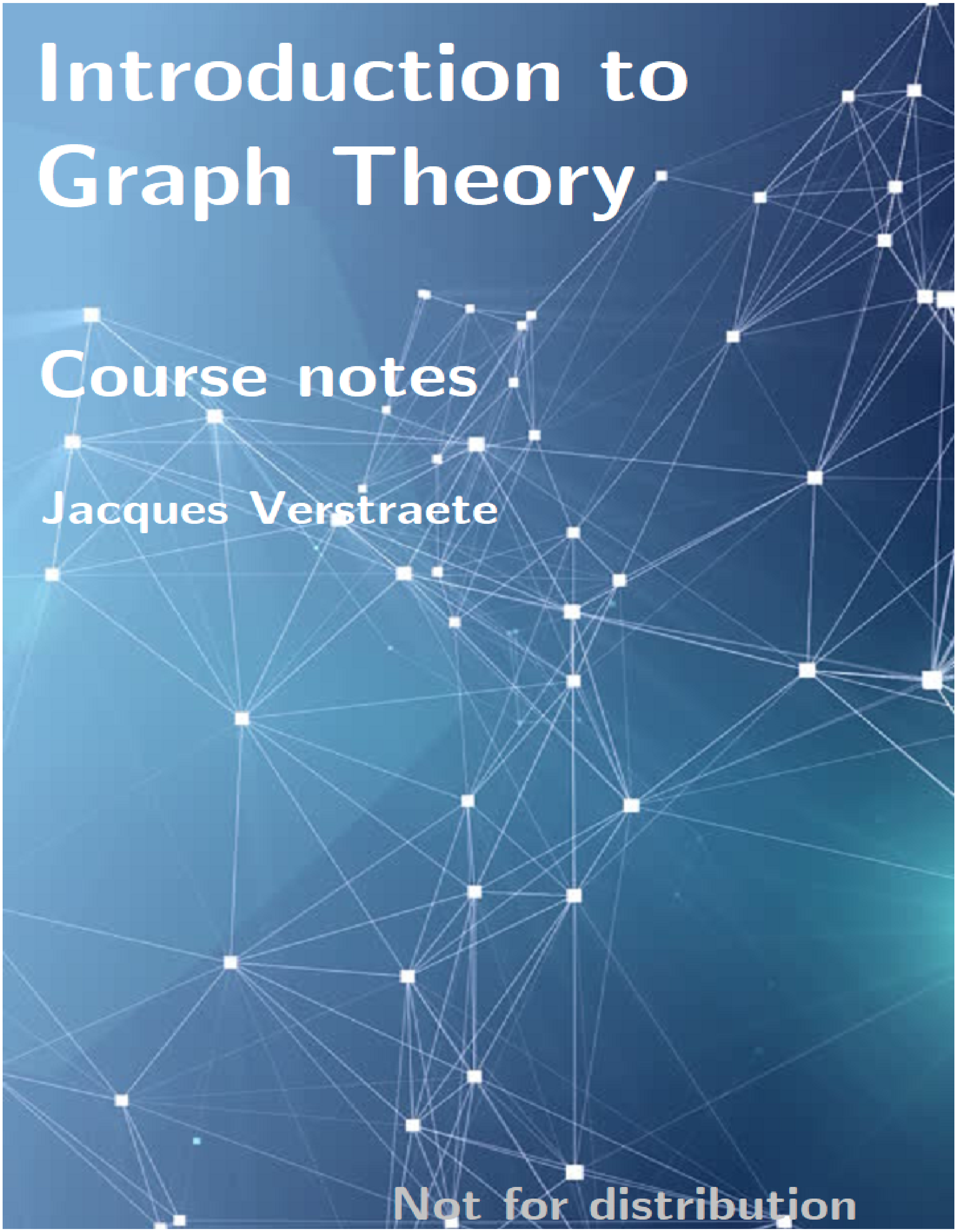

Introduction to Graph Theory

Course notes

Jacques Verstraete

1 Introduction to Graph Theory

A ***graph*** G is a pair (V, E) where V is a set and E is a set of unordered pairs[1] of elements of V. The elements of V are called ***vertices*** and V is called the ***vertex set*** of the graph, and the elements of E are called ***edges***, and E is called the ***edge set*** of the graph. If G is a graph, we let $V(G)$ denote its vertex set and $E(G)$ its edge set. If u and v are two vertices of a graph $G = (V, E)$, then we say u and v are ***adjacent*** if $\{u, v\} \in E$ – in other words $\{u, v\}$ is an edge of G. For instance, the vertex set of a graph might be $V = \{1, 2, 3\}$, and the edge set might be $\{\{1, 2\}, \{1, 3\}\}$. The graph itself would be denoted $G = (\{1, 2, 3\}, \{\{1, 2\}, \{1, 3\}\})$. The definition of graphs given above is often not the best way to represent a graph. In general, it may be convenient to represent any graph $G = (V, E)$ by drawing V as a set of points in the plane, and draw a straight line between any two adjacent vertices in V.

1.1 Examples of graphs

Example 1. Consider the graph $G = (V, E)$ where $V = \{1, 2, 3\}$ and $E = \{\{1, 2\}, \{1, 3\}\}$. Then the drawing below represents this graph:

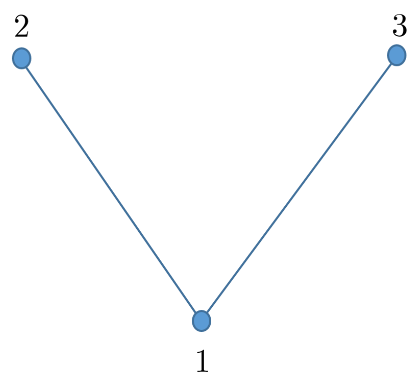

Figure 1: The graph $G = (\{1, 2, 3\}, \{\{1, 2\}, \{1, 3\}\})$

Example 2. Let $V = \{p_1, p_2, p_3, p_4, p_5, p_6\}$ be a set of six people at a party, and suppose that p_1 shook hands with p_2 and p_4, p_3 shook hands with p_4, p_5 and p_6, and p_5 and p_6 shook hands. Let $G = (V, E)$ be the graph with edge set E consisting of pairs of people who shook hands. Then

$$E = \{\{p_1, p_2\}, \{p_1, p_4\}, \{p_3, p_4\}, \{p_3, p_5\}, \{p_3, p_6\}, \{p_5, p_6\}\}.$$

A drawing of G is given in Figure 2 below:

[1]We denote sets using braces, for instance $\{1, 2, 3\}$ is the set whose elements are $1, 2$ and 3, and we write $1 \in \{1, 2, 3\}$ to say "1 is an element of the set $\{1, 2, 3\}$." Note that a set precludes "repeated elements".

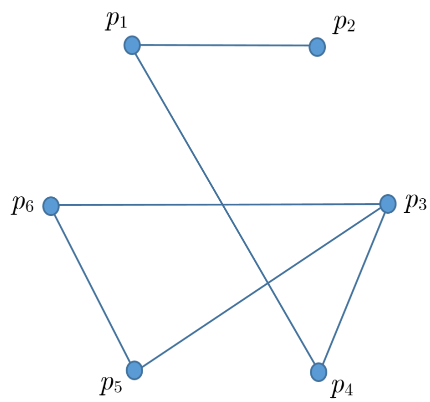

Figure 2: The handshake graph G.

Example 3. Let $\mathbb{Z}$ denote the set of integers[2] and let

$$V = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} : 0 \leq x \leq 2, 0 \leq y \leq 2\}.$$

Then V is just the set of points in the plane with integer co-ordinates between 0 and 2. Now suppose $G = (V, E)$ is the graph where E is the set of pairs of vertices of V at distance 1 from each other. In other words, (x, y) and (x', y') are adjacent if and only if $(x-x')^2+(y-y')^2 = 1$. We check that the edge set is

$$\begin{aligned} E \quad = \quad & \{\{(0,0),(0,1)\},\{(0,0),(1,0)\},\{0,1),(0,2)\},\{(1,0),(2,0)\},\{(1,0),(1,1)\}, \\ & \{(0,1),(1,1)\},\{(0,2),(1,2)\},\{(2,0),(2,1)\},\{(2,1),(2,2)\},\{(1,2),(2,2)\}\}. \end{aligned}$$

This is a cumbersome way to write the edge set of G, as compared to the drawing of G in Figure 3 below, which is much easier to absorb.

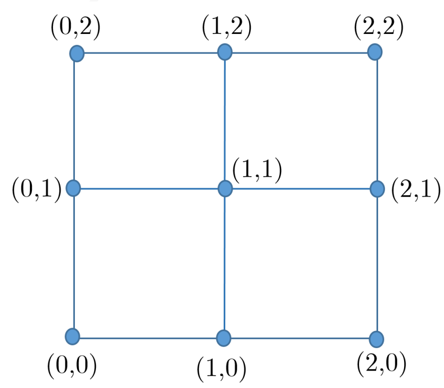

Figure 3: The grid graph G.

[2]Thus $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, 3, \ldots\}$. Then $\mathbb{Z} \times \mathbb{Z}$ is the ***Cartesian product***, which is the set of pairs (x, y) such that $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$.

Example 4. Let V be the set of binary strings of length three, so

$$V = \{000, 001, 010, 100, 011, 101, 110, 111\}.$$

Then let E be the set of pairs of strings which differ in one position. Then

$$E = \{\{000, 001\}, \{010, 000\}, \{100, 000\}, \ldots, \{111, 101\}, \{111, 110\}, \{111, 011\}\}.$$

The reader should fill in the rest of the edges as an exercise. Once again, this graph Q actually has a very nice drawing (which explains why it is sometimes called the cube graph).

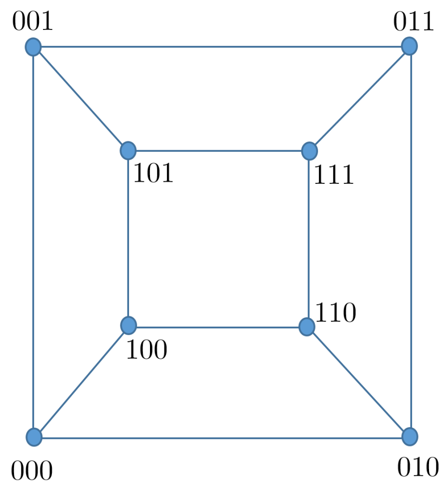

Figure 4: The cube graph Q.

1.2 Graphs in practice

Graphs appear in many theoretic and practical applications, including statistical physics, chemistry, the world-wide-web, broadcasting and networks, circuit design, computational complexity, coding and information theory, algorithm design, probability theory, algebra, number theory and geometry, and chemistry, to mention a few. We give a four examples in this section:

The web graph. Let V denote the set of websites on the internet, and let E denote the set of pairs of websites with a link between them. The web graph is growing all the time, and due to its size, difficult to analyze. In the figure below, two ***induced subgraphs*** of the web graph are shown.

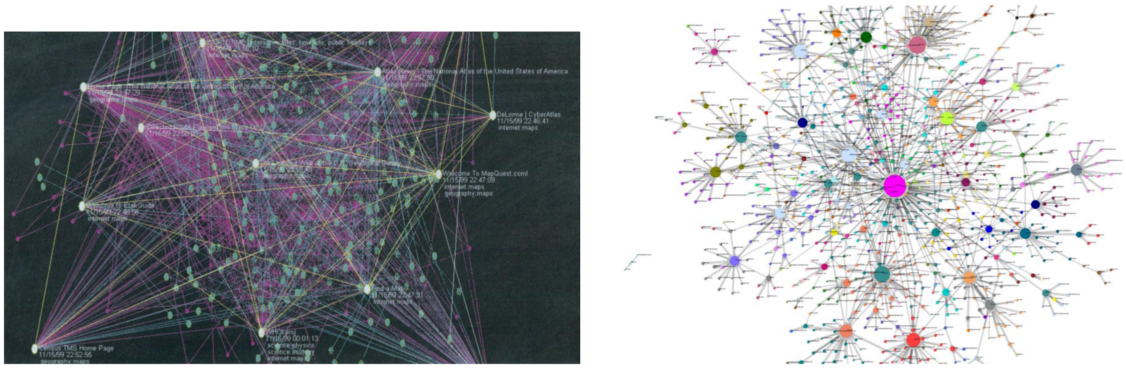

Figure 5: Induced subgraphs of the web graph.

Natural questions related to searching are whether the web graph is ***connected***, the ***radius*** and ***diameter*** of the web graph, and so on.

Planar graphs and geometry. [Notes Chapter 5] A graph is planar if it can be "drawn" in the plane or on a sphere without any edges crossing. If we consider an abstract map, then we may represent it as a planar graph by representing each country by a vertex, and drawing an edge between countries which share a border. If we consider a three-dimensional polyhedron, then it has a natural embedding on a sphere without crossing edges. Similarly, we can consider planar lattices such as the integer lattice, hexagonal lattice (honeycomb lattice) and triangular lattice in the Euclidean plane.

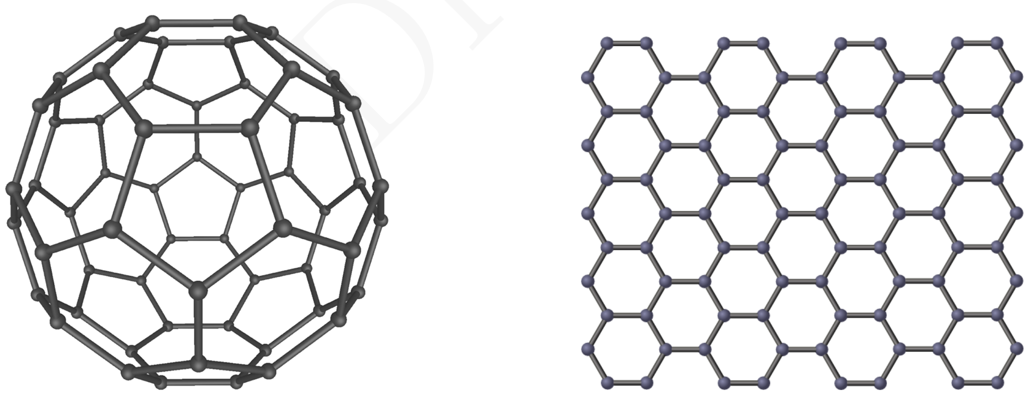

Figure 6: Carbon C_{60} fullerene and hexagonal lattice

One of the famous problems in graph theory is to color the regions of a map (in other words, color the vertices of a planar graph) so that no two adjacent regions receive the same color. A coloring of the world map with four colors is shown below:

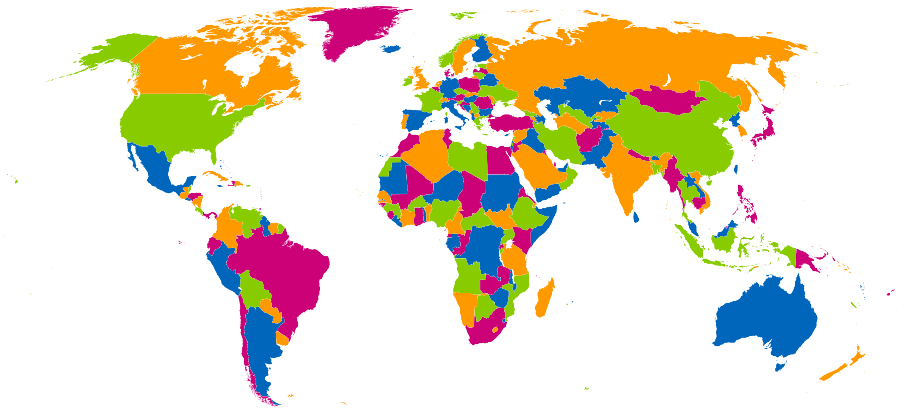

Figure 7: 4-Coloring of the world map

The famous ***4-color theorem*** says every map can be colored with at most four colors so that no adjacent countries have the same color. The integer lattice is an example of a ***unit distance graph***: a graph whose vertices are points in the plane and whose edges are pairs of points at distance 1. Other examples of unit distance graphs are shown below. The big open problem of Erdős is to determine the maximum number of edges in an n-vertex unit distance graph.

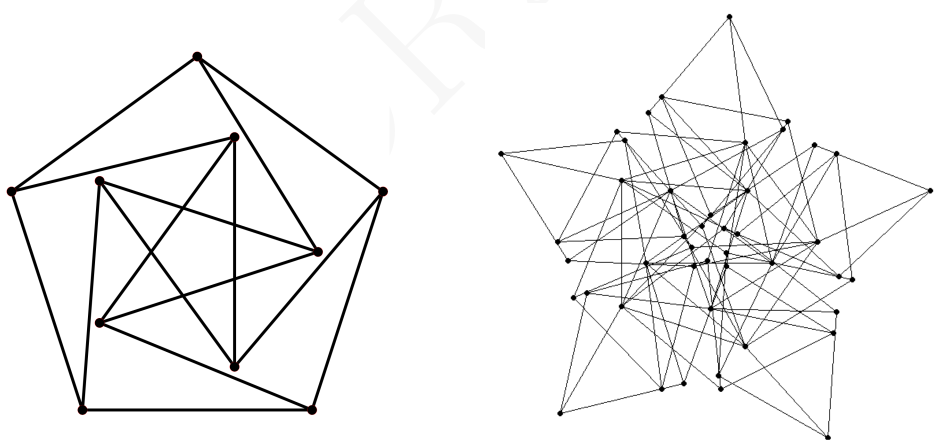

Figure 8: Unit distance graphs

Percolation and automata. Let G be the graph whose vertex set is a set of organisms, and put an edge between two data points if they can communicate a virus between them. For each vertex v in the graph, let $r(v)$ denote the minimum number of infected neighbors of v required in order for v to become infected. If X is the set of vertices initially infected, one may ask whether the infection spreads to the entire graph. In addition, perhaps after a certain time a vertex v becomes uninfected, and the same question remains. In fact, this is an instance of the famous ***Conway's game of life***. The game of life is on cells of the integer lattice, according to the following rules, with cells being in two states, infected or dormant:

- Infected cells with at most one/more than three infected neighbours becomes dormant
- Dormant cells with exactly three infected neighbours becomes infected.

In all other cases, the cells preserve their state. The question is whether the infection dies out or spreads forever, and what the set of infected cells looks like at any time. For example, if the cells initially infected form the white cells in the left frame of the picture below, it takes 130 generations for the infection to die out. Some of these generations are shown in the figure.

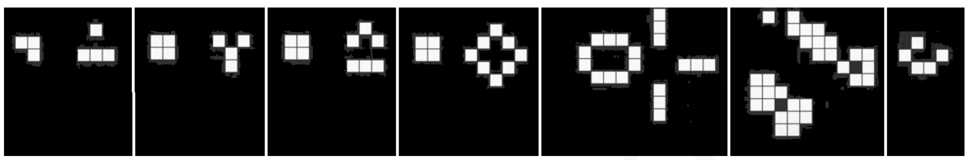

Figure 9: Conway's game of life

These kinds of questions fall into the realm of ***percolation*** and ***cellular automata***.

Connectivity and matchings. [Notes Chapter 1 and 2] Given a graph G, how many vertices or edges must be removed to disconnected the graph (split it into connected pieces)? This is the fundamental connectivity problem in graphs, addressed by Menger's Theorems. Given a graph G, how many vertex-disjoint edges (a ***maximum matching***) can we find in the graph? This is addressed by the ***König-Ore Theorem*** and ***Hall's Theorem*** for bipartite graphs, and ***Tutte's 1-Factor Theorem*** and ***Tutte-Berge Formula*** in general graphs. Furthermore, the maximum matching can be found efficiently (see the last section of Chapter 2). Let $G = (V, E)$, and let $s, t \in V$ be vertices designated as ***source*** and ***sink***. Suppose each edge of the graph has a ***capacity***, denoting the maximum number of units of fluid that the edge can carry between its ends. If fluid flows through the network from the source to the sink, we assume that the flow in to each vertex other than s or t is equal to the flow out of the vertex. Given the capacities, the question is the maximum flow can be transmitted from s to t (flow occurs simultaneously in all edges). This is completely answered by the famous max-flow min-cut theorem, together with an efficient algorithm for finding a maximum flow. Many generalizations of this theorem exist, with wide applicability. The theorem turns out to imply Hall's Theorem on matchings and Menger's Theorem on connectivity. The matching problem, for instance, is very natural in practical applications, such as scheduling and job assignment. Given a set $A = \{a_1, a_2, \ldots, a_k\}$ of people and a set $B = \{b_1, b_2, \ldots, b_l\}$ of jobs, and for each person a list of jobs in B that they can do, we would like to assign as many people to jobs without having one job done by two people or two jobs done by one person. The natural graph has vertex set $A \cup B$, where a_i is joined to b_j by an edge if a_i can do job b_j. Then we are asking for a maximum matching, and an efficient algorithm exists, even if we put a weight on each edge $\{a_i, b_j\}$ to denote how much a_i would like to do job b_j. An example with three people and four jobs is below:

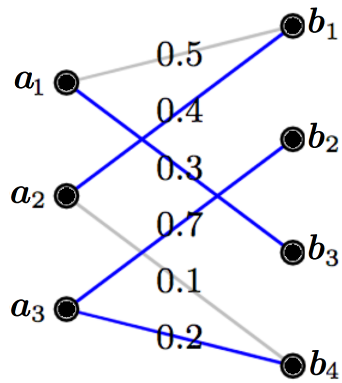

Figure 10: Maximum matching

Random graphs. The classical ***Erdős-Rényi model*** of random graphs takes n vertices and then for each pair of vertices, we place an edge with probability p and no edge with probability $1-p$ (in other words, the edge set are decided by $\binom{n}{2}$ coin flips). In this way we generate a graph $G_{n,p}$. When $p = 0$, it is the empty graph, and when $p = 1$, it is the complete graph. In the figure below, we show examples of $G_{64,p}$ when $p \in \{0, \frac{1}{256}, \frac{1}{64}, \frac{1}{16}, \frac{1}{4}, 1\}$.

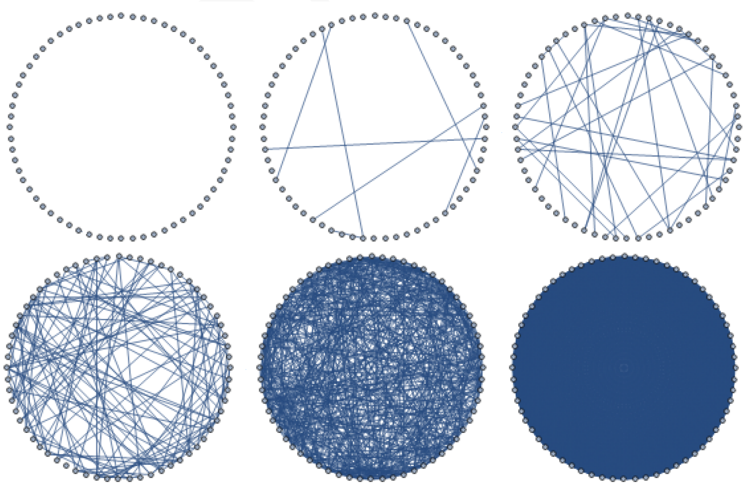

Figure 11: Random graphs.

There are very many other types of random graphs, for example the ***preferential attachment*** graph used to model the web graph, or ***random regular graphs***, or ***random geometric graphs***. The graph below is a random geometric graph in the unit square: the vertices are uniformly randomly chosen points in the unit square, and the edges correspond to pairs of points at most a certain distance from each other.

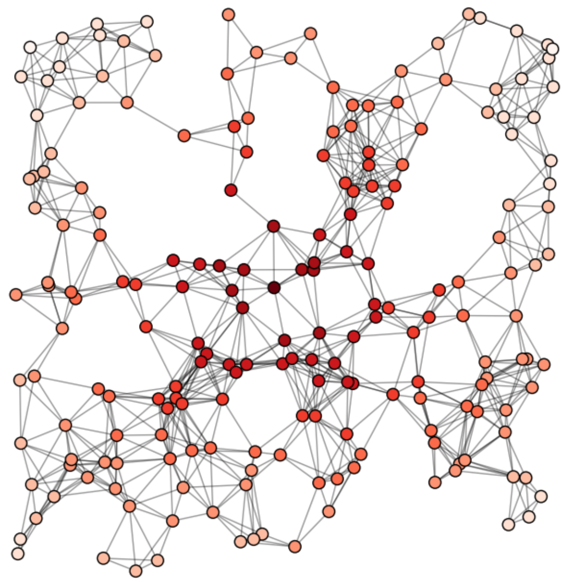

Figure 12: Random geometric graph.

1.3 Basic classes of graphs

There are some graphs which we shall encounter very frequently, and we describe these here.

Complete Graphs. The ***complete graph*** on n vertices, denoted K_n is the graph consisting of all possible edges on n vertices (in other words, every pair of vertices is adjacent). The ***empty graph*** on n vertices has no edges. In Figure 13, drawings of K_n for $2 \leq n \leq 6$ are given:

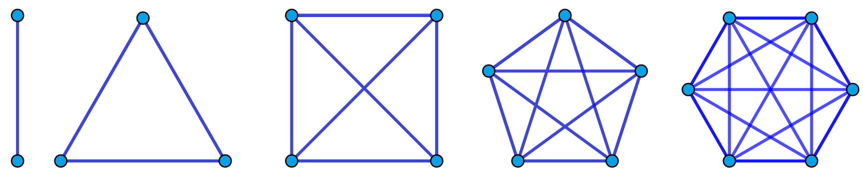

Figure 13: The complete graphs

Since the number of pairs of vertices in K_n is $\binom{n}{2}$, and every pair is an edge, the number of edges in K_n is $\binom{n}{2}$.

Bipartite graphs. Recall a ***partition*** of a set V consists of pairwise disjoint non-empty subsets whose union is V. A ***bipartite graph*** is a graph $G = (V, E)$ such that for some partition of V into two sets A and B – we call these the ***parts*** of G – every edge of G has the form $\{a, b\}$ with $a \in A$ and $b \in B$ (or in other words, no two vertices in A are adjacent, and no two vertices in B are adjacent). When $|A| = r$ and $|B| = s$ and all possible edges $\{a, b\}$ with $a \in A$ and $b \in B$ are included, then G is called the ***complete bipartite graph***, and denoted $K_{r,s}$. In Figure 14, we draw the graphs $K_{2,3}$ and $K_{2,5}$.

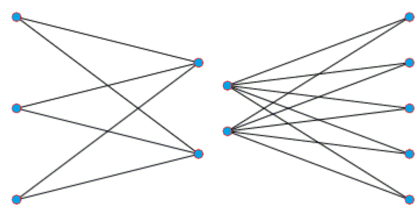

Figure 14: Complete bipartite graphs $K_{2,3}$ and $K_{2,5}$

Note that the number of edges in a complete bipartite graph K_{rs} is exactly rs.

Paths, walks, and Cycles. For $k \geq 3$, a ***k-cycle*** is the graph C_k with vertex set $\{1, 2, \ldots, k\}$ and edge set

$$\{\{1, 2\}, \{2, 3\}, \{3, 4\}, \ldots, \{k-1, k\}, \{k, 1\}\}.$$

For $k \geq 1$, a ***k-path*** is the graph P_k with vertex set $\{1, 2, \ldots, k+1\}$ and edge set

$$\{\{1, 2\}, \{2, 3\}, \{3, 4\}, \ldots, \{k-1, k\}, \{k, k+1\}\}.$$

Note that a k-cycle has k edges and a k-path has k edges, and we often refer to the number k as the ***length*** of the cycle or path. In Example 1, P_2 is drawn, and in Figure 15, we draw C_3 and C_6.

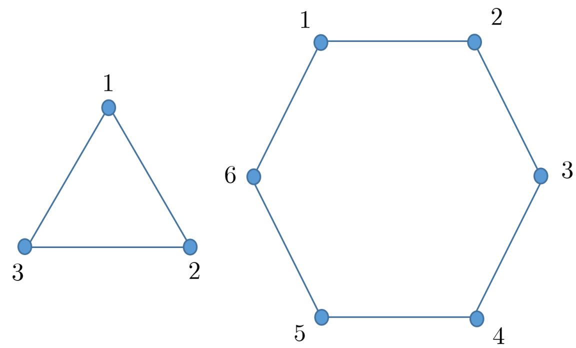

Figure 15: Cycles C_3 and C_6

A ***walk*** in a graph $G = (V, E)$ is an alternating sequence of vertices and edges, whose first and last elements are vertices, and such that each edge joins the vertices immediately preceding it and succeeding it in the sequence. For example,

$$a\{a,d\}d\{d,e\}e\{e,a\}a\{a,d\}d$$

is a walk in the graph in Figure 16. Since there is no ambiguity, we denote a walk by a sequence of vertices, so the above walk is (a, d, e, a, d). Note that if the vertices of a walk are all distinct, then the walk is a path. The ***length*** of a walk is the number of steps taken in the walk. A ***closed walk*** is a walk whose first and last vertices are the same. If a closed walk has no repeated vertices except the first and the last, then we observe it is a cycle. If the first and last vertices of a walk are u and v, then we say the walk is a uv-***walk***. We refer similarly to a uv-***path***. The vertices u and v are called the ***ends*** of the path or walk.

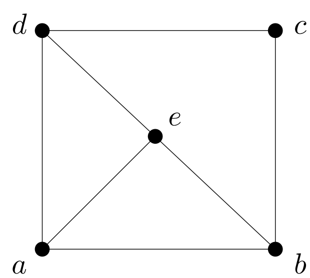

Figure 16: Walks

Lemma 1.3.1 *Let u, v be distinct vertices in a graph G, and let W be a shortest uv-walk in G. Then W is a path.*

Proof ▷ Suppose $W = v_0e_0v_1e_1 \dots v_{k-1}e_{k-1}v_k$, where $v_0, v_1, \dots, v_k$ are vertices of G with $v_0 = u$ and $v_k = v$, and $e_0, e_1, \dots, e_k$ are edges of G. If W is not a path, then $v_i = v_j$ for some

$i < j$ with $(v_i, v_j) \neq (u, v)$. Define the new walk

$$W' = v_0 e_0 v_1 e_1 \dots v_i e_j v_{j+1} \dots e_{k-1} v_k.$$

Then the length of W' is less than the length of W, a contradiction. So W is a path. $\square$

1.4 Degrees and Neighbourhoods

The ***neighborhood*** of a vertex v in a graph $G = (V, E)$, denoted $N_G(v)$, is the set of vertices of G which are adjacent to v. The ***degree*** of a vertex v in a graph G, denoted $d_G(v)$, is $|N_G(v)|$. When it is clear which graph G we are referring to, we write $d(v)$ and $N(v)$ instead of $d_G(v)$ and $N_G(v)$. The ***degree sequence*** of a graph G is the sequence of degrees of vertices of G in non-increasing order. For example, the degree sequence of the graph in Figure 16 is $(3, 3, 3, 3, 2)$, whereas the degree sequence of the graph in Figure 3 is $(4, 3, 3, 3, 3, 2, 2, 2, 2)$. A vertex of degree zero is called an ***isolated vertex***.

We write $\delta(G) = \min\{d_G(v) : v \in V\}$ and $\triangle(G) = \max\{d_G(v) : v \in V\}$ for the ***minimum degree*** and ***maximum degree*** of G, respectively. For the examples in the last section, we note $\delta(G) = 1$ and $\triangle(G) = 2$ for Figure 1, $\delta(G) = 2$ and $\triangle(G) = 4$ for 3, $\delta(G) = 1$ and $\triangle(G) = 3$ for 2, and $\delta(Q) = \triangle(Q) = 3$ for the cube graph in Figure 4. If all vertices in a graph have the same degree r, then the graph is said to be ***r-regular***. For instance, the graph Q is 3-regular (all the degrees are 3). Sometimes, 3-regular graphs are also referred to as ***cubic*** graphs.

1.5 The handshaking lemma

An important fact involving the degrees of a graph G, which we will use on numerous occasions, is the ***handshaking lemma***:

Lemma 1.5.1 (Handshaking Lemma) *For any graph $G = (V, E)$,*

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

Proof $\triangleright$ When we add up the degrees of vertices of G, every edge of G is counted twice, so the sum of the degrees is twice the number of edges. $\square$

The handshaking lemma gives an easy way to count the number of edges in a graph: it is just half the sum of the degrees of the vertices. Note if G is r-regular and has n-vertices, then the number of edges in G is $nr/2$, by the handshaking lemma (check this for the cube graph Q in the last section). A consequence of the handshaking lemma is that the number of vertices of odd degree in any graph must be even – otherwise the sum on the left above would be odd whereas the right hand side is even:

Lemma 1.5.2 *For any graph $G = (V, E)$, the number of vertices of odd degree is even.*

The reader may check that this is satisfied for the graphs in Examples 1 – 4. Consider the complete graph K_n. Every vertex of K_n is adjacent to every other vertex of K_n, so the degree of every vertex of K_n is $n-1$ – in other words, K_n is $(n-1)$-regular. By the handshaking lemma, the number of edges in K_n is $\frac{1}{2} \cdot n \cdot (n-1) = \binom{n}{2}$, as we already knew. Next, consider Figure 3 in the last section (the grid graph). The degree sequence of this graph is $(4, 3, 3, 3, 3, 2, 2, 2, 2)$. Therefore by the handshaking lemma, the number of edges in the grid graph is ◁◁

$$\frac{1}{2}(4 + 3 + 3 + 3 + 3 + 2 + 2 + 2 + 2) = 12.$$

A manual count of the edges in Figure 3 confirms this. The reader should check how many edges the n by n grid graph has (the vertex set is $V = \{(x, y) \in \mathbb{Z} \times Z : 0 \leq x < n, 0 \leq y < n\}$ and the edge set is the set of pairs of vertices at distance 1 from each other.) ◁◁

Example 5. The ***n*-cube**, denoted Q_n, is the graph whose vertex set is the set of binary strings of length n, and whose edge set consists of all pairs of strings differing in one position. The cube graph Q_3 in Example 4 is the 3-cube. Let us see how many edges Q_n has as a formula in n. Since there are 2^n binary strings of length n, there are 2^n vertices in Q_n. Now each vertex v is adjacent to n other vertices – namely flip one position in the string v to get each string adjacent to v, and there are n possible positions in which to do a flip. So every vertex of the n-cube has degree n (in other words, it is n-regular), and so the number of edges in Q_n is

$$\frac{1}{2} \sum_{v \in V} d_{Q_n}(v) = \frac{1}{2} \cdot 2^n \cdot n = n2^{n-1}.$$

A manual count of the edges confirms this for Q_4, which is drawn below:

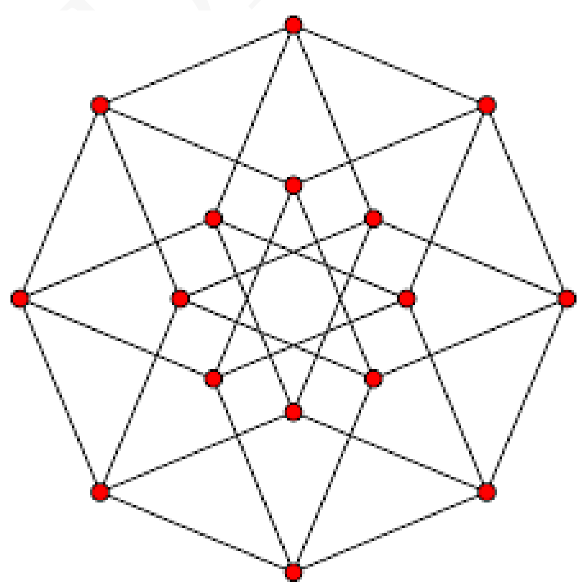

Figure 17: The 4-cube Q_4.

1.6 Subgraphs

If H and G are graphs and $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$, then H is called a ***subgraph*** of G. To denote that H is a subgraph of G, we write $H \subseteq G$. If in addition $V(H) = V(G)$ then H is called a ***spanning subgraph*** of G.

Example 6. For instance, the reader will check that the graph $G = P_2$ shown in Figure 1 is a subgraph of the graphs in Figures 2 – 4. The graph in Figure 2 is not a subgraph of any of the others, since it contains a triangle but none of the others contains a triangle. The graph G in Figure 3 is not a subgraph of the cube graph Q in Figure 4 since it has a vertex of degree four, whereas Q is 3-regular. We note that every graph with at most n vertices is a subgraph of K_n, and every graph with n vertices is a spanning subgraph of K_n. The path P_{k-1} is a spanning subgraph of C_k.

We now define how to remove edges and vertices from a graph G. If X is a set of vertices of G, we denote by $G - X$ the graph with vertex set $V(G)\backslash X$ and edge set $E = \{e \in E(G) : e \cap X = \emptyset\}$. If $L \subseteq E(G)$, we denote by $G - L$ the graph with vertex set $V(G)$ and edge set $E(G)\backslash L$.

Example 7. For instance, if we remove one edge e from a cycle C_k, we get the path P_{k-1}, which we write as $C_k - e = P_{k-1}$. If we remove one vertex v from a cycle C_k, we get the path P_{k-2}, which we write as $C_k - v = P_{k-2}$. If we remove the vertex 1 from the graph in Figure 1, we get a graph consisting of two ***isolated vertices***. If we remove $X = \{101, 100, 111, 110\}$ from the graph Q in Figure 4, we get C_4, so we may write $Q - X = C_4$. If instead we remove $X = \{001, 101, 110\}$ we get the graph shown below in Figure 18:

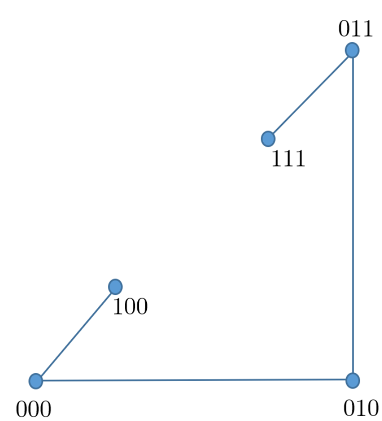

Figure 18: The graph $Q - \{001, 101, 110\}$.

The subgraph of G ***induced*** by a set $X \subseteq V(G)$, denoted $G[X]$, is precisely $G - (V\backslash X)$. A subgraph H of G is an ***induced subgraph*** if for some $X \subseteq V(G)$, $H = G[X]$. If L is a set of edges of G, then the subgraph of G ***spanned by*** L is the graph with edge set L and vertex set $\bigcup_{e \in L} e$. The graph in Figure 18 is an induced subgraph of Q, whereas P_{k-1} is not an induced subgraph of C_k.

2 Connected graphs

A graph is ***connected*** if any pair of vertices in the graph are joined by at least one path. If a graph is not connected, we say it is disconnected. The ***components*** of a graph $G = (V, E)$ are the maximal connected subgraphs of G – that is, the connected subgraphs such that no edge of G not already in the subgraph can be added while still preserving connectivity. For instance, the graph below in Figure 19 has three components:

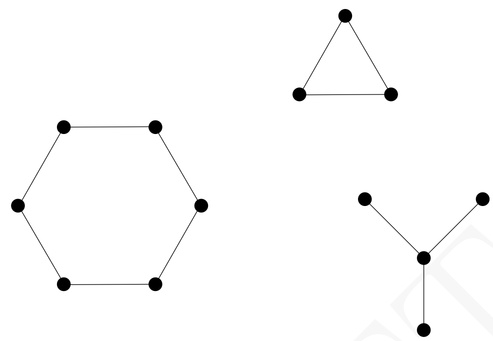

Figure 19: Components.

2.1 Bridges and trees

A ***tree*** is a connected graph without cycles – a connected ***acyclic*** graph. To describe the structure of trees, we define the notion of a bridge. A ***bridge*** of a graph G is an edge $e \in E(G)$ such that $G - e$ has more components than G. For example, in Figure 2, the edges $\{p_1, p_2\}, \{p_1, p_4\}$ and $\{p_4, p_3\}$ are bridges, whereas the remaining edges are not bridges. It is easy to spot the bridges of a graph, using the following lemma:

Lemma 2.1.1 *An edge $e \in E(G)$ is a bridge of G if and only if e is not contained in any cycle in G.*

Proof ▷ If e is contained in a cycle C of G, then $C - e$ is a path joining the ends of e. But that means $G - e$ is connected, so e could not have been a bridge. □

Since a tree has no cycles, every edge of a tree must be a bridge. We can now characterize which graphs are trees in a few ways.

Proposition 2.1.2 *Each of the following is equivalent to a graph G being a tree:*

1. *The graph G is connected and acyclic.*
2. *The graph G is connected and every edge of G is a bridge.*
3. *The graph G is connected and has $|V(G)| - 1$ edges.*

Proof ▷ Clearly Proposition 2.1.2.1 is the definition of G being a tree. Since a connected graph is acyclic if and only if every edge of the graph is a bridge, by the last lemma, Proposition 2.1.2.1

and Proposition 2.1.2.2 are equivalent. We proved Proposition 2.1.2.1 implies Proposition 2.1.2.3 by strong induction on the number of vertices in the tree, so it remains to show that Proposition 2.1.2.3 implies Proposition 2.1.2.1. To see that, if G is connected with $|V(G)| - 1$ edges, we remove an edge of any cycle and that does not disconnect G, by Lemma 2.1.1. We continue removing edges of G in cycles until all the cycles are gone. But then the remaining graph T is connected and acyclic, so must be a tree. Since it has $|V(G)|$ vertices, we know it must have $|V(G)| - 1$ edges. But G itself has $|V(G)| - 1$ edges, so $G = T$. □

The last part of the proof of this proposition is important. It says that in any connected graph G, while there is a cycle, pick an edge of the cycle and remove it. By Lemma 2.1.1, we did not disconnect the graph, so if we repeat this procedure we eventually obtain a spanning subgraph of G which is acyclic and connected – a tree. We call this a ***spanning*** tree of the graph. A spanning tree of the cube graph Q is given below in bold edges (the reader should find other spanning trees of Q): ◁◁

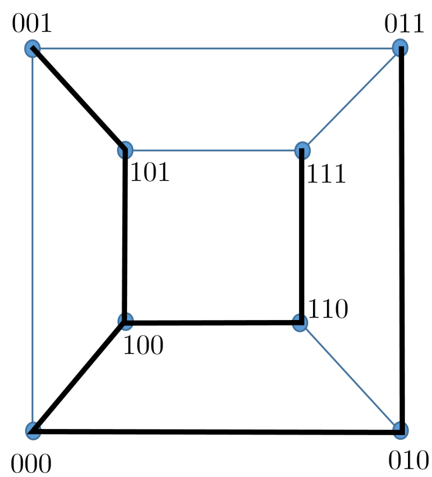

Figure 20: Spanning tree

In general, a graph has many spanning trees.

Proposition 2.1.3 *Any connected graph contains a spanning tree.*

The proof gives a fairly quick way to find a spanning tree of a graph: search for a cycle and remove an edge of the cycle, and repeat until there are no cycles left. We next discuss an algorithmic way for finding a spanning tree.

2.2 Breadth-first search

One of the simplest things to check is whether a connected graph is bipartite. Namely, pick any vertex of the graph and place it in A. Then, all the neighbors of that vertex are forced to be in B. Then all their neighbors must be in A, and so on. We repeat this procedure until all the vertices of the graph have been placed in A and in B. Applying this procedure to the graphs in Examples 1 – 4, the reader may check that only the graph in Example 2 is not bipartite. There is a systematic way to check whether a graph is bipartite, at the same time as producing a spanning tree in G. To describe this algorithm, we need the notion of distance in graphs.

The ***distance*** between vertices u and v in a connected graph G, denoted $\mathrm{dist}_G(u, v)$, is the length of a shortest uv-path. For instance, we check in Figure 16 that

$$\mathrm{dist}_G(a, c) = \mathrm{dist}_G(e, c) = \mathrm{dist}_G(b, d) = 2$$

and any two other vertices are adjacent, so they are at distance 1. The maximum distance between any two vertices in a connected graph is called the ***diameter*** of G. The minimum r such that every vertex of G is at distance at most r from some vertex of G is called the ***radius*** of G. For instance, the graph in Figure 3 has radius 2 but diameter 4, since every vertex is at distance at most 2 from $(1, 1)$, whilst the shortest path from $(0, 0)$ and $(2, 2)$ has length four. Similarly, the graph in Figure 2 has radius 2 and diameter 4. For complete graphs, the radius and diameter are both 1.

If v is a vertex in a connected graph G, we let $N_i(v)$ denote the set of vertices at distance exactly i from v, so that $N_1(v)$ is exactly the neighborhood of v and $N_0(v) = \{v\}$. Order the vertices of G. We build a tree T by first adding v to T, and then adding all vertices of $N_1(v)$ in increasing order, and then the vertices of $N_2(v)$ in increasing order, and so on, until all vertices have been added. The edges of T are described as follows: an edge $\{u, w\} \in E(G)$ with $u \in N_i(v)$ and $w \in N_{i+1}(v)$ is present in T if there exists no vertex t less than u in $N_i(v)$ adjacent to w. An example is given in Figure 21, where the vertices are ordered $1, 2, 3, 4, 5, 6, 7, 8, 9$ and $v = 1$. The arrowed edges denote T, whereas the blue lines denote edges of the graph G. Then T can be encoded by the sequence $(1, 3, 9, 2, 6, 8, 4, 7, 5)$, where $N_1(v) = \{3, 9\}$, $N_2(v) = \{4, 6, 8\}$ and $N_3(v) = \{5, 7\}$.

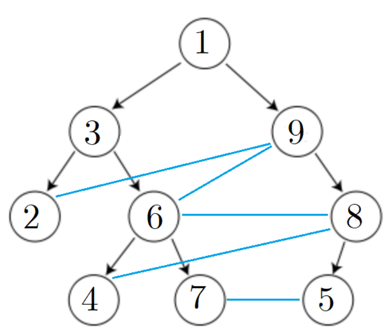

Figure 21: Breadth-first search.

Lemma 2.2.1 *Let G be a connected graph and $v \in V(G)$. Then T is a spanning tree of G such that $dist_T(v, w) = dist_G(v, w)$ for all $w \in V(G)$.*

The last statement in this lemma says that T preserves distances from v to all other vertices. The tree T is called a ***breadth-first search tree rooted at*** v. The sets $N_i(v)$ are sometimes called the ***layers*** of T, and the ***height*** of T is the maximum distance of any vertex from v. In Figure 21, we have a tree with four layers and height three.

Example 8. The famous ***Petersen graph*** is drawn below, with vertices labelled 1 through 10. Let us apply the breadth first search algorithm to find a spanning tree in G rooted at vertex 1. Of course, we start by adding 1 to the tree.

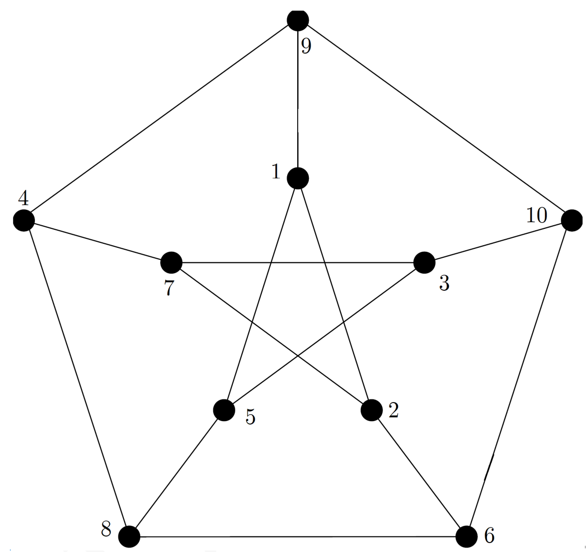

Figure 22: Petersen graph

We then add the neighbors of 1 in increasing order, namely, 2, 5 and then 9. So far our tree has edges $\{1, 2\}, \{1, 5\}$ and $\{1, 9\}$. Now we move on to the first vertex added in $N_1(v)$, namely 2. We add first the vertex 6 and then the vertex 7, with edges $\{2, 6\}$ and $\{2, 7\}$. Then move to the next added vertex in $N_1(v)$, namely 5. We add 3 and 8 and the edges $\{5, 3\}$ and $\{5, 8\}$. Finally, we move to the vertex 9, and add the vertices 4 and 10 and the edges $\{9, 4\}$ and $\{9, 10\}$. Then we stop since there are no vertices left to add. The tree has edge set $\{\{1, 2\}, \{1, 5\}, \{1, 9\}, \{2, 6\}, \{2, 7\}, \{5, 3\}, \{5, 8\}, \{9, 4\}, \{9, 10\}$, and the order in which vertices were added is $(1, 2, 5, 9, 6, 7, 3, 8, 4, 10)$. The layers of the tree are $N_0(v) = \{1\}$, $N_1(v) = \{2, 5, 9\}$, and $N_2(v) = \{6, 7, 3, 8, 4, 10\}$. The reader can check that the Petersen graph has diameter and radius equal to two. ◁◁

We now use breadth first search to prove a lemma characterizing bipartite graphs. Via this lemma, the Petersen graph in Figure 22 is not bipartite, since it contains a cycle of length five, for instance with vertex set $\{1, 2, 7, 3, 5\}$ (in fact there are many cycles of length 5, 7 and 9). ◁◁

Lemma 2.2.2 *A graph G is bipartite if and only if it does not contain any odd cycles.*

Proof ▷ Since an odd cycle is not bipartite, bipartite graphs cannot contain odd cycles. Conversely, if a graph has no odd cycles, let T be a breadth-first search tree in G, rooted at some vertex v. We claim that $A = N_0(v) \cup N_2(v) \cup \dots$ and $B = N_1(v) \cup N_3(v) \cup \dots$ do not contain any edges of G, and therefore they are the parts in a bipartition of G. Suppose there exists an edge $\{x, y\}$ in A. Since edges of T connect consecutive layers, $\{x, y\}$ is not in T. Let P be a path in T connecting $\{x, y\}$. Then P together with $\{x, y\}$ forms a cycle C. On the other hand, P must have even length, since if $x \in N_{2i}(v)$ and $y \in N_{2j}(v)$, for if $N_h(v)$ is the lowest layer that P intersects, then P has length $(2i-h)+(2j-h) = 2i+2j-2h$. But then C has odd length, which is a contradiction (this is evident for instance in Figure 21, with $x = 6$ and $y = 8$, $i = j = 2$ and $h = 0$, and P is the path with edge set $\{\{6,3\},\{3,1\},\{1,9\},\{9,8\}\}$). Similarly, B does not contain any edges of G, so A and B are the parts of G. □

2.3 Eulerian graphs

A graph is ***eulerian*** if all its vertices have even degree. A ***trail*** in a graph is a walk with no repeated edges, and a ***tour*** in a graph is a closed walk with no repeated edges. An ***eulerian tour*** in a graph G is a tour which contains every edge of G and an ***eulerian trail*** is a trail that contains all the edges of G. In the graph shown below, an example of a tour is the walk $(v_1, v_2, v_4, v_1, v_5, v_6, v_1)$. This graph has an eulerian tour, namely $(v_1, v_2, v_3, v_4, v_5, v_6, v_1, v_5, v_2, v_4, v_1)$. Roughly speaking, the presence of an eulerian tour in a graph means that the graph can be drawn on paper without lifting your pen and without retracing edges.

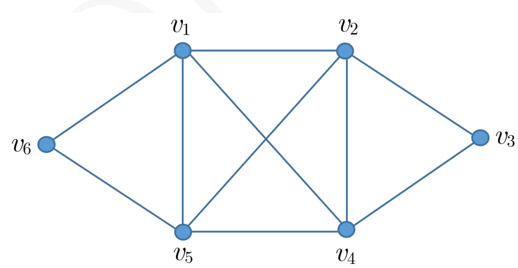

Figure 23: Petersen graph

The problem of existence of eulerian tours was first studied by Euler, in his famous "bridges of Königsberg problem". The following theorem is responsible for the existence of an eulerian tour in the above graph.

Theorem 2.3.1 *A connected graph G has an eulerian tour if and only if all of the vertices of G have even degree.*

Proof ▷ If G has an eulerian tour, say $(v_1, v_2, \dots, v_m, v_1)$ (in this sequence, note that some vertices can be repeated), let i denote the first index such that $v_i = v_1$. Then the edges

$\{v_1, v_2\}, \{v_2, v_3\}, \dots, \{v_{i-1}, v_i\}, \{v_i, v_1\}$ form a cycle C in G. If $i = m$, then $G = C$, and all vertices of G have degree two. Otherwise, $G - E(C)$ has the eulerian tour $(v_1, v_{i+1}, v_{i+2}, \dots, v_m, v_1)$, and therefore all degrees of $G - E(C)$ have degree two. Adding back the edges of C increases degrees by zero or two, so all degrees in G are even, as required.

Now suppose all vertices of G have even degree. Let $\tau = (v_1, v_2, \dots, v_k)$ be the longest possible trail in G. If $v_k \neq v_1$, then as in the first part of the proof given above, the reader will check that an odd number of edges of τ contain each of v_1 and v_k, so there is an edge $\{v_k, v_{k+1}\}$ of G ◁◁ that is not traversed by τ. Now $(v_1, v_2, \dots, v_k, v_{k+1})$ is a longer trail than τ, a contradiction. We conclude $v_k = v_1$ and τ is a tour in G. Since G is connected, there is an edge e not in the trail τ, say $\{v_i, v\} \in E(G)$. If v is not a vertex of the trail, then

$$(v_i, v_{i+1}, \dots, v_k, v_1, v_2, \dots, v_{i-1}, v_i)$$

is a tour of the same length as τ in G. If we add the edge $\{v_i, v\}$, we get the trail

$$(v_i, v_{i+1}, \dots, v_k, v_1, v_2, \dots, v_{i-1}, v_i, v)$$

which is longer than τ. If v is a vertex on the trail, say $v = v_j$ where $j < i$, then consider the trail $(v_i, v_{i+1}, \dots, v_k, v_1, \dots, v_{j-1}, v_j, v_i, v_{i-1}, \dots, v_{j+1}, v_j)$ is a trail using the edge e and is one longer than τ. This contradiction completes the proof. $\square$

2.4 Block Decomposition

We just gave three equivalent characterizations of trees in Proposition 2.1.2. In general, we would like to describe how to build connected graphs. The main result in this section will be the block decomposition theorem. We require some definitions. A ***cut vertex*** of a graph G is a vertex G such that $G - \{x\}$ is disconnected. A ***block*** of a graph is a maximal connected subgraph with no cut vertex – a subgraph with as many edges as possible and no cut vertex. So a block is either K_2 or is a graph which contains a cycle. For example in a tree, every block is K_2. The block decomposition of a graph is just the set of all the blocks of the graph. An example of a block decomposition is shown below.

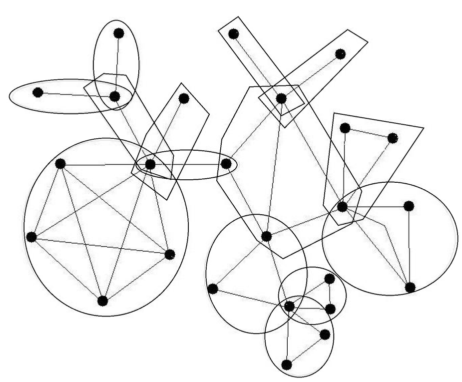

Figure 24: Blocks.

In the picture, there are fourteen blocks. Seven of the blocks are K_2, four of the blocks are triangles, one of the blocks is K_5, and there are two other blocks. The block decomposition theorem says that block decompositions of graphs have a "tree-like structure". To make this precise, given a graph G, we form a new graph $\mathcal{B}$ where the vertices of $\mathcal{B}$ consist of all cut vertices of G and also all blocks of G, and where a block is joined to all cut vertices of G contained in it. An example of this graph is shown below for the figure above:

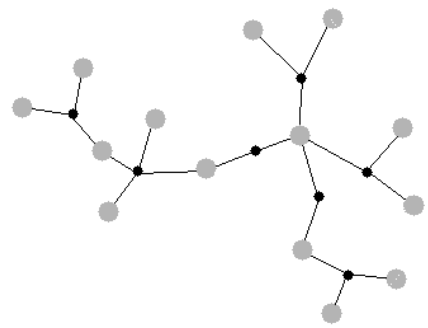

Figure 25: The graph $\mathcal{B}$.

In the figure, the black vertices represent cut vertices of G, and the grey vertices represent blocks of G. Here is the block decomposition theorem:

Theorem 2.4.1 *Let G be a connected graph. Then $\mathcal{B}$ is a tree.*

Proof $\triangleright$ By adding edges inside the blocks of G, we do not change $\mathcal{B}$, so we can assume every block of G is a complete graph. Since G is connected, clearly $\mathcal{B}$ is connected too. Now we show $\mathcal{B}$ has no cycles. The vertices of a cycle $\mathcal{C} \subseteq \mathcal{B}$ are alternately blocks of G and cut vertices of G, by definition of $\mathcal{B}$. This is shown in the figure below:

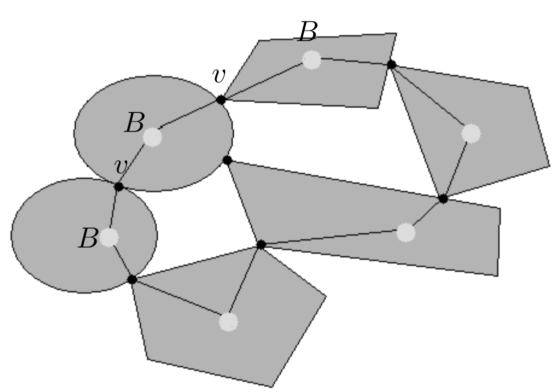

Figure 26: Cycle in $\mathcal{B}$.

In the figure, the blocks are shown as grey dots and labelled B and the cut vertices are black dots labelled v. Let the cut vertices of G in order along $\mathcal{C}$ be $v_0, v_1, \ldots, v_k, v_0$. Then $v_0 v_1 v_2 \ldots v_k v_0$ is a cycle $C \subseteq G$. If $B \in \mathcal{C}$, then $B \cup C$ is a subgraph of G consisting of the complete graph B together with the cycle C containing an edge of B and at least one edge not in B. Therefore $B \cup C$ has no cut vertex, and must be a block of G. However, this contradicts the definition that B is block. $\square$

Using this theorem, we give a first example of a ***structure theorem*** in graph theory. We say that a uv-path P in a graph G is ***internally disjoint*** from a subgraph H of G if $V(P) \cap V(H) = \{u, v\}$. Define a ***theta graph*** to be any graph consisting of the union of three pairwise internally disjoint paths between two vertices.

Proposition 2.4.2 *Let G be a connected graph containing no theta graph. Then every block of G is a cycle or K_2 and G is a tree of cycles and K_2s, as shown in Figure 27 below.*

Proof $\triangleright$ Let B be a block of G. If $B \neq K_2$, then B contains a cycle, C. If $B \neq C$, then there is a path P in B such that $P \cup C$ is a theta graph: namely, pick a shortest path in $B - E(C)$ between two vertices of C. Therefore $B = K_2$ or B is a cycle. We know by the last result that G is then a tree of cycles and K_2s. $\square$

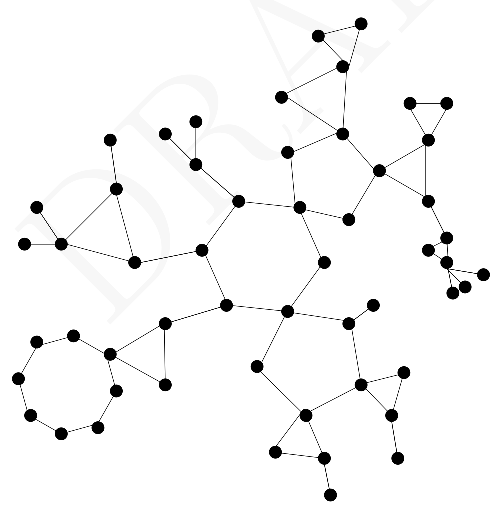

Figure 27: Tree of cycles and K_2

2.5 Decomposing blocks into paths and cycles

In this section we will give method for decomposing blocks, called ***ear-decomposition***. Let $G \neq K_2$ be a block and $P \subset G$ a path all of whose internal vertices have degree two in G and whose ends have degree at least three in G. Then P is called an ear of G (see Figure 28). Note that an ear can be a single edge.

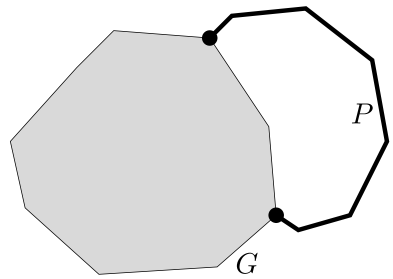

Figure 28: Ear decomposition.

The main theorem in this section says that blocks can be built from a cycle by adding ears. More precisely, a graph G has an ear decomposition if there is sequence of subgraphs of G, say $G_0 \subset G_1 \subset \cdots \subset G_t$ such that G_0 is a cycle, $G_t = G$, and G_i is obtained from G_{i+1} by removing the internal vertices of some ear in G_{i+1} or, if the ear is a single edge, deleting this edge.

Theorem 2.5.1 (Ear decomposition)
A graph $G \neq K_2$ is a block if and only if it has an ear-decomposition.

We prove Theorem 2.5.1 using the notion of ***equivalence relations***.

Definition 2.5.2 *An equivalence relation on a set S is a set R of ordered pairs of elements of S with the following properties:*

1. $(a, a) \in R$
2. *if* $(a, b) \in R$ *then* $(b, a) \in R$.
3. *if* $(a, b), (b, c) \in R$ *then* $(a, c) \in R$.

The properties 1, 2 and 3 of an equivalence relation are called ***reflexivity***, ***symmetry*** and ***transitivity***, respectively. If $(a, b) \in R$, we say that a and b are equivalent under R.

Example 9. For instance, if $G = (V, E)$ is a graph and

$$R = \{(u, v) \in V \times V : u \text{ and } v \text{ are joined by a path}\},$$

then R is an equivalence relation, and any two vertices in a component of G are equivalent under R. To prove this, the main thing to check is transitivity: that if u and v are joined by a path and v and w are joined by a path then also u and w are joined by a path. It is convenient, rather than writing $(u, v) \in R$, to write $u \sim v$.

For the proof of Theorem 2.5.1, we define an equivalence relation $\sim$ on the edge set of a graph $G = (V, E)$ as follows: for $e, f \in E$, $e \sim f$ if and only if $e = f$ or some cycle in G contains both e and f. The following lemma says that $\sim$ is indeed an equivalence relation:

Lemma 2.5.3 *For any graph G, the relation $\sim$ is an equivalence relation on $E(G)$.*

Proof $\triangleright$. By definition we know $e \sim e$ for any edge $e \in E(G)$, and $e \sim f$ is clearly the same as $f \sim e$. It remains to verify transitivity: we have to prove that if some cycle $C \subset G$ contains e and f, and some cycle $D \subset G$ contains f and g, then some cycle in G contains both e and g. Consider the path $P = D - f$. then there is a path $Q \subseteq P$ containing g whose first and last vertices u, v are in $V(C)$ but with no other vertices in C. Clearly $C \cup Q$ is a theta graph containing e and g, consisting of internally disjoint uv-paths Q, R and S such that $R \cup S = C$. Then either $Q \cup S$ or $Q \cup R$ is the required cycle containing both e and g. $\square$

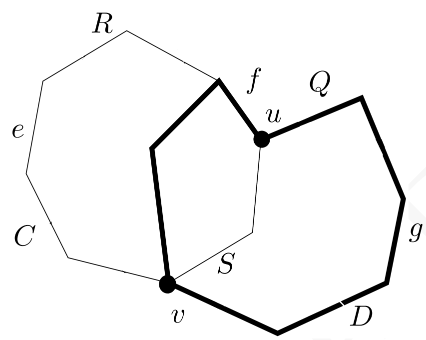

Figure 29: Transitivity of $\sim$

Theorem 2.5.4
For a graph G with at least three vertices and no isolated vertices, the following three statements are equivalent:

1. *G is a block*
2. *every two edges of $E(G)$ are in a common cycle*
3. *any two vertices of $V(G)$ are in a common cycle.*

Proof $\triangleright$. We show first that 1 implies 2. Let $e_0 = \{x_0, x_1\}$ and $e_k = \{x_k, x_{k+1}\}$ be edges of G. We have to show $e_0 \sim e_k$. Since G is connected, there is a path $P \subset G$ of the form $x_0 e_0 x_1 e_1 x_2 e_2 \dots x_k e_k x_{k+1}$. Since $G - \{x_i\}$ is connected, there is a path in $G - \{x_i\}$ from x_{i-1} to x_{i+1}, which means that e_{i-1} and e_i are contained in a common cycle in G, for all i. In other words, $e_{i-1} \sim e_i$ for all i. But by transitivity, this means $e_0 \sim e_k$, as required. So 1 implies 2. To prove that 2 implies 3, let $u, v \in V(G)$ and select an edge e containing u and an edge $f \neq e$ containing v (this edge exists because G has at least three vertices). Then $e \sim f$ by assumption, so some cycle in G contains both u and v, as required. So 2 implies 3. Finally, to show 3 implies 1, $G - \{x\}$ is connected for any $x \in V(G)$, otherwise we get the contradiction that two vertices in different components of $G - \{x\}$ are not on a cycle in G. $\square$

Proof $\triangleright$ OF THEOREM 2.5.1 Suppose G is a block, and let H be a maximal subgraph of G with an ear decomposition. Since G contains a cycle, H certainly exists. Suppose, for a contradiction, that $H \neq G$. Then there exists an edge $e \in E(G)\backslash E(H)$. If e joins two vertices of H, then $H + e$ has an ear decomposition, contradicting the maximality of H. Therefore e has an end not in H. Let f be any edge of H. Then e and f are contained in a common cycle, C, by Theorem 2.5.4 part 2. In particular, C contains at least two vertices of H, so there is a path $P \subset C$, internally disjoint from H, and with both ends in H. But then P is an ear of $H \cup P$, contradicting the maximality of H. We conclude that $H = G$. The proof of the converse statement is left as an exercise. □ ◁◁

The theorem on ear decomposition is very useful for proving statements about blocks by induction.

2.6 Decomposing bridgeless graphs

Here we prove an ear-decomposition theorem for graphs with no bridges. It cannot be the same as for blocks, since the graph consisting of the union of two cycles sharing exactly one vertex is not a block and does not have an ear decomposition in the sense of the last section. The new ear decomposition is described as follows: an ear decomposition of a graph G is a sequence of subgraphs of G, say $G_0 \subset G_1 \subset \cdots \subset G_t$ such that G_0 is a cycle, $G_t = G$, and $G_{i+1} = G_i \cup P$ for a path P internally disjoint from G_i with both ends in $V(G_i)$, or $G_{i+1} = G_i \cup C$ for a cycle C with exactly one vertex in common with G_i.

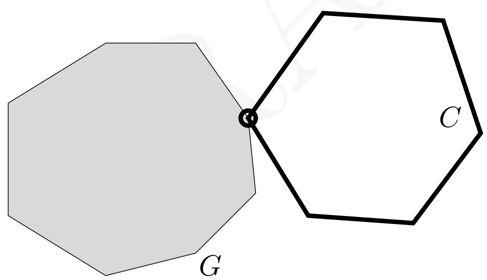

Figure 30: New ear decomposition.

The proof of the ear-decomposition theorem is similar to that of Theorem 2.5.1:

Theorem 2.6.1
A graph is bridgeless if and only if it has an ear decomposition.

Proof ▷ Suppose G has an ear decomposition. Let e be an edge of G. It is sufficient to prove that e is contained in a cycle – then e cannot be a bridge of G by Lemma 2.1.1. If e is not in a cycle, then e is a bridge of G. Let P be an ear of G. Since e is a bridge of G, P can never contain e, since there is a cycle in G containing all edges of P. Therefore e survives our procedure, but then e must be in a cycle, a contradiction. So G must be 2-edge-connected. Now suppose that G is 2-edge-connected. Then G contains a cycle, so G has a subgraph with an ear decomposition. So we can take a maximal subgraph H of G so that H has an ear decomposition. We'll show $H = G$. If $H \neq G$, then there is an edge e of G joining a vertex of H to a vertex of G not in H, otherwise $H + e$ has an ear decomposition. This edge is in a cycle C, by Lemma 2.1.1. If C contains only one vertex of H, then $H \cup C$ has an ear-decomposition, a contradiction. So C contains two vertices in H, and we find a shortest path between two vertices of H in G to add to H, a contradiction. So $H = G$. □

2.7 Menger's Theorem

An ***edge cut*** of a graph G is a set of edges whose removal from G gives a disconnected graph. A graph G is ***k-edge-connected*** if every edge cut has size at least k. A ***vertex cut*** of a graph G is a set of vertices whose removal from G gives a disconnected graph. A graph G is ***k-connected*** if every vertex cut has size at least k.

In the last two sections, we gave a structural characterization of 2-connected graphs (blocks) and 2-edge-connected graphs (bridgeless graphs). There is no good structural characterization of k-connected graphs in general. From the results of the last section, it is possible to show that any two vertices in a 2-connected graph are the ends of two internally disjoint paths, and any two vertices in a 2-edge-connected graph are ends of two edge-disjoint paths. In this section, we generalize this to k-edge-connnected and k-connected graphs via ***Menger's Theorems***.

Let u and v be vertices in a graph G, and let P and Q be uv-paths. Then P and Q are internally disjoint if the only vertices they have in common are u and v. A ***uv-separator*** is a set $W \subset V(G)\backslash\{u, v\}$ such that u and v are in distinct components of $G - W$. For non-adjacent vertices $u, v \in V(G)$, let $\kappa(u, v)$ denote the minimum size of a uv-separator. We prove the vertex form of Menger's Theorem:

Theorem 2.7.1 (Menger's Theorem - Vertex Form) *The minimum size $\kappa(u, v)$ of a uv-separator in a graph G is equal to the maximum number of pairwise internally disjoint uv-paths in G. In particular, a graph is k-connected if and only if each pair of its vertices is connected by k pairwise internally disjoint paths.*

Proof ▷ For $k = 1$ the theorem is just the definition of a connected graph. Now suppose $k \geq 2$. If there are k internally disjoint uv- paths in G, then clearly k vertices are required to separate u from v, as at least one vertex is required to destroy each of the internally disjoint uv-paths. Now suppose k vertices are required to separate u from v. Let G be a counterexample to the theorem with the smallest possible value of k, and with the smallest number of edges. Now $N(u) \cap N(v) = \emptyset$ otherwise, for any $x \in N(u) \cap N(v)$, $G - \{x\}$ is a counterexample to the theorem with $k - 1$ vertices separating u from v but at most $k - 2$ internally disjoint uv-paths. Now let W be a set of k vertices separating u from v. We consider two cases.

Case 1. $W \not\subseteq N(u)$ and $W \not\subseteq N(v)$. Let H_u and H_v be the components of $G - W$ containing u and v respectively. Note that H_u and H_v have each at least two vertices, since $N(u) \cap N(v) = \emptyset$, so $E(H_u) \neq \emptyset$ and $E(H_v) \neq \emptyset$. Let G_u be obtained from $G - V(H_u)$ by adding a vertex w adjacent to all neighbors of H_u in W. By the minimality of the number of edges in G as a counterexample, there are k internally disjoint wv-paths $P_1, P_2, \ldots, P_k$. Similarly, the graph G_v obtained from $G - V(H_v)$ by adding a vertex x adjacent to all neighbors of H_v in W has k internally disjoint xu-paths $Q_1, Q_2, \ldots, Q_k$. Suppose $W = \{w_1, w_2, \ldots, w_k\}$ and P_i starts with the edge $\{w, w_i\}$ and Q_i starts with the edge $\{x, w_i\}$ (see Figure 31. Then $Q_i - \{x\}$ together with $P_i - \{w\}$ is a uv-path $R_i \subseteq G$, and the paths $R_1, R_2, \ldots, R_k$ are internally disjoint, as required (in Figure 31, R_i is shown, with $Q_i - \{x\}$ in green and $P_i - \{x\}$ in blue).

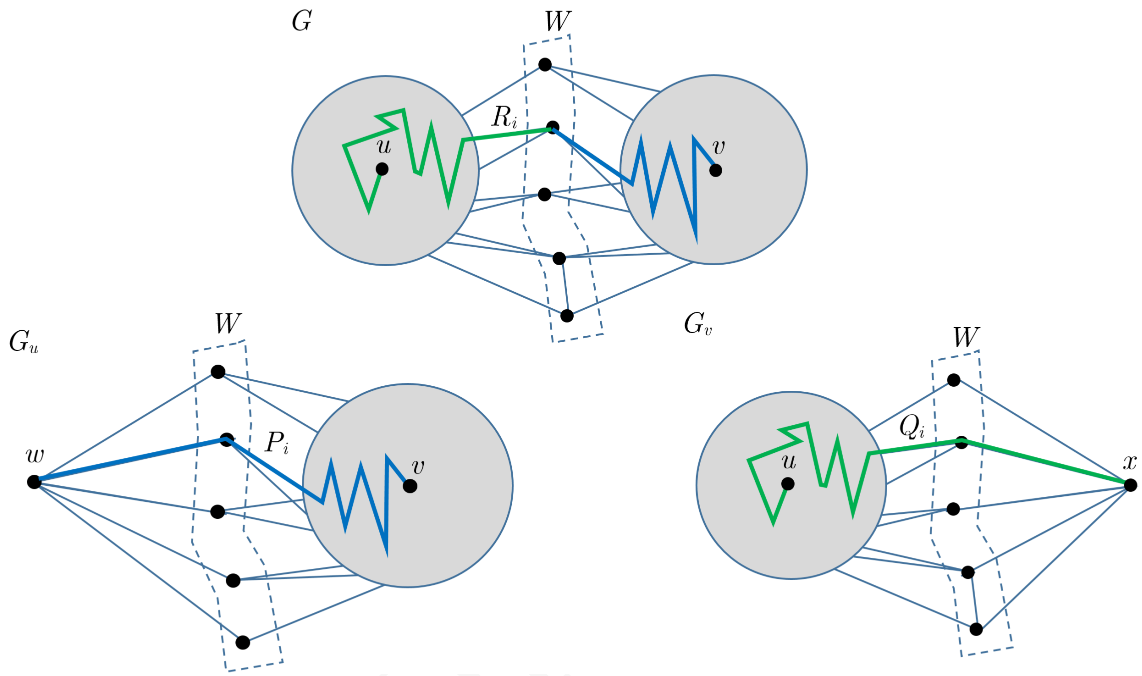

Figure 31: Case 1 in Menger's Theorem

Case 2. $W \subseteq N(u)$ **or** $W \subseteq N(v)$. We reduce this case to Case 1. Let P be a shortest uv-path with edges $\{u, u_1\}, \{u_1, u_2\}, \ldots, \{u_{i-1}, u_i\}, \{u_i, v\}$. Since $N(u) \cap N(v) = \emptyset$, P has length at least three. In particular, $u_1 \notin N(v)$ and $u_2 \notin N(u)$. Let $e = \{u_1, u_2\}$. Then every uv-separator in $G - e$ has size at least $k - 1$. If every uv-separator has size at least k then, by minimality of G, we have k internally disjoint uv-paths in $G - e$, and therefore in G, which is a contradiction. So $G - e$ has a uv-separator W_0 of size $k - 1$. Then $W_1 = W_0 \cup \{u_1\}$ and $W_2 = W_0 \cup \{u_2\}$ are uv-separators of size k in G. Since $N(u) \cap N(v) = \emptyset$, $W_0 \not\subseteq N(u)$ or $W_0 \not\subseteq N(v)$. If $W_0 \not\subseteq N(u)$, then $W_1 \not\subseteq N(u)$ and $W_1 \not\subseteq N(v)$ since $u_1 \notin N(v)$. If $W_0 \not\subseteq N(v)$, then $W_2 \not\subseteq N(v)$ and $W_2 \not\subseteq N(u)$ since $u_2 \notin N(u)$. So Case 1 applies to W_1 or W_2. $\square$

A set $L \subset E(G)$ is a ***uv-edge-separator*** if u and v are in different components of $G - L$. Let $\lambda(u, v)$ denote the minimum size of a uv-edge-separator. The edge form of Menger's Theorem for k-edge-connected graphs is as follows:

Theorem 2.7.2 (Menger's Theorem - Edge Form) *The minimum size $\lambda(u, v)$ of a uv-edge-separator in a graph G equals the maximum size of a set of pairwise edge-disjoint uv-paths in G. In particular, a graph is k-edge-connected if and only if each pair of its vertices is connected by k pairwise edge-disjoint paths.*

2.8 Fan Lemma

If A and B are sets of vertices in a graph G, then an AB-path is a path with one end in A and the other in B, and no other vertices in $A \cup B$. We leave the following lemma as an exercise: $\lhd\lhd$

Lemma 2.8.1 *Let G be a k-connected graph and let A be a set of at least k vertices in G. Then the graph obtained from G by adding a new vertex adjacent to all vertices in A is k-connected.*

Corollary 2.8.2 (Fan Lemma) *Let G be k-connected with at least $k+1$ vertices. Then*

1. *for any $X \subset V(G)$ of size k and $u \in V(G) \backslash X$, there are k paths from u to X with only the vertex u in common.*
2. *for any sets $A, B \subset V(G)$ of size k, there exist k vertex-disjoint AB-paths.*

Proof $\rhd$. To prove (2), let G^* be the graph obtained from G by adding a vertex x adjacent to all vertices in X. Since $|X| \geq k$, Lemma 2.8.1 shows G^* is k-connected. By Menger's Theorem, there exist k internally disjoint paths between u and x in G^*. Removing x from all of these paths, we have k paths from u to X with only u in common.

To prove (3), let G^{**} be obtained from G by adding a vertex a adjacent to all vertices in A and a vertex b adjacent to all vertices in B. Since $|A| \geq k$ and $|B| \geq k$, G^{**} is k-connected, via two applications of Lemma 2.8.1. By Menger's Theorem, there are k internally disjoint ab-paths in G^{**}. Removing a and b from these paths gives k vertex-disjoint AB paths in G. $\square$

2.9 Dirac's Theorem

Dirac's Theorem says that through any k vertices in a k-connected graph we can find a cycle, when $k \geq 2$.

Theorem 2.9.1 (Dirac's Theorem) *Let G be a k-connected graph, where $k \geq 2$, and let X be a set of k vertices of G. Then there exists a cycle in G containing X.*

Proof $\rhd$ By induction on k. If $k = 2$, then every pair of vertices of G is joined by two internally disjoint paths by Menger's Theorem, so every pair of vertices is contained in a cycle (this is also Theorem 2.2.4 (3)).

Now let G be a k-connected graph, where $k > 2$, and let $X = \{x_1, x_2, \dots, x_k\}$ be a set of k vertices of G. Since G is also $k-1$ connected, there is a cycle C containing $\{x_1, x_2, \dots, x_{k-1}\}$. We can assume that the order in which these vertices appear on C is $x_1, x_2, \dots, x_{k-1}$. We consider first the case that C has length $k-1$. Since $|V(G)| \geq k+1$, there is a vertex $x \in V(G) \backslash X$. By the Fan Lemma (2), there are k paths from x_k to $\{x_1, x_2, \dots, x_{k-1}, x\}$ with only the vertex x_k in common. Now if P_i is the path from x_k to x_i, then $C \cup P_i \cup P_{i+1}$ is a cycle containing X, as required. Next we consider the case $|V(C)| \geq k$. If $x_k \in V(C)$, we are done, so we assume $x_k \notin V(C)$. Then for any $x \in V(C) \backslash \{x_1, x_2, \dots, x_{k-1}\}$, there are k paths from x_k to $\{x_1, x_2, \dots, x_{k-1}, x\}$ with only the vertex x_k in common, by the Fan Lemma (2). Let these paths be $P_1, P_2, \dots, P_{k-1}, P_k$. Let y_i denote the first vertex of P_i on C and let

$Q_i \subset P_i$ denote the path from x_k to y_i. For some i, j, there is a path $P \subset C$ joining y_i to y_j containing none of the vertices $\{x_1, x_2, \ldots, x_{k-1}\}$ (see Figure 32). Now delete the vertices of P between y_i and y_j from C to get a path $Q \subset C$. Then $Q \cup Q_i \cup Q_j$ is a cycle containing X (define $Q_{k+1} = Q_1$). This proves the result. $\square$

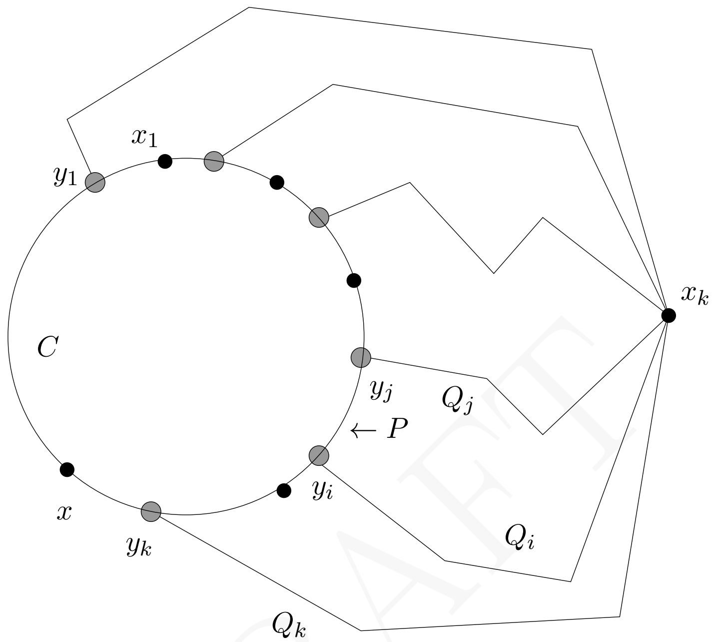

Figure 32: Dirac's Theorem.

2.10 Vertex and edge connectivity

Let G be a graph. We define $\lambda(G)$, the ***edge-connectivity*** of G, to be the minimum size of an edge cut in G: it is the minimum size of $L \subset E(G)$ such that $G - L$ is disconnected. Thus a graph is ℓ-edge-connected if and only if $\lambda(G) \geq \ell$, and

$$\lambda(G) = \min\{\lambda(u, v) : u, v \in V(G)\}.$$

If G is not a complete graph, then we define $\kappa(G)$, the ***vertex-connectivity*** of G, to be the minimum size of a vertex cut in G: it is the minimum size of a set $S \subset V(G)$ such that $G - S$ is disconnected. Thus a graph is k-edge-connected if and only if $\kappa(G) \geq k$, and

$$\kappa(G) = \min\{\kappa(u, v) : u, v \in V(G), \{u, v\} \notin E(G)\}.$$

If $G = K_n$, we define $\kappa(G) = n - 1$. It should be intuitively clear that $\kappa(G) \leq \lambda(G)$ for every graph G, since we do more "damage" by removing vertices than by removing edges. The quickest proof is via Menger's Theorem:

Corollary 2.10.1 *For any graph G, $\kappa(G) \leq \lambda(G) \leq \delta(G)$.*

Proof ▷ Since the edges containing a vertex of minimum degree form an edge cut, $\lambda(G) \le \delta(G)$. Now we prove $\kappa(G) \le \lambda(G)$. For $u, v \in V(G)$, let $k(u,v)$ is the maximum number of pairwise internally disjoint uv-paths, and $\ell(u,v)$ is the maximum number of pairwise edge-disjoint uv-paths. Then by the edge form of Menger's Theorem:

$$\lambda(G) = \min\{\lambda(u,v) : u,v \in V(G)\} = \min\{\ell(u,v) : u,v \in V(G)\}.$$

Now $k(u,v) \le \ell(u,v)$ for all $u,v \in V(G)$, since internally disjoint paths are also edge-disjoint paths, and therefore

$$\begin{aligned}
\kappa(G) &= \min\{\kappa(u,v) : u,v \in V(G), \{u,v\} \notin E(G)\} \\
&= \min\{k(u,v) : u,v \in V(G)\} \\
&\le \min\{\ell(u,v) : u,v \in V(G)\} \; = \; \lambda(G).
\end{aligned}$$

The reader should check why the first two lines are equal here. □ ◁◁

This corollary can be proved directly, without Menger's Theorems. It is also the case that for any three positive integers $d \ge \ell \ge k$, there exists a graph G with $\kappa(G) = k$, $\lambda(G) = \ell$ and $\delta(G) = d$. ◁◁

3 Matchings and Factors

A ***matching*** in a graph is a set of pairwise vertex-disjoint edges of the graph. In this section we are interested in determining the size of a ***maximum matching*** in a given graph and when a graph has a ***perfect matching*** or ***1-factor*** – that is, a matching covering all its vertices. For bipartite graphs, this question will be completely answered by Hall's Theorem and the König-Ore formula. For general graphs, Tutte's 1-Factor theorem and the Tutte-Berge formula apply.

3.1 Independent Sets and Covers

An ***independent set*** in a graph G is a set X of vertices no pair of which is an edge of G – in other words the subgraph $G[X]$ induced by X has no edges. The maximum size of an independent set in a graph G is denoted $\alpha(G)$. The maximum size of a matching in a graph G is denoted $\alpha'(G)$. A ***vertex cover*** of G is a set of vertices $X \subset V(G)$ such that $e \cap X \neq \emptyset$ for every edge $e \in E(G)$ – in other words, a set of vertices which intersects every edge of G. The minimum size of a vertex cover of G is denoted $\beta(G)$. An ***edge cover*** of G is a set of edges covering every vertex of G – that is a set $E \subset E(G)$ such that for every vertex $v \in V(G)$, there is an edge of E containing v. The minimum size of an edge-cover is denoted $\beta'(G)$.

Example 10. The Petersen graph P is shown in the figure below. This graph has a perfect matching, for instance the edges $\{1,9\},\{3,10\},\{2,6\},\{5,8\},\{7,4\}$ form a perfect matching. Therefore $\alpha'(P) = 5$. An example of a maximum independent set is $\{2,4,5,10\}$, and therefore $\alpha(P) = 4$. A minimum vertex cover is $\{1,3,6,7,8,9\}$ and so $\beta(P) = 6$. Finally, a perfect matching is by definition a minimum edge cover, so $\beta'(G) = 5$.[3]

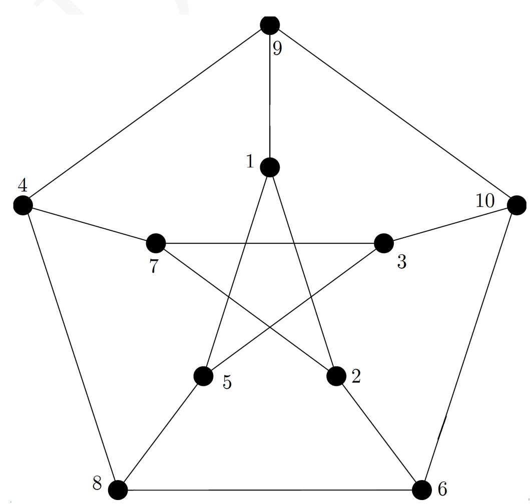

Figure 33: Covers, matchings and independent sets.

[3]Find examples of other perfect matchings, maximum independent sets, and minimum covers in the Petersen graph.

Lemma 3.1.1 *For any graph G, $\alpha(G) + \beta(G) = |V(G)|$.*

Proof $\triangleright$. If I is an independent set of vertices in G, then $V(G)\backslash I$ is a vertex cover: every edge of G has at least one end in $V(G)\backslash I$ since no edges have both ends in I. Conversely, if C is a vertex cover, then every edge is incident with C so no edges have both ends in $V(G)\backslash C$. Therefore $V(G)\backslash C$ is an independent set of G. We conclude that $\alpha(G) + \beta(G) = |V(G)|$. $\square$

Lemma 3.1.2 (Gallai's Lemma) *Let G be a graph with no isolated vertices. Then $\alpha'(G) + \beta'(G) = |V(G)|$.*

Proof $\triangleright$. Let M be a matching in G of size $\alpha'(G)$ – a maximum matching. Then no edge of G has both ends in $V(G)\backslash V(M)$, so $V(G)\backslash V(M)$ is an independent set of vertices. Now let us choose one edge incident with each vertex in $V(G)\backslash V(M)$ and all edges of M. The set of edges we get, say F, is an edge-cover of size $|E(M)| + |V(G)\backslash V(M)| = |V(G)| - \alpha'(G)$. Therefore $\beta'(G) \leq |V(G)| - \alpha'(G)$.

Conversely, let F be an edge-cover of G of size $\beta'(G)$ – a minimum edge-cover. Then $F - e$ is not an edge cover for any $e \in E(F)$. This means that each edge of F must cover one of its ends uniquely, so every edge of F has an end of degree one in F. In particular, every component of F is a star – a $K_{1,t}$ for some $t \geq 1$ (see Figure 34). Pick one edge from each component of F to get a matching M with $|E(M)|$ equal to the number of components of F. Since all components of F are stars,

$$\beta'(G) = |E(F)| = |V(F)| - |E(M)| = |V(G)| - |E(M)| \geq |V(G)| - \alpha'(G).$$

This completes the proof. $\square$

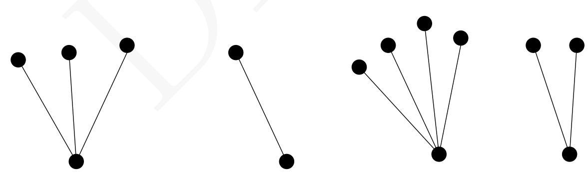

Figure 34: Structure of minimal edge-covers

Theorem 3.1.3 (König's Theorem) *If $G(A, B)$ is a bipartite graph with no isolated vertices, then $\alpha'(G) = \beta(G)$ and $\beta'(G) = \alpha(G)$.*

Proof $\triangleright$. It is sufficient to show $\alpha'(G) = \beta(G)$, by the last two lemmas. We give a proof of the theorem using Menger's Theorem. Add a vertex a adjacent to all vertices in A and a vertex b adjacent to all vertices in B. By Menger's Theorem, the maximum number of internally disjoint ab-paths in this new graph equals the minimum number of vertices required to separate a from b. The maximum is $\alpha'(G)$, and the minimum is $\beta(G)$, so $\alpha'(G) = \beta(G)$. $\square$

3.2 Hall's Theorem

Let X be a set of vertices in a graph G. We define $N(X)$ to be the ***neighbourhood of*** X, namely

$$\{y \in V(G)\backslash X : \{x, y\} \in E(G) \text{ for some } x \in X\}.$$

In other words, it is the set of vertices not in X adjacent to some vertex in X. Hall's Theorem gives a necessary and sufficient condition for a bipartite graph to have a perfect matching – and in fact a matching covering all vertices of one part. There are many proofs of Hall's Theorem; we give two proofs.

Theorem 3.2.1 (Hall's Theorem) *Let $G(A, B)$ be a bipartite graph. Then G has a perfect matching if and only if*

$$|N(X)| \geq |X| \quad \text{(Hall's Condition)}$$

for every set $X \subset A$ and every set $X \subset B$.

Proof $\triangleright$. The first proof we give is based on Menger's Theorem. First we observe that $|A| = |B|$, since $|B| \geq |N(A)| \geq |A|$ and $|A| \geq |N(B)| \geq |B|$. Let H be the graph obtained from G by adding a vertex a adjacent to all vertices of A and a vertex b adjacent to all vertices of B. Then G has a perfect matching if and only if there exist $|A|$ internally disjoint paths from A to B. By Menger's Theorem, such paths exist if and only if the minimum number of vertices required to separate a from b is $|A|$. For a contradiction, suppose that there exists a set S separating a from b with $|S| < |A|$. Let $X = A\backslash S$ and $Y = B\backslash S$. Now there are no edges of G between X and Y, so $N(X) \subset S \cap B$ and $N(Y) \subset S \cap A$. Therefore

$$\begin{array}{ccccc} |A\backslash S| & \leq & |N(X)| & \leq & |S \cap B| \\ |B\backslash S| & \leq & |N(Y)| & \leq & |S \cap A| \end{array}$$

Since $|S| < |A|$ and $|A| = |B|$, we have

$$|A\backslash S| + |B\backslash S| = |A| + |B| - |S| > |A|.$$

So if we add the two preceding inequalities, we obtain $|A| < |S|$, a contradiction. $\square$

Proof $\triangleright$. The second proof is a direct proof. By induction on $|A|$: we prove that if $|N(X)| \geq |X|$ for every set $X \subset A$ in a graph G, then G contains a matching saturating all vertices of A. Note that this proves Hall's Theorem, since we could apply the same statement to B to get a matching saturating all vertices of B. If $|A| = 1$, then the statement is true. Suppose $|A| > 1$. We consider two cases. The first case is that $|N(X)| > |X|$ for all $X \subset A$. In this case, pick any edge of G and remove both ends of that edge, say a and b. Then we obtain the bipartite graph $H(A\backslash\{a\}, B\backslash\{b\})$. In this bipartite graph, Hall's Condition holds in $A\backslash\{a\}$, and therefore H has a matching, M, saturating all vertices of $A\backslash\{a\}$. Now $M \cup \{a, b\}$ is the required matching in G. The second case is that for some $X \subset A$ or $X \subset B$, $|N(X)| = |X|$. Let $Y = N(X)$. In this case, consider the graph $G_1(A_1, B_1)$ obtained from G by removing all vertices of $X \cup Y$, and the graph $G_2(X, Y)$ consisting of all edges between X and Y. Then Hall's Condition holds in G_1 and also in G_2. To see that it holds in G_1, take any set $S \subset A_1$. Then:

$$|N_{G_1}(S)| + |N_{G_2}(X)| \geq |N_G(X \cup S)| \geq |X \cup S| = |X| + |S|$$

and since $|N_{G_2}(X)| = |N_G(X)| = |X|$, we have $|N_{G_1}(S)| \geq |S|$ for any $S \subset A_1$. By induction G_1 and G_2 have matchings, say M_1 and M_2, saturating all their vertices in A, and $M_1 \cup M_2$ is a matching in G saturating all vertices of A. □

A ***1-factorization*** of a graph G is a collection of edge-disjoint 1-factors $M_1, M_2, \dots, M_r$ such that $G = M_1 \cup M_2 \cup \dots \cup M_r$. For example, for even values of n, the complete graph K_n has a 1-factorization.

Corollary 3.2.2 *Let $G(A, B)$ be a k-regular bipartite graph, where $k \geq 1$. Then G has a 1-factorization.*

Proof ▷. It suffices to prove that G has a perfect matching. To see this, we apply Hall's Theorem. For a set $X \subset A$ or $X \subset B$, there are $k|X|$ edges of G incident with exactly one vertex of X. There are also $k|N(X)|$ edges incident with exactly one vertex of $N(X)$. This set of edges contains all edges incident with X, so $k|N(X)| \geq k|X|$ and Hall's Condition is satisfied. Therefore by Hall's Theorem G has a 1-factor. □

3.3 König-Ore Formula

A vertex not contained by any edge of a given matching is called ***unsaturated*** or ***exposed*** by the matching, and those vertices which are contained in edges of the matching are called ***saturated*** by the matching. Hall's Theorem gives a formula for finding $\alpha'(G)$ in a bipartite graph. For a bipartite graph $G(A, B)$, define $\text{ex}(G, A) = |A| - \alpha'(G)$: this is the number of vertices of A exposed by a maximum matching. Hall's Theorem gives a formula for $\text{ex}(G, A)$:

Theorem 3.3.1 (König-Ore Formula) *Let $G(A, B)$ be a bipartite graph. Then*

$$\text{ex}(G, A) = \max_{S \subset A} \{|S| - |N(S)|\}.$$

Proof ▷. Let d be the right hand side of the identity above. Add d vertices to B, all adjacent to all vertices of A. Then Hall's Condition – namely $|N(X)| \geq |X|$ for all $X \subset A$ – is satisfied in this new graph, so it has a matching covering all vertices of A, by Hall's Theorem. It follows that G has a matching of size at least $|A| - d$. Therefore $\text{ex}(A) \leq d$. Conversely, if M is a matching of size $|A| - \text{ex}(A)$, then each set $S \subset A$ has at least $|S| - \text{ex}(A)$ neighbours in B. In other words, $|N(S)| \geq |S| - \text{ex}(A)$ for all S so $d = \max\{|S| - |N(S)|\} \leq \text{ex}(A)$, as required. □

As an exercise, one can prove that a bipartite graph $G(A, B)$ of minimum degree δ and maximum degree $\triangle$ contains a matching of size at least $\delta|A|/\triangle$. ◁◁

3.4 Tutte's 1-Factor Theorem

There is a natural condition for a graph G to have a perfect matching: if S is a set of vertices of G and $H_1, H_2, \ldots, H_r$ are the ***odd components*** of $G - S$ – that is the components with an odd number of vertices – then none of the H_i can have a perfect matching, so each sends at least one edge of a perfect matching to S (see Figure 35). In particular $|S| \geq r$, so we have for all $S \subset V(G)$,

$$|S| \geq \text{odd}(G - S).$$

Note that if $S = \emptyset$, this asserts that G has an even number of vertices. Tutte's Theorem shows, remarkably, that this is also a sufficient condition.

Theorem 3.4.1 (Tutte's 1-Factor Theorem)
Let G be a graph. Then G has a perfect matching if and only if

$$|S| \geq \text{odd}(G - S) \qquad (\text{Tutte's Condition})$$

for every set $S \subset V(G)$.

Proof $\triangleright$. The proof we give is by induction on $|V(G)|$, the case $|V(G)| = 2$ being trivial. Let S be the largest subset of G such that equality holds in $|S| \leq \text{odd}(G - S)$. Such an S exists, because $|V(G)|$ is even, and so $G - s$ has at least one odd component for each $s \in V(G)$. Let F and H denote generic odd and even components of $G - S$.

Claim 1. *The graph H has a 1-factor.*
For any $R \subset V(H)$, as required we have:

$$\text{odd}(H - R) + \text{odd}(G - S) = \text{odd}(G - R \cup S) \leq |R| + |S|.$$

By induction H has a 1-factor.

Claim 2. *The graph $F' = F - v$ has a 1-factor for any $v \in V(F)$.*
By induction, if this is false, then there exists a set $Q \subset V(F')$ such that $\text{odd}(F' - Q) > |Q|$. Now for any set $R \subset V(F)$,

$$\text{odd}(F - R) + |R| \equiv |V(F)| \equiv 1 \mod 2$$

since F has an odd number of vertices (this step is really key to the proof). Therefore $\text{odd}(F' - Q) \geq |Q| + 2$, and so if $T = S \cup \{v\} \cup Q$, then

$$\begin{aligned} |T| &\geq \text{odd}(G - T) \\ &= \text{odd}(G - S) - 1 + \text{odd}(F' - Q) \\ &\geq |S| + |Q| + 1. \end{aligned}$$

This shows $\text{odd}(G - T) = |T|$, contradicting the maximality of S, and the claim is proved.

Claim 3. *Let $G(S,C)$ be the bipartite graph formed from G by contracting each odd component of $G-S$ to a single vertex, and taking all edges with one end in S and one end in the set C of contracted vertices. Then $G(S,C)$ has a perfect matching.*

To prove this, we use Hall's Theorem: for every set $X \subset C$,

$$|X| = \mathrm{odd}(G - N(X)) \leq |N(X)|$$

as required. Since $|S| = |C| = \mathrm{odd}(G-S)$, there is a 1-factor in $G(S,C)$.

To complete the proof of Tutte's 1-Factor Theorem, put together all the 1-factors that we found in Claims 1–3. Let $M_1, M_2, \ldots, M_r$ be 1-factors in the even components of G. Now let M be a 1-factor in $G(S,C)$. Then the edges of M form a matching in G, and for each odd component H_i of $G-S$, for $i \in \{1,2,\ldots,s\}$ where $s = \mathrm{odd}(G-S)$, there is exactly one vertex of H, say v_i, incident with an edge of M. Now Claim 2 gives a 1-factor N_i in $H - v_i$. Then

$$M \cup M_1 \cup \cdots \cup M_r \cup N_1 \cup N_2 \cup \cdots \cup N_s$$

is a perfect matching of G. □

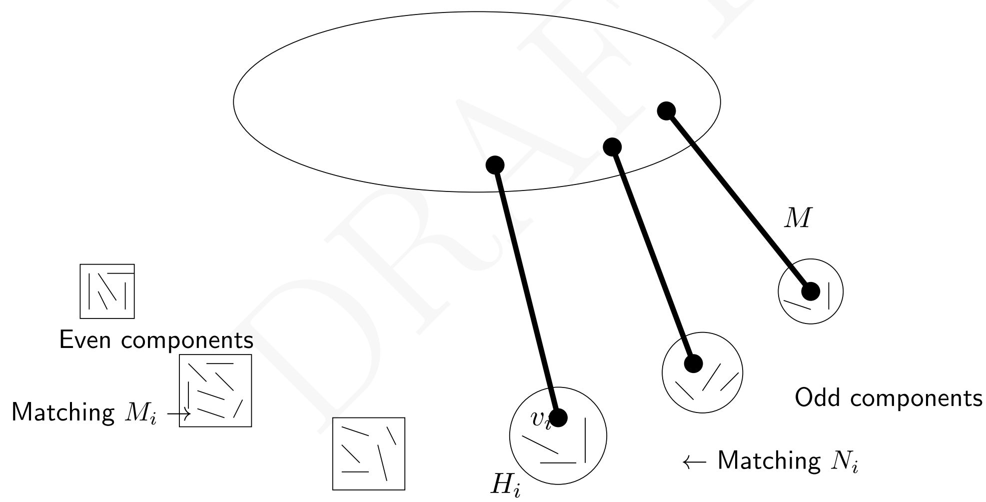

Figure 35: The proof of Tutte's Theorem.

From Tutte's 1-Factor Theorem, we obtain the following condition for a cubic (3-regular) graph to have a perfect matching:

Theorem 3.4.2 (Petersen's Theorem) *Any cubic bridgeless graph has a 1-factor.*

Proof ▷. We have to check Tutte's Condition. Pick a set $S \subset V(G)$. If $S = \emptyset$, then Tutte's Condition holds since G has an even number of vertices and is connected. Then there are at least two edges from S to each odd component of $G-S$. If H is an odd component of $G-S$, then it contains an even number of vertices of degree three, so it must send to S an odd number of edges. It must send at least three edges. So we have $3r$ edges out of odd components. On the other hand, G is cubic so $|S| \geq r$, as required. □

3.5 Tutte-Berge Formula

The Tutte-Berge Formula is the analog of the König-Ore Formula for non-bipartite graphs, and gives a method for finding $\alpha'(G)$. We define $\mathrm{ex}(G)$ to be the minimum number of vertices of G exposed by a matching of G – thus $\mathrm{ex}(G) = |V(G)| - 2\alpha'(G)$.

Theorem 3.5.1 (Tutte-Berge Formula) *For any graph G,*

$$\mathrm{ex}(G) = \max_{S \subset V(G)} \{\mathrm{odd}(G - S) - |S|\}.$$

The proof of this theorem is left as an exercise. The theorem can be used to give lower bounds ◁◁ on $\alpha'(G)$ for various graphs. For example, we apply the Tutte-Berge Formula to cubic graphs – graphs where all the vertices have degree three – to get a lower bound on $\alpha'(G)$:

Theorem 3.5.2
Let G be a cubic graph on n vertices. Then G has a matching of size at least $\frac{7n}{16}$.

Proof ▷. It may be assumed that G is connected, otherwise we pass to the components of G. We have to find an upper bound for $\mathrm{ex}(G)$, namely $\mathrm{ex}(G) \leq n/8$. By the Tutte-Berge formula, this is the same as showing $\mathrm{odd}(G - X) - |X| \leq n/8$ for all sets $X \subset V(G)$. Let $X \subset V(G)$ have size γ, and let α be the number of odd components of $G - X$ with at most three vertices, and β be the number of odd components of $G - X$ with at least five vertices. Let's call these α-components and β-components, respectively. Then $\mathrm{odd}(G - X) - |X| = \alpha + \beta - \gamma$. Now each α-component H of G is K_1 or K_3 or a path on three vertices. In each case, since G is cubic, $e(V(H), X) \geq 3$. Each β-component F of G has $e(V(F), X) \geq 1$. On the other hand, $e(X, V(G)\backslash X) \leq 3|X|$, since every vertex of X has degree three. Therefore

$$3\alpha + \beta \leq 3\gamma.$$

Next we observe that there are $n - \gamma$ vertices in $G - X$, but also at least $\alpha + 5\beta$ vertices in $G - X$, so

$$\alpha + 5\beta \leq n - \gamma.$$

We want to maximize $\alpha + \beta - \gamma$ subject to the above two inequalities. It is not hard to see that we must have $\alpha = 0$, $\beta = 3n/16$ and $\gamma = n/16$, in which case $\mathrm{ex}(G) = \alpha + \beta - \gamma = n/8$, as required. $\square$

Theorem 3.5.2 is best possible: the graph shown in Figure 36 is cubic with $n = 16$ vertices with no matching of size more than $7 = 7n/16$.

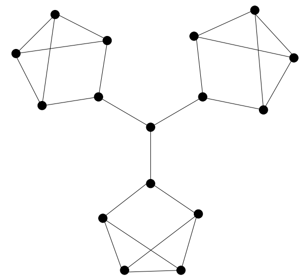

Figure 36: A cubic graph with no perfect matching

3.6 Matching Algorithms

In bipartite graphs, König's Theorem gives a practical way to find a maximum matching, using the notion of an augmenting path. An ***alternating path*** in a graph G with a matching M is a path whose every alternate edge is in M. An ***augmenting path*** for a matching M is an alternating path whose first and last edges are not in M.

Theorem 3.6.1 (Berge) *A matching M in a graph G is a maximum matching if and only if M does not admit any augmenting paths.*

Proof $\triangleright$. If M is a maximum matching, it clearly admits no augmenting path. Conversely, suppose M is a matching which does not admit an augmenting path, and $|M| < |N|$ for some maximum matching N. Then $M \cup N$ is a graph of maximum degree at most two, and so all the components of $M \cup N$ are paths or cycles (a linear forest). But since $|N| > |M|$, and any cycle in $M \cup N$ has as many edges of M as of N, there must be a path P such that $|E(P) \cap M| < |E(P) \cap N|$. This means that the first and last edge of the path are in N, and so the path augments M, a contradiction. $\square$

A version of ***König's matching algorithm*** (also known as the ***Hungarian Method***) for finding a maximum matching in bipartite graphs G is as follows:

König's Matching Algorithm.

1. identify the parts A and B of the bipartite graph G.

2. pick an arbitrary matching M in the graph.

3. let U be the set of exposed vertices with respect to M.

4.1. If two vertices of U are adjacent, add this edge to M and repeat. Otherwise go to 4.2.

4.2. U is an independent set of G. Starting at each vertex of U, grow an ***alternating path*** with respect to M.

4.2.1. If some alternating path is an augmenting path, say the path has edges $e_1, e_2, \ldots, e_{2k+1}$, we remove $e_2, e_4, \ldots, e_{2k}$ from M, and add $e_1, e_3, \ldots, e_{2k+1}$ to the matching, to obtain a matching larger than M. Then return to Step 3 with M equal to this new matching.

4.2.2. If no alternating path is augmenting, then M is a maximum matching.

It is not hard to see this algorithm runs in time at most $mn/2$ for a bipartite graph with m edges and n vertices, since there are m possible augmenting paths and at most $n/2$ iterations of the procedure. It turns out with more careful analysis that this algorithm runs in time at most $2\sqrt{n}m$.

Example 11. Consider the grid graph below. We use the matching algorithm to find a maximum matching in the grid, starting with the given matching $\{1, 2\}, \{5, 6\}$ shown in bold.

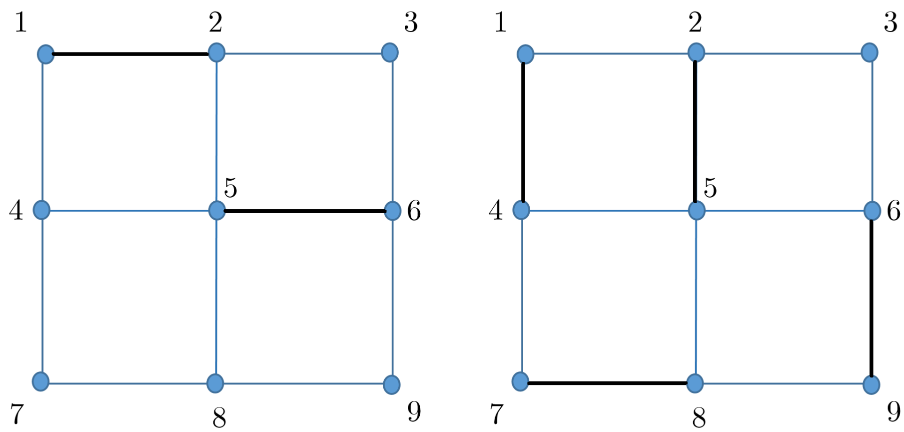

Figure 37: A matching in the grid graph.

First we identify the parts A and B of the grid graph. We may let $A = \{1, 3, 5, 7, 9\}$ and $B = \{2, 4, 6, 8\}$, and we are starting with the matching $M = \{\{1, 2\}, \{5, 6\}\}$.

The set U of exposed vertices with respect to M is $U = \{3, 4, 7, 8, 9\}$. Since $\{7, 8\}$ is an edge contained in U, we can add $\{7, 8\}$ to the matching M to get a new matching $M = \{\{1, 2\}, \{5, 6\}, \{7, 8\}\}$.[4] Now the set U of exposed vertices is $\{3, 4, 9\}$, and this is an independent set. We now grow alternating paths with respect to M, starting at vertices of U. Starting at $9 \in U$, an alternating path P has edges $\{9, 6\}, \{6, 5\}, \{5, 2\}, \{2, 1\}, \{1, 4\}$ and we stop since we arrived at a vertex in U, namely 4. This means P is an augmenting path: we take the edges $E(P) \cap E(M)$ out of M, and add the edges of $E(P)\backslash E(M)$ to M: so we take $\{5, 6\}, \{1, 2\}$ out of M and add the edges $\{9, 6\}, \{5, 2\}, \{1, 4\}$ to M. So now $M = \{\{9, 6\}, \{5, 2\}, \{1, 4\}, \{7, 8\}\}$ (see Figure 37.

Then we restart the algorithm with this matching M. The set U of exposed vertices is just $U = \{3\}$. We grow an alternating path starting at 3: the edges of such a path could be $\{3, 6\}, \{6, 9\}, \{9, 8\}, \{8, 7\}, \{7, 4\}, \{4, 1\}, \{1, 2\}$. All vertices of the graph are in this path, and the path did not end in U. Therefore M is a maximum matching (it was clear since there are 9 vertices in the graph, so at least one must be exposed by every maximum matching).

There is an algorithm for maximum matchings in general graphs, called ***Edmonds' matching algorithm***, but it is beyond the scope of this course.

[4] We could equally have added $\{4, 7\}$ or $\{8, 9\}$.

4 Vertex and Edge-Coloring

A ***proper k-edge-coloring*** of a graph G is a function $\chi : E(G) \to \{1, 2, \ldots, k\}$ such that if $e, f \in E(G)$ intersect, then $\chi(e) \neq \chi(f)$. In other words, any two edges which share a vertex must receive different colors (it is convenient to refer to the elements of $\{1, 2, \ldots, k\}$ as colors). The minimum k for which G has a proper k-edge-coloring is denoted $\chi'(G)$, and referred to as the ***edge-chromatic number of G***. Another way of saying it is: $\chi'(G)$ is the minimum number of matchings which partition $E(G)$, since the set of edges of any particular color is a matching. A graph G is ***k-edge colorable*** if $\chi'(G) \leq k$, and ***k-edge-chromatic*** if $\chi'(G) = k$. It is left as an exercise to verify that $\chi'(K_n) = n - 1$ when n is even and $\chi'(K_n) = n$ if n is odd. The main theorem we prove on edge coloring is ***Vizing's Theorem***, which states that $\chi'(G)$ is either $\triangle(G)$ or $\triangle(G) + 1$, where $\triangle(G)$ denotes the maximum degree of G.

A ***proper k-coloring*** of a graph G is a function $\chi : V(G) \to \{1, 2, \ldots, k\}$ such that if $u, v \in V(G)$ are adjacent, then $\chi(u) \neq \chi(v)$. So we color the vertices with k colors in such a way that no two adjacent vertices have the same color. The ***chromatic number*** of G is denoted $\chi(G)$, and is the minimum k for which G has a proper k-coloring. Thus $\chi(G)$ is the minimum number of independent sets which partition $V(G)$. For example, $\chi(K_n) = n$, and a graph G is bipartite if and only if $\chi(G) \leq 2$. We say that a graph is ***k-colorable*** if $\chi(G) \leq k$ and ***k-chromatic*** if $\chi(G) = k$. The main theorem on vertex coloring is Brook's Theorem, which states that $\chi(G) \leq \triangle(G)$ when G is not an odd cycle or a complete graph (for those graphs one has $\chi(G) = \triangle(G) + 1$).

4.1 König's Theorem

For any graph, it is clear that $\chi'(G) \geq \triangle(G)$ – all the edges incident with a vertex of degree $\triangle(G)$ must have different colors in a proper coloring. The main theorem we prove on edge-coloring is Vizing's Theorem. Before proceeding to Vizing's Theorem, we discuss edge-colorings of bipartite graphs. König's Theorem states that $\chi'(G) = \triangle(G)$ for any bipartite graph G – thus determining $\chi'(G)$ in bipartite graphs is easy:

Theorem 4.1.1 (König's Theorem) *For any bipartite graph G, $\chi'(G) = \triangle(G)$.*

Proof ▷ The first proof we give relies on Hall's Theorem: we know by Corollary 3.2.2 that every k-regular bipartite graph has a k-coloring. So if we can show that G is contained in a $\triangle(G)$ regular bipartite graph, then we are done. To prove this, take two copies of G, say $G_1(A, B)$ and $G_2(A, B)$, and if $y \in A \cup B$ has degree d, add $\triangle(G) - d$ multiple edges between the vertex of $G_1(A, B)$ corresponding to y and the vertex in $G_2(A, B)$ corresponding to y. Then we obtain a graph J which is $\triangle(G)$-regular, so $\chi'(J) = \triangle(G) = \chi'(G)$. □

Proof ▷ The second proof we give is by induction on $|E(G)|$. If $|E(G)| = 0$ then the theorem is clear. Suppose $|E(G)| > 0$ and let $e = \{x, y\} \in E(G)$. By induction, the graph $G - e$ is $\triangle(G)$-edge-colorable. If there is a color i which is not used on any edges incident with x or y, then we can assign color i to $\{x, y\}$ to get a $\triangle(G)$-edge-coloring of G. So we may assume that the colors at x are $1, 2, \ldots, \triangle(G)-1$ and the colors at y are $2, 3, \ldots, \triangle(G)$. Let H be the subgraph of G spanned by edges of colors 1 and $\triangle(G)$. Then the component of H containing x is a path or a cycle. It cannot be a cycle, otherwise x would be incident with an edge of color 1 and color $\triangle(G)$ in the cycle, contradicting that $\triangle(G)$ is missing at x. So the component of H containing x is a path, P. If P ends at y, then since P has odd length we would have an edge of color 1 at y, a contradiction. So P ends at a vertex $z \neq y$. Now z is not incident with any edge of color 1 or $\triangle(G)$ in $G - E(P)$, otherwise we could extend the path or the edge is incident with a vertex w of the path, but then the coloring would not be a proper edge-coloring. Now interchange colors 1 and $\triangle(G)$ along the path P, to obtain a proper coloring of $G - e$ where the color 1 does not appear at x. Finally, assign e color 1 to get a proper coloring of G. $\square$

4.2 Vizing's Theorem

The next remarkable theorem tells us that $\chi'(G)$ is either the maximum degree of G or one more than that. For example, for the complete graph K_n, we have $\chi'(K_n) = n-1$ if n is even and $\chi'(K_n) = n$ if n is odd (the first statement does require a proof – it is equivalent to saying we can partition K_n into $n-1$ edge-disjoint matchings when n is even – this is left as an exercise). Perhaps surprisingly, it is known to be difficult to determine whether $\chi'(G) = \triangle(G)$ or $\chi'(G) = \triangle(G)+1$ for a given graph G.

Theorem 4.2.1 (Vizing's Theorem) *For every graph G of maximum degree $\triangle$, $\chi'(G) = \triangle$ or $\chi'(G) = \triangle + 1$.*

Proof $\triangleright$. Since $\triangle$ different colors are needed at a vertex of degree $\triangle$ in G, $\chi'(G) \geq \triangle$. Now we prove by induction on $|E(G)|$ that G is $\triangle+1$-colorable, which gives $\chi'(G) \leq \triangle + 1$. If $|E(G)| = 0$, then the theorem is clearly true. Suppose $|E(G)| > 0$, and let $\{x, y_1\} \in E(G)$ be any edge of G. By induction, $G_1 = G - \{x, y_1\}$ is $\triangle + 1$-colorable. Now if there is a color, say color c_1, missing at y_1 and missing at x, then we can assign edge $\{x, y_1\}$ the color c_1. So we can assume that an edge on x, say $\{x, y_2\}$ has color c_1. Let c be a color missing at x – we know c appears on y_1 otherwise $\{x, y_1\}$ could be colored with color c. In general, we construct a maximal sequence $y_1, y_2, \ldots, y_k$ of neighbours of x such that c_i is missing at y_i and $\{x, y_{i+1}\}$ has color c_i for all $i < k$, and color c_k is missing at y_k and does not appear on any edge $\{x, y\}$ for $y \notin \{y_1, y_2, \ldots, y_k\}$.

<u>Case 1.</u> For all $i < k$, $c_k \neq c_i$. In this case, a proper edge-coloring of G is found by recoloring $\{x, y_j\}$ with color c_j for all $j \leq k$. Note that the coloring is proper since color c_j is missing at y_j for all $j \leq k$. An illustration is provided in Figure 38.

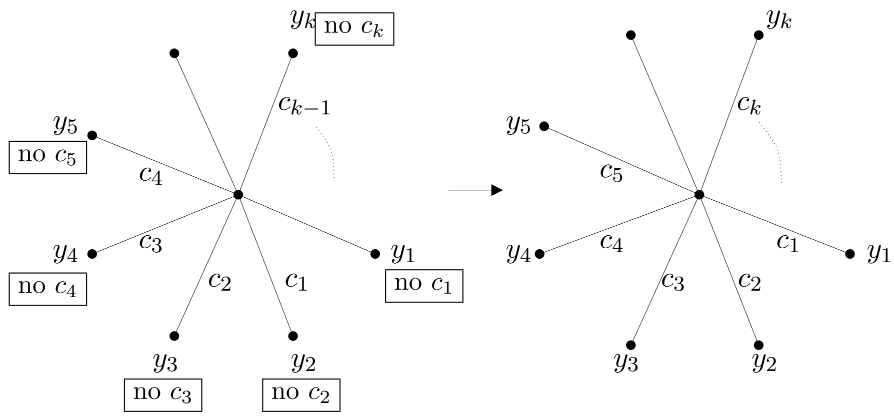

Figure 38: Rotate colors around x

<u>Case 2.</u> For some $i < k$, $c_k = c_i$. In this case, recolor all edges $\{x, y_j\}$ for $j \leq i$ with color c_j – so far we still have a proper coloring since color c_j is missing at y_j for $j \leq i$. This is shown for $i = 4$ in the left diagram in Figure 39. Then $\{x, y_{i+1}\}$ is the new uncolored edge, since the edge $\{x, y_1\}$ has now received color c_1. Now let H denote the subgraph of G consisting of edges of color c and edges of color c_k. Then the components of H are paths and cycles, since H has maximum degree at most two. Also x, y_{i+1}, y_k all have degree one in H, so either x, y_{i+1} are in different components of H or x, y_k are in different components of H. We consider these cases separately. If x, y_{i+1} are in different components of H, then we interchange colors c and c_i in the component of H containing y_{i+1}. In this new coloring, color c is missing at x and missing at y_{i+1}, so we can assign the edge $\{x, y_{i+1}\}$ the color c (see Figure 39). If x, y_k are in different components of H, then recolor the edge $\{x, y_j\}$ for $i < j < k$ with color c_j, so that $\{x, y_k\}$ is the new uncolored edge. Then H is unchanged (we never recolored edges of color c or c_i) so we may interchange the colors c and c_k in the component of H containing y_k. In doing so, c becomes a missing color at x and y_k, so the uncolored edge $\{x, y_k\}$ can be colored with color c. This completes the proof. $\square$

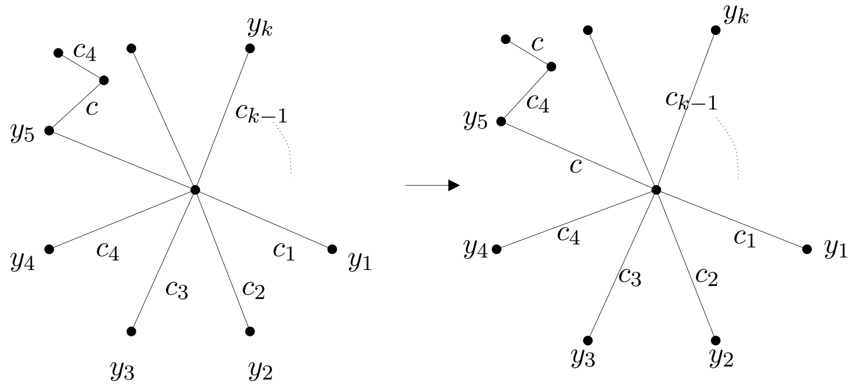

Figure 39: Interchanging colors in components of H

4.3 Brook's Theorem

The chromatic number of a graph G is the minimum number of colors which can be assigned to the vertices of G so that no two adjacent vertices have the same color. This number is denoted $\chi(G)$. Unlike in the case of edge-coloring, $\chi(G)$ can be arbitrarily small relative to $\triangle(G)$: for example $\chi(G) \leq 2$ if and only if G is a bipartite graph. One also notices that $\chi(G) = \triangle(G)+1$ is possible, since $\chi(K_n) = n$ and $\chi(C) = 3$ when C is an odd cycle. In fact these are the only cases where $\chi(G) = \triangle(G) + 1$:

Theorem 4.3.1 (Brook's Theorem) *Let G be a connected graph of maximum degree $\triangle$. Then $\chi(G) \leq \triangle(G)$, unless G is an odd cycle or a complete graph.*

To prove this theorem, we will first prove a proposition which often gives a better bound for $\chi(G)$ than $\triangle(G)$. The idea is to remove vertices of small degree from the graph and to notice that whenever we remove a vertex of degree k from a graph G and obtain a graph with a proper $(k+1)$-coloring, then we can reinsert the vertex and color it to obtain a proper $(k+1)$-coloring of G.

Proposition 4.3.2 *Let $d(G)$ denote the largest possible value of $\delta(H)$ taken over all subgraphs of G. Then $\chi(G) \leq d(G) + 1$.*

Proof ▷. By definition, G has a vertex v of degree at most $d(G)$. When we remove this vertex from G, we obtain a graph G' and clearly $d(G') \leq d(G)$. Therefore G' is $(d(G) + 1)$-colorable. Now v has $d(G)$ neighbours, so we can definitely assign v a color from the $d(G)+1$ colors used to color G', to obtain a proper coloring of G. □

A graph is called ***d*-degenerate** if $d(G) \leq d$, and we refer to $d(G)$ as the ***degeneracy of the graph***. So the proposition states that a d-degenerate graph is $(d + 1)$-colorable. We will use this fact later on when coloring planar graphs.

Proof ▷ of Brook's Theorem. Since a graph G of maximum degree $\triangle$ is definitely $\triangle$-degenerate, $\chi(G) \leq \triangle + 1$, by the last proposition. We omit the proof that if $\chi(G) = \triangle + 1$ then G is an odd cycle or a complete graph. □

The ***contraction*** of a pair of vertices $\{a, b\} \subset V(G)$ is the graph $G/\{a, b\}$ obtained from G by identifying the vertices a and b and joining the new vertex ab to all neighbours of a and all neighbours of b. Suppose a and b are not adjacent in G; then a k-coloring of $G/\{a, b\}$ gives a k-coloring of G: all vertices of G have the same color as $G/\{a, b\}$, except the vertices a and b are both assigned the color of ab. So the set of all proper colorings of $G/\{a, b\}$ is in 1-1 correspondence with the set of all proper colorings of G in which a and b receive the same color. Similarly, the set of colorings of $G - e$ for an edge $e \in E(G)$ is in 1-1 correspondence with the colorings of G in which the ends of e get different colors. These two notions are key reductions throughout the theory of graph coloring.

5 Planar Graphs

Roughly speaking, a graph is planar if and only if it can be drawn in the plane without any two of its edges crossing. More formally, an ***embedding*** of a graph $G = (V, E)$ is a function $f : V \cup E \to \mathbb{R}^2 \cup \mathcal{C}$, where $\mathcal{C}$ is the set of continuous curves in $\mathbb{R}^2$, such that f is one-to-one, $f(v)$ is a point in $\mathbb{R}^2$ for each $v \in V$, and $f(\{u, v\})$ is a continuous curve in $\mathbb{R}^2$ with ends u and v when $\{u, v\} \in E$. The graph G is ***planar*** if we can choose f so that the curves $f(e) : e \in E$ meet only at their ends – that is no curve meets itself and any point in the intersection of two distinct curves is an endpoint of both of the curves. A drawing of G without crossings is called a ***plane embedding*** of G, or a ***plane graph***. Thus a graph is planar if and only if it has a plane embedding.

The main theorem of this section, which we will prove, is a necessary and sufficient condition for a graph to be planar – and a characterization of planar graphs. A ***subdivision*** of a graph G is any graph obtained from G by replacing each edge of G with a path with the same ends as the edge, such that paths may meet only at their ends.

Theorem 5.0.1 (Kuratowski's Theorem) *A graph is planar if and only if it contains no subdivision of K_5 and no subdivision of $K_{3,3}$.*

5.1 Euler's Formula

Throughout this section, we deal only with connected graphs. If G is a plane graph, then $\mathbb{R}^2 \backslash G$ consists of a union of disjoint connected plane regions, which are called ***faces*** of G. The ***boundary*** ∂F of a face F of G is the set of points in the topological closure of F which are not in the interior of F. Each plane graph has a face which is infinite, which we refer to as the ***infinite face***. The ***boundary walk*** of a face F, which we also denote by ∂F, is the closed walk consisting of edges and vertices in the boundary of F. We denote by $F(G)$ the set of faces of a plane graph G. The ***degree*** of a face $F \in F(G)$ is the length of the walk ∂F, and denoted $\deg(F)$.

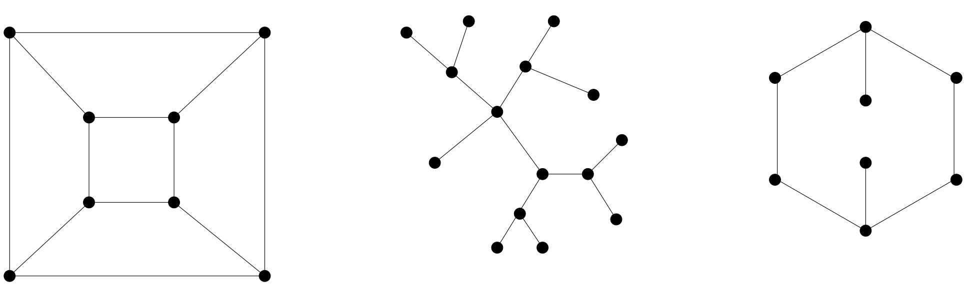

Figure 40: Faces of a plane graph

The graph on the left in Figure 40 has six faces, all boundary walks of which are cycles of length four – so every face has degree four. The tree in the centre has only one face – the infinite face – and since a tree on n vertices has $n-1$ edges and the boundary walk goes through each edge twice, the degree of the infinite face is $2(n-1)$. In the graph on the right, there are two faces, one of degree six and one of degree ten.

The degrees of the faces in a plane graph depend very much on the way the graph is drawn in the plane: for example, the graph on the right in Figure 40 can be redrawn as a new plane graph by flipping one of the bridges into the infinite face, thereby producing two new faces, both of degree eight.

There is a very useful analog of the handshaking lemma for face degrees in a plane graph. If we add up the degree of every face $F \in F(G)$, we observe that every edge of the graph is counted exactly twice. This is true since an edge in a cycle is counted once for each of the faces on either side of it, and an edge which is not in a cycle is a bridge (Lemma 2.1.1), and therefore counted twice in one boundary walk. These observations give the following useful fact:

Theorem 5.1.1 *Let G be a plane graph. Then*

$$\sum_{F \in F(G)} deg(F) = 2|E(G)|.$$

For the examples in Figure 40, one checks that the theorem holds. In general, note the a bridge on the boundary of a face is counted twice in the boundary walk of that face, whereas all other edges in the boundary are counted once.

Theorem 5.1.1 is very useful in conjunction with Euler's Formula and the handshaking lemma for proving non-existence of planar graphs with given face and vertex degrees. Euler's Formula relates $|F(G)|$, $|E(G)|$ and $|V(G)|$ as follows:

Theorem 5.1.2 (Euler's Formula) *Let G be a connected plane graph. Then*

$$|V(G)| - |E(G)| + |F(G)| = 2.$$

Proof ▹ Proceed by induction on $|E(G)|$. The minimum value of $|E(G)|$ is $|V(G)| - 1$, by Proposition 2.1.2. In that case, $|F(G)| = 1$ and Euler's Formula is satisfied. So we may assume that $|E(G)| > |V(G)| - 1$ and G contains a cycle C. Let e be an edge of C. By Lemma 2.1.1, $G - e$ is connected, since e is not a bridge. By induction,

$$|V(G-e)| - |E(G-e)| + |F(G-e)| = 2.$$

We now observe $|E(G-e)| = |E(G)| - 1$ and $|F(G-e)| = |F(G)| - 1$ and $|V(G-e)| = |V(G)|$. It follows that

$$|V(G)| - (|E(G)| - 1) + (|F(G)| - 1) = 2$$

and this gives Euler's Formula. □

A useful application is to give a sufficient condition for non-planarity, rather than trying to draw the graph in every possible way:

Theorem 5.1.3 *Let G be a planar graph containing a cycle. Then $|E(G)| \leq \frac{g}{g-2}(|V(G)|-2)$, where g is the length of a shortest cycle in G. In particular, for any planar graph G, $|E(G)| \leq 3|V(G)| - 6$, and therefore G is 5-degenerate.*

Proof $\triangleright$. Since every face has degree at least g, Theore 5.1.1 gives $g|F(G)| \leq 2|E(G)|$. Putting this in Euler's Formula, we get

$$|V(G)| - |E(G)| + \frac{2}{g}|E(G)| \geq 2$$

which, rearranged, gives the required bound on $|E(G)|$. The right side of the formula is maximized when $g = 3$, in which case we get $|E(G)| \leq 3|V(G)| - 6$ for all planar graphs G. By the handshaking lemma, if all vertices of G had degree at least six, then $|E(G)| \geq 3|V(G)|$, a contradiction to what we just proved. So every planar graph has a vertex of degree at most five. Since every subgraph of G is also planar, this means that every subgraph of G has a vertex of degree at most five, so G is 5-degenerate. $\square$

By Theorem 5.1.3, any graph satisfying $|E(G)| > \frac{g}{g-2}(|V(G)|-2)$ can't be planar. In particular, K_5 is not planar since $|E(K_5)| = 10$ and $g = 3$, and $K_{3,3}$ is not planar since $|E(K_{3,3})| = 9$ and $g = 4$.

5.2 Coloring Planar Graphs

Euler's Formula also can be applied to vertex-coloring of planar graphs. Recall that a graph is d-degenerate if every subgraph of G (including G itself) has minimum degree at most d. Also, any d-degenerate graph is $(d+1)$-colorable, by Proposition 4.3.2. By Theorem 5.1.3, every planar graph is 5-degenerate, so this means that every planar graph is 6-colorable. Here is another example: suppose we have a planar graph G of girth at least six. Then $|E(G)| \leq \frac{3}{2}(|V(G)|-2)$ by Theorem 5.1.3, so every subgraph of G must have a vertex of degree at most two, by the handshaking lemma. Therefore G is 2-degenerate, which means that G is 3-colorable. In fact, every planar graph is 5-colorable.

Theorem 5.2.1 *Every planar graph is 5-colorable.*

Proof $\triangleright$. Proceed by induction on $|V(G)|$. If $|V(G)| \leq 5$, then the theorem is true: just assign all vertices different colors. Now suppose $|V(G)| > 5$. If G has a vertex v of degree at most four, then $G - \{v\}$ is 5-colorable by induction, and we can extend this coloring to v by assigning to v a color which does not appear on any of its neighbours, since there were five colors but at most four neighbours of v. So now we assume G has no vertex of degree at most four.

Since G is 5-degenerate, by Theorem 5.1.3, G has a vertex v of degree exactly five. If all neighbours of v are adjacent to each other, then they form a K_5, but K_5 is not planar (as we saw from Theorem 5.1.3), so that is a contradiction. Therefore we can pick neighbours a and b of v which are not adjacent. Now consider the graph H obtained from $G - \{a, b, v\}$ by adding a new vertex w joined to every vertex in $N(a) \cup N(b) \cup N(v)$ – see Figure 3. This graph H is still planar. By induction, H has a 5-coloring $\chi : V(H) \to [5]$. For each $u \in V(G) \backslash \{a, b, v\}$, let assign it color $\chi(u)$. Assign a and b the color $\chi(w)$ – we can do that since a and b are not adjacent. Finally, the number of colors used by neighbors of v is at most four, since a and b got the same color. So there is a color i not used by any neighbor of v, and we let $\chi(v) = i$ to get a 5-coloring of G. □

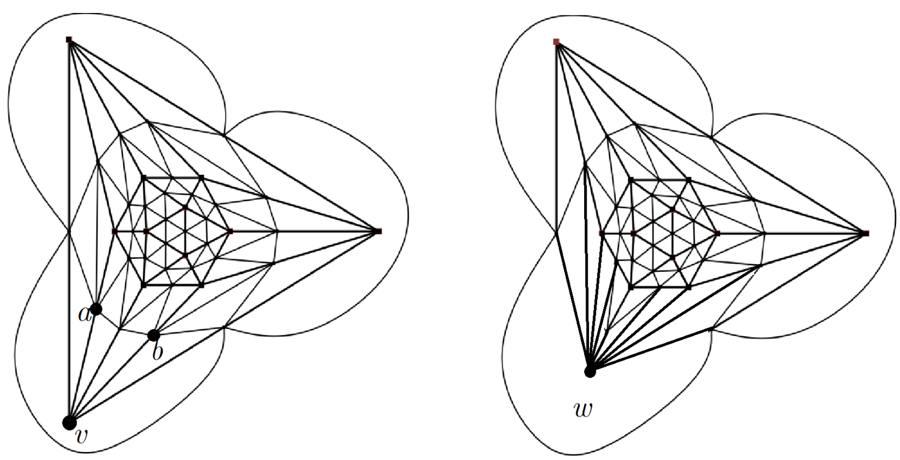

Figure 41: Proof of the 5-Color Theorem

Perhaps the most famous theorem in all of graph theory is the 4-color theorem, proved by Appel and Haken (1976): every planar graph is 4-colorable. Unfortunately, there is no proof known which is not computer assisted. The shortest proof is currently the one in Robertson and Seymour (1997).

Theorem 5.2.2 (4-Colour Theorem)
Every planar graph is 4-colorable.

5.3 Drawing Planar Graphs

We remarked earlier that there are many plane embeddings for a given planar graph G; even the degrees of the faces can change with different embeddings. In fact, we can go from any plane embedding of G to any other plane embedding of G using the notion of stereographic projection. In particular, we can make any face of a plane embedding the infinite face:

Proposition 5.3.1 *Every face of a plane embedding G_0 of a graph G is the infinite face of some plane embedding of G. Furthermore, if every edge of G_0 is a straight line, then we can ensure that every edge of the new embedding is also a straight line.*

Proof ▷. Let $\mathbb{S}$ denote a sphere of diameter one placed so that the xy-plane is tangent to $\mathbb{S}$ at the origin. Then wrap the plane embedding G_0 of G around the sphere. Formally, consider the function f which maps a point (x, y) to the point $(x, y, z) \in \mathbb{S}$ which is at height $1 - 1/(1 + x^2 + y^2)$ in the plane defined by the line through the origin and (x, y) and the z-axis. Note that f is a bijection between $\mathbb{R}^2$ and $\mathbb{S}\backslash N$, where $N = (0, 0, 1)$ denotes the north pole of $\mathbb{S}$. Let H be the image of G_0 under f. Keeping H fixed, rotate the sphere until some face F of H contains the north pole of $\mathbb{S}$. Now apply f^{-1} to get an embedding of G, namely $f^{-1}(H)$, with the property that the face F of $f^{-1}(H)$ is the infinite face. The second statement of the theorem is left as an excercise. □

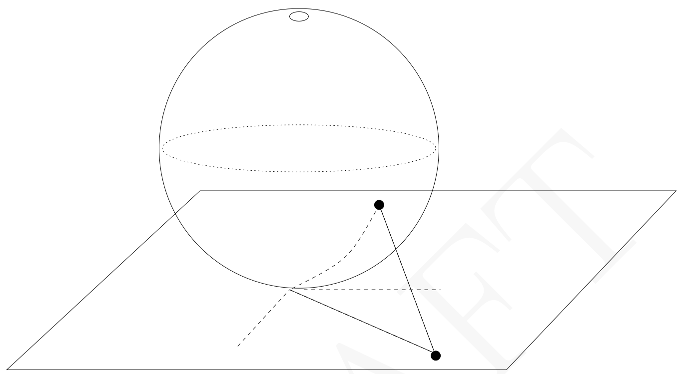

Figure 17 : Stereographic projection

Perhaps the most natural embedding would be to try to draw the edges as striaght lines. This can be done, by the following theorem:

Theorem 5.3.2 (Fary's Theorem) *Every simple planar graph has a plane embedding in which all edges are straight line segments.*

Proof ▷. By Theorem 5.1.3, every planar graph is 5-degenerate. Now we proceed by induction on $|V(G)|$, the number of vertices in a planar graph G. If $|V(G)| \leq 3$, then the result is obvious. Suppose G is a planar graph with $|V(G)| > 3$. We may assume that G is maximal planar – so $G + e$ is not planar anymore for any edge e. Then if G_0 is a plane embedding of G, all faces of G_0 have degree three, otherwise we could add a diagonal edge in some face. Now let v be a vertex of degree at most five in G_0. Then $G_0 - v$ has a plane embedding, call it H, such that all the edges are straight lines. Let $v_1, v_2, \ldots, v_k$ be the neighbours of v, where $k \leq 5$. Then there is a cycle $C \subset H$ such that $V(C) = \{v_1, v_2, \ldots, v_k\}$ – since every face of G_0 is of degree three, every face of H containing non-neighbours of v on its boundary is a triangle. This means that C is the boundary of a face of H. By Proposition 5.3.1, we can make C the boundary of a finite face of H. Now place v in the interior of C so that v sees all vertices of C – that is, we can draw a straight line segment from v to each vertex $v_1, v_2, \ldots, v_k$. This is an embedding of G in which all edges are straight line segments. $\square$

Concerning properties of the drawing of a planar graph, we have seen that there are, in general, plane embeddings with different face degrees (Figure 16). Furthermore, we can't ensure that the faces are convex, even for 2-connected planar graphs, for example $K_{2,3}$ has no embedding in which all faces are convex. However, Tutte showed that every 3-connected planar graph can be drawn with convex faces and straight line edges, and Whitney's Theorem states that the embedding is unique.

Theorem 5.3.3 (Tutte-Whitney Theorem)
Every 3-connected planar graph has a unique embedding in the plane in which every face is a convex polygon and every edge is a straight line segment.

A natural physical interpretation is to nail down the edges of a cycle which is a face in a plane embedding of G, and replace the edges with rubber bands. Then, allowing this dynamical system to reach equilibrium in terms of the laws of physics, Tutte proved that the plane embedding at equilibrium is a convex straight line embedding. We do not prove this or Theorem 5.3.3 here.

5.4 Duality

Let G be a plane graph, and let G^* denote the graph obtained by placing a vertex v_f in the interior of each face $f \in F(G)$ and whose edges are defined as follows: (1) for each bridge on the boundary of a face $f \in F(G)$, join v_f to v_f with a loop in G^* passing through the bridge. (2) for each edge $e \in E(G)$ on the boundary of distinct faces $f, g \in F(G)$, join v_f and v_g by an edge in G^* which crosses e. Then G^* is referred to as the ***plane dual*** or ***combinatorial dual*** of G. Examples of duals are shown in Figure 42:

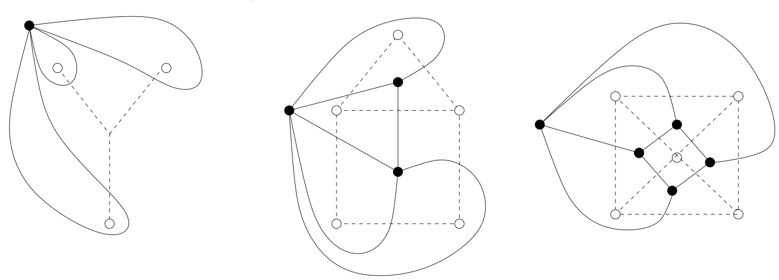

Figure 42: Duality

There are many uses of duality in planar graphs, but for brevity we mention one example in coloring. The map coloring problem is to color the faces of a plane graph in such a way that whenever two faces share an edge, they have different colors. Now by drawing the dual of a planar graph, we see that the map coloring problem on a plane graph is equivalent to the vertex coloring problem in the dual, except that we have to remove all loops in the dual. By the 4-color theorem, this means that the regions of any map can be colored in four colors in such a way that adjacent regions have different colors. In fact, even more is true: if we want to prove the 4-color theorem for plane graphs, it is sufficient to consider maximal plane graphs on at least three vertices(i.e. if we add any edge we get a non-planar graph). In a maximal plane graph, all faces are bounded by triangles (exercise) , and therefore the dual of a maximal $\triangleleft\triangleleft$ plane graph is a cubic graph. It can also be checked that every maximal plane graph, except a triangle, is 3-connected, and that the dual is therefore also 3-connected. The oldest approach to the 4-color theorem is to try to prove that every cubic graph is 3-edge-colorable: in fact this is equivalent to the 4-color theorem.

Theorem 5.4.1 *Every planar graph is 4-colorable if and only if every cubic planar 3-connected graph has edge-chromatic number three.*

Proof $\triangleright$. Let G be a planar graph and let G_0 be a plane embedding of G. Then G_0 is contained in a maximal plane graph G_1. If every planar graph is 4-colorable, then G_1 is 4-colorable which means that the map G_1^* is 4-face-colorable and cubic. Since G_1 is 3-connected, no edge of G_1^* is a bridge so every edge of G_1^* is on the boundary of exactly two faces. Now assign edge-color 1 to those edges of G_1^* on the boundary of faces of color 1 and 2, or color 3 and 4, assign edge-color 2 to those edges of G_1^* on the boundary of faces of colors 1 and 3, or colors 2 and 4, and assign edge-color 3 to all remaining edges of G. One checks that this is a proper 3-edge-coloring of G^*, as required.

Define G, G_0, G_1, G_1^* as in the first part of the proof. If every cubic planar graph is 3-edge-colorable, then G_1^* has a proper 3-edge-coloring, with colors 1, 2 and 3. That is to say that $G_1^* = M_1 \cup M_2 \cup M_3$ where M_i is the perfect matching consisting of edges of color i. Then $H_1 = M_1 \cup M_2$ is a plane graph and $H_2 = M_1 \cup M_3$ is a plane graph. Colour the faces of H_1 with colors 1 and 2, and color the faces of H_2 with colors 1' and 2'. To get a coloring of the faces of G_1^*, and hence a color of G, color a face F with color (i, j') if it is contained in a region of color i in H_1 and a region of color j' in H_2. Then the number of colors we used is four, and one checks that this a proper coloring of the faces of G_1^*. $\square$

Tait conjectured that all 3-connected cubic planar graphs are hamiltonian – i.e. contain a spanning cycle – but this is false, as a counterexample of Tutte on forty-six vertices showed (Figure 43). Tutte's counterexample is shown below. If Tait's conjecture had been true, then we could color the hamiltonian cycle red and blue, and the remaining matching with green to get a proper 3-coloring of every cubic 3-connected graph.

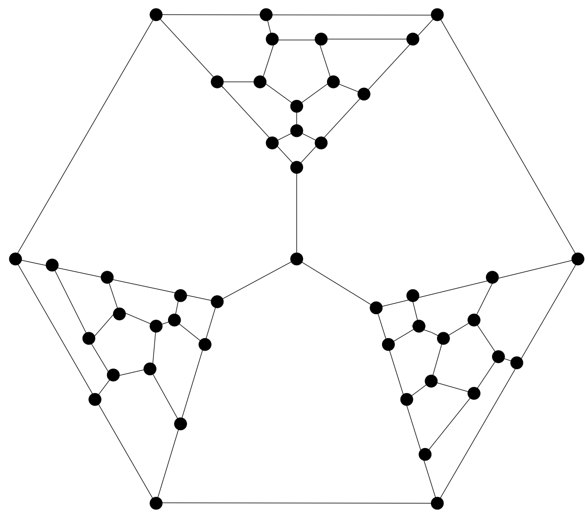

Figure 43: Tutte's Graph

5.5 Kuratowski's Theorem

In this section, we'll give a proof of Kuratowski's Theorem. There are many proofs of this theorem; we give a fairly recent proof by Makarychev (1998). First, recall from Chapter 1 (see Section 2.2) that a graph G which does not contain a theta-graph is a tree of cycles and bridges. We refer to a cycle or bridge containing only one cutvertex of G as an endblock. The second ingredient is the following lemma, whose slightly technical proof is left as an exercise. Let $\mathcal{S}$ denote the set of all graphs containing a subdivision of K_5 or $K_{3,3}$.

Lemma 5.5.1 *For any graph G and $e \in E(G)$, $G/e \in \mathcal{S}$ implies $G \in \mathcal{S}$.*

The final notion is the following: if C is a cycle in a plane graph G, then $\mathrm{int}(C)$ denotes the set of vertices of G inside C, and $\mathrm{ext}(C)$ denotes the set of vertices of G outside C.

Proof ▷ of Kuratowski's Theorem. If G is planar then $G \notin \mathcal{S}$. Now let $G \notin \mathcal{S}$ be a minimal non-planar graph. By case checking, we see that $|V(G)| > 6$. Furthermore, every proper subgraph of G is planar, and G/e is planar, by Lemma 5.5.1. Clearly $d(v) \geq 2$ for all $v \in V(G)$, otherwise $G - \{v\}$ is planar which implies G is planar. Also, $d(v) > 2$ for all $v \in V(G)$: otherwise with $N(v) = \{u, w\}$, the graph $G/\{u, v\}$ is planar. Since G is a subdivision of $G/\{u, v\}$ – we insert v into $\{u, w\}$ – G is also planar, a contradiction.

<u>Part 1</u> *For $\{u, v\} \in E(G)$, $G_{uv} := G - \{u\} - \{v\}$ contains no theta-graph.*
Suppose $T \subseteq G_{uv}$ is a theta-graph and let $C \subset T$ be a cycle. By Proposition 5.3.1, $G/\{u, v\}$ has a plane embedding H with $u, v \in \mathrm{int}(C)$ and $\mathrm{ext}(C) \neq \emptyset$. Now $H - \mathrm{int}(C) = G - \mathrm{int}(C)$ is a plane graph in which C is a face boundary. Also $H - \mathrm{ext}(C)$ is a plane graph with C as the infinite face boundary, and since $G - \mathrm{ext}(C)$ is planar, there is a plane embedding I of

$G - \text{ext}(C)$ in which C is the infinite face boundary (careful : this key step is subtle). Gluing I and $H - \text{int}(C)$ along C, we get a plane embedding of G, a contradiction.

<u>Part 2</u> *For $\{u, v\} \in E(G)$, G_{uv} has at most one leaf.*
Let X be a set of two leaves of G_{uv} and $Y = V(G_{uv})\backslash X$. Notice that $e(X, Y) = 2$ and since $|V(G)| > 6$, $|Y| > 2$. Since $d(x) \geq 3$ for $x \in X$, $u, v \in N(x)$ for $x \in X$. This implies $G - Y$ contains a theta-graph, and Part 1 shows $E(G[Y]) = \emptyset$. Since $d(y) \geq 3$ for $y \in Y$, and y sends at most two edges to $\{u, v\}$, $e(X, Y) \geq |Y| > 2$, a contradiction.

Now we complete the proof. Part 1 and Proposition 2.4.2 show that G_{uv} is a tree of cycles and bridges. By Part 2, G_{uv} has at most one leaf, so some endblock $C \subseteq G_{uv}$ is a cycle. Let $P \subseteq C$ be a path of length two; if $|V(C)| > 3$ then P can be chosen to contain no cutvertex of G_{uv}. Since $|V(G)| > 6$ and G_{uv} has at most one leaf, we can find an edge $\{w, x\} \in G_{uv}$ (see Figure 44) vertex-disjoint from P. Now each vertex of P is adjacent to u or v, since G has minimum degree at least three. This implies $G[V(P) \cup \{u, v\}]$ contains a theta-graph, vertex-disjoint from $\{w, x\}$. This contradicts Part 1 applied to G_{wx}. $\square$

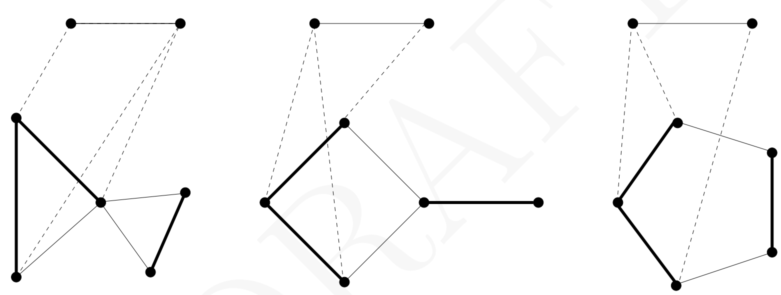

Figure 44: The path $P \subset C$ and edge $\{w, x\}$

5.6 Graphs on Surfaces

In this section, we study drawing graphs on general surfaces without crossings. Rather than assume background in general topology, we will define everything in elementary terms, keeping in mind that everything can be made rigorous through topology. The surfaces we look at will all be orientable: these are closed surfaces which consist of a sphere with a finite number of handles (or tube) attached. For example, the torus consists in attaching one handle to the sphere, and the double torus consists in attaching two handles to the sphere. The number of handles attached to $\mathbb{S}$ is called the genus of $\mathbb{S}$, and denoted $\gamma(\mathbb{S})$. The Euler characteristic of $\mathbb{S}$ is $\chi(\mathbb{S})$, defined by $2 - 2\gamma(\mathbb{S})$. For example, the Euler characteristic of the sphere is two, whereas the euler characteristic of the torus is zero. Let $G = (V, E)$ be a graph. An embedding of G without crossings on a surface $\mathbb{S}$ is a function $f : V \cup E \to \mathbb{S} \cup \mathcal{C}$, where $\mathcal{C}$ is the set of continuous curves in $\mathbb{S}$, such that f is one-to-one, $f(v)$ is a point in $\mathbb{S}$ for each $v \in V$, and $f(\{u, v\})$ is a continuous curve in $\mathbb{S}$ with ends u and v when $\{u, v\} \in E$, and none of the curves in $\mathcal{C}$ cross internally. Generally we identify G with the image of $V \cup E$ under f. Throughout this section, we assume that the connected regions (the faces) of $\mathbb{S} \backslash G$ are homeomorphic to discs. For example, an embedding of K_5 on the torus is shown below, where all faces are homeomorphic to discs. It is convenient to call a graph toroidal if it can be embedded on the torus without crossings. As an exercise, find an embedding of K_4 on the torus with two faces, $\triangleleft\triangleleft$ but such that one of the faces is not homeomorphic to an open disc.

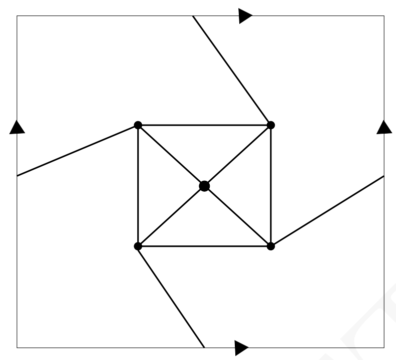

Figure 45: Toroidal embedding of K_5.

Thus the definition of faces in a drawing of a graph on a surface $\mathbb{S}$ without crossings is the same as in plane graphs. The degree of a face is again the shortest closed walk which traverse every edge on the boundary of the face (since the faces are essentially discs these definitions are the same as for plane graphs, but we could define it even if the faces were not discs). For example, in the toroidal embedding of K_5 in Figure 45, there are five faces, four having degree three and one having degree four. It is a good exercise to try to embed other graphs, for example $K_{4,4}$, in the torus. A note on drawings: take a square $[0,1] \times [0,1]$ and identify $(0,a)$ with $(1,a)$ and $\triangleleft\triangleleft$ identify $(b,0)$ with $(b,1)$ for all $a, b \in [0,1]$. Then we obtain the torus (the reverse procedure corresponds to cutting a torus along a cross section and then cutting the resulting cylinder along its height). A drawing of a graph without crossings in the rectangle with opposite points identified then gives a toroidal drawing of the graph. Roughly speaking, an edge can pass through the top (respectively, left) of the square and emerge from the vertically opposite (respectively, horizontally opposite) point at the bottom (respectively, right).

The genus of a graph G, denoted $\gamma(G)$, is the minimum possible value of γ such that G embeds without crossings in a surface of genus γ. So $\gamma(K_5) = 1$. The Euler characteristic of G is $\chi(G) = 2 - 2\gamma(G)$)not to be confused with the chromatic number). The Euler-Poincaré formula is an analog of Euler's formula for surfaces, and it says that the number of faces in any embedding of a graph on a surface of characteristic χ does not depend on the embedding (again with the proviso that the faces behave like discs):

Theorem 5.6.1 (Euler-Poincaré Formula) *Let G be a connected graph of Euler characteristic χ embedded without crossings in a surface $\mathbb{S}$ of characteristic χ, where all faces of G are homeomorphic to discs. Then*

$$|V(G)| - |E(G)| + |F(G)| = \chi.$$

The proof of this theorem can be achieved by induction on e, similar to Euler's Formula. Actually it is convenient rather to prove that $|V(G)|-|E(G)|-|F(G)| = \chi(G)+c(G)-1$ where $c(G)$ is the number of components of G. For instance, if G has a bridge e, then $\chi(G) = \chi(G-e)$ and $|F(G)| = |F(G-e)|$ and $c(G-e) = c(G)+1$ so the induction step works. Now suppose G has no bridges. Then for an edge $e \in E(G)$ such that $\chi(G-e) = \chi(G)$, we have $|F(G-e)| = |F(G)| - 1$ and again the induction step works. Finally if $\chi(G-e) = \chi(G) - 1$, then e must have been the only edge on some handle of $\mathbb{S}$. Then one verifies that $|F(G-e)| = |F(G)| + 1$, so the induction step works. The details are left to the reader. The formula is called the Euler-Poncaré formula since Poincaré gave a very general topological generalization of it, where edges are replaced by simplices. The first proof of the Euler-Poincaré formula for general χ was given by Lhuilier.

A natural consequence of the Euler-Poincaré formula is that $|E(G)| \leq 3|V(G)| - 3\chi$ whenever G is a graph which can be embedding in a surface of characteristic χ without crossings. This shows, for example, that K_8 cannot be embedding on the torus without crossings, since $|E(K_8)| = \binom{8}{2} = 28$ whereas $3|V(K_8)| - 3\chi = 24$. Another consequence is that the degeneracy of a graph G embedded in $\mathbb{S}$ without crossings satisfies

$$d(G) \leq \lfloor 6 - \frac{6\chi}{|V(G)|} \rfloor.$$

So we can extend our result about planar graphs being 6-colorable to higher genus surfaces.

Theorem 5.6.2 (Heawood's Map-Coloring Theorem)
The chromatic number of a graph embedding without crossings in a surface of characteristic $\chi \leq 0$ is at most

$$h(\chi) = \lfloor \frac{1}{2}(7 + \sqrt{49 - 24\chi}) \rfloor$$

Furthermore, if G is a minimal $h(\chi)$-chromatic graph drawn on a surface of characteristic $\chi \neq -2$, then $G = K_{h(\chi)}$.

Proof $\triangleright$. First we prove the upper bound $h(\chi) \leq \lfloor \frac{1}{2}(7 + \sqrt{49 - 24\chi}) \rfloor$. Let G be embedded on a surface $\mathbb{S}$ of characteristic χ without crossings. We may assume that G is a minimal k-chromatic graph on $\mathbb{S}$, where $k = h(\chi)$. By Proposition 4.3.2 (or by just deleting a vertex of small degree) it follows that $\delta(G) \geq k-1$. Therefore $|E(G)| \geq \frac{k-1}{2}|V(G)|$, by the handshaking lemma. On the other hand $|E(G)| \leq 3|V(G)| - 3\chi$, by the Euler-Poincaré Formula. Therefore

$$\frac{k-1}{2}|V(G)| \leq 3|V(G)| - 3\chi$$

which is the same as

$$(k-7)|V(G)| + 6\chi \leq 0.$$

Since we assumed $\chi \leq 0$, $k \geq 7$ follows from the fact that K_7 can be embedded on the torus. Now $|V(G)| \geq k$ since $\delta(G) \geq k-1$, so

$$k^2 - 7k + 6\chi \leq 0.$$

This gives $2k \leq 7 \pm \sqrt{49 - 24\chi}$, and we clearly must take the positive square root. This proves the upper bound on $h(\chi)$.

The second part of the proof is to show $G = K_k$. If $G \neq K_k$ then $|V(G)| \geq k+2$, and it is an excercise to prove that if $|V(G)| = k+2$, then G consists of a pentagon disjoint from K_{k-3} together with all edges between the pentagon on the K_{k-3}. In particular, $|E(G)| = \binom{k+2}{2} - 5$. This is greater than $3(|V(G)| - \chi)$, contradicting the Euler-Poincaré Formula. Therefore $|V(G)| \geq k+3$. Now $\delta(G) \geq k-1 \geq 6$, and by Brook's Theorem, G is not $k-1$ regular, otherwise it would be $k-1$ colorable. Therefore

$$|E(G)| > \frac{|V(G)|(k-1)}{2}.$$

This gives $k^2 - 4k - 20 + 6\chi \leq 0$ and so $k \leq 2 + \sqrt{24 - 6\chi}$. This is not possible unless $\chi = -2$. □

Note that we do not include the case of plane graphs, where $h(\chi) = 4$. Surprisingly, this case also agrees with the above formula for $h(\chi)$. Also, we avoided $\chi = -2$. The full classification, for all χ and even for all surfaces (not only orientable ones) was given by Ringel and Youngs. It should be noted that better results can be obtained for a graph of girth $g > 3$: in that case Euler's Formula can be used to give

$$|E(G)| \leq \frac{g}{g-2}(|V(G)| - \chi)$$

and then one can color these graphs with fewer colors, by repeating the pattern of the proof above.

5.7 Characterization of Graph on Surfaces

We give the following quote from the book of Mohar and Thomassen:

> *"Graphs on surfaces form a natural link between discrete and continuous mathematics. They enable us to understand both graphs and surfaces better. It would be difficult to prove the celebrated classification theorem for (compact) surfaces without the use of graphs. Map color problems are usually formulated and solved as problems concerning graphs."*

Which graphs are embed in a surface of genus γ without crossings? A consequence of one of the deepest theorems in mathematics, called the Graph Minors Theorem, due to Robertson and

Seymour, is that for any γ, there exists a finite list of graphs $\mathcal{H}_\gamma$ such that any graph which does not embed on a surface of genus γ contains a graph in $\mathcal{H}_\gamma$ as a minor. Kuratowski's Theorem is such a result for planar graphs, where $\mathcal{H}_0 = \{K_5, K_{3,3}\}$. Unfortunately, for surfaces of higher genus, the known list is prohibitively long, and the smallest possible list is not known. The shortest proof that there is a finite list is due to Thomassen (it is much simpler than the graph minors theorem).

A very natural question on embedding graphs is the following: in the last section, we gave embeddings of graphs on surfaces where some of the faces were a bit unorthodox: they are not homeomorphic to plane disks. While it is true that every 2-connected planar graph has an embedding where every face is homeomorphic to a disk, it is not known if every 2-connected graph can be embedded in some surface in such a way that all faces are homeomorphic to disks. This is known as the strong embeddability conjecture, and it remains open It has many other consequences in graph theory; for example, does every 2-connected graph contain a set of cycles such that every edge is in either one or two of the cycles? For planar graphs, this is clear: the boundaries of the faces will do, and then every edge is in exactly two cycles. The conjecture also has implications for the existence of nowhere zero 5-flows. For more information on graphs on surfaces, see Archdeacon or Mohar and Thomassen.

6 Extremal Graph Theory

If F and G are graphs, then we say G is ***F-free*** if G does not contain F as a subgraph. For instance, we know that bipartite graphs are F-free whenever F is any odd cycle (Lemma 2.2.2). Let $\mathrm{ex}(n, F)$ denote the maximum number of edges that an n-vertex F-free graph can have. These are known as the ***extremal number*** or ***Turá number*** for F, and in this section we study the value of $\mathrm{ex}(n, F)$ for various F. An F-free graph on n vertices with $\mathrm{ex}(n, H)$ edges is called an ***extremal graph***.

Let's start by observing $\mathrm{ex}(n, K_2) = 0$ and $\mathrm{ex}(n, K_{1,2}) = \lfloor n/2 \rfloor$ – for the last statement, notice that a graph with no $K_{1,2}$ must be a matching, and the largest matching on n vertices has $\lfloor n/2 \rfloor$ edges. The reader should check more generally that

$$\mathrm{ex}(n, K_{1,s}) = \lfloor -1)n/2 \rfloor$$

for all $n \geq s+1$, using the handshaking lemma. For $n \geq s$, we observe $\mathrm{ex}(n, K_{1,s}) = \binom{n}{2}$. This is a general trend: if n is less than $|V(F)|$, then $\mathrm{ex}(n, F) = \binom{n}{2}$.

Example 12. Let us consider another example, where F is a matching of size two. If $n \leq 3$, then $\mathrm{ex}(n, F) = \binom{n}{2}$, but if $n \geq 4$, then we claim $\mathrm{ex}(n, F) = n - 1$. To prove such a statement, we have to give an extremal F-free graph on n vertices with $n-1$ edges, and we have to show any n-vertex F-free graph has at most $n-1$ edges. The first is easy to do: take $K_{1,n-1}$, which is clearly F-free with $n-1$ edges. Now we show $\mathrm{ex}(n, F) \leq n - 1$. If G is an n-vertex graph and $|E(G)| \geq n \geq 4$, then by Proposition 2.1.2, G is not a tree so G contains a cycle, C. If C has length at least four, then C contains F. So C has length three. Since $n \geq 4$, there is an edge e in G that is not on the cycle C. But then we pick an edge of the cycle disjoint from e, so that e and f form a copy of F in G. Therefore $F \subset G$, and so $\mathrm{ex}(n, F) \leq n - 1$.

We are usually interested in the behavior of $\mathrm{ex}(n, H)$ for large values of n. The first case of interest which we study is the case $H = K_r$.

6.1 Turán's Theorem

To go about constructing a graph on n vertices with many edges and no complete graph on $r+1$ vertices, take disjoint sets $V_1, V_2, \ldots, V_r$, where $|V_1| + |V_2| + \ldots |V_r| = n$ and join all vertices in V_i to all vertices in V_j for all $i \neq j$ and $i, j \in \{1, 2, \ldots, r\}$. This graph is called a ***complete r-partite graph*** – for $r = 2$ it is a complete bipartite graph with parts V_1 and V_2. Note that a complete r-partite graph cannot possibly contain K_{r+1}, since $\chi(K_{r+1}) = r + 1$. The number of edges an r-partite graph is

$$\sum_{i \neq j} |V_i||V_j|.$$

Since $|V_1| + |V_2| + \cdots + |V_r| = n$, this expression is maximized when all the V_is are as equal in size as possible, so $|V_i| = \lfloor n/r \rfloor$ or $|V_i| = \lceil n/r \rceil$ for all $i \in \{1, 2, \ldots, r\}$.

The Turán graph, denoted $T_r(n)$, is the unique r-partite graph all of whose parts have sizes as equal as possible. For $r = 2$, this corresponds to a complete bipartite graph with parts of size $\lfloor n/2 \rfloor$ and $\lceil n/2 \rceil$. So the number of edges in $T_2(n)$ is exactly $\lfloor n^2/4 \rfloor$. In general, we let $t_r(n)$ denote the number of edges in $T_r(n)$ – it is not a very nice number to determine, but it is roughly $(1 - \frac{1}{r})\binom{n}{2}$. Turán's Theorem states that $\mathrm{ex}(n, K_{r+1}) = t_r(n)$ when $n \geq r$. But it says even more: the only graph with $t_r(n)$ edges and no K_{r+1} is $T_r(n)$ – so $T_r(n)$ is the unique extremal graph. The inductive proof we give is fairly subtle. It relies on two facts: first that we can't add any edges to $T_r(n)$ without creating a K_{r+1}, and second, the degrees of $T_r(n)$ are as close together as possible amongst all n-vertex graphs with $t_r(n)$ edges.

Theorem 6.1.1 (Turán's Theorem) *Let G be an n-vertex graph containing no K_{r+1}, where $n \geq r$. Then $|E(G)| \leq t_r(n)$, with equality if and only if $G = T_r(n)$.*

Proof $\triangleright$. We prove the theorem by induction on n, the number of vertices in G, starting with $n = r$. The statement we prove is that if G is an n-vertex graph with no K_{r+1} and $|E(G)| \geq t_r(n)$, then $G = T_r(n)$. This proves the theorem: if $|E(G)| > t_r(n)$ then delete edges until $t_r(n)$ edges remain, but then the graph is $T_r(n)$, and we can't add any edges to $T_r(n)$ without creating a K_{r+1}. If $n = r$, then $t_r(n) = t_r(r) = \binom{r}{r}$, and clearly $|E(G)| \geq \binom{r}{2}$ implies $G = K_r = T_r(r)$, as required. Now suppose $n > r$, and let G be a graph on n vertices containing no K_{r+1} and with at least $t_r(n)$ edges. Delete edges from G until $|E(G)| = t_r(n)$. Now $T_r(n)$ is a graph with $t_r(n)$ edges with the largest minimum degree amongst all graphs with $t_r(n)$ edges. Therefore $\delta(G) \leq \delta(T_r(n))$. Now every vertex of $T_r(n)$ has degree $n - \lfloor n/r \rfloor$ or $n - \lceil n/r \rceil$, so $\delta(T_r(n)) = n - \lceil n/r \rceil$. Furthermore, if x has degree $\delta(T_r(n))$ in $T_r(n)$, then $T_r(n) - \{x\} = T_r(n-1)$. Therefore $t_r(n) - \delta(T_r(n)) = t_r(n-1)$. This shows that if v is a vertex of smallest degree in G, then

$$|E(G - \{v\})| = |E(G)| - \delta(G) \geq t_r(n) - \delta(T_r(n)) = t_r(n-1).$$

By induction, $G - \{v\} = T_r(n-1)$. This means that v has degree exactly $\delta(T_r(n)) = n - \rceil n/r \rceil$. Now since G has no K_{r+1}, v is joined to vertices in $r-1$ parts of $T_r(n-1)$. But since v has degree exactly $n - \lceil n/r \rceil$, v must be joined to the $r-1$ smallest parts of $T_r(n-1)$, which means that $G = T_r(n)$, as required. $\square$

6.2 Erdős-Simonovits-Stone Theorem

If F is any graph of chromatic number $r \geq 3$, then a complete $(r-1)$-partite graph cannot contain F. We deduce that for all F with chromatic number r,

$$\mathrm{ex}(n, F) \geq \mathrm{ex}(n, K_r).$$

The Erdős-Simonovits-Stone Theorem gives a nearly matching upper bound for $\mathrm{ex}(n, F)$.

Theorem 6.2.1 (Erdős-Simonovits-Stone Theorem) *Let F be a graph of chromatic number $r \geq 2$. Then*

$$\lim_{n\to\infty} \frac{\mathrm{ex}(n, H)}{\mathrm{ex}(n, K_r)} = 1.$$

The proof is beyond the scope of the course, but we do give one example of finding the extremal number for a non-complete graph of chromatic number three:

Proposition 6.2.2 *Let G be a graph with $\lfloor n^2/4 \rfloor + 2$ edges and $n \geq 5$ vertices. Then G contains two triangles intersecting in one vertex (a bowtie).*

Proof $\triangleright$. Proceed by induction on n, starting with $n = 5, 6, 7$. We omit the verification that the proposition holds for $n = 5, 6, 7$, which is rather technical. Now suppose $n > 7$ and let G be an n-vertex graph with at least $\lfloor n^2/4 \rfloor + 2$ edges. By deleting some edges from G, we can assume that G has exactly $\lfloor n^2/4 \rfloor + 2$ edges. Now by the handshaking lemma:

$$\frac{1}{2}\delta(G)n \leq \left\lfloor \frac{n^2}{4} \right\rfloor + 2.$$

This means $\delta(G) \leq \frac{2}{n}\lfloor \frac{n^2}{4} \rfloor + \frac{4}{n}$. If n is even, say $n = 2m$ where $m > 3$, then

$$\delta(G) \leq m + \frac{2}{m}$$

and since $m > 3$ and $\delta(G)$ is an integer, we have $\delta(G) \leq \lfloor n/2 \rfloor$. Now suppose $n = 2m + 1$ where $m > 3$. Then

$$\delta(G) \leq \frac{2m^2 + m}{2m+1} + \frac{m+4}{2m+1} = m + \frac{m+4}{2m+1}$$

so again $\delta(G) \leq \lfloor n/2 \rfloor$ (note that $(m + 4)/(2m + 1) < 1$ if $m > 3$). Now let v be a vertex of degree $\delta(G)$ in G. Then

$$|E(G - \{v\})| \geq \left\lfloor \frac{n^2}{4} \right\rfloor + 2 - \left\lfloor \frac{n}{2} \right\rfloor = \left\lfloor \frac{(n-1)^2}{4} \right\rfloor + 2.$$

By induction, $G - \{v\}$ contains a bowtie, so G contains a bowtie. $\square$

As an exercise, the reader should construct an n-vertex graph with $\lfloor n^2/4 \rfloor + 1$ edges and no bowtie (an extremal graph), for all $n \geq 5$. As another exercise, try to determine $\mathrm{ex}(n, H)$ for the graph H shown below: $\triangleleft\triangleleft$

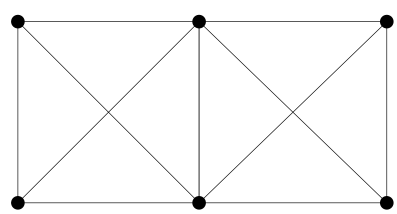

Figure 46: Two K_4s joined along an edge

In general it is hard to work out $\mathrm{ex}(n, F)$ precisely, but the above proof method works when F has chromatic number three or more. the main difficulty lies in checking the base cases; for small graphs, such as two triangles sharing an edge, establishing the base case is feasible.

6.3 Bipartite Graphs

The Erdős-Simonovits-Stone Theorem tells us asymptotically that $\mathrm{ex}(n, K_r)$ and $\mathrm{ex}(n, F)$ are the same, provided F has chromatic number at least three. This leaves the case of chromatic number two – bipartite graphs. Unfortunately, $\mathrm{ex}(n, F)$ is a mystery for a general bipartite graph F, so we restrict our attention to $F = K_{r,s}$, the complete bipartite graph with parts of sizes r and s. Before we prove the main theorem, we state a numerical lemma:

Lemma 6.3.1 *Let $a_1, a_2, \ldots, a_n$ and r be positive integers and let $a = \sum_{i=1}^n a_i$. Then*

$$\sum_{i=1}^{n} \binom{a_i}{r} \le n\binom{a/n}{r}.$$

Theorem 6.3.2 (Kövari-Sós-Turán Theorem) *Let r, s be positive integers, and suppose $r \le s$. Then*

$$\mathrm{ex}(n, K_{r,s}) \le cn^{2-1/r}$$

where c is a constant depending only on r and s.

Proof $\triangleright$. Let G be an n-vertex graph not containing $K_{r,s}$. The number of sets of r vertices of G is exactly $\binom{n}{r}$. Now for a vertex $v \in V(G)$ of degree $d(v)$, there are exactly $\binom{d(v)}{r}$ sets of size r in the neighbourhood of v. So the total number of sets of size r which are in the neighbourhood of a vertex is

$$\sum_{v \in V(G)} \binom{d(v)}{r}.$$

Note that we might have counted some sets of size r more than once. But what we do know is that no set of size r could have been counted s times, otherwise that set of size r would be in the neighbourhood of s vertices in $V(G)$, and that would give a $K_{r,s}$ in G. So

$$\sum_{v \in V(G)} \binom{d(v)}{r} \le (s-1)\binom{n}{r}.$$

Applying Lemma 6.3.1,

$$n\binom{2|E(G)|/n}{r} \le (s-1)\binom{n}{r}.$$

After a calculation, we get $|E(G)| \le cn^{2-1/r}$. $\square$

6.4 Constructions of extremal graphs

One of the main difficulties in determining the Turán numbers $\mathrm{ex}(n, F)$ is in giving a construction of n-vertex F-free graphs with many edges. This was not so hard for $F = K_r$, but when F is bipartite, this problem is notoriously difficult. In this section, we consider the case that F is a 4-cycle. We prove the following.

Theorem 6.4.1 *There exists an infinite number of positive integers n such that*

$$\mathrm{ex}(n, C_4) \geq \frac{1}{2}(n^{3/2} - n - \sqrt{n} + 1).$$

Proof $\triangleright$ Let p be an odd prime number, and define a graph G_p with vertex set

$$V = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} : 0 \leq x < p, 0 \leq y < p\}.$$

For a positive integer m, let $(m)_p$ denote the remainder when m is divided by p. Let $A = \{(z, (z^2)_p : 1 \leq z < p\}$. The edge set E of G_p will be the set of pairs $\{(x, y), (x', y')\}$ such that $(x + x')_p \in A$ and $(y + y')_p \in A$. For instance, if $p = 3$, then

$$V = \{(0, 0), (0, 1), (1, 0), (1, 1), (2, 1), (2, 2), (1, 2), (0, 2), (2, 0)\}$$

and $A = \{(1, 1), (2, 1)\}$, and the graph G_3 is drawn below:

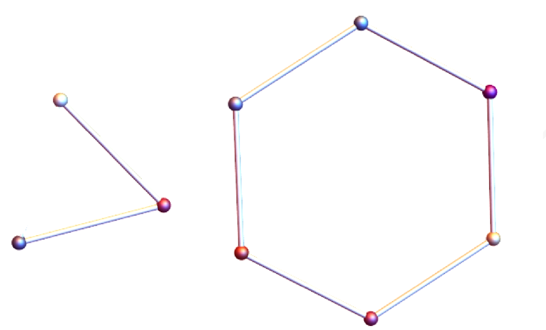

Figure 47: Graph G_3

The reader should label the vertices with the elements of V. We are now going to check that G_p does not contain a 4-cycle. Suppose, for a contradiction, there is a 4-cycle in G_p, with edges

$$\{(a, b), (c, d)\}, \{(c, d), (e, f)\}, \{(e, f), (g, h)\}, \{(g, h), (a, b)\}.$$

Let us write $m \equiv n$ if the remainder when m is divided by p equals the remainder when n is divided by p. One of the basic facts we use is that if α is not a multiple of p, then $\alpha x \equiv \alpha y$ implies $x = y$ – in other words we can divide by non-zero elements. By definition, there are elements $x, y, z, w \in \{1, 2, \ldots, p - 1\}$ such that

$$a + c \equiv x \quad b + d \equiv x^2 \tag{1}$$
$$c + e \equiv y \quad d + f \equiv y^2 \tag{2}$$
$$e + g \equiv z \quad f + h \equiv z^2 \tag{3}$$
$$g + a \equiv w \quad h + b \equiv w^2. \tag{4}$$

If we subtract the (1) from (2) we get $e - a \equiv y - x$ and $f - b \equiv y^2 - x^2$, and if we subtract (3) from (4) we get $a - e \equiv w - z$ and $b - f \equiv w^2 - z^2$. Therefore

$$y - x \equiv z - w \tag{5}$$
$$y^2 - x^2 \equiv z^2 - w^2. \tag{6}$$

If $y = x$ then from (1) and (2), $(a, b) = (e, f)$. This is a contradiction. So $y \neq x$, and since $y^2 - x^2 \equiv (y - x)(y + x)$, we get $y + x \equiv z + w$ from (5) and (6). Adding this to (5), we get $2y \equiv 2z$ which implies $y = z$. But then from (2) and (3), $(c, g) = (e, h)$, a contradiction. Therefore G_p does not contain a 4-cycle.

Now we determine the number of edges in G_p. Clearly G_p has $n = p^2$ vertices. Each $(x, y) \in V$ is adjacent to all vertex (a, b) such that $x + a \equiv z$ and $y + b \equiv z^2$. However, (x, y) is joined to (x, y) if $x + x \equiv z$ and $y + y \equiv z^2$, and this happens exactly when $2x \equiv z$ and $2y \equiv z^2$. That means there are $p - 1$ vertices (x, y) with only $p - 2$ neighbors, whereas all other vertices have $p - 1$ neighbors. Therefore, by the handshaking lemma,

$$|E(G_p)| = \frac{1}{2}((p^2 - p + 1)(p - 1) + (p - 1)(p - 2)) = \frac{1}{2}(p^3 - p^2 - p + 1).$$

Therefore with $n = p^2$, $\mathrm{ex}(n, C_4) \geq \frac{1}{2}(n^{3/2} - n - \sqrt{n} + 1)$. This proves the theorem. $\square$

Two illustrations of G_5 are shown in the figures below:

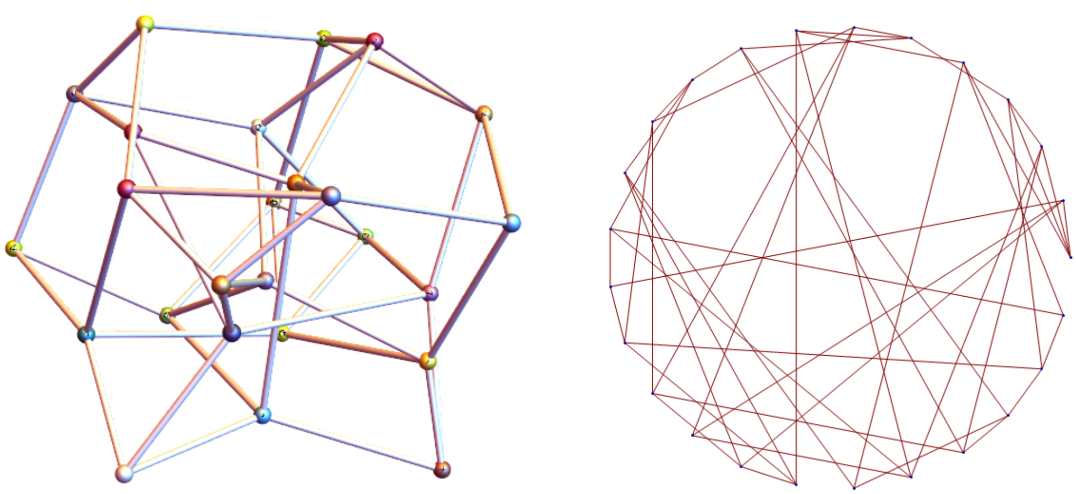

Figure 48: Graph G_5

Theorem 6.4.1 compares well with the bound given for $r = s = 2$ in Theorem 6.3.2: the reader will check $\mathrm{ex}(n, C_4) \leq (n/4)(1 + \sqrt{4n - 3})$ from that proof. Using the distribution of primes, these two theorems can be combined to obtain

$$\lim_{n \to \infty} \frac{\mathrm{ex}(n, C_4)}{n^{3/2}} = \frac{1}{2}.$$