MATH5425 Graph Theory. Semester 2, 2015

Summary of lecture notes by Catherine Greenhill

This is a summary of the MATH5425 lecture content. Most of the content is based on Diestel's textbook *Graph Theory* [3]. The other two textbooks are mentioned in the bibliography at the end of the notes. For ease of reference, all numbered sections, numbered theorems etc, matching the numbering in Diestel, unless otherwise specified. The main purpose of these notes is to capture the important definitions and statements of results covered in the course. Proofs are omitted from this summary: in almost all cases they are either left as an exercise (possibly on the problem sheets) or given during lectures.

1 Introduction

The main reference for this section is Diestel [3, Chapter 1].

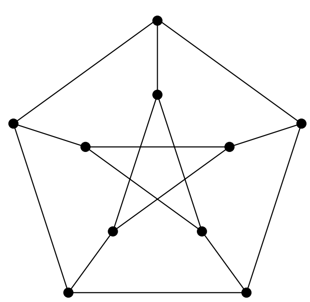
A *graph* $G = (V, E)$ is a set V of *vertices* and a set E of unordered pairs of distinct vertices, called *edges*. Sometimes we write $V(G)$ and $E(G)$ for the vertex set and edge set of a graph G, respectively.

Write vw or $\{v, w\}$ for the edge joining v and w, and say that v and w are *neighbours* or that they are *adjacent*.

In this course, unless otherwise stated, graphs are

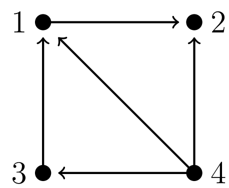
- *finite*: so $|V| \in \mathbb{N}$.
- *labelled*: we can distinguish vertices from each other. Usually $V = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$.
- *undirected*, since edges are *unordered* pairs of vertices.
- *simple*: no loops (that is, no edges $\{v, v\}$) or multiple edges (since E is a *set*, not a multiset).

Usually we draw a graph like this, with vertices as black dots and edges as a line between them. (We often don't display the vertex labels.) This is the *Petersen graph*:

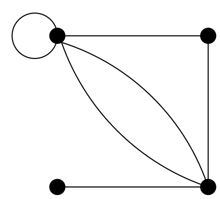


Related objects

Directed graphs have ordered edges (v, w), drawn as an arrow from v to w.



Multigraphs are undirected but have loops and multiple edges allowed.



A *hypergraph* $H = (V, E)$ consists of a set of vertices V and a set E of *hyperedges*, where each hyperedge is a nonempty subset of V. These are harder to draw.

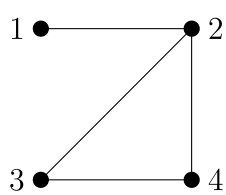
Another representation: adjacency matrices

A graph G with vertex set $\{v_1, \ldots, v_n\}$ can be represented by its *adjacency matrix* $A(G)$, which is an $n \times n$ matrix $A(G) = (a_{ij})$ where

$$a_{ij} = \begin{cases} 1 & \text{if } v_i v_j \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Hence $A(G)$ is a symmetric 0-1 matrix (every entry is either 0 or 1) with zero diagonal. For example:

G



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

Similarly, a multigraph corresponds to a symmetric matrix over $\mathbb{N}$ and a directed graph corresponds to a square 0-1 matrix with a zero diagonal.

Hypergraphs are described by their *incidence matrix*. If $H = (V, E)$ where $V = \{v_1, \ldots, v_n\}$ and $E = \{e_1, \ldots, e_m\}$ then the incidence matrix $B(H)$ of H is the $n \times m$ 0-1 matrix $B(H) = (b_{ve})$, where

$$b_{ve} = \begin{cases} 1 & \text{if } v \in e, \\ 0 & \text{otherwise.} \end{cases}$$

In this course we do not work much with the adjacency/incidence matrices.

Usefulness of graphs

We claim that graphs are useful in the real world.

- **Computer or communications networks**. Let each node in a computing network be a vertex in the graph, and put an edge between two vertices if the corresponding computers are joined by a cable. This can also be used to model social networks, or the internet. How many computers would have to fail before the network is disconnected? This is related to the *connectivity* of the graph.

- **Gene sequencing.** Smash many copies of a gene into small segments. Let V be the set of these segments and place an edge between two segments if they have a large enough overlap. Then use graph theoretical techniques to sequence the gene.

- A mobile phone company wants to build transmitters to cover the sites in some set V. Each transmitter can transmit for r km. What is the least number of transmitters required, and where should they be built? Model the problem as a graph with vertex set V, and with an edge from v to w if the distance from v to w is at most r. Then look for the smallest subset U of vertices such that every vertex in V which is not in U is *adjacent* to a vertex in U. In graph-theoretic language, we seek a *minimal dominating set* of the graph.

  Other optimization problems on graphs include the Travelling Salesman problem: design an efficient route for a salesman to visit a given set of towns. Each road (with distance) is modelled as a *labelled* edge in the graph. The aim is to minimise the sum of the distances along a *Hamilton cycle* or *Hamilton path*: a cycle (or path) which visits each vertex exactly once. This problem is NP-hard.

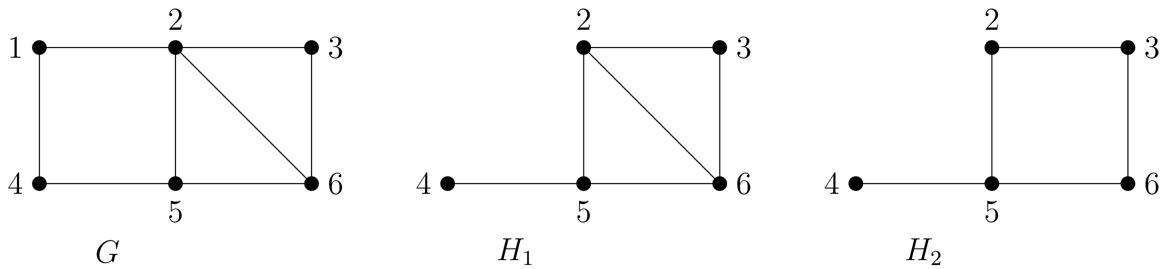
- **Random graphs**. Often in applications we want to know how an algorithm or hypothesis will work on a "typical" instance of the problem. If the problem can be modelled using graphs, then we want a *random graph* with certain properties.

At the end of the course we will look at random graphs. We will also use probabilistic methods to prove some deterministic results during the course. Only very basic discrete probability theory will be used, and the necessary concepts will be introduced when needed.

More definitions

The *trivial graph* has at most one vertex. Hence it has no edges. (We often ignore the graph with zero vertices.)

A *subgraph* of a graph $G = (V, E)$ is a graph $H = (W, F)$ such that $W \subseteq V$ and $F \subseteq E$. The figure below shows a graph G together with two subgraphs, H_1 and H_2.



We say that H is an *induced subgraph* if for all $v, w \in W$, if $vw \in E(G)$ then $vw \in E(H)$. So H is induced if it contains all edges of G which are subsets of W. In this case we write $H = G[W]$, and say that H is the subgraph of G *induced by* the vertex set W. In the above example, H_1 is an induced subgraph of G but H_2 is not. We can write $H_1 = G[2, 3, 4, 5, 6]$. (For ease of notation we drop the curly brackets around the vertices when writing the set out explicitly.)

The number of vertices, written $|G| = |V(G)|$, is called the *order* of G and the number of edges, written $\|G\| = |E(G)|$, is called the *size* of G. (I tend to write $|E(G)|$ though.)

Two graphs $G = (V, E)$ and $H = (W, F)$ are *isomorphic* if there exists a bijection $\varphi : V \to W$ such that $\varphi(v)\varphi(w) \in F$ whenever $vw \in E$. The map φ is called a *graph isomorphism*, or just *isomorphism*. We often say that two graphs are *equal* if they are equal up to isomorphism, which is the same as saying that we can relabel the vertices of one graph to obtain the other.

Any graph property which is invariant under isomorphism is a *graph invariant*. For example, the number of edges is a graph invariant.

1.2 The degree of a vertex

If $v \in e$ where v is a vertex and e is an edge, then we say that e is *incident with* v.

The *degree* $d_G(v)$ of vertex v in a graph G is the number of edges of G which are incident with v. We write $d(v)$ if the graph G is clear from the context.

Let $N_G(v)$ (or just $N(v)$) be the set of all neighbours of v in G. Then $d(v) = |N(v)|$.

Lemma (Handshaking lemma)**.**
In any graph $G = (V, E)$,

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

Let $\delta(G) = \min_{v \in V} d(v)$ be the *minimum degree* and $\Delta(G) = \max_{v \in V} d(v)$ be the *maximum degree* in G.

A vertex of degree 0 is an *isolated vertex*.

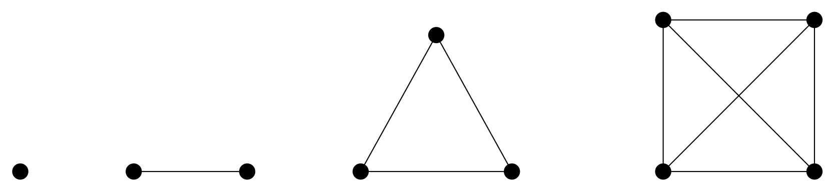
Some special graphs

A graph is *k-partite* if there exists a partition of its vertex set

$$V = V_1 \cup V_2 \cup \cdots \cup V_k$$

into k nonempty disjoint subsets (parts) such that there are no edges between vertices in the same part. We say *bipartite* and *tripartite* instead of 2-partite and 3-partite, respectively.

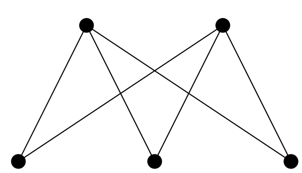
The *complete graph* on r vertices, denoted K_r, has all $\binom{r}{2}$ edges present. (We say "the" complete graph K_r because it is unique up to isomorphism.) The first few complete graphs are shown below.



K_1 K_2 K_3 K_4

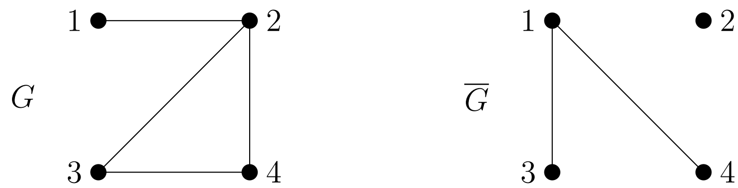
The *complete bipartite graph* $K_{r,s}$ has r vertices in one part of the vertex bipartition, s vertices in the other, and all rs edges present. For example, $K_{2,3}$ is pictured below.

$K_{2,3}$



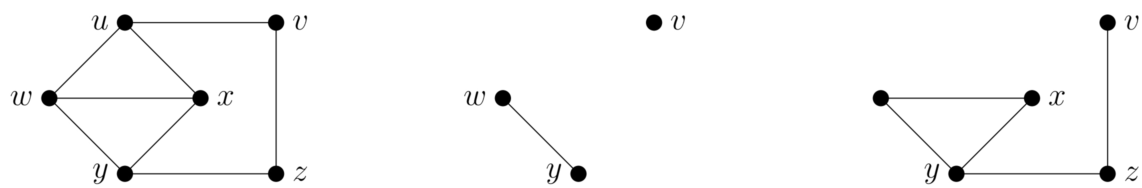
A graph is *regular* if every vertex has the same degree. If every vertex of a graph has degree d then we say that the graph is d-regular. Observe that K_r is regular (of what degree?). When is $K_{r,s}$ regular?

The *complement* of a graph G is the graph $\overline{G} = (V, \overline{E})$ where $vw \in \overline{E}$ if and only if $vw \notin E$. An example of a graph G and its complement $\overline{G}$ is given below.



Note that $\overline{K_n}$ is the graph with n vertices and on edges.

If $G = (V, E)$ and $X \subset V$, then $G - X$ denotes the graph obtained from G by deleting all vertices in X and alledges which are incident with vertices in X. We write $G - v$ instead of $G - \{v\}$, if $v \in E$. Also write $G - H$ instead of $G - V(H)$, if H is a subgraph of G.



G $\qquad$ $G - \{u, x, z\}$ $\qquad$ $G - u$

Similarly, if F is any subset of the edges of G then $G - F$ denotes the graph $(V, E - F)$ obtained from G by deleting those edges in F. Write $G - e$ or $G - xy$ when $F = \{e\}$ and $e = xy$.

Also write $G + xy$ or $G + e$ to denote the graph obtained from G by adding the edge e. That is, $G + e = (V, E \cup \{e\})$. (If e is already present as an edge of G then $G + e = G$.)

1.3 Paths and cycles

A *walk* in the graph G is a sequence of vertices

$$v_0\, v_1\, v_2\, \cdots\, v_k$$

such that $v_i v_{i+1} \in E$ for $i = 0, 1, \ldots, k-1$. Sometimes we write this walk as

$$v_0\, e_0\, v_1\, e_1\, \cdots\, v_{k-1}\, e_{k-1}\, v_k$$

where $e_i = v_i v_{i+1}$. The *length* of this walk is k.

The walk is *closed* if $v_0 = v_k$.

An *Euler tour* is a closed walk in a graph which uses every edge precisely once. A graph is *Eulerian* if it has an Euler tour.

The following theorem, proved by Euler in 1736, may be the first theorem in graph theory. Read about the Seven Bridges of Königsberg on Wikipedia:
http://en.wikipedia.org/wiki/Seven_Bridges_of_Konigsberg

Theorem. 1.8.1 (Euler, 1736)
A connected graph is Eulerian if and only if every vertex has even degree.

Proof. Exercise (see Problem Sheet 1). □

It is a nice fact that the number of walks of length k from v to w in G is $(A(G)^k)_{vw}$ (the (v, w) entry of the kth power of $A(G)$). This can be proved by induction on k (exercise).

A walk is called a *path* if it does not visit any vertex more than once. So a path is a sequence of *distinct* vertices, with subsequent vertices joined by an edge. A path $v_0\, v_1 \cdots v_k$ with k edges is called a k-path, and has length k. A k-path can be thought of as a graph with $k + 1$ vertices and k edges.

If $k \geq 3$ and $P = v_0\, v_1 \cdots v_{k-1}$ is a path of length $k - 1$ then $C = P + v_0 v_{k-1}$ is a *cycle* of length k. It is a closed walk which visits no internal vertex more than once. We may write $v_0\, v_1 \cdots v_{k-1}\, v_0$ to denote the cycle.

An edge which joins two vertices of a cycle C, but which is not an edge of C, is called a *chord*. An *induced cycle* is a cycle which has no chords.

Proposition. 1.3.1
Every graph G contains a path of length $\delta(G)$ and a cycle of length at least $\delta(G) + 1$, if $\delta(G) \geq 2$.

Proof. Proof given in lectures. □

The minimum length of a cycle in G is the *girth* of G, denoted by $g(G)$.

Given $x, y \in V$, let $d_G(x, y)$ be the length of a shortest path from x to y in G, called the *distance* from x to y in G. Set $d_G(x, y) = \infty$ if no such path exists.

We say that G is *connected* if $d_G(x, y)$ is finite for all $x, y \in V$. Check that $d_G(\cdot, \cdot)$ defines a metric on V if G is connected.[1]

Let the *diameter* of G be

$$\mathrm{diam}(G) = \max_{x,y \in V} d_G(x, y).$$

Proposition. 1.3.2
Every graph G which contains a cycle satisfies

$$g(G) \leq 2\ \mathrm{diam}(G) + 1.$$

Proof. Proof given in lectures. □

[1]Although we use d to denote degree and distance, this should not cause confusion as degree has only one argument while distance has two.

Graphs for which equality holds (that is, where $g(G) = 2\,\mathrm{diam}(G) + 1$) are called *Moore graphs*. Two examples are odd cycles (that is, cycles of odd length) and the Petersen graph (check). Singleton proved in 1968 that any graph G with diameter k and girth $2k + 1$ must be d-regular for some d, with order

$$|G| = 1 + \sum_{i=0}^{k} d(d-1)^i.$$

Only a handful of Moore graphs are known to exist: this is a topic of ongoing research.

Fact: a graph is bipartite if and only if it has no odd cycles. (The proof is left as an exercise: see Problem Sheet 1.)

1.4 Connectivity

We have seen that a graph G is connected if there is a path from v to w in G, for all $v, w \in V$.

A maximal connected subgraphs of G is called a *component* (or *connected component*) of G. Components are nonempty.

Proposition. 1.4.1
The vertices of a connected graph can be labelled $v_1, v_2, \ldots, v_n$ such that $G_n = G$ and $G_i = G[v_1, \ldots, v_n]$ is connected for all i.

Proof. Proof given in lectures. □

Now let $A, B \subseteq V$ be sets of vertices. An (A, B)-*path* in G is a path $P = x_0\, x_1 \cdots x_k$ such that

$$P \cap A = \{x_0\}, \qquad P \cap B = \{x_k\}.$$

(Note, this notation differs slightly than that used by Diestel, who writes "$A - B$ path", but we avoid this as it looks too much like "A minus B".)

Let $A, B \subseteq V$ and let $X \subseteq V \cup E$ be a set of vertices and edges. We say that X *separates* A and B in G if every (A, B)-path in G contains a vertex or edge from X.

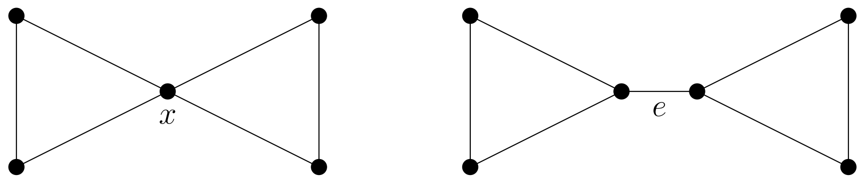
Here we do not assume that A and B are disjoint. If X separates A and B then $A \cap B \subseteq X$.

More generally, we say that X *separates* G, and call X a *separating set* for G, if X separates two vertices of $G - X$ in G.

If $X = \{x\}$ is a separating set for G, where $x \in V$, then we say that x is a *cut vertex*. In this case, $G - x$ has more components than G.

If $G - e$ has more components than G, where $e \in E$, then we say that the edge e is a *bridge*. In this case e separates its end vertices in G.

In the figure below, the graph on the left has a cut vertex x and the graph on the right has a bridge e.



Fact: an edge is a bridge if and only if it does not lie on a cycle.

The unordered pair (A, B) is a *separation* of G if $A \cup B = V$ and G has no edge between $A - B$ and $B - A$. The second condition is equivalent to saying that $A \cap B$ separates A from B in G.

If both $A - B$ and $B - A$ are nonempty then the separation is *proper*. The *order* of the separation is $|A \cap B|$.

Definition. Let $k \in \mathbb{N}$. The graph G is *k-connected* if $|G| > k$ and $G - X$ is connected for all subsets $X \subseteq V$ with $|X| < k$.

The *connectivity* $\kappa(G)$ of G is defined by $\kappa(G) = \max\{k : G \text{ is } k\text{-connected}\}$.

This says that it is impossible to disconnect a k-connected graph by deleting fewer than k vertices.

Every graph with at least one vertex is 0-connected. The graph G is 1-connected if and only if it is connected and has at least 2 vertices.

Therefore $\kappa(G) = 0$ if and only if G is trivial or G is disconnected.

Fact: $\kappa(K_n) = n - 1$ for all positive integers n.

Definition. Let $\ell \in \mathbb{N}$ and let G be a graph with $|G| \geq 2$. If $G - F$ is connected for every set $F \subseteq E$ of fewer than ℓ edges (that is, for all $F \subseteq E$ with $|F| < \ell$) then we say that G is *ℓ-edge-connected*.

The *edge connectivity* $\lambda(G)$ is defined by $\lambda(G) = \max\{\ell : G \text{ is } \ell\text{-edge-connected}\}$.

If G is disconnected then $\lambda(G) = 0$.

Proposition. 1.4.2
If $|G| \geq 2$ then

$$\kappa(G) \leq \lambda(G) \leq \delta(G).$$

Proof. Exercise (see Problem Sheet 1). □

From this it follows that high connectivity implies high minimum degree. However, large minimum degree is not enough to ensure high connectivity, or even high edge-connectivity: think of some examples that illustrate this.

On the other hand, we can link connectivity with the *average degree* of a graph G, which equals $2|E(G)|/|V(G)|$.

Theorem. 1.4.3 (Mader, 1973)
Let k be a positive integer. Every graph G with average degree at least $4k$ has a $(k+1)$-connected subgraph H with

$$\frac{|E(H)|}{|V(H)|} > \frac{|E(G)|}{|V(G)|} - k.$$

Proof. Proof given in lectures. □

(This theorem says that a graph G with sufficiently high average degree contains a subgraph H which is k-connected and almost as dense as G.)

1.5 Trees and forests

A graph with no cycles is a *forest* (also called an *acyclic* graph). A connected graph with no cycles is a *tree*.

The vertices of degree 1 in a tree or forest are called the *leaves*.

Fact: Every nontrivial tree has at least two leaves. (Consider a longest path.)

Theorem. 1.5.1
The following are equivalent for a graph T:

(i) T is a tree;

(ii) Any two vertices of T are linked by a *unique* path in T;

(iii) T is *minimally connected*: that is, T is connected but $T - e$ is disconnected for every $e \in E(T)$;

(iv) T is *maximally acyclic*: that is, T is acyclic but $T + xy$ has a cycle for any two nonadjacent vertices x, y in T.

Proof. Exercise (see Problem Sheet 1). □

A subgraph H of G is a *spanning subgraph* if every vertex of G is incident with at least one edge of H.

If H is a spanning subgraph which is a tree then H is a *spanning tree*.

Corollary.
If G is connected then G has a spanning tree.

Proof. Proof given in lectures. □

Corollary. 1.5.2
The vertices of a tree can be labelled as $v_1, \ldots, v_n$ so that for $i \geq 2$, vertex v_i has a unique neighbour in $\{v_1, \ldots, v_{i-1}\}$.

Proof. Proof given in lectures. □

We now show that we can characterise trees by counting edges.

Corollary. 1.5.3
A connected graph with n vertices is a tree if and only if it has $n - 1$ edges.

Proof. Proof given in lectures. □

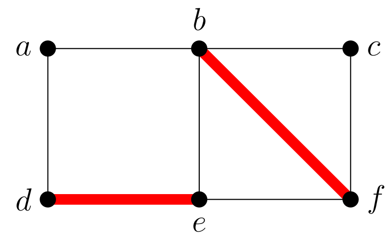
Corollary. 1.5.4
If T is a tree and G is any graph with $\delta(G) \geq |T| - 1$ then G has a subgraph isomorphic to T.

Proof. Proof given in lectures. □

2 Matchings and Hamilton cycles

The reference for the material on matchings is Diestel [3, Chapter 2], and for Hamilton cycles it is Diestel [3, Chapter 10].

Two edges in a graph are called *independent* if they have no vertices in common. A set M of pairwise independent edges in a graph is called a matching. For example, $\{bf,\, de\}$ is a matching in the graph shown below.



Given $G = (V, E)$, say that $M \subseteq E$ is a *matching of* $U \subseteq V$ if M is a matching and every vertex in U is incident with an edge of M. In the example above, M is a matching of $\{b, d, e, f\}$, and it is also a matching of $\{b, d\}$. We say that the vertices in U are *matched* by M, and that the vertices not incident with any edge of M are *unmatched*.

A matching M is *maximal matching* of G if $M \cup \{e\}$ is not a matching for any $e \in E - M$. A *maximum matching* of G is a matching of G such that no set of edges with size greater than $|M|$ is a matching. The matching $\{bf,\, de\}$ in our example is maximal, but not maximum (check!).

A *perfect matching* of G is a matching of G which matches every vertex of G. It is a 1-regular spanning subgraph of G, also called a *1-factor* of G.

The graph on the left cannot have a perfect matching, as it has an odd number of vertices. The graph on the right has a perfect matching as shown: any others?

More generally, a *k-factor* is a k-regular spanning subgraph. For example, a 2-factor in a graph is a union of disjoint cycles which covers all the vertices.

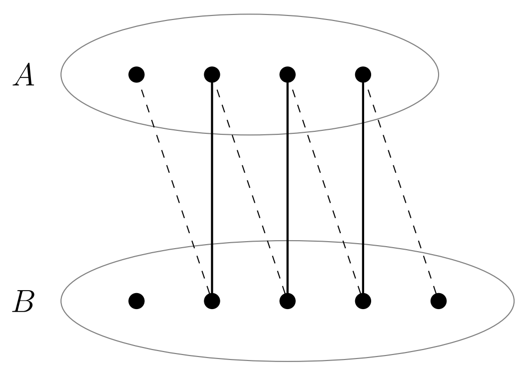
Our focus in this section is: when does a graph have a perfect matching?

2.1 Matchings in bipartite graphs

Warning: if you look at Diestel [3, Section 2.1], he explains *once* that G is bipartite throughout that section, and does not continually remind you that G is bipartite. So please bear in mind that any theorem or lemma stated in this section of Diestel *only holds for bipartite graphs*.

Let $G = (V, E)$ b a fixed bipartite graph with vertex bipartition $V = A \cup B$ (where A and B are nonempty disjoint sets). We use the convention that vertices called $a, a', a'', \ldots$ belong to A, while vertices called $b, b', b'', \ldots$ belong to B. We want to find a matching in G which contains as many vertices of A as possible.

Consider an arbitrary matching M in G. A path in G which starts at an *unmatched* vertex of A and contains, alternately, edges from $E - M$ and from M, is called an *alternating path* with respect to M. If an alternating path P ends in an unmatched vertex of B then it is called an *augmenting path*, and we can use it to turn M into a larger matching, as illustrated below.



This figure shows an augmenting path P of length 7. Replacing the three edges of M (the solid lines) by the four edges of $E(P) - M$ (the four dashed lines) gives rise to a matching of size four. In general, the symmetric difference of M with $E(P)$ is a matching and the set of matched vertices is increased by two.

Fact: If you start from any matching and repeatedly “flip” augmenting paths in this manner until no more augmenting paths exist, then the result is a maximum matching of G: that is, you reach a matching with the largest possible number of edges. The proof is left as an exercise (see Problem Sheet 2).

We can characterise maximum matchings using a kind of duality condition.

Definition. A set $U \subseteq V$ is a *cover* (or a *vertex cover* of G) if every edge of G is incident with a vertex in U.

Theorem. 2.1.1 (König, 1931)
Let G be a bipartite graph. The size of a maximum matching in G is equal to the size of a minimum (i.e., smallest) vertex cover of G.

Proof. Proof given in lectures. $\square$

For a bipartite graph to have a perfect matching we need $|A| = |B|$. Instead we will ask when G has a matching which matches A. For a subset $S \subseteq A$, let $N(S) = \cup_{v \in S} N(v)$ be the set of vertices in B which are neighbours of some set in A.

Theorem. 2.1.2 (Hall 1935)
Let G be a bipartite graph. Then G contains a matching of A if and only if

$$|N(S)| \geq |S| \qquad \text{for all } S \subseteq A. \tag{1}$$

(This condition is called "Hall's condition".)

Proof. Proof given in lectures. $\square$

Diestel presents three proofs of this theorem: one is the one we saw in lectures (which uses vertex covers), one uses augmenting paths and one is a first principles argument using induction.

Corollary.
Let G be a bipartite graph and let $d \in \mathbb{N}$. If $|N(S)| \geq |S| - d$ for all $S \subseteq A$ then G has a matching of size $|A| - d$.

Proof. Proof given in lectures. $\square$

Corollary. 2.1.3
If G is a k-regular bipartite graph then G has a perfect matching.

Proof. Proof given in lectures. $\square$

Hall's Theorem is one of the most frequently applied results in graph theory. It gives a very nice proof of this very early graph theory result.

Corollary. 2.1.5 (Petersen, 1891)
Every regular graph of positive even degree has a 2-factor.

Proof. Proof given in lectures. $\square$

10. Hamilton cycles

Inspired by Petersen's result on 2-factors, we take a slight detour now to discuss Hamilton cycles. A *Hamilton cycle* is a connected 2-factor. That is, it is a cycle which includes every vertex. Deciding whether or not a graph has a Hamilton cycle is an NP-complete problem, which means that it is unlikely that a polynomial-time algorithm exists for this problem. (Contrast this with Theorem 1.8.1, which gives a very easy criterion for the existence of an Euler tour.)

Clearly if the graph G contains a Hamilton cycle (we also say that G is *Hamiltonian*) then G must be connected and the minimum degree of G is at least two. But no lower bound on $\delta(G)$ is sufficient to ensure a Hamilton cycle: exercise (see Problem Sheet 2). However, we do have the following theorem, proved by Gabriel Dirac, who was the stepson of the physicist Paul Dirac.

Theorem. 10.1.1 (Dirac, 1952)
Every graph with $n \geq 3$ vertices and with minimum degree at least $n/2$ has a Hamilton cycle.

Proof. Proof given in lectures. □

2.2 Matchings in general graphs

We return now to the topic of matchings, but this time for general graphs (which need not be bipartite).

Given a graph G, let C_G be the set of its components and let $q(G)$ denote the number of *odd components* (that is, the number of components of odd order).

Theorem. 2.2.1 (Tutte, 1947)
A graph G has a perfect matching if and only if

$$q(G - S) \leq |S| \qquad \text{for all } S \subseteq V(G). \tag{2}$$

Proof. Proof given in lectures. □

We say a graph is *cubic* if it is 3-regular.

Corollary. 2.2.2 (Petersen, 1891)
Every bridgeless cubic graph has a perfect matching.

Proof. Proof given in lectures. □

3 The probabilistic method

The main reference for this section is Alon & Spencer *The Probabilistic Method* [1].

The probabilistic method is a non-constructive existence proof method which was pioneered by Paul Erdős in the 1950s, and which led to the study of *random graphs*. The idea is as follows: to prove that some element of a set Ω has a desired property, define a probability distribution on Ω and show that a *random* element of Ω satisfies the desired property *with positive probability.* (See the handout for basic probability theory definitions.)

Let Ω be a finite set and let $\pi : \Omega \to [0, 1]$ be a probability distribution on Ω. Then (Ω, π) is a probability space on Ω. Often we work with the *uniform distribution*, where $\pi(z) = 1/|\Omega|$ for all $z \in \Omega$.

Example. Let Ω be the set of all graphs on the vertex set $\{1, 2, \ldots, n\}$. Then $|\Omega| = 2^{\binom{n}{2}}$. Define $\pi(G) = 2^{-\binom{n}{2}}$ for all $G \in \Omega$. This is the *uniform model of random graphs.*

We often just write $\Pr(\cdot)$ instead of $\pi(\cdot)$.

Given a random variable $X : \Omega \to \mathbb{R}$ on a probability space (Ω, π), the *expected value* $\mathbb{E}X$ of X is

$$\mathbb{E}X = \sum_{z\in\Omega} \Pr(z)\, X(z).$$

It is a weighted average of the values of X over Ω, weighted by the probability of each element under π.

Lemma.
The expected number of edges in a uniformly chosen graph on the vertex set $\{1, 2, \ldots, n\}$ is $\frac{1}{2}\binom{n}{2}$.

Proof. Proof given in lectures. □

There is a much easier way to perform this and similar calculations, using indicator variables and linearity of expectation. The *indicator variable* for an event $A \subseteq \Omega$ is the map $I_A : \Omega \to \{0, 1\}$ where

$$I_A(z) = \begin{cases} 1 & \text{if } z \in A, \\ 0 & \text{otherwise.} \end{cases}$$

(It is also called the *characteristic function* of A, when A is viewed as a set.) Note that $\mathbb{E}I_A = \Pr(A)$.

Lemma.
Let Ω be the set of all subsets of some given set S, where $|S| = n$. (That is, Ω is the power set of S.) Define a *random set* $X \subseteq S$ by setting $\Pr(x \in X) = \frac{1}{2}$, *independently* for each $x \in S$. (You can picture this as an experiment where you flip a fair coin for each $x \in S$, independently, and put x into X if and only if the coin flip for x comes up heads.)

Then $\Pr(X = A) = 2^{-n}$ for all $A \subseteq S$, so this gives the *uniform probability space* on Ω.

Proof. Proof given in lectures. □

Exercise: Describe the uniform model of random graphs using coin flips. (See Problem Sheet 2.)

Now we come to our first application of the probabilistic method.

Theorem. (Alon & Spencer [1, Theorem 2.2.1])
Let G be a graph with n vertices and m edges. Then G contains a bipartite subgraph with at least $m/2$ edges.

Proof. Proof given in lectures. □

An *independent set* in a graph G is a subset $U \subseteq V$ such that if $v, w \in U$ then $vw \notin E(G)$. Let $\alpha(G)$ be the size of a maximum (i.e., largest) independent set in G, called the *independence number*.

Theorem. (Alon & Spencer [1, Theorem 3.2.1])
Let G have n vertices and $nd/2$ edges, where $d \geq 1$. (Here d is the average degree of G.) Then $\alpha(G) \geq \frac{n}{2d}$.

Proof. Proof given in lectures. □

4 Graph colourings

The reference for this section is Diestel [3, Chapter 5].

How many different radio frequencies do you need to be able to assign a radio frequency to each radio station in such a way that nearby stations do not interfere with each other's broadcasts? This question can be answered using graph colourings.

A *vertex colouring* (or *colouring*) of a graph $G = (V, E)$ is a function $c : V \to S$ such that $c(u) \neq c(v)$ whenever $uv \in E$. Here S is the set of available *colours*, usually $S = \{1, 2, \ldots, k\}$ for some positive integer k. A *k-colouring* of G is a colouring $c : V \to \{1, 2, \ldots, k\}$.

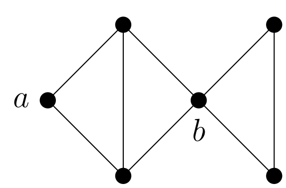
Often we want the smallest value of k for which a k-colouring of G exists. This smallest value of k is called the *chromatic number* of G, denoted by $\chi(G)$. If $\chi(G) = k$ then G is said to be *k-chromatic*. If $\chi(G) \leq k$ then G is said to be *k-colourable*.

The set of all vertices in G with a given colour under c is called a *colour class*. Each colour class is an *independent set*. (Recall, an independent set is a set of vertices which contains no edge of G.) So a k-colouring is just a partition of $V(G)$ into k independent sets. For example, G is 2-colourable if and only if G is bipartite.

The use of colouring terminology arises from the famous *Four Colour Theorem* for planar graphs: more later.

You can check that $\chi(K_r) = r$, so the chromatic number of a complete graph is the same as its order.

A *clique* in a graph G is a complete subgraph of G. The order of the largest clique in G is called the *clique number* of G, denoted by $\omega(G)$. In the graph below, $\omega(G) = 3$ but $\omega(G + ab) = 4$.



Clearly $\chi(G) \geq \omega(G)$.

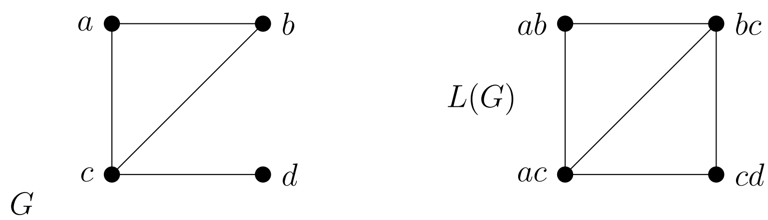
Also $\chi(G) \geq n/\alpha(G)$ where $\alpha(G)$ is the *independence number* of G (which is the size of the largest independent set in G). Exercise: See Problem Sheet 4.

An *edge colouring* of G is a map $c : E \to S$ such that $c(e) \neq c(f)$ whenever e and f share an endvertex (that is, whenever e and f are adjacent). If $S = \{1, 2, \ldots, k\}$ then c is a *k-edge-colouring* and G is *k-edge-colourable*.

Let $\chi'(G)$ be the smallest positive integer k for which G is k-edge-colourable: this value is the *chromatic index* of G (also called the *edge-chromatic number* of G).

A colour class in an edge colouring is a matching of G. Hence an edge colouring displays $E(G)$ as a union of disjoint matchings.

The *line graph*, denoted $L(G)$ has vertex set $E(G)$ and $e, f \in E(G)$ form an edge of $L(G)$ if and only if e, f share an endvertex in G. An example of a graph and its line graph is shown below.



Every edge-colouring of G is a vertex colouring of $L(G)$, and vice-versa. So $\chi'(G) = \chi(L(G))$.

5.2 Vertex colourings

Proposition. 5.2.1
If G has m edges then

$$\chi(G) \leq \frac{1}{2} + \sqrt{2m + \frac{1}{4}}.$$

Proof. Proof given in lectures. $\square$

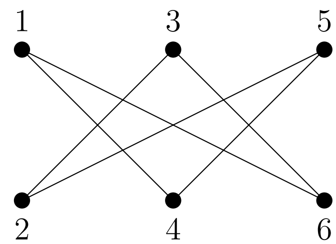
How should we colour a graph, in practice? There is a very simple but useful **greedy algorithm**:

Given a graph G, fix an ordering $v_1, v_2, \ldots, v_n$ on the vertices of G, and colour them one by one in this order. To start with, all vertices are uncoloured. At step i, where $i = 1, 2, \ldots, n$, colour vertex i with the first *available* colour: that is, the least positive integer which has not been used to colour any of the neighbours of v_i in $\{v_1, \ldots, v_{i-1}\}$.

Since v_i has at most $\Delta(G)$ neighbours in $v_1, \ldots, v_{i-1}$, this produces a k-colouring of G with $k \leq \Delta(G) + 1$. That is,

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

This greedy algorithm may produce wasteful colourings if it colours the vertices in a bad order. For example, applying the greedy algorithm to this graph with the given vertex labelling would use 3 colours, whereas the graph is bipartite.



This example can be extended to produce a bipartite graph for which the greedy algorithm will use r colours, for any integer $r \geq 4$.

Fact: $\chi(G) = \Delta(G) + 1$ if G is a complete graph or an odd cycle. (See Problem Sheet 4).

Theorem. 5.2.4 (Brooks, 1941)
Let G be a connected graph. If G is neither complete nor an odd cycle then $\chi(G) \leq \Delta(G)$.

Proof. Proof given in lectures. □

Calculating $\chi(G)$ (or more precisely, determining whether $\chi(G) \leq k$ for some fixed positive integer k) is a NP-complete problem.

One easy way to force a high chromatic number is to construct graphs with subgraphs isomorphic to K_r for large r. However, having large dense substructures is not a necessary condition for a high chromatic number. Using the probabilistic method, Erdős proved in 1959 that there exists graphs with arbitrarily high chromatic number and yet arbitrarily high girth. Such graphs are "locally 2-colourable" (if you look in a small local area, you see a tree!). This shows that high chromatic number can sometimes be a global phenomenon. We will prove Erdős' famous result at the end of this course.

5.3 Edge colourings

By considering a vertex of maximum degree, we see that the chromatic index $\chi'(G)$ satisfies $\chi'(G) \geq \Delta(G)$ for all graphs G.

Proposition. 5.3.1 (König, 1916)
If G is bipartite then $\chi'(G) = \Delta(G)$.

Proof. Proof given in lectures. □

Unlike $\chi(G)$, we can pin $\chi'(G)$ down to a very small range.

Theorem. 5.3.2 (Vizing, 1964)
Every graph G satisfies

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

We omit the proof: it use alternating path arguments which are very similar to those used in the proof of Brooks' Theorem (Theorem 5.2.4) and Proposition 5.3.1. If you are interested, you can read the proof in Diestel [3].

Even so, determining whether a given graph G has $\chi'(G) = \Delta$ (that is, G is "class 1") or $\chi'(G) = \Delta + 1$ (that is, G is "class 2") is an NP-complete problem.

5 Connectivity

The reference for this section is Diestel [3, Chapter 3].

Recall that a graph G is k-connected if $|G| > k$ and G cannot be disconnected by deleting fewer than k vertices.

We will prove an alternative characterisation, called Menger's Theorem (1927)): A graph is k-conneted if and only if any two vertices can be joined by k independent paths (that is, paths with no common internal vertices). First we consider 2-connectivity and 3-connectivity.

3.1 2-connected graphs

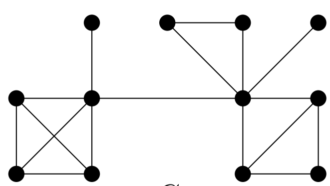
Let G be a graph. A maximal connected subgraph of G with no cut vertex is called a *block*. Every block of G is either a maximal 2-connected subgraph of G or a bridge or an isolated vertex. Conversely, every such subgraph of G is a block.

By maximality, different blocks of G overlap in at most one vertex, which must be a cut vertex in G. Hence every edge of G lies in a unique block, and G is the union of its blocks. (Blocks are the 2-connected ananalogues of components.)

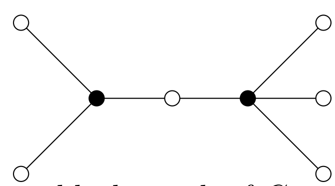
Let A be the set of cut vertices in G and let $\mathcal{B}$ be the set of the blocks of G. Form the bipartite graph on $A \cup \mathcal{B}$ with edge set

$$\{aB : a \in A,\ B \in \mathcal{B} \text{ and } a \in B\}.$$

This bipartite graph is the *block graph* of G. An example of a graph G and its block graph is shown below. The graph has 6 blocks and 2 cut vertices.



G



block graph of G

Proposition. 3.1.2
The block graph of a connected graph is a tree.

Proof. Exercise (see Problem Sheet 5). □

Let H be a subgraph of a graph G. An *H-path* is a path in G which intersects H only in its endvertices.

The following result is useful in inductive proofs for 2-connected graphs.

Proposition. 3.1.3
A graph is 2-connected if and only if it can be constructed from a cycle by successively adding H-paths (or "ears") to graphs H already constructed.

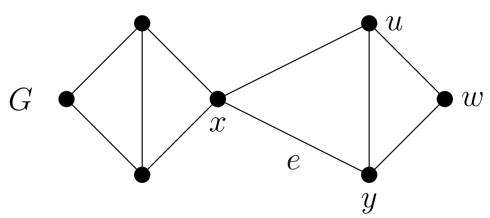
Proof. Proof given in lectures. $\square$

3.2 3-connected graphs

Let $e = xy \in E(G)$. Define the graph G/e by $G/e = (V', E')$ where

$$V' = (V - \{x, y\}) \cup \{v_e\},$$
$$E' = \{uw \in E(G) : \{u, w\} \cap \{x, y\} = \emptyset\} \cup \{v_e w : xw \in E(G) \text{ or } yw \in E(G)\}.$$

We say that G/e is formed by *contracting* the edge e in G. This creates a new vertex v_e which replaces the endvertices of e.



As illustrated in this example, if both x and y are neighbours of some vertex (here, u) then G/e has the edge $v_e u$, but only with multiplicity 1 (since G/e is a graph, not a multigraph).

Lemma.
Let G be a 3-connected graph with $|G| \geq 5$. Then G has an edge e such that G/e is 3-connected.

Proof. Proof given in lectures. $\square$

We can reverse this process to construct all 3-connected graphs, starting with K_4.

Theorem. 3.2.2 (Tutte, 1961)
A graph G is 3-connected if and only if there exists a sequence $G_0, G_1, \ldots, G_r$ of graphs such that

(i) $G_0 = K_4$ and $G_r = G$,

(ii) G_{i+1} has an edge xy with degrees $d(x), d(y) \geq 3$ such that $G_i = G_{i+1}/xy$, for $i = 0, \ldots, r - 1$.

The proof of this result is not too difficult, but we don't have time to cover it. See Diestel [3] if interested.

3.3 Menger's Theorem

Diestel describes Menger's Theorem as "one of the cornerstones of graph theory", and presents three proofs of this result. We will use his second proof.

Recall that if $A, B \subseteq V$ then an (A,B)-path is a path $P = x_0 \dots x_k$ such that $A \cap P = \{x_0\}$ and $B \cap P = \{x_k\}$. A set $S \subset V$ separating A from B in G is called an (A,B)-separator. (This means that every (A,B)-path intersects S, and in particular $A \cap B \subseteq S$.)

Let $\mathcal{P}$, $\mathcal{Q}$ be sets of disjoint (A,B)-paths in G. Say that $\mathcal{Q}$ *exceeds* $\mathcal{P}$ if the set of vertices in A which belong to paths in $\mathcal{P}$ is a *proper subset* of the set of vertices in A which belong to paths in $\mathcal{Q}$, and similarly for B.

If $P = x_0x_1 \cdots x_k$ then we write Px_i for the subpath $x_0 \cdots x_i$, and we write x_iP for the subpath $x_ix_{i+1} \cdots x_k$.

Theorem. 3.3.1 (Menger's Theorem, 1927)
Let $G = (V, E)$ be a graph and $A, B \subseteq V$. Then the minimum number of vertices separating A from B in G equals the maximum number of disjoint (A,B)-paths in G. (Here "disjoint" means *including their endpoints*.)

Proof. Proof given in lectures. □

Corollary. 3.3.5
Let a, b be distinct vertices of G.

(i) If $ab \notin E$ then the minimum number of vertices (distinct from a, b) separating a from b is equal to the maximum number of *independent* (a,b)-paths in G. (Recall that two paths are *independent* if they share no internal vertex.)

(ii) The minimum number of edges separating a from b in G equals the maximum number of edge-disjoing (a,b)-paths in G.

Proof. Proof given in lectures. □

Theorem. 3.3.5 (Global version of Menger's Theorem)

(i) A graph is k-connected if and only if it has order at least 2 and there are k independent paths between any two distinct vertices.

(ii) A graph is k-edge-connected if and only if it has k edge-disjoint paths betwen any two distinct vertices.

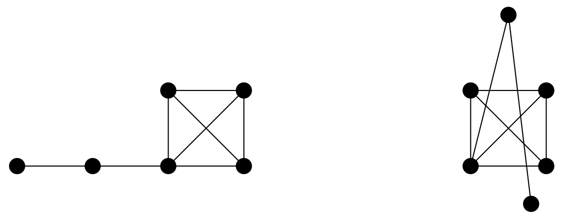
Proof. Proof given in lectures. □

Note: if you apply Menger's Theorem (Theorem 3.3.1) to a bipartite graph with bipartition A, B, you obtain the statement of König's Theorem (Theorem 2.1.1), since a maximum matching is a maximum set of disjoint (A, B)-paths, and a minimum vertex cover is a minimum set which separates A and B. Hence König's Theorem can be obtained as a special case of Menger's Theorem.

6 Planar graphs

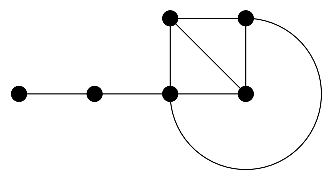
The reference for this section is Diestel [3, Chapter 4].

What is the best way to draw a graph on a page? Is the left picture clearer than the right one?



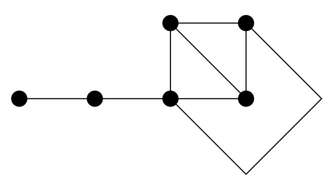
Which do you prefer?

A graph which is drawn in the plane so that no two edges meet except at common endvertices is called a *plane graph*. (We will define this concept more formally below.) An abstract graph which can be drawn in this way is called *planar*.



Graph drawing

A graph is drawn in the Euclidean plane $\mathbb{R}^2$ by representing each vertex by a point and each edge by a curve between these points. To avoid complications we restrict to curves which are *piecewise linear*. For example:



4.2 Plane graphs

We skip the proofs of some topological statements (since this isn't a course on topology, and the statements seem fairly straightforward.) You can consult Diestel [3, Chapter 4] if you wish to see these proofs.

An *arc* (or *polygonal arc*) is a subset of $\mathbb{R}^2$ which is the union of finitely many straight line segments, which is homeomorphic to $[0, 1]$ (so it doesn't cross itself).

A *plane graph* is a pair (V, E) of finite sets (with elements of V called *vertices* and elements of E called *edges*) such that

(i) $V \subseteq \mathbb{R}^2$,

(ii) Every edge is an arc between two vertices (that is, no loops);

(iii) Different edges have different sets of endvertices (that is, no repeated edges);

(iv) The interior of an edge contains no vertex and no point of any other edge.

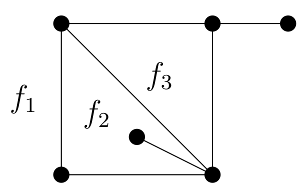
Here the *interior* of an edge/arc e, denoted by $\mathring{e}$, is te arc minus its endpoints: if e is the arc from x to y then $\mathring{e} = e - \{x, y\}$.

A plane graph defines a graph G in a natural way. If no confusion arises we will use the name G for the abstract graph, the plane graph, and the *point set*

$$V \cup \left(\bigcup_{e \in E} e \right) \subseteq \mathbb{R}^2.$$

The point set of a plane graph G is a closed set in $\mathbb{R}^2$, and $\mathbb{R}^2 - G$ is open. Two points in an open set O are *equivalent* if they can be linked by an arc in O. This defines an equivalence relation. The equivalence classes of $\mathbb{R}^2 - G$ are open connected regions, called the *faces* of G.

For example, the following plane graph has three faces.



Since G is bounded (that is, it lies within some sufficiently large disc $D \subseteq \mathbb{R}^2$), exactly *one* face of G is unbounded: it is the face that contains $\mathbb{R}^2 - D$. We call this the *outer face* of G. All other faces of G are called *inner faces*. (In the example above, f_1 is the outer face.)

Let $F(G)$ be the set of faces of G. The *boundary* of a face f is called the *frontier* of f. (It is the set of all points $y \in \mathbb{R}^2$ such that every neighbourhood of y meets both f and $\mathbb{R}^2 - f$.)

Lemma. 4.2.1.
Let G be a plane graph with subgraph $H \subseteq G$ and face $f \in F(G)$.

(i) There is a face $f' \in F(H)$ which contains f (that is, $f \subseteq f'$).

(ii) If the frontier of f lies in H then $f' = f$.

Proof. Proof given in lectures. □

We omit the proof of the next two results, but you can find them in Diestel [3].

Lemma. 4.2.2
Let G be a plane graph and let e be an edge of G.

(i) If X is the frontier of a face of G then either $e \subseteq X$ or $X \cap \mathring{e} = \emptyset$.

(ii) If e lies on a cycle $C \subseteq G$ then e lies on the frontier of exactly two faces of G, and these are contained in the distinct faces of C.

(iii) If e does not lie on a cycle then e lies on the frontier of exactly one face of G.

Corollary. 4.2.3
The frontier of a face of a plane graph G is always the point set of a subgraph of G.

The subgraph of G whose point set is the frontier of a face f is said to *bound* f and is called the *boundary* of f. Denote this subgraph by $G[f]$. A face is said to be *incident* with the vertices and edges of its boundary.

By Lemma 4.2.2(ii), every face of G is also a face of its boundary.

Proposition. 4.2.4
A plane forest has exactly one face.

Proof. Exercise: Use induction on the number of edges. □

Lemma. 4.2.5
If a plane graph has two distinct faces with the same boundary then the graph is a cycle.

Proof. Proof given in lectures. □

Proposition. 4.2.6
In a 2-connected plane graph, every face is bounded by a cycle.

Proof. Proof given in lectures. □

A plane graph G is *maximally plane* (or just *maximal*) if we cannot add a new edge to form a new plane graph G' with $V(G') = V(G)$ such that $E(G')$ strictly contains $E(G)$.

Call G a *plane triangulation* if every face of G (including the outer face) is bounded by a triangle.

Proposition. 4.2.8
A plane graph of order at least 3 is maximally plane if and only if it is a plane triangulation.

Proof. Proof given in lectures. □

We can now prove Euler's Formula for the plane. (Higher genus versions of this formula exist, e.g. for graphs drawn on a torus.)

Theorem. 4.2.9 (Euler's Formula, 1752)
Let G be a connected plane graph with n vertices, m edges and ℓ faces. Then

$$n - m + \ell = 2.$$

Proof. Proof given in lectures. □

Warning: Euler's Formula only works for *connected* graphs. Somehow this is easy to overlook.

Corollary. 4.2.10
A plane graph with $n \geq 3$ vertices has at most $3n - 6$ edges. Every plane triangulation has $3n - 6$ edges.

Proof. Proof given in lectures. □

As a consequence, we see that K_5 is not planar: K_5 has 5 vertices and 10 edges, but

$$10 > 3 \cdot 5 - 6.$$

Hence by Corollary 4.2.10, the complete graph K_5 is not planar.

Similarly, $K_{3,3}$ is not planar. This does not follow immediately from Corollary 4.2.10, but note that $K_{3,3}$ is bipartite and hence contains no triangle. Since $K_{3,3}$ is 2-connected, every face is bounded by a cycle of at least 4 edges. Adapting the proof of Corollary 4.2.10 to this case shows that in any planar graph with no triangles, $2m \geq 4\ell$, and then Euler's Formula implies that

$$m \leq 2n - 4.$$

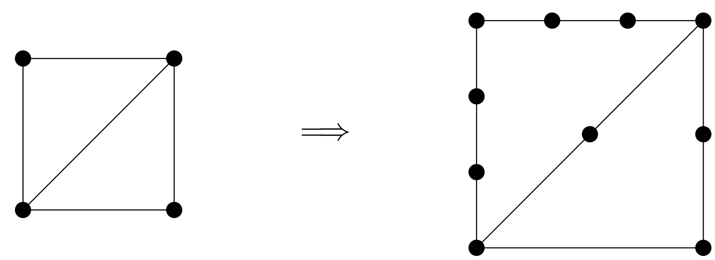
(Check the details yourself.) However, $K_{3,3}$ has 6 vertices and 9 edges, but

$$9 > 2 \cdot 6 - 4.$$

Therefore $K_{3,3}$ is not planar.

Note: **Kuratowski's Theorem** is one of the highlights of graph theory. We now describe this theorem briefly.

A *subdivision* of a graph G is obtained by replacing each edge of G by an independent path between its endvertices. For example:



Kuratowski's Theorem (1930) says that a graph G is planar if and only if no subgraph of G is a subdivision of K_5 or $K_{3,3}$.

Wagner (1937) reformulated Kuratowski's theorem in terms of *graph minors*, rather than subdivisions. I will not define graph minors here, but note that this is a very active topic of research.

Unfortunately, the proof of Kuratowski's Theorem is beyond the scope of this course.

5.1 Colouring maps

This is the famous Four Colour Theorem:

Theorem. 5.1.1
Every planar graph is 4-colourable. (That is, there exists a proper 4-colouring of the vertices of any planar graph.)

A very brief history:

- This question arose in 1852 for the problem of colouring maps of (contiguous") countries with the smallest number of colours. This corresponds to colouring *faces* of a plane graph. By planar graph duality (which exchanges vertices with faces), this is equivalent to asking whether planar graphs have a (vertex) 4-colouring.

- In 1976, Appel and Haken gave a "computer-assisted" proof, with 1476 cases to check. (Some mathematicians refuse to acknowledge that this constitutes a proof, since it cannot be checked by hand.)

- In 1996, Robertson, Sanders, Seymour and Thomas reduced the number of "unavoidable configurations" to be checked down to 633. This is still a computer-assisted proof, however.

Do you think that a computer-assisted proof is a valid mathematical proof? You can read more about the Four Colour Theorem on Wikipedia

http://en.wikipedia.org/wiki/Four_color_theorem

or on Robin Thomas' page http://www.math.gatech.edu/~thomas/FC/

Given the amount of cases to check we will not prove the Four Colour Theorem here. However, we will do the next best thing.

Proposition. 5.1.2
Every planar graph is 5-colourable.

Proof. Proof given in lectures. □

To finish this section, we mention (without proof) this nice result.

Theorem. 5.1.3 (Grötzsch, 1959)
Every planar graph which does not contain a triangle is 3-colourable.

7 Ramsey Theory

The main references for this chapter are Diestel [3, Chapter 9] and Bollobás [2, Chapter 6], but we also use a proof from Alon & Spencer [1, Chapter 1].

How many people do need at a party before you are *guaranteed* to find either 3 people who already knew each other, or 3 people who were all strangers to one another (when they arrived)?

If there are n people at the party, you can model this using the complete graph K_n. Each vertex represents a person, and the edge ij is coloured red if person i and person j already knew each other before the party, or coloured blue if they did not.

How large must n be before, in *any* red-blue colouring of the edges of K_n, you can find either a red triangle or a blue triangle?

(Note: These are *not* "proper" edge colourings in the sense discussed earlier, as incident edges may have the same colour.)

More generally, for integers $s, t \geq 2$, let $R(s,t)$ be the least positive integer n such that any red-blue colouring of K_n has either a red copy of K_s or a blue copy of K_t. Set $R(s,t) = \infty$ if no such n exists. The numbers $R(s,t)$ are called *Ramsey numbers*. Write $R(s)$ instead of $R(s,s)$ (this is the "diagonal" case).

Exercise: Show that $R(3) = 6$ and that $R(s,2) = R(2,s) = s$ for all $s \geq 2$. See Problem Sheet 6.

Upper bounds

George Szekeres was a Hungarian mathematician who was a Professor of Pure Mathematics at UNSW for many years, from 1963 through his retirement in 1975 and beyond. (He was still coming into his office regularly after I joined UNSW in 2003, aged more than 90!). Szekeres proved these very nice upper bounds on the Ramsey numbers with Paul Erdős, a very prolific and eccentric mathematician. To read more:

http://en.wikipedia.org/wiki/George_Szekeres
http://en.wikipedia.org/wiki/Paul_Erdos

The reference for the next theorem is Bollobás [2, Chapter 6, Theorem 1].

Theorem. (Erdős & Szekeres, 1935)
For all integers $s, t \geq 2$, the Ramsey number $R(s,t)$ is finite. If $s > 2$ and $t > 2$ then

$$R(s,t) \leq R(s-1,t) + R(s,t-1) \tag{3}$$

and

$$R(s,t) \leq \binom{s+t-2}{s-1}. \tag{4}$$

Proof. Proof given in lectures. □

Very few Ramsey numbers are known precisely:

$$\begin{array}{rrr} R(3)=6, & R(3,4)=9, & R(3,5)=14, \\ R(3,6)=18, & R(3,7)=23, & R(3,8)=28, \\ R(3,9)=36, & R(4)=18, & R(4,5)=25. \end{array}$$

There is a nice anecdote about how difficult it is to compute Ramsey numbers: search for the word "alien" on this Wikipedia page:

http://en.wikipedia.org/wiki/Ramsey's_theorem

Since Ramsey numbers are hard to compute, our interest turns to asymptotics, especially for the diagonal Ramsey numbers $R(s)$: how fast do they grow, as $s \to \infty$? A key tool for asymptotic analysis of combinatorial quanties is *Stirling's inequalities*: for any positive integer n,

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n \le n! \le e^{1/(12n)}\sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

From this we can obtain the weaker upper bound

$$n! \le 1.1\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

which we will use below. Stirling's inequalities also imply that

$$\binom{n}{k} \le \frac{n^k}{k!} \le \left(\frac{en}{k}\right)^k$$

for all positive integers k and n.

We now derive an asymptotic upper bound for $R(s)$, starting from (4). If $s \ge 2$ then

$$\begin{aligned} R(s) \le \binom{2s-2}{s-1} = \frac{(2s-2)!}{(s-1)!^2} &\le \frac{1.1 \times \sqrt{4\pi(s-1)}\,\left(\frac{2s-2}{e}\right)^{2s-2}}{2\pi(s-1)\,\left(\frac{s-1}{e}\right)^{2(s-1)}} \\ &= \frac{1.1 \times 2^{2(s-2)}}{\sqrt{\pi(s-1)}} \\ &\le \frac{4^s}{4\sqrt{s}}. \end{aligned}$$

(Check!) Thomason (1988) gave an improved bound

$$R(s) \le \frac{4^s}{s},$$

which was the best known upper bound until 2009, when Conlon proved a more complicated result which implies that the ratio $4^s/R(s)$ is larger than any polynomial in s. That is, for any positive integer a there exists a positive constant C_a such that

$$R(s) \le C_a \frac{4^s}{s^a}.$$

Lower bounds

We use the probabilistic method to prove lower bounds on Ramsey numbers. The reference for the next theorem is Alon & Spencer [1, Proposition 1.1.1].

Theorem. (Erdős, 1947)
If

$$\binom{n}{s} 2^{1-\binom{s}{2}} < 1$$

then $R(s) > n$. Hence $R(s) > \lfloor 2^{s/2} \rfloor$ for $s \geq 3$.

Proof. Proof given in lectures. $\square$

Note, it is possible to give a sharper analysis and obtain the lower bound

$$R(s) > \frac{s}{\sqrt{2e}} 2^{s/2}$$

for $s \geq 3$. Further improvements in the constant are possible by applying the Lovász Local Lemma. (See Problem Sheet 7.) But combining the best known bounds, we see that there is still a huge gap in the exponent of $R(s)$, despite a lot of effort.

$$\ln \sqrt{2} \leq \lim_{s \to \infty} \frac{\ln R(s)}{s} \leq \ln 4.$$

Tim Gowers, the Fields Medalist, wrote about this problem in his essay "The Two Cultures of Mathematics", which is well worth a read.

https://www.dpmms.cam.ac.uk/~wtg10/2cultures.pdf

If you are looking for good news regarding Ramsey numbers, consider $R(3,t)$. Atjai, Komlós and Szemerédi proved in 1980 that

$$R(3,t) = O\left(\frac{t^2}{\log t}\right).$$

This means that there exists a positive constant C such that

$$R(3,t) \leq C \frac{t^2}{\log t}$$

for all sufficiently large positive integers t. Shearer improved this in 1983, showing that

$$R(3,t) < (1 + o(1)) \frac{t^2}{\log t}.$$

Here $o(1)$ stands for some function of t which tends to zero as t tends to infinity.

Erdős had proved a lower bound in 1961 which was too small by a factor of some constant times $\log t$. This was the best known lower bound until a 1995 result by Kim, who proved a lower bound of the right asymptotic order. Then in 2013, Bohman and Keevash proved that

$$R(3,t) > \left(\tfrac{1}{4} - o(1)\right) \frac{t^2}{\log t},$$

which is within a factor 4 of Shearer's upper bound, asymptotically. Great!

Graph Ramsey Theory

The main reference for this subsection is Bollobás [2, Chapter 6]. You can also find some material in Diestel [3, Section 9.2].

Let H_1, H_2 be fixed graphs with no isolated vertices, and let $R(H_1, H_2)$ be the least positive integer n such that in every red-blue colouring of the edges of K_n, there is either a red copy of H_1 or a blue copy of H_2. Write $R(H) = R(H,H)$ and note that $R(K_s, K_t) = R(s,t)$, the Ramsey number. So the *graph Ramsey numbers* $R(H_1, H_2)$ generalise the Ramsey numbers.

There are very few exact results known, but we will prove one. Write ℓK_2 for a set of ℓ independent edges.

Theorem. (Bollobás [2, Chapter 6, Theorem 10])
For $\ell \geq 1$ and $p \geq 2$,

$$R(\ell K_2,\ K_p) = 2\ell + p - 2.$$

Proof. Proof given in lectures. □

For a graph G, let $c(G)$ be the number of vertices in the largest component of G, and let $u(G)$ be the *chromatic surplus* of G, which is the minimum size of the smallest colour class of G, taken over all $\chi(G)$-colourings of G (here we mean proper vertex colourings with $\chi(G)$ colours). Note that

$$u(C_{2k}) = k, \qquad u(C_{2k+1}) = 1.$$

Theorem. (Bollobás [2, Chapter 6, Theorem 11])
For all graphs H_1, H_2 with no isolated vertices, we have

$$R(H_1, H_2) \geq (\chi(H_1) - 1)\,(c(H_2) - 1) + u(H_1).$$

In particular, if H_2 is connected then

$$R(H_1, H_2) \geq (\chi(H_1) - 1)\,(|H_2| - 1) + 1.$$

Proof. Proof given in lectures. □

As a sample application, let $H_2 = T$ be a tree with t vertices, where $t \geq 2$. The above theorem says that

$$R(K_s, T) \geq (s-1)(t-1) + 1$$

since T is connected and $\chi(K_s) = s$. In fact for any tree T we actually have *equality*:

$$R(K_s, T) = (s-1)(t-1) + 1,$$

but we will not prove this here. (The proof involves Corollary 1.5.4.)

8 Random graphs

The main reference for this section is Diestel [3, Chapter 11].

Let Ω_n be the set of all graphs with vertex set $\{1, 2, \dots, n\}$. Then $|\Omega_n| = 2^{\binom{n}{2}}$. Consider the uniform probability space on Ω_n. Recall that we can choose a random element from Ω_n by flipping a fair coin, independently for each pair of distinct vertices $\{i, j\}$, and letting $ij \in E$ if and only if the coin comes up heads. This is the *uniform model* of random graphs. We saw earlier that the expected number of edges in this model is $\frac{1}{2}\binom{n}{2}$ (see Problem Sheet 3).

Now take a coin which comes up heads with probability p, where $p \in [0, 1]$ is fixed, and perform the same procedure: for each pair of distinct vertices $\{i, j\}$, independently, flip the coin and let $ij \in E$ if and only if the coin comes up heads. Then $\Pr(ij \in E) = p$, independently for each $i \neq j$. This gives a random graph model called the *binomial model* and denoted $G(n, p)$. The binomial model was introduced by Gilbert in 1959. Note that $G(n, \frac{1}{2})$ is the uniform model.

We write $G \in G(n, p)$ to mean that G is a random graph chosen from the binomail model. For a fixed $G_0 \in \Omega_n$, the probability that the random graph G equals G_0 is

$$\Pr(G = G_0) = p^{|E(G_0)|} \, (1 - p)^{\binom{n}{2} - |E(G_0)|}$$

which depends only on $|E(G_0)|$. The proof of this follows from Problem Sheet 3, Question 3: check this!

You should also make sure that you can see why, for $G \in G(n, p)$, the expected number of edges of G is $p\binom{n}{2}$.

Note: for a fixed $p \in [0, 1]$ we have defined a *sequence* of probability spaces,

$$\left(G(n, p)\right)_{n \in \mathbb{Z}^+} .$$

We can also let p be a function of n, where $p(n) \in [0, 1]$ for all $n \in \mathbb{Z}^+$. This gives the sequence of probability spaces

$$\left(G(n, p(n))\right)_{n \in \mathbb{Z}^+} .$$

For example, we could take $p = 1/n$. (We often just write p instead of $p(n)$, even when p depends on n.) For example, when $G \in G(n, \frac{1}{n})$, the expected number of edges of G is $\frac{1}{n}\binom{n}{2} = \frac{n-1}{2}$.

Recall that $\omega(G)$ is the clique number of G, and $\alpha(G)$ is the independence number of G.

Lemma. 11.1.2
Let $G \in G(n,p)$. Then for any integer $k \geq 2$,

$$\Pr(\omega(G) \geq k) \leq \binom{n}{k} p^{\binom{k}{2}},$$
$$\Pr(\alpha(G) \leq k) \leq \binom{n}{k} (1-p)^{\binom{k}{2}}.$$

Proof. Proof of the first statement given in lectures. Proof of the second statement is an exercise (Problem Sheet 8). □

For $a \in \mathbb{R}$ and $r \in \mathbb{N}$, let

$$(a)_r = a(a-1)\cdots(a-r+1)$$

denote the *falling factorial*.

Lemma. 11.1.5
Let $k \geq 3$ be an integer. The expected number of k-cycles in $G \in G(n,p)$ is

$$\frac{[n]_k}{2k} p^k.$$

Proof. Proof given in lectures. □

Exercise: find the expected number of k-paths in $G \in G(n,p)$. (See Problem Sheet 8.)

For a given graph property $\mathcal{P}$, we can ask how $\Pr(G \in \mathcal{P})$ behaves for $G \in G(n,p)$ as $n \to \infty$. If $\Pr(G \in \mathcal{P}) \to 1$ as $n \to \infty$ then we say that $G \in \mathcal{P}$ holds *asymptotically almost surely*, abbreviated to "a.a.s.". (Note, Diestel says "*almost every $G \in G(n,p)$ has $\mathcal{P}$*".)

Proposition. 11.3.1
For fixed $p \in (0,1)$ and every graph H, a.a.s. $G \in G(n,p)$ has an induced subgraph which is isomorphic to H.

Proof. Proof given in lectures. □

Given $i, j \in \mathbb{N}$, let $\mathcal{P}_{ij}$ be the property that given any disjoint vertex sets U, W with $|U| \leq i$ and $|W| \leq j$, the graph contains a vertex $v \notin U \cup W$ that is adjacent to all vertices in U but to none in W.

Lemma. 11.3.2
For every constant $p \in (0,1)$ and all $i, j \in \mathbb{N}$, let $G \in G(n,p)$. Then a.a.s. $G \in \mathcal{P}_{ij}$.

Proof. Proof given in lectures. $\square$

Using this result we can prove that many natural graph properties hold with probability tending to 1 in $G(n,p)$.

Corollary. 11.3.3
For every constant $p \in (0,1)$ and all $k \in \mathbb{N}$, a.a.s. $G \in G(n,p)$ is k-connected.

Proof. Proof given in lectures. $\square$

Most graphs have quite a high chromatic number. As elsewhere in this course, ln means the natural logarithm

Proposition. 11.3.4
For every constant $p \in (0,1)$ and all $\varepsilon > 0$, a.a.s. $G \in G(n,p)$ satisfies

$$\chi(G) \geq \frac{\ln(1/q)\, n}{(2+\varepsilon)\ln n},$$

where $q = 1 - p$.

Proof. Proof given in lectures. $\square$

Now we start to lead up to the proof of Erdős' result from 1959, which says that there exist graphs with arbitrarily high girth and arbitrarily high chromatic number. This surprising result brought the probabilistic method to many people's attention.

Recall that the girth of a graph is the length of its smallest cycle.

Lemma. 11.2.1
Let k be a positive integer and let $p = p(n)$ be a function of n such that $p(n) \in [0,1]$ and

$$p(n) \geq \frac{6k\, \ln n}{n}$$

for sufficiently large n. Then for $G \in G(n,p)$, a.a.s. $\alpha(G) < \frac{n}{2k}$.

Proof. Proof given in lectures. $\square$

Recall Markov's inequality: if $X : \Omega \to \mathbb{N}$ is a nonnegative integer random variable on a set Ω and $k > 0$, then

$$\Pr(X \geq k) \leq \frac{\mathbb{E}X}{k}.$$

We can now prove Erdős' famous result, which is a masterful application of the probabilistic method.

Theorem. 11.2.2. (Erdős, 1959)
For every integer $k \geq 3$ there exists a graph H with girth $g(H) \geq k$ and chromatic number $\chi(H) \geq k$.

Proof. Proof given in lectures. □

References

[1] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley, 2000.

[2] B. Bollobás, *Modern Graph Theory*, Cambridge University Press, Cambridge, 1998.

[3] R. Diestel, *Graph Theory* (4th edn.), Springer, Berlin, 2010.
Available online: http://diestel-graph-theory.com