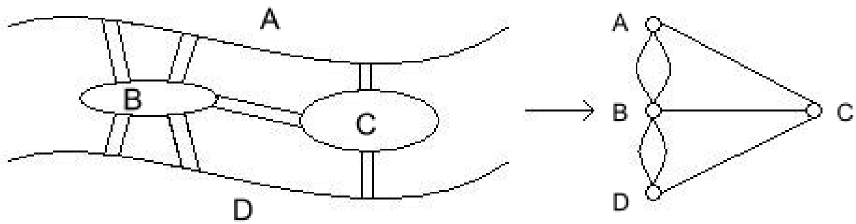
CS 70 Discrete Mathematics and Probability Theory
Summer 2014 James Cook **Note 9**

An Introduction to Graphs

A few centuries ago, residents of the city of Königsberg, Prussia were interested in a certain problem. They wanted to know whether it was possible to walk through their city, represented in the image below, by crossing each bridge only once.



This problem is called The Seven Bridges of Königsberg, and in 1736 the brilliant mathematician Leonhard Euler proved that such a task was impossible. Euler did this by modeling the arrangement of bridges as a graph (or, more accurately, a *multigraph* since there can be multiple edges between the same pair of vertices). For this reason, Euler is generally hailed as the inventor of graph theory.

The relevant features of the seven bridge problem are the two islands (B and C), the two banks (A and D), and the bridges between these areas. All of these features can be represented by a *graph*, as shown on the right in the image above. The two islands and each bank are represented by the 4 little circles, which are called *vertices*. The bridges are represented by the lines between the vertices, called the *edges* of the graph. (Include labels for vertices/edges?) This graph happens to be *undirected*, since the bridges can be crossed in either direction. If we instead had one-way bridges, the graph would be *directed*.

More formally, a directed graph $G(V,E)$ consists of a finite set of vertices V and a set of edges E. E is a set of pairs (u,v) of vertices[1]. An edge (v,w) in a directed graph is usually indicated by drawing a line between v and w, with an arrow pointing towards w. *Undirected graphs* may be regarded as special kinds of directed graphs, in which $(u,v) \in E$ if and only if $(v,u) \in E$. Thus in an undirected graph the directions of the edges are unimportant, so an edge of an undirected graph is an *unordered* pair of vertices $\{u,v\}$ and is indicated by a line between u and v with no arrow.

As we have defined them, graphs are allowed to have *self-loops*; i.e, edges of the form (u,u) that go from a vertex u to itself. Usually, however, graphs are assumed to have no self-loops unless otherwise stated, and we will assume this from now on. We will also be working only with graphs, and not multigraphs such as the one representing the bridges of Königsberg.

A *path* in a directed graph $G = (V,E)$ is a sequence of edges $(v_1,v_2),(v_2,v_3),...,(v_{n-2},v_{n-1}),(v_{n-1},v_n)$. In this case we say that there is a path *between* v_1 *and* v_n. A path in an undirected graph is defined similarly. Usually a path is assumed to be *simple*, meaning $v_1,\dots,v_n$ are distinct. A path with repeated vertices will be called a *walk*. A *cycle* (or *circuit*) is a sequence of edges $(v_1,v_2),(v_2,v_3),,(v_{n-2},v_{n-1}),(v_{n-1},v_n),(v_n,v_1)$, where $v_1,\dots,v_n$ are distinct. A *tour* is a walk which starts and ends at the same vertex. A graph is said to be

[1] Another way to say this is that E is a subset of $V \times V$, where $V \times V$ is the *Cartesian product* of V with itself.

connected if there is a path between any two distinct vertices.

We say that an edge (u,v) is *incident* to vertices u and v, and vertices u and v are *neighbors* or *adjacent*. If $G=(V,E)$ is an undirected graph then the *degree* of vertex $u \in V$ is the number of edges incident to u, i.e, $degree(u) = |\{v \in V : \{u,v\} \in E\}|$. A vertex u whose degree is 0 is called an *isolated* vertex, since there is no edge which connects u to the rest of the graph. In a directed graph, the *in-degree* of a vertex u is the number of edges from other vertices to u, and the *out-degree* of u is the number of edges from u to other vertices.

Eulerian Walks and Tours

We can now restate the problem above in the language of graphs: is there a walk in the graph that uses each edge exactly once? This type of walk is called an *Eulerian walk*. An *Eulerian tour* is a tour that uses each edge exactly once.

Let an even degree graph be a graph in which all vertices have even degree. The next theorem gives necessary and sufficient conditions for a graph to have an Eulerian tour.

Euler's Theorem: An undirected graph $G=(V,E)$ has an Eulerian tour if and only if the graph is connected (except possibly for isolated vertices) and even degree.

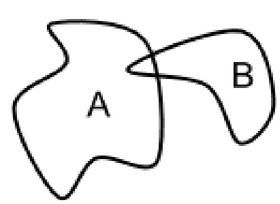
Proof ($\Longrightarrow$): Assume that the graph has an Eulerian tour. This means every vertex that has an edge adjacent to it (i.e, every non-isolated vertex) must lie on the tour, and is therefore connected with all other vertices on the tour. This proves that the graph is connected (except for isolated vertices).

Next note that, for each vertex in the tour except the first (and last), the tour leaves it in the next step after entering it. Thus, every time the tour visits a vertex, it traverses exactly two edges incident to it. Since the Eulerian tour uses each edge exactly once, this implies that every vertex except the first has even degree. And the same is true of the first vertex v as well because the first edge leaving v can be paired up with the last edge returning to v. This completes the proof of this direction.

Proof ($\Longleftarrow$): Assume that $G=(V,E)$ is connected and even degree. We will show how to construct an Eulerian tour in G. We will begin walking from a vertex u, never repeating edges. Since the graph is even degree, each time the walk enters a vertex, it must be able to exit as well. Our walk can therefore only end on u; it will be a tour. The claim below formalizes this intuition:

> **Claim 1** In an even degree graph, a walk starting from a vertex u can only get stuck at u.

Let's call the resulting tour A. Is it Eulerian? Not necessarily; our tour may not encounter all edges in the graph. Consider the image below- our algorithm could have resulted in tour A, which does not traverse edges along B and is therefore not Eulerian.



To continue, we will remove A from G and create a new tour B on the remaining edges. We can do this because the remaining graph (consisting of the remaining edges) still has even degree:

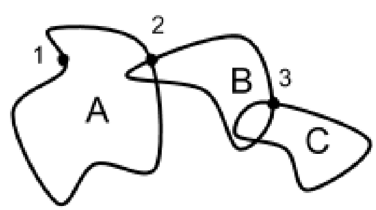
Claim 2 Removing a tour from an even degree graph will result in an even degree graph.

However, since our eventual goal is to create a single Eulerian tour. So we must splice tours A and B together into a single tour that traverses each of them. We can only do this if they intersect — there must be some edge of B such that one of its endpoints lies on A. The following claim says there is always an untraversed edge "hanging" from A (we will use this edge as our starting point to discover tour B):

Claim 3 Let A be a tour in a connected (except for isolated vertices) graph G. If A does not contain all edges in G, there exists an edge $\{u,v\}$ such that A passes through u but does not contain $\{u,v\}$.

We can now begin walking from u, using only edges that did not occur in A. By Claim 2, the removal of edges of A results in an even degree graph and therefore by Claim 1, this new walk will get stuck at u, creating tour B. We can combine A with B by splicing the two together: we start walking along A until we reach u, and then walk along B until we return to u, and then return to walking along A until we finish. If this new tour does not include all edges in G, Claim 3 implies that there exists an untraversed edge which is connected to the tour. We can then repeat the process.

Here is a possible scenario:



If the above image is the graph in question, our first tour could be tour A. Our second tour could be tour B. We splice the first two tours together by starting at point 1, walking along A until we reach point 2, walking along B until we return to point 2, and finishing our walk along A, ending back at point 1. Call this tour T. We then create another tour C. To splice T and C together, we walk along T until reaching point 3, then walk along C until returning to point 3, and then finish our walk along T. The final walk would be an Eulerian tour.

We now prove all the claims above.

Proof of Claim 1

First consider a walk from u to v. For a vertex w, let $n(w)$ be the number of edges on the walk incident to w. Let's say we get stuck at v. Then $n(v)$ is odd; we have entered v but not exited. However, since v is even degree, there must be at least one unused edge incident to v, which we can use to exit v. Therefore, we cannot get stuck at v.

Proof of Claim 2

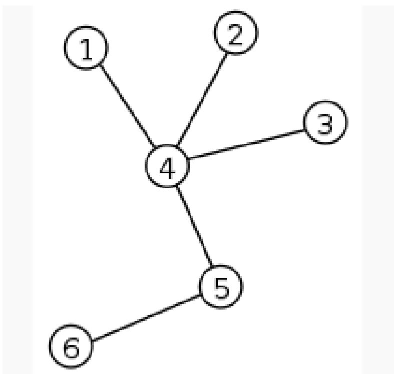
If we removed the edges traversed by a tour, we are decreasing the degree of each vertex w by $n(w)$. Since a tour exits each vertex as often as it enters and our tour does not have any repeated edges, $n(w)$ must be even for all w. Therefore, the degree of each vertex remains even.

Proof of Claim 3

Suppose A does not contain all edges in the graph G. Let $\{x,y\}$ be an edge not in the tour. If x is on A then we are done. Otherwise, since the graph is connected, there is a path from vertex a on tour A to vertex x. This path starts with a vertex on A and ends with a vertex not on A. As we traverse the path from a to x, there must be a first time it touches a vertex v not on the tour A (you can formally prove this by induction!). The previous vertex u must be on the tour, and so $\{u,v\}$ is the desired edge.

Trees

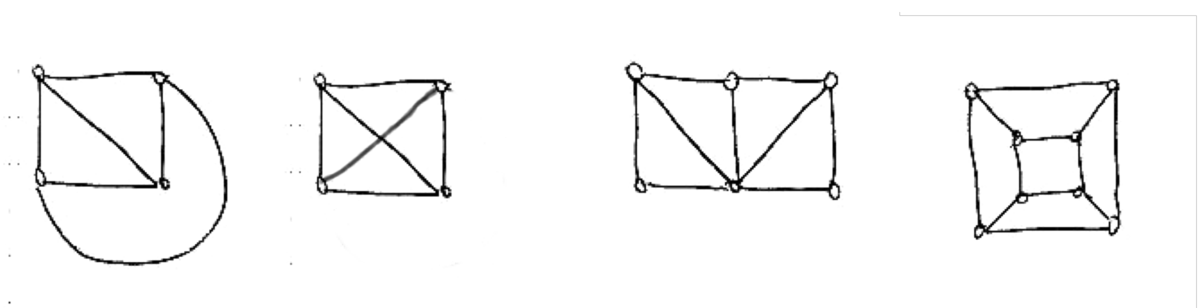
A graph is a *tree* if it is connected and acyclic (contains no cycles). There are many other equivalent definitions. For example, a tree is a connected graph where the number of vertices is one more than the number of edges. Or, a tree is a connected graph such that if you delete any edge it becomes disconnected.

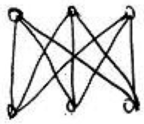
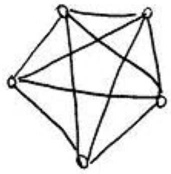
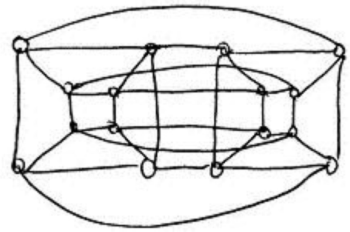


Planar Graphs

A graph is *planar* if it can be drawn on the plane without crossings. For example, the first four graphs shown below are planar. Notice that the first and second graphs are the same, but drawn differently. Even though the second drawing has crossings, the graph is still considered planar since it is possible to draw it without crossings.

The other three graphs are not planar. The first one of them is the infamous "three houses-three wells graph," also called $K_{3,3}$. The second is the complete graph with five nodes, or K_5. The third is the four-dimensional cube. We shall soon see how to prove that all three graphs are non-planar.



When a planar graph is drawn on the plane, one can distinguish, besides its vertices (their number will be denoted v here) and edges (their number is e), the *faces* of the graph (more precisely, of the drawing). The faces are the regions into which the graph subdivides the plane. One of them is infinite, and the others are finite. The number of faces is denoted f. For example, for the first graph shown $f = 4$, and for the fourth (the cube) $f = 6$.

One basic and important fact about planar graphs is *Euler's formula*, $v + f = e + 2$ (check it for the graphs above). It has an interesting story. The ancient Greeks knew that this formula held for all polyhedra (check it for the cube, the tetrahedron, and the octahedron, for example), but could not prove it. How do you do induction on polyhedra? How do you apply the induction hypothesis? What is a polyhedron minus a vertex, or an edge? In the 18th century Euler realized that this is an instance of the inability to prove a theorem by induction *because it is too weak*, something that we saw time and again when we were studying induction. To prove the theorem, one has to generalize polyhedra. And the right generalization is *planar graphs*.

Can you see why planar graphs generalize polyhedra? Why are all polyhedra (without "holes") planar graphs?

Theorem 9.1: (Euler's formula) For every connected planar graph, $v + f = e + 2$

Proof: By induction on e. It certainly holds when $e = 0$, and $v = f = 1$. Now take any connected planar graph. Two cases:

- If it is a tree, then $f = 1$ (drawing a tree on the plane does not subdivide the plane), and $e = v - 1$ (check homework).
- If it is not a tree, find a cycle and delete any edge of the cycle. This amounts to reducing both e and f by one. By induction the formula is true in the smaller graph, and so it must be true in the original one.

■

What happens when the graph is not connected? How does the number of connected components enter the formula?

Take a planar graph with f faces, and consider one face. It has a number of *sides*, that is, edges that bound it clockwise. Note that an edge may be counted twice, if it has the same face on both sides, as it happens for example in a tree (such edges are called bridges). Denote by s_i the number of sides of face i. Now, if we add the s_i's we are going to get $2e$, because each edge is counted twice, once for the face on its right and once for the face on its left (they may coincide if the edge is a bridge). We conclude that, in any planar graph,

$$\sum_{i=1}^{f} s_i = 2e. \tag{1}$$

Now notice that, since we don't allow parallel edges between the same two nodes, and if we assume that there are at least two edges, every face has at least three sides, or $s_i \geq 3$ for all i. It follows that $3f \leq 2e$.

Solving for f and plugging into Euler's formula we get

$$e \leq 3v - 6.$$

This is an important fact. First it tells us that planar graphs are *sparse*, they cannot have too many edges. A 1,000-vertex connected graph can have anywhere between a thousand and half a million edges. This inequality tells us that for planar graphs the range is very small, between 999 and $2,994$.

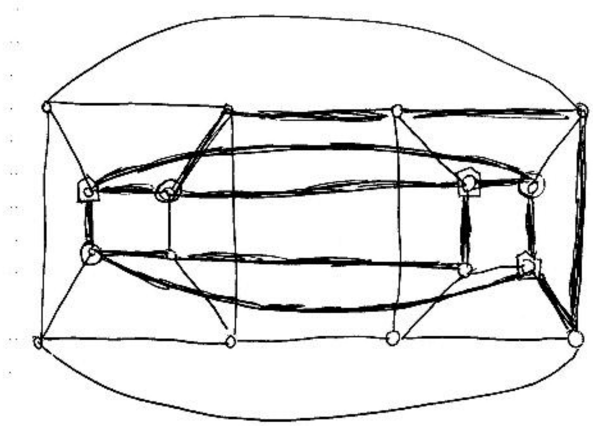
It also tells us that K_5 is not planar: Just notice that it has five vertices and ten edges.

$K_{3,3}$ has $v = 6, e = 9$ so it passes the planarity test with flying colors. We must think a little harder to show that $K_{3,3}$ is non-planar. Notice that, if we had drawn it on the plane, there would be no triangles. Because a triangle means that two wells or two houses are connected together, which is false. So, Equation (1) now gives us $4f \leq 2e$, and solving for f and plugging into Euler's formula, $e \leq 2v - 4$, which shows that $K_{3,3}$ is non-planar.

So, we have established that K_5 and $K_{3,3}$ are both non-planar. There is something deeper going on: In some sense, these are *the only non-planar graphs*. This is made precise in the following famous result, due to the Polish mathematician Kuratowski (this is what "K" stands for).

Theorem 9.2: A graph is non-planar if and only if it contains K_5 or $K_{3,3}$.

"Contains" here means that one can identify nodes in the graph (five in the case of K_5, six in the case of $K_{3,3}$) which are connected as the corresponding graph through paths (possibly single edges), and such that no two of these paths share no vertex. For example, the 4-cube shown below is non-planar, because it contains $K_{3,3}$, as shown.



Can you find K_5 in the same graph?

One direction of Kuratowski's theorem is obvious: If a graph contains one of these two non-planar graphs, then of course it is itself non-planar. The other direction, namely that in the absence of these graphs we can draw any graph on the plane, is difficult. For a short proof you may want to type "proof of Kuratowski's theorem" in your favorite search engine.

Duality and Coloring

There is an interesting *duality* between planar graphs. For example, the Greeks knew that the octahedron and the cube are "dual" to each other, in the sense that the faces of one can be put in correspondence with the vertices of the other (think about it). The tetrahedron is self-dual. And the dodecahedron and the icosahedron (look for images in the web if you don't know these pretty things) are also dual to one another.

What does this mean? Take a planar graph G, and assume it has no bridges and no degree-two nodes. Draw a new graph G^*: Start by placing a node on each face of G. Then draw an edge between two faces if they touch at an edge — draw the new edge so that it crosses that edge. The result is G^*, also a planar graph. Notice now that, if you construct the dual of G^*, it is the original graph: $(G^*)^* = G$.

Duality is a convenient consideration when thinking about planar graphs. Also, it tells us that "coloring a political map so that no two countries who share a border have the same color" is the same problem as "coloring the vertices of a planar graph (the dual of the political map) so that no two adjacent vertices have the same color." A famous theorem states that four colors are always enough! (Search for "four color theorem".) We shall prove something weaker:

Theorem 9.3: Every planar graph can be colored with five colors.

Proof: Induction on v. The base case is not worth talking about, so we go directly to the inductive step. Let G be a planar graph. I claim there is a node of degree five or less. In proof, consider the inequality $e \leq 3v - 6$. If all vertices had degree six or more, then e would be at least $3v$.

So, consider a node u of degree five or less. If it has degree four or less, we are done: Remove u, color the remaining graph with 5 colors (by induction), and then put u back in and color it by a color that is missing from its neighbors.

So, u must have 5 vertices, and in the coloring of $G - u$ they all got different colors. Look at them clockwise around u, and call them u_1, u_2, u_3, u_4, u_5, and their colors $1, 2, 3, 4, 5$. Now try to change the color of u_2 to color 4. If you can legally do it, we are done, because then you color u_2 by color 4, and color u by color 2. But this may not be possible, because there is a neighbor of u_2 colored 4. So we try to color that node 2. And so on. If this process ever succeeds, you are done by a chain of color changes. The only way that it can fail is if it ends up at the neighbor u_4 of u. That is, there is path from u_2 to u_4 of nodes colored 2 and 4.

Similarly, we can try to change the color of u_1 to 3, and this will succeed unless there is a path from u_1 to u_3 colored 1 and 3.

If both of these attempts fail, then we get two paths: one from u_1 to u_3 colored 1 and 3, and the other from u_2 to u_4 colored 2 and 4. But planarity says that these two paths must intersect at some vertex. What is the color of this vertex? ■