theorem:

**Theorem** (Schur's Theorem). Let  $n$  be a positive integer. Then if  $p$  is a sufficiently large positive integer, whenever  $\{1, 2, \dots, p\}$  is partitioned into  $n$  parts, we can solve  $a + b = c$  with  $a, b, c$  all in some part.

# 1 Ramsey Theory

**Definition** (Graph). A **graph** is an ordered pair  $(V, E) = G$  where  $V$  is a finite set and  $E$  is a set of unordered pairs of distinct elements of  $V$ . We call elements of  $V$  **vertices** of  $G$  and elements of  $E$  **edges**. We often write  $v \in G$  to mean  $v \in V$  and sometimes, where clear,  $e \in G$  to mean  $e \in E$ . Often denote  $\{u, v\} \in E$  by  $uv$ . Note  $uv = vu$ .

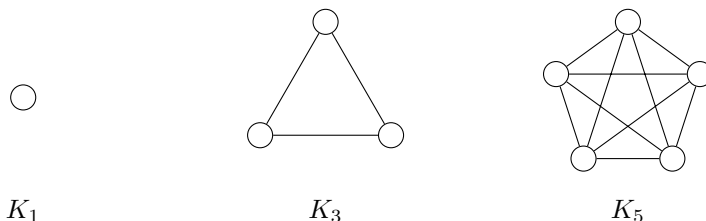
**Example.** Here's an example of a **graph**  $G = (\{1, 2, 3, 4, 5, 6, 7\}, \{12, 23, 13, 14, 67\})$ , but it's often easier to represent a graph by a drawing: take a point for each **vertex**, join two vertices if they are in an **edge**.



**Definition** (Isomorphism). Let  $G = (V, E)$  and  $G' = (V', E')$  be **graphs**. An **isomorphism** from  $G$  to  $G'$  is a bijection  $\phi : V \rightarrow V'$  such that for all  $u, v \in V$ , we have  $\phi(u)\phi(v) \in E'$  if and only if  $uv \in E$ . If such an isomorphism exists, we say  $G$  is **isomorphic** to  $G'$ .

**Definition** (Subgraph). Suppose also  $H = (W, F)$  is a **graph**. We say  $H$  is a **subgraph** of  $G$  and write  $H \subset G$  if  $W \subset V$  and  $F \subset E$ . Often, we say ' $H$  is a subgraph of  $G$ ' to mean ' $H$  is **isomorphic** to a subgraph of  $G$ '.

**Definition** (Complete graph of order  $n$ ). The **complete graph of order  $n$** ,  $K_n$  has  $n$  vertices with every pair forming an edge.



Looking at the diagram, we see  $K_3$  looks like a triangle, and we will often just refer to  $K_3$  as a **triangle**. In addition, note that if  $m \leq n$ , then  $K_m \subset K_n$ .

Recall **Schur's Theorem**:

**Theorem** (Schur's Theorem reformulated). Let  $k$  be a positive integer. Then there is a positive integer  $n$  such that if the set  $[n] = \{1, 2, \dots, n\}$  is coloured with  $k$  colours, we can find  $a, b, c$  with  $a + b = c$  and  $a, b, c$  the same colour.

Can we prove this directly? First try  $k = 1$ , where  $n = 2$  immediately works because  $1 + 1 = 2$ .

For  $k = 2$ , use colours **red** and **green**, wlog we can take **1 red**. As  $1 + 1 = 2$ , we must have a **green 2**. Similarly, **4** must be **red** otherwise  $2 + 2 = 4$ . Next, we have  $1 + 3 = 4$ , so **3** must be **green**. Finally, notice that  $1 + 4 = 2 + 3$  so  $n = 5$  suffices.

This case analysis worked for  $k = 2$ , but is likely to get a lot more fiddly for larger  $k$ . Even for  $k = 3$  this case analysis turns out to be a lot harder, so can we come up with a ‘better’ proof for the  $k = 2$  case, that is ‘more likely to generalise’?

*Proof.* ( $k = 2$  of Schur’s Theorem, improved) Suppose  $[5] = \{1, 2, 3, 4, 5\}$  are coloured **red** and **green**. Then some three of these are the same colour, and without loss of generality  $i < j < k$  are **red**. If  $j - i$  is **red** we are done, since  $i + (j - i) = j$ . Similarly if  $k - i$  or  $k - j$  is **red**, we are done. If not, all of  $j - i$ ,  $k - i$ ,  $k - j$  are **green**, but then  $k - i = (j - i) + (k - j)$ , and we are done.  $\square$

Let’s try a similar approach for  $k = 3$ , and I claim  $n = 16$  works.

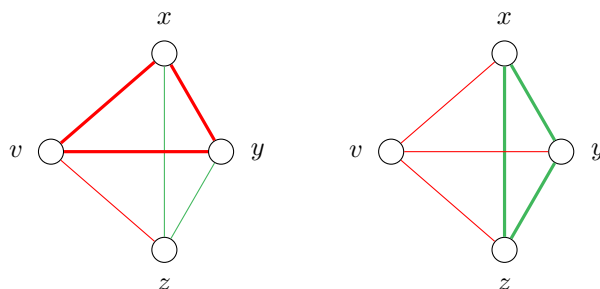
*Proof.* (Schur’s theorem,  $k = 3$ ) Suppose  $[16]$  are coloured **red**, **green** and **blue**. By the pigeonhole principle, some six numbers are the same colour, without loss of generality  $x_1 < x_2 < \dots < x_6$  are **red**. If  $x_j - x_i$  is **red** for any  $i < j$  then we are done:  $x_i + (x_j - x_i) = x_j$ . So assume all  $x_j - x_i$  are **blue** or **green**. Consider the five numbers  $x_2 - x_1, x_3 - x_1, x_4 - x_1, x_5 - x_1, x_6 - x_1$ . By the pigeonhole principle, some three of these are the same colour: say  $x_i - x_1, x_j - x_1, x_k - x_1$  are **green**, for  $i < j < k$ .

If  $x_j - x_i$  is **green**, we are done:  $(x_i - x_1) + (x_j - x_i) = x_j - x_1$ , similarly if  $x_k - x_i$  or  $x_k - x_j$  is **green**. Otherwise, all of  $x_j - x_i, x_k - x_i, x_k - x_j$  are **blue**, and we have  $(x_j - x_i) + (x_k - x_j) = x_k - x_i$ , so we are done.  $\square$

This last paragraph looks very similar to the proof of the  $k = 2$  case, so can we use induction? It’s unclear, because it wasn’t exactly the previous case. But notice how we always seem to be interested in differences of two numbers, yet we never seem to use facts like  $7 - 3 = 9 - 5$ . So, consider a graph where the vertices are numbers, and the edges refer to the difference between two vertices, and we colour edges.

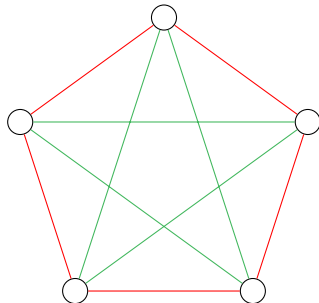
Suppose the edges of a **complete graph**  $K_6$  are coloured blue and yellow. Then we can find a monochromatic **triangle**.

*Proof.* Pick  $v \in K_6$ .  $v$  is in five edges, so some three are the same colour, without loss of generality call them  $vx, vy, vz$  and say they are **red**. If any of  $xy, xz, yz$  is **red**, we have a **red triangle** with  $v$ . If not,  $xyz$  is a **green triangle**.



$\square$

**Remark.** Schur for  $k = 2$  now follows: Suppose  $[5]$  are coloured **red**/**green**. Take a  $K_6$  with vertices  $\{0, 1, \dots, 5\}$ . Colour edge  $ij$ , ( $i < j$ ) with the colour of  $j - i$ . By the above proof, we have a monochromatic triangle  $abc$  ( $a < b < c$ ). Then  $(c - b) + (b - a) = c - a$ , and  $c - b, b - a, c - a$  are all the same colour. We cannot necessarily find a monochromatic triangle in  $K_5$ , so 6 vertices are necessary.



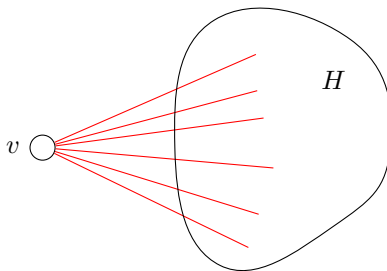
In this graph form, the statement for more colours is now amenable to an induction proof.

**Proposition 1.** Let  $k$  be a positive integer. Then there is a positive integer  $n$  such that whenever the edges of  $K_n$  are coloured with  $k$  colours we can find a monochromatic triangle.

*Proof.* Induction on  $k$ . For  $k = 1$ ,  $n = 3$  works, so consider  $k > 1$ .

By the induction hypothesis, there exists  $m$  such that if  $K_m$  is  $(k - 1)$ -coloured, then there is a monochromatic triangle. Let  $n = k(m - 1) + 2$ .

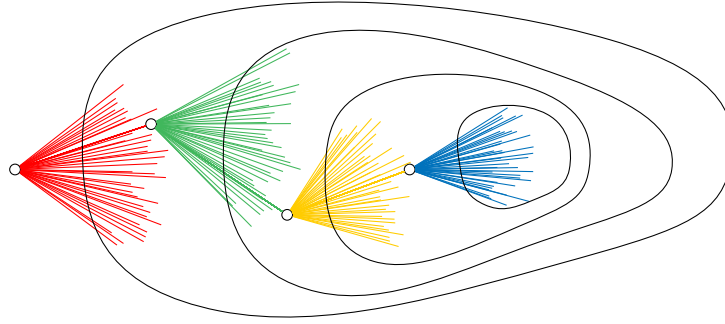
Now  $k$ -colour the edges of  $K_n$ . Pick vertex  $v$ . The number of edges containing  $v$  is  $n - 1 = k(m - 1) + 1$ . So some  $m$  of them are the same colour, without loss of generality blue. Let  $H$  be a  $K_m$  joined to  $v$  by blue edges. If  $H$  contains a blue edge, it makes a blue triangle with  $v$ . If not,  $H$  is a  $(k - 1)$ -coloured  $K_m$  so by definition of  $m$ , it contains a monochromatic triangle.



□

**Remark.**

1. Can think in terms of the following picture (WIP). Eventually, we will run out of colours and get a monochromatic triangle.



2. How big do we need to take  $n$  to be? Let  $f(k)$  be the smallest  $n$  that works with  $k$  colours. Clearly  $f(1) = 3$ . Our proof says for  $k > 1$ ,  $f(k) \leq k(f(k-1) - 1) + 2 \leq kf(k-1)$ . So by induction,  $f(k) \leq 3k!$ .
3. [Schur's Theorem reformulated](#) follows immediately in the same way as for the  $k = 2$  case.

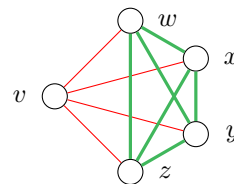
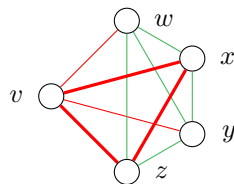
Can we find larger monochromatic subgraphs, for instance a monochromatic  $K_4$ ? Think about two colours at first for simplicity. Yes, but we need an intermediate step first.

**Example.**

1. Let  $K_{10}$  have edges coloured red/green. Then there is a red triangle or a green  $K_4$ .

*Proof.* Take a vertex  $v$ . It's in 9 edges so either  $v$  is in 4 red edges, or in 6 green edges.

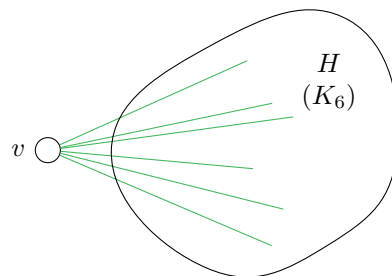
- $v$  is in 4 red edges



If any edge between  $w, x, y, z$  is red, we have a red triangle.

If not, they are all green and we have a green  $K_4$ .

- $v$  is in 6 green edges



We know  $H$  contains a monochromatic triangle. If it is **red**, we are done. If it is **green**, then it makes a **green**  $K_4$  with  $v$  as required.  $\square$

2. Let  $K_{20}$  be coloured **red**/**green**. Then  $\exists$  a monochromatic  $K_4$ .

*Proof.* Pick a vertex  $v$ . It's in 19 edges, so some 10 of them are the same colour, without loss of generality **red**.

Let  $H$  be a  $K_{10}$  joined to  $v$  by **red** edges.

Either  $H$  has a **red triangle**, making a **red**  $K_4$  with  $v$ , or  $H$  contains a **green**  $K_4$ .  $\square$

**Definition** (Ramsey number). Let  $s, t \geq 2$ . The Ramsey number  $R(s, t)$  is the least  $n$  such that whenever  $K_n$  has edges coloured **red**/**green** there must be a **red**  $K_s$  or a **green**  $K_t$  (if such an  $n$  exists). We also write  $R(s) = R(s, s)$ .

We have seen  $R(3) = 6$ ;  $R(3, 4) \leq 10$ ;  $R(4) \leq 20$ .

**Theorem 2** (Ramsey's Theorem).  $R(s, t)$  exists for all  $s, t \geq 2$ . Moreover, if  $s, t \geq 2$  then  $R(s, t) \leq R(s-1, t) + R(s, t-1)$ .

*Proof.* Induction on  $s + t$ .

For  $s = 2$ , we have  $R(2, t) = t$ : If all edges of a  $K_t$  are **green**, then we have a **green**  $K_t$ , otherwise there is a **red edge**, which is exactly a **red**  $K_2$ . Similarly  $R(s, 2) = s$ .

In the case  $s, t > 2$ , let  $a = R(s-1, t)$  and  $b = R(s, t-1)$  which exist by the induction hypothesis. Let  $n = a + b$ . Suppose  $K_n$  has edges coloured **red**/**green**. Pick  $v \in K_n$ , which is in  $a + b - 1$  edges, so either  $v$  is in **a red edges**, or it is in **b green edges**.

- If  $v$  is in **a red edges**, let  $H$  be the  $K_a$  joined to  $v$  by **red** edges. Now  $a = R(s-1, t)$ , so either  $H$  has a **red**  $K_{s-1}$ , making a **red**  $K_s$  with  $v$ , or  $H$  has a **green**  $K_t$ , so done.
- If  $v$  is in **b green edges**, and we can make the same argument with colours reversed.

$\square$

**Corollary 3.** For all  $s, t \geq 2$ ,  $R(s, t) \leq 2^{s+t}$ , so  $R(s) \leq 4^s$ .

*Proof.* Induction on  $s + t$ . Base cases  $s = 2$ ,  $R(2, t) = t \leq 2^{2+t}$  and similarly for  $t = 2$ ,  $R(s, 2) = s \leq 2^{s+2}$ . For  $s, t \geq 2$

$$\begin{aligned} R(s, t) &\leq R(s-1, t) + R(s, t-1) \\ &\leq 2^{s-1+t} + 2^{s+1-t} \\ &= 2^{s+t}. \end{aligned}$$

$\square$

**Remark.**

- A more careful version of the same proof gives  $R(s) = \mathcal{O}\left(\frac{4^s}{\sqrt{s}}\right)$  (Erdős-Szekeres 1935)
- A harder proof gives an improvement to  $R(s) = \mathcal{O}\left(\frac{4^s}{s}\right)$  (Thomason 1988)
- An even harder proof gives  $R(s) = \mathcal{O}\left(\frac{4^s}{s^k}\right)$  for any  $k$  (Conlon 2009)
- Is  $R(s) = \mathcal{O}((3.999999999)^s)$ ? Unknown.



**Theorem 4** (Multicolour Ramsey Theorem). Let  $k \geq 1$  and  $s \geq 2$ . Then there exists some  $n$  such that whenever the edges of  $K_n$  are coloured with  $k$  colours, we can find a monochromatic  $K_s$ .

*Proof.* Induction on  $k$ . For the base case  $k = 1$ , we can take  $n = s$ .

For  $k > 1$ , by the induction hypothesis we can find  $m$  such that if  $K_m$  is  $(k-1)$ -coloured, then there is a monochromatic  $K_s$ . Let  $n = R(s, m)$  and colour  $K_n$  with  $k$  colours, including **red** but not green. Re-colour by turning all non-red edges **green**. By definition of  $n$ , we have either

- A **red**  $K_s$ , so done
- A **green**  $K_m$ . Then in the original colouring, this  $K_m$  was  $(k-1)$ -coloured, so by definition of  $m$ , it contains a monochromatic  $K_s$ .

□

**Remark.**

1. We write  $R_k(s)$  for the smallest  $n$  that works. In particular  $R_1(s) = s$  and  $R_2(s) = R(s)$ . Proof of theorem 4 will give  $R_k(s) \leq 4^{s^{k-1}}$ .
2. What do we really mean by colouring? We have informally talked about ‘red edges’ etc. More formally, we colouring of the edges of  $G = (V, E)$  with  $k$  colours is a function  $c : E \rightarrow [k]$ . We mostly refer to the colours as ‘red’, ‘green’ and so on. We say  $H = (W, F)$  is monochromatic if  $c|_F$  is constant.

With any number of colours, we can find arbitrarily large monochromatic subgraphs. What about infinite ones?

**Definition** (Infinite graph). An **infinite graph** is an ordered pair  $G = (V, E)$  where  $V$  is an infinite set and  $E$  is a set of unordered pairs of elements of  $V$ . Note, in our terminology, an infinite graph is not a graph.

**Definition** ((Possibly infinite) graph). A **(possibly infinite) graph** is a graph or an infinite graph.

**Definition** (Infinite complete graph).  $K_\infty$  is the infinite graph with a countably infinite vertex set and every pair of vertices forming an edge.

Where it makes sense to do so, we carry terminology over from graphs to infinite graphs. Suppose edges of  $K_\infty$  are coloured blue/yellow. What can we find? By [Ramsey’s Theorem](#), we can find a monochromatic  $K_s$  for every finite  $s = 1, 2, 3, \dots$ . Note this doesn’t immediately give a monochromatic  $K_\infty$ :

In the diagram, we see that for any finite  $s$  we have a blue  $K_s$  but no blue  $K_\infty$ . (However in this example, there is a yellow  $K_\infty$ )

**Theorem 5** (Infinite Ramsey Theorem). Let  $k \geq 1$ . Whenever the edges of  $K_\infty$  are  $k$ -coloured, we have a monochromatic  $K_\infty$  subgraph.

*Proof.* Take  $v_1 \in K_\infty$ . The vertex  $v_1$  is in infinitely many edges, so infinitely many edges from  $v_1$  are the same colour. Let  $A_1$  be an infinite set of vertices of  $K_\infty$  such that for all  $u \in A_1$ ,  $v_1u$  has colour  $c_1$ . Now pick  $v_2 \in A_1$ . Similarly, we can find an infinite  $A_2 \subset A_1$  such that all edges  $v_2u$  for  $(u \in A_2)$  have colour  $c_2$ . Keep going. We get infinite sequences  $v_1, v_2, v_3, \dots$  of vertices,  $c_1, c_2, c_3, \dots$  of colours and  $A_1 \supset A_2 \supset A_3 \supset \dots$  such that

- for  $i \geq 2$ ,  $v_i \in A_{i-1}$
- for  $i \geq 1$ , for all  $u \in A_i$ ,  $v_i u$  is an edge of colour  $c_i$ .

In particular, if  $i < j$  then  $v_i v_j$  has colour  $c_i$ . Now, infinitely many of the  $c_i$  are the same. Say  $i_1 < i_2 < i_3 < \dots$  such that  $c_{i_1} = c_{i_2} = c_{i_3} = \dots$ . Consider  $v_{i_1}, v_{i_2}, v_{i_3}, \dots$ . Any edge between two of these vertices has colour  $c_{i_1}$ . So we have a monochromatic  $K_\infty$ .  $\square$

**Remark.** This is often called a ‘two-pass proof’. On the first pass, we found a sequence of colours, and on the second pass we found infinitely many colours the same.

**Corollary 6.** Any bounded sequence has a convergent subsequence.

*Proof.* Any bounded monotone sequence converges, so it is enough to show that any real sequence  $(x_n)_{n \geq 1}$  has a monotone subsequence. Let  $G$  be a  $K_\infty$  with vertex set  $\{1, 2, 3, \dots\}$ . Colour  $ij$ , with  $i < j$  blue if  $x_i < x_j$  and yellow if not.

By infinite Ramsey, there is a monochromatic subgraph  $H \cong K_\infty$ . Let the vertices of  $H$  be  $n_1 < n_2 < n_3 < \dots$ . If  $H$  is blue, then  $(x_{n_j})_{j \geq 1}$  is decreasing, whereas if  $H$  is yellow then  $(x_{n_j})_{j \geq 1}$  is increasing.  $\square$

## 1.1 Basic Terminology

Let  $G = (V, E)$  be a graph. Sometimes write  $V(G) = V$ , and  $E(G) = E$ . The **order** of  $G$  is  $|G| = |V(G)|$ . We denote the number of edges of  $G$  by  $e(G) = |E(G)|$ , sometimes called the size of  $G$ .

Note that the **complete graph of order  $n$**  has order  $n$ .

**Definition** (Neighbourhood). Let  $v \in G$ . Then **neighbourhood** of  $v$  is the set

$$\Gamma(v) = \{w \in G \mid vw \in E(G)\}$$

If  $w \in \Gamma(v)$ , then  $w$  is a **neighbour** of  $v$ , or  $w$  is **adjacent** to  $v$ , we write  $w \sim v$ .

**Definition** (Degree). The **degree** of  $v$  is  $d(v) = |\Gamma(v)|$ , the number of vertices **adjacent** to  $v$ .

The **maximum degree** of  $G$  is  $\Delta(G) = \max_{v \in G} d(v)$

The **minimum degree** of  $G$  is  $\delta(G) = \min_{v \in G} d(v)$

The **average degree** of  $G$  is  $\frac{1}{|G|} \sum_{v \in G} d(v)$

Clearly  $\delta(G) \leq \text{average degree of } G \leq \Delta(G)$ .

**Definition** (Regular). If every vertex in  $G$  has the same degree, we say  $G$  is **regular**.

If this degree is  $r$ , say  $G$  is  **$r$ -regular**.

**Example.** How do vertex-degrees relate to  $e(G)$ ? We can count ‘ends of edges’ in two different ways. In particular, each edge has two ends, so there are  $2e(G)$  ends. A vertex  $v$  is in  $d(v)$  ends, so there are  $\sum_{v \in G} d(v)$  ends, and hence  $\sum_{v \in G} d(v) = 2e(G)$ .

Hence also the average degree of  $G$  is  $\frac{2e(G)}{|G|}$ .

**Definition** (Path). Let  $G$  be a graph. A **path** in  $G$  is a finite sequence  $v_0, v_1, \dots, v_l$  of distinct vertices of  $G$  with  $v_{i-1} \sim v_i$  for  $1 \leq i \leq l$ .



We say this path has length  $l$  and goes from  $v_0$  to  $v_l$ .

If  $v, w \in G$  we write  $v \rightarrow w$  to mean there is a path from  $v$  to  $w$ .

The relation  $\rightarrow$  is an equivalence relation (Exercise.)

**Definition** (Components). The equivalence classes of  $\rightarrow$  are called the **components** of  $G$ . If  $G$  has only one component, we say  $G$  is **connected**.

**Definition** (Cycle). A **cycle** is a sequence  $v_0, v_1, \dots, v_l$  of vertices of  $G$  with  $v_0, \dots, v_{l-1}$  distinct,  $v_l = v_0$ ,  $v_{i-1} \sim v_i$  for  $1 \leq i \leq l$  and  $l \leq 3$ . We say that the length of the cycle is  $l$ .

**Definition** (Forest). A graph with no cycles is called a **forest**. A **tree** is a connected forest. Each component of a forest is a tree.

**Definition** (Disjoint union). Suppose  $G, H$  are graphs with  $V(G) \cap V(H) = \emptyset$ . The **disjoint union** of  $G, H$  is the graph  $G \cup H$  with  $V(G \cup H) = V(G) \cup V(H)$  and  $E(G \cup H) = E(G) \cup E(H)$ .

We often write  $G \cup H$  even if  $V(G) \cap V(H) \neq \emptyset$ , this means take graphs  $G', H'$  with  $G' \cong G$ ,  $H' \cong H$ ,  $V(G') \cap V(H') = \emptyset$  then take  $G' \cup H'$ .

**Example.**

1.  $K_3 \cup K_5$
2. Any graph is the disjoint union of its components  $(G_1 \cup \dots \cup G_k)$ . For instance, a forest is a **disjoint union** of trees.

**Definition** (Induced subgraph). Let  $G = (V, E)$  be a graph, and let  $W \subset V$ . The **induced subgraph** on  $W$  is the graph  $G[W]$  with  $V(G[W]) = W$  and, for  $x, y \in W$ ,  $xy \in E(G[W]) \iff xy \in E$ .

**Definition** (Complement). Let  $G = (V, E)$  be a graph. The **complement** of  $G$  is the graph  $\overline{G}$  with  $V(\overline{G}) = V$ , and for distinct  $x, y \in V$ ,  $xy \in E(\overline{G}) \iff xy \notin E$ .

**Example.**

- Suppose edges of  $K_n$  are coloured blue/yellow. We can think about the ‘blue subgraph’ with  $n$  vertices and blue edges only. The complement is the yellow subgraph.
- 

**Notation.** Let  $f, g : \mathbb{N} \rightarrow (0, \infty)$ . We say

- $f = \mathcal{O}(g)$  if  $f < Ag$  for some constant  $A$ .
- $f = \Omega(g)$  if  $g = \mathcal{O}(f)$
- $f = \Theta(g)$  if  $f = \mathcal{O}(g)$  and  $f = \Omega(g)$ .
- $f = o(g)$  means  $f/g \rightarrow 0$  (as  $n \rightarrow \infty$ )
- $f = \omega(g)$  means  $f/g \rightarrow \infty$
- $f \sim g$  means  $f/g \rightarrow 1$ .

## 2 Extremal Graph Theory

**Problem** (from (link) section 1): How large must  $n$  be so that if edges of  $K_n$  are coloured blue/yellow then always get a monochromatic  $K_s$ ?

Let  $G$  be a ‘blue subgraph’ of  $K_n$ , consisting of all vertices, and only the blue edges.  $K_n$  has a blue  $K_s \iff K_s \subset G$ , and  $K_n$  has a yellow  $K_s \iff G$  has an induced  $\overline{K_s}$  subgraph.

**rephrase** How large must  $n$  be to force every graph of order  $n$  to have  $K_s$  or  $\overline{K_s}$  as an induced subgraph?

This is a typical example of an **extremal problem**: How large must some parameter of  $G$  be to force  $G$  to have a certain property?

Alternatively, how big can the parameter be with  $G$  not having the property?

### 2.1 Forbidden Subgraph Problem

Fix a graph  $H$  with at least one edge. Let  $n \geq |H|$ . Clearly  $H \subset K_n$  but  $H \not\subset \overline{K_n}$ .

How many edges must a graph  $G$  of order  $n$  have to force  $H \subset G$ ? Alternatively, if  $|G| = n$  and  $H \subset G$ , how large can  $e(G)$  be?

**Definition.** Define

$$\text{ex}(n; H) = \max \{ e(G) \mid |G| = n, H \not\subset G \}$$

**Problem** Can we determine  $\text{ex}(n; H)$ ?

#### 2.1.1 Triangles

Let  $H = K_3$ . The goal is to find  $G$  with  $|G| = n$ ,  $e(G)$  large and  $\Delta \not\subset G$ .

The idea is to partition  $V(G) = X \cup Y$  with all edges going from  $X$  to  $Y$ .

**Definition.** A graph  $G$  is **bipartite** (with bipartition  $(X, Y)$ ) if  $V(G)$  can be partitioned as  $X \cup Y$  in such a way that if  $e \in E(G)$  then  $e = xy$  for some  $x \in X, y \in Y$ .

Bipartite graphs have no  $\Delta$ s (and indeed no cycles of odd length). In fact the converse is also true:

**Theorem 7.** A graph is bipartite iff it contains no odd cycles.

*Proof.* May return to this later, an exercise for now. □

Which bipartite graph is best?

Clearly it is a good idea to include all possible edges from  $X$  to  $Y$ .

**Definition.** Let  $s, t \geq 1$ . The **complete bipartite graph**  $K_{s,t}$  has bipartition  $(X, Y)$  with  $|X| = s, |Y| = t$  and  $xy \in E(K_{s,t}) \forall x \in X, y \in Y$ .

We have  $|K_{s,t}| = s + t$  and  $e(K_{s,t}) = st$ , and aim to maximise  $st$  subject to  $s + t = n$ , that is, maximise  $s(n - s)$ . So, the best bipartite graph is  $K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}$

But what if some non-bipartite graph is better? In fact, the bipartite graph always wins:

**Theorem 8** (Mantel’s Theorem). Let  $n \geq 3$ . Suppose  $|G| = n, e(G) \geq \lfloor \frac{n^2}{4} \rfloor$  and  $\Delta \not\subset G$ . Then  $G \cong K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}$

*Proof.* Induction on  $n$ . Start with the  $n = 3$  case:

$n = 3$  : Consider  $K_{2,1}$ , it satisfies  $|G| = 3$ ,  $e(G) \geq 2$ ,  $\Delta \subset G$ , as required.

$n > 3$  : Let  $|G| = n$ ,  $e(G) \geq \lfloor \frac{n^2}{4} \rfloor$ ,  $\Delta \subset G$ .

First, remove edges from  $G$  if necessary to get  $H$  with  $|H| = n$ ,  $e(H) = \lfloor \frac{n^2}{4} \rfloor$ . Clearly  $\Delta \not\subset H$ .

Let  $v \in H$  with  $d(v) = \delta(H)$  and let  $K = H - v$  (ie  $H$  with vertex  $v$  and all edges including  $v$  removed). Now,  $|K| = n - 1$ ,  $\Delta \not\subset K$  and  $e(K) = \lfloor \frac{n^2}{4} \rfloor - \delta(H)$ .

Suppose  $n$  is even. Then  $\delta(H) \leq$  average degree of  $H = \frac{2e(H)}{n} = \frac{n^2/2}{n} = \frac{n}{2}$ . Hence  $e(K) \geq \frac{n^2}{4} - \frac{n}{2} = \frac{n^2-2n}{4} = \frac{n^2-2n+1}{4} - \frac{1}{4} = \frac{(n-1)^2}{4} - \frac{1}{4} = \lfloor \frac{(n-1)^2}{4} \rfloor$ . Similarly if  $n$  odd, also get  $e(K) \geq \lfloor \frac{(n-1)^2}{4} \rfloor$ . Hence by the induction hypothesis,  $K \cong K_{\lceil \frac{n-1}{2} \rceil, \lfloor \frac{n-1}{2} \rfloor}$ .

Also  $d(v) = e(H) - e(K)$ . If  $n$  even,  $d(v) = \frac{n^2}{4} - \frac{n^2-2n}{4} = \frac{n}{2}$ .  $H$  is formed by adding a vertex  $v$  to  $K \cong K_{\frac{n}{2}, \frac{n-2}{2}}$  and joining  $v$  to  $\frac{n}{2}$  vertices of  $K$ , without creating a  $\Delta$ .

If  $K$  has bipartition  $(X, Y)$ ,  $v$  cannot be joined both to a vertex in  $X$  and a vertex in  $Y$ . So  $v$  must be joined to all vertices in the larger of  $X, Y$ .

Thus  $H \cong K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}$  and similarly if  $n$  is odd.

We recover  $G$  by adding edges to  $H$  without making a  $\Delta$ . But any new edge creates a  $\Delta$ , so  $G \cong H$ .  $\square$

Hence if  $|G| = n$ ,  $e(G) > \lfloor \frac{n^2}{4} \rfloor$  then  $\Delta \subset G$ . Therefore  $\text{ex}(n; \Delta) = \lfloor \frac{n^2}{4} \rfloor$ .

### 2.1.2 Complete graphs

**Definition.** A graph  $G$   **$r$ -partite** if we can partition  $V(G) = X_1 \cup \dots \cup X_r$  in such a way that if  $xy \in E(G)$  then  $x \in X_i, y \in X_j$  for some  $i \neq j$ . We say  $G$  is **complete  $r$ -partite** if whenever  $x \in X_i, y \in X_j$  with  $i \neq j$  then  $xy \in E(G)$ .

**Definition.** The **Turán graph**  $T_r(n)$  is the complete  $r$ -partite graph with  $n$  vertices and vertex-classes as equal as possible. Write  $t_r(n) = e(T_r(n))$ .

**Example.**

- $T_2(n) = K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil}$
- The diagram shows  $T_3(7)$ .

Properties of the Turán graph

1.  $K_{r+1} \not\subset T_r(n)$ . If we add an edge to  $T_r(n)$  we create a  $K_{r+1}$ .
2.  $T_r(n)$  has more edges than any other  $r$ -partite graph on  $n$  vertices. Suppose  $G$  is a complete  $r$ -partite graph with some two vertex classes  $X, Y$  with  $|Y| > |X| + 1$ . Moving a vertex from  $Y$  to  $X$  loses  $|X|$  edges but gains  $|Y| - 1 > |X|$  edges.
3. If  $r \mid n$  then all classes of  $T_r(n)$  are same. If  $r \nmid n$  then we have ‘small’ classes, with  $k$  vertices, say, and ‘big’ classes with  $k + 1$  vertices.
4. If  $r \mid n$  then  $T_r(n)$  is regular. If  $r \nmid n$  then if  $v$  in big class,  $d(v) = \delta(T_r(n))$  and if  $v$  is in the small class  $d(v) = \Delta(T_r(n)) = \delta(T_r(n)) + 1$ . So in general, vertex degrees in  $T_r(n)$  differ by at most 1.

5. Deleting a vertex of minimal degree from  $T_r(n)$  removes a vertex from a big class, thus making  $T_r(n-1)$ . So  $t_r(n-1) = t_r(n) - \delta(T_r(n))$ .
6. Suppose we have  $T_r(n-1)$ . Add a vertex  $v$ . How many edges can we add without creating a  $K_{r+1}$ ? We can't join  $v$  to a vertex in every class so must miss some class. So the best is to join  $v$  to all vertices except those in a particular small class  $X$ . Same as adding  $v$  to class  $X$ , so we've made  $T_r(n)$ . So # of edges we can add is at most  $t_r(n) - t_r(n-1)$  and only way to achieve this makes  $T_r(n)$ .

**Theorem 9** (Turán's Theorem). Let  $r \geq 2$  and  $|G| = n \geq r+1$ . If  $e(G) \geq t_r(n)$  and  $K_{r+1} \not\subset G$  then  $G \cong T_r(n)$ .

*Proof.* Induction on  $n$ .  $n = r+1$ .  $T_r(r+1)$  has one class of 2 vertices, rest with 1 vertex each. So  $T_r(r+1)$  is  $K_{r+1}$  with 1 edge removed. So  $G \cong T_r(r+1)$ .

$n > r+1$ . First delete edges from  $G$  to form subgraph  $H$  with  $|H| = n$ ,  $e(H) = t_r(n)$  and  $K_{r+1} \not\subset H$ . Let  $v \in H$  have minimal degree and let  $K = H - v$ . We know  $|H| = |T_r(n)|$  and  $e(H) = e(T_r(n))$  so  $H$  and  $T_r(n)$  have same average degree. But degrees in  $T_r(n)$  are as equal as possible by (4).

So  $\delta(H) \leq \delta(T_r(n))$ . Thus  $|K| = n-1$ ,  $K_{r+1} \not\subset K$  and  $e(K) = e(H) - \delta(H) \geq e(H) - \delta(T_r(n)) = t_r(n) - \delta(T_r(n)) = t_r(n-1)$  by 5.

So by induction hypothesis,  $K \cong T_r(n-1)$ . And  $d(v) = e(H) - e(K) = t_r(n) - t_r(n-1)$ . To recover  $H$ , we must add a vertex and  $t_r(n) - t_r(n-1)$  edges to  $K$  without creating a  $K_{r+1}$ . So by 6,  $H \cong T_r(n)$ . To recover  $G$ , add edges to  $H$  without creating a  $K_{r+1}$ . But by (1) we can't add any edges. So,  $G \cong T_r(n)$ .  $\square$

So if  $r \geq 2$ ,  $n \geq r+1$ ,  $\text{ex}(n, K_{r+1}) = t_r(n)$ .

How big is  $t_r(n)$  approximately? Each vertex is joined to a proportion approximately  $1 - \frac{1}{r}$  of all vertices, so we have a proportion approx  $1 - \frac{1}{r}$  of all possible edges. So  $\text{ex}(n; K_{r+1}) \approx (1 - \frac{1}{r})\binom{n}{2}$ . More precisely,

**Corollary 10.** Let  $r \geq 2$ . As  $n \rightarrow \infty$ ,  $\text{ex}(n; K_{r+1}) \sim (1 - \frac{1}{r})\binom{n}{2}$

### 2.1.3 Bipartite graphs

Can think of cycles and paths as graphs (ignore direction and start point). The graph  $C_n$  is cycle of length  $n$  eg  $C_5$ .

**Definition.** The **cyclic graph of order  $n$** , is the cycle of length  $n$ , called  $C_n$ .

**Definition.** The **path graph of order  $n$**  is the path of length  $n$ , called  $P_n$ .

So  $|C_n| = n$ ,  $|P_n| = n+1$ , and  $e(C_n) = e(P_n) = n$ .

**Example.** What can we say about  $\text{ex}(n; C_4)$ ? Note  $C_4 \cong K_{2,2}$ .  $C_4$  is made from two copies of  $P_2$ . Suppose  $G$  is a graph,  $|G| = n$ ,  $e(G) = m$ , and  $C_4 \not\subset G$ . How many  $P_2$  subgraphs does  $G$  have?

Given  $v \in G$ ,  $v$  is the middle vertex of  $\binom{d(v)}{2}$   $P_2$ s. So the total number of  $P_2$ s is  $\sum_{v \in G} \binom{d(v)}{2}$ .

Instead, observe that if  $\{x, y\} \subset V(G)$  with  $x \neq y$  then there is at most one  $P_2$  where  $x, y$  are end vertices (because  $C_4 \not\subset G$ ). So, the number of  $P_2$ s is  $\leq \binom{n}{2}$ .

Hence

$$\sum_{v \in G} \binom{d(v)}{2} \leq \binom{n}{2}$$

Recall Jensen's Inequality: for a convex function  $f$ ,

$$\frac{1}{n} \sum_{i=1}^n f(x_i) \geq f\left(\frac{1}{n} \sum_{i=1}^n x_i\right)$$

Apply Jensen to the earlier bound to give

$$\binom{n}{2} \geq \sum_{v \in G} \binom{d(v)}{2} \geq n \binom{\frac{1}{n} \sum_{v \in G} d(v)}{2} = n \binom{\frac{2m}{n}}{2}$$

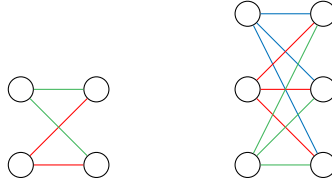
Hence  $n(n-1) \geq n \frac{2m}{n} (\frac{2m}{n} - 1)$ , so  $n^2(n-1) \geq 2m(2m-n)$  so  $4m^2 - 2mn - n^2(n-1) \leq 0$ .

This is a quadratic in  $m$  with roots

$$\begin{aligned} m &= \frac{2n \pm \sqrt{4n^2 + 16n^2(n-1)}}{8} \\ &= \frac{n}{4} (1 \pm \sqrt{1 + 4(n-1)}) \\ &= \frac{n}{4} (1 \pm \sqrt{4n-3}) \end{aligned}$$

Hence  $m \leq \frac{n}{4} (1 + \sqrt{4n-3})$ , so  $\text{ex}(n; C_4) \leq \frac{n}{4} (1 + \sqrt{4n-3})$ , and so  $\text{ex}(n; C_4) = \mathcal{O}(n^{\frac{3}{2}})$ .

A similar approach turns out to give a bound on  $\text{ex}(n; K_{t,t})$  in general:



**Definition.** A **t-fan** in a graph  $G$  is an ordered pair  $(v, W)$  where  $v \in V(G)$ ,  $W \subset V(G)$ ,  $|W| = t$  and  $\forall w \in W, v \sim w$ .

**Example.** 2-fans correspond to  $P_2$  subgraphs.

**Theorem 11.** Let  $t \geq 2$ . Then  $\text{ex}(n; K_{t,t}) = \mathcal{O}(n^{2-\frac{1}{t}})$ .

*Proof.* Let  $|G| = n$ ,  $e(G) = m = \text{ex}(n; K_{t,t})$  and  $K_{t,t} \not\subset G$ . How many  $t$ -fans are there in  $G$ ? Each  $v \in G$  is in  $\binom{d(v)}{t}$   $t$ -fans. This total number is  $\sum_{v \in G} \binom{d(v)}{t}$ . Given  $W \subset V(G)$  with  $|W| = t$ ,  $W$  is in at most  $t-1$   $t$ -fans (as  $K_{t,t} \not\subset G$ ).

So the total number of  $t$ -fans is  $\leq \binom{n}{t}(t-1)$ .

Hence

$$\begin{aligned} \binom{n}{t}(t-1) &\geq \sum_{v \in G} \binom{d(v)}{t} \\ &\geq n \binom{\frac{1}{n} \sum_{v \in G} d(v)}{t} \text{ by Jensen} \\ &= n \binom{\frac{2m}{n}}{t} \end{aligned}$$

So  $\frac{n^t}{t!}(t-1) \geq \frac{n}{t!}(\frac{2m}{n} - t)^t$  so  $n^t(t-1) \geq n(\frac{2m}{n} - t)^t$ .

If  $n$  is sufficiently large (as we may assume) then  $m \geq nt$  and so  $\frac{m}{n} \geq t$  and so  $\frac{2m}{n} - t \geq \frac{m}{n}$ . Thus  $n^t(t-1) \geq n(\frac{m}{n})^t$  so  $m^t \leq n^{2t-1}(t-1)$  so  $m \leq (t-1)^{\frac{1}{t}} n^{2-\frac{1}{t}} = \mathcal{O}(n^{2-\frac{1}{t}})$ .  $\square$

We will see lower bounds later.

Closely related is the **Problem of Zarankiewicz**. Let  $z(n, t)$  be the largest possible number of edges in a bipartite graph  $G$  with  $n$  vertices in each class and  $K_{t,t} \not\subset G$ .

**Theorem 12.** Let  $t \geq 2$ . Then  $z(n, t) = \mathcal{O}(n^{2-\frac{1}{t}})$ .

*Proof.* Let  $G$  be bipartite with classes  $X, Y$  with  $|X| = |Y| = n$ ,  $e(G) = m = z(n, t)$  and  $K_{t,t} \not\subset G$ .

Count number of  $t$ -fans with vertex in  $X$  and set in  $Y$ . Similar to the proof of [theorem 11](#),

$$\binom{n}{t}(t-1) \geq \sum_{v \in X} \binom{d(v)}{t}$$

Now  $\sum_{v \in X} d(v) = m$ , so a similar calculation to [theorem 11](#) gives  $m = \mathcal{O}(n^{2-\frac{1}{t}})$ .  $\square$

#### 2.1.4 General graphs

What is  $\text{ex}(n; H)$ ? This is a hard question, and we could not even find an exact answer for  $H = C_4$ , so instead consider some asymptotics. Usually  $\text{ex}(n; H) \rightarrow \infty$  as  $n \rightarrow \infty$ .

Instead what about  $\lim_{n \rightarrow \infty} \frac{\text{ex}(n; H)}{\binom{n}{2}}$ ? Does it even exist?

**Proposition 13.** Let  $H$  be a graph with at least one edge, and for  $n \geq |H|$ , let  $x_n = \frac{\text{ex}(n; H)}{\binom{n}{2}}$ . Then  $(x_n)$  converges.

*Proof.* Let  $n > |H|$ , let  $|G| = n$ ,  $e(G) = \text{ex}(n; H) = x_n \binom{n}{2}$  and  $H \not\subset G$ . For any  $v \in G$ ,  $|G - v| = n - 1$  and  $H \not\subset G - v$ , so  $e(G - v) \leq \text{ex}(n - 1; H) = x_{n-1} \binom{n-1}{2}$ . Each edge  $xy \in E(G)$  in  $G - v$  for all  $v \neq x, y$ . Hence  $(n - 2)x_n \binom{n}{2} = (n - 2)e(G) = \sum_{v \in G} e(G - v) \leq nx_{n-1} \binom{n-1}{2}$ . So  $x_n \leq x_{n-1}$ , so  $(x_n)$  is decreasing and bounded below by zero.  $\square$

Write  $\text{ex}(H) = \lim_{n \rightarrow \infty} \frac{\text{ex}(n; H)}{\binom{n}{2}}$  which exists by [proposition 13](#). We have seen  $\text{ex}(K_{r+1}) = 1 - \frac{1}{r}$ , by [corollary 10](#). From [theorem 11](#),  $\text{ex}(n, K_{t,t}) = \mathcal{O}(n^{2-\frac{1}{t}}) = o(n^2) = o(\binom{n}{2})$ , so  $\text{ex}(K_{t,t}) = 0$ .

**Fact.** Turán says that a proportion  $1 - \frac{1}{r}$  of edges is enough to guarantee a  $K_{r+1}$ . Increasing this proportion by a tiny amount gives *much* more.

**Definition.** Write  $K_r(t)$  for the complete  $r$ -partite graph with  $t$  vertices in each class (so  $K_r(t) = T_r(rt)$ ).

**Theorem 14** (Erdős-Stone Theorem). Let  $r, t \geq 1$  be integers, and let  $\epsilon > 0$  be real. Then  $\exists n_0$  such that  $\forall n \geq n_0$ ,  $|G| = n$ ,  $e(G) \geq (1 - \frac{1}{r} + \epsilon) \binom{n}{2} \implies K_{r+1}(t) \subset G$ .

*Proof.* Coming soon.  $\square$

This is enough to determine  $\text{ex}(H)$  for all  $H$ .



**Definition** (Chromatic number). If  $H$  is a graph, the **chromatic number** of  $H$ , denoted  $\chi(H)$ , is the least  $r$  such that  $H$  is  $r$ -partite.

**Example.**  $\Psi(K_r(t)) = r$ .

**Corollary 15.** Let  $H$  be a graph with at least one edge. Then

$$\text{ex}(H) = 1 - \frac{1}{\chi(H) - 1}.$$

*Proof.* Let  $r = \chi(H) - 1$ . Then  $H$  is  $(r+1)$ -partite so we can find  $t$  such that  $H \subset K_{r+1}(t)$  (for instance  $t = |H|$  suffices). Let  $\epsilon > 0$ , and take  $n_0$  as in [theorem 14](#). Then for all  $n \geq n_0$ ,

$$\begin{aligned} |G| = n, \quad e(G) \geq \left(1 - \frac{1}{r} + \epsilon\right) \binom{n}{2} &\implies K_{r+1}(t) \subset G \\ &\implies H \subset G. \end{aligned}$$

So, for all  $n \geq n_0$ ,  $\text{ex}(n; H) < \left(1 - \frac{1}{r} + \epsilon\right) \binom{n}{2}$ . Hence

$$\text{ex}(H) = \lim_{n \rightarrow \infty} \frac{\text{ex}(n; H)}{\binom{n}{2}} \leq 1 - \frac{1}{r} + \epsilon.$$

But  $\epsilon > 0$  was arbitrary, so  $\text{ex}(H) \leq 1 - \frac{1}{r}$ .

On the other hand, for all  $n$ ,  $H \not\subset T_r(n)$  as  $H$  is not  $r$ -partite. So  $\text{ex}(n, H) \geq t_r(n)$ , and

$$\frac{t_r(n)}{\binom{n}{2}} \rightarrow 1 - \frac{1}{r} \quad \text{so} \quad \text{ex}(H) \geq 1 - \frac{1}{r}.$$

□

So if  $H$  not bipartite,  $\text{ex}(H) = 1 - \frac{1}{\chi(H)-1} \neq 0$ , and  $\text{ex}(n; H) \sim \left(1 - \frac{1}{\chi(H)-1}\right) \binom{n}{2}$ . This essentially solves the forbidden subgraph problem for non-bipartite  $H$ . What about for  $H$  bipartite? [Corollary 15](#) just says  $\text{ex}(H) = 0$ , i.e.  $\text{ex}(n; H) = o(n^2)$ , but doesn't tell us for instance that  $\text{ex}(n; H)$  grows like a certain power of  $n$ . We do however have upper bounds from [theorem 11](#):  $\text{ex}(n, K_{t,t}) = \mathcal{O}(n^{2-\frac{1}{t}})$ .

Another application:

**Definition** (Density). We can define the **density** of a graph  $G$  to be

$$D(G) = \frac{e(G)}{\binom{|G|}{2}} \in [0, 1]$$

**Example.**

- $D(T_r(n)) \rightarrow 1 - \frac{1}{r}$  as  $n \rightarrow \infty$
- If  $\alpha > \text{ex}(H)$  then if  $|G|$  is sufficiently large and  $D(G) = \alpha$ , we have  $H \subset G$ .

Could we define the **density** of an infinite graph? We cannot take the same definition, as we would get  $\infty$ . Instead try something like the 'limit of densities of large finite subgraphs'.

**Definition.** The upper density of an infinite graph  $G$  is

$$\text{ud}(G) = \lim_{n \rightarrow \infty} \sup \{ D(H) \mid H \subset G, |H| = n \}.$$

**Remark.**

- Again  $\text{ud}(G) \in [0, 1]$ .
- If  $\alpha < \text{ud}(G)$  then for sufficiently large  $n$ ,  $G$  has subgraphs of order  $n$  and density  $\geq \alpha$ .
- If  $\alpha > \text{ud}(G)$  then for sufficiently large  $n$ ,  $G$  has no subgraphs of order  $n$  and density  $\geq \alpha$ .
- A priori, it seems like  $\text{ud}(G)$  could take any value in  $[0, 1]$ . But surprisingly:

**Corollary 16.** For any infinite graph  $G$ ,

$$\text{ud}(G) \in \{0, 1, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots\}.$$

*Proof.* It is enough to show that for  $r = 1, 2, 3, \dots$ ,

$$\text{ud}(G) > 1 - \frac{1}{r} \implies \text{ud}(G) \geq 1 - \frac{1}{r+1}.$$

Suppose  $\text{ud}(G) > 1 - \frac{1}{r}$ . Pick  $\alpha$  such that  $\text{ud}(G) > \alpha > 1 - \frac{1}{r}$  and fix  $n$ . For sufficiently large  $N$ ,  $G$  has subgraphs of order  $N$  and density  $\geq \alpha > 1 - \frac{1}{r} = \text{ex}(T_{r+1}(n))$ . So  $T_{r+1}(n) \subset G$ , but  $D(T_{r+1}(n)) \rightarrow 1 - \frac{1}{r+1}$  as  $n \rightarrow \infty$ . We can do this for every  $n$ , hence  $\text{ud}(G) \geq 1 - \frac{1}{r+1}$ .  $\square$

### 2.1.5 Proof of Erdős-Stone

**Theorem 14** says that if a graph  $G$  has enough edges, then  $K_{r+1}(t) \subset G$ .

This condition on  $e(G)$  is equivalent to a condition on the average degree. However, given  $d$  and  $\delta > 0$  then it turns out that any large graph of average degree  $\geq d$  has a large subgraph of minimum degree  $\geq d - \delta$ . If  $G$  has average degree  $d$  and  $|G| = n$  large, we can find  $|H| = \Omega(n)$  and  $\delta(H) \geq d - \delta$  (not proved here. Delete vertices of minimum degree and it works). So, we have an equivalent version of Erdős-Stone with minimum degree condition:

**Theorem 14a.** Let  $r, t \geq 1$  be integers and  $\epsilon > 0$  be real. Then  $\exists n_0 \forall n \geq n_0$ ,

$$|G| = n, \quad \delta(G) \geq (1 - \frac{1}{r} + \epsilon)n \implies K_{r+1}(t) \subset G.$$

*Proof.* Induction on  $r$ . Fix  $T$  such that

Then choose  $n_0$  such that for all  $n \geq n_0$ ,

$$|G| = n, \quad \delta(G) \geq (1 - \frac{1}{r} + \epsilon)n \implies K_{r+1}(t) \subset G. \tag{1}$$

How? For  $r = 1$ ,  $K_1(T) \subset G \iff |G| \geq T$ . For  $r > 1$ ,  $1 - \frac{1}{r} > 1 - \frac{1}{r-1}$  so it follows by inductive hypothesis.

Suppose the result is not true. Then we can find arbitrarily large  $n$  and graphs  $G$  with  $|G| = n$  and  $\delta(G) \geq (1 - \frac{1}{r} + \epsilon)n$ ,  $K_{r+1} \not\subset G$ . Pick such an  $n$ ,  $G$  with  $n \geq n_0$  and also

Then we can find  $K_r(T) \subset G$ , say with vertex classes  $X_1, \dots, X_r$ . Let

$$A = \{ (W, v_1, \dots, v_r) \mid |W| = t, W \subset V(G), \forall i, \forall v_i \in X_i, \forall w \in W, v_i \sim w \}$$

What can we say about  $|A|$ ? First, given  $v_1 \in X, \dots, v_r \in X_r$ , we can check from the minimum degree condition that

$$|\Gamma(v_1) \cap \dots \cap \Gamma(v_r)| \geq r\epsilon n$$

So there are at least  $\binom{r\epsilon n}{t}$  choices for  $W$ . Hence  $|A| \geq T^r \binom{r\epsilon n}{t}$ .

On the other hand, given the set  $W$ , as  $K_{r+1}(t) \not\subset G$ , we know that there is some  $Y_i$  containing at most  $t-1$  vertices joined to all of  $W$ . Hence  $|A| \leq \binom{n}{t}(t-1)T^{r-1}$ , thus

$$T^r \binom{r\epsilon n}{t} \leq \binom{n}{t}(t-1)T^{r-1}.$$

Now,

$$\text{RHS} \leq \frac{n^t}{t!}(t-1)T^{r-1}$$

while

$$\text{LHS} \geq T^r \frac{1}{t!}(r\epsilon n - t)^t \geq T^r \frac{1}{t!} \left(\frac{r\epsilon n}{2}\right)^t.$$

Combining, we get

$$T^r \left(\frac{r\epsilon n}{2}\right)^t \leq n^t(t-1)T^{r-1}$$

hence  $T \leq \left(\frac{2}{r\epsilon}\right)^t(t-1)$ , a contradiction.  $\square$

## 2.2 Hamiltonian graphs

**Definition** (Hamiltonian). A **Hamiltonian cycle** in a graph  $G$  is a cycle of length  $|G|$ , i.e., going through all vertices of  $G$ . If  $G$  has a Hamiltonian cycle, we say  $G$  is **Hamiltonian**.

If  $|G| = n$ , how big does  $e(G)$  need to be to force  $G$  **Hamiltonian**? This turns out to be uninteresting, we can have  $e(G)$  almost  $\binom{n}{2}$  and have no Hamiltonian cycle.

More interestingly, is there a condition on  $\delta(G)$ ?

**Theorem 17** (Dirac's Theorem). Let  $|G| = n \geq 3$  and  $\delta(G) \geq \frac{n}{2}$ . Then  $G$  is Hamiltonian.

**Remark.**

1. This result is best possible:
2. Of course there are Hamiltonian graphs of smaller minimum degree:
3. In general, given a graph  $G$  it is hard computationally to determine if  $G$  has a Hamiltonian cycle.

*Proof.* First, observe  $G$  is connected. Indeed, if  $x \neq y$ ,  $x \sim y$ , then  $|\Gamma(x) \cup \Gamma(y)| \leq n-2$ , but  $|\Gamma(x)| + |\Gamma(y)| \geq \frac{n}{2} + \frac{n}{2} = n$  so  $\Gamma(x) \cap \Gamma(y) \neq \emptyset$ . Let  $v_0, v_1, \dots, v_n$  be a path of maximal length in  $G$ , say length  $k \leq n-1$ . By maximality,  $\Gamma(v_0) \subset \{v_1, \dots, v_n\}$ . Similarly,  $\Gamma(v_k) \subset \{v_0, \dots, v_{k-1}\}$ . If we have some situation like in the diagram, we get a cycle. To be more precise, if

$$A = \{i \in [k] \mid v_0 \sim v_i\} \quad \text{and} \quad B = \{i \in [k] \mid v_k \sim v_{i-1}\}$$

then  $A \cap B \neq \emptyset \implies$  we have a cycle. But  $|A \cup B| \leq k < n$  while  $|A| + |B| \geq \frac{n}{2} + \frac{n}{2} = n$ . Hence  $\exists i \in A \cap B$  so we have a cycle  $C = v_0 v_1 \dots v_{i-1} v_k v_{i-1} \dots v_i v_0$  of length  $k+1$ .

If  $k = n - 1$ , we have a Hamiltonian cycle as required.

If  $k < n - 1$ , relabel the cycle  $C = u_0 u_1 \dots u_k u_0$ . By connectedness, we have some  $u_i \in C$  and  $w \notin C$  with  $w \sim u_i$ . Then  $w u_i u_{i+1} \dots u_k u_0 \dots u_{i-1}$  is a path of length  $k + 1$ , contradicting maximality.  $\square$

**Remark.** The same proof gives the following: Let  $n > k \geq 3$  and let  $G$  be a connected graph  $|G| = n$  with  $\delta(G) \geq \frac{k}{2}$ . Then  $G$  contains a path of length  $k$  and a cycle of length at least  $\frac{k+2}{2}$ .

The Königsberg bridge problem (link) seems superficially similar, but in fact much less interesting.

**Definition** (Euler circuit). A **circuit** of a graph  $G$  is a sequence  $v_0 v_1 \dots v_n$  of vertices of  $G$ , not necessarily distinct with  $v_0 = v_n$ , where if  $1 \leq i \leq k$  then  $v_{i-1} \sim v_i$  and if  $1 \leq i < j \leq k$  then edges  $v_{i-1} v_i$  and  $v_{j-1} v_j$  are distinct. It is an **Euler circuit** if for every  $e \in E(G)$ , there is some  $i$  with  $e = v_{i-1} v_i$ . If  $G$  has an Euler circuit we say  $G$  is **Eulerian**.

**Proposition 18.** Let  $G$  be a connected graph. Then  $G$  Eulerian  $\iff \forall v \in G, d(v)$  is even.

*Proof.*  $\implies$  is obvious: an **Eulerian circuit** must go in and out of a given vertex the same number of times.

$\impliedby$  : use induction on  $e(G)$ . For  $e(G) = 0$  it is clearly true.

Consider  $e(G) > 0$ . Let  $v_0 v_1 \dots v_k = C$  be a circuit in  $G$  of maximal length. If  $C$  uses all edges of  $G$  then we are done. If not, delete all edges used in  $C$  from  $G$  to form  $H$ . In  $H$ , every vertex still has even degree. Let  $H_1$  be a component of  $H$  with at least one edge.

By induction hypothesis,  $H_1$  has an Euler circuit  $D$ . Certainly  $C, D$  meet at some vertex  $v$ . Join them at  $v$  to produce a longer circuit in  $G$ , a contradiction. (Walk along  $C$  until we get to  $v$ , then walk all round  $D$  starting/ending at  $v$ , then walk along the rest of  $C$ ).  $\square$

## 3 Graph Colouring

### 3.1 Planar Graphs

**Definition** (Graph drawing). A **drawing** of  $G$  is an ordered pair  $(f, \gamma)$  where  $f : V \rightarrow \mathbb{R}^2$  is an injection and  $\gamma : E \rightarrow C([0, 1], \mathbb{R}^2)$  such that

- (i) If  $uv \in E$  then  $\{\gamma(uv)(0), \gamma(uv)(1)\} = \{f(u), f(v)\}$ .
- (ii) If  $e, e' \in E$  with  $e \neq e'$  then  $\gamma(e)((0, 1)) \cap \gamma(e')((0, 1)) = \emptyset$ .
- (iii) If  $e \in E$  then  $\gamma(e)$  is injective.
- (iv) If  $e \in E$  and  $v \in V$  then  $f(v) \notin \gamma(e)((0, 1))$

That is, vertices  $\leftrightarrow$  points and edges  $\leftrightarrow$  continuous curves between end vertices, with no unnecessary intersections. If  $G$  has a drawing, we say  $G$  is **planar**.

Are there topological problems? Perhaps an edge is represented by a space-filling curve. But,

**Fact.** Any planar graph can be drawn with piecewise linear edges - i.e. each edge is finitely many line segments. Henceforth, assume all drawings satisfy this.

**Example.**

1.  $K_3$  is planar
2.  $K_4$
3.  $K_5$  Suppose we have a drawing of  $K_5$ . It must contain a closed curve, say through vertices  $u, v, w, x, y$  in that order. The curve divides the plane into inside and outside. Suppose edge  $vy$  is inside. Then  $ux, uw$  are outside. So  $yu, vx$  inside, so they intersect, a contradiction. Similarly,  $vy$  cannot be outside, hence  $K_5$  is not planar.
4.  $K_{3,3}$  is not planar by similar argument (exercise).

Clearly if  $K_5 \subset G$  or  $K_{3,3} \subset G$  then  $G$  is not planar. Remarkably, these are the only obstacles to planarity. Also, eg

**Definition** (Subdivision). Let  $G$  be a graph. A **subdivision** of  $G$  is a graph formed by repeatedly selecting  $vw \in E(G)$ , removing  $vw$  and adding vertex  $u$  and edges  $uv, uw$ .

**Theorem 19** (Kuratowski's Theorem). Let  $G$  be a graph. Then  $G$  planar iff  $G$  contains no subdivision of  $K_5$  or  $K_{3,3}$ .

Recall that a tree is a connected acyclic graph.

**Definition.** A **leaf** of a tree is a vertex of order 1.

**Proposition 20.** Every tree of order at least 2 has a leaf.

*Proof.* Let  $T$  be a tree  $|T| \geq 2$  and let  $v_0 v_1 \dots v_k$  be a path of maximum length  $T$ . Now  $v_k \sim v_{k-1}$ , but  $v_k$  has no other neighbours in path (as  $T$  acyclic) and  $v_k$  has no neighbours outside path (by maximality). Hence  $v_k$  is a leaf.  $\square$

Why do we care? If  $T$  is a tree,  $|T| \geq 2$ , then  $T$  has a leaf  $v$  and  $T - v$  is a tree of smaller order. This is particularly useful for induction, for instance

**Proposition 21.** Let  $T$  be a tree,  $|T| = n \geq 1$ . Then  $e(T) = n - 1$ .

*Proof.* Use induction on  $n$ . For  $n = 1$ ,  $e(T) = 0$  as required.

For  $n > 1$ , let  $v$  be a leaf. Then  $T - v$  is a tree with  $|T - v| = n - 1$  so by the induction hypothesis,  $e(T - v) = n - 2$ , so  $e(T) = n - 1$ .  $\square$

**Proposition 22.** Every tree is planar.

*Proof.* Let  $T$  be a tree,  $|T| = n$  and use induction on  $n$ . For  $n = 0, 1$  we are done immediately.

For  $n > 1$ , let  $v \in T$  be a leaf. By the induction hypothesis,  $T - v$  can be drawn. Let  $u \in T$  be the neighbour of  $v$ . Take a small circle around  $u$  in the drawing; so the circle contains only the centre and some radii in the drawing. Hence, we can easily add  $v$  and  $uv$  to the drawing.  $\square$

**Definition (Faces).** If we have a drawing of a graph, it divides the plane into connected regions called **faces**. Precisely one of these regions, the **infinite face** is unbounded.

**Remark.** Important warning: Graphs don't have faces, drawings have faces.

**Theorem 23 (Euler's Formula).**  $G$  connected and planar. Take  $|G| = n \geq 1$ ,  $e(G) = m$  with  $l$  faces. Then  $n - m + l = 2$ .

*Proof.* Induction on  $m$ . If  $G$  is a tree:  $m = n - 1$ ,  $l = 1$ , and we are done.

Otherwise,  $G$  has a cycle. Pick an edge  $e$  in the cycle and consider  $G - e$ . Then  $|G - e| = n$ ,  $e(G - e) = m - 1$ . Moreover, in our drawing, removing  $e$  combines two faces so we have drawn  $G - e$  with  $l - 1$  faces. So, by the induction hypothesis  $n - (m - 1) + (l - 1) = 2$ , so  $n - m + l = 2$ .  $\square$

**Corollary 24.** Let  $G$  be planar,  $|G| = n \geq 3$ . Then  $e(G) \leq 3n - 6$ .

*Proof.* Let  $e(G) = m$ . Draw  $G$  with  $l$  faces. Without loss of generality,  $G$  is connected (if not, add edges to make it so).  $\square$

**Proposition 25 (Six colour theorem).** Any planar graph is 6-colourable.

*Proof.* Let  $G$  be planar,  $|G| = n$ . Induction on  $n$ . For  $n \leq 6$ , give every vertex a different colour, as required.

For  $n > 6$ , by [corollary 24](#),  $e(G) \leq 3n - 6$ . Hence,

$$\delta(G) \leq \frac{2e(G)}{|G|} \leq \frac{6n - 12}{n} = 6 - \frac{12}{n} < 6.$$

So,  $\delta(G) \leq 5$ .

Pick  $v \in G$  with  $d(v) \leq 5$ . By inductive hypothesis,  $G - v$  can be 6-coloured. Some colour is missing from  $\Gamma(v)$ , use this colour to colour  $v$ .  $\square$

A lot more work gives:

**Theorem 26 (Five colour theorem).** Any planar graph is 5-colourable.

*Proof.* Let  $G$  be planar,  $|G| = n$ . Use induction on  $n$ . The base case  $n = 5$  is trivial, so take  $n > 5$ . As in [proposition 25](#), we can find  $v \in G$  with  $\delta(v) \leq 5$  and 5-colour  $G - v$ . If there is a colour missing from  $\Gamma(v)$ , use that colour at  $v$ . Otherwise, draw  $G$ ; wlog  $v$  has neighbours  $x_1, \dots, x_5$  in clockwise order around  $v$  with colours  $1, \dots, 5$  respectively.

Call a path in  $G - v$  an  $ij$ -path if all its vertices have colour  $i$  or  $j$ . Given  $x \in G - v$ , the  $i - j$  component of  $x$  consists of those vertices reachable from  $x$  along  $ij$ -paths.

If  $x_1, x_3$  are in different 13 components, then swap the colours of 1, 3 on the 13 component of  $x$  and give  $v$  colour 1. If not, then  $x_2, x_4$  are in different 24 components as we see in the diagram, so swap colours 2, 4 on 24 component of  $x_2$  and give  $v$  colour 2.  $\square$

**Theorem** (Four colour theorem). Any planar graph is 4-colourable.

*Proof.* Let  $G$  be planar,  $|G| = n$  and use induction on  $n$ . For  $n = 4$ , we are immediately done.

For  $n > 4$ , pick  $v \in G$  with  $d(v) \leq 5$ , then draw  $G$  and 4-colour  $G - v$ . If some colour is missing on  $\Gamma(v)$  then we are done. If not, there are three cases.

Case 1:  $d(v) = 4$ . There cannot be both a 13-path from  $x_1$  to  $x_3$  and a 24 path from  $x_2$  to  $x_4$ , so done as in the proof of [theorem 26](#).

Case 2:  $d(v) = 5$ ,  $v$  has neighbours  $x_1, \dots, x_5$  clockwise with (wlog) colours 1, 2, 3, 4, 1. Then there is either no 24 path from  $x_2$  to  $x_4$  or no 13 path from  $x_3$  to  $x_1$ , so again we are done.

Case 3:  $d(v) = 5$ ,  $\square$

The four colour theorem was proved in 1976, with 1936 cases checked.

## 3.2 General Graphs

We know  $G$  planar  $\implies \chi(G) \leq 4$ .

What can we say about  $\chi(G)$  if not given  $G$  planar? Clearly  $\chi(G)$  can be as large as we like, e.g.  $\chi(K_k) = k$ .

### 3.2.1 Lower bounds

- (a) If  $K_k \subset G$  then  $\chi(G) \geq k$ . The **clique number** of  $G$  is  $\omega(G) = \max \{ k \mid K_k \subset G \}$ . Then  $\chi(G) \leq \omega(G)$ . Sometimes this bound is not good. For instance, we can find  $G$  with  $\omega(G) = 2$  and  $\chi(G)$  arbitrarily large (see example sheet 2).

- (b) Suppose we have a colouring of  $G$ . Let  $W \subset V(G)$  be a colour class. Then  $W$  is an **independent set** - there are no edges within  $W$ . The **independence number** of  $G$  is

$$\alpha(G) = \max \{ |W| \mid W \subset V(G), W \text{ independent} \}.$$

Then  $\chi(G) \geq \frac{|G|}{\alpha(G)}$ . Sometimes this bound is not good. For instance, take  $G = K_k \cup \overline{K_k}$ . Then  $\chi(G) = k$ ,  $\frac{|G|}{\alpha(G)} = \frac{2k}{k+1} \approx 2$ .