Part II – Graph Theory

Based on lectures by Prof. P. Russell
Notes taken by Bhavik Mehta

Michaelmas 2017

Contents

0 Introduction **3**
0.1 Preliminary . . . 3
0.2 Informal definitions . . . 3
0.3 Where do such structures arise? . . . 3

1 Ramsey Theory **6**
1.1 Basic Terminology . . . 12

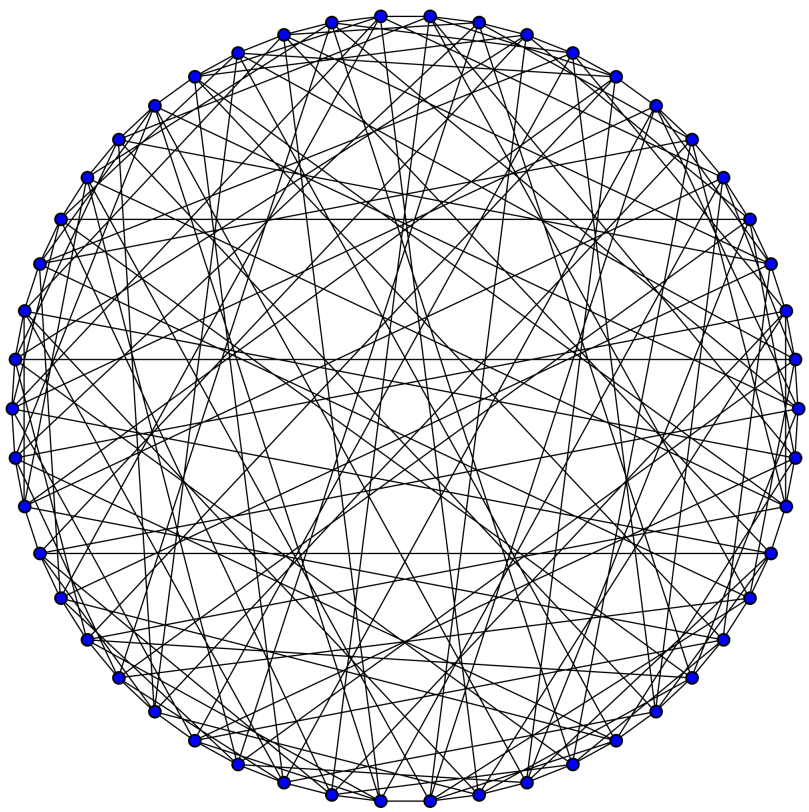
2 Extremal Graph Theory **16**
2.1 Forbidden Subgraph Problem . . . 16
2.1.1 Triangles . . . 16
2.1.2 Complete graphs . . . 17
2.1.3 Bipartite graphs . . . 19
2.1.4 General graphs . . . 21
2.1.5 Proof of Erdős-Stone (non-examinable) . . . 24
2.2 Hamiltonian graphs . . . 25

3 Graph Colouring **28**
3.1 Planar Graphs . . . 28
3.2 General Graphs . . . 33
3.3 Graphs on surfaces . . . 35
3.4 Edge Colouring . . . 36

4 Connectivity **38**
4.1 The Marriage Problem . . . 38
4.2 Connectivity . . . 39
4.3 Edge connectivity . . . 41

5 Probabilistic Techniques **42**
5.1 The Probabilistic Method . . . 42
5.2 Modifying a Random Graph . . . 42
5.3 The Structure of Random Graphs . . . 45

6 Algebraic Methods **49**
6.1 The Chromatic Polynomial 50
6.2 Eigenvalues 51
6.3 Strongly Regular Graphs 53



The unique 7-regular Moore graph

0 Introduction

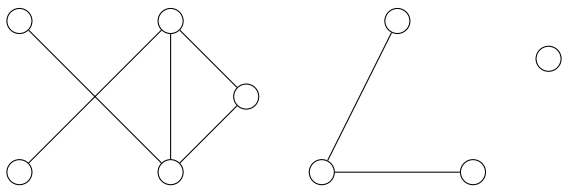
0.1 Preliminary

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. Bollobás.

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'.



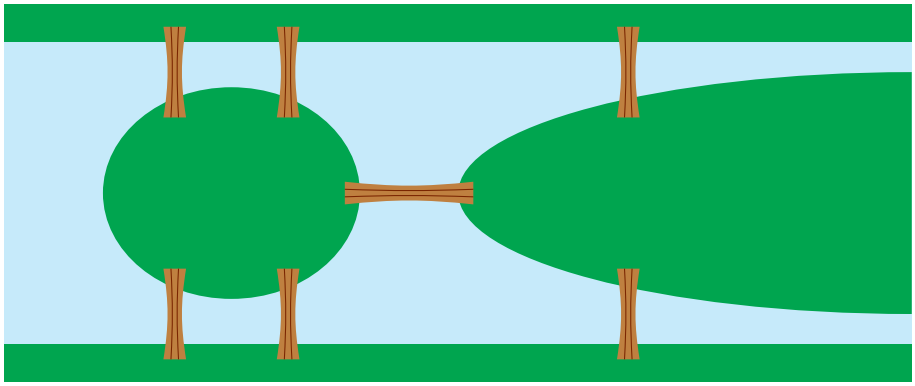
In this course, we make the following assumptions:

- 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.

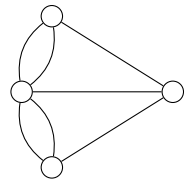
0.3 Where do such structures arise?

Example.

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

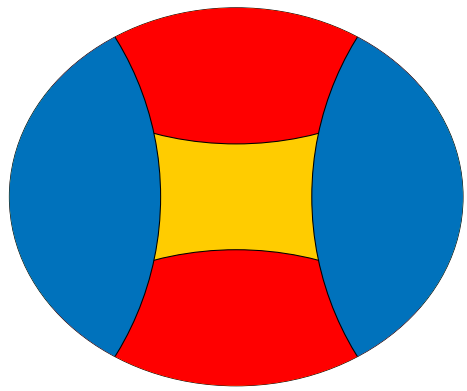


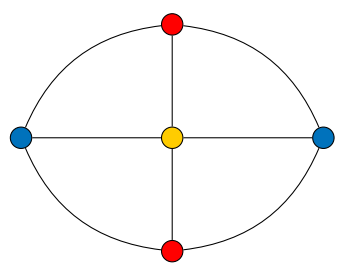
We turn the problem into a graph problem - in particular, a multigraph to allow multiple edges between vertices. The vertices of this graph are the four bits of the city, while the edges are the seven bridges. In terms of this graph, the problem becomes to walk around the edges of this graph precisely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





   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, $\exists a_1, \ldots, a_n \in G$ such that the left cosets of H are $a_1H, \ldots a_nH$. Similarly, $\exists b_1, \ldots, b_n \in G$ such that the right cosets of H are $Ha_1, \ldots, Ha_n$. Can we do this simultaneously? In particular, are there $a_1, \ldots, a_n \in G$ such that $a_1H, \ldots, a_nH \in G$ are the left cosets, and $Ha_1, \ldots, Ha_n \in G$ are the right cosets?

   In the case where $H \lhd G$ this is easy since we have $aH = Ha \ \forall a \in G$. However it is less obvious if $H \ntriangleleft G$. 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 nth roots an element can have. So, G can be partitioned into $g_1H, g_2H, \ldots, g_mH$ for some $g_1, g_2, \ldots, g_m$ 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

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

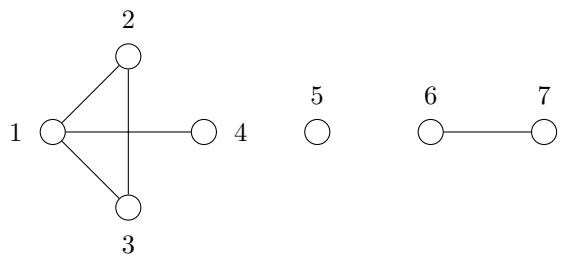
   It finally remains to show Schur's Theorem:

   Theorem (Schur's Theorem)**.** Let n be a positive integer. Then if p is a sufficiently large positive integer, whenever $\{1, 2, \ldots, p\}$ 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 $uv = vu$.

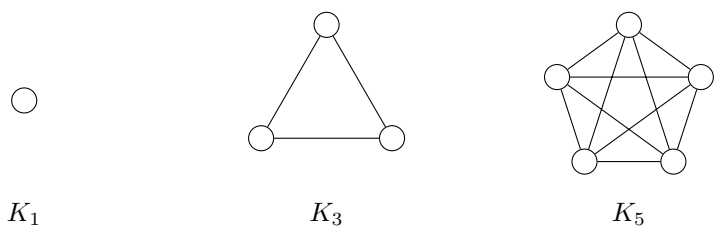
Example. Here's an example of a graph $G = (\{1, 2, 3, 4, 5, 6, 7\}, \{12, 23, 13, 14, 67\})$, but it's often easier to represent a graph by a drawing: take a point for each vertex, join two vertices if they are in an edge.



Definition (Isomorphism)**.** Let $G = (V, E)$ and $G' = (V', E')$ be graphs. An **isomorphism** from G to G' is a bijection $\phi : V \to V'$ 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 $[n] = \{1, 2, \ldots, n\}$ 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. □

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

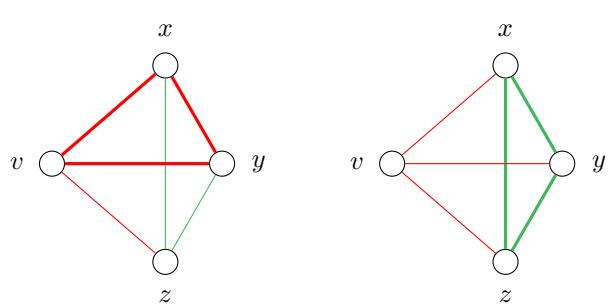
Proof. (Schur's theorem, $k = 3$) Suppose $[16]$ are coloured red, green and blue. By the pigeonhole principle, some six numbers are the same colour, without loss of generality $x_1 < x_2 < \cdots < x_6$ are red. If $x_j - x_i$ is red for any $i < j$ then we are done: $x_i + (x_j - x_i) = x_j$. So assume 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. □

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.

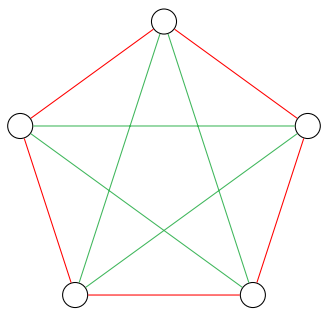
Suppose the edges of a complete graph K_6 are coloured red and green. Then we can find a monochromatic triangle.

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.



□

Remark. Schur for $k = 2$ now follows: Suppose $[5]$ are coloured red/green. Take a K_6 with vertices $\{0, 1, \ldots, 5\}$. 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.



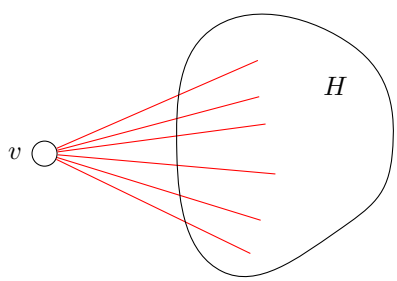
In this graph form, the statement for more colours is now amenable to an induction proof.

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 $n = k(m - 1) + 2$.

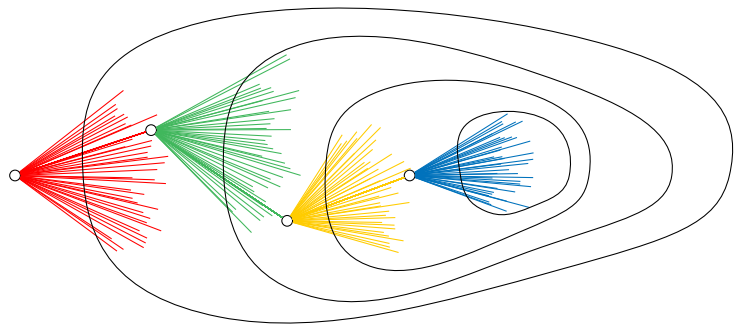
Now k-colour the edges of K_n. Pick vertex v. The number of edges containing v is $n - 1 = k(m - 1) + 1$. So some m of them are the same colour, without loss of generality blue. Let H be a K_m joined to v by blue edges. If H contains a blue edge, it makes a blue triangle with v. If not, H is a $(k-1)$-coloured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



□

Remark.

1. Can think in terms of the following picture. 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.

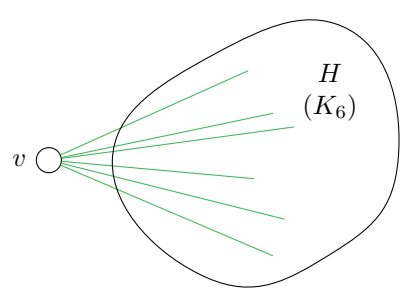
1. Let K_{10} have edges coloured red/green. Then there is a red triangle or a green K_4.

   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 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 $R(s) = \mathcal{O}\left(3.999999999^s\right)$? 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 $c : E \to [k]$. We mostly refer to the colours as 'red', 'green' and so on. We say $H = (W, F)$ is monochromatic if $c|_F$ is constant.

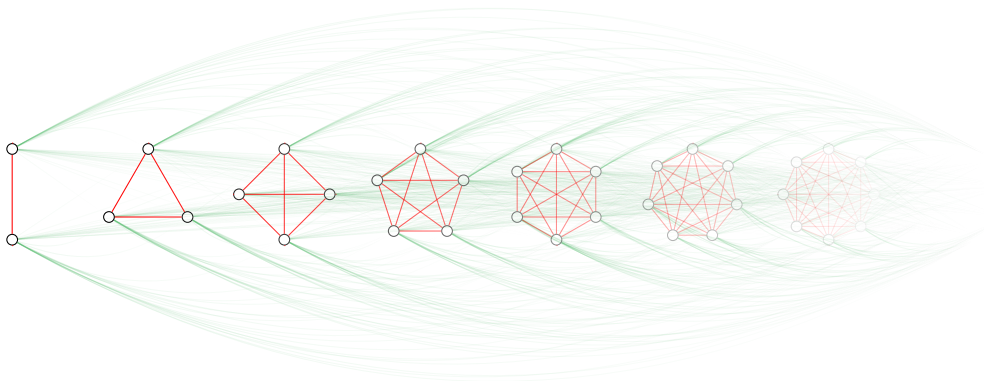
With any number of colours, we can find arbitrarily large monochromatic subgraphs. What about infinite ones?

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_∞, the **infinite complete graph**, 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_∞ are coloured red/green. What can we find? By Ramsey's Theorem, we can find a monochromatic K_s for every finite $s = 1, 2, 3, \ldots$. Note this doesn't immediately give a monochromatic K_∞:



In the diagram, we see that for any finite s we have a red K_s but no red K_∞. (However in this example, there is a green K_∞)

Theorem 5 (Infinite Ramsey Theorem)**.** Let $k \geq 1$. Whenever the edges of K_∞ are k-coloured, we have a monochromatic K_∞ 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_∞ such that for all $u \in A_1$, $v_1 u$ 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 $v_2 u$ for $(u \in A_2)$ have colour c_2. Keep going. We get infinite sequences $v_1, v_2, v_3, \ldots$ of vertices, $c_1, c_2, c_3, \ldots$ of colours and $A_1 \supset A_2 \supset A_3 \supset \ldots$ 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 $i_1 < i_2 < i_3 < \ldots$ such that $c_{i_1} = c_{i_2} = c_{i_3} = \ldots$. Consider $v_{i_1}, v_{i_2}, v_{i_3}, \ldots$. Any edge between two of these vertices has colour c_{i_1}. So we have a monochromatic K_∞. $\square$

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_∞ with vertex set $\{1, 2, 3, \ldots\}$. 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 $n_1 < n_2 < n_3 < \ldots$. 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 $\delta(G) \leq$ 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 $\sum_{v \in G} d(v) = 2e(G)$.

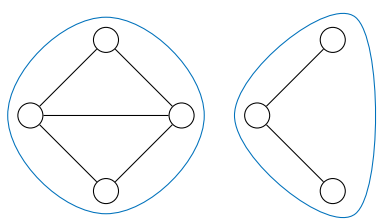
Hence also the average degree of G is $\frac{2e(G)}{|G|}$.

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

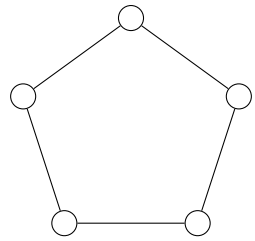
We say this path has length l and goes from v_0 to v_l. If $v, w \in G$ we write $v \to w$ to mean there is a path from v to w.

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

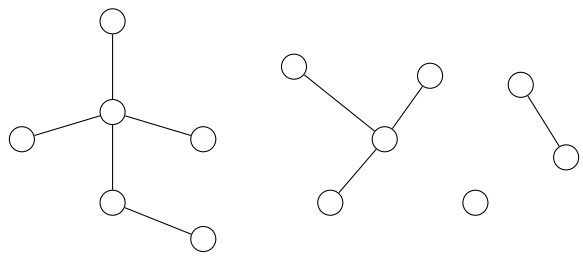
Definition (Components)**.** The equivalence classes of $\to$ are called the **components** of G. If G has only one component, we say G is **connected**.



Definition (Cycle)**.** A **cycle** is a sequence $v_0, v_1, \ldots, v_l$ of vertices of G with $v_0, \ldots, v_{l-1}$ 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.

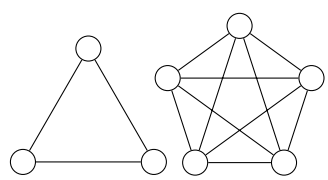


Definition (Disjoint union)**.** Suppose G, H are graphs with $V(G) \cap V(H) = \varnothing$. The **disjoint union** of G, H is the graph $G \cup H$ with $V(G \cup H) = V(G) \cup V(H)$ and $E(G \cup H) = E(G) \cup E(H)$.

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

Example.

1. $K_3 \cup K_5$



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

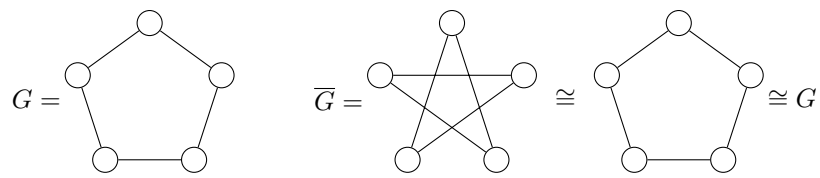
Definition (Induced subgraph)**.** Let $G = (V, E)$ be a graph, and let $W \subset V$. The **induced subgraph** on W is the graph $G[W]$ with $V(G[W]) = W$ and, for $x, y \in W$, $xy \in E(G[W]) \iff xy \in G$.

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 red/green. We can think about the 'red subgraph' with n vertices and red edges only. The complement is the green subgraph.

- Isomorphic to its complement:



Notation. Let $f, g : \mathbb{N} \to (0, \infty)$. 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 $f/g \to 0$ (as $n \to \infty$)
- $f = \omega(g)$ means $f/g \to \infty$
- $f \sim g$ means $f/g \to 1$.

2 Extremal Graph Theory

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

Let G be a 'red subgraph' of K_n, consisting of all vertices, and only the red edges. K_n has a red $K_s \iff K_s \subset G$, and K_n has a green $K_s \iff G$ has an induced $\overline{K_s}$ subgraph.

Rephrased: 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 (Extremal number)**.** Define

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

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

2.1.1 Triangles

Let $H = K_3$. The goal is to find G with $|G| = n$, $e(G)$ large and $\triangle \not\subset G$. The idea is to partition $V(G) = X \cup Y$ with all edges going from X to Y.

Definition (Bipartite graph)**.** 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 $\triangle$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 (Complete bipartite graph)**.** 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 $\triangle \not\subset G$. Then $G \cong K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}$.

Proof. Induction on n. Start with the $n = 3$ case: Consider $K_{2,1}$, it satisfies $|G| = 3$, $e(G) \geq 2$, $\triangle \subset G$, as required.

$n > 3$: Let $|G| = n$, $e(G) \geq \lfloor \frac{n^2}{4} \rfloor$, $\triangle \not\subset G$. First, remove edges from G if necessary to get H with $|H| = n$, $e(H) = \lfloor \frac{n^2}{4} \rfloor$. Clearly $\triangle \not\subset H$. Let $v \in H$ with $d(v) = \delta(H)$ and let $K = H - v$ (i.e. H with vertex v and all edges including v removed). Now, $|K| = n - 1$, $\triangle \not\subset K$ and $e(K) = \lfloor \frac{n^2}{4} \rfloor - \delta(H)$.

Suppose n is even. Then $\delta(H) \leq$ average degree of $H = \frac{2e(H)}{H} = \frac{n^2/2}{n} = \frac{n}{2}$. Hence

$$e(K) \geq \frac{n^2}{4} - \frac{n}{2} = \frac{n^2 - 2n}{4} = \frac{n^2 - 2n + 1}{4} - \frac{1}{4} = \frac{(n-1)^2}{4} - \frac{1}{4} = \left\lfloor \frac{(n-1)^2}{4} \right\rfloor$$

Similarly if n odd, also get $e(K) \geq \left\lfloor \frac{(n-1)^2}{4} \right\rfloor$.

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 triangle.

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 $\triangle$. But any new edge creates a $\triangle$, so $G \cong H$. $\square$

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

2.1.2 Complete graphs

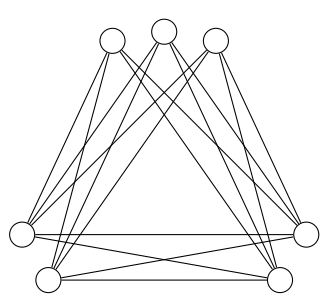
What about if we are trying to avoid a K_4? Try a complete tri-partite graph.

Definition (r-partite graph)**.** A graph G is **r-partite** if we can partition $V(G) = X_1 \cup \cdots \cup X_r$ 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 (Turán graph)**.** 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 $T_3(7)$.



Remark (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 the 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. Take first $n = r + 1$. $T_r(r+1)$ 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 $G \cong T_r(r+1)$.

Next consider $n > r + 1$. 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 property 4 earlier.

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 property 1 we can't add any edges. So, $G \cong T_r(n)$. $\square$

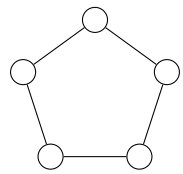
So if $r \geq 2$, $n \geq r + 1$, $\mathrm{ex}(n; K_{r+1}) = t_r(n)$. How big is $t_r(n)$ approximately? Each vertex is joined to a proportion approximately $1 - \frac{1}{r}$ of all vertices, so we have a proportion approx $1 - \frac{1}{r}$ of all possible edges. So $\mathrm{ex}(n; K_{r+1}) \approx (1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

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

2.1.3 Bipartite graphs

Can think of cycles and paths as graphs (ignore direction and start point).

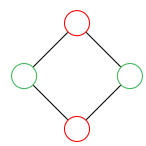
Definition (Cyclic graph)**.** The **cyclic graph of order** n, is the cycle of length n, called C_n.



Definition (Path graph)**.** The **path graph** of order n is the path of length n, called P_n.

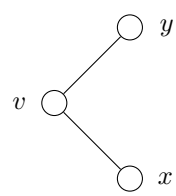
So $|C_n| = n$, $|P_n| = n + 1$, and $e(C_n) = e(P_n) = n$.

Example. What can we say about $\mathrm{ex}(n; C_4)$? 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_2s. So the total number of P_2s is $\sum_{v \in G} \binom{d(v)}{2}$.

On the other hand, 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_2s 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,

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

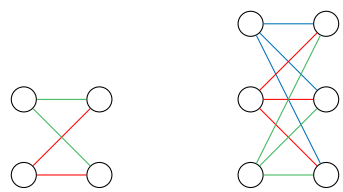
Apply Jensen to the earlier bound to give

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

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

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

Hence $m \leq \frac{n}{4}\left(1+\sqrt{4n-3}\right)$, so $\mathrm{ex}(n; C_4) \leq \frac{n}{4}\left(1+\sqrt{4n-3}\right)$, and so $\mathrm{ex}(n; C_4) = \mathcal{O}\left(n^{\frac{3}{2}}\right)$. A similar approach turns out to give a bound on $\mathrm{ex}(n; K_{t,t})$ in general:



Definition (t-fan)**.** 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 $\mathrm{ex}(n; K_{t,t}) = \mathcal{O}\left(n^{2-\frac{1}{t}}\right)$.

Proof. Let $|G| = n$, $e(G) = m = \mathrm{ex}(n; K_{t,t})$ 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

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

So $\frac{n^t}{t!}(t-1) \geq \frac{n}{t!}(\frac{2m}{n} - t)^t$, taking the highest factor from the LHS and lowest factor from the RHS so

$$n^t(t-1) \geq n\left(\frac{2m}{n} - t\right)^t.$$

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

$$\begin{aligned} n^t(t-1) &\geq n\left(\frac{m}{n}\right)^t \\ m^t &\leq n^{2t-1}(t-1) \\ m &\leq (t-1)^{\frac{1}{t}} n^{2-\frac{1}{t}} = \mathcal{O}(n^{2-\frac{1}{t}}). \end{aligned}$$

□

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 $m = \mathcal{O}(n^{2-\frac{1}{t}})$. □

2.1.4 General graphs

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

$$\lim_{n\to\infty} \frac{\mathrm{ex}(n;H)}{\binom{n}{2}}?$$

Does it even exist?

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

Proof. Let $n > |H|$. Let $|G| = n$, $e(G) = \mathrm{ex}(n;H) = x_n\binom{n}{2}$ and $H \not\subset G$. For any $v \in G$, $|G - v| = n - 1$ and $H \not\subset G - v$, so

$$e(G - v) \leq \mathrm{ex}(n-1;H) = x_{n-1}\binom{n-1}{2}.$$

Each edge $xy \in E(G)$ is 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. □

Definition (Asymptotic extremal number)**.** Write

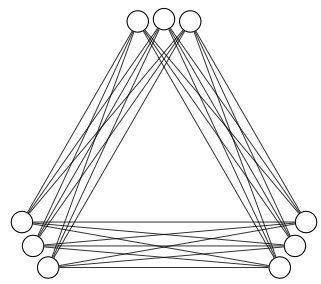
$$\mathrm{ex}(H) = \lim_{n\to\infty} \frac{\mathrm{ex}(n;H)}{\binom{n}{2}}$$

which exists by Proposition 13.

Example.

1. From Corollary 10, we have seen $\mathrm{ex}(K_{r+1}) = 1 - \frac{1}{r}$.
2. From Theorem 11, $\mathrm{ex}(n, K_{t,t}) = \mathcal{O}(n^{2-\frac{1}{t}}) = o(n^2) = o(\binom{n}{2})$, so $\mathrm{ex}(K_{t,t}) = 0$.

Fact. Turán's Theorem, and in particular Corollary 10 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. For example, with a $\frac{1}{2}$ the possible edges, we can guarantee a $\triangle$, but with proportion $\frac{1}{2}+\epsilon$ we can (eventually) guarantee a 'blown up triangle':



Definition (Complete r-partite graph)**.** 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 $\forall n \geq n_0$,

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

Proof. Coming up in Section 2.1.5, non-examinable. □

This is enough to determine $\mathrm{ex}(H)$ 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. $\chi(K_r(t)) = r$.

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

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

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 $n \geq n_0$,

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

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

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

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

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

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

So if H not bipartite, $\mathrm{ex}(H) = 1 - \frac{1}{\chi(H)-1} \neq 0$, and $\mathrm{ex}(n; H) \sim (1 - \frac{1}{\chi(H)-1})\binom{n}{2}$. This essentially solves the forbidden subgraph problem for non-bipartite H.

What about if H bipartite? Corollary 15 just says $\mathrm{ex}(H) = 0$, i.e. $\mathrm{ex}(n; H) = o(n^2)$, but doesn't tell us for instance that $\mathrm{ex}(n; H)$ grows like a certain power of n. We do however have upper bounds from Theorem 11: $\mathrm{ex}(n, K_{t,t}) = \mathcal{O}(n^{2-\frac{1}{t}})$.

Another application of Erdős-Stone:

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.

- $D(T_r(n)) \to 1 - \frac{1}{r}$ as $n \to \infty$
- If $\alpha > \mathrm{ex}(H)$ 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 $\frac{\infty}{\infty}$. Instead try something like the 'limit of densities of large finite subgraphs'.

Definition (Upper density)**.** The upper density of an infinite graph G is

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

Remark.

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

Corollary 16. For any infinite graph G,

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

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

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

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

2.1.5 Proof of Erdős-Stone (non-examinable)

We now return to prove the Erdős-Stone Theorem, which 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 $\exists n_0$ such that $\forall n \geq n_0$,

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

Proof. Induction on r. Fix T such that

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

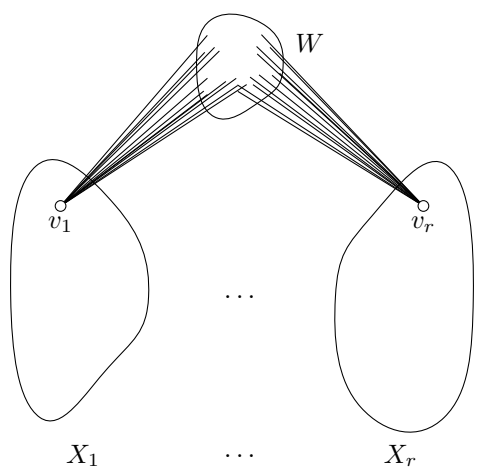
Then choose n_0 such that for all $n \geq n_0$,

$$|G| = n, \quad \delta(G) \geq \left(1 - \frac{1}{r} + \epsilon\right) n \implies K_r(T) \subset G.$$

(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 $\delta(G) \geq (1 - \frac{1}{r} + \epsilon)n$, $K_{r+1}(t) \not\subset G$. Pick such an n, G with $n \geq n_0$ and also $n \geq \frac{2t}{r\epsilon}$.

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



Let

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

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

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

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 X_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

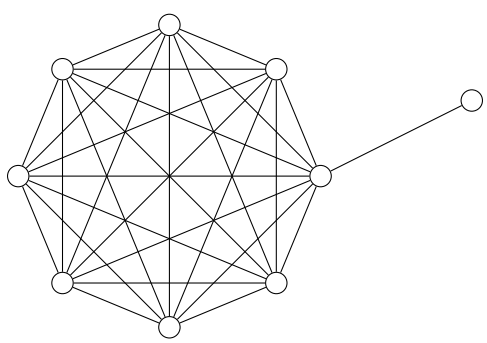
$$T^r \left(\frac{r\epsilon n}{2}\right)^t \leq n^t(t-1)T^{r-1}$$

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

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.



K_{n-1}

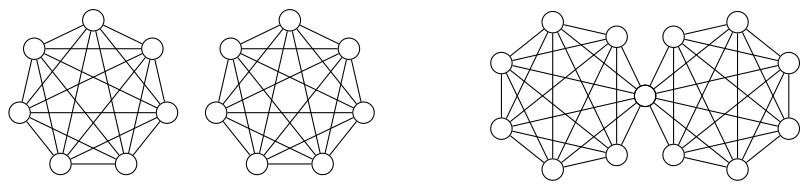
More interestingly then, is there a condition on $\delta(G)$?

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

Remark.

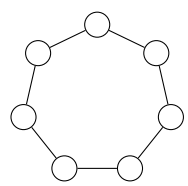
1. This result is best possible:

n even, $\delta(G) = \frac{n}{2} - 1$ $\qquad$ n odd, $\delta(G) = \frac{n-1}{2}$



$K_{\frac{n}{2}}$ $\qquad$ $K_{\frac{n}{2}}$ $\qquad$ $K_{\frac{n-1}{2}}$ $\qquad$ $K_{\frac{n-1}{2}}$

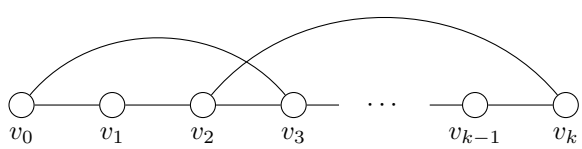
2. Of course there are Hamiltonian graphs of smaller minimum degree:



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 $x \neq y$, $x \nsim y$, then $|\Gamma(x) \cup \Gamma(y)| \leq n-2$, but $|\Gamma(x)| + |\Gamma(y)| \geq \frac{n}{2} + \frac{n}{2} = n$ so $\Gamma(x) \cap \Gamma(y) \neq \varnothing$.

Let $v_0, v_1, \ldots, v_k$ be a path of maximal length in G, say length $k \leq n-1$. By maximality, $\Gamma(v_0) \subset \{v_1, \ldots, v_k\}$. Similarly, $\Gamma(v_k) \subset \{v_0, \ldots, v_{k-1}\}$.



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 $A \cap B \neq \varnothing \implies$ 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 $C = v_0 v_1 \ldots v_{i-1} v_k v_{k-1} \ldots v_i v_0$ of length $k+1$.

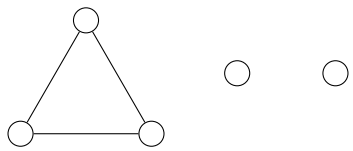
If $k = n-1$, we have a Hamiltonian cycle as required. If $k < n-1$, relabel the cycle $C = u_0 u_1 \ldots u_k u_0$. By connectedness, we have some $u_i \in C$ and $w \notin C$ with $w \sim u_i$. Then $w u_i u_{i+1} \ldots u_k u_0 \ldots u_{i-1}$ is a path of length $k+1$, contradicting maximality. $\square$

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 Konigsberg bridge problem seems superficially similar, but in fact is much less interesting.

Definition (Euler circuit)**.** A **circuit** of a graph G is a sequence $v_0v_1 \dots v_n$ 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**.

We are only interested in connected graphs: We could have a non-connected Eulerian graph, but only because of the degree zero vertices.



Proposition 18. Let G be a connected graph. Then

$$G \text{ Eulerian if and only if } \forall v \in G,\ d(v) \text{ is even.}$$

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

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

Consider $e(G) > 0$. Let $v_0v_1 \dots v_k = C$ 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

Think about the map-colouring problem. How many colours do we need to colour graphs with no crossings, such that adjacent vertices have different colours.

Definition (Colouring)**.** A k**-colouring** of a graph G is a function $c : V(G) \to [k]$. In proofs we often say 'red', 'green' for 1,2, etc.

Note G has a k-colouring if and only if G is k-partite. So, $\chi(G)$ is the least number of colours needed to colour G.

3.1 Planar Graphs

Definition (Graph drawing)**.** A **drawing** of G is an ordered pair (f, γ) where $f : V \to \mathbb{R}^2$ is an injection and $\gamma : E \to C([0,1], \mathbb{R}^2)$ 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 $\gamma(e)((0,1)) \cap \gamma(e')((0,1)) = \varnothing$.

(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,

$$\begin{aligned} \text{vertices} &\longleftrightarrow \text{points} \\ \text{edges} &\longleftrightarrow \text{continuous curves between end vertices,} \end{aligned}$$

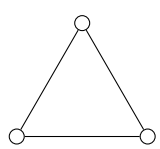
with no unnecessary intersections. If G has a drawing, we say G is **planar**.

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

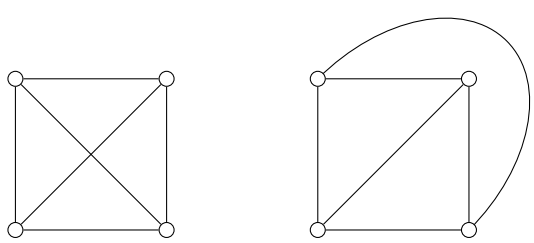
Fact. Any planar graph can be drawn with piecewise linear edges - i.e. each edge is finitely many line segments. Henceforth, assume all drawings satisfy this.

Example.

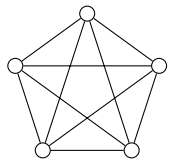
1. K_3 is planar



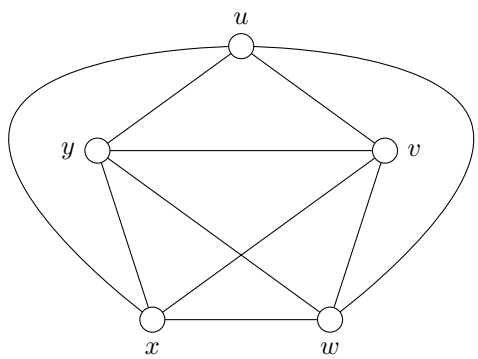
2. K_4 may appear non-planar but it is in fact planar.



3. K_5 is not planar.



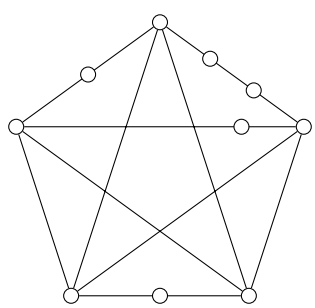
Suppose we have a drawing of K_5. It must contain a closed curve, say through vertices u, v, w, x, y in that order. The curve divides the plane into inside and outside.



Suppose edge vy is inside. Then ux, uw are outside. So yu, vx inside, so they intersect, a contradiction. Similarly, we can show 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, e.g.



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 (Leaf)**.** 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 in T. Now $v_k \sim v_{k-1}$, but v_k has no other neighbours in the path (as T acyclic) and v_k has no neighbours outside path (by maximality). Hence v_k is a leaf. □

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 $e(T) = n - 1$. □

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. □

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

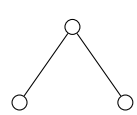
Proof. Induction on m. If G is a tree: $m = n - 1$, $l = 1$, and we are done.

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 $n - m + l = 2$. □

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).

We have a special case where

$G =$ 

with $n = 3$, $m = 2$ and $3n - 6 = 3 \geq 2$.

Otherwise, we know $n - m + l = 2$ and each face has at least 3 edges in its boundary, and each edge is in the boundary of at most 2 faces. So, $l \leq \frac{2m}{3}$. Thus, $n - m + \frac{2}{3}m \geq 2$ so $\frac{1}{3}m \geq n - 2$ so $m \geq 3n - 6$. □

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. □

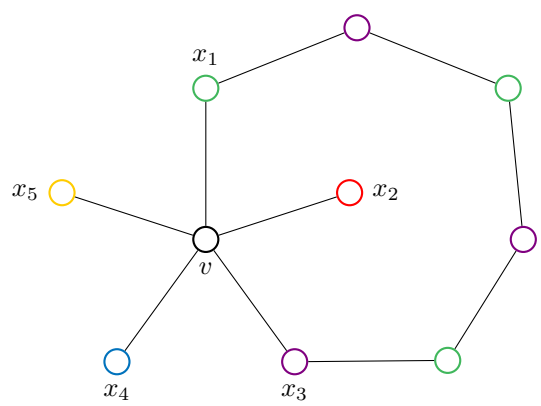
A lot more work gives:

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 $n \leq 5$ 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 $x_1, \ldots, x_5$ in clockwise order around v with colours $1, 2, 3, 4, 5$ 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 ij**-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_1 and give v colour 1.



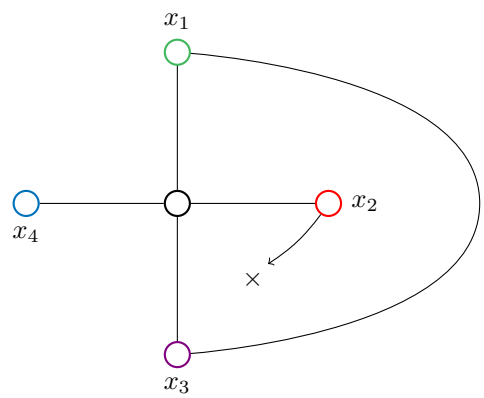
If not, then x_2, x_4 are in different 24 components as we see in the diagram, so swap colours $2, 4$ on the 24 component of x_2 and give v colour 2. □

Theorem 27 (Four colour theorem)**.** Any planar graph is 4-colourable.

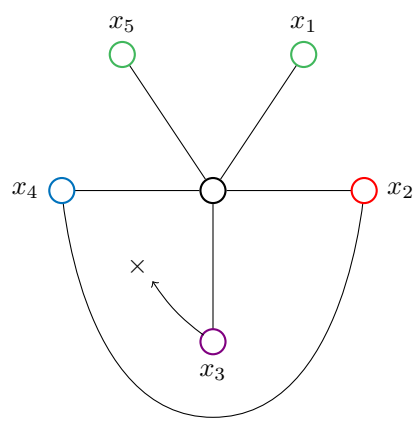
Proof. Let G be planar, $|G| = n$ and use induction on n. For $n = 4$, we are immediately done.

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.

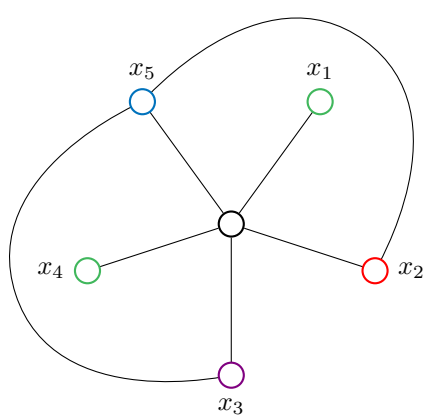


Case 2: $d(v) = 5$, v has neighbours $x_1, \ldots, x_5$ clockwise with colours $1, 2, 3, 4, 1$.



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: $d(v) = 5$, colours $1, 2, 3, 1, 4$.



If there is no 24 path from x_2 to x_5, we are done. If there is no 34 path from x_3 to x_5, we are done.

Otherwise, swap colours $1, 3$ on the 13 component of x_1 and swap colours $1, 2$ on the 12 component of x_4. Then, use colour 1 at v, but this is false. $\square$

The Theorem 27 was proved correctly in 1976, with 1936 cases checked.

3.2 General Graphs

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

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

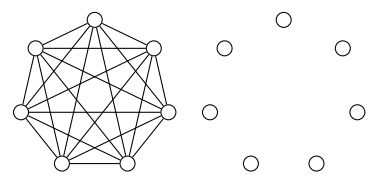
Lower bounds

(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 $\chi(G) = k$, $\frac{|G|}{\alpha(G)} = \frac{2k}{k+1} \approx 2$.



Simple upper bound

List the vertices in some order, colour each in turn using the least colour not already used on its neighbours. This is often called a '**greedy algorithm**'.

When we colour v, there are at most $d(v) \leq \Delta(G)$ colours unavailable. Thus

$$\chi(G) \leq \Delta(G) + 1 \tag{$*$}$$

But e.g. if $G = K_{1,t}$ then $\chi(G)$ then $\chi(G) = 2$ but $\Delta(G) = t$.

For which graphs G is $(*)$ achieved?

- G complete: $\chi(K_k) = k$, $\Delta(K_k) = k - 1$.
- G an odd cycle, e.g. a pentagon: $\chi(G) = 3$, $\Delta(G) = 2$.

Theorem 28 (Brooks' theorem)**.** Let G be a connected graph that is neither complete nor an odd cycle. Then $\chi(G) \leq \Delta(G)$.

Proof. Induction on $|G|$. Write $\Delta = \Delta(G)$. We cannot have $\Delta = 0, 1$ as $G \ncong K_1, K_2$. If $\Delta = 2$ then G is a path or an even cycle, so $\chi(G) = 2$. So assume $\Delta \geq 3$. Pick $v \in G$ and let H be a component of $G - v$.

Either $\Delta(H) < \Delta$, in which case by greedy algorithm bound $(*)$ we have

$$\chi(H) \leq \Delta(H) + 1 \leq \Delta.$$

Or $\Delta(H) = \Delta$. Then H is connected and not an odd cycle (as $\Delta \geq 3$). Moreover, $\exists u \in H$ with $u \sim v$ in G. In G, $d(u) \leq \Delta$ so in H, $d(u) \leq \Delta - 1$. So H is not regular, so not complete. Hence by induction hypothesis, $\chi(H) \leq \Delta$.

Do this for each component of $G - v$ to obtain a Δ-colouring c of $G - v$. If there is a colour missing from $\Gamma(v)$, then use that colour at v, as required. So assume that is not the case.

So have $\Gamma(v) = \{x_1, ..., x_\Delta\}$ with $\forall i, c(x_i) = i$. We can also assume:

(i) if $i \neq j$ then there is an ij-path P_{ij} from x_i to x_j,

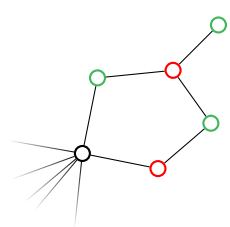
(ii) if $i \neq j$ then P_{ij} is entire ij-component containing x_i, x_j, and

(iii) if i, j, k distinct then P_{ij}, P_{ik} meet only at x_i.

(Why? If any of these fail then it is easy to check that the colouring c can be modified to change the colour of x_i and allowing us to use colour i at v).

(i) is as in previous proofs.

(ii) if P_{ij} is not the entire ij-component, then at some point in the path we get a point (wlog colour j) which 'branches off'.



Then this vertex has 3 green neighbours, so there are at most $\Delta - 3$ colours used in the rest of its neighbours, so there is a colour not in the colours of its neighbours$\cup\{i, j\}$. We can swap it to that, breaking the chain P_{ij}.

(iii) if P_{ij}, P_{ik} meet at some point which must have colour i, then that point has 2 j neighbours and 2 k neighbours, so as before there is a colour other than i that none of its neighbours have, which we can swap it to, breaking P_{ij} and P_{ik}.

As $G \ncong K_{\Delta+1}$, there are some i, j with $i \neq j$, $x_i \nsim x_j$. As $\Delta \geq 3$, pick $k \in [\Delta] \setminus \{i, j\}$. Let u be the neighbour of x_i of colour j. Swap colours i, k on the ik-component of x_i (i.e. on P_{ik}). This gives a new colouring c', with $c'(x_i) = k$, $c'(x_k) = i$.

Also, if $w \in P_{ij}$ with $w \neq x_i$ then $c'(w) = c(w)$ so $c'(x_j) = c'(u) = j$. c' must satisfy conditions (i),(ii),(iii) as before. Then by (i) there is a kj-path from x_i to x_j , P'_{kj} . By (ii), $u \in P'_{kj}$.

By (i), there is a ji-path P'_{ji} from x_j to x_k. By (ii), $u \in P'_{ji}$. But now P'_{kj} and P'_{ji} meet at u, contradicting (iii). $\square$

3.3 Graphs on surfaces

We've seen if G can be drawn in the plane then $\chi(G) \leq 4$. What about other surfaces?

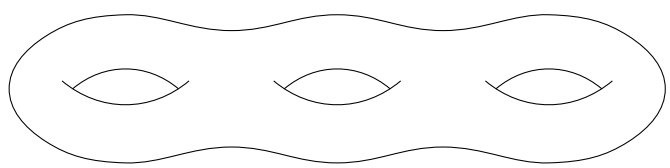
On the sphere, it is just the same. But, we can draw K_5 on the torus (missing picture), and $\chi(K_5) = 5$.

There is a classification theorem for surfaces. The closed surfaces without boundary are:

- For $g \geq 0$, the **orientable surface of genus** g, the torus with g holes called T_g.
- For $g \geq 1$, the **non-orientable surface of genus** g, the sphere with g discs cut out and g Möbius strips sewn in called S_g.

Example.

- T_0 is the sphere, T_1 is the torus
- T_3 is shown here



- S_1 is the real projective plane
- S_2 is the Klein bottle

Definition (Chromatic number of surface)**.** Given a surface S, the **chromatic number of** S is

$$\chi(S) = \max\{\, \chi(G) \mid G \text{ can be drawn on } S \,\}$$

The Four colour theorem says that $\chi(T_0) = 4$. Also $\chi(T_1) = 5$. We use the **Euler-Poincaré Formula**: If G can be drawn on S with l faces, $|G| = n$, $e(G) = m$ then $n - m + l \geq E$, where E is the Euler characteristic of the surface given by

$$S = T_g \implies E = 2 - 2g$$
$$S = S_g \implies E = 2 - g$$

Theorem 29 (Heawood's Theorem)**.** Let S be a closed boundaryless surface of Euler characteristic $E \leq 1$. Then

$$\chi(S) \leq \left\lfloor \frac{7 + \sqrt{49 - 24E}}{2} \right\rfloor$$

Proof. Let $\chi = \chi(S)$. Let G be a minimal χ-colourable graph that can be drawn on S - i.e. $\chi(G) = \chi$ but

$$H \subset G, H \neq G \implies \chi(G) \leq \chi - 1.$$

Clearly G is connected and $|G| \geq \chi$.

Let $|G| = n \geq \chi$, $e(G) = m$ and draw G on S with l faces. Then by Euler-Poincaré, $n - m + l \geq E$. As before, $l \leq \frac{2}{3}m$ so $n - \frac{1}{3}m \geq E$ so $m \leq 3n - 3E$. Hence

$$\delta(G) \leq \frac{2m}{n} \leq \frac{6n - 6E}{n} = 6 - \frac{6E}{n}. \qquad (*)$$

On the other hand, if $v \in G$ then $G - v$ is $(\chi - 1)$-colourable and this colouring does not extend to a $(\chi - 1)$-colouring of G so $d(v) \geq \chi - 1$. Hence $\delta(G) \geq \chi - 1$.

Combining this with $(*)$: If $E \leq 0$: $\chi - 1 \leq \delta(G) \leq 6 - \frac{6E}{n} \leq 6 - \frac{6E}{\chi}$ as $n \geq \chi$. Hence $\chi^2 - 7\chi + 6E \leq 0$ and so $\chi \leq \frac{7+\sqrt{49+24E}}{2}$.

If $E = 1$, $\delta(G) \leq 6 - \frac{6}{n} < 6$ so $\delta(G) \leq 5$. Hence $\chi - 1 = 5$ so $\chi \leq 6 = \frac{7+\sqrt{49+24}}{2}$. □

Remark. The condition $E \leq 1$ takes out only the sphere. The theorem is true for the sphere (Four colour theorem), but this proof doesn't work if $E = 2$.

Example. Suppose $E = 0$, so $\chi(S) \leq 7$. There are 2 possible surfaces:

(i) The torus T_1. We've seen K_5 can be drawn on the torus. In fact, a K_7 can also be drawn on the torus, and so $\chi(T_1) = 7$.

(ii) The Klein bottle S_2. It turns out K_6 can be drawn on the Klein bottle, but not K_7. So $6 \leq \chi(S_2) \leq 7$.

Can we find G on Klein bottle with $\chi(G) = 7$? Suppose we can. Let G be a minimal 7-chromatic graph on S_2.

Examine the proof of Heawood's Theorem. First, G is connected. Also $\delta(G) \geq 6$. Then $6 \leq \delta(G) \leq \frac{2m}{n} \leq 6$. So equality holds throughout. Thus $\delta(G) = 6$ and average degree 6, so G is 6-regular. In particular, G is connected, $\Delta(G) = 6$ and $\chi(G) = 7$. By Brooks' theorem, $G \cong K_7$, a contradiction. So, $\chi(S_2) = 6$.

In fact the bound from Heawood's Theorem is tight for all surfaces except the Klein bottle. That is if $S \neq S_2$,

$$\chi(S) \leq \left\lfloor \frac{7 + \sqrt{49 - 24E}}{2} \right\rfloor = \chi.$$

Why? It turns out that K_χ can be drawn on S (hard).

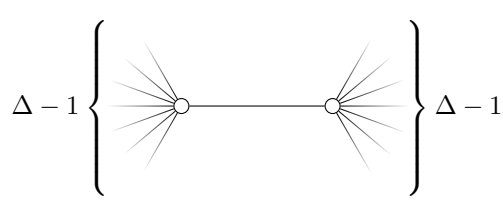
3.4 Edge Colouring

Definition (Edge colouring)**.** A k-edge colouring of a graph $G = (V, E)$ is a function $\varphi : E \to [k]$ such that if $e, e' \in E$ with precisely one common vertex then $\varphi(e) \neq \varphi(e')$.

Definition (Edge chromatic number)**.** The **edge-chromatic number** of G is

$$\chi'(G) = \min \{ k \mid G \text{ has a } k\text{-edge colouring} \}.$$

Clearly $\Delta(G) \leq \chi'(G) \leq 2\Delta(G) - 1$, where the lower bound comes from a greedy algorithm: the worst case looks like



where we need to colour the central edge. In fact, we can tighten the upper bound:

Theorem 30 (Vizing's theorem)**.** Let G be a graph. Then

$$\chi'(G) \leq \Delta(G) + 1$$

Proof. Induction on $e(G)$. The $e(G) = 0$ case is immediate.

For $e(G) > 0$, let $\Delta = \Delta(G)$. Pick an edge xy. By the induction hypothesis we can find an $(\Delta + 1)$-edge colouring of $G - xy_1$, call it φ.

As $\Delta(G) < \Delta + 1$, there some colour 'missing' at each vertex. Let c_1 be missing at y_1. If c_1 is also missing at x, then colour xy_1 with colour c, so done.

If not, let $y_2 \in \Gamma(x)$ with $\varphi(xy_2) = c_1$ and let c_2 be missing at y_2. Continue inductively $(*)$: Given distinct $y_1, \ldots, y_n \in \Gamma(x)$, distinct colours $c_1, \ldots, c_k$ such that c_i is missing at y_i (for $1 \leq i \leq k$) and $\varphi(xy_{i+1}) = c_i$ (for $1 \leq i \leq k-1$). If c_k is missing at x, re-colour xy_i in colour c_i. Otherwise, let $y_{k+1} \in \Gamma(x)$ with $\varphi(xy_{k+1}) = c_k$. Let c_{k+1} be missing at y_k. If $c_{k+1} \notin \{c_1, \ldots, c_k\}$, continue as at $(*)$.

Otherwise assume WLOG $c_{k+1} = c_1$. (If instead $c_{k+1} = c_j$ for some $j > 1$, then uncolour xy_j, recolour xy_i in colour c_i for $1 \leq i < j$ and relabel $y_j, y_{j+1}, \ldots$ as $y_1, y_2, \ldots$). Let c be a colour missing at x. Consider the cc_1 subgraph of G, i.e. only edges coloured c or c_1. This subgraph has maximum degree ≤ 2 so each component is a path or a cycle. Moreover, x, y_1, y_{k+1} have degree ≤ 1 in this subgraph so x_1, y_1, y_{k+1} are not all in the same component.

If x, y_1 in different cc_1 component, then swap c, c_1 on the component of y_1 and give xy_1 colour i. Otherwise xy_{k+1} in different components. In this case, uncolour xy_{k+1} and recolour xy_i with colour i for $1 \leq i \leq k$. Then swap colours c, c_1 on the cc_1-component of y_{k+1} and give xy_{k+1} colour c. $\square$

4 Connectivity

4.1 The Marriage Problem

Legal note: This problem was first posed in 1935. Note 1935 ¡ 2014, so marriage has 1 man and 1 woman.

Suppose we have n women and n men. For each woman, we have a list of 'suitable' husbands. Can we marry everyone off? Reformulate this as a bipartite graph: X is the set of women, Y is the set of men, and an edge denotes suitability.

Definition (Matching)**.** Let G be a bipartite graph with bipartition (X, Y). A **matching** from X to Y is a set $M \subset E(G)$ such that $\forall x \in X$, $\exists$ unique $e \in M$ with $x \in e$ and for all $y \in Y$ there is at most one $e \in M$ with $y \in e$.

When can we find a matching? Obviously need $\forall x \in X$, $|\Gamma(x)| \geq 1$. Also need if $x, y \in X$ distinct then $|\Gamma(x) \cup \Gamma(y)| \geq 2$. In general, if $A \subset X$ then we need

$$|\Gamma(A)| := \left| \bigcup_{x \in A} \Gamma(x) \right| \geq |A|.$$

Surprisingly, this is enough.

Theorem 31 (Hall's Marriage Theorem)**.** Let G be a bipartite graph with bipartition (X, Y). Then G has a matching from X to Y iff G satisfies **Hall's condition**:

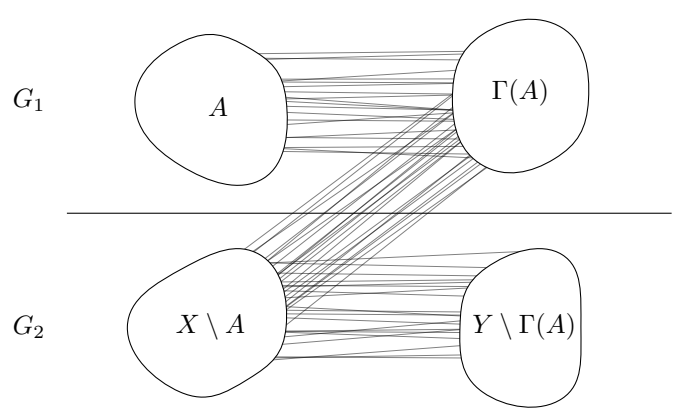
$$\forall A \subset X, |\Gamma(A)| \geq |A|$$

Proof. ($\Rightarrow$) is obvious. For ($\Leftarrow$), use induction on $|X|$. $|X| = 0, 1$ are immediate.

For $|X| \geq 2$, there are two cases. The easy case has $|\Gamma(A)| > |A|$ for all $A \subset X$ with $A \neq \varnothing, X$. Pick any $x \in X$. $|\Gamma(x)| \geq 1$ so pick $y \in \Gamma(x)$. Look at $G - \{x, y\}$; this graph satisfies Hall's condition, so has a matching from $X - \{x\}$ to $Y - \{y\}$. Add edge xy, and done.

The harder case has $|\Gamma(A)| = |A|$ for some $A \neq 0, X$. Let

$$G_1 = G[A \cup \Gamma(A)]$$
$$G_2 = G[(X \setminus A) \cup (Y \setminus \Gamma(A))].$$



First, G_1 obviously satisfies Hall's condition so $\exists$ a matching A to $\Gamma(A)$. What about G_2? Take $B \subset X \setminus A$. Then

$$|\Gamma_{G_2}(B)| = |\Gamma(A \cup B) \setminus \Gamma(A)| = |\Gamma(A \cup B)| - |\Gamma(A)| \geq |A \cup B| - |A| = |B|.$$

Hence G_2 also satisfies Hall and has a matching from $X \setminus A$ to $Y \setminus \Gamma(A)$. Combine the two matchings to get a matching from X to Y in G. □

Definition (Independent set)**.** Let $G = (V, E)$ be a graph. A set $F \subset E$ is **independent** if no two edges of F share a vertex.

Note: A matching from X to Y is just a set of $|X|$ independent edges.

Corollary 32 (Defect Hall)**.** Let G be a bipartite graph with bipartition (X, Y) and let $d \geq 1$. Then G contains $|X| - d$ independent edges if and only if $\forall A \subset X,\ |\Gamma(A)| \geq |A| - d$.

Proof. ($\Rightarrow$) is immediate. ($\Leftarrow$). In marriage terminology: Introduce d imaginary perfect men, suitable husbands for all the women. This satisfies Hall's condition so has matching from X to Y. In real life, at most d women unmarried. □

Corollary 33 (Polyandrous Hall)**.** Let G be a bipartite graph, bipartition (X, Y), $d \geq 2$. Then G contains a set of $d|X|$ edges, each vertex in X in precisely d of them, each vertex in Y in at most one $\iff \forall A \subset X, |\Gamma(A)| \geq d|A|$.

Proof. ($\Rightarrow$) is immediate. ($\Leftarrow$). In marriage terminology: clone each woman $d - 1$ times so there are d copies of each. This satisfies Hall's condition so have matching from X to Y. Destroy the clones. □

Further extensions are possible, e.g.

(i) (Variable polyandrous) Take women $x_1, \ldots, x_n$ and numbers $d_1, \ldots, d_n \geq 1$. Give woman x_i d_i husbands.

(ii) (Defect polyandrous) Aim to each each woman d husbands but allow d' 'missing' husbands.

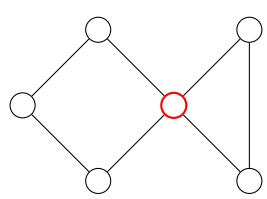
(iii) (Same sex) (Harder) If G is a graph, when can we find a 1-factor in G: a set of $\frac{|G|}{2}$ independent edges.

4.2 Connectivity

Definition (k-connectivity)**.** Let $k \geq 1$. We say a graph G is **k-connected** if whenever $W \subset V(G)$ with $|W| < k$ then $G - W$ is connected.

(Assume for now that G is not a complete graph).

Example. G is 1-connected $\iff$ G is connected. G is 2-connected $\iff$ G is connected and G has no 'cut vertex'.

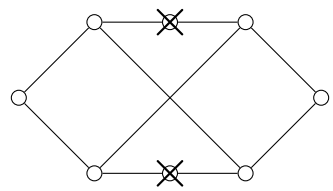


The highlighted vertex is a cut vertex, and the graph is not 2-connected.

Definition (Independent paths)**.** Let G be a graph and $a, b \in V$ be distinct. A collection of paths from a to b is **independent** if the paths meet only at a and b.

Suppose for all distinct $a, b \in G$ there are k independent paths from a to b. Then if $W \subset V(G)$, $|W| < k$ then $G - W$ is still connected. So G is k-connected.

What about a converse? If G is k-connected and we can only find $k-1$ independent ab-paths for some $a, b \in G$ then we would like to say 'delete' one vertex from each path to disconnect the graph, reaching a contradiction. But this is false, e.g.



In fact, the converse is true, but not obvious.

Definition (AB-path)**.** Let G be a graph and $A, B \subset V(G)$. An AB**-path** is a path that meets A in its first vertex and nowhere else, and meets B in its last vertex and nowhere else. A set $W \subset V(G)$ is an AB-separator if $G - W$ contains no AB-path.

Remark.

- A is an AB-separator. Also, B is an AB-separator.
- Suppose G is k-connected. Then as long as $|A|, |B| \geq k$ we have that any AB-separator W satisfies $|W| \geq k$.
- Note that we do not suppose $A \cap B = \varnothing$. If $v \in A \cap B$ then v is an AB-path (of length zero). Thus for any AB-separator W, $A \cap B \subset W$.

Theorem 34. Let G be a graph and $A, B \subset V(G)$. Let

$$k = \min \{ |W| \mid W \text{ is an } AB\text{-separator} \}.$$

Then G contains k vertex-disjoint AB-paths.

Note **vertex-disjoint** means the paths don't even meet inside A or B.

Proof. Coming soon. □

Corollary 35 (Menger's Theorem)**.** Let G be an incomplete k-connected graph and let $a, b \in V(G)$, $a \neq b$. Then G contains k independent ab-paths.

Proof. Suppose first $a \nsim b$. Let $A = \Gamma(a)$ and $B = \Gamma(b)$. We have G k-connected and $G - A, G - B$ disconnected, so $|A|, |B| \geq k$.

Hence, any AB-separator W has $|W| \geq k$, so by Theorem 34, there are k vertex-disjoint AB paths. Extend these to a,b.

If instead $a \sim b$: $G - ab$ is $(k-1)$-connected, so has $k-1$ independent ab paths by first part. ab is another, as required. □

Definition (Connectivity)**.** If G is an incomplete graph, the **connectivity** of G is

$$\kappa(G) \coloneqq \max\left(\{\, k \geq 1 \mid G \text{ is } k\text{-connected} \,\} \cup \{0\}\right).$$

In light of Corollary 35, define $\kappa(K_n) = n - 1$ for $n \geq 2$.

We now return to prove Theorem 34: that if every AB-separator has order $\geq k$ then there are k vertex-disjoint AB-paths.

Proof of Theorem 34. Induction on $e(G)$. If $e(G) = 0$, the smallest AB-separator is $A \cap B$, and each vertex of $A \cap B$ gives an AB-path (of length zero), so done.

If $e(G) > 0$, pick $xy \in E(G)$. Let W be an AB-separator of minimum order in $G - xy$. If $|W| \geq k$ then by the induction hypothesis there are k vertex-disjoint AB-paths in $G - xy$ and so also in G.

So assume $|W| < k$. Then $W \cup \{x\}$ is an AB-separator in G. Hence $|W \cup \{x\}| \geq k$ so $|W| \geq k - 1$ so $|W| = k - 1$. Write $W = \{w_1, w_2, \ldots, w_{k-1}\}$. As $|W| < k$, $G - W$ contains an AB-path. This path must include the edge xy. Assume WLOG x comes before y in this path (if not, swap x and y).

Let $X = W \cup \{x\}$ and $Y = W \cup \{y\}$. **missing end of proof** □

In fact, Hall's Marriage Theorem is also a consequence of Theorem 34. Let G be a bipartite graph with bipartition (X, Y) such that

$$\forall A \subset X, |\Gamma(A)| \geq |A|.$$

Let W be an XY-separator. Then $\Gamma(X \setminus W) \subset Y \cap W$ so

$$\begin{aligned} |W| &= |X \cap W| + |Y \cap W| \\ &\geq |X \cap W| + |\Gamma(X \setminus W)| \\ &\geq |X \cap W| + |X \setminus W| = |X|. \end{aligned}$$

Hence by Theorem 34 we have $|X|$ vertex-disjoint XY-paths, i.e. a matching from X to Y.

4.3 Edge connectivity

Definition (l-edge connected)**.** Let G be a graph with $|G| \geq 2$ and let $l \geq 1$. We say G is **l-edge connected** if whenever $D \subset E(G)$ with $|D| < l$ we have $G - D$ connected. The **edge-connectivity** of G is

$$\lambda(G) \coloneqq \max\left(\{\, l \geq 1 \mid G \text{ is } l\text{-edge connected} \,\} \cup \{0\}\right).$$

Corollary 36 (Edge Menger)**.** Let G be l-edge connected and $a, b \in V(G)$ be distinct. Then G has l edge-disjoint ab-paths.

Proof. The **line graph** of $G = (V, E)$ is the graph $L(G) = (E, F)$ where

$$F = \{\, ee' \mid e, e' \in E,\ e, e' \text{ share precisely one vertex} \,\}.$$

Then $L(G)$ is l-connected. Add extra vertices a', b' to $L(G)$ with

- a' joined to each $e \in E$ with $a \in e$ and
- b' joined to each $e \in E$ with $b \in e$.

By Corollary 35, $L(G)$ has l independent $a'b'$-paths. This yields l edge-disjoint ab-paths in G. □

5 Probabilistic Techniques

5.1 The Probabilistic Method

Recall the Ramsey number $R(s)$. From Corollary 3, we have $R(s) = \mathcal{O}(4^s)$. What about lower bounds? On example sheet 1, we get

- $R(s) = \Omega(s^2)$ (simple colouring)
- $R(s) = \Omega(s^3)$ (much harder).

A similar but much harder still method gives $R(s) = \Omega(s^4)$. But this doesn't even give, say, $R(s) = \Omega(1.000001^s)$. Remarkably,

Theorem 37 (Erdős)**.**

$$R(s) = \Omega(\sqrt{2}^s)$$

Proof. Fix n, s. Colour each edge of K_n red/green at random, independently, each colour equally likely. Given $H \subset K_n$ with $H \cong K_s$,

$$\mathbb{P}(H \text{ monochromatic}) = 2\mathbb{P}(H \text{ red}) = 2 \times \left(\frac{1}{2}\right)^{\binom{s}{2}}.$$

So

$$\begin{aligned}
\mathbb{P}(K_n \text{ has a monochromatic } K_5) &\leq \binom{n}{s} \cdot 2 \cdot \left(\frac{1}{2}\right)^{\binom{s}{2}} \\
&\leq \frac{n^s}{s!} \cdot 2 \left(\frac{1}{2}\right)^{\binom{s}{2}} \\
&\leq n^s \cdot 2^{-\frac{s(s-1)}{2}} \\
&= \left(\frac{n}{\sqrt{2}^{s-1}}\right)^s < 1 \quad \text{if } n < \sqrt{2}^{s-1}.
\end{aligned}$$

So if $n < \sqrt{2}^{s-1}$ then there is *some* colouring with no monochromatic K_5. So $R(s) \geq \sqrt{2}^{s-1}$. $\square$

5.2 Modifying a Random Graph

In the proof of Theorem 37, we randomly coloured K_n red/green, showed that

$$\mathbb{P}(\text{no mono } K_s) > 0$$

and deduced there must be a colouring with no monochromatic K_s. In general, we do something random, show

$$\mathbb{P}(\text{thing we want happens}) > 0 \tag{$*$}$$

and deduce that sometimes the thing we want happens.

Sometimes it is too hard to prove $(*)$ directly and we have to work a bit harder. Instead, show $\mathbb{P}(\text{thing we want 'almost happens'}) > 0$. Take an example where the thing 'almost happens', and modify it somehow to get what we want.

Recall the Problem of Zarankiewicz. Theorem 12 gave an upper bound: if $t \geq 2$ then $z(n,t) = \mathcal{O}(n^{2-\frac{1}{t}})$. What about lower bounds?

Theorem 38. If $t \geq 2$ then $z(n,t) = \Omega(n^{2-\frac{2}{t+1}})$.

Proof. Strategy: Given n, p, let G be a random bipartite graph with vertex classes X, Y with $|X| = |Y| = n$, where for each $x \in X$, $y \in Y$ we have $xy \in E(G)$ with probability p, independently. Let $A = e(G)$ and B be the number of copies of $K_{t,t}$ in G (so A,B are random variables).

Aim: Given n, try to choose p such that $\mathbb{E}(A - B)$ is large, specifically $\mathbb{E}(A - B) = \Omega(n^{2-\frac{2}{t+1}})$. Then we can find a specific graph G with $A - B = \Omega(n^{2-\frac{2}{t+1}})$. Remove an edge from each $K_{t,t}$ in G to form a graph H with no $K_{t,t}$ and $e(H) = \Omega(n^{2-\frac{2}{t+1}})$, as required.

Now, $\mathbb{E}(A) = n^2 p$ and $\mathbb{E}(B) = \binom{n}{t}^2 p^{t^2} \leq \frac{1}{(t!)^2} n^{2t} p^{t^2}$. So $\mathbb{E}(A - B) \geq n^2 p - \frac{1}{t!^2} n^{2t} p^{t^2}$.

We want $n^2 p = n^{2t} p^{t^2}$ i.e. $p = n^{\frac{2-2t}{t^2-1}} = n^{-\frac{2}{t+1}}$. So, take $p = n^{-\frac{2}{t+1}}$. Then

$$\mathbb{E}(A - B) \geq \left(1 - \frac{1}{(t!)^2}\right) n^{2-\frac{2}{t+1}}.$$

□

Recall from Example sheet 2 that there is a $\triangle$-free graph of large chromatic number (a hard result). What if we also ban cycles of length 4? This is much harder still. In fact:

Theorem 39. Let $g \geq 3$, $k \geq 2$. Then there is a graph G with no cycles of length $\leq g$ and $\chi(G) \geq k$.

Proof. Strategy: Fix n and p. Let G be a random graph with n vertices, each possible edge present independently with probability p. Let X be the number of short cycles in G. Recall that $\chi(G) \geq \frac{|G|}{\alpha(G)}$ where $\alpha(G)$ is the independence number of G. For cycles, by 'short', we mean of length $\leq g$.

Aim: Pick n and p such that

1. $\mathbb{P}(X > \frac{n}{2}) < \frac{1}{2}$ and
2. $\mathbb{P}(\alpha(G) \geq \frac{n}{2k}) > \frac{1}{2}$.

Then $\mathbb{P}(X > \frac{n}{2}$ or $\alpha(G) \geq \frac{n}{2k}) < 1$ so can pick a specific G such that $X \leq \frac{n}{2}$ and $\alpha(G) < \frac{n}{2k}$. Remove a vertex from each short cycle to get H with $|H| \geq \frac{n}{2}$ and $\alpha(H) < \frac{n}{2k}$ so $\chi(H) > \frac{n/2}{n/(2k)} = k$, as required.

1. For $3 \leq i \leq g$, let X_i be the number of cycles of length i appearing as subgraphs of G. Then
$$\mathbb{E}(X_i) = \binom{n}{i} \frac{i!}{2i} p^i \leq (np)^i.$$
Now $X = \sum_{i=3}^{g} X_i$ so
$$\mathbb{E}(X) \leq \sum_{i=3}^{g} (np)^i < g(np)^g \text{ as long as } np \geq 1. \qquad (*)$$
By Markov,
$$\mathbb{P}\left(X > \frac{n}{2}\right) \leq \frac{\mathbb{E}(X)}{n/2} < 2gn^{g-1}p^g \leq \frac{1}{2}$$
if $p \leq (\frac{1}{4g})^{\frac{1}{g}} n^{\frac{1}{g}-1}$. Take $p = (\frac{1}{4g})^{\frac{1}{g}} n^{\frac{1}{g}-1}$. Then $np = \left(\frac{1}{4g}\right)^{\frac{1}{g}} n^{\frac{1}{g}} \geq 1$ if n sufficiently large, satisfying the condition of $(*)$.

2.

$$\begin{aligned}\mathbb{P}\left(\alpha(G) \geq \frac{n}{2k}\right) &\leq \binom{n}{\frac{n}{2k}}(1-p)^{\binom{n/2k}{2}} \\ &\leq n^{\frac{n}{2k}} e^{-p\frac{n^2}{16k^2}} \\ &= \exp\left\{\frac{2}{2k}\log n - \frac{n^2}{16k^2}\left(\frac{1}{4g}\right)^{\frac{1}{g}} n^{\frac{1}{g}-1}\right\} \to 0 \text{ as } n \to \infty.\end{aligned}$$

So if n is sufficiently large, $\mathbb{P}(\alpha(G) \geq \frac{n}{2k}) < \frac{1}{2}$. □

Properties of Expectation

We used:

1. **Linearity of expectation**.

$$\mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{E}(X_i).$$

For example, we had G bipartite, n vertices in each class and each possible edge present independently with probability p; B the number of $K_{t,t}$s in G. We said

$$\mathbb{E}B = \binom{n}{t}^2 p^{t^2}.$$

Why? Let $H_1, \ldots, H_N$ be the possible $K_{t,t}$s, where $N = \binom{n}{t}^2$. Let

$$X_i = \begin{cases} 1 & \text{if } H_i \text{ appears in } G \\ 0 & \text{if not} \end{cases}$$

so $\mathbb{E}X_i = \mathbb{P}(H_i \text{ appears}) = p^{t^2}$. Then $B = \sum_{i=1}^{N} X_i$ so

$$\mathbb{E}B = \sum_{i=1}^{N} \mathbb{E}X_i = N\mathbb{E}X_1 = \binom{n}{t}^2 p^{t^2}$$

as desired.

2. **Markov's inequality**. If X is a nonnegative random variable with finite mean and $\lambda > 0$ then

$$\mathbb{P}(X \geq \lambda) \leq \frac{\mathbb{E}X}{\lambda}.$$

Proof: $\lambda \mathbb{1}_{X \geq \lambda} \leq X$. Take expectations, giving $\lambda \mathbb{P}(X \geq \lambda) \leq \mathbb{E}X$.

3. **Chebyshev's inequality**. Let X be a random variable with finite mean μ and finite variance σ^2. Then

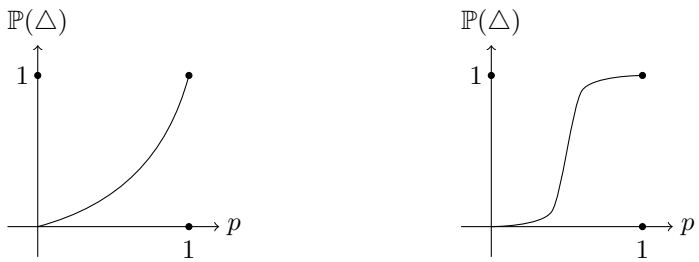
$$\forall \epsilon > 0 \quad \mathbb{P}(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}.$$

Proof: Apply Markov's inequality to the random variable $(X - \mu)^2$ with mean σ^2.

5.3 The Structure of Random Graphs

Given $n \geq 1$ and $p \in [0,1]$ the probability space $\mathcal{G}(n,p)$ consists of all graphs G with vertex set $[n] = \{1, 2, \ldots, n\}$ where for $i, j \in [n]$ distinct, $\mathbb{P}(ij \in E(G)) = p$, independently for different edges. What does a typical $G \in \mathcal{G}(n,p)$ look like?

For example, does G contain a $\triangle$? Yes usually if p large, no usually if p small. We might expect a curve as on the left, but in fact it more closely resembles the right.



That is, we get a 'sharp threshold', so to say.

First, consider an easier example: Does G contain an edge? Let $G \in \mathcal{G}(n,p)$ and let $X = e(G)$. Let $\mu = \mathbb{E}X = \binom{n}{2}p = Np$, writing $N = \binom{n}{2}$. By Markov,

$$\mathbb{P}(G \text{ has an edge}) = \mathbb{P}(X \geq 1) \leq \mathbb{E}X = Np.$$

So if $\mu = Np \to 0$ as $n \to \infty$, then $\mathbb{P}(G \text{ has an edge}) \to 0$. That is, if $p = o(\frac{1}{N}) = o(n^{-2})$ then $\mathbb{P}(G \text{ has an edge}) \to 0$.

What if p is bigger, say $p = \omega(n^{-2})$? Then $\mu = Np \to \infty$, and we need to think about variance. What is $\mathrm{Var}(X)$? Let $e_1, e_2, \ldots, e_N$ be the possible edges of G and let

$$X_i = \begin{cases} 1 & \text{if } e_i \text{ appears in } G \\ 0 & \text{if not} \end{cases}$$

then $X = \sum_{i=1}^{N} X_i$. Now $\mathbb{E}X_i = p$ and $\mathbb{E}X_i^2 = p$ so $\mathrm{Var}(X_i) = p - p^2$. The X_is are independent so

$$\mathrm{Var}(X) = \sum_{i=1}^{N} \mathrm{Var}(X_i) = N(p - p^2) =: \sigma^2.$$

By Chebyshev,

$$\begin{aligned} \mathbb{P}(X = 0) \leq \mathbb{P}(|X - \mu| \geq \mu) &\leq \frac{\sigma^2}{\mu^2} \\ &= \frac{N(p - p^2)}{(Np)^2} = \frac{1}{Np} - \frac{1}{N} \\ &= \frac{1}{\mu} - \frac{1}{N} \leq \frac{1}{\mu} \to 0 \text{ as } n \to \infty. \end{aligned}$$

since $\mu \to \infty$ as $n \to \infty$. So if $p = \omega(n^{-2})$ then $\mathbb{P}(X = 0) \to 0$ so $\mathbb{P}(X \geq 1) \to 1$ i.e. $\mathbb{P}(G \text{ has an edge}) \to 1$.

We have proved: $p = n^{-2}$ is a 'sharp threshold' for $G \in \mathcal{G}(n,p)$ in the sense that

- if $p = o(n^{-2})$ then almost every $G \in \mathcal{G}(n,p)$ has no edge, whereas

- if $p = \omega(n^{-2})$ then almost every $G \in \mathcal{G}(n,p)$ has an edge.

Here, we say 'almost every $G \in \mathcal{G}(n,p)$ has property Q' to mean

$$\mathbb{P}(G \in \mathcal{G}(n,p) \text{ has } Q) \to 1 \text{ as } n \to \infty.$$

(Note this is different from terminology usually used by probabilists.)

More generally, let $A_1, A_2, \ldots, A_n$ be events, X is the number of A_i that occur. Then $X = \sum_{i=1}^n X_i$ where

$$X_i = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & \text{if not} \end{cases}$$

so $\mathbb{E}X_i = \mathbb{P}(A_i)$. Hence by linearity of expectation,

$$\boxed{\mathbb{E}X = \sum_{i=1}^n \mathbb{P}(A_i)}$$

We know $\mathrm{Var}(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2$. Now,

$$\begin{aligned}
\mathbb{E}X^2 &= \mathbb{E}\left[\left(\sum_{i=1}^n X_i\right)^2\right] = \mathbb{E}\left[\left(\sum_{i=1}^n X_i\right)\left(\sum_{j=1}^n X_j\right)\right] \\
&= \mathbb{E}\sum_{i=1}^n\sum_{j=1}^n X_iX_j = \sum_{i=1}^n\sum_{j=1}^n \mathbb{E}(X_iX_j) \\
&= \sum_{i=1}^n\sum_{j=1}^n \mathbb{P}(A_i \cap A_j).
\end{aligned}$$

and

$$(\mathbb{E}X)^2 = \left(\sum_{i=1}^n \mathbb{P}(A_i)\right)^2 = \sum_{i=1}^n\sum_{j=1}^n \mathbb{P}(A_i)\mathbb{P}(A_j).$$

Hence

$$\mathrm{Var}(X) = \sum_{i=1}^n\sum_{j=1}^n (\mathbb{P}(A_i \cap A_j) - \mathbb{P}(A_i)\mathbb{P}(A_j))$$

Note: Supposing A_i and A_j are independent, $\mathbb{P}(A_i \cap A_j) - \mathbb{P}(A_i)\mathbb{P}(A_j) = 0$. So independent pairs do not contribute to this sum.

Proposition 40. $p = \frac{1}{n}$ is a sharp threshold for $G \in \mathcal{G}(n,p)$ to contain a $\triangle$, in the sense that:

- if $p = o(\frac{1}{n})$ then almost every $G \in \mathcal{G}(n,p)$ has no $\triangle$, whereas
- if $p = \omega(\frac{1}{n})$ then almost every $G \in \mathcal{G}(n,p)$ has a $\triangle$.

Proof. Let $G \in \mathcal{G}(n,p)$ and let X be the number of $\triangle$s in G. Let $\mu = \mathbb{E}X$ and $\sigma^2 = \mathrm{Var}(X)$. Then $\mu = \binom{n}{3}p^3 \sim \frac{1}{6}(np)^3$. Also,

$$\sigma^2 = \binom{n}{3}(p^3 - p^6) + \binom{n}{3} \cdot 3 \cdot (n-3)(p^5 - p^6) \leq n^3p^3 + n^4p^5.$$

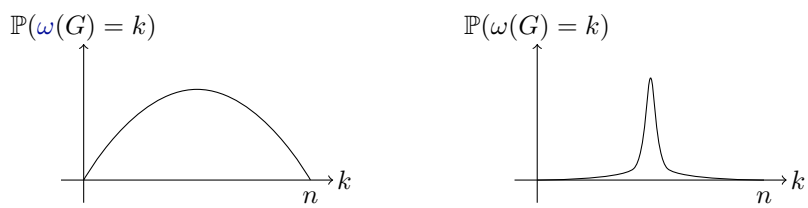
Suppose first $p = \wr(\frac{1}{n})$, i.e. $np \to 0$. Then by Markov,

$$\mathbb{P}(X = 0) = 1 - \mathbb{P}(X \geq 1) \geq 1 - \mu \to 1 \text{ as } n \to \infty.$$

Suppose instead $p = \omega(\frac{1}{n})$ so $np \to \infty$. Then with Chebyshev,

$$\frac{\sigma^2}{\mu^2} \leq \frac{1}{\mu^2}(n^3p^3 + n^4p^5) \sim \frac{36}{n^6p^6}(n^3p^3 + n^4p^5) = \frac{36}{(np)^3} + \frac{36}{n \cdot np} \to 0 \text{ as } n \to \infty. \qquad \square$$

We next wonder: What is the clique number $\omega(G)$ for a typical $G \in \mathcal{G}(n,p)$?



We might guess a picture on the left, but in fact it is closer to the right: we have a 'sharp concentration', so to say.

Theorem 41. There exists a function $d : \mathbb{N} \to \mathbb{N}$ such that a.e. $G \in \mathcal{G}(n,p)$ has $\omega(G) \in \{d-1, d, d+1\}$ (where $d = d(n)$).

Proof sketch. **(Currently missing).** $\square$

Remark.

1. This also shows that a.e. G in $\mathcal{G}(n,p)$ has $\omega(G) \sim 2\log_2 n = \frac{2\log n}{\log 2}$.
2. The same proof works to give a similar result for general p. Then $\omega(G) \sim \frac{2\log n}{\log \frac{1}{p}}$.

Corollary 42. Almost every $G \in \mathcal{G}(n,p)$ has

$$\chi(G) \geq (1 + o(1))\frac{n \log \frac{1}{q}}{2\log n}$$

where $q = 1 - p$.

Proof. Let $G \in \mathcal{G}(n,p)$. Then $\overline{G} \in \mathcal{G}(n,q)$. So by Theorem 41, with probability tending to 1 as $n \to \infty$ we have

$$\begin{aligned} \omega(\overline{G}) &\sim \frac{2\log n}{\log \frac{1}{q}} \\ \implies \alpha(G) &\sim \frac{2\log n}{\log \frac{1}{q}} \\ \implies \chi(G) &\geq \frac{|G|}{\alpha(G)} \sim \frac{n \log \frac{1}{q}}{2\log n}. \end{aligned}$$

□

In fact, Bollobás (1988) proved a.e. $G \in \mathcal{G}(n,p)$ has $\chi(G) \sim \frac{n \log \frac{1}{q}}{2\log n}$, the upper bound is much harder.

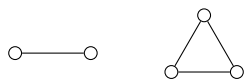
6 Algebraic Methods

Definition (Distance)**.** Let G be a connected graph, $u, v \in G$. The **distance** from u to v is $d(u, v)$, the length of the shortest path from u to v.

Definition (Diameter)**.** The **diameter** of a connected graph G is

$$\max_{u,v \in G} d(u, v).$$

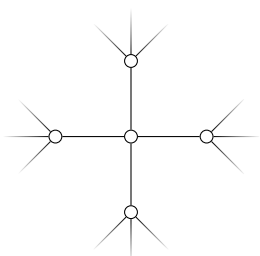
Example. What are the diameter 1 graphs?



In fact, it is just the complete graphs K_n (for $n \geq 2$).

What about diameter 2? There are many more possibilities.

How many vertices can a graph of diameter 2 have if say $\Delta(G) = k$? As diameter 2, if we have $x \in G$: $V(G) = \{x\} \cup \Gamma(x) \cup \Gamma(\Gamma(x))$. So $|G| \leq 1 + k + k(k-1) = k^2 + 1$.

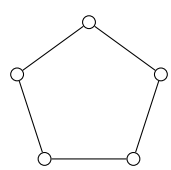


Can we actually achieve this?

Definition (Moore graph)**.** A **Moore graph** is a graph G such that for some k, $|G| = k^2+1$, $\Delta(G) = k$, diameter of G is 2.

For which k can we find a Moore graph? For $k = 1$ we try

but this has the wrong diameter. For $k = 2$:

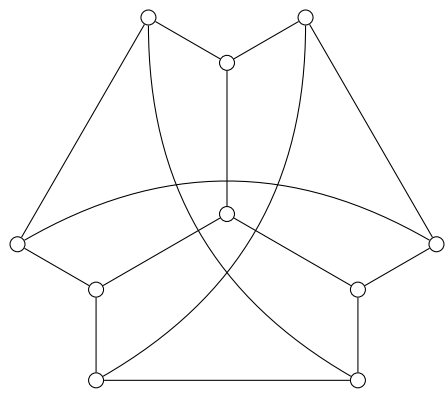


and we see C_5 works.

In general, if u, v are distinct for G Moore,

- If $u \sim v$ then u, v have no common neighbours, i.e. there is no $\triangle$.
- If $u \nsim v$ then u, v have exactly 1 common neighbour, i.e. there are no 4-cycles.

For $k = 3$:



We see the Petersen graph works.

For $k = 4$ we try and fail to construct one, there is no Moore graph with $k = 4$. We might guess that Moore graphs only exist for $k = 2, 3$.

6.1 The Chromatic Polynomial

Suppose G is a k-colourable graph. How many k-colourings does G have?

Example.

- For $G = K_n$ and $k < n$, there are none. For $k \geq n$, we have $k(k-1)\cdots(k-n+1)$.
- If G is a tree, we can repeatedly remove leaves until we get down to a single vertex. Reverse this order and do a greedy algorithm, giving $k(k-1)^{|G|-1}$ colours.

Write $f_G(k)$ for the number of k-colourings of G.

Definition (Contraction)**.** Let G be a graph and $e = uv \in E(G)$. The **contraction of** G **over** e is the graph G/e formed from G by deleting vertices u, v, adding a new vertex e^* with $\Gamma(e^*) = \Gamma(x) \cup \Gamma(y)$.

Theorem 43 (Cut-fuse relation)**.** Let G be a graph, $e \in E(G)$, $k \geq 1$. Then $f_G(k) = f_{G-e}(k) - f_{G/e}(k)$.

Proof. Let $e = uv$. Let c be a k-colouring of $G - e$. If $c(u) \neq c(v)$ then c is a k-colouring of G and every k-colouring of G arises uniquely like this. If $c(u) = c(v)$ then c yields a k-colouring of G/e and every k-colouring of G/e arises uniquely like this.

$$f_{G-e}(k) = f_G(k) + f_{G/e}(k). \qquad \square$$

Corollary 44. Let G be a graph. Then f_G is a polynomial.

Proof. Induction on $e(G)$. For $e(G) = 0$, $f_G(k) = k^{|G|}$. For $e(G) > 0$, pick $e \in E(G)$. Then f_{G-e}, $f_{G/e}$ are polynomials by the induction hypothesis and hence $f_G = f_{G-e} - f_{G/e}$ is a polynomial. $\square$

In light of this, we call f_G the **chromatic polynomial** of G. The original motivation to study f_G was for the Four colour theorem, which is equivalent to saying that if G is planar then $f_G(4) \neq 0$. The idea was to do something algebraic to prove this. The aim was to prove that if G is planar and α a real root of f_G then $\alpha < 4$. Unfortunately, no one managed to do this.

But, f_G carries information about G, e.g.:

Corollary 45. If $|G| = n$, $e(G) = m$ then

$$f_G(X) = X^n - mX^{n-1} + \dots$$

Proof. Induction on $e(G)$. If $e(G) = 0$, then $f(X) = X^n$, as required. For $e(G) > 0$, pick $e \in E(G)$. Then

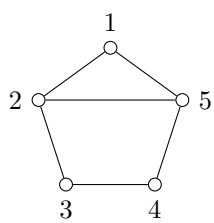
$$\begin{aligned} f_G(X) &= f_{G-e}(X) - f_{G/e}(X) = (X^r(m-1)X^{n-1} + \dots) - (X^{n-1} + \dots) \\ &= X^n - mX^{n-1} + \dots \end{aligned}$$

□

6.2 Eigenvalues

Definition (Adjacency matrix)**.** Let G be a graph with $V(G) = \{1, 2, \dots, n\}$. The **adjacency matrix** of G is the $n \times n$ matrix A where

$$A_{ij} = \begin{cases} 1 & \text{if } i \sim j \\ 0 & \text{if } i \nsim j. \end{cases}$$

$G =$ 

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix}$$

Note $i \sim j \iff j \sim i$ so A is real, symmetric. $i \nsim i$ so diagonal is zero.

What is A^2? We have

$$(A^2)_{ij} = \sum_{k=1}^{n} A_{ik} A_{kj} = |\Gamma(i) \cup \Gamma(j)|,$$

the number of common neighbours of i, j. In particular, $(A^2)_{ii} = d(i)$ so $\mathrm{Tr}(A^2) = 2e(G)$.

More generally,

$$(A^l)_{ij} = \sum_{k_i=1}^{n} \cdots \sum_{k_{l-1}=1}^{n} A_{ik_1} A_{k_1 k_2} \cdots A_{k_{l-2} k_{l-1}} A_{k_{l-1} j}.$$

Definition (Walk)**.** Define a **walk** of length l from u to v to be a sequence

$$u = u_0, u_1, \dots, u_l = v$$

of (not necessarily distinct) vertices with $u_{i-1} \sim u_i$ for $1 \leq i \leq l$.

Then $(A^l)_{ij}$ is the number of walks of length l from i to j.

Definition (Eigenvalues)**.** If G is a graph, the **eigenvalues** of G are the eigenvalues of its adjacency matrix.

Remark. Change of labelling of vertices corresponds to a reordering of basis, so the eigenvalues do not depend on what labelling we choose.

Also, A is real symmetric, so it has real eigenvalues, is diagonalizable and we have an orthonormal basis of eigenvectors. Sometimes it's easy to spot eigenvalues/eigenvectors.

Example.

1) $G = K_n$. Then

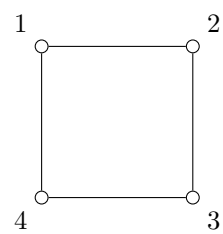
$$A \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} n-1 \\ n-1 \\ \vdots \\ n-1 \end{pmatrix}$$

so $n-1$ is an eigenvalue. Also,

$$A + I = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$$

has rank 1, so nullity $n-1$. So -1 is an eigenvalue with multiplicity $n-1$.

2) $G = C_4$.



$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

A has rank 2 so nullity 2, so 0 is an eigenvalue with multiplicity 2. Also observe

$$A \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \end{pmatrix}, \quad A \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \\ -2 \\ 2 \end{pmatrix}$$

so 2 and -2 are eigenvalues, multiplicity 1 each.

Let A have an orthonormal basis $e_1, \dots, e_n$ of eigenvectors with eigenvalues

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n.$$

Given $x \in \mathbb{R}^n$, we can write $x = \sum_{i=1}^n x_i e_i$ where $\sum x_i^2 = \|x\|^2$. Then $Ax = \sum_i \lambda_i x_i e_i$ so $Ax \cdot x = \sum_i \lambda_i x_i^2$ hence

$$\lambda_1 = \max_{\|x\|=1} Ax \cdot x \quad \text{and} \quad \lambda_n = \min_{\|x\|=1} Ax \cdot x.$$

Theorem 46. Let G be a graph, $\Delta(G) = \Delta$, λ an eigenvalue of G. Then $|\lambda| \leq \Delta$. Moreover if G is connected then Δ is an eigenvalue $\iff$ G is Δ-regular; in this case Δ has multiplicity 1 and eigenvector $\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$.

Proof. Let A be the adjacency matrix and x an eigenvector with eigenvalue λ. Let i be such that $|x_i|$ is maximal. Without loss of generality, $x_i = 1$ and $\forall j, |x_j| \leq 1$. Then

$$\begin{aligned} |\lambda| = |\lambda x_i| = |(Ax)_i| &= \left| \sum_{j=1}^n A_{ij} x_j \right| = \left| \sum_{j \in \Gamma(i)} x_j \right| \\ &\leq \sum_{j \in \Gamma(i)} |x_j| \leq d(i) \leq \Delta. \end{aligned}$$

Assume now G is connected. ($\Leftarrow$). If G is Δ-regular then clearly $(1, \dots, 1)^T$ is an eigenvector with eigenvalue Δ.

($\Rightarrow$). Suppose Δ is an eigenvalue. Then taking $\lambda = \Delta$ in previous:

$$\Delta = (Ax)_i = \sum_{j \in \Gamma(i)} x_j.$$

Hence $d(i) = \Delta$ and $\forall j \in \Gamma(i), x_j = 1$. Repeat: $\forall j \in \Gamma(i)$ we have $d(j) = \Delta$ and $\forall k \in \Gamma(j)$, $x_k = 1$. Continuing, as G connected, $\forall k, d(k) = \Delta$ and $x_k = 1$. $\square$

6.3 Strongly Regular Graphs

Definition (Strongly regular graph)**.** Let $k, b \geq 1$ and $a \geq 0$. A graph G is (k, a, b)**-strongly regular** if G is k-regular and, for all $x, y \in G$ with $x \neq y$

- $x \sim y \implies |\Gamma(x) \cap \Gamma(y)| = a$,
- $x \nsim y \implies |\Gamma(x) \cap \Gamma(y)| = b$,

Remark. A strongly regular graph is connected. Indeed, it has diameter ≤ 2.

A Moore graph is $(k, 0, 1)$-strongly regular. Conversely, suppose G is $(k, 0, 1)$-strongly regular with $k \geq 2$. Then $\Delta(G) = k$ and G has diameter 2. Pick $x \in G$, then

$$|G| = |\{x\}| + |\Gamma(x)| + |\Gamma(\Gamma(x)) \setminus \{x\}| = 1 + k + k(k-1) = k^2 + 1$$

so G is a Moore graph.

Let G be (k, a, b)-strongly regular. Let A be the adjacency matrix of G. Let $|G| = n$. By Theorem 46, k is an eigenvalue of multiplicity 1 with eigenvector $(1, 1, \ldots, 1)^T = v$.

What about other eigenvalues? Let $\lambda \neq k$ be an eigenvalue with eigenvector x. Now

$$(A^2)_{ij} = \begin{cases} k & \text{if } i = j \\ a & \text{if } i \sim j \\ b & \text{if } i \nsim j \end{cases}$$

Thus $A^2 = kI + aA + b(J - I - A)$ where J is the matrix with a 1 in every place. Applying this to x, and noting that $x \perp v$, giving $Jx = 0$, we get $\lambda^2 x = kx + a\lambda x - bx - b\lambda x$. Also $x \neq 0$, so $\lambda^2 + (b - a)\lambda + (b - k) = 0$.

So the remaining eigenvalues are

$$\lambda = \frac{(a-b) + \sqrt{(b-a)^2 - 4(b-k)}}{2} \text{ and } \mu = \frac{(a-b) - \sqrt{(b-a)^2 - 4(b-k)}}{2}$$

with multiplicities r, s say, respectively.

Now A is diagonalizable so

$$r + s + 1 = n \tag{1}$$

and $\operatorname{Tr} A = 0$ so

$$\lambda r + \mu s + k = 0 \tag{2}$$

Now take $\lambda \times (1) - (2)$: $(\lambda - \mu)s = \lambda(n - 1) + k$.

We have $\lambda - \mu = \sqrt{(b-a)^2 - 4(b-k)}$ so

$$s = \frac{1}{2}\left\{(n-1) + \frac{(a-b)(n-1) + 2k}{\sqrt{(b-a)^2 - 4(b-k)}}\right\}$$

and so

$$r = \frac{1}{2}\left\{(n-1) - \frac{(a-b)(n-1) + 2k}{\sqrt{(b-a)^2 - 4(b-k)}}\right\}$$

Theorem 47 (Rationality condition)**.** Let G be (k, a, b)-strongly regular. Then

$$\frac{1}{2}\left\{(n-1) \pm \frac{(a-b)(n-1) + 2k}{\sqrt{(b-a)^2 - 4(b-k)}}\right\} \in \mathbb{Z}.$$

Proof. See discussion above. □

Example. Moore graphs have $a = 0$, $b = 1$, $n = k^2 + 1$. So

$$\frac{1}{2}\left\{k^2 \pm \frac{2k - k^2}{\sqrt{4k - 3}}\right\} \in \mathbb{Z}.$$

There are two possibilities:

(i) $2k - k^2 = 0$, so $k = 2$.

(ii) $4k-3=u^2$ for some $u \in \mathbb{N}$, with $u|2k-k^2$. Then $k=\frac{u^2+3}{4}$ so $k^2-2k=ut$ for some $t \in \mathbb{Z}$. Hence

$$\left(\frac{u^2+3}{4}\right)^2 - 2\left(\frac{u^2+3}{4}\right) = ut$$
$$\implies u^4+6u^2+9-8u^2-24=16ut \implies u^4-2u^2-16ut-15=0.$$

Hence $u \mid 15$, so $u=1,3,5,15$. Hence $k=1,3,7,57$.

Can we actually find these? For $k=1$ we try

but this has the wrong diameter.

For $k=2$: We saw C_5 works. Similarly for $k=3$ we saw the Petersen graph works. For $k=7$, we have $n=50$, and such a graph does exist (see Page 2). For $k=57$? $n \approx 3000$, and this is in unsolved problem.