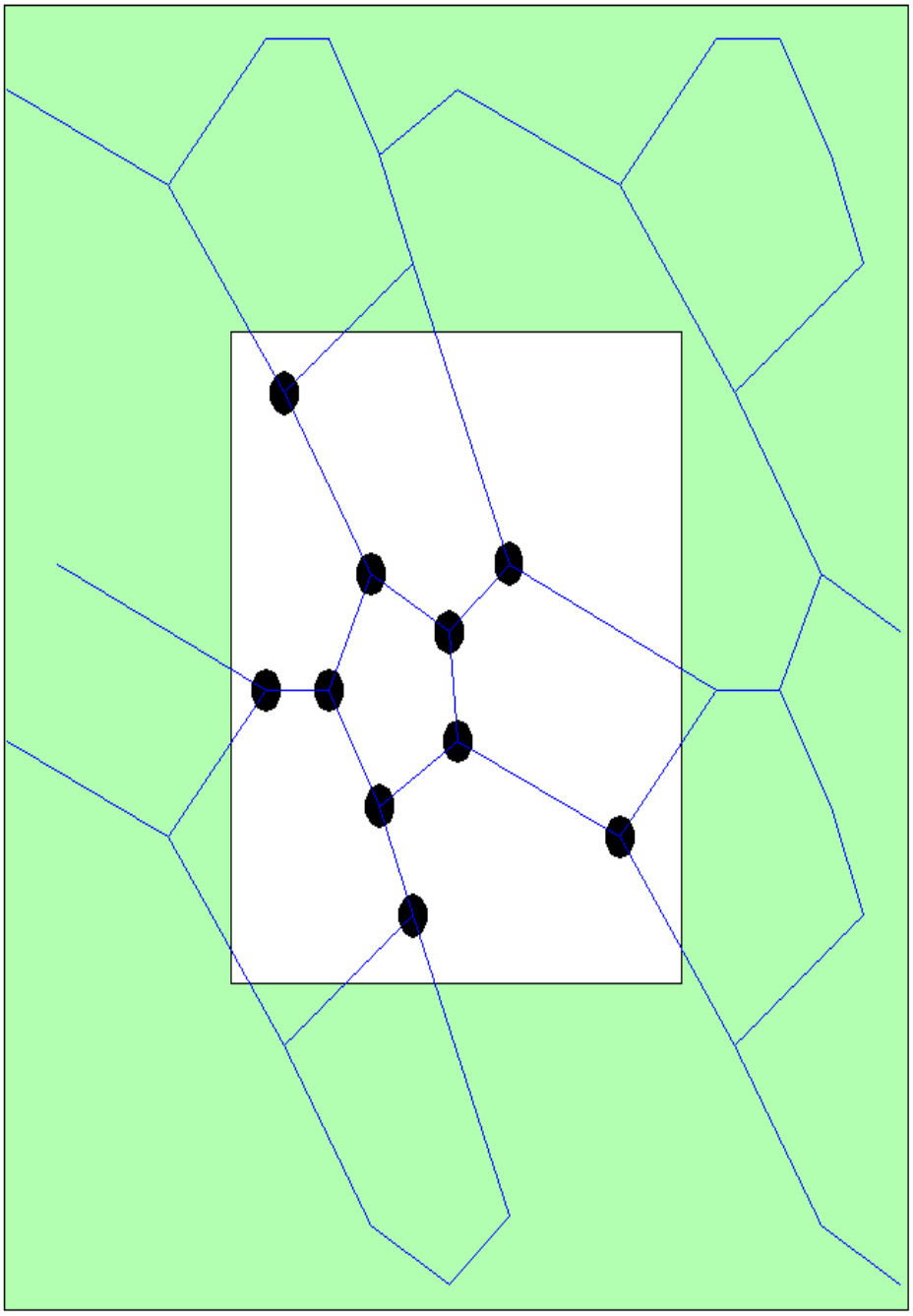
Math 415 Graph Theory



Cape Breton University Math415 Notes

James Preen

September 20, 2007

Contents

Some recommended books: Introduction to Graph Theory, R.J.Wilson
Graph Theory, Reinhard Diestel, Graph Theory and its Applications, Bondy & Murty

1 Introduction to Graph Theory

A *graph* is defined as a (non-empty) set of *vertices*, some of which are joined by *edges*. Mathematically, the graph G in figure 1 can be described by

$$V(G) := \{a, b, c, d, e, f, g, h\} \text{ and } E(G) := \{ad, af, bc, ce, df, dg, dh, fg, gh\}.$$

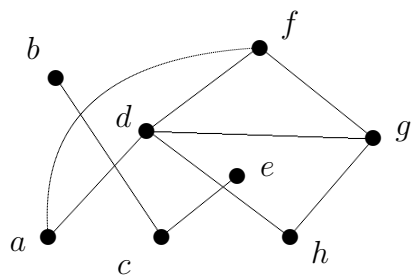


Figure 1: A graph with 8 vertices and 9 edges

Note that this set of letters and this particular way of drawing the graph is not a unique one. What we are interested in is the way they are arranged with respect to each other, not how they are placed or named on the page. Vertex a could be a person, a computer or yoghurt. An edge could indicate whether two people know each other, whether the computers talk the same language or if that object is one of a person's favourite foods. Our aim in this class is to see what we can say about graphs, whatever they are about. Similarly, it is not significant that edge af is curved, we just drew it like that to make it clearer. Where edge ce crosses edge dh there is not a vertex, as otherwise we would have marked it by a black circle and a letter.

We say that two vertices are *adjacent* if there is an edge joining them and a vertex is *incident* with an edge if it is one of the edge's two vertices. Two edges are incident if they have a vertex in common. We can also classify vertices by their valencies: the standard way of doing this for a vertex v is to count the number of edges *incident* with v, a quantity called the *valency* of v, $\rho(v)$. For instance, in figure 1, we have $\rho(a) = 2$, $\rho(b) = 1$, $\rho(c) = 2$, $\rho(d) = 4$, $\rho(e) = 1$, $\rho(f) = 3$, $\rho(g) = 3$ and $\rho(h) = 2$ and the *valency sequence* of G is (4, 3, 3, 2, 2, 2, 1, 1) [we normally order from high to low, but there is nothing significant in the order]. We have special names for vertices with small valencies: if $\rho(v) = 0$ then v is an *isolated* vertex, if $\rho(v) = 1$ then v is a *terminal* vertex and otherwise v is an *internal* vertex. If, $\forall v \in V(G)$, $\rho(v) = r$ then we say that G is *r-regular*. The *total valency* of a graph G is $\sum_{v \in V(G)} \rho(v)$.

Exercise 1 *Add/remove edges from G (from figure 1) to make it respectively, 2 regular, 3 regular and 4 regular.*

Lemma 1 *The number of vertices of odd valency in a graph G is even.*

Proof: In counting the total valency of a graph G each edge is counted twice; once for each end of each edge. Hence this figure is even and so, since the sum of the even valencies is even, so must be the sum of the odd valencies. $\diamond$

Theorem 1.1 *In any graph G there is always at least one pair of vertices of the same valency.*

Proof: Let $V(G) := \{v_1, v_2, \ldots v_p\}$ so that $|V(G)| = p$. As each vertex has at most $p-1$ choices for neighbours, $\rho(v_i) \leq p-1$ for $1 \leq i \leq p$. Thus there are p different possible valencies (0, 1,..., $p-1$) for the p vertices in G. However, if $\rho(v) = 0$ for some $v \in G$ then there cannot be a vertex $w \in G$ for which $\rho(w) = p-1$ and a similar argument applies in reverse. Thus there are just $p-1$ different valencies for the p vertices in G and thus the result is proved by the pigeon-hole principle. $\diamond$

1.1 The smallest graphs.

To get the feel of graphs we shall look at the smallest, in terms of numbers of vertices. There is, of course, just one graph on one vertex, and it has no edges since in a (simple) graph there cannot be loops. Similarly, with two vertices there are just two graphs, either with an edge between them or not.

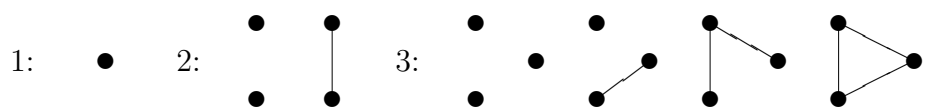


Figure 2: The smallest graphs

For three vertices some interesting features of graphs start to emerge: there are actually eight ways to assign edges to three vertices but four of these are redundant since the resulting graph in those cases "looks the same" as one of the others. This concept is known as *isomorphism* and, formally, two graphs G and H are *isomorphic* ($G \cong H$) if there exists a pair of bijections $\theta : V(G) \to V(H)$ and $\psi : E(G) \to E(H)$. Note that this implies that, for instance, the set of valencies of the vertices of G is the same as that of H, but this condition is not enough as is shown by the two graphs in figure 3, which both have valency sequence (3, 2, 2, 2, 1).

Figure 3: Two non-isomorphic graphs with the same valency sequence

1.2 Isomorphism

In general it is a difficult thing to tell whether two graphs are isomorphic, and it is largely a matter of practicing until you get good at it. However, there are several things you can do initially which will prove that no isomorphism exists between two graphs:

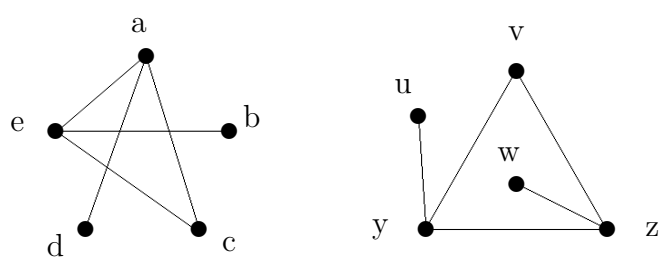
1. count the number of vertices and edges in each graph
2. determine both the valency sequences
3. check other easily found parameters for each graph and compare them

In each case, if any observation which is independent of how the graph is drawn differs between two graphs then the two graphs *cannot* be isomorphic.

More advanced techniques follow similar ideas but can involve groups of vertices, edges, or combinations of the two. For instance in figure 3 we can see that the graphs are different by considering the (lone) vertex of valency three: in the left hand graph it is adjacent to the three vertices of valency two, whereas this is not the case in the right hand one. Hence the two graphs are different.

Exercise 2 *Show that the two graphs in figure 3 are not isomorphic by considering the vertices of valency 2.*

To show that two graphs are isomorphic it is required to give an actual mapping between the vertices of the two graphs and verify that the edge set are preserved, as shown in figure 4 for the two graphs. Note that the mapping is not a unique one (b and d and e and a could be swapped) but that does not matter. What is important is that a mapping exists in which the edge sets correspond as shown.



a→ y	ae, ad, ac	yz, yu, yv
b→w	be	wz
c→ v	ca, ce	vy, vz
d→ u	da	uy
e→ z	eb, ea, ec	zw, zy, zv

Figure 4: Two Isomorphic Graphs and an Isomorphic Mapping

1.3 Valency Sequences

We say that a valency sequence is *graphical* if there exists a graph which has that sequence. The following theorem can be used to determine whether or not a sequence is graphical:

Theorem 1.2 *The valency sequence $D = (d_1, d_2, \ldots, d_p)$ with $p - 1 \geq d_1 \geq d_2 \geq \ldots \geq d_p$ is a graphical valency sequence of and only if the modified sequence $D' = (d_2 - 1, d_3 - 1, \ldots d_{d_1+1} - 1, d_{d_1+2}, \ldots, d_p)$ is a graphical valency sequence.*

Proof: If D' is a graphical sequence, then surely D is, since we can just add a vertex to a graph G' with D' as its sequence and add d_1 edges from it to the first d_1 vertices in G'.

Let G be a graph with valency sequence D. If a vertex v of valency d_1 is adjacent to vertices of valency d_k for $k = 2, \ldots, d_1 + 1$ then we can just remove v from G to give a graph G' with valency sequence D'.

So suppose there is no such vertex in G. We will show that from G we can always form another graph with valency sequence D which does have such a vertex, and so we can proceed as above. Let us suppose that the vertices are labelled v_i with $\rho(v_i) = d_i$ and that v_1 has valency d_1 and the sum of the valencies of its neighbours is maximal.

Since v_1 is not adjacent to the next d_1 vertices there must be two vertices v_i and v_j with $d_i > d_j$ such that v_1v_j is an edge but v_1v_i isn't. Since $d_i > d_j$ there must be a vertex which isn't adjacent to v_j but is to v_i, v_k suppose. So we can then remove the two edges v_1v_j and v_iv_k and add v_1v_i and v_jv_k, which gives a different graph which still has valency sequence D. But in this graph the sum of the valencies of the neighbours of v_1 is greater than before, contradicting the supposition of its maximality. Thus we can repeat this edge switching operation (a finite number of times) until we get a graph with vertex v_1 with the desired property. $\diamond$

To use theorem 1.2 we proceed as follows:

1. Consider the valency sequence; if it is easily seen to be graphical (all zeroes or an even number of ones etc.) or can be seen to be non-graphical (includes a negative number, sum of valencies is odd etc.) then stop.

2. Take a valency in the list (usually the largest) with value ρ, say, and remove it from the sequence and subtract one from the next ρ other valencies to make a new valency sequence

3. return to step 1

Note, though, that if you start with a sequence with an even sum then at each step you will have an even sum since we are subtracting a total of 2ρ from the sequence sum. If you get an odd sum halfway through then you know you have made an error.

Exercise 3 *Test the following sequences using theorem 1.2 to see if they are graphical or not: (if they are graphical draw an exemplar graph having that valency sequence)*

1. $(6, 5, 5, 4, 3, 3, 2, 2)$

2. $(6, 6, 4, 4, 3, 3, 2, 2)$

3. $(8, 4, 4, 4, 3, 3, 2, 2)$

4. $(5, 5, 5, 4, 3, 3, 2, 2)$

Note that this procedure (if it works) finds only one graph with the valency sequence, in order to find all graphs with the same valency sequence we have to be a little more rigourous as shown in this example:

Given the sequence $(3, 2, 2, 2, 2, 1)$ we remove the 3, and the possible sequences left are (after re-arrangement) $(2, 1, 1, 1, 1)$ or $(2, 2, 1, 1, 0)$. Repeating this procedure for these two

sequences, or by observation, we see that these are, uniquely, the graphs G_1 and G_2 in figure 5. To get a graph with the sequence S from G_1 we have to add the three edges to three vertices of valency 1, which can be done in the two ways shown as G_3 and G_4. For G_2 the edges must go to both 1s and the 0 and so that graph must be G_5. Hence there are *e*xactly three graphs with valency sequence S.

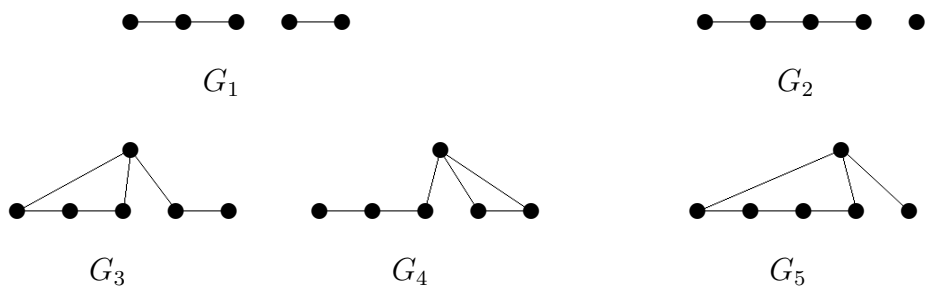


Figure 5: Finding the graphs with valency sequence (3,2,2,2,2,1)

1.4 Basic Families of Graphs

We are now in place to define several families of graphs which we will be seeing many times throughout this course. To start with we define the *complete* graph (on n vertices) K_n. This has an edge between every pair of vertices and thus is unique for each value of n. At the other end of the spectrum we define $\overline{K_n}$ to be the *null* graph on n vertices (notation to be explored further in section 1.6). Similarly we define C_n to be the circuit on n vertices; we take n vertices and add a sequence of n edges to form a cycle. Similarly P_n is the path on n vertices (and can be formed from C_n by deleting one edge from it). Finally, the *wheel* graph W_n is formed from C_{n-1} by adding a vertex (normally drawn in the middle) with "spokes" to all other vertices of the cycle.

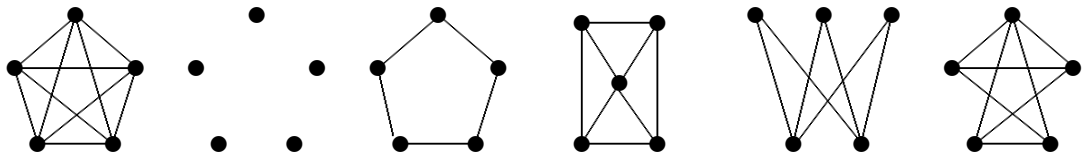
Exercise 4 *For which n is K_n a cycle? a path? neither? a wheel?*

You may also have noticed in figure 2 earlier that the triangle graph (C_3, K_3) is the first which we have seen whose vertices cannot be split into two sets, A and B say, such that no vertex in one set is adjacent to another in that set. A graph whose vertices can be divided into two such sets is called a *bipartite* graph. We define $K_{i,j}$ as the *complete bipartite* graph with two sets of size m and n all of which are joined to those in the other set. We can similarly define *complete tripartite* graphs $K_{i,j,k}$ (three sets of sizes i, j and k in which every vertex is adjacent to every vertex not in its set), and so on...

Exercise 5 *Which K_n, P_n, C_n are bipartite? tripartite? quadripartite? k-partite?*

The numbers of (non-isomorphic) graphs on a particular number of vertices actually grows very quickly: for four vertices there are 11, for five vertices there are 34, six there are 156, seven 1044 and eight 12346 (I used a computer to calculate these ;)

Note that K_n (the largest graph on n vertices) has $\frac{n(n-1)}{2}$ edges since each vertex has the maximum valency $n-1$ and so the total valency of such a graph is $n(n-1)$. Hence the number of edges is half that. This is therefore the maximum number of edges in a graph with n vertices.



K_5 $\quad$ $\overline{K_5}$ $\quad$ C_5 $\quad$ W_5 $\quad$ $K_{2,3}$ $\quad$ $K_{2,2,1}$

Figure 6: Some graphs with five vertices

1.5 Operations on Graphs

Let G be a graph and S a set of vertices and/or edges. If S contains only edges and vertices of G then $G - S$ is formed by deleting, from G, all edges in S and all vertices in S, together with all the edges incident with the vertices in S (since if those edges weren't removed they wouldn't have two end vertices). If all edges in S have both of their vertices also in S then S is a graph and we write $S \subseteq G$ and say S is a *subgraph* of G.

More specifically, if a and b are vertices in G and ab is an edge between them then $G - a := G - \{a\}$ and $G - ab := G - \{ab\}$, but note that $G - \{a, b\}$ is not necessarily the same as $G - \{ab\}$. If S is simply a subset of the vertices of G then $G - S$ is denoted by $G[S']$, where $S' := G \setminus S$, and is the subgraph *induced* by S'.

$G \cup S$ is formed by adding, to G, the edges and vertices in S. $G \cup a := G \cup \{a\}$ and $G \cup ab := G \cup \{ab\}$, assuming that a and b are vertices in G, since otherwise we would have an edge without two end vertices. Let G_1 and G_2 be subgraphs of G: G_1 and G_2 are *disjoint* if they have no vertices or edges in common, and *edge-disjoint* if they have no edges in common. We (ab)use the previous notation to define $G_1 \cup G_2$ as the subgraph of G whose vertices are in either G_1 or G_2, but again do not create multiple edges if G_1 and G_2 are not edge-disjoint. Similarly, we define $G_1 \cap G_2$ as the subgraph of G whose vertices and edges are in both G_1 and G_2.

Exercise 6 *Which of the graphs in figure 6 are subgraphs of other graphs in the figure?*

With regard to combining two graphs there are several possible useful ways to do this as described below and shown in figure 7:

- Union: $G_1 \cup G_2$ has vertex set $V_1 \cup V_2$ and its edge set includes all the edges in G_1 and G_2, and no others.
- Join: $G_1 + G_2$ is simply formed from $G_1 \cup G_2$ by adding an edge from every vertex of G_1 to every vertex in G_2 (and hence vice-versa).

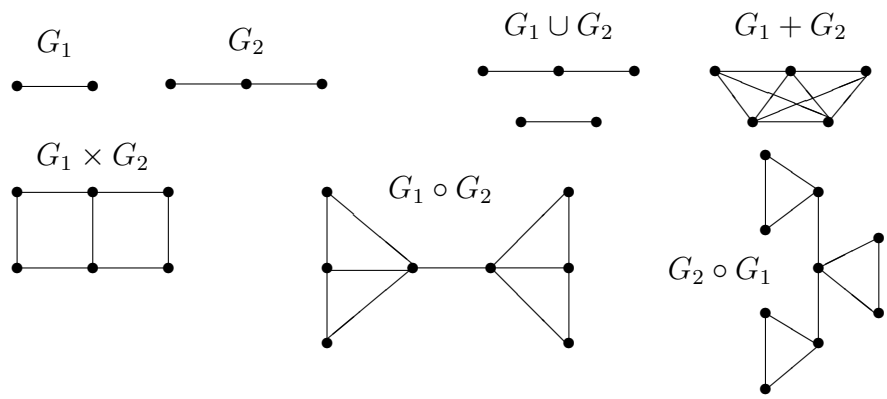


Figure 7: Examples of union, join, product and corona

- Product: $G_1 \times G_2$ has vertex set $\{v_1, v_2 : v_1 \in V_1 \text{ and } v_2 \in V_2\}$. If $u = \{u_1, u_2\}$ and $v = \{v_1, v_2\}$ are vertices of $G_1 \times G_2$ then u and v are adjacent if $u_1 = v_1$ and $u_2v_2 \in E(G_2)$ or $u_2 = v_2$ and $u_1v_1 \in E(G_1)$.
- Corona: $G_1 \circ G_2$ is obtained by taking one copy of G_1 with n_1 vertices and n_1 copies of G_2 and joining the ith vertex of G_1 to every vertex in the ith copy of G_2. Note that $G_1 \circ G_2$ is not necessary the same graph as $G_2 \circ G_1$.

1.6 Complements

Recall that the smallest graph on n vertices has no edges. This is the null graph on n vertices and was denoted by $\overline{K_n}$. This notation is chosen because it is the *complement* of K_n. In general, for a graph G:

$$\begin{aligned} V(\overline{G}) &:= V(G) \\ E(\overline{G}) &:= \{ab : a, b \in V(G), ab \notin E(G)\} \end{aligned}$$

Thus in $\overline{G}$ the valency of v will be related to the valency in G using $(n-1) - \rho(v)$. The complement is a useful operation in several ways; for instance the number of graphs with n vertices and m edges can be easily proven to be the same as the number with n vertices and $\frac{n(n-1)}{2} - m$ edges. Note also that the complement of a regular graph is also regular.

Exercise 7 *What is the complement of $K_{i,j}$?*

Thus the complement of G_1 in figure 8 is G_2 (and vice versa, of course). However, the graph H has, as its complement, itself. Such a graph is called *self-complementary*.

This is thus the only self-complementary graph on four vertices since any such graph must have three edges and those in figure 8 can be shown to be the only graphs on four vertices with three edges.

Theorem 1.3 *If G is a self-complementary graph then $|V(G)| \equiv 0, 1 \bmod 4$.*

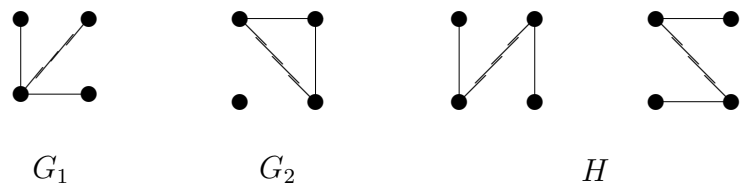


Figure 8: Small Complementary Graphs

Proof: Suppose $|V(G)| = n$. Then, if G is a self-complementary graph, since K_n has $\frac{n(n-1)}{2}$ edges, G must have $\frac{n(n-1)}{4}$ edges, and this is integral if and only if n is congruent to zero or one modulo four. $\diamond$

Exercise 8 *Show that a self-complementary graph cannot have a vertex connected to either all or none of the other vertices.*

In investigating self-complementary graphs it is necessary to consider which vertices will map to which others in the complement. The basic fact for any self-complementary graph is that it must have a "balanced" valency sequence (so that if you reverse the sequence and add it to the original you get (n-1, n-1, . .,.n-1). For instance, figure 9 shows a self-complementary graph on nine vertices.

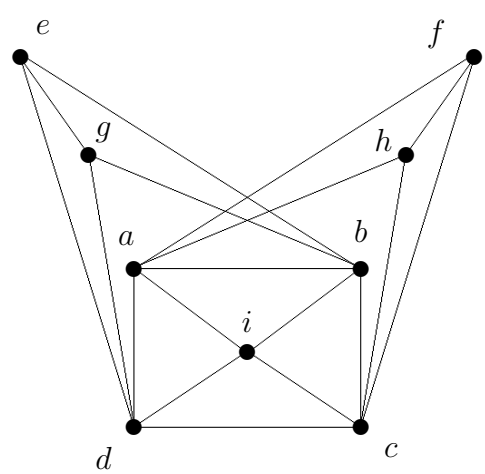


Figure 9: A Self-Complementary Graph on Nine Vertices

It has valency sequence (5, 5, 5, 5, 4, 3, 3, 3, 3) and so, in the isomorphism between G and $\overline{G}$, the vertex of valency 4 must map to itself and each vertex of valency 5 must map to one of valency 3 and vice-versa. In particular, the subgraph induced by the vertices of valency 5 in G must be the complement of the subgraph of G induced by the vertices of valency 3. Even with these conditions, it is still hard to find the right way to join the vertices of valencies 5 and 3 together, but it is possible, in at least one way. One suitable isomorphism is shown in table 1.

Vertex in G	Vertex in $\overline{G}$	Neighbours in G	Neighbours in $\overline{G}$
a	e	b,d,f,h,i	a,c,f,h,i
b	f	a,c,e,g,i	b,d,e,g,i
c	g	b,d,f,h,i	a,c,f,h,i
d	h	a,c,e,g,i	b,d,e,g,i
e	d	b,d,g	b,f,h
f	a	a,c,h	c,e,g
g	b	b,d,e	d,f,h
h	c	a,c,f	a,e,g
i	i	a,b,c,d	e,f,g,h

Table 1: An isomorphism for the self-complementary graph above

1.7 The Reconstruction Conjecture

This is one of the most tantalizing conjectures in modern graph theory. Posed just fifty years ago by Ulam it has been proven for many classes of graphs, including both regular graphs and trees (which we shall soon be meeting). It can be posed in several different ways, but the following is the one from which the name arises: If G is a graph with vertex set $\{v_1, \ldots, v_n\}$ then the *deck* of G, $D(G)$, is the set of n graphs $\{G - v_i : 1 \leq i \leq n\}$. A *legitimate deck* is a deck which can be obtained from some graph G.

Conjecture 1.4 *Given a legitimate deck it is possible to reconstruct G if $n \geq 3$.*

Various properties of G can be proven to be *reconstructable*, such as the number of vertices (just count the number of cards). To find the number of edges we can use a special case of a result known as Kelly's lemma:

Lemma 2 *If $H \subset G$ and H has less vertices than G then the number of different copies of H in G is equal to the number of copies of H in $D(G)$ divided by $|V(G)| - |V(H)|$.*

Corollary 1.5

$$|E(G)| = \frac{\sum_{X \in D(G)} |E(X)|}{|V(G)| - 2}$$

Exercise 9 *Which graph has the deck in figure 10?*

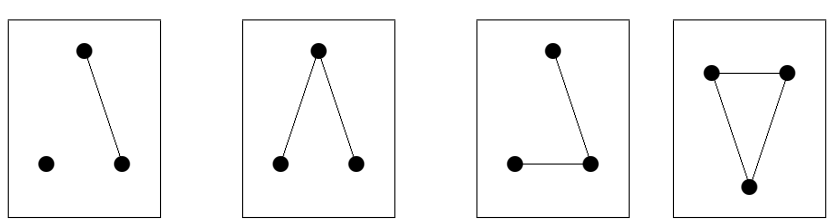


Figure 10: The deck of a graph

2 Walks, paths and trails.

2.1 Introduction

A *walk* W of length l ($l \in \mathbb{N}$) is a sequence of $l+1$ vertices and l edges

$$a_0, e_1, a_1, e_2, \ldots, a_{n-1}, e_n, a_n,$$

where e_i is an edge joining a_{i-1} and a_i ($1 \leq i \leq n$). W *connects* a_0 to a_n and is designated a *pseudopath* if no edge occurs more than once in it. If, in addition, no vertices are repeated in W then it is a *path*. If $n \geq 1$ and $a_0 = a_n$ then W is a *closed walk*, a *pseudocircuit* if no edge is repeated and a *circuit* if no vertices from a_1 to a_n are repeated.

Two vertices in a graph G are *connected* if there exists some walk (or equivalently a path - exercise?) between them. Clearly connectedness is an equivalence relation ($v \sim v$, $v \sim w \Rightarrow w \sim v$ and $v \sim w$ and $w \sim u \Rightarrow v \sim u$) and so the equivalence classes formed (including the edges between these vertices) we shall call the *components* of G. If G has just one component then we say that it is *connected*.

The *distance* between two vertices u and v in a component of G is defined as the length of a shortest walk from u to v and is written $d(u, v)$. Such a walk is called a $u - v$ *geodesic*. The distance between two vertices in different components is customarily defined as ∞. We define the *diameter* of G as the length of any longest geodesic. Similarly, the *girth* of a graph G is the length of a shortest circuit in G and the *circumference* is the length of a longest circuit. Both terms are defined as ∞ if G contains no circuits.

2.2 Bipartite Graphs

We now prove a theorem providing an alternate definition of a bipartite graph:

Theorem 2.1 *A graph G is bipartite if and only if it contains no closed walk of odd length.*

Proof: Suppose that G is bipartite with partite sets A and B: any closed walk W in G must start in one set, say A, and then go to B, and then back to A, and so on. Thus, since W is closed it must end in the set in which it started, and this involves transversing an even number of edges. Thus W cannot have odd length.

Now suppose that G contains no closed walk of odd length. We use induction on the number of edges in G, the inductive statement being P_k : if G_k is a graph with k edges with no closed walk then G_k is bipartite. It is easy to see that P_0, P_1 and P_2 are true since G_k in these cases is always bipartite and there are no closed walks of of odd length.

So let G_m be a graph with $m(\geq 3)$ edges, and no closed odd walk and suppose that P_{m-1} holds. We choose any edge vw of G_m and consider $G_{m-1} := G_m - uv$. G_{m-1} has $m-1$ edges and, since the deletion of an edge cannot create any new walks, no closed odd walk. Hence, by the induction hypothesis, G_{m-1} is bipartite and has partite sets A and B, say.

Case i). $v \in A$ and $w \in B$, say. We can add the edge vw to G_{m-1} and we still have a bipartite decomposition, and so P_m is true.

Case ii). $v, w \in A$, say. We note that there cannot exist a walk W between v and w in G_{m-1} since W would be of even length as G_{m-1} is bipartite. This would imply

that, in G_m, there would be a closed path of odd length, using vw and W, contrary to our supposition. Thus v and w are in different components of G_{m-1} and so we label the set of vertices in A which are in the same component as v as A'. The remaining vertices in this component we group into a set B'. The vertices in both of these sets are all non-adjacent as G_{m-1} is bipartite. Thus we can divide the vertices of G into the two sets $A'' := A \cup B'$ and $B'' := B \cup A'$, and these give us a situation as in case i) since $v \in B''$ and $w \in A''$. $\diamond$

2.3 Trees and Connectedness

A *cut-vertex* (respectively *cut-edge*) of a component C is a vertex (edge) in C whose removal disconnects C. For instance, in figure 11, v, w and x are all cut vertices and e and f are cut edges.

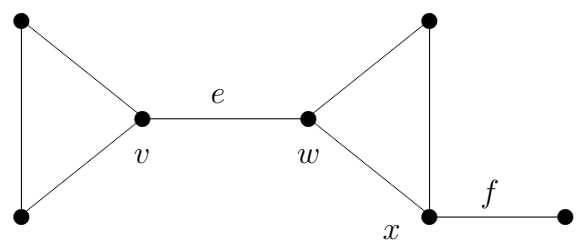


Figure 11: Cut Vertices and Edges

A graph G is *k-connected* ($k \geq 1$) if $|V(G)| \geq k+1$ and G is connected and cannot be disconnected by the removal of any set of $k-1$ (or fewer) vertices. We note that, outside this definition, disconnected is equivalent to 0-connected and K_n is better than n-connected since it can never be disconnected.

A connected graph with no circuits is called a *tree*. A graph in which all the components are trees is called, originally enough, a *forest*.

Theorem 2.2 *A (finite) graph with no isolated or terminal vertices has a circuit, and so is not a tree.*

Proof: Suppose G is such a graph. Let P be a path in G of maximal possible length, say $P : a_0, e_1, a_1, \ldots, a_{n-1}, e_n, a_n$ (there must exist such a path since G is finite). Since every vertex has valency at least two, a_n must be adjacent to some other vertex than a_{n-1} in G, and, moreover, it must be one of the vertices in P (a_i, say) since otherwise P would not be of maximal length. But then there exists a circuit $a_i, e_{i+1}, \ldots, e_n, a_n, e, a_i$. $\diamond$

Exercise 10 (G.A. Dirac, 1952) *A finite graph G which has $\rho(v) \geq d \geq 2\, \forall v \in G$ has a circuit of length at least $d+1$.*

Exercise 11 *If T is a (finite) tree then $|V(G)| = |E(G)| + 1$. (Hint: use induction and theorem 2.2)*

It is possible to characterise trees in a number of different ways: these seven statements can be proved to be equivalent to each other.

1. G is a tree on n vertices.
2. G has no circuits and $n-1$ edges.
3. G is connected and has $n-1$ edges.
4. Every pair of vertices in G is connected by a unique path.
5. G is connected and deletion of any edge disconnects it.
6. G is circuit-free but addition of any edge creates a circuit.
7. G is connected but addition of any edge creates exactly one circuit.

2.4 Spanning Trees

A *spanning* subgraph of a graph G is a subgraph which includes all the vertices of G. A *spanning tree* of a graph G is a spanning subgraph which is a tree. For example, the diagrams in figure 12 show four different spanning trees for the same graph:

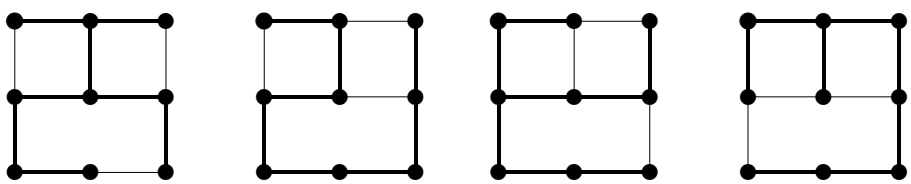


Figure 12: Four spanning trees of a graph (shown in bold)

Lemma 3 *Every finite connected graph has a spanning tree.*

Proof: Suppose that W is a closed walk in G (if no such walk exists then G is already a tree). Remove one edge from W (so that G remains connected) and we have a graph G' with one less edge than G. Obviously, we can continue this process until we have a tree. $\diamond$

Corollary 2.3 *The number of edges removed from a connected graph G to obtain a spanning tree is the same whichever spanning tree is formed and is equal to $\gamma(G) := |E(G)|-|V(G)|+1$, the* circuit rank *of G.*

It is sometimes useful to associate number with edges which indicate "cost" of using that edge. Usually this will be something like the length of a road or the difference between the genes of two species. In this case it may be useful to find a spanning tree which has the smallest sum of numbers:

Theorem 2.4 *Kruskal's algorithm (the greedy algorithm) Let G be a connected graph with n vertices and a cost associated with each edge of G. The following gives a least cost spanning tree for G:*

i) *let e_1 be an edge of least cost $c(e_1)$*

ii) *having chosen i edges $e_1, \cdots, e_i$, choose a new edge e_{i+1} of smallest cost so long as it doesn't complete a circuit in G using the edges $e_1, \ldots, e_i$.*

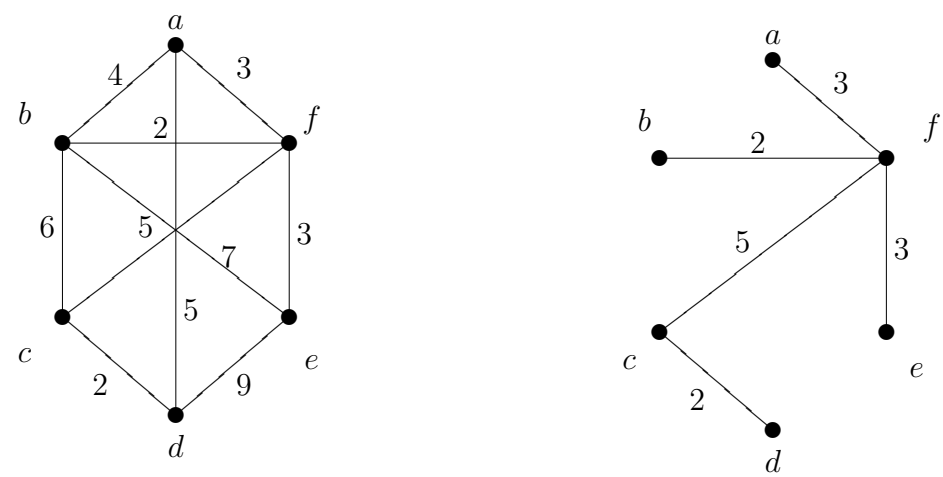
Repeat ii) for $2 \leq i \leq n-2$.

Proof: This process gives a spanning tree T of G since it is connected, it has no circuits, has $n-1$ edges and n vertices. We still have to prove that T is a spanning tree of minimum cost. To do this we assume that S is another spanning tree with a smaller cost than T and achieve a contradiction. Let e_k be the first edge in T which is not in S, so that $e_1, \ldots, e_{k-1}$ are edges both in T and S. Since S is a tree, adding e_k to S produces a unique circuit C (as in (f) above), containing e_k. But C must contain an edge e which is not in T, as otherwise T would not be a tree. If $c(e) < c(e_k)$ then we would have chosen e before e_k when forming T since the edges $e_1, \ldots, e_{k-1}, e$ are all in S and so do not contain a circuit. This contradicts what would actually have happened when carrying out the algorithm so we can assume $c(e) \geq c(e_k)$.

Thus we can consider the graph $S' := \{S \cup e_k\} - e$ which is also a spanning tree of G, but has at most the same cost as S but has one more edge in common with T than S. We can then repeat the process with S' and then S'' etc., each time increasing the number of edges in common and keeping less cost than T. But this process must terminate (when there are no edges in $S^{(p)}$ different from those in T, i.e. $S^{(p)} := T$!). But we then have

$$c(T) > c(S) \geq c(S') \geq \ldots \geq c(S^{(p)}) = c(T), \tag{1}$$

which is a clear contradiction. Hence T is a minimal spanning tree. $\diamond$



Example:

Edges	Cost	Chosen	Edges	Cost	Chosen
bf	2	1st	cf	5	5th
cd	2	2nd	ad	5	not
af	3	3rd	bc	6	not
ef	3	4th	be	7	not
ab	4	not	ef	9	not

2.5 Eccentricity

We define the *eccentricity* of a vertex in a (connected) graph G as follows:

$$\epsilon(v) := \max\{d(u, v) : u \in V(G)\}.$$

Thus it is the distance to the vertex furthest away from v. The vertices with the smallest eccentricity (the *radius* of G) in G we group together as the *centre* of G, and call the vertices *central*. Similarly, the largest eccentricity in a graph is the diameter and all vertices with this eccentricity are called *peripheral* vertices and they are grouped into the *periphery* of the graph. An eccentric vertex for a vertex v is one at distance $\epsilon(v)$ from v.

The eccentricities of the vertices in a graph are another useful tool in discriminating between graphs which may look identical but are not isomorphic. Usually, when finding the eccentricities of a graph, we write the numbers representing the eccentricity of a vertex by the vertex. In order to find this number it is simply a matter of counting the lengths of the shortest paths to each of the other vertices in the graph. Note that any peripheral vertex must, by definition, have a second vertex at distance equal to $\text{diam}(G)$ and hence there are always at least two vertices in the periphery.

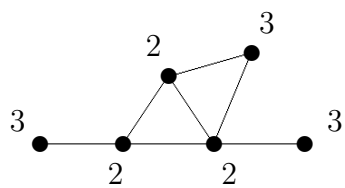


Figure 13: The eccentric labeling of a graph

In figure 13 we have six vertices, three of which have eccentricity 2, while the remaining three have eccentricity 3. This can be verified by taking any particular vertex, and then marking which vertices are adjacent to it, then taking this set of vertices, and marking all its neighbours. When there are no more vertices unmarked, the number of steps taken is the eccentricity of the vertex. Note that the centre of the graph in figure 13 is K_3. For trees we can prove there are only two different centres:

Theorem 2.5 *The centre of a tree is either isomorphic to K_2 or K_1.*

Proof: We shall proceed by induction. For the small trees K_1 and K_2 the theorem is clearly true, and these are the only trees with no internal vertices. For each vertex v in a tree T only an end vertex can be an eccentric vertex since, in a tree, any longest path must end in an end vertex. We perform a "pruning" operation on T to form a sub-tree T' as follows: remove every end vertex in T, together with its edge. This leaves a graph with fewer edges, and the eccentricities of the remaining vertices must decrease by exactly 1, since all the eccentric vertices were removed. But this leaves the centre of the tree unchanged and so the centre of T' is the centre of T. We can repeat this operation until either K_1 or K_2 is left and this graph is the centre of T, as per the theorem. $\diamond$

2.6 Matrices and graphs.

We now introduce two matrices which can be used to represent a graph G with n vertices $a_1, \ldots, a_n$ and m edges $e_1, \ldots, e_m$: the *adjacency matrix* of G is a $n \times n$ matrix A where $A_{i,j}$ is the number of edges between a_i and a_j. the *incidence matrix* of G is a $m \times n$ matrix M with $M_{i,j}$ the number of times a_i and e_j are incident. Note that the adjacency matrix is normally the smaller of the two and so is the one most usually used to store graphs in computers for example. In addition, both matrices have entries which are all either zero or one. Note that we can prove the relation $MM^T = A + D$, where D is an $n \times n$ diagonal matrix with $D_{i,i} = \rho(a_i)$.

Theorem 2.6 *The number of walks of length k from a_i to a_j is the (i,j)th entry of A^k.*

Proof: Clearly it is true for $k = 0$ ($A = I$, and the only walks of length 0 are the trivial ones from a_i to itself for all i) and $k = 1$ (the only walks of length one use just one edge of G). We shall now use induction on P_k (the statement of the theorem). Assuming P_x is true we consider $A^{x+1} = AA^x$. By the induction hypothesis, the (i,j)th entry of A^x is the number of walks of length x from a_i to a_j. But, by the rules of matrix algebra, the (c,d)th entry of A^{x+1} is $\sum A_{c,y} A^x_{y,d}$. This is equivalent to P_{x+1} since any path of length $x+1$ from a_c to a_d must first go to a neighbouring vertex a_y of a_c (which it can do in $A_{c,y}$ ways) and then, from a_y, there are $A^x_{y,d}$ paths of length x to a_d. All such paths are necessarily distinct. $\diamond$

We finish this round of definitions with a summary of the names we shall be using for graph-theoretical objects in this course. Note that many different terms for the same thing are used in books and so it is important to know what the particular author means in every case. For instance, what we shall be calling a pseudopath, Bondy and Murty call a trail.

vertex	point, node, junction
edge	line, arc
valency	degree, order
terminal vertex	pendant vertex
bipartite	even, bichromatic
walk	edge-sequence
pseudopath	chain, trail, path
path	simple path, chain, arc, way
pseudocircuit	cyclic path, circuit, closed path/trail
circuit	simple circuit, cycle
cut-vertex	point of articulation
cut-edge	bridge, isthmus

3 Planarity and Colouring.

3.1 Introduction

A graph is *planar* if it has a proper embedding (one in which no edges meet except at vertices) in the plane (or equivalently, the sphere). For example we have the two embeddings of K_4 in figure 14: the right hand one is a proper embedding but both graphs are planar (since they are isomorphic).

Figure 14: Two drawings of K_4

We note that if G is a planar graph then the area of the plane is then split up into regions which are bounded by the edges of G. We shall call these regions *faces*. For instance, in figure 14 there are four faces, the three triangles inside, plus the outside face. In fact, for any planar graph the number of faces is immediately known, thanks to the following result:

Theorem 3.1 *(Euler, 1752.) If a plane graph G has n vertices, m edges and f faces and c components then $n - m + f = 1 + c$.*

Proof: Consider the graph G_0 which has the vertices of G but no edges: the theorem is true for G_0 since in this case $c = n$, $f = 1$ and $m = 0$. We now use induction on k, the number of edges in G_k, supposing the theorem is true for G_{k-1}. Add an edge e (which is not already in G_{k-1}) from G to G_{k-1} to get G_k. There are two cases to consider:

i) e connects two vertices of G_{k-1} which did not have a path between them. In this case c is decreased but m is increased and f remains the same and so the statement is true for G_k too.

ii) e connects two vertices of G_{k-1} which did have a path between them. This operation will divide the face containing both the vertices and so f is increased but c remains the same and so, again, the statement is true for G_k too.

$\diamond$

Using the concepts we can also get relations between the number of vertices and edges as follows:

Lemma 4 *If G is a plane graph with $m = |E(G)| \geq 2$ then $2m \geq 3f$, where f is the number of faces in G as usual.*

Proof: Every face of G must be bounded by at least three edges. Letting x be the sum of the number of edges bounding each face of G we get $x \geq 3f$. But, since this sum must count every edge in G exactly twice, $x = 2e$, and hence we get the required relation. $\diamond$

Corollary 3.2 *Let G be a plane graph, $G \not\simeq K_1$ or K_2, with $n = |V(G)|$. Then $m \leq 3n - 6$.*

Proof: By lemma 4 we see that $\frac{2m}{3} \geq f$. Using theorem 3.1 we see $n - m + f = c + 1 \geq 2$ (as $c \geq 1$) and so

$$\begin{array}{rcl} n - m + \dfrac{2m}{3} & \geq n - m + f & \geq 2 \\ \Rightarrow & n - \frac{m}{3} & \geq 2 \\ \Rightarrow & m & \leq 3n - 6. \end{array}$$

$\diamond$

There is an exercise in the problems which generalizes this result to the following:

Theorem 3.3 *If the length of the smallest circuit in a graph G is g, $m \leq \frac{g}{g-2}(n-2)$.*

Finally, we can deduce something about the valencies in a plane graph:

Corollary 3.4 *Any plane graph has a vertex of valency at most 5.*

Proof: We suppose there exists a graph G in which there are no vertices of valency 5 or less. Then we can deduce that the total valency is at least $6n$ and so $m \geq 3n$. This contradicts corollary 3.2 which states $m \leq 3n - 6$. $\diamond$

3.2 Planarity Testing

Using corollary 3.2 we see that, since K_5 has 5 vertices and 10 edges, it cannot be planar. Similarly, $K_{3,3}$ has $g = 4$, $n = 6$ but $m = 9$, contradicting theorem 3.3. Hence, neither K_5 and $K_{3,3}$ can be drawn in the plane and, in fact, these are the two smallest non-planar graphs and are shown in figure 15.

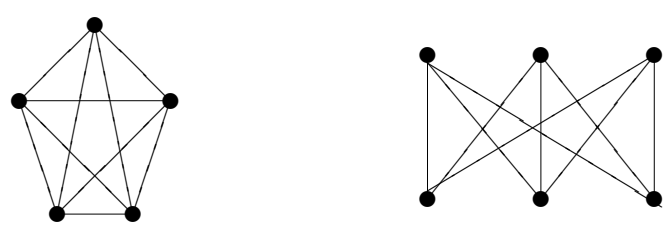


Figure 15: K_5 and $K_{3,3}$

Before we completely characterize non-planar graphs we must first make one definition: Two graphs are *homeomorphic* if they are isomorphic up to vertices of valency two. That is, we can either add or remove vertices (but not the edges adjacent to those vertices) of valency two from one graph and get the other. Two graphs homeomorphic to each other are shown in figure 16.

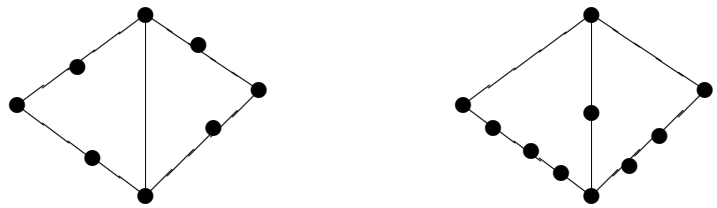


Figure 16: Two graphs homeomorphic to each other

Theorem 3.5 *(Kuratowski, 1930) A finite graph is planar if and only if it contains no subgraph homeomorphic to K_5 or $K_{3,3}$.*

The proof of this theorem is quite involved and so it shall be left unproven. Thus, in order to find a whether or not a graph is planar it is a matter of either getting a planar embedding or finding such a homeomorphic subgraph, but in practice, it is rare to find a non-planar graph without $K_{3,3}$ as a homeomorphic subgraph so it normally suffices to check for that once you are satisfied that no planar embedding exists.

Drawing Planar Graphs

So we think that the graph in figure 17 might be planar. We want to draw it as a planar graph but first we need to define a couple of new terms:

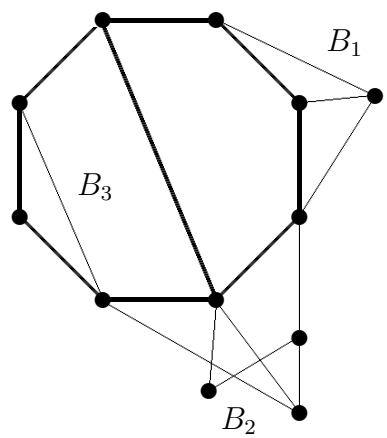


Figure 17: A possibly planar graph and its bridges

A *bridge* B of a subgraph H of G is either an edge of G which is not in H, or a component of $G - H$ together with all the edges joining it to G.

The vertices which are in $V(B) \cap V(H)$ are called the *vertices of attachment*. We say that B is *compatible* with a face F of H if all of its vertices of attachment come from the border of F. For example, in the graph in figure 17 H has three faces. B_2 is only compatible with the outer face whereas the other two bridges are also compatible with one of the inner faces.

We have failed to find a Kuratowski subgraph, so we proceed to apply algorithm 1.

Algorithm 1

1. *Choose a circuit in G as big as possible and draw it as a polygon in the plane.*
2. *If H is the graph we have so far in the plane and $H \ncong G$ then find a bridge B compatible with one face of F. If no such B exists choose any bridge for B and let F be one of its compatible faces.*
3. *Choose a path in B between two of its vertices of attachment and draw it in F then return to the previous step.*

If we follow this algorithm properly then we will either find a plane embedding (and it just remains to show that this graph is the same as the one we started with) or we find that the algorithm breaks down (and so we have to assume it is non-planar and go back to looking for Kuratowski subgraphs, using the information gleaned from the algorithm).

3.3 Topological Graph Theory

We also briefly divert towards topology, in considering other surfaces (other than the plane) that we can embed graphs upon. Technically these are are compact connected (perhaps non-) orientable manifolds but we shall just refer to them as surfaces. For instance, from the sphere we have its closest relation, the torus (doughnut) which can be imagined to just be a sphere with a hole all through it. From this we can consider the double-torus, treble torus, etc., to just be the sphere with however many holes in it.

Alternatively we can also have non-orientable surfaces such as the projective plane (which can be thought of as a Möbius cylinder) and the Klein bottle. These objects can be represented in the plane using the "arrowed polygon" notation shown in figure 18. You have to imagine the polygon as a sheet of paper and then you join the edges of the piece of paper which have matching arrows. This can be seen for the torus by doing it, but it is impossible to create the projective plane in three dimensions...

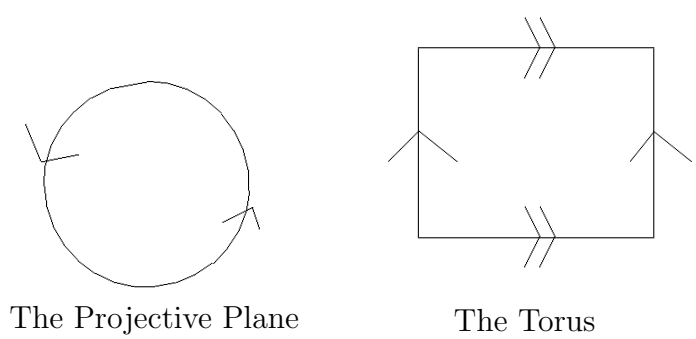


Figure 18: Plane Representation of Surfaces

The arrows on the sides of the polygons represent in what orientation the opposite sides are identified, in the sense that an edge hitting the bottom edge of the torus would continue

from the point vertically above it on the top edge. For the projective plane the same happens relative to the arrows, which coincides with reappearing at the opposite side of the circle.

For example, figure 19 shows embeddings of K_5 on the torus and $K_{3,3}$ on the projective plane

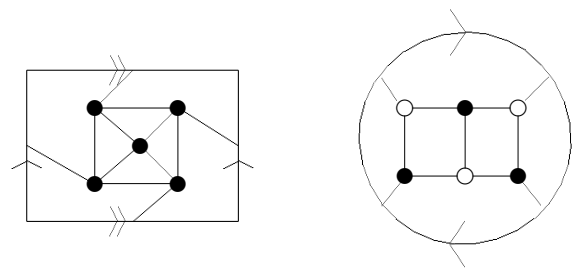


Figure 19: Some embeddings on higher surfaces

Exercise 12 *Embed K_5 on the projective plane.*

For each surface S there is a constant $\chi = \chi(S)$, the Euler-Poincaré characteristic, and it is defined as the number such that any connected graph embedded in S (such that every face is homeomorphic to an open disc in the plane) satisfies

$$n - m + f = \chi.$$

The sphere/plane has $\chi = 2$, torus $\chi = 0$ and the n-torus $\chi = 2 - 2n$. The projective plane has $\chi = 1$ and the Klein bottle $\chi = 0$ and this sequence continues similarly. We proved this relation for the plane in theorem 3.1.

Lemma 5 *Let G be a plane graph that is d-regular, such that every face has k edges bounding it. Then*

$$nd = |V(G)| \times d = fk = |E(G)| \times 2 = 2m.$$

Proof: The total valency of G is nd, and is equal to $2m$ as in section 1. As in lemma 4 this again counts every face each edge is a part of the boundary of and so must also equal fk. ◇

Theorem 3.6 *There are exactly five regular polyhedra.*

Proof: Any polyhedron can be seen to have a representation as a planar graph G since its projection onto a sphere can be transferred simply to the plane. Thus G must be d-regular and each face has k edges bounded, where d and k are integers. We can see that $d \geq 3$ since $d = 1$ gives only the graphs xK_2 and $d = 2$ implies that G is the circuit C_k neither of which are polyhedra. As before $k \geq 3$ for all d since all faces must be at least triangular. Since G is connected we have $n - m + f = 2$ and by lemma 5 we have $nd = 2m$ and $fk = 2m$ and so, since $m, d, k \geq 0$, ◇

$$x = \frac{1}{d} - \frac{1}{2} + \frac{1}{k} = \frac{1}{m} \geq 0.$$

Table 2: The Regular Polyhedra

d	k	x	m	n	f	name
3	3	$\frac{1}{6}$	6	4	4	tetrahedron
	4	$\frac{1}{12}$	12	8	6	cube
	5	$\frac{1}{30}$	30	20	12	dodecahedron
	6	0	∞	∞	∞	(hexagon-tiling)
	> 6	< 0	-	-	-	no polyhedra
4	3	$\frac{1}{12}$	12	6	8	octahedron
	4	0	∞	∞	∞	(square-tiling)
	> 4	< 0	-	-	-	no polyhedra
5	3	$\frac{1}{30}$	30	12	20	icosohedra
	> 3	< 0	-	-	-	no polyhedra
6	3	0	∞	∞	∞	(triangle-tiling)
	> 3	< 0	-	-	-	no polyhedra

3.4 Chromatic Number and Polynomials

Given any graph G we define its chromatic number as the minimum number of colours required to give each vertex of G a colour in such a way that no two adjacent vertices share the same colour. We denote this graph theoretical constant as $\chi(G)$. We can easily evaluate it for most of the basic families of graphs such as K_n, $\overline{K_n}$, trees, C_n: $\chi(K_n) = n$ since every pair of vertices is adjacent and so all vertices need a new colour; $\chi(\overline{K_n}) = 1$ because no two vertices are adjacent; $\chi(T) = 2$ since trees are bipartite, and two colours are needed unless $T = K_1$; $\chi(C_n) = 2$ or 3 depending on whether n is even or odd.

Theorem 3.7 *Any planar graph is five colourable; i.e. if G is planar then $\chi(G) = 5$.*

Proof: We prove that any triangulation can be coloured with a certain number k of colours then, since any planar graph G can be turned into a triangulation T (by adding edges across faces), we can do the following: colour T with k colours and then remove the edges added to leave G, and the colouring thus formed is still a proper one. We use induction on the number of vertices of the graph, supposing that every planar graph with less than n vertices is 5-colourable. This is certainly true for any planar graph with less than six vertices so we have the base of the induction.

By corollary 3.4 we know that there must exist some vertex v in G with valency at most five. If we consider the graph $G - v$ this has a 5-colouring by the induction hypothesis and so we take any such colouring. If $\rho(v) \leq 4$ then it has at most four neighbours and these can have at most four of the five colours available so we can simply replace v and colour it with any colour which is not used by its neighbours. We can proceed similarly if $\rho(v) = 5$ but one colour is not used in colouring v's five neighbours. Thus we just need to consider the case in which we have all five colours for v's neighbours and so the situation is as shown in figure 20

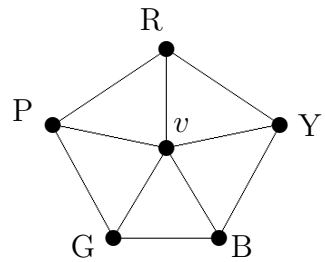


Figure 20: The difficult case in the five-colour theorem

To resolve this problem we consider subgraphs of $G - v$ induced by two of the five colours and the components of these graphs are called Kempe chains. In figure 20 we first consider the red-green induced subgraph and consider whether the red and green vertices shown are in the same Kempe chain. If they are not then we swap the colours in the chain involving the red vertex, red for green and vice versa. This operation still leaves a valid colouring but now v has no red neighbour and so can be coloured red in G, giving it a proper colouring.

If there is a red-green chain joining the red and green neighbours of v then we consider all blue-purple chains. This induced subgraph cannot have one component containing both neighbours of v since $G - v$ is planar and the red-green chain splits the graph into two separate components. Thus we can change one of the two chains as before to allow v to be coloured purple. $\diamond$

Chromatic Polynomials

We define the function $P(G, t)$ as the number of ways in which one can colour the vertices of G using t colours, where each vertex is labeled so that any problems with isomorphism are suppressed. For instance, $P(K_3, 2) = 0$ but $P(K_3, 3) = 6$. It is first necessary to establish that this function is indeed a polynomial:

Theorem 3.8 *If $|V(G)| = n$ then $P(G, t)$ is a monic polynomial of degree n.*

By inspection we can note that $P(K_n, t) = \frac{t!}{(t-n)!}$(since we have t colours available for the first vertex, the next can choose any one of the $t - 1$ others, and this continues since every vertex is adjacent to every other. Similarly, $P(\overline{K_n}, t) = t^n$ since every vertex can have any of the t colours. For trees we have $P(T, t) = t(t - 1)^{n-1}$ since the first vertex can have any of the t colours and we choose the colours of the remaining vertices recursively by considering neighbours of those already coloured. We always have $t - 1$ colours to choose from since any vertex has exactly one previously coloured neighbour as there are no circuits.

In order to calculate chromatic polynomials there are several very useful more advanced techniques which we shall not prove in this course:

1. If $G = G_1 \cup G_2$ (in that every edge of G is in either G_1 or G_2 or perhaps both) and $G_1 \cap G_2 = K_i$ then we use Complete Intersection:

$$P(G, t) = \frac{P(G_1, t)P(G_2, t)}{P(K_i)}$$

2. If e is an edge of G then we have Deletion-Contraction:

$$P(G,t) = P(G-e,t) - P(G \circ e,t)$$

or equivalently, if there is no edge e, Addition-Identification:

$$P(G,t) = P(G+e,t) + P(G \circ e,t)$$

3. Finally, if v is adjacent to every vertex of G then

$$P(G,t) = t \times P(G-v,t-1)$$

Exercise 13 *Find the chromatic polynomials of some of the graphs in the course*

4 Eulerian and Hamiltonian circuits.

4.1 Eulerian circuits

An *Euler-circuit* (Eulerian pseudocircuit) of a graph G is a pseudocircuit that covers the whole graph; i.e. one which includes every edge of G (exactly once) and every vertex (probably more than once each). We say that G is *Eulerian* if it has an Euler-circuit or it has only one vertex; it is then clearly finite and connected.

Theorem 4.1 *(Euler, 1736) A finite connected graph G is Eulerian if and only if every vertex of G has even valency.*

Proof: "only if" is clear: if a vertex occurs k times in the pseudocircuit then its valency is $2k$.
"if": Suppose that every vertex of G has even valency. The result is obvious if $|V(G)| = 1$; so suppose that $|V(G)| >= 2$. Let P, connecting a to b, be a pseudopath of maximum length in G. If $b \neq a$ then P must use an odd number of edges at b, and so there must be an unused edge at b; this edge can be used to increase the length of P, a contradiction. Thus $b = a$ and P is actually a pseudocircuit. We prove that P is actually an Euler-circuit of G.

Let cd be an edge in G that is not in P: since G is connected there is a path from c to a, say, and the first vertex of P along the path, say e, is incident to an edge of G not in P. Let Q be a maximal length pseudocircuit from e to f, using only edges not in P. By the same argument as before, $f = e$ and Q is a pseudocircuit. We can thus add Q to P in the obvious way to form a longer pseudocircuit than P. This contradiction shows that no such edge as cd can exist and so P is an Euler-circuit of G. $\diamond$

4.2 Algorithms for Finding Euler-Circuits

Algorithm 2 Fleury's algorithm (1921). *Let G be an Eulerian graph. Then the following procedure is always possible and will lead to the construction of and Eulerian pseudocircuit of G.*
Start out from any vertex and proceed along the edges in any manner subject to the following rules:

(a) *erase the edges as they are used, along with any vertex that would become isolated if it were not erased.*

(b) *never use an edge if its removal at that moment would disconnect the graph (except for isolating the initial vertex of the edge).*

In short: "erase the edges as you go along and don't disconnect other edges".

4.3 Hamiltonian Circuits

A *Hamiltonian* circuit in a graph is defined as a circuit which includes every vertex in the graph. This may seem a similar concept to Euler-circuits but, as we shall see, it is much harder to even tell whether or not a graph has a Hamiltonian circuit (in which case we say it is Hamiltonian), let alone finding one. However, we can say some things:

Theorem 4.2 *A graph with a cut-vertex is not Hamiltonian*

Proof: Suppose we have a cut-vertex v which splits a graph G into two parts, G_1 and G_2. If G is Hamiltonian there exists a cycle C including each vertex in G. We suppose C exists and trace it around G: without loss of generality we can start in G_1. We move along C in G_1 until we come to v; since v is the only vertex in both G_1 and G_2 we must now cross into G_2 as otherwise we can never include any vertex in G_2 in C. But now, in order to return the the vertex with which we started, we must cross back into G_1 again, and that would mean using v again, which is impossible as C goes through every vertex in G exactly once. Thus C cannot exist and so G is not Hamiltonian. $\diamond$

In practice, we normally find that a graph is Hamiltonian by finding a Hamiltonian circuit in it. This can be done by trial and error, guided by a couple of small observations and a little inside knowledge:

Firstly, we know that any vertex of valency 2 has to imply that both of its incident edges are used in any Hamiltonian cycle. This then gives us information about which edges its neighbours use. Similarly to this, we can try certain pairs of edges at vertices of valency higher than two and then deduce from this guess, restrictions on which edges can be used at several other vertices. We normally show these restrictions by the use of markers called *transitions*, which are indicated as curved of coloured lines.

For instance, we try to find a Hamiltonian cycle in figure 21:
the two transitions at the vertices of valency two are marked. We use the symmetry of the graph to note that x and its mirror image can't both use the same type of transition (the ones inside form a 4-cycle, the ones outside form a 6-cycle). Therefore, we can assume x uses the inside transition and goes to y where the other vertex uses the other transition and goes to z via the other vertex of valency 2. From here it is an easy matter to deduce that this graph is indeed Hamiltonian, the rest of the cycle traveling around the triangle from y to z.

To prove a graph non-Hamiltonian we proceed along similar lines; for instance, consider the graph formed from figure 21 by deleting the edge ab. We now have transitions at both of these vertices and these transitions already form a 6-cycle. Since transitions must form

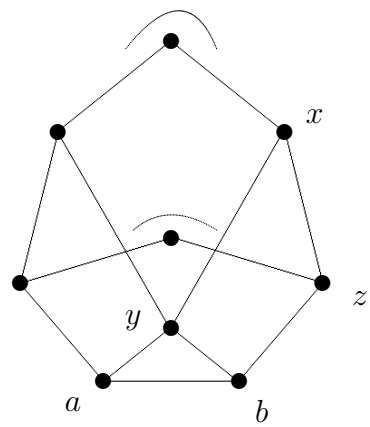


Figure 21: Transitions in a graph

only an n-cycle we know that it is now non-Hamiltonian. More complicated cases are dealt with by using symmetry and breaking the proof down into cases.

Exercise 14 *Prove whether or not the Petersen graph is Hamiltonian.*

The following theorem is a sufficient condition, but any graph which satisfies the condition has so many edges that it is almost always much easier to find a Hamiltonian circuit just by tracing randomly through the graph:

Theorem 4.3 *Let G be a graph with n vertices. If, for any two non-adjacent vertices u and v are such that $\rho(u) + \rho(v) \geq n$ then G is Hamiltonian.*

Taking this idea further we define the closure of a graph to be the graph formed by adding in an edge between two vertices of a graph such that they satisfy the conditions in theorem 4.3, and repeating this operation until it is impossible to continue. It has been proved that this graph is unique, and we call it the *closure* of G, $C(G)$, and it is now possible to state a known result:

Theorem 4.4 *G is Hamiltonian if and only if $C(G)$ is.*

Although this result now gives us a fairly good method of determining whether a graph is Hamiltonian there are some problems with graphs in which $G \cong C(G)$ and so the theorem gives us no new information in this case. Normally, however, the closure of a Hamiltonian graph will turn out to be a complete graph or some other graph easily seen to be Hamiltonian.

We conclude the course with the statement and demonstration of a powerful theorem for investigating whether planar graphs are non-Hamiltonian. Let C be a Hamiltonian circuit in a graph G and let a_i denote the number of faces of size i outside C and b_i the number of size i inside.

Theorem 4.5 (Grinberg, 1968)

$$\sum_{i=3}^{\infty}(i-2)(a_i - b_i) = 0.$$

To apply this theorem we assume the existence of a Hamiltonian circuit and then consider the equation and whether it can have a solution in integers. For instance, if there were a Hamiltonian graph with faces of size 5, 8 and 7 but only one of size 7, we can immediately see that we must have

$$3(a_5 - b_5) + 6(a_6 - b_6) = 5(b_7 - a_7).$$

But the left hand side clearly is divisible by three whereas the right cannot be since it is equal to either 5 or -5. Hence no such Hamiltonian graph can exist.

Sometimes, though, Grinberg's theorem gives us false hope; it can give rise to a set of soluble equations but the graph itself is not Hamiltonian so that the theorem cannot be used to check Hamiltonicity surely.

Exercise 15 *Construct a non-Hamiltonian planar graph which has Grinberg solutions.*

4.4 Matchings and Independent Sets

We define two more graph theoretic parameters in this part of the course; the *matching* number of G is the maximum number of vertex disjoint edges it is possible to fit into G. Similarly, the *independence* number of G is the maximum number of non adjacent vertices it is possible to fit into G. The independence number of G is denoted by $\alpha(G)$.

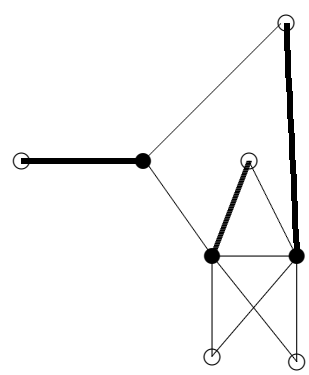


Figure 22: Independence and Matching sets

For instance, in the graph in figure 22 we have independence number of 5, shown by the hollow circles, and matching number 3, as shown by the bold lines. That these values are no more is quite an intricate task to explain, as was the case when we considered chromatic number and connectivity. In this case we try as follows:

If the matching number was 4 then every vertex in the set would have a bold edge in it. In particular, the two bottom-most vertices would need edges covering them and these edges

would then have to match the two vertices above them. This leaves 4 vertices to match with two edges, but one is isolated from the other three, making this impossible.

Similarly, we suppose the independence number is 6. This would mean that only two of the vertices would not be in the set, and since there is a triangle in the graph, at most one of these three vertices must in any independent set. This implies that all the other 5 vertices must be in a set but it is easy to see that set would not be an independent one.

In general it is much easier to calculate the matching number thanks to this theorem which involves the concept of an *alternating path*, which is defined as a subgraph of the graph which is a path whose edges are in the matching set, then out, then in, etc.

Theorem 4.6 *A matching is maximal if it contains no alternating path with non-matched ends.*

Proof: If such an alternating path exists then we simply switch the roles of the edges in the path and leave the other edges alone and thus generate a matching which is larger than the original one. ◇

It is possible to prove that this theorem can be stated as "if and only if" but time shall probably prevent our covering this topic. This theorem can be used to generate an algorithm to find a maximum matching in any graph.

Edge Covering Sets

It can be noted that, in figure 22, the filled vertices form a set which has a different property; every edge in G is incident with at least one of these vertices. Such a set is called an *edge cover* and a smallest such set is a minimal edge covering set.

Theorem 4.7 *The size of a minimal edge covering set is $n - \alpha(G)$.*

Proof: Suppose we have a set of vertices C which is an edge covering set which is of size larger than $n - \alpha(G)$. Then the set $I := V(G) - C$ is an set of vertices of size less than $\alpha(G)$. We now prove that I is independent. If it were not, there would be two vertices in I which had an edge e between them, but then that would imply that C didn't cover e and so C wasn't an edge covering set. Thus I would be a smaller independent set than the minimal one, a contradiction.

Now we just have to show that a set of size $n - \alpha(G)$ exists and we form it as mentioned above, from an independent set, J. Considering the vertices not in J, they form an edge covering set as if they didn't J wouldn't be independent. ◇