

Graph Theory

Adam Kelly (ak2316@cam.ac.uk)

August 22, 2021

Graph theory is an area of combinatorics which has lots of fun problems and plenty of interesting theorems.

This article constitutes my notes for the ‘Graph Theory’ course, held in Lent 2021 at Cambridge. These notes are *not a transcription of the lectures*, and differ significantly in quite a few areas. Still, all lectured material should be covered.

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 2 |
| 1.1 | Definitions | 2 |
| 1.1.1 | Common Graphs | 3 |
| 1.1.2 | Subgraphs | 4 |
| 1.1.3 | Graph Isomorphism | 5 |
| 1.1.4 | Connectivity | 5 |
| 1.1.5 | Edges and Distance | 7 |
| 1.2 | Trees | 8 |
| 1.3 | Bipartite Graphs | 11 |
| 2 | Hall’s Theorem | 13 |
| 2.1 | Matchings | 13 |
| 2.2 | Matching in Bipartite Graphs | 14 |
| 2.3 | Hall’s Theorem | 15 |
| 2.4 | Corollaries of Hall’s Theorem | 16 |
| 3 | Connectivity | 19 |
| 3.1 | Measuring Connectivity | 19 |
| 3.2 | Preparing for Menger’s Theorem | 21 |
| 3.3 | Actually Menger’s Theorem | 22 |
| 3.4 | Edge Connectivity | 23 |
| 4 | Planar Graphs | 25 |
| 4.1 | Defining Planar Graphs | 26 |
| 4.2 | What Graphs are Planar? | 27 |
| 5 | Graph Colouring | 29 |
| 5.1 | Basic Concepts | 29 |
| 5.2 | Brooks’ Theorem | 31 |
| 5.3 | Colouring Planar Graphs | 32 |
| 5.4 | Colouring Graphs on (Other) Surfaces | 33 |

| | | |
|----------|--|-----------|
| 5.5 | Edge Colouring | 35 |
| 6 | Extremal Graph Theory | 36 |
| 6.1 | Eulerian Circuits and Hamiltonian Cycles | 36 |
| 6.2 | Complete Subgraphs and Turan's Theorem | 39 |
| 6.3 | The Zarankiewicz Problem | 41 |
| 6.4 | The Erdős-Stone Theorem | 42 |
| 7 | Ramsey Theory | 43 |
| 7.1 | Ramsey Numbers | 44 |
| 7.2 | Infinite Ramsey | 45 |
| 7.3 | Ramsey Numbers for Hypergraphs | 46 |
| 7.4 | Lower Bounds for Ramsey Numbers | 46 |
| 8 | The Probabilistic Method | 48 |
| 8.1 | Random Graphs | 48 |
| 8.2 | Random Graphs and Chromatic Number | 49 |
| 8.3 | Structure of Random Graphs | 50 |
| 9 | Algebraic Methods in Graph Theory | 53 |
| 9.1 | A Motivating Problem | 53 |
| 9.2 | The Adjacency Matrix | 54 |
| 9.3 | Moore Graphs | 56 |

§1 Introduction

For many people, 'Graph Theory' is a first course in combinatorics. It's an area with a big focus on problem solving, and it can give a perspective on many other areas of mathematics.

§1.1 Definitions

We will begin this handout on graph theory naturally by defining what a graph is.

Definition 1.1 (Graph)

A **graph** is an ordered pair $G = (V, E)$ where V is the set of **vertices**, and $E \subseteq \{\{x, y\} \mid x, y \in V, x \neq y\}$ is a set of unordered pairs of vertices called **edges**.

We have a natural way of drawing a graph. For each vertex we have a point in the plane, and for each edge we draw a line between the corresponding pair of vertices.

Example 1.2 (Example of a Graph)

The ordered pair (V, E) where $V = \{1, 2, \dots, 6\}$ and $E = \{\{1, 2\}, \{2, 3\}, \dots, \{5, 6\}\}$ is a graph.



This graph is known as P_6 , a path on 6 vertices.

§1.1.1 Common Graphs

There are some graphs that will appear repeatedly when working on problems involving graph theory, and we will define them now.

Definition 1.3 (Path)

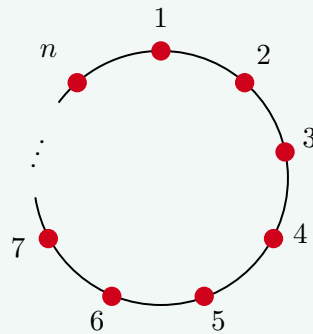
We define P_n to be the graph $V = \{1, \dots, n\}$, $E = \{\{1, 2\}, \{2, 3\}, \dots, \{n-1, n\}\}$ as shown.



We call this a **path** on n vertices, and say it has **length** $n - 1$.

Definition 1.4 (Cycle)

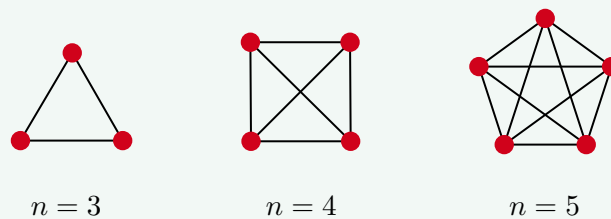
We define C_n (for $n \geq 3$) to be the graph $V = \{1, \dots, n\}$, $E = \{\{1, 2\}, \dots, \{n-1, n\}, \{n, 1\}\}$ as shown.



We call this the **cycle** on n vertices.

Definition 1.5 (Complete Graph)

The **complete graph** on n vertices K_n is the graph $\{1, \dots, n\}$ and $E = \{\{i, j\} \mid i \neq j \in V\}$.



Note that there is an edge between every pair of vertices.

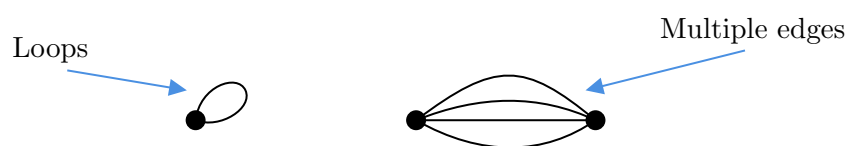
Definition 1.6 (Empty Graph)

We define the **empty graph** on n vertices \overline{K}_n to have $V = \{1, \dots, n\}$ but $E = \emptyset$.

Remark. In our definition of a graph, we *don't allow*¹ loops, and there *cannot* be

¹These limitations are inherent in our definition, where we use sets rather than multisets.

multiple edges between the same set of vertices.



You can define graphs where such things are allowed, but for now we will outlaw them. We also note that edges are *unordered pairs*, so for now edges have no direction.

To be slightly more succinct, we will use some shorthand notation.

Notation. If $G = (V, E)$ is a graph, and we have some edge $\{x, y\} \in E$, we will denote it by xy . We will also define $|G| = |V|$, and $e(G) = |E|$.

Example 1.7 (Vertices and Edges of K_n)

Consider the graph K_n . We have $|K_n| = n$, and $e(K_n) = \binom{n}{2}$, as there is an edge between any pair of vertices.

§1.1.2 Subgraphs

Now we will define the notion of a *subgraph*, in the natural way.

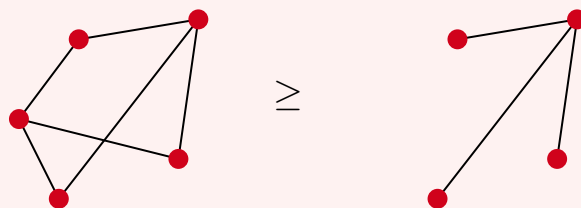
Definition 1.8 (Subgraph)

We say that $H = (V', E')$ is a **subgraph** of $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$.

Informally, H is a subgraph of G if we can remove vertices and edges from G to get H . Let's look at some examples.

Example 1.9 (Example of a Subgraph)

The graph on the right is a subgraph of the graph on the left.



We are also going to use some notation for removing an edge or a vertex from a graph. Of course, when removing a vertex you also have to remove the edges connecting to it.

Notation (Adding/Removing Vertices & Edges). For an edge xy or a vertex x , we define $G - xy$ to be the graph G with the edge xy removed, and $G - x$ to be G with vertex x removed, along with all edges incident to x . We will also define $G + xy$ to be G with the edge xy , and $G + x$ to be G with the vertex x .

An easy way to get a subgraph is by taking a subset of the vertices and seeing what edges you get from the original graph.

Definition 1.10 (Inducted Subgraph)

If $G = (V, E)$ is a graph and $X \subseteq V$, the **subgraph induced by X** is defined to be $G[X] = (X, \{xy \in E \mid x, y \in X\})$.

§1.1.3 Graph Isomorphism

Now that we have defined graphs, it's natural to define some notion of isomorphism.

Definition 1.11 (Graph Isomorphism)

Let $G = (V, E)$ and $H = (V', E')$ be graphs. We say that $f : V \rightarrow V'$ is a **graph isomorphism** if $f(u)f(v) \in E' \iff uv \in E$.

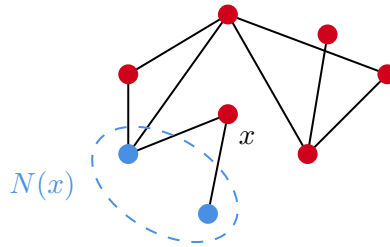
If there is a graph isomorphism between G and H then we say they are **isomorphic**.

Now for the following discussion, fix some graph $G = (V, E)$, and let $x \in V$.

Definition 1.12 (Neighbourhood)

If $xy \in E$, then we say that x and y are **adjacent**. We define the **neighborhood** of x to be the set $N(x) = \{y \in V \mid xy \in E\}$ of all vertices adjacent to x .

Note that as in the diagram below, x is not in its own neighborhood.

**Definition 1.13 (Degree)**

We define the **degree** of a vertex x to be $d(x) = |N(x)|$. This is equal to the number of edges that are incident to x .

Definition 1.14 (Regularity)

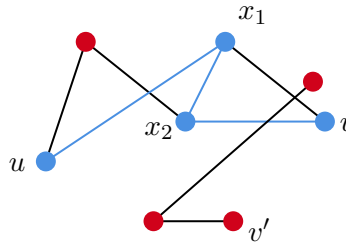
A graph G is said to be **regular** if all of the degrees are the same. We say G is k -regular if $d(x) = k$ for all $x \in V$.

Example 1.15 (Regular and Non-Regular Graphs)

The graphs K_n is $n - 1$ regular, and C_n is 2-regular. The graph P_n is not regular.

§1.1.4 Connectivity

We now want to define some notion of *connectivity*, where a vertex u is connected to vertex v if you can follow some path in the graph to get from u to v .



For example, in the graph above we want to say somehow that u and v are connected, but u and v' are not. To do this, we will introduce some more definitions.

Definition 1.16 (uv Path)

A **uv path** is a sequence x_1, x_2, \dots, x_l where x_1, \dots, x_l are distinct, $x_1 = u$, $x_l = v$ and $x_i x_{i+1} \in E$.

In the example above, ux_1x_2v is a uv path.

The slight subtlety in this condition is the *distinctness* condition. For example, if $x_1 \dots x_l$ is a uv path and $y_1 \dots y_{l'}$ is a vw path, then $x_1 \dots x_l y_1 \dots y_{l'}$ may *not* be a uw path since we may have reused an edge. Of course, we can just not reuse edges by avoiding cycles.

Proposition 1.17 (Joining Paths)

If $x_1 \dots x_l$ is a uv path and $y_1 \dots y_{l'}$ is a vw path, then $x_1 \dots x_l y_1 \dots y_{l'}$ contains a uw path.

Proof. Choose a minimal subsequence $w_1 \dots w_r$ of $x_1 \dots x_l y_1 \dots y_{l'}$ such that

1. $w_i w_{i+1} \in E$.
2. $w_1 = u$, $w_r = w$.

We now claim that $w_1 \dots w_r$ is a uw path. If this was not the case, then it must fail on distinctness, so there would exist some z such that the sequence is

$$w_1 \dots w_a z w_{a+2} \dots w_b z w_{b+2} \dots w_r,$$

but now note that

$$w_1 \dots w_a z w_{b+2} \dots w_r$$

also satisfies the conditions for the subsequence, but is strictly shorter length. This contradicts the minimality condition. \square

Now given $G = (V, E)$, let's define an equivalence relation \sim on V , where

$$x \sim y \iff \text{there exists an } xy \text{ path in } G.$$

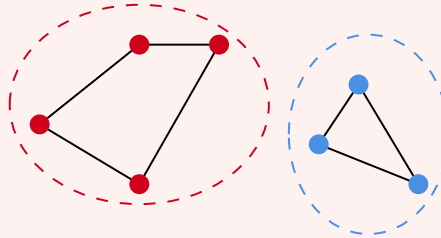
Proposition 1.18

\sim is an equivalence relation.

Proof. Note that \sim is reflexive and symmetric, and we get transitivity from our previous proposition. \square

Example 1.19

In the graph below, the vertices that are the same colour are in the same equivalence class under \sim .



Definition 1.20 (Connected Graph)

If there is a path between any two vertices in G then we say that G is **connected**.

Definition 1.21 (Connected Components)

We call the equivalence classes of \sim on G the **components** or **connected components** of G .

§1.1.5 Edges and Distance

We can also introduce some useful definitions relating to the edges of a graph.

Definition 1.22 (Minimum/Maximum Degree)

Let G be a graph. The **maximum degree** of G , $\Delta(G)$ is defined to be $\Delta(G) = \max_{x \in V} d(x)$. Similarly, we define the **minimum degree** of G , $\delta(G)$ to be $\delta(G) = \min_{x \in V} d(x)$.

In a k -regular graph as mentioned above, we have $\Delta(G) = \delta(G) = k$.

Definition 1.23 (Graph Distance)

Let $G = (V, E)$ be a graph. The associated **graph distance** $d : V \times V \rightarrow \mathbb{R}^{\geq 0} \cup \{\infty\}$ is defined so that $d(x, y)$ is the minimum path length from x to y if it exists, and ∞ otherwise.

Proposition 1.24 (Graph Distance is a Metric)

Let $G = (V, E)$ be a connected graph and let d be the associated graph distance. Then (V, d) defines a metric space.

Proof Sketch. We have $d(x, x) = 0$, and $d(x, y) = d(y, x)$ (taking the shortest path

in the opposite direction), and $d(x, z) \leq d(x, y) + d(y, z)$ as we can find path from x to z by taking paths from x to y and y to z and adjoining them, and this puts an upper bound on $d(x, z)$. \square

§1.2 Trees

We will now discuss a special class of graph called *trees*. This class is quite restrictive (yet is quite useful), and they have some nice properties.

To define what a tree is, we first need a notion of when a graph is acyclic.

Definition 1.25 (Acyclic)

A graph G is said to be **acyclic** if it does not contain any subgraph isomorphic to a cycle, C_n .

Example 1.26 (Example of Acyclic/Non-Acyclic Graphs)

In the example below, the two graphs are both *acyclic*.



Two *non-acyclic* graphs are shown below. The subgraphs isomorphic to C_4 and C_3 are highlighted.

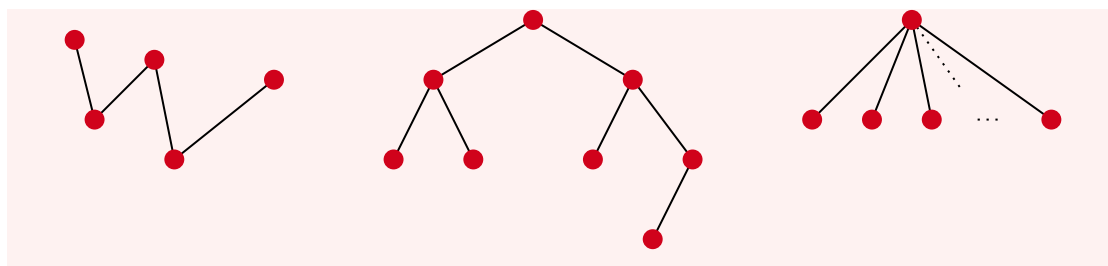


Definition 1.27 (Tree)

A **tree** is a connected, acyclic graph.

Example 1.28 (Examples of Trees)

The following three graphs are trees.



Proposition 1.29 (Characterising Trees)

The following are equivalent.

- (a) G is a tree.
- (b) G is a maximal acyclic graph (adding any edge creates a cycle).
- (c) G is a minimal connected graph (removing any edge disconnects the graph).

Proof. (a) \implies (b). By definition G is acyclic. Let $x, y \in V$ such that $xy \notin E$. As G is connected, there is an xy path P . So xPy then defines a cycle.

(b) \implies (a). By definition G is acyclic. So for a contradiction assume G is not connected and let x, y be vertices from different components. Now note $G + xy$ is acyclic, but this contradicts the claim that G is maximally acyclic.

(a) \implies (c). By definition G is connected. Suppose, for a contradiction, that there exists some vertices $x, y \in E$ with $x \neq y$ and $G - xy$ is connected. But then there is some xy path P that does not use the edge xy , so xPy is then a cycle, contradicting that G is acyclic.

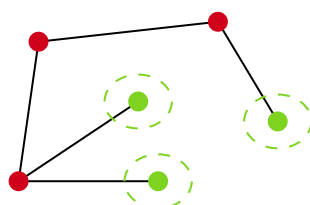
(c) \implies (a). By definition G is connected. Again for a contradiction, assume that G contains a cycle C . Then let xy be an edge on C . We claim $G - xy$ is still connected. If $u, v \in V(G - xy)$ then let P be a path in G from u to v . If xy does not appear as consecutive vertices on this path, then u is connected to v . Otherwise, we can consider a new path where we replace x, y with the other vertices in $C - xy$ in order. Thus u and v are still connected. This contradicts the minimal connectedness of G .

□

Definition 1.30 (Leaf)

Let G be a graph. A vertex $v \in V(G)$ is a **leaf** if $d(v) = 1$.

For example, the tree below has three leaves.



In general, trees has a leaf.

Proposition 1.31 (Trees Have Leaves)

Every tree T with $|T| \geq 2$ has a leaf.

Proof. Let T be a tree with $|T| \geq 2$, and let P be a path of maximum length in T , with $P = x_1 \dots x_k$. We claim that $d(x_k) = 1$. Observe that $\deg(x_k) \geq 1$, since $x_k x_{k-1} \in E$. If x_k is adjacent to another vertex $y \neq x_{k-1}$, then either $y \in \{x_1, \dots, x_{k-2}\}$, which would imply that T contains a cycle, or $y \notin \{x_1, \dots, x_{k-2}\}$, then $x_1 \dots x_k y$ is a path longer than P , which violates its maximality. \square

Remark. This proof gives us two leaves in T , which is the best we can hope for considering P_n is a tree with exactly two leaves.

Proposition 1.32 (Edges of a Tree)

Let T be a tree. Then $e(T) = |T| - 1$.

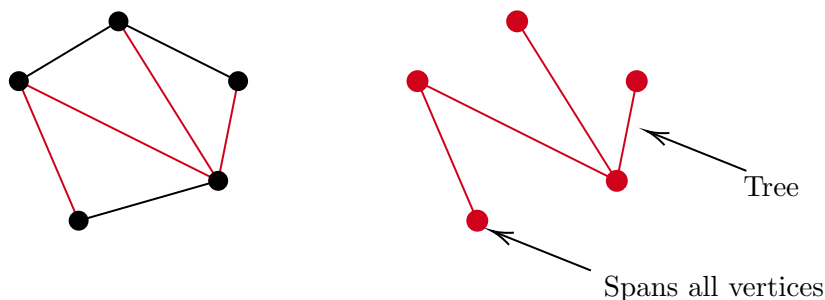
Proof. We will do induction on $n = |T|$. If $n = 1$, this is trivial as there is only one edge. Now given T with at least 2 vertices, let x be a leaf in T , and define $T' = T - x$.

T' must be acyclic, since we have only removed vertices. T' must also be connected since for all $u, v \in V(T')$ there exists a path from u to v in T that does not use x , so it is also a path from u to v in T' . Thus T' is a tree. Thus by induction, T' has $n - 2$ edges, and $e(T) = e(T') + 1 = |T| - 1$. \square

Now let's think about trees as subgraphs of other graphs.

Definition 1.33 (Spanning Tree)

Let G be a graph. We say T is a **spanning tree** of G if T is a tree on $V(G)$ and is a subgraph of G .



Spanning trees are useful in a number of contexts, one of which is giving a sensible ordering to the vertices of a graph. They are particularly useful because of the following result.

Proposition 1.34 (Connected Graphs have Spanning Trees)

Every connected graph contains a spanning tree.

Proof. A tree is a minimal connected graph. So take the connected graph and remove edges until it becomes a minimal connected graph. Then this will be a subgraph of the original graph, and will thus be a spanning tree. \square

§1.3 Bipartite Graphs

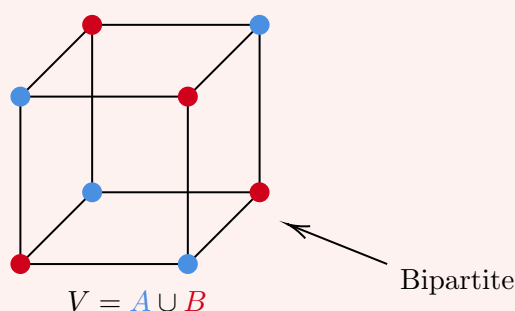
The next type of graph we will look at is *bipartite* graphs.

Definition 1.35 (Bipartite Graphs)

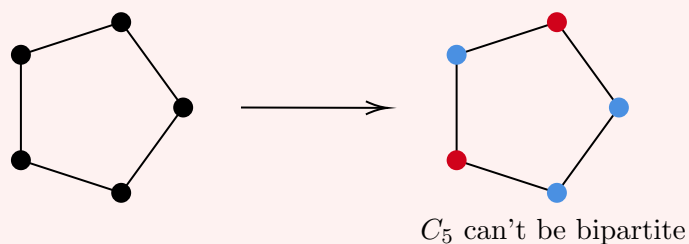
A graph $G = (V, E)$ is **bipartite** if $V = A \cup B$ where $A \cap B = \emptyset$ and all edges $xy \in E$ have either $x \in A, y \in B$ or $x \in B, y \in A$.

Example 1.36 (Example of Bipartite Graphs)

The graph below is bipartite, with vertices in the set A being coloured red and vertices in the set B being coloured blue.



An example of a non-bipartite graph is C_5 . To see this, we can start by choosing a vertex to be in A (without loss of generality), then the adjacent vertices must be in B , but then their adjacent vertices must be in A , but then there is an edge between two vertices in A . This is shown below.



The argument given for C_5 works in general.

Proposition 1.37 (Bipartite Cyclic Graphs)

The cycle C_{2k+1} is *not* bipartite, and the cycle C_{2k} *is* bipartite.

Proof. Assume that C_{2k+1} is bipartite. Then there must be disjoint sets A and B , and as $2k + 1$ is odd, we must have (without loss of generality), $|A| > |B|$. Now let's count the edges between A and B . This must be $2|A|$ and also $2|B|$, as every vertex has degree 2. But then $|A| = |B|$, which is a contradiction.

For C_{2n} , we can let $v_i \in A$ if i is even and $v_i \in B$ if i is odd. Then a vertex i only has edges to vertices $i - 1$ and $i + 1 \pmod{2}$, which have opposite parity. Thus all edges are between A and B , as required. \square

There is then a natural question: given some arbitrary graph G , how do we determine if a given graph is bipartite? It turns out that there is a nice check for ‘bipartness’. We will state the result and then do some setup before we prove it.

Proposition 1.38 (Bipartite Criterion)

A graph G is bipartite if and only if G contains no odd cycles.

We need to first develop some theory regarding *circuits*. Informally, a circuit is like a cycle where we can revisit vertices.

Definition 1.39 (Circuit)

A circuit is a sequence $x_1 \dots x_l$ where $x_1 = x_l$ and $x_i x_{i+1} \in E$. The **length** of the circuit is $l - 1$, the number of edges traversed in the circuit.

Definition 1.40 (Odd Circuits)

If the length of a circuit is odd, then we say it is an **odd circuit**.

Proposition 1.41

An odd circuit contains an odd cycle.

Proof. We will prove this by induction on the length of the circuit. For a circuit of length 3, the circuit must be a cycle. In general, let $C = x_1 \dots x_l$ be our circuit. If x_1, \dots, x_{l-1} are distinct, then C is a cycle and we are done.

Otherwise, there exists some $z \in C$ that is repeated. We write

$$C = x_1 \dots x_a z x_{a+2} \dots x_b z x_{b+2} \dots x_l.$$

We define $C' = x_1 \dots x_a z x_{b+2} \dots x_l$ and $C'' = z x_{a+2} \dots x_b z$. The length of C' and C'' is strictly less than the length of C . One of these circuits must have odd length, and by induction that odd circuit contains an odd cycle. \square

We can now prove our original bipartness criterion, that a graph is bipartite if and only if it contains no odd cycles.

Proof (Bipartite Criterion). If G was bipartite and contained an odd cycle, then there exists an odd cycle that is bipartite. But this is a contradiction.

Now if G is not bipartite, we can induct on the number of vertices. For $|G| = 1$, this holds. Now if G is not connected, let C_1, \dots, C_k be the components of G . We may now apply our induction to each component of G to obtain a bipartition $V(C_i) = A_i \cup B_i$ for each $i \in 1, \dots, k$. Then $A = A_1 \cup \dots \cup A_k$ and $B = B_1 \cup \dots \cup B_k$ is a bipartition for the whole graph.

We may now assume without loss of generality that G is connected. Fix some vertex $v \in V$, and define

$$\begin{aligned} A &= \{u \in V \mid d(u, v) \text{ is odd}\} \\ B &= \{u \in V \mid d(u, v) \text{ is even}\} \end{aligned}$$

We claim that $A \cup B$ is a bipartition.

Assume (for a contradiction) that u_1 is adjacent to u_2 and $d(u_1, v) \equiv d(u_2, v) \pmod{2}$. Then there exists paths P_1 from v to u_1 and P_2 from u_2 to v with $|P_1| \equiv |P_2|$. But this implies that $vP_1u_1u_2P_2$ defines a odd circuit in G . Therefore, by our previous proposition, G contains an odd cycle, which is a contradiction. \square

§2 Hall's Theorem

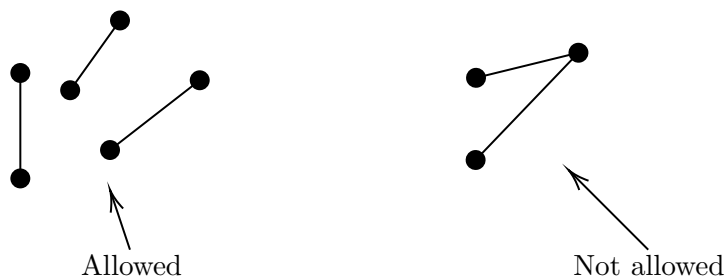
In this section, we will build up our knowledge of *matchings* so that we can prove the first theorem of the course – Hall's theorem.

§2.1 Matchings

An appropriate place to start is probably by defining what a matching is.

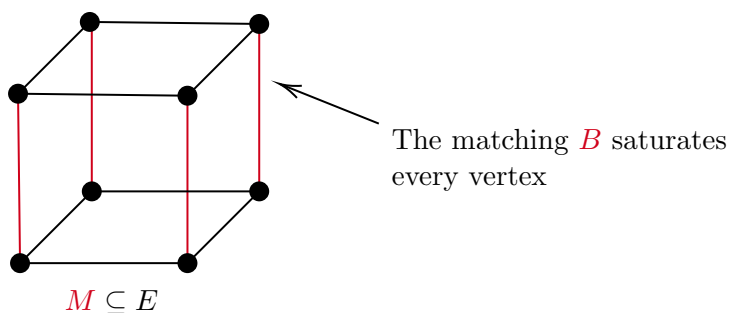
Definition 2.1 (Matching)

A **matching** of G is a collection of edges $M \subseteq E$ so that $\forall e_1, e_2 \in M$ with $e_1 \neq e_2$ has $e_1 \cap e_2 = \emptyset$.



Definition 2.2 (Saturated)

Given a graph $G = (V, E)$ and a matching M in G , we say that a vertex $v \in V$ is **saturated by M** if there exists an edge in M containing v .

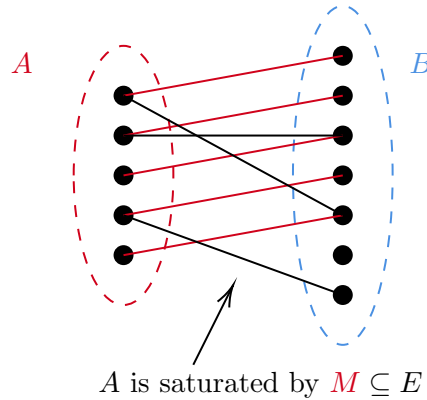


A matching where every vertex is saturated (such as the above) is known as a **per-**

fect matching. However, there is no restriction in general on how many vertices are saturated by a matching.

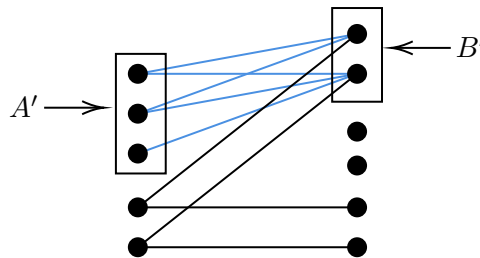
§2.2 Matching in Bipartite Graphs

Matchings are particularly interesting in bipartite graphs. We will be interested in the following question: given a bipartite graph $G = (A \cup B, E)$, when can I find a matching saturating A ?



This question is the same as asking when is there a function $f : A \rightarrow B$ where $xf(x) \in E$ that is an injection.

In trying to answer this question, we might try and think about why it may not be possible. The simplest reason is when B isn't large enough to have an injection, when $|B| \leq |A|$. In a similar way, we might have a graph is big enough, but that isn't true for a small part of the graph, like below.



What Hall's theorem says is that this issue is the only obstruction to creating such a matching.

Definition 2.3 (Neighbourhood of a Vertex Set)

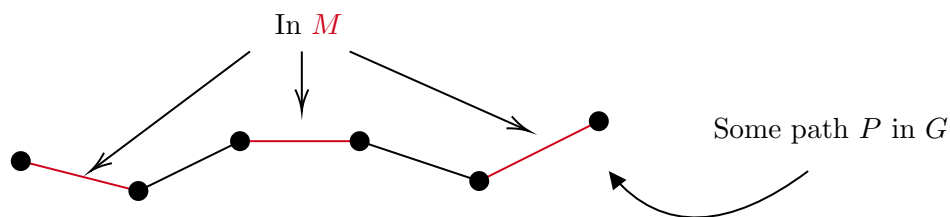
If $X \subseteq V$, then we define $N(X)$, the **neighbourhood of X** to be $N(X) = \bigcup_{x \in X} N(x)$.

With the notion of the neighborhood of a set of vertices, we can rephrase the issue mentioned about as $|N(A')| < |A'|$ for some subset $A' \subseteq A$. We will use this notation in our statement for Hall's theorem.

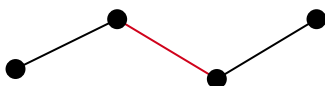
Before we prove the theorem, we will need to prepare a little bit.

Definition 2.4 (M -Alternating Path)

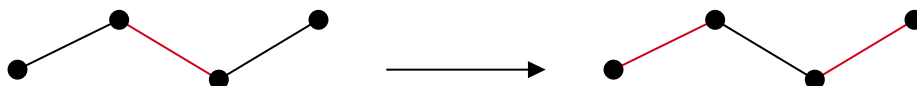
Let $G = (V, E)$ be a graph with a matching M in G . We say that a path $P = x_1 \dots x_l$ is **M-alternating** if $x_i x_{i+1}$ is alternately in M and not in M .



Another example of an M -alternating path is the one below.



If we saw the path above in the graph, and we knew that the end vertices were not saturated, then we could change the edges that are in M , so we would still have a matching. This move will be key in our proof of Hall's theorem.



We are going to call this configuration an *augmented path*.

Definition 2.5 (M -Augmenting Path)

Given a graph $G = (V, E)$ and a matching M in G , an M -alternating path $P = x_1 \dots x_l$ is said to be **M -augmenting** if x_1 and x_l are not saturated.

Proposition 2.6

If M is a matching in G of maximum size, then there are no M -augmenting paths.

Proof. If there is an M -augmenting path in G , then we can flip the edges of M along P to find a strictly larger matching. \square

An important observation is that an M -alternating path in a bipartite graph $G = (A \cup B, E)$ with $P = x_1 \dots x_l$, with $x_1 \in A$ and $x_1 x_2 \in M$, then $x_{2k+1} x_{2k+2} \in M$, and vice versa.

§2.3 Hall's Theorem

We are now ready to state and prove Hall's matching theorem, which formalises the ideas mentioned in the previous section.

Theorem 2.7 (Hall's Theorem)

Let $G = (A \cup B, E)$ be a bipartite graph. Then there exists a matching saturating A if and only if every subset $A' \subseteq A$ satisfies $|N(A')| \geq |A'|$.

Proof. First we prove the forward (and easy) direction. Let $A' = \{x_1, \dots, x_t\} \subseteq A$. We have matching edges $x_1y_1, \dots, x_ty_t \in M \subseteq E$, and thus $\{y_1, \dots, y_t\} \subseteq N(A')$, and thus $|N(A')| \geq |A'|$.

Now for the other (harder) direction, which is that this condition implies the existence of such a matching. Choose a matching M in G with $|M|$ maximized. For a contradiction, assume there is some vertex $a_0 \in A$ that is not saturated by M .

We inductively define sets $A_i \subseteq A$, $B_i \subseteq B$ by setting $A_0 = \{a_0\}$, $B_0 = \emptyset$. We will maintain, for all t , that

1. $|A_t| = t + 1$, $|B_t| = t$.
2. Every vertex in $A_t \cup B_t$ is the endpoint of an alternating path that started at a_0 .
3. $A_t \setminus \{a_0\}$ is matched to B_t .

So given A_t , B_t , we need to define A_{t+1} , B_{t+1} .

First consider $N(A_t)$. We have $|N(A_t)| \geq |A_t| = t + 1 > |B_t|$. So $N(A_t) \setminus B_t$ is non-empty. So let $b_{t+1} \in N(A_t) \setminus B_t$. Observe that b_{t+1} can be reached along an alternating path which started at a_0 . Call $y \in A_t$ such that $y b_{t+1} \in E$. Then y can be reached along an alternating path starting at a_0 . Let $P = a_0 x_1 \dots x_l y$ be such an alternating path. Since $a_0 x_1 \notin M$, $x_l y \in M$. Hence $P b_{t+1}$ is an alternating path.

Now, b_{t+1} is saturated by M , as otherwise $P b_{t+1}$ would be an M -augmenting path, which is a contradiction. So let $a_{t+1} b_{t+1} \in M$. We claim $a_{t+1} \notin A_t$, since A_t is matched to B_t , and $b_{t+1} \notin B_t$. So we may define $A_{t+1} = A_t \cup \{a_{t+1}\}$ and $B_{t+1} = B_t \cup \{b_{t+1}\}$. We can check that what we claimed before holds.

Since $a_{t+1} \notin A_t$, and $b_{t+1} \notin B_t$, we have $|A_{t+1}| = |A_t| + 1$ and $|B_{t+1}| = |B_t| + 1$. Also every vertex is the endpoint of an alternating path starting at a_0 by construction. Also $A_{t+1} \setminus \{a_0\}$ is matched to B_{t+1} since $A_t \setminus \{a_0\}$ is matched to B_t , and by construction.

This completes the construction of A_t , B_t for all t . Then if $t > |A|$, then $|A_t| > |A|$, but $A_t \subseteq A$, which is a contradiction. \square

§2.4 Corollaries of Hall's Theorem

So Hall's theorem gives us a nice way to guarantee the existence of matchings that saturate a part of a bipartite graph. We are going to use Hall's theorem to prove some interesting results.

Corollary 2.8

A k -regular bipartite graph contains a perfect matching.

Proof. Let $G = (A \cup B, E)$, and let $A' \subseteq A$. We want to show that $|N(A')| \geq |A'|$ so that we may apply Hall's theorem.

We will count the number of edges between A' and $N(A')$ in two different ways. We know that each vertex in A' has degree k , so there is $k|A'|$ edges. But on the

other hand, the number of edges in $N(A')$ is $|N(A')k|$. Thus $|N(A')k| \geq |A'|k$, so $|N(A')| \geq |A'|$, and by Hall's theorem there exists a matching saturating A .

We claim that $|B| = |A|$. Indeed, the number of edges between A and B is $k|A|$ and is also $k|B|$, and thus $|A| = |B|$. So a matching saturating A also saturates B . So there exists a perfect matching. \square

We can also consider an extension of Hall's theorem. Hall's theorem tells us when, in a bipartite graph $G = (A \cup B, E)$, there is a matching saturating A , but what if we only wanted a matching of size k ?

Let's say that a matching in G has **deficiency** d if it saturates $|A| - d$ vertices.

Corollary 2.9

Let G be a bipartite graph. Then G contains a matching saturating $|A| - d$ vertices in A if and only if for all $A' \subseteq A$, we have $|N(A')| \geq |A'| - d$.

Proof. If we have a matching with deficiency d , then this clearly holds.

Now if this condition holds for some graph $G = (A \cup B, E)$, we define a new graph \tilde{G} by setting $\tilde{B} = B \cup \{z_1, \dots, z_d\}$ where z_1, \dots, z_d are distinct from $x \in A \cup B$, and we define $\tilde{E} = E \cup \{e_i a \mid i \in \{1, \dots, d\}, a \in A\}$. So $\tilde{G} = (A \cup \tilde{B}, \tilde{E})$.

This is a bipartite graph, and we observe that it satisfies Hall's condition. Thus \tilde{G} has a matching M saturating A . So if we define M' by removing all of the edges with an endpoint in $\{z_1, \dots, z_d\}$, then M' saturates all but at most d vertices in G . \square

Another way of stating Hall's theorem is with set systems.

Definition 2.10 (System of Distinct Representatives)

Gives sets $S_1, \dots, S_n \subset X$ where S_i is finite for all i . Then we say that $x_1, \dots, x_n \in X$ is a **system of distinct representatives** (or SDR) if they are distinct and $x_i \in S_i$.

The question is, for what set systems does there exist a system of distinct representatives? The answer is if they satisfy some Hall-like condition.

Corollary 2.11 (Existence of SDRs)

Given $S_1, \dots, S_n \subseteq X$, with S_i finite, then S_1, \dots, S_n has a system of distinct representatives if and only if

$$\left| \bigcup_{i \in I} S_i \right| \geq |I|,$$

for all $I \in \{1, \dots, n\}$.

Proof. If such an SDR exists, then this condition clearly holds.

In the other direction, given that this condition holds, we define a graph G by setting $A = \{S_1, \dots, S_n\}$ and $B = \bigcup_{i=1}^n S_i$, with $E = \{\{S_i, x\} \mid x \in A, i \in \{1, \dots, n\}, x \in S_i\}$. Then $G = (A \cup B, E)$. Now observe that a SDR is exactly a matching in G .

that saturates A .

We check Hall's condition. Given $A' \subseteq A$, then $N(A') = \bigcup_{S_i \in A'} S_i$. Thus

$$|N(A')| = \left| \bigcup_{S_i \in A'} S_i \right| \geq |A'|,$$

by our condition. So such a matching exists. \square

Remark. The existence of SDRs is equivalent to Hall's theorem.

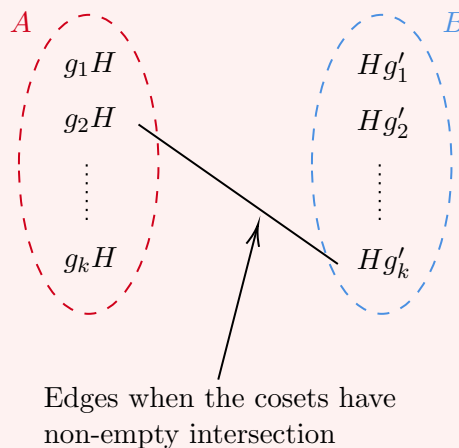
The last example of using Hall's theorem will be an application to group theory.

Example 2.12 (Left and Right Coset Representatives)

Let G be a group and let $H < G$, with G finite. We have the left cosets g_1H, \dots, g_kH and right cosets Hg'_1, \dots, Hg'_k .

We want to know does there exists h_1, \dots, h_k such that h_1H, \dots, h_kH are all of the left cosets and Hh_1, \dots, Hh_k are all of the right cosets. It turns out that this question is entirely combinatorial, and we can use Hall's theorem.

Let's form the bipartite graph $(A \cup B, E)$ with $A = \{g_1H, \dots, g_kH\}$ and $B = \{Hg'_1, \dots, Hg'_k\}$ so that there is an edge between g_iH and Hg'_j if the cosets have non empty intersection.



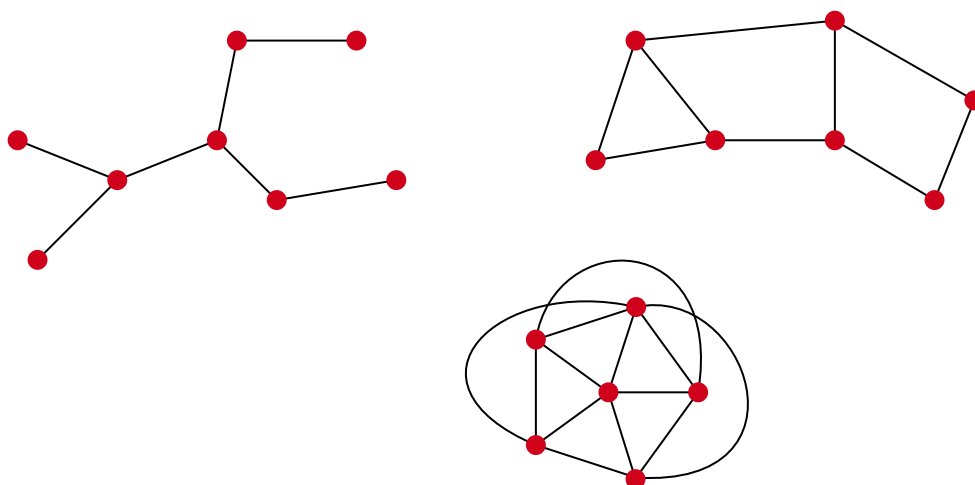
If we had a perfect matching in this graph, then we could pick an element h_i in the (non-empty) intersection of the matched left and right cosets, and then h_1, \dots, h_k would be a set of coset representatives for all of the k left and right cosets.

So we want to show that Hall's theorem is satisfied. Given $I \subseteq \{1, \dots, k\}$, we want to show that the number of right cosets intersecting $\bigcup_{i \in I} g_iH$ is at least $|I|$.

Observe that $|\bigcup_{i \in I} g_iH| = |H||I|$. Since the right cosets partition G , and each coset has size $|H|$, so the number of right cosets intersecting $\bigcup g_iH$ is at least $|I|$. Thus $|N(\{g_iH \mid i \in I\})| \geq |I|$, and thus Hall's theorem is satisfied, and we have a perfect matching in the graph.

§3 Connectivity

We have already defined what it means for a graph to be connected, but consider the following connected graphs:



Clearly these are connected, but they are all ‘connected to different extents’. For example, in the first graph, removing any vertex disconnects the graph. In the second graph, any vertex could be removed and the graph would stay connected. We can also see that in the first graph, there’s only one path from one vertex to another, whereas in the third graph there is many. This also seems to correlate with ‘how connected’ a graph is.

§3.1 Measuring Connectivity

So looking at the graphs above, we two natural notions for ‘how connected a graph is’ can be informally described as

1. A ‘deletion notion’ of connectivity, where we consider how connected the graph is after some vertices are removed.
2. A ‘paths notion’ of connectivity, where we consider how many independent paths there is between vertices.

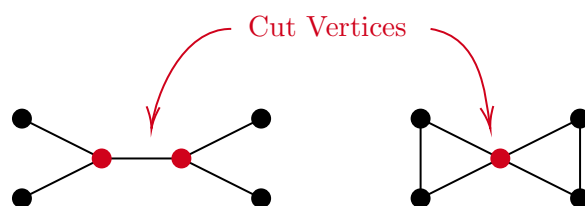
One of the main goals of this section will be turning this vague notion into a concrete concept, and proving an interesting result about how the two notions relate to each other.

Notation (Removing Vertex Sets). In this section (and for the rest of this article), if $G = (V, E)$ is a graph and $S \subseteq V$, we define $G - S = G - x_1 - x_2 - \dots - x_l$ where $S = \{x_1, \dots, x_l\}$.

We will begin with a definition.

Definition 3.1 (Cut Vertex)

Let $G = (V, E)$ be a connected graph. We say that $v \in V$ is a **cut vertex** if $G - v$ is disconnected.

**Definition 3.2 (Separator)**

If $G = (G, V)$ is a connected graph, we say that a subset $S \subseteq V$ is a **separator** (or separating set) if $G - S$ is disconnected.

With these concepts defined, we can define our ‘deletion’ notion of connectivity.

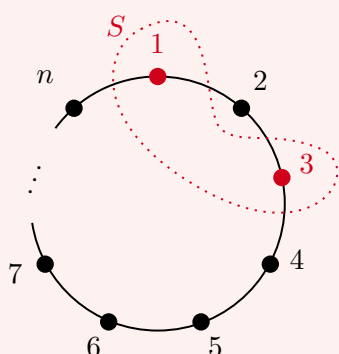
Definition 3.3 (Connectivity)

Let $G = (V, E)$ be a graph. The **connectivity** of G , denoted $\kappa(G)$, is the size of the smallest set $S \subseteq V$ such that $G - S$ is disconnected, or just a single vertex^a.

^aThis is needed for the case of a complete graph.

Example 3.4 (Connectivity of C_n)

Consider the graph C_n for $n \geq 3$.

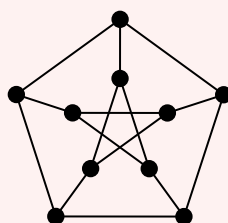


Removing one vertex will never disconnect C_n .
However, removing two vertices can disconnect it.

$$\implies \kappa(C_n) = 2$$

Example 3.5 (Connectivity of the Petersen Graph)

We define the **Petersen graph** with 10 vertices and 15 edges as shown below.



We can see that the connectivity is at most 3, since that is the degree of each vertex, and also removing two vertices won't disconnect the graph. Thus the connectivity of the Petersen graph is exactly 3.

Definition 3.6 (k -Connected)

We say that a graph G is **k -connected** if $\kappa(G) \geq k$. In particular, any set $S \subseteq V(G)$ with $|S| < k$ will have $G - S$ connected.

We note this immediately implies that G is 1-connected if and only if it's connected, and it is 2-connected if and only if it has no cut vertex.

We can note some basic properties of connectivity.

Lemma 3.7 (Increasing/Reducing Connectivity)

If $G = (V, E)$ is a k -connected graph and $v \in V$ then $G - v$ is $(k - 1)$ -connected^a. Also, if we have some $e \in E$, then $G - e$ is $(k - 1)$ -connected.

^aNote that it is also possible that $G - v$ is $(k + 1)$ -connected, so connectivity can also increase by deleting a vertex.

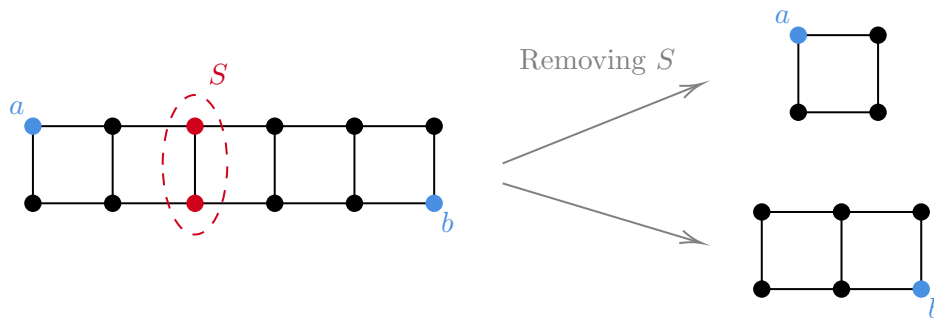
Proof Sketch. Check definitions. □

§3.2 Preparing for Menger's Theorem

We are going to prove Menger's theorem about connectivity measures, but we have to do some setting up beforehand.

Definition 3.8 (ab -Seperator)

If $G = (V, E)$ is a graph and $a, b \in V$ are distinct vertices, we say that S is an **ab -seperator** if a and b lie in different components of $G - S$ (so $a, b \notin S$).

**Definition 3.9** (Seperator)

If $G = (V, E)$ is a graph and $F \subseteq E$, we say that F is a **seperator** of G if $G - F$ is disconnected.

We then have a number of useful lemmas, that we will employ later on.

Notation. We let $\kappa_{a,b}(G)$ be the minimum size of an ab separator.

Lemma 3.10

Let $G = (V, E)$ be a graph. Then $\kappa_{a,b}(G) \geq \kappa_{a,b}(G - v)$, where $v \in V$ and $v \neq a, b$.

Proof. If $G - v$ has an ab separator S then $S \cup \{v\}$ is an ab separator in G . \square

Lemma 3.11

For a graph $G = (V, E)$, $\kappa_{a,b}(G - e) \geq \kappa_{a,b}(G) = 1$, for $e \in E$ and $a \sim b$.

Proof. If $G - e$ has an ab separator S then $S \cup \{x\}$ and $S \cup \{y\}$ are ab separators in G , where $e = xy$. \square

Lemma 3.12

Let $G = (V, E)$ be a graph with distinct non-adjacent vertices $a, b \in V$. Also let $\kappa_{a,b}(G) \geq k$. Let S be an ab separator in G , and say $G - S = A \cup C$, where A is the connected component containing a . Then define \tilde{G} as the induced graph $G[A \cup S]$ with a vertex x joined to all of S . Then $\kappa_{a,x}(\tilde{G}) \geq k$.

Proof Sketch. Check definitions. \square

We can now formalize our ‘independent paths’ notion of connectivity.

Definition 3.13 (Independent Paths)

We say that P_1, \dots, P_k are **independent ab paths** if each of P_1, \dots, P_k are ab -paths and all of the vertices in these paths are distinct, apart from a and b .

§3.3 Actually Menger’s Theorem

Finally, we can state and prove Menger’s theorem.

Theorem 3.14 (Menger’s Theorem, First Form)

Let $G = (V, E)$ be a graph, with distinct and non-adjacent $a, b \in V$. If every ab separator in G has size at least k then we can find k independent ab paths.

Proof. Suppose for a contradiction, assume this is false and let G be the counterexample that

- (1) Minimizes $\kappa_{a,b}(G)$
- (2) Subject to (1), minimizes the number of edges in the graph.

Then we observe that $\kappa_{a,b}(G - e) = \kappa_{a,b}(G) - 1$ for any edge $e \in E$. We will let $k = \kappa_{a,b}(G)$.

Claim. There exists an ab separator S with $|S| = k$ so that $S \not\subseteq N(a)$ and $S \not\subseteq N(b)$.

We first observe that $N(a) \cap N(b)$ is empty. To see this, let $x \in N(a) \cap N(b)$ and $G' = G - x$. Then $\kappa_{a,b}(G') \geq \kappa_{a,b}(G) - 1$, thus there are $k - 1$ independent paths P_1, \dots, P_{k-1} in G' , then P_1, \dots, P_{k-1}, axb are k independent paths in G , and our graph would then not be a counterexample.

Now choose a shortest path $P = ax_1 \dots x_lb$. Let $G' = G - x_1x_2$. We know that $\kappa_{a,b}(G') = k - 1$. Therefore there exists an ab separator S' of size $k - 1$. Let A be the component of a in $G' - S'$, and B be the component of b in $G' - S'$. Note that x_1x_2 must have $x_1 \in A$ and $x_2 \in B$.

Note $x_1 \sim a$, and $x_2 \not\sim a$ since P is the shortest path. Also note that $x_2 \neq b$, since $N(a) \cap N(b) = \emptyset$.

If $S' \subseteq N(a)$, then $S' \cup \{x_2\}$ is an ab separator of size k , and $S \not\subseteq N(a), N(b)$. If $S' \subseteq N(b)$, then $S = S' \cup \{x_1\}$ is an ab separator of size k , and $S \not\subseteq N(b), N(a)$. Lastly, if $S' \not\subseteq N(a)$ and $S' \not\subseteq N(b)$, then $S = S' \cup \{x_1\}$ is an ab separator of size k that is not in $N(a)$ or $N(b)$. So our claim holds.

So let S be an ab separator with $|S| = k$ and $S \not\subseteq N(a), N(b)$. Define A, B to be the components containing a and b in $G - S$. We also define \tilde{G}_a as $G[A \cup S]$ with a vertex x that joins to all of S . define \tilde{G}_b likewise. We now have

$$e(\tilde{G}_a) < e(G), \quad e(\tilde{G}_b) < e(G).$$

We also have $\kappa_{a,b}(\tilde{G}_a) = k = \kappa_{a,b}(\tilde{G}_b)$. Thus \tilde{G}_a and \tilde{G}_b satisfy the theorem, by minimality.

So we can find independent ax paths $P_1, \dots, P_k \in \tilde{G}_a$ and yb paths $Q_1, \dots, Q_k \in \tilde{G}_b$. Thus we can find k independent ab paths by concatenation and reordering in G , as desired. \square

Remark. It should be noted that we need the non-adjacent condition, otherwise there is no ab separator. Before we write down the proof, we will isolate some notable facts about connectivity. Also, this result implies Hall's theorem.

Another form of Menger's theorem is more common.

Theorem 3.15 (Menger's Theorem, Second Form)

Let $G = (V, E)$ be a graph. Then G is k -connected if and only for all $u, v \in V$ with $u \neq v$, there exists k independent uv -paths.

Proof. If u is not adjacent to v , then apply Menger's theorem (first form) to find k -independent uv paths. If they are adjacent, then $G' = G - uv$ is $k - 1$ connected, thus Menger's theorem (first form) tells us that there are uv independent paths P_1, \dots, P_{k-1} in $G - uv$. Thus P_1, \dots, P_{k-1}, uv are k independent paths. The other direction is straightforward. \square

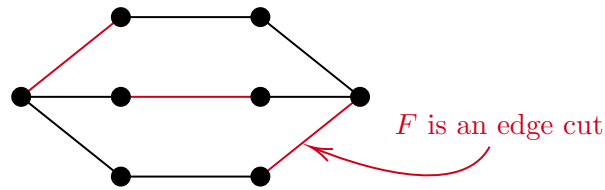
§3.4 Edge Connectivity

We have so far been looking at connectivity related to removing vertices and considering independent paths. We will now look at connectivity related to removing edges, and we will see that Menger's theorem is still useful.

Notation. If $G = (V, E)$ is a graph and $F \subseteq E$, define $G - F = (V, E \setminus F)$.

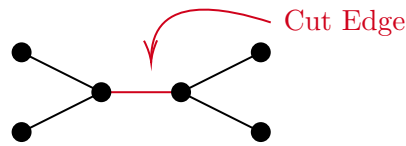
Definition 3.16 (Edge Cut)

An **edge cut** in a graph $G = (V, E)$ is a set $F \subseteq E$ so that $G - F$ is disconnected.



Definition 3.17 (Cut Edge)

A **cut edge** $e \in E$ is an edge so that $G - e$ is disconnected.

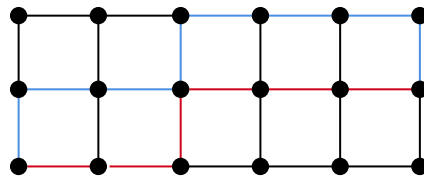


Now similarly to how we had independent paths before (that didn't share vertices), we can define a notion of edge independent paths.

Definition 3.18 (Edge Disjoint Paths)

We say that the uv paths P_1, \dots, P_k are **edge disjoint** if $E(P_i) \cap E(P_j) = \emptyset$ for all $i \neq j$.

The **blue** and **red** paths are edge disjoint



We can then define edge connectivity (our deletion notion).

Definition 3.19 (Edge Connectivity)

Define the **edge connectivity** of G , $\lambda(G)$ to be the smallest $|S|$ with $S \subseteq E$ such that $G - S$ is disconnected.

Definition 3.20 (k -Edge-Connected)

We say a graph G is **k -edge-connected** if $\lambda(G) \geq k$. In other words, $G - F$ is connected for all $|F| \leq k - 1$, with $F \subseteq E$.

We note that G is 1-edge-connected if and only if it is connected, and it is 2-edge-connected if and only if there is no cut edge.

We have a Menger's theorem for this notion of connectivity also.

Theorem 3.21 (Menger's Theorem, Edge Version)

Let $G = (V, E)$ be a graph, and u, v be distinct vertices of G . If every set of edges $F \subseteq E$ that separates u from v has size greater than or equal to k , then there exists k edge disjoint paths from u to v .

We also have a similar second version.

Theorem 3.22 (Menger's Theorem, Edge Version Two)

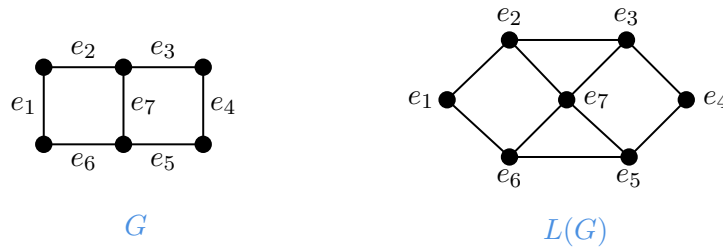
Let $G = (V, E)$ be a graph. Then G is k -edge-connected if and only if for every $u, v \in V$ with $u \neq v$ there exists k edge disjoint uv -paths P_1, \dots, P_k .

We are going to prove this by constructing a graph that we can then get the required result from by applying vertex Menger. The construction will be based on the idea of a *line graph*.

Definition 3.23 (Line Graph)

Given a graph $G = (V, E)$, we define $L(G)$ to be the **line graph** as follows. $V(L(G)) = E$, and for $e, f \in E$, we have $ef \in E(L(G))$ if $e \cap f \neq \emptyset$.

An example of a graph and its corresponding line graph is shown below.



We can now prove the edge version of Menger's theorem.

Proof (Menger's Theorem, Edge Version). We are given a graph $G = (V, E)$ and distinct vertices $u, v \in V$ such that every u, v separator has size $\geq k$.

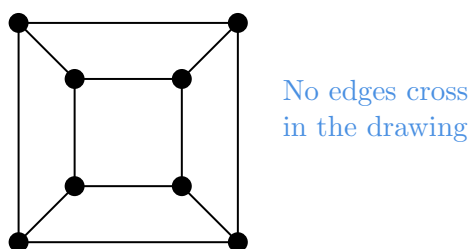
Construct the graph \tilde{G} which is $L(G)$ along with vertices u and v with u joined to all $e \in V(L(G))$ such that $u \in e$, and likewise for v .

Then applying Menger's theorem (the vertex version, form one) to \tilde{G} with u, v as the distinguished vertices. \square

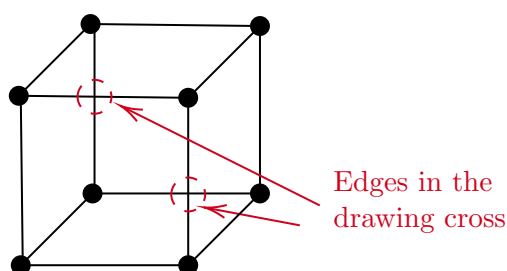
§4 Planar Graphs

Informally, a graph is **planar** if it can be drawn in the plane without any pair of edges crossing.

For example the cube graph we mentioned earlier is planar, as we can draw it as below.



Of course, a graph is planar only if *there is some drawing* where the edges don't intersect. For example, we could draw the cube graph as below (where edges intersect), and the graph would still be planar.



§4.1 Defining Planar Graphs

Let's formalize the notion of 'planar graphs' a bit. First we will define (a somewhat obvious notion) what we mean for a graph to be in a plane.

Definition 4.1 (Plane Graph)

A **plane graph** is a finite set of points $V \subseteq \mathbb{R}^2$ and a collection of disjoint polygonal curves (representing the edges) with start and endpoints in V .

Then we can define what it means for a graph to be *planar*.

Definition 4.2 (Planar Graph)

A graph G is **planar** if there exists a graph isomorphism from G to some plane graph.

And again, informally this says that a graph is planar if there is some way to draw it in the plane so that edges don't intersect.

We can also define the notion of 'faces' of a plane graph, by looking at the components.

Definition 4.3 (Faces)

The **faces** of a plane graph G are the connected components of $\mathbb{R}^2 - G$.

With this we get a nice relation between the vertices, edges and faces of a plane graph.

Theorem 4.4 (Euler's Formula)

Let G be a connected plane graph with V vertices, E edges and F faces. Then

$$V - E + F = 2.$$

Proof. We apply induction on the number of edges of G . The base case is $E = 0$, and then since G is connected, $V = 1$, and $F = 1$, and the formula holds. Now if G contains a cycle C , then let e be on C . Then $G' = G - e$ is still connected, and the number of faces increases by 1. The face enclosed is lost. So considering G' and applying Euler's formula, $V - (E - 1) + (F - 1) = 2$, and $V - E + F = 2$, as required.

If G does not contain a cycle, then G is acyclic and connected and is then a tree. Then $E = V - 1$, $F = 1$, and $V - E + F = 2$, as required. \square

Corollary 4.5 (Planar Graphs are Sparse)

Let $G = (V, E)$ be a planar graph, with $|V| \geq 3$. Then $|E| \leq 3|V| - 6$. This bound is also sharp^a.

^aBy considering plane graphs where every face is a triangle.

Proof. We may assume without loss of generality that G is connected (if not, then we can add edges until it is connected). Also draw G in the plane so that there is no edge crossings. Then by Euler's formula we have $V - E + F = 2$.

Now since every face has at least three edges on its boundary and every edge on the boundary is incident to at most two faces, we obtain

$$3F \leq |\{(e', f') \mid e' \in E, f' \text{ is a face, and } e' \text{ is on the boundary of } f'\}| \leq 2E.$$

Thus $3F \leq 2E$, so $3(2 - V + E) \leq 2E$, and $E \leq 3V - 6$, as required. \square

§4.2 What Graphs are Planar?

In this section we really care about the following question:

What graphs are planar?

The theorems proved at the end of the last section already give us a small amount of control over planar graphs. For example, we can show that K_5 is not planar.

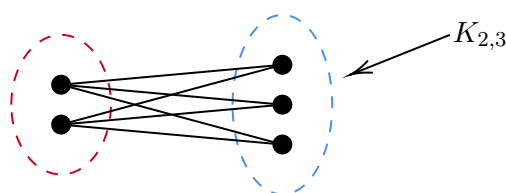
Example 4.6 (K_5 is Non-Planar)

We will show that K_5 is not planar. Since the number of edges is 10, and the number of vertices is 5, by the previous result we would need $10 \leq 3 \cdot 5 - 6$, which is not the case.

Let's consider another type of graph.

Definition 4.7 (Bipartite Complete Graph)

Define the **bipartite complete graph** $K_{n,m}$ to be the graph with vertex set $V = A \cup B$ where $A \cap B = \emptyset$ so that $|A| = n$, $|B| = m$, and $E(K_{n,m}) = \{ab \mid a \in A, b \in B\}$.



If you play around a bit, you can see that $K_{3,2}$ is planar, but we get a problem if we try use $K_{3,3}$.

Example 4.8 ($K_{3,3}$ is Non-Planar)

We will show that $K_{3,3}$ is not planar. Note that there is 9 edges, and 6 vertices. Then by the result in the previous section, we would need $9 \leq 3 \cdot 6 - 6$, which holds.

However, recall in the proof that of our result that we had each face having at least three edges on its boundary. But that for bipartite graphs, we can get something stronger, as each face has at least four edges on its boundary (since there is no cycles of length 3).

So if we repeat the proof of the previous theorem with the stronger bound, we obtain $E \leq 2V - 4$, which does not hold for this graph. Thus $K_{3,3}$ is not planar.

Now these two examples were interesting, but of course we care about whether *any* graph is planar. It turns out though that these are (in some sense) the *only* fundamentally non-planar graphs, in that any non planar graph will have one of these graphs ‘behind it’.

To look at this formally, we need to look at the idea of a subdivision.

Definition 4.9 (Subdivision)

A **subdivision** of a graph G is a graph \tilde{G} , obtained by replacing the edges of G with disjoint paths.

Lemma 4.10

If G is non-planar, then a subdivision of G is non-planar also.

Proof. Given a plane drawing of the subdivided graph, then by disregarding the vertices on the subdivided paths, we obtain a plane drawing of G (which is a contradiction). \square

Corollary 4.11

Subdivisions of $K_{3,3}$ and K_5 are non-planar.

Proof. Follows directly from the previous lemma. \square

What ties all of this together is Kuratowski’s theorem, which gives us a nice necessary and sufficient condition for a graph to be planar.

Theorem 4.12 (Kuratowski's Theorem)

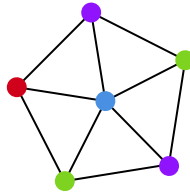
A graph G is planar if and only if G contains no subdivisions of $K_{3,3}$ or K_5 .

Proof. Omitted.

We will leave the topic of planar graphs here (for now), but we note that there are some other interesting notions that are related to planarity. For example, for what graphs is it possible to draw on a torus with no edge crossings?

§5 Graph Colouring

Informally, a graph colouring is just a way of colouring in different vertices of a graph, so that adjacent vertices are different colours. An example of a graph colouring is shown below.



§5.1 Basic Concepts

Of course, we need to define what this means in a slightly more mathematical sense, so we will define a colouring as follows.

Notation. We will write $[n] = \{1, 2, \dots, n\}$.

Definition 5.1 (r -Colouring)

Let $G = (V, E)$ be a graph. An r -colouring of G is a function $c : V \rightarrow [r]$ that satisfies $xy \in E \implies c(x) \neq c(y)$.

An r -colouring divides up a graph into r different **colour classes**, where there is only edges between the different colour classes.

Definition 5.2 (Chromatic Number)

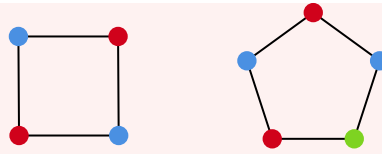
The **chromatic number** of a graph G , denoted $\chi(G)$ is the smallest r for which there exists an r -colouring.

Example 5.3 (Examples of Chromatic Numbers)

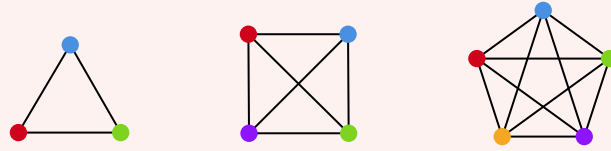
In the graph P_n , we have $\chi(P_n) = 2$ for $n \geq 2$, and $\chi(P_1) = 1$.



The graph C_n has $\chi(C_n) = 2$ if n is even, and $\chi(C_n) = 3$ if n is odd.



The complete graph K_n has $\chi(K_n) = n$.



The case for C_{2n} follows from a more general fact about bipartite graphs.

Proposition 5.4 (Chromatic Number of Bipartite Graphs)

If G is a bipartite graph, then $\chi(G) \leq 2$, and indeed $\chi(G) = 2$ unless $E = \emptyset$.

Proof Sketch. Colour the vertices in each part separately. □

Indeed, one way to think about chromatic number is as a generalization of bipartite graphs to multiple parts.

A straightforward observation to make is that the maximum degree of a graph (Δ) puts a bound on the chromatic number.

Proposition 5.5 (Degree Bound for Chromatic Number)

For a graph G , we have $\chi(G) \leq \Delta(G) + 1$. This bound is also sharp^a.

^aTake the complete graph or odd cycles. These are the only such examples though

To prove this, we are going to use a type of greedy algorithm.

Definition 5.6 (Greedy Colouring)

Given a graph $G = (V, E)$ with vertices $V = \{v_1, v_2, \dots, v_n\}$, the **greedy colouring** of G is a function $c_g : V \rightarrow \mathbb{N}$ defined inductively by

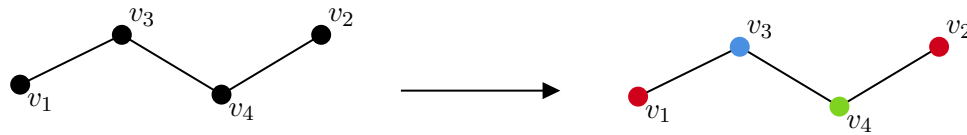
- $c_g(v_1) = 1$,
- Given coloured v_1, \dots, v_t , we have $c_g(v_{t+1}) = \min(\mathbb{N} \setminus \{c(v_i) \mid v_i \sim v_{t+1}, i \leq t\})$.

Proof of Proposition 5.5. Apply the greedy colouring to G with an arbitrary vertex ordering v_1, \dots, v_n . Then we note that

$$|\{c(v_i) \mid v_i \sim v_{t+1}, i \leq t\}| \leq \Delta(G),$$

and thus $c_g(v_{t+1}) \in [\Delta + 1]$. □

The ‘greedy’ approach need not give you any colouring that’s in any way optimal. For example, if we have the graph P_4 with vertices labelled as below, we get the following colouring from our greedy approach.



This is clearly not optimal.

§5.2 Brooks' Theorem

One step up from the greedy approach to obtaining a colouring comes in the form of Brooks' Theorem. Before we look at that, we will make an observation.

Proposition 5.7

Let G be a connected graph for which $\delta(G) < \Delta(G)$. Then $\chi(G) \leq \Delta(G)$.

Proof. We find a better ordering to apply the greedy colouring to. First define $v_n = v$, where $d(v) \leq \Delta(G) - 1$. Now choose an ordering of v_1, \dots, v_{n-1} so that

$$d(v_1, v) \geq d(v_2, v) \geq \dots \geq d(v_{n-1}, v).$$

We claim that each vertex v_i with $i \in [n-1]$ has at most $\Delta(G) - 1$ neighbors in v_1, \dots, v_{i-1} .

This is true for v_n by definition. For v_i with $i \neq n$, we observe that a shortest path P from v_i to v_n contains a neighbor v_j of v_i with $d(v_j, v) < d(v_i, v)$. Thus v_j comes later in the ordering (that is, $j > i$).

So the greedy colouring gives each vertex one of $\Delta(G)$ colours. \square

We can now extend this idea to prove Brook's theorem.

Theorem 5.8 (Brooks' Theorem)

Let G be a connected graph that is not complete or an odd cycle. Then $\chi(G) \leq \Delta(G)$.

Proof. Let G be a counterexample with a minimal number of edges. We may assume G is a regular graph (Δ -regular), and also that $\Delta \geq 3$.

Since G is not complete, there exists a vertex $v \in V(G)$ so that $G[N(v)]$ is not complete, and let $x, y \in N(v)$ where $x \sim y$ and $x \neq y$.

If G is 3-connected then $G' = G - x - y$ is connected. In this case, we order the vertices of G' by $v_1 = x$, $v_2 = y$, and $v_n = v$. Then we define v_3, \dots, v_{n-1} by ordering the vertices of G' so that

$$d_{G'}(v_3, v) \geq d_{G'}(v_4, v) \geq \dots \geq d_{G'}(v_{n-1}, v).$$

As we saw before, the greedy colouring with the ordering v_1, \dots, v_n will give us a Δ -colouring of G .

If G has a cut vertex $w \in V$, then let $G - w = C_1 \cup \dots \cup C_k$ be the components,

and $G_i = G[C_i \cup \{w\}]$. The maximum degree in G_i is bounded by Δ , and therefore G_i has a Δ -colouring for each i , by the minimality of the counter example. Note that G_i cannot be a complete graph on $\Delta + 1$ vertices since $d_{G_i}(w) \leq \Delta - 1$. By permuting the colours, we can assume each colouring gives the vertex w colour 1. Thus we have a colouring of G with Δ colours, which is a contradiction.

If $G - \{w_1, w_2\}$ is disconnected for $w_1 \neq w_2$ with $w_1, w_2 \in V$, then let C_1, \dots, C_k be the components of $G - \{w_1, w_2\}$. We may assume that G has no cut vertex by the previous case. Now let $G_i = G[C_i \cup \{w_1, w_2\}] + w_1w_2$. Note that $e(G_i) < e(G)$ since w_1, w_2 both send at least one edge to each component C_1, \dots, C_k . Also $d_{G_i}(w_1) \leq \Delta$ and likewise $d_{G_i}(w_2) \leq \Delta$, thus the maximum degree $G_i \leq \Delta$ and thus we can find a Δ -colouring of G_i for each i . That gives w_1, w_2 different colours. After permuting these colours, I can assume w_1 is coloured 1 in each, and w_2 is coloured 2 in each. Thus we can put all of the colourings together to obtain a colouring of G . \square

§5.3 Colouring Planar Graphs

We will now consider colourings on graphs that are planar. It is in this section that the theorem ever-present in the popular maths psyche is kept (without proof).

Theorem 5.9 (Four-Colour Theorem)

Let G be a planar graph. Then $\chi(G) \leq 4$.

Or informally: any map can be coloured with only four colours, where neighboring regions have different colours. Note that this is the dual form of the theorem above. This theorem was proved in 1976 by Appel & Haken and centers around reducing the theorem to a large number of cases, that were checked by computer.

We are going to prove two slightly weaker versions, that are still quite interesting.

Theorem 5.10 (Six Colour Theorem – Warmup)

Let G be a planar graph. Then $\chi(G) \leq 6$.

Proof. We will use induction on $n = |G|$. For $n = 1$, this is trivial. Now inductively, we claim that there is a vertex v with $\deg(v) \leq 5$. We note

$$\frac{1}{n} \left[\sum_{x \in V} d(x) \right] = \frac{2E}{n} \leq 6 - \frac{12}{n} < 6.$$

Thus there is a vertex v with $\deg(v) \leq 5$. By induction, $G - v$ is 6-colourable, and since v has at most 5 neighbors, there is a colour in $[6]$ that does not appear in $N(v)$. Colouring v with this colour, we get that the graph is 6-colourable. \square

Now we are going to kick it up a notch shortly, by introducing one more ingredient.

Definition 5.11

Given a graph G and an r -colouring of G , let $v \in V(G)$, and define the **$\{i, j\}$ -component of v** to be all of the vertices that can be reached starting at v along a

path using only colours i and j .

We make the following observation, which gives us an extra ‘move’ to use in the stronger proof

Proposition 5.12

Given a graph G with an r -colouring c , and for $i, j \in [r]$ with $i \neq j$, we can *swap* the colour on an $\{i, j\}$ -component to obtain a new colouring.

Proof Sketch. This works because of the ‘being reached’ condition in the $\{i, j\}$ -component definition. \square

Now we can prove the five colour theorem, using a lot of the ideas from the proof of the six colour theorem.

Theorem 5.13 (Five Colour Theorem)

Let G be a planar graph. Then $\chi(G) \leq 5$.

Proof. By induction on $n = |G|$, we note $n = 1$ is trivial. Now we proceed with the induction step. Let v be a vertex with $d(v) \leq 5$. Then apply induction to get a 5-colouring of $G - v$. Let $N(v) = \{x_1, \dots, x_5\}$, where x_1, \dots, x_5 are arranged in a clockwise manner. We may assume that $c(x_i) = i$ (otherwise we can colour v with the missing colour).

Now consider the $\{1, 3\}$ -component containing x_1 . If x_3 is not in this component, we can swap colours on this $\{1, 3\}$ component so that x_1 is colored with 3, and then colour v with 1.

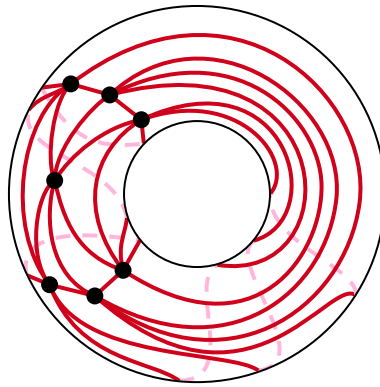
So we may assume there exists a path $x_1 \rightarrow x_3$ using colours 1 and 3 only. By the same argument there exists a path $x_2 \rightarrow x_4$ using colours 2 and 4 only. But then these paths must share a vertex, which is a contradiction. \square

§5.4 Colouring Graphs on (Other) Surfaces

Following on from the last section, we will consider a related guiding question:

If G is a graph, drawn on the torus, what can we say about $\chi(G)$?

More generally, if G is a graph drawn on a surface of genus g , what can we say about $\chi(G)$? For example, the graph K_7 can be drawn on a torus without edge crossings.



We may wonder if there is an Euler's formula for surfaces, and indeed there is, but instead of equality, we get a bound.

Theorem 5.14 (Euler's Formula for Surfaces)

If G is drawn in a surface of genus g , then $V - E + F \geq 2 - 2g$, where F is the number of connected components of $(\text{Surface} - G)$.

Proof Sketch. A similar inductive proof to the planar case. □

Remark. By a surface of genus g , we mean a compact orientable surface of genus g . Informally this is the surface formed from taking a sphere and adding g 'handles' to it. Also note that $2 - 2g$ is the **Euler characteristic** of a surface of genus g .

We can then use this to get a bound on the edges of a graph drawn on a surface with no edge crossings.

Proposition 5.15

If $G = (V, E)$ is a graph drawn on a surface of genus g , then $|E| \leq 3(|G| - (2 - 2g))$.

Proof Sketch. We have $3F \leq 2E$, then apply Euler's formula for surfaces. □

We can now get a bound (similarly to the planar case) on the chromatic number of a graph drawn on a surface.

Theorem 5.16 (Heawood's Theorem)

If G is a graph drawn on a surface of Euler characteristic E , with $E \leq 0^a$, then

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

^aThis condition really is needed, as otherwise for the planar case we get $\chi(G) \leq 4$, which is correct but this is not a proof for it!

Proof. Let $G = (V, E)$ be a given graph with $\chi(G) = k$. We may assume that G has the minimum number of edges, subject to $\chi(G) = k$. Observe that $\delta(G) \geq k - 1$, as otherwise there would be a vertex v with $d(v) = k - 1$, and thus $\chi(G - v) = k - 1$, and thus $\chi(G) = k - 1$. Also $k \leq n$, where $n = |V|$. Now the average degree of each vertex is

$$\frac{1}{n} \left[\sum_{v \in V} d(v) \right] = \frac{2e}{n} \leq 6 \left(1 - \frac{E}{n} \right),$$

where e is the number of edges and E is the Euler characteristic.

Thus

$$k - 1 \leq \delta(G) \leq \text{avg degree} \leq 6 \left(1 - \frac{E}{n} \right) \leq 6 \left(1 - \frac{E}{k} \right),$$

and so $k^2 - k \leq 6(k - E)$, then solving gives the required result. \square

Remark. It turns out that this estimate is sharp. Calling $H(E) = \left\lfloor \frac{7 + \sqrt{49 - 24E}}{2} \right\rfloor$, we find that $K_{H(E)}$ can be drawn on a surface of Euler characteristic E .

§5.5 Edge Colouring

Now instead of assigning vertices colours, we will assign them to edges.

Definition 5.17 (Edge Colouring)

Let $G = (V, E)$ be a graph. A **k -edge colouring** is a function $c : E \rightarrow [k]$ so that $c(e) \neq c(f)$ for all $e, f \in E$ such that $e \cap f \neq \emptyset$.

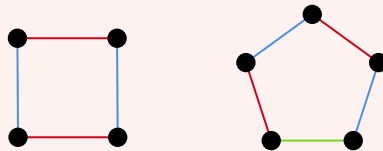
We have a similar notion to chromatic number for edge colourings.

Definition 5.18 (Edge Chromatic Number/Chromatic Index)

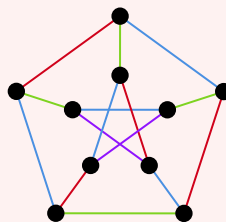
The **edge chromatic number** or **chromatic index** of a graph G is $\chi'(G)$, the minimum k for which a k -edge colouring exists.

Example 5.19 (Examples of Edge Chromatic Number)

The graph C_n has $\chi'(C_n) = 2$ if n is even, and $\chi'(C_n) = 3$ for n odd.



The Petersen graph G has $\chi'(G) = 4$.



In contrast to the chromatic number, it is much easier to get a handle on the edge chromatic number, as we will see in the following theorem.

Theorem 5.20 (Vizing's Theorem)

Let G be a graph. Then $\Delta(G) \leq \chi'(G) \leq \Delta(G) + 1$.

Proof. We proceed by induction on $e(G)$. The basis step is trivial, then for the inductive step we are given a graph G , and let's assume for a contradiction that $\chi'(G) > \Delta + 1$.

By induction, $G - e$ has a $\Delta + 1$ colouring, so let's write $e = xy$ with $x \neq y \in G$. We define vertices y_1, \dots, y_k inductively by setting $y_1 = y$. Now assume y_1, \dots, y_t are defined, and the colours missing from y_1, \dots, y_t are c_1, \dots, c_t respectively. Then if $c_t \notin \{c_1, \dots, c_{t-1}\}$, then let y_{t+1} be so that xy_{t+1} receives colour c_t . Note that such a vertex exists as otherwise we could recolour xy_1 with c_1 , xy_2 with c_2 , and so on until xy_t is with c_t , to obtain a $\Delta + 1$ colouring, which is a contradiction.

Now if $c_t \in \{c_1, \dots, c_{t-1}\}$, then stop. Say we stop after k steps, so I have defined y_1, \dots, y_k with missing colours c_1, \dots, c_k and $c_k = c_i$ for some $i < k$. We may assume that $i = 1$ (otherwise uncolour the edge xy_i and recolour xy_1 with c_1 , and so on until xy_{i-1} with c_{i-1}).

Let's call the colour missing at x , c_0 . Consider the $\{c_0, c_1\}$ component C , containing y_1 . If $x \notin C$, then we can flip colours on C so that c_0 is missing at y_1 , then colour xy_1 with c_0 . Likewise, the $\{c_0, c_1\}$ -component containing y_k must contain x , as otherwise we flip colours on this component so that the colour c_0 is missing at y_k . Then recolour xy_k to c_0 and xy_i to c_i for $i < k$.

Thus $x, y_1, y_k \in C$, the $\{c_0, c_1\}$ -component. But this is impossible since x, y_1, y_k all have one of the colours $\{c_0, c_1\}$ missing, thus $d_C(x), d_C(y_1), d_C(y_k) \leq 1$. But this is impossible for a path or cycle. \square

Remark. This theorem is *not* true if we generalize to multigraphs, which are graphs that have multiple edges between vertices.

§6 Extremal Graph Theory

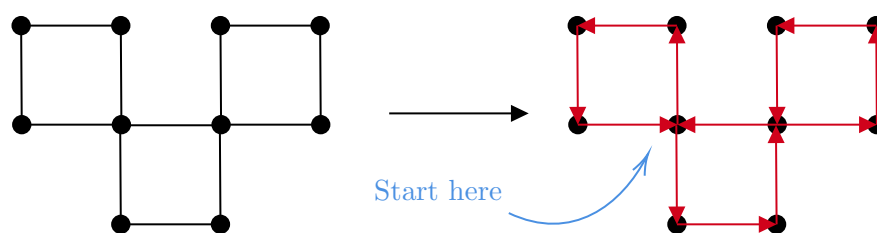
Welcome to the section on extremal graph theory, where we think about problems involving things with an 'extreme' flavour.

§6.1 Eulerian Circuits and Hamiltonian Cycles

We are going to begin our chapter on extremal graph theory with something that's not *really* extremal graph theory. Still, it will lead us nicely into the more extreme parts of this chapter and the course. Let's state a definition.

Definition 6.1 (Eulerian Circuit)

An **Eulerian circuit** is a circuit in a graph G that crosses each edge exactly once. If a G has an Eulerian circuit, we say that the graph is **Eulerian**.



What's nice is that there is a straightforward characterisation of Eulerian graphs.

Theorem 6.2 (Euler's Theorem)

A connected graph has an Eulerian circuit if and only if every vertex has even degree.

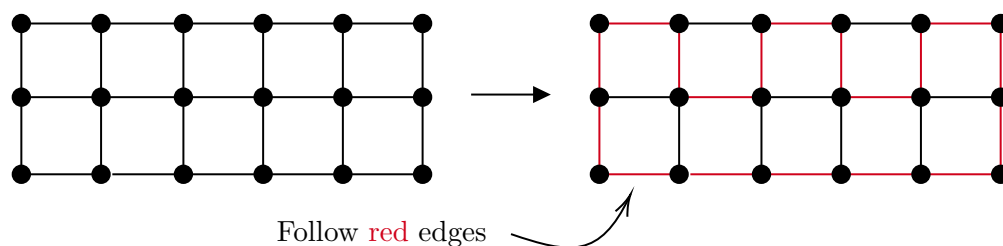
Proof. If a graph G has an Eulerian circuit, then the degree of each vertex must be even. This is because a fixed vertex x is entered and exited a fixed number of times.

Now if each vertex of a graph G has even degree, we will apply induction on $e(G)$. If $e(G) = 0$, then we are done. Now if $d(x) \geq 1$ for all vertices x , then $d(x) \geq 2$, and G contains a cycle C . Define $G' = G - E(C)$. Let G_1, \dots, G_k be the components of G' . The degree of each G_i has all degrees even. Thus by induction, there is an Eulerian circuit W_1, \dots, W_k for each G_1, \dots, G_k respectively. Thus we can combine C with W_1, \dots, W_k to obtain an Eulerian circuit for all of G . \square

Now we are going to define a similar looking notion, but it will not end up being so well behaved as Eulerian circuits.

Definition 6.3 (Hamiltonian Cycle)

Let G be a graph, then a **Hamiltonian cycle** in G is a cycle that visits each vertex exactly once. We say G is **Hamiltonian** if it contains a Hamiltonian cycle.



Our first question might be 'is there an iff type condition for Hamiltonian cycles', and so far there is no such property known. However there are some results that give us information about whether a graph is Hamiltonian.

Theorem 6.4 (Dirac's Theorem)

Let G be a graph with $n \geq 3$ such that $\delta(G) \geq n/2$. Then G contains a Hamiltonian cycle.

Proof. For a contradiction, assume that G is a counterexample to the theorem. Note that G is connected, since $x \not\sim y$ then $N(x), N(y) \subset V \setminus \{x, y\}$ and $|N(x)|, |N(y)| \geq \frac{n}{2}$, and thus $N(x) \cap N(y) \neq \emptyset$, and thus $d(x, y) \leq 2$.

Let x_1, \dots, x_ℓ be the longest path in G . Note that x_1, \dots, x_ℓ does not form a cycle, as otherwise if $\ell = n$, then we have a contradiction, and if $\ell < n$ then there exists $y \notin \{x_1, \dots, x_\ell\}$ that is $y \sim x_i$. Thus we can find a longer path (which is also a contradiction).

Now if there exists $i \in \{1, \dots, \ell-1\}$ so that $x_i \sim x_\ell$ and $x_{i+1} \sim x_1$, then the vertices x_1, \dots, x_ℓ form a cycle. This contradicts the above.

Define

$$N^+(x_\ell) = \{x_i \mid x_{i-1} \in N(x_\ell), 2 \leq i \leq \ell\}.$$

We have $N^+(x_\ell) \cap N(x_1) = \emptyset$, but this is impossible since $|N^+(x_\ell)| \geq n/2$, $|N(x_1)| \geq n/2$, and $N^+(x_\ell), N(x_1) \subseteq \{x_2, \dots, x_\ell\}$. Thus we have a contradiction. \square

Remark. We never *really* used this $\delta(G) \geq n/2$ condition fully. It suffices to have $d(x) + d(y) \geq n$ for $x \not\sim y$.

We can use this same argument to prove a more general result about paths.

Proposition 6.5

Let G be a connected graph. Let $k < n$ and assume $\delta(G) \geq k/2$. Then $G \geq P_{k+1}$.

Proof Sketch. Similar to Dirac's theorem. Choose a longest path x_1, \dots, x_ℓ in G . They don't form a cycle, and then we can force a configuration like $x_1 \sum x_{i+1}$, $x_\ell \sum x_i$ by using the minimum degree condition. \square

We may wonder in the above if it's possible to replace 'path' with 'cycle' in the above. The answer is, sadly, no.

A natural question is to wonder how many edges are needed to 'force' a (for example) triangle. For a Hamilton cycle, this sort of question would have been a bit strange, and a more natural question would be what is the minimum value of $\delta(G)$ to ensure that G contains a Hamilton cycle. In the next theorem, we will give a result that answers a simpler sort of question.

Theorem 6.6

Let G be a graph. If $e(G) > \frac{n}{2}(k-1)$, then G contains a path of length k .

Proof. We will prove the contrapositive: If G is P_{k+1} free, then $e(G) \leq \frac{n}{2}(k-1)$.

We will apply induction on n . This is true for $n = 2$. Then given a graph G on $|G| = n \geq 3$ vertices, if G is disconnected let G_1, \dots, G_k be the components of G . Then each of these has $G_i \not\supseteq P_{k+1}$. Then by induction we have $e(G_i) \leq \frac{n(G_i)(k-1)}{2}$. Thus

$$e(G) = \sum_{i=1}^k e(G_i) \leq \sum_{i=1}^k \frac{n(G_i)(k-1)}{2} = \left(\frac{k-1}{2}\right) n,$$

and we are done.

Now if there is a vertex v with $d(v) \leq \frac{k-1}{2}$, then consider $G - v$. Then $e(G - v) =$

$e(G) - d(v)$, and also $G - v$ does not contain P_{k+1} . Thus $e(G - v) \leq \frac{n-1}{2}(k-1)$. So

$$e(G) = e(G - v) + d(v) \leq \frac{n-1}{2}(k-1) + \frac{k-1}{2} \leq \frac{n}{2}(k-1).$$

Thus we can assume G is connected and $d(v) \geq \frac{k}{2}$. We can apply [Proposition 6.5](#) to find a path of length k , that is, $P_{k+1} \subseteq G$, which is a contradiction.

[We may also assume $k < n$ for if $k = n$ then $e(G) \leq \frac{k}{2}(n-1)$, which reduces to showing $e(G) \leq \binom{n}{2}$, which is trivial]. \square

§6.2 Complete Subgraphs and Turan's Theorem

Let's return to the question of how many edges are needed in a graph to guarantee the existence of a triangle. Later on, we will look more generally at the question of the number of edges needed to guarantee a K_n subgraph.

Theorem 6.7 (Mantel's Theorem)

If $e(G) > \frac{n^2}{4}$, then $G \supset K_3$, and this is sharp.

Proof. Given a K_3 -free graph G and $x, y \in V$ such that $x \sim y$ and $d(x) + d(y) \leq n$, and let $m = e(G)$. Then summing we get

$$\sum_{xy \in E} d(x) + d(y) \leq mn.$$

We have that this is $\sum_x \sum_y d(x) \mathbb{1}(xy \in E) = \sum_{x \in V} (d(x))^2$. Now $\sum_{x \in V} d(x) = 2m$, and by Cauchy-Schwarz we have $(\sum d(x))^2 \leq n \sum d(x)^2$, and thus

$$mn^2 \geq (2m)^2 \implies \frac{n^2}{4} = e(G).$$

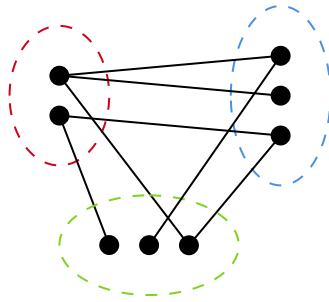
To see that this is sharp, consider the complete bipartite graph with n vertices. Then this has $n^2/4$ edges, but no triangle. Thus this is sharp. \square

The next question we are going to answer is how many edges are needed to guarantee the existence of a K_n subgraph. In the previous theorem, to see that our result was sharp, we considered the complete bipartite graph on n vertices. To discuss this more general question, a natural thing to look at is a generalisation of bipartite graphs.

Definition 6.8 (r -Partite Graph)

An **r -partite graph** is a graph of the form $G = (V, E)$ where the vertices are partitioned into r subsets $V = V_1 \cup \dots \cup V_r$ with $V_i \cap V_j = \emptyset$ if $i \neq j$, and $xy \notin E$ for all $x, y \in V_i$.

For example, the graph below is 3-partite.



It's worth noting that saying a graph is r -partite is the same as saying the graph has chromatic number of at most r . We care particularly about the case of an r -partite graph where all possible edges are included.

Definition 6.9 (Complete r -Partite Graph)

We say an r -partite graph is **complete** if no edge can be added with the graph remaining r -partite.

Now if we consider a complete r -partite graph G on n vertices, with all parts of size n/r , then we would have

$$e(G) = \binom{n}{r}^2 \binom{r}{2} = \left(1 - \frac{1}{r}\right) \frac{n^2}{2}.$$

The graph G can also have no K_{r+1} subgraph, since that would imply there's an edge between two vertices in the same part. This gives us a lower bound on the number of edges needed to have a K_{r+1} subgraph, and it turns out that bound is sharp.

Theorem 6.10 (Turan's Theorem)

If $e(G) > \left(1 - \frac{1}{r}\right) \frac{n^2}{2}$, then $G \supset K_{r+1}$, and this is sharp.

Proof. Suppose we have some graph G on n vertices that had no K_{r+1} subgraph, and also that the result holds up to n, r . If $r \geq n$ then we are clearly done, so we may suppose that $n > r$. Let A be a K_r subgraph in G , which must exist by our assumption. Let $B = V - A$. Then we have $e(G) = e(A) + e(B) + e(A, B)$, where $e(A, B)$ denotes the number of edges between vertices in A and B . We have an upper bound of

$$e(G) \leq \binom{r}{2} + \left(1 - \frac{1}{r}\right) \frac{(n-r)^2}{2} + (n-r)(r-1),$$

where $e(A, B) \leq (n-r)(r-1)$ since each vertex in B has at most $r-1$ neighbours in A . We can rewrite this as

$$e(G) \leq \frac{1}{2} \left(1 - \frac{1}{r}\right) (r^2 + (n-r)^2 + (n-r)r) = \frac{1}{2} \left(1 - \frac{1}{r}\right) n^2,$$

as required. □

There's another non-examinable proof called Zykov Symmetrisation

§6.3 The Zarankiewicz Problem

Turan's theorem answers the question 'how many edges can a graph have without having a K_t subgraph'. A natural follow on question is 'how many edges can a bipartite graph have without having a $K_{t,t}$ subgraph'. This is the Zarankiewicz problem, and unfortunately does not yet have a definite answer like the previous question. We can however obtain some bounds.

Definition 6.11

We define $Z(n, t)$ to be the maximum number of edges in a bipartite graph with n vertices in each part and no $K_{t,t}$ subgraph.

We really care about $Z(n, t)$ where t is fixed and n is large. We will prove the following theorem.

Theorem 6.12

We have $Z(n, t) \leq t^{1/t} n^{2-1/t} + tn$ for all n .

Proof. Given a graph $G = (A \cup B, E)$ where $|A| = |B| = n$, and $G \not\supseteq K_{t,t}$, we want to show that $m = e(G) \leq t^{1/t} n^{2-1/t} + tn$.

For some set $S \subseteq A$, and $|S| = t$, we have

$$\left| \bigcap_{x \in S} N(x) \right| \leq t - 1,$$

as otherwise we would have a $K_{t,t}$ subgraph. Then averaging over such subsets S , we have

$$\binom{n}{t}^{-1} \sum_{\substack{S \subseteq A, \\ |S|=t}} \left| \bigcap_{x \in S} N(x) \right| \leq t - 1.$$

This sum can be written as

$$\begin{aligned} \sum_{\substack{S \subseteq A, \\ |S|=t}} \left| \bigcap_{x \in S} N(x) \right| &= \sum_{\substack{S=\{x_1, \dots, x_t\}, \\ S \subseteq A}} \sum_y \mathbb{1}(y \sim x_1, y \sim x_2, \dots, y \sim x_t) \\ &= \sum_y \sum_{\substack{S=\{x_1, \dots, x_t\}, \\ S \subseteq A}} \mathbb{1}(y \sim x_1, y \sim x_2, \dots, y \sim x_t) \\ &= \sum_y \binom{d(y)}{t}. \end{aligned}$$

We may assume that $d(y) \geq t - 1$ for all y , as otherwise we can add an edge incident with y and not create a $K_{t,t}$ subgraph. Then by convexity and since $d(y) \geq t - 1$, we have

$$\sum_y \binom{d(y)}{t} \geq \sum_{y \in B} \binom{d}{t} = n \binom{d}{t},$$

where $d = \frac{1}{n} \sum_{y \in B} d(y) = m/n$. Combining this inequality with what we obtained previously, we get the bound

$$t - 1 \geq \frac{n \binom{d}{t}}{\binom{n}{t}} = n \frac{d(d-1) \cdots (d-t+1)}{(n-1) \cdots (n-t+1)} \geq \frac{(d-t+1)^t}{n^{t-1}},$$

thus $(t-1)^{1/t} n^{t-1} \geq d-t+1$, and $t^{1/t} n^{2-1/t} + tn \geq m$, as required. \square

While we do have this upper bound, we don't know $Z(n, t)$ for most values of n . We do know, for example, that

$$\begin{aligned} cn^{3/2} &\leq Z(n, 2) \leq 2n^{3/2}, \\ c'n^{5/2} &\leq Z(n, 3) \leq 2n^{5/3}, \end{aligned}$$

for all large n , and some constants c, c' . Still we don't even know $t = 4$. The constructions are based on finite geometry. As an example, we will sketch the construction that shows the bound $cn^{3/2} \leq Z(n, 2)$.

Theorem 6.13

For infinitely many n we have $Z(n, 2) \geq cn^{3/2}$ for $c > 0$.

Proof Sketch. Let p be a prime, and consider^a $(\mathbb{Z}/p\mathbb{Z})^2$. Define a line $L = \{(x, ax + b) \mid x \in \mathbb{Z}/p\mathbb{Z}\}$, for some $a, b \in \mathbb{Z}/p\mathbb{Z}$ with $a \neq 0$.

We need the following straightforward facts:

1. Two distinct lines intersect in at most one point.
2. Each line contains p points.
3. There are p^2 points, $p(p-1) \approx p^2$ lines^b.

We are going to construct a graph bipartite graph $G = (A \cup B, E)$ where every vertex in A corresponds to a line and every vertex in B corresponds to a point. There's approximately p^2 vertices in each of these parts, so we will take n to be about p^2 . Then for $\ell \in A$ and $x \in B$, we define $\ell \sim x$ if $x \in \ell$. Then $e(G) \approx p^2 \cdot p = p^3 = n^{3/2}$. To see this graph works, if G contained a $K_{2,2}$ then there would exist two lines ℓ_1, ℓ_2 with $|\ell_1 \cap \ell_2| \geq 2$, which is a contradiction. \square

^aYou can consider this to be like a torus.

^bTo make this a proper proof, we would have to fix all of the approximations

§6.4 The Erdős-Stone Theorem

So far we have been interested in questions of the form 'how many edges do I need in a graph before I force a subgraph H '. We've been able to prove a variety of interesting but relatively scattered results of this form, but it turns out that if we consider the question asymptotically, we can obtain a more general unifying result.

We begin by defining the following function.

Definition 6.14 (Extremal Function)

Let H be a fixed graph. We define the **extremal function** $\text{ex}(n, H)$ to be

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

Using this function we can restate some of the results we previously obtained, such as Mantel's theorem giving us $\text{ex}(n, K_3) \leq n^2/4$, and Turan's theorem giving us $\text{ex}(n, K_{r+1}) \leq (1 - 1/r)n^2/2$.

Looking a bit more at Turan's theorem, if we consider things asymptotically on the scale of n^2 , we get a weaker version of Turan that looks like

$$\lim_{n \rightarrow \infty} \frac{\text{ex}(n, K_{r+1})}{\binom{n}{2}} = \left(1 - \frac{1}{r}\right).$$

The Erdős-Stone Theorem allows us to get these types of results generally, using just the chromatic number of the graph.

Theorem 6.15 (Erdős-Stone Theorem)

Let H be a graph with $\chi(H) = r$, and $r \geq 2$. Then

$$\lim_{n \rightarrow \infty} \frac{\text{ex}(n, H)}{\binom{n}{2}} = \left(1 - \frac{1}{r-1}\right).$$

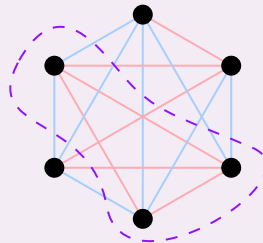
Proof. Omitted

§7 Ramsey Theory

We are going to jump into Ramsey theory, an area of graph theory that has a similar flavour to previous section. Ramsey theory is often described as the collection of results showing 'complete disorder is impossible', but its easiest to see what we study with an example.

Proposition 7.1 (Monochromatic Triangles)

Colour each of the edges of K_6 red or blue^a. Then there must be a monochromatic triangle.



^aNote that we don't have any restrictions on the colouring.

Proof. Consider some vertex v in the graph. We know that at least three incident edges must be the same colour, so suppose (without loss of generality) that they are red, with endpoints v_1, v_2, v_3 . Then either $v_1v_2v_3$ is a monochromatic blue triangle, or some v_iv_j is red, in which case vv_iv_j is a monochromatic red triangle. \square

Straight away we might wonder about K_4 . Does there exist some large enough n such that every 2-colouring of K_n contains a monochromatic K_4 ? This is the type of problem we will consider in Ramsey theory.

§7.1 Ramsey Numbers

We will begin our discussion of Ramsey theory by introducing the *Ramsey numbers*.

Definition 7.2 (Ramsey Numbers)

For $t \in \mathbb{N}$, we define $R(t)$ the **t th Ramsey number** to be the smallest n for which every 2-colouring of K_n contains a monochromatic K_t .

We already know that $R(3) = 6$ (since we showed there's always a monochromatic triangle in K_6 , and it's not hard to find a colouring of K_5 that doesn't have one). We haven't yet proved though that such an $R(t)$ exists though, so we will clear this up first.

Theorem 7.3 (Ramsey's Theorem)

For all $t \in \mathbb{N}$, $R(t)$ is finite and $R(t) \leq 4^t$.

Proof. We will begin by proving a small lemma. For $s, t \geq 2$, we define the Ramsey number $R(s, t)$ to be the smallest n such that every red/blue colouring of K_n either contains a red K_s or a blue K_t .

We will prove that for all $s, t \geq 2$,

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

Claim. Assume that $R(s-1, t)$ and $R(s, t-1)$ exist. Then $R(s, t)$ exists and $R(s, t) \leq R(s-1, t) + R(s, t-1)$.

Let $a = R(s-1, t)$ and $b = R(s, t-1)$ and let c be a red/blue colouring of K_{a+b} . Let x be a vertex in this graph, and we note $d(x) = a + b - 1$. Then x has either

- (i) a red neighbors
- (ii) b blue neighbors.

In case (i), let N_r be the red neighbors of x , and note $|N_r| \geq a$. Then the colouring induced on N_r contains either a K_{s-1} in red or a K_t in blue. In the latter case we are done, and in the former case we can add x to K_{s-1} to finish. Case (ii) is symmetric, and thus our claim is true.

Now we can return to showing (\dagger) . We are going to induct on $s + t$. Note that $R(s, 2) = s$ and $R(2, t) = t$. Inductively assume that $R(s-1, t), R(s, t-1)$ exist

and satisfy this relation, then

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

as desired. Taking $R(t) = R(t, t)$ gives $R(t) \leq \binom{2t-2}{t-1} \leq 4^t$. \square

Ramsey numbers are generally quite mysterious. We know for example that $R(3, 3) = 6$, $R(4, 4) = 18$, and $R(3, 7) = 23$, but even something like $R(5, 5)$ is unknown, and our best bounds are $43 \leq R(5, 5) \leq 48$. The numbers $R(3, t)$ are quite well understood, and for $\epsilon > 0$ we have (for sufficiently large t)

$$\left(\frac{1}{4} + \epsilon\right) \frac{t^2}{\log t} \leq R(3, t) \leq (1 + \epsilon) \frac{t^2}{\log t}.$$

A big question though is getting a lower bound on $R(t)$, which is currently nowhere as close as our upper bound.

§7.2 Infinite Ramsey

Now instead of finite complete graphs, we are going to consider an infinite graph.

Definition 7.4 (Complete Countable Graph)

We define the **complete countable graph** $G = (\mathbb{N}, \mathbb{N}^{(2)})$.

Notation. For a set X , we let $X^{(r)} = \{A \subseteq X \mid |A| = r\}$, the set of r element subsets of X .

Our motivating question will be the following: given a 2-colouring of the complete countable graph, what can we say about the monochromatic structures that appear in the colouring? Obviously by Ramsey's theorem we can find arbitrarily large monochromatic complete graphs, but is it possible to find a monochromatic complete countable subgraph?

Example 7.5 (Complete Countable Graph with Complete Countable Subgraph)

Consider the colouring of a complete countable graph where xy is red if $x+y$ is even, and blue if $x+y$ is odd. Then the subgraph induced on the vertices $\{2, 4, 8, \dots\}$ is a monochromatic complete countable subgraph.

Now consider instead the colouring where $x+y$ is red if it has an even number of prime factors, and $x+y$ is blue if it has an odd number of prime factors. Can you find a monochromatic complete countable subgraph?

It turns out that by another theorem of Ramsey, it is always possible to find such a subgraph.

Theorem 7.6 (Infinite Ramsey's Theorem)

Let $G = (\mathbb{N}, \mathbb{N}^{(2)})$, the complete countable graph. Then for every 2-colouring of G there exists an infinite set $X \subseteq \mathbb{N}$ so that $X^{(2)}$ is monochromatic.

Proof. We begin by inductively defining vertices x_1, x_2, \dots, x_t and colours c_1, c_2, \dots, c_t such that $N_{c_i}(x_i) \supset \{x_{i+1}, \dots, x_t\}$, and $\bigcap_{i=1}^t N_{c_i}(x_i)$ is infinite, where $N_{c_i}(x_i)$ is the neighborhood of x_i in the colour c_i .

We can do this by taking x_1, \dots, x_t and then choosing any $x_{t+1} \in \bigcap_{i=1}^t N_{c_i}(x_i)$. Then we can define c_i accordingly, since at least one of $N_{\text{red}}(x_{t+1}) \cap (\bigcap_{i=1}^t N_{c_i}(x_i))$ is infinite, or $N_{\text{blue}}(x_{t+1}) \cap (\bigcap_{i=1}^t N_{c_i}(x_i))$ is infinite.

This then gives us an infinite collection x_1, x_2, \dots so that $x_i x_j$ with $i < j$ are given colour c_i . Then since both

$$\{x_i \mid c_i = \text{red}\} \quad \text{and} \quad \{x_i \mid c_i = \text{blue}\}$$

are monochromatic cliques, at least one of them is infinite, giving a monochromatic complete countable subgraph. \square

Remark. By keeping track of the sizes of sets at each step, we can adapt this proof to give us a different proof of the finite case.

§7.3 Ramsey Numbers for Hypergraphs

Before, we were colouring $X^{(2)}$, the pairs of a set X (the vertices of a graph). Of course, we might wonder what happens when we colour r -sets?

Definition 7.7 (Hypergraph Ramsey Numbers)

We define $R^{(r)}(s, t)$ to be the smallest n such that every red/blue colouring of $[n]^{(r)}$ either contains a set $X \subseteq [n]$ with $|X| = s$ and $X^{(r)}$ red, or there exists $Y \subseteq [n]$ with $|Y| = t$ and $Y^{(r)}$ blue.

And of course, it's not guaranteed that these exist, but it is possible to show that they do.

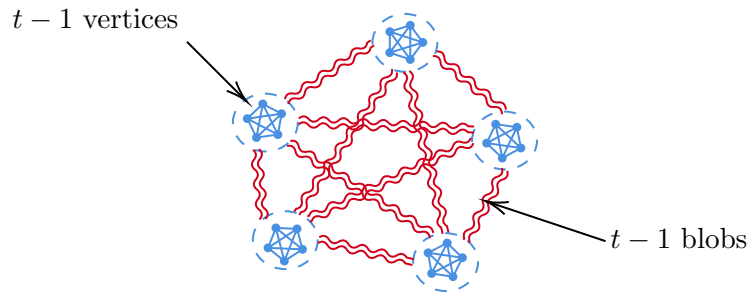
Theorem 7.8 (Hypergraph Ramsey Theorem)

For all $r, s, t \geq 2$, $R^{(r)}(s, t)$ exists.

Proof. Omitted.

§7.4 Lower Bounds for Ramsey Numbers

We said and the end of [subsection 7.1](#) that we were left to get a good lower bound on the Ramsey numbers $R(t)$. One way to do that might be by considering the following extremal-seeming construction:



This clearly doesn't have a monochromatic K_t subgraph, and we can compute the number of vertices is $(t-1)^2$, giving

$$R(t) \geq (t-1)^2.$$

However, it turns out that this quadratic bound is, quite frankly, awful. To show something better, we are going to need a new idea – considering random graphs and using probability.

Theorem 7.9 (Erdős)

$$R(t) \geq 2^{\frac{t}{2}}.$$

Proof. We will show that for a random colouring of the edges of K_n ,

$$\mathbb{P}(\text{there exists a monochromatic } K_t \subset K_n) \leq 2 \binom{n}{t} 2^{-\binom{t}{2}}.$$

Once we have shown this, we will then note for $n = 2^{t/2}$ that the RHS of this inequality is strictly less than 1, which implies that is *some* colouring of K_n with no monochromatic K_t .

To prove this inequality, we first fix some $A \subseteq K_n$ with $|A| = t$, and note that

$$\mathbb{P}(A^{(2)} \text{ is monochromatic}) = 2 \cdot 2^{-\binom{t}{2}}.$$

Applying the union bound then gives us our desired result

$$\mathbb{P}(\text{there exists a monochromatic } K_t \subset K_n) \leq \sum_{A \in K_n^{(t)}} 2 \cdot 2^{-\binom{t}{2}} = 2 \binom{n}{t} 2^{-\binom{t}{2}}.$$

Then to show that this is less than 1 for $n = 2^{t/2}$, we can take

$$2 \binom{n}{t} 2^{-\binom{t}{2}} \leq 2 \frac{n^t}{t!} 2^{\frac{t(t-1)}{2}} = \left(\frac{2^{1/t}}{t^{1/t}} \cdot n 2^{-\frac{t-1}{2}} \right)^t,$$

and thus it is indeed less than 1 for $n = 2^{t/2}$. \square

The proof idea we employed is known as the *probabilistic method*, and is an immensely powerful technique that can be used to show all kinds of interesting results.

Getting much better bounds on $R(t)$ is still a major open problem, and surprisingly the bounds we have proved so far aren't far from state of the art – the base of the

exponential term is still the same in best known results. Also, there's currently no known construction that gives an exponential lower bound directly.

§8 The Probabilistic Method

At the end of our discussion on Ramsey numbers, we used an argument that combines probability and random graphs to prove a result. In this section we are going to look a bit more closely at the type of argument that we used.

§8.1 Random Graphs

We first are going to come up with a probability space involving a graph where each edge is included independently at random with probability p .

Definition 8.1 (Binomial Random Graph)

Let $n \in \mathbb{N}$ and $p \in [0, 1]$, let $\mathcal{G}(n, p)$ be the probability space on all graphs with vertex set $\{1, 2, \dots, n\}$, with $\mathbb{P}(ij \in E(\mathcal{G})) = p$, and every edge included/excluded independently.

Remark. For our ' $R(t) \geq 2^{t/2}$ ' result, we were working with $G(n, 1/2)$, which is the same as sampling uniformly on a graph on $\{1, \dots, n\}$.

Another use of random graphs is in the Zarankiewicz problem. Recall our definition of $Z(n, t)$. We proved already that $Z(n, t) \leq 2n^{2-1/t}$, and we hadn't yet proved a lower bound. We are able to come up with one using the probabilistic method.

Theorem 8.2

We have $Z(n, t) \geq \frac{1}{4}n^{2-\frac{2}{t+1}}$.

Proof. Let G be a random bipartite graph with parts A, B and $|A| = |B| = n$. Then, define \tilde{G} to be G with an edge removed from each $K_{t,t}$ in G . Then

$$e(\tilde{G}) \geq e(G) - \#(K_{t,t}\text{'s in } G),$$

and by linearity of expectation

$$\mathbb{E}[e(\tilde{G})] \geq \mathbb{E}[e(G)] - \mathbb{E}[\#(K_{t,t}\text{'s in } G)].$$

We have $\mathbb{E}[e(G)] = pn^2$, and also $\mathbb{E}[\#(K_{t,t}\text{'s in } G)]$ can be written using indicator functions as

$$\mathbb{E}\left[\sum_{A', B'} \mathbb{1}(A', B' \text{ induce a } K_{t,t})\right] = \binom{n}{t}^2 p^{t^2},$$

where we sum over $A' \leq A$, $B' \leq B$ with $|A'| = |B'| = t$. Thus

$$\mathbb{E}[e(\tilde{G})] \geq pn^2 - \binom{n}{t}^2 p^{t^2}. \quad (*)$$

We want the RHS of this to be positive, so it's enough to have

$$pn^2 \geq n^{2t}p \iff n^{-\frac{2}{t+1}} \geq p,$$

so if $p = n^{-2/(t+1)}$ then $(*) \geq pn^2/2 \geq n^2/2 - 2/(t+1)$, as desired. \square

This proof is an example of the ‘modification method’.

§8.2 Random Graphs and Chromatic Number

One big picture question in graph theory is ‘what makes $\chi(G)$ large?’ A possible answer is that G contains a large clique, but that's not the whole story. For example, there's a triangle free graph with arbitrarily large χ . Another possible answer is that G has lots of short cycles, but it turns out that this is false too.

Definition 8.3 (Girth)

The **girth** of a graph G is the length of the shortest cycle in G , denoted $\text{girth}(G)$.

Definition 8.4 (Independent Set)

An **independent set** in a graph is a subset of vertices no two of which are adjacent. We will denote the size of the largest independent set $\alpha(G)$.

We can prove that graphs with large girth can also have arbitrarily large chromatic number. We will first need to note that if G is a graph, then

$$\chi(G) \geq \frac{n}{\alpha(G)}.$$

Theorem 8.5

For every $k, g \in \mathbb{N}$, there exists a graph G with $\text{girth}(G) \geq g$ and $\chi(G) \geq k$.

Proof. Let G be a random graph sampled from $\mathcal{G}(n, p)$ where $p = n^{\frac{1}{g}-1}$. We form the new graph \tilde{G} be the graph where we remove a vertex from each cycle of length $g-1$.

We are going to do a few steps.

1. We want to show that $\mathbb{P}(|\tilde{G}| \geq n/2) \rightarrow 1$ as $n \rightarrow \infty$.
2. We want to use $\chi(\tilde{G}) \geq |\tilde{G}|/\alpha(\tilde{G})$, noting that $\alpha(\tilde{G}) \leq \alpha(G)$. To do this, we will prove $\mathbb{P}(\alpha(G) \leq n/2k) \rightarrow 1$ as $n \rightarrow \infty$.

Step 1. Let X_i be the number of cycles of length i in G , and let $X = \sum_{i=1}^{g-1} X_i$. We can write

$$X_i = \sum_{x_1, \dots, x_i \in [n]} \mathbb{1}(x_1 \dots x_i \text{ are a cycle in } G).$$

Taking expectations,

$$\mathbb{E}[X_i] = \sum_{x_1, \dots, x_i \in [n]} \mathbb{P}(x_1 \dots x_i \text{ are a cycle in } G) \leq n^i p^i.$$

Using this we have

$$\mathbb{E}[\# \text{ of cycles of length } \leq g-1] = \mathbb{E}[X] = \sum_{i=1}^{g-1} (np)^i \leq gn^{\frac{g-1}{g}}.$$

Using Markov's inequality,

$$\mathbb{P}(X > n/2) \leq \frac{\mathbb{E}[X]}{n/2} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

thus $\mathbb{P}(|\tilde{G}| \geq n/2) \rightarrow 1$, as we needed.

Step 2. Let Y_t be the number of independent sets of size $t = n/2k$. Then

$$Y_t = \sum_{I \in [n]^{(t)}} \mathbb{1}(I \text{ is independent}),$$

and taking expectations

$$\mathbb{E}[Y_t] = \sum_{I \in [n]^{(t)}} \mathbb{P}(I \text{ is independent}) = \binom{n}{t} (1-p)^{\binom{t}{2}} \leq n^t e^{-p \binom{t}{2}} = \left(n e^{-p \frac{t-1}{2}} \right)^t,$$

and with $p = n^{1/g-1}$ and $t = n/2k$, so the RHS tends to 0 as n tends to infinity. Then

$$\mathbb{P}(Y_t \geq 1) \leq \mathbb{E}[Y_t] \rightarrow 0,$$

and $\mathbb{P}(Y_t \geq 1) = \mathbb{P}(\alpha(G) \geq t) = \mathbb{P}(\alpha(G) \geq n/2k)$. Putting that together, $\chi(\tilde{G}) \geq \frac{|\tilde{G}|}{\alpha(\tilde{G})} \geq \frac{|\tilde{G}|}{\alpha(G)} \geq \frac{n/2}{n/2k} \geq k$ holds with probability tending to 1 as n tends to infinity, finishing our proof. \square

§8.3 Structure of Random Graphs

So what does a typical graph sampled from $\mathcal{G}(n, p)$ look like? What does its structure look like as p varies from 0 to 1? As a first example, we will consider the following more precise question:

For what p does G sampled from $\mathcal{G}(n, p)$ contain a K_t ?

Theorem 8.6

Let G be sampled from $\mathcal{G}(n, p(n))$, and let $t \in \mathbb{N}$. Then $\lim_{n \rightarrow \infty} \mathbb{P}(G \text{ contains a } K_t) \rightarrow 1$ if $p(n)n^{\frac{2}{t-1}} \rightarrow \infty$. Moreover, $\lim_{n \rightarrow \infty} \mathbb{P}(G \text{ contains } K_t) \rightarrow 0$ if $p(n)n^{\frac{2}{t-1}} \rightarrow 0$.

Proof. Define X_t to be the number of K_t s in G . Then

$$X_t = \sum_{A \in [n]^{(t)}} \mathbb{1}(G[A] \text{ is complete}),$$

and

$$E[X_t] = \sum_{A \in [n]^{(t)}} \mathbb{P}(G[A] \text{ is complete}) = \binom{n}{t} p^{\binom{t}{2}}.$$

We note that $\binom{n}{t} p^{\binom{t}{2}} \leq n^t p^{\binom{t}{2}} = \left(n^{\frac{2}{t-1}} p\right)^{\binom{t}{2}}$, and this tends to 0 or ∞ if $n^{\frac{2}{t-1}} p$ tends to 0 or ∞ respectively. So

$$\mathbb{P}(G \text{ contains a } K_t) = \mathbb{P}(X_t \geq 1) \leq \mathbb{E}[X_t] \rightarrow 0 \quad \text{if} \quad n^{\frac{2}{t-1}} p \rightarrow 0.$$

We now need to show that $\mathbb{P}(G \supseteq K_t) \rightarrow 1$ if $p n^{\frac{2}{t-1}} \rightarrow \infty$. We will use $\mathbb{P}(G \not\supseteq K_t) = \mathbb{P}(X_t = 0) \leq \mathbb{P}(|X_t - \mathbb{E}[X_t]| \geq \mathbb{E}[X_t])$. By Chebyshev, we have

$$\mathbb{P}(|X_t - \mathbb{E}[X_t]| \geq \mathbb{E}[X_t]) \leq \frac{\text{Var}(X_t)}{\mathbb{E}[X_t]^2},$$

so it suffices to show that this tends to 0.

We calculate the variance as

$$\begin{aligned} \text{Var}(X_t) &= \mathbb{E}[X_t^2] - (\mathbb{E}[X_t])^2 \\ &= \sum_{A, B \in [n]^{(t)}} [\mathbb{P}(G[A], G[B] \text{ complete}) - \mathbb{P}(G[A] \text{ complete})\mathbb{P}(G[B] \text{ complete})], \end{aligned}$$

which we calculate by hand as

$$\text{Var}(X_t) = \sum_{s=1}^t \binom{n}{t} \binom{t}{s} \binom{n-t}{t-s} \left[p^{2\binom{t}{2} - \binom{s}{2}} - p^{2\binom{t}{2}} \right],$$

and noting that $\binom{n-t}{t-s} / \binom{n}{t} \leq t! / n^s$ for n large, we have

$$\begin{aligned} \frac{\text{Var}(X_t)}{(\mathbb{E}[X_t])^2} &= \sum_{s=1}^t \frac{\binom{n}{t} \binom{t}{s} \binom{n-t}{t-s} \left[p^{2\binom{t}{2} - \binom{s}{2}} - p^{2\binom{t}{2}} \right]}{\binom{n}{t}^2 p^{2\binom{t}{2}}} \\ &\leq 2^{t+1} t! \sum_{s=1}^t \frac{1}{n^s} \cdot \frac{1}{p^{\binom{s}{2}}} \\ &\leq (t!)^2 \sum_{s=1}^t \left(\frac{1}{n^{\frac{2}{s-1}}} p \right)^{\binom{s}{2}}. \end{aligned}$$

Since $n^{\frac{2}{t-1}} p \rightarrow \infty$, then $n^{\frac{2}{s-1}} p \rightarrow \infty$ for each $s \leq t$, this goes to zero as required. \square

This is an example of a threshold result, where almost suddenly when p goes to a certain value, we start to see some consistent structure in the random graphs. Another somewhat similar question is:

What is the threshold for the graph to be connected?

We will find that this ‘transition’ will be much faster than with the previous example.

Theorem 8.7

Let G be sampled from $\mathcal{G}(n, p(n))$, then $\lim_{n \rightarrow \infty} \mathbb{P}(G \text{ is connected}) = 1$ if $p > (1 + \epsilon) \frac{\log n}{n}$ and $\lim_{n \rightarrow \infty} \mathbb{P}(G \text{ is disconnected}) = 0$ if $p < (1 - \epsilon) \frac{\log n}{n}$.

Proof. For the ‘first direction’ we will show that there is an isolated vertex with probability going to 1 if $p < (1 - \epsilon) \frac{\log n}{n}$. Let I be the number of isolated vertices. We are going to use a similar strategy to before, where we calculate the variance and apply Chebyshev.

We have

$$\mathbb{E}[I] = n(1 - p)^{n-1} \quad \text{and} \quad \mathbb{E}[I^2] = \sum_{u, v \in [n]} \mathbb{P}(u, v \text{ are isolated}),$$

which gives us the variance

$$\begin{aligned} \text{Var}(I) &= \sum_{u, v \in [n]} [\mathbb{P}(u, v \text{ isolated}) - \mathbb{P}(u \text{ isolated})\mathbb{P}(v \text{ isolated})] \\ &= \left(\sum_{u \in [n]} \mathbb{P}(u \text{ iso}) - \mathbb{P}(u \text{ iso})^2 \right) + \left(\sum_{u \neq v \in [n]} \mathbb{P}(u, v \text{ iso}) - \mathbb{P}(u \text{ iso})\mathbb{P}(v \text{ iso}) \right) \\ &= (n(1 - p)^{n-1}) + \left(\sum_{u \neq v \in [n]} (1 - p)^{2(n-1)-1} - (1 - p)^{2(n-1)} \right) \\ &= n(1 - p)^{n-1} + n(n - 1)(1 - p)^{2(n-1)-1}p. \end{aligned}$$

Then we can apply Chebyshev to get

$$\begin{aligned} \mathbb{P}(I = 0) &\leq \mathbb{P}(|I - \mathbb{E}[I]| \geq \mathbb{E}[I]) \leq \frac{\text{Var}(I)}{\mathbb{E}[I]^2} \\ &= \frac{n(1 - p)^{n-1} + n(n - 1)(1 - p)^{2(n-1)-1}}{n^2(1 - p)^{2(n-1)}} \\ &= \frac{1}{(1 - p)^{n-1}n} + p \rightarrow 0, \end{aligned}$$

as required.

Now for the ‘other direction’, we want to show that the $\mathbb{P}(G \text{ is disconnected}) \rightarrow 0$. We can write

$$\begin{aligned} \mathbb{P}(G \text{ is disconnected}) &= \mathbb{P} \left(\bigcup_{A \subseteq [n], 0 < |A| \leq n/2} \{\text{no edges between } A \text{ and } A'\} \right) \\ &\leq \sum_{A \subseteq [n], 0 < |A| \leq n/2} \mathbb{P}(\text{no edges between } A \text{ and } A') \end{aligned}$$

$$= \sum_{s=1}^{n/2} \binom{n}{s} (1-p)^{s(n-s)},$$

so we want to show that this tends to zero. We can break up this sum as

$$\begin{aligned} \sum_{s=1}^{n/2} \binom{n}{s} (1-p)^{s(n-s)} &\leq \sum_{s=1}^{\epsilon n/2} \left(n e^{-p(n-s)} \right)^s + \sum_{\epsilon n/2 \leq s \leq n/2} \binom{n}{s} e^{-ps(n-s)} \\ &\leq \sum_{s=1}^{\epsilon n/2} \left(\frac{1}{n^{\epsilon/4}} \right)^s + \left(\sum_{\epsilon n/2 \leq s \leq n/2} \binom{n}{s} \right) e^{-\epsilon n \log n/8} \\ &\leq \sum_{s=1}^{\epsilon n/2} \left(\frac{1}{n^{\epsilon/4}} \right)^s + \left(2e^{-\epsilon \log n/8} \right)^n \rightarrow 0, \end{aligned}$$

thus $\mathbb{P}(G \text{ is disconnected}) \rightarrow 0$, as required. \square

§9 Algebraic Methods in Graph Theory

The final aspect of graph theory we are going to look at is the use of linear algebraic methods to solve problems. This is a naturally occurring technique, as will be shown by the problem below.

§9.1 A Motivating Problem

The **diameter** of a graph is the maximum distance between two vertices in a graph. We are going to try and attack the problem below:

How many vertices can a graph with fixed Δ have, given its diameter is at most 2?

By just thinking a little bit, it's not hard to come up with the bound

$$n \leq 1 + \Delta + (\Delta - 1)\Delta = 1 + \Delta^2,$$

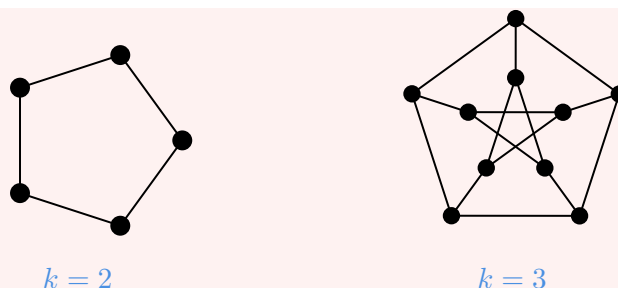
but is it possible to have a graph that attains this bound? We are going to call such graphs *Moore graphs*.

Definition 9.1 (Moore Graph)

A **Moore graph** is a k -regular graph on $k^2 + 1$ vertices that has diameter 2.

Example 9.2 (Examples of Moore Graphs)

Below are Moore graphs for $k = 2$ and $k = 3$.



Note that the graph for $k = 3$ is the Petersen graph. Playing around a bit, you might find that such a graph for $k = 4$ *does not exist*, so it seems that Moore graphs are somewhat special.

Remark. A Moore graph is a k -regular graph for which any $x \neq y$ with $x, y \in V(G)$ has exactly one path of length ≤ 2 between x and y .

So Moore graphs don't always exist, leading to a natural question – does there exist infinitely many of them?

§9.2 The Adjacency Matrix

One way to encode a graph is using an adjacency matrix.

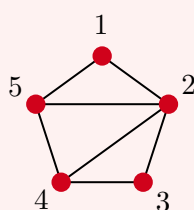
Definition 9.3 (Adjacency Matrix)

Given a graph $G = (V, E)$ on n vertices, we define the **adjacency matrix** of G to be the $n \times n$ matrix

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

Example 9.4 (Example of an Adjacency Matrix)

An graph and its associated adjacency matrix are given below.



Adjacency Matrix

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

Now while this is just a straightforward representation of a graph, the fact that it's a matrix means that we can start asking questions about it, even if it's not immediately obvious that they have any meaning. For example, what do the eigenvalues of this matrix look like? What about the eigenvectors?

We can observe that the adjacency matrix is symmetric, which we know implies that it has real eigenvalues $\lambda_1 \geq \dots \geq \lambda_n$, and an associated eigenbasis v_1, \dots, v_n which is orthonormal.

We can also observe that for a graph G with adjacency matrix A , since there's no edge

between a vertex and itself, that there is only zeros along the diagonal of A , so it has zero trace and

$$\sum_{i=1}^n \lambda_i = \text{tr}(A) = 0.$$

Also, if we call $\lambda_1 = \lambda_{\max}$ and $\lambda_n = \lambda_{\min}$, then

$$\lambda_{\max} = \max_{x:|x|^2=1} x^T A x, \quad \text{and} \quad \lambda_{\min} = \min_{x:|x|^2=1} x^T A x$$

Somewhat surprisingly, we are going to be able to connect all of these algebraic facts back to our original graph.

Proposition 9.5

Let $G = (V, E)$ be a graph with adjacency matrix A . Then

- (i) $\frac{1}{n} \sum_{x \in V} d(x) \leq \lambda_{\max}(G) \leq \Delta(G)$.
- (ii) $\lambda_{\max}(G) = \Delta(G)$ if and only if G is Δ -regular.
- (iii) $\lambda_{\min}(G) = \Delta(G)$ if and only if G is Δ -regular and bipartite.

Proof. Part (i). For the lower bound, let $w = \frac{1}{\sqrt{n}}(1, \dots, 1)$. Then we observe by interpreting the adjacency matrix in a graph theoretic sense that

$$\lambda_{\max} = \max_{x:|x|^2=1} x^T A x \geq w^T A w = \frac{1}{n} \sum_{x \in V} d(x).$$

For the upper bound, let $x = (x_1, \dots, x_n)$ be an eigenvector corresponding to λ_{\max} . Assume WLOG that $|x_1| \geq |x_i|$, for all i . Then

$$\lambda_{\max} x_1 = (\lambda_{\max} x)_1 = (Ax)_1 = \sum_{i: v_i \sim v_1} x_i,$$

which gives us

$$|\lambda_{\max} x_1| \leq \sum_{i: v_i \sim v_1} |x_i| \leq |x_1| \Delta,$$

so $|\lambda_{\max}| \leq \Delta$, and $\lambda_{\max} \leq \Delta$.

Part (ii). If G is Δ -regular, then consider $\mathbb{1} = (1, \dots, 1)$. Then

$$A\mathbb{1} = \Delta\mathbb{1},$$

thus $\lambda_{\max} \geq \Delta$, so $\lambda_{\max} = \Delta$ by our previous result.

If $\lambda_{\max} = \Delta$, then let $x = (x_1, \dots, x_n)$ be an associated eigenvector, and as before assume $|x_1| \geq |x_i|$ for all i . Then

$$\Delta x_1 = (\Delta x)_1 = (Ax)_1 = \sum_{i: v_i \sim v_1} x_i,$$

therefore $x_i = x$ for all i where $v_i \sim v_1$. So now we repeat the argument for all $v_i \in N(v_1)$ to learn that all x_j with $d(v_j, v_1) \leq 2$ has $x_j = x_1$ and so on, until we

get $x_1 = x_i$ for all v_j in the component of v_i . We then repeat for each component, to see that x is constant on each component, and thus G must be a regular graph.

Part (iii). Omitted. \square

Another fact that will bring us closer to Moore graphs concerns the interpretation of the square of an adjacency matrix.

Proposition 9.6

Let G be a graph and let A be its adjacency matrix. Then $(A^2)_{ij}$ is the number of walks of length 2 between i, j .

Proof Sketch. Consider how A^2 is calculated. \square

§9.3 Moore Graphs

Now we can apply our newfound algebraic tools to our question about Moore graphs and graph diameter. We will see that our result really does show that Moore graphs are pretty special.

Lemma 9.7 (Necessary Condition for the Existence of Moore Graphs)

Let G be a Moore graph of degree k , then

$$\frac{1}{2} \left(k^2 \pm \frac{k^2 - 2k}{\sqrt{4k - 3}} \right) \text{ are integers.}$$

Proof. Let G be a Moore graph of degree k on n vertices, and let A be the adjacency matrix of G . Then we have

$$(A^2)_{ij} = \begin{cases} k & \text{if } i = j, \\ 1 & \text{if } i \neq j \text{ and } ij \notin E, \\ 0 & \text{if } ij \in E. \end{cases}$$

Thus $A^2 = (J - A) + (k - 1)I$, where J is the matrix of all 1s. So we have

$$A^2 + A - (k - 1)I = J.$$

Now let λ_{max} be a maximum eigenvalue of A , and let x be a corresponding eigenvector. Since G is k -regular, $\mathbb{1}$ is an eigenvector of A , and thus we know that $x \perp \mathbb{1}$, since $\lambda_{max} \neq \Delta$. Then

$$(A^2 + A - (k - 1)I)x = Jx \implies (\lambda_{max}^2 + \lambda_{max} - (k - 1))x = 0,$$

so $\lambda_{max}^2 + \lambda_{max} - (k - 1) = 0$. Thus

$$\lambda_{max} = \left(\frac{-1 \pm \sqrt{1 + 4(k - 1)}}{2} \right) = \left(\frac{-1 \pm \sqrt{4k - 3}}{2} \right).$$

So A has eigenvalues k with multiplicity 1, λ_{max} with multiplicity say r , and μ with multiplicity s , with $r + s = n - 1$. We also know that

$$\sum_{i=1}^n \lambda_i = 0 = k + r\lambda + s\mu,$$

and solving this for integers r and s and using $n = k^2 + 1$ gives the desired result. \square

With this lemma, we can heavily restrict the possible Moore graphs.

Theorem 9.8 (Hoffman-Singleton Theorem)

Let G be a Moore graph of degree k . Then $k \in \{2, 3, 7, 57\}$.

Proof. We need to find out when

$$\frac{1}{2} \left(k^2 \pm \frac{k^2 - 2k}{\sqrt{4k - 3}} \right)$$

are integers. One case is when $k^2 - 2k = 0$, implying that $k = 2$. Another case is when $4k - 3 = t^2$, where $t \mid k^2 - 2k$. Then we have

$$t \mid \left(\frac{t^2 + 3}{4} \right)^2 - 2 \left(\frac{t^3 + 3}{4} \right),$$

and multiplying by 16,

$$t \mid t^4 - 2t^2 - 15,$$

so $t \mid 15$, and thus $t \in \{3, 5, 15\}$, and since $4k - 3 = t^2$, this shows that $k \in \{3, 7, 57\}$. \square

Remark. For $k = 2, 3$ and 7 , the graphs are known to exist, but a graph for $k = 57$ is has not yet been found.