Math 158 Lecture Notes (Professor: Jacques Verstraete)

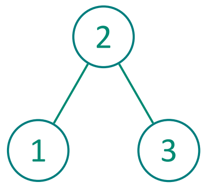
Isabelle Mills

March 24, 2024

Lecture 1: 1/9/2024

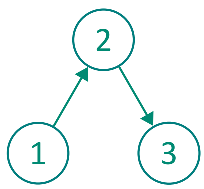
A graph is a pair (V, E) where V is a set of vertices and E is a set of unordered pairs of elements of V called edges. For $u, v \in V$, we say u and v are adjacent if $\{u, v\} \in E$.

For example: $G = (\{1, 2, 3\}, \{\{1, 2\}, \{2, 3\}\})$



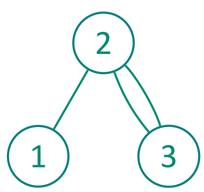
A directed graph (a.k.a a digraph) is a pair (V, E) where V is a set of vertices and E is a set of ordered pairs of elements of V.

For example: $G = (\{1, 2, 3\}, \{(1, 2), (2, 3)\})$



A multigraph is a pair (V, E) where V is a set of vertices and E is a multiset of unordered pairs of elements of V.

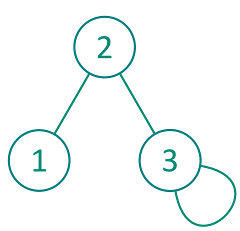
For example: $G = (\{1, 2, 3\}, \{\{1, 2\}, \{2, 3\}, \{2, 3\}\})$



A pseudograph is like a graph and multigraph except that the pairs in E are multisets.

Essentially, an element $\{a, a\}$ can belong to E in a pseudograph. This type of edge is called a loop.

For example: $G = (\{1, 2, 3\}, \{\{1, 2\}, \{2, 3\}, \{3, 3\}\})$



If $G = (V, E)$ and $v \in V$, the neighborhood of v is $N_G(v) = \{w \in V \mid \{v, w\} \in E\}$.

The degree of v is $d_G(v) = |N_G(v)|$. Or in other words, v's degree is equal to the number of edges connecting to v.

The Handshaking lemma states that for any graph (V, E):

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

The reason for this is that each edge increments the degrees of exactly two vertices. So the above sum counts every edge twice.

Lemma: Every graph has an even number of vertices with odd degrees.

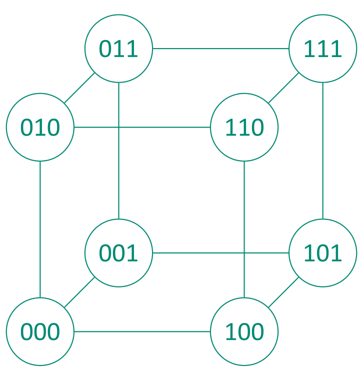
Proof: We can split the vertices of any graph into two categories: those with odd degrees, and those with even degrees.

Now recall that an even number plus an even number always equals an even number, as does an odd number plus an odd number. However, an odd number plus an even numbers equals an odd number. Based on this fact, we can guarentee that the sum of even degrees in any graph is even. And since the sum of even degrees plus the sum of odd degrees must be even as it equals $2|E|$ by the Handshaking lemma, we thus know that the sum of odd degrees must be even. Hence, it must be the case that there are an even number of vertices with odd degree because otherwise the sum of their degrees won't be even.

A graph is called r-regular if all of its vertices have degree r.

Note that the number of edges in any n-vertex r-regular graph is $\dfrac{rn}{2}$.

An r-dimensional cube graph, denoted as Q_r, is a graph such that $V(\mathrm{Q}_r)$, the set of vertices in Q_r, is equal to the set of binary strings of length r; and $E(\mathrm{Q}_r)$, the set of edges in Q_r, is equal to the set of pairs of binary strings which differ in only one position.



$$|V(\mathrm{Q}_r)| = 2^r$$

$$|E(\mathrm{Q}_r)| = \tfrac{2^r r}{2} = 2^{r-1}r$$

Note that Q_r is r-regular.

If $G = (V, E)$, then $H = (W, F)$ is a subgraph of G if $W \subseteq V$ and $F \subseteq E$.

If $W = V$, then H is a spanning subgraph of G (meaning that H has the same vertices as G but is lacking some of G's edges)

We define subtracting a set of vertices from a graph as follows:
For $G = (V, E)$ and $X \subset V$, we define...

$$G - X = (V \setminus X, \{\{u, v\} \in E \mid \{u, v\} \cap X = \emptyset\})$$

We define subtracting a set of edges from a graph as follows:
For $G = (V, E)$ and $L \subset E$, we define...

$$G - L = (V, E \setminus L)$$

Lecture 2: 1/11/2024

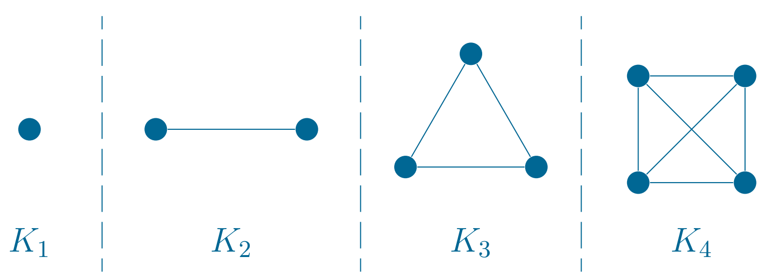
We shall notate that H is a subgraph of G by writing $H \subseteq G$.

An induced subgraph of $G = (V, E)$ is a subgraph $G[X] = G - (V \setminus X)$ where $X \subseteq V$. Alternatively, this is called the subgraph induced by X.

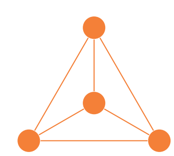
Given $G = (V, E)$ and $F \subseteq E$, the subgraph spanned by F is the subgraph whose edge set is F and whose vertex set is $\bigcup_{e \in F} e$.

Here are some basic classes of graphs:

- Complete graphs / cliques, denoted K_n, are graphs where every possible edge is present between n vertices.



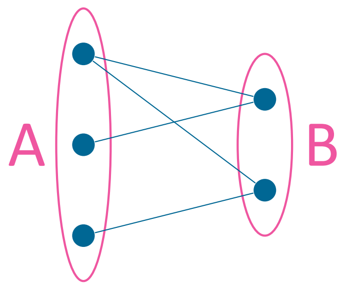
Note we can also draw K_4 such that there are no edge interceptions as follows:



$$|V(K_n)| = n$$

$$|E(K_n)| = \binom{n}{2} = \frac{n(n-1)}{2}$$

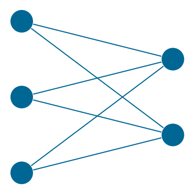
- A graph $G = (V, E)$ is bipartite if there exists a partition (A, B) of V such that every edge in E has one end in A and one end in B.



The partition (A, B) is called the bipartition of G. Then A and B are called the parts of G.

- A Complete bipartite graphs $K_{s,t}$, is the bipartite graph with parts A and B where $|A| = s$, $|B| = t$, and all possible edges between A and B exist.

For example, $K_{3,2}$:

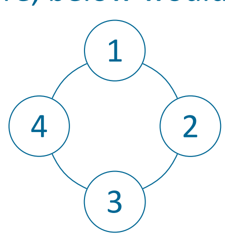


- A path P_k of length k has a vertex set $V = \{v_1, v_2, \ldots, v_k, v_{k+1}\}$ and an edge set $E = \{\{v_1, v_2\}, \{v_2, v_3\}, \ldots, \{v_{k-1}, v_k\}, \{v_k, v_{k+1}\}\}$.

Note that $|V(P_k)| = k + 1$ and $|E(P_k)| = k$.
Therefore, below would be P_3...

- A cycle C_k of length k has a vertex set $V = \{v_1, v_2, \ldots, v_k\}$ and an edge set $E = \{\{v_1, v_2\}, \{v_2, v_3\}, \ldots, \{v_{k-1}, v_k\}, \{v_k, v_1\}\}$.

Note that $|V(C_k)| = k$ and $|E(C_k)| = k$.
Therefore, below would be C_4...



Here is some terminology before the next lemma. For the graph $G = (V, E)$...

- $\delta(G) = \min\{d_G(v) \mid v \in V\}$ is the minimum degree of G.
- $\Delta(G) = \max\{d_G(v) \mid v \in V\}$ is the maximum degree of G.
- The degree sequence of G is the sequence of degrees of vertices G in non-increasing order.

Lemma (part 1): If $G = (V, E)$ is a graph of minimum degree $k \geq 2$, then G contains a cycle of length at least $k + 1$.

Proof: Let P be a longest possible path in G, say:

$$V(P) = \{v_1, v_2, \ldots, v_r\}$$

Then $N(v_r) \subseteq V(P)$. After all, if this were not the case, we'd be able to extend the path to the vertex in $N(v_r)$ but not in $V(P)$, thus contradicting the fact that P is a longest path.

Let v_i be the first neighbor of v_r along the path from v_1 to v_r. Then $\{v_i, v_{i+1}, \ldots, v_r\}$ are the vertices of a cycle C.

Now note that because $N(v_r) \subseteq P$ and v_i was the first element in the path P to belong to $N(v_r)$, we know that C contains all the elements of P that $N(v_r)$ also has. So, $N(v_r) \subseteq C$.

But now note that $|N(v_r)| \geq \delta(G) = k$. Plus, v_r itself is not in $N(v_r)$. Combining these facts together, we can say that the cycle C has at least $k + 1$ vertices.

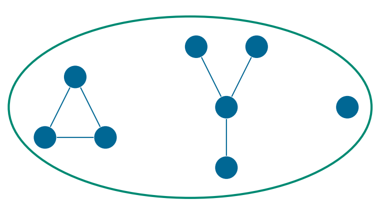
Lemma (part 2): The cycle length $k + 1$ is the longest we can guarentee based on the minimum degree of the graph being k.

Proof: Take the graph K_{k+1} which has a minimum degree k. Obviously, the longest cycle in K_{k+1} is the cycle containing all $k + 1$ elements of K_{k+1}. Thus, we have shown that there are graphs with minimum degree k which don't have cycles of length greater than $k + 1$.

A connected graph is a graph in which any two vertices are the ends of a path.

The components of a graph are the maximal connected subgraphs. For example:

Let us define G as:



As can be seen, G has three components.

A tree is a connected graph with no cycles (a.k.a it is acyclic).

Some examples of small trees include: K_1, K_2, $K_{1,2}$, P_3, and $K_{1,3}$.

Lemma: Every tree with n vertices has exactly $n-1$ edges.

Proof: We shall proceed by induction.

If $n=1$, the tree is K_1, meaning that it has $0=n-1$ edges.

Now assume the lemma is true for all trees with n vertices, and let T be a tree with $n+1$ vertices. Then, we shall remove a vertex v of T with degree 1. (Note that we know such a vertex must exist since otherwise the minimum degree of T would be at least 2 and that would guarentee a cycle exists of at least length 3. This of course contradicts the fact that T is acyclic.)

Then $T-\{v\}$ is a tree with n vertices as it must be acyclic and connected. So by induction it has $n-1$ edges. And because v has degree 1, we know that
$|E(T)| = 1 + |E(T-\{v\})| = 1 + (n-1) = n$.

Lemma: Any connected graph with finite vertices has a spanning tree.

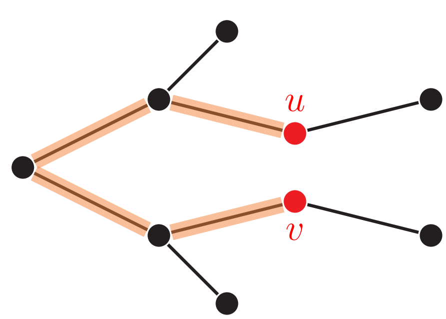
Proof:

Firstly consider the case that the graph G has no cycle. Then, it is a tree by definition.

Now, consider if G has a cycle C. Then for any edge $e \in E(C)$, we have that $G-\{e\}$ is still connected. So, we can now go back to the top of the proof and ask: does $G-\{e\}$ have any cycles? We can repeatedly do this until the graph has no cycles since taking away edges does not remove any vertices.

This actually acts as an algorithm for finding a spanning tree of any connected graph.

If u and v are two vertices in a connected graph, the distance from u to v is the length of a shortest path with ends at u and v.



Let $d_G(u,v)$ be the distance between u and v.

Distance is a metric, meaning:

1. $d_G(u,v) = 0 \iff u = v$
2. $d_G(u,v) = d_G(v,u)$
3. $\forall w \in V, \;\; d_G(u,v) \leq d_G(u,w) + d_G(w,v)$

The diameter of a connected graph G is the maximum distance between any two vertices of G. Or in other words, $\max\{d_G(u,v) \mid u,v \in V(G)\}$.

The radius of G is equal to $\min\{\max\{d_g(u,v) \mid u \in V(G)\} \mid v \in V(G)\}$. What that means is that the radius of G measures the smallest distance path one could limit themselves to drawing while still being able to have that path have one end at some fixed vertex and its other end at any arbitrary vertex in the graph.

Examples:

1. The radius of K_n is 1. The diameter of K_n is n.

2. The diameter of P_k is k. The radius, can be computed as follows:

   The middle vertex of a path will have the fastest access to either end of the path. So, we shall measure the radius from the vertex: $v_{\lceil \frac{k+1}{2} \rceil}$. Then, we can see that v_{k+1} is going to be a farthest element from $v_{\lceil \frac{k+1}{2} \rceil}$. So the radius of P_k equals $k + \lceil \frac{k+1}{2} \rceil$.

   Now you can consider what happens when k is even and odd. But what's important is that it works out that the radius is $\lceil \frac{k}{2} \rceil$.

We can use a search tree to more generally find the radii and diameters of graphs.

Breadth-First-Search

Here's how to find a spanning tree in a connected graph with a root vertex v such that the tree "preserves" all distances from v. (This tree is called a BFS tree).

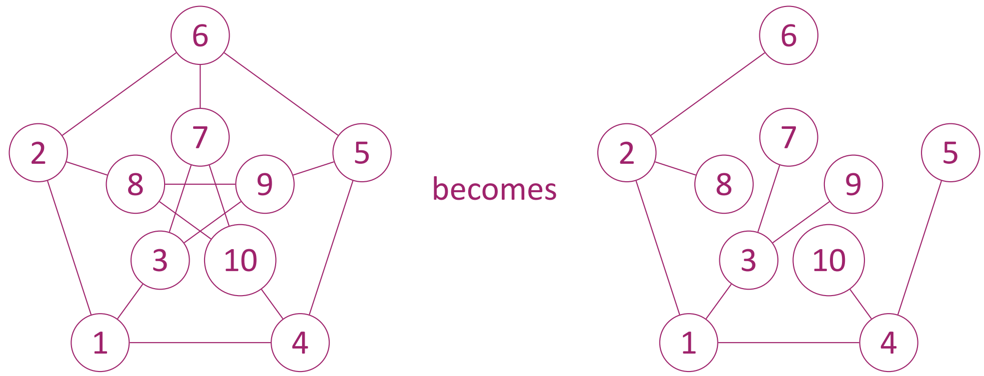
Let G be a connected graph and let $(v_1, v_2, v_3, \ldots, v_n)$ be any ordering of the vertices of G.

Pick a vertex $v = v_1$ to be the root of the BFS tree.

Now, at any stage in constructing this tree, we will have a vertex set $V(T) = \{v_1, v_2 \ldots, v_k\}$ (when we first start, $V(T)$ will only contain v_0. So don't worry about that). Now if $V(T) = V(G)$ we can stop. Otherwise though, we can say that there is a smallest integer i such that for $v_i \in V(T), \;\; N(v_i) \setminus V(T) \neq \emptyset$. Choose v_{k+1} to be the smallest neighbor (by the ordering of $V(G)$) of v_i not in T and add the edge $\{v_i, v_{k+1}\}$ to T. Then we repeat this paragraph.

Beware the ordering we are creating in our tree will often be different from the order of the graph you started with.

Worked Demonstration:



Properties of BFS:

- If the root is v, then $d_T(v, w) = d_G(v, w)$. In otherwords, a BFS tree preserves distances from its root.
- The Tree with root v has layers $N_i(v) = \{w \in V(G) \mid d_G(v, w) = i\}$. Furthermore all edges in the original graph either stay inside a single layer $N_i(v)$ or go between adjacent layers (i.e. from $N_i(v)$ to $N_{i+1}(v)$).

  If an edge did "jump over" a layer, that violate the fact that distance is a metric.
- The diameter of G equals the maximum number of layers of all BFS trees (not including the 0-layer).
- The radius of G equals the minimum number of layers of all BFS trees (also not including the 0-layer).

Lecture 3: 1/16/2024

Note that a tree is "minimally connecting" as subtracting any edge from a tree will produce a disconnected graph.

We know this is the case because if we could remove an edge and still have the graph be connected, then that would imply the existence of a path between two neighboring vertices that doesn't go through their shared edge. But then, we'd be able to make a cycle subgraph by adding their shared edge to that path.

Depth-First-Search

Here is alternate algorithm for generating a spanning tree of a connected graph. A resulting tree of this algorithm is called a DFS tree.

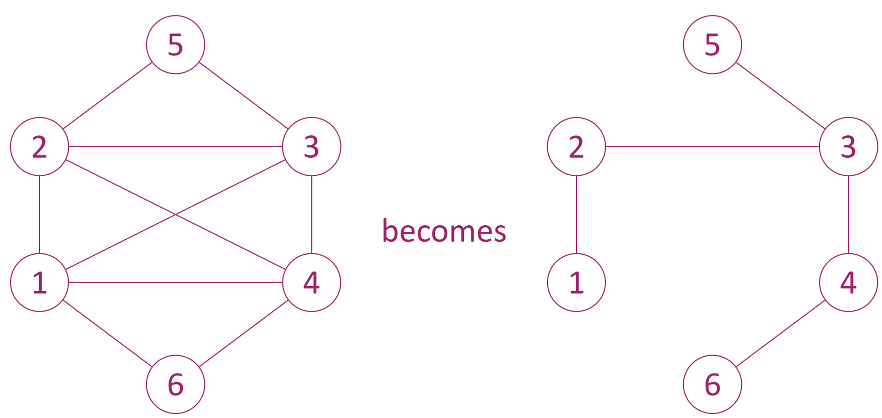
Let G be a connected graph and let $(v_1, v_2, \ldots, v_n)$ be any ordering of the vertices of G.

Pick a vertex $v = v_1$ to be the root of the DFS-tree.

Now, at any stage in constructing this tree, we will have a vertex set $V(T) = \{v_1, v_2, \ldots, v_k\}$. If $V(T) = V(G)$, we can stop. Otherwise though, we select i to be the largest integer such that for $v_i \in V(T)$, $N(v_i) \setminus V(T) \neq \emptyset$. Then, choose v_{k+1} to be the smallest neighbor (by the ordering of $V(G)$) of v_i not in $V(T)$ and add the edge $\{v_i, v_{k+1}\}$ to T. Then we repeat this paragraph.

Once again beware the ordering we are creating in our tree will typically be different from the order of the graph you started with.

Worked Demonstration:



Theorem: A graph is bipartite if and only if it contains no odd cycles.

Proof:

($\Longrightarrow$) First note that an odd cycle isn't bipartite. Thus, any graph containing an odd cycle is not bipartite.

($\Longleftarrow$) Now supposed we are given some graph G with no odd cycles.
Then, assuming G is connected (if G isn't connected, we can break G up into its component subgraphs and do this process for each component), we can construct a BFS-tree in G rooted at some $v \in V(G)$. Let us name this tree T.

Now as noted before, T will have layers L_i where each $L_i = \{u \in V(G) \mid d_G(v,u) = i\}$. Using those layers, we can partition T into two subsets A and B where A is the union of all L_i where i is even and B is the union of all L_j where j is odd. So, T is clearly bipartite.

Now, let's reinsert the removed edges from G back into T. Note that for each re-inserted edge e, it must be the case that either e is a subset of some L_i or that e goes between some L_i and L_{i+1}. Importantly, edges of the latter case do not violate our partition. So, if all the edges in $E(G) \setminus E(T)$ go between layers, then we can conclude that G is definitely bipartite just like T.

With that, we now intend to show that an edge G having an edge belonging to a single layer L_i guarentees that G contains an odd cycle.

Assume the graph G has an edge $\{u, w\} \subseteq L_i$ where L_i is the ith layer of a BFS tree rooted at v. Then, we know that there exists a path P_1 contained in that BFS tree going from v to u and a path P_2 contained in that BFS tree going from v to w. In order to draw a cycle from this information, let x be the vertex of some L_j such that $x \in V(P_1)$, $x \in V(P_2)$, and j is as large as possible. That way, by defining the subpaths ${P_1}'$ going from x to u and ${P_2}'$ going from x to w, we can get the following cyclic subgraph of G :

$$C = (V({P_1}') \cup V({P_2}'), E({P_1}') \cup E({P_2}') \cup \{u, w\})$$

However, now note that $|E({P_1}')| = |E({P_2}')| = i - j$. Hence, $|E(C)| = 2(i - j) + 1$, which in turn means that C has an odd number of edges. So, we have shown that if a graph G contains an edge within a single layer L_i, then we can give an example of an odd cycle within G.

So in conclusion, if we assume G has no odd cycles, then G can't have any edges which are subsets of a single layer L_i. But that means that every edge in G respects the partition we made to show that T is bipartite. So, G must also be bipartite with the same partition as T. ■

A Hamiltonian cycle is a spanning cycle of a graph. We say a graph is Hamiltonian if it contains such a cycle.

A Hamiltonian path is a spanning path of a graph. We say a graph is traceable if it has a hamiltonian path.

A walk is a sequence of vertices and edges: i.e. $(v_1, \{v_1, v_2\}, v_2, \{v_2, v_3\}, \ldots)$

Note that a walk can go over the same edge or vertex multiple times.

A trail is a walk with no repeated edge.

Interestingly, all paths are trails and all trails are walks. So a trail is kind of a middle concept between being a walk or a path.

A tour is a trail with the same first and last vertex.

So, all cycles are tours and all tours are walks.

An Eulerian tour of a graph is a tour which contains all the edges of the graph.

For context, the name Eulerian is in honor of Leonhard Euler because he was the first mathematician to ask when a graph would have an Eulerian tour (look up the Seven Bridges of Königsburg problem).

We call a graph Eulerian if for every vertex v in the graph: $d(v)$ is even.

Lecture 4: 1/18/2024

Before finding conditions for the existence of an eulerian tour of a graph, let's establish some terminology for digraphs so that we can study Eulerian tours in digraphs as well.

Firstly, a walk, trail, and tour are defined almost identically in a digraph as in a graph. The one difference is that given some edge (u, v) of a digraph, a walk, trail, and path are only allowed to traverse that edge going from u to v.

Given a digraph (V, E), the "out" and "in" neighborhoods of a vertex $v \in V$ are:

out: $N^+(v) = \{w \in V \mid (v, w) \in E\}$

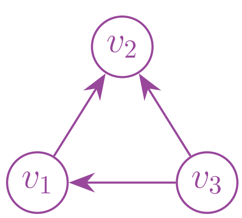
in: $N^-(v) = \{w \in V \mid (w, v) \in E\}$

Similarly, the out-degree and in-degree of $v \in V$ are:

out: $d^+(v) = |N^+(v)|$

in: $d^-(v) = |N^-(v)|$

A digraph is called Eulerian if for each vertex v , $d^+(v) = d^-(v)$.



$d^+(v_1) = 1 \quad d^+(v_2) = 0 \quad d^+(v_3) = 2$

$d^-(v_1) = 1 \quad d^-(v_2) = 2 \quad d^-(v_3) = 0$

This digraph is not Eulerian.

An orientation of a graph G is a digraph with the same vertices as G but where each $\{u, v\} \in E(G)$ is replaced with either (u, v) or (v, u).

The underlying graph (or multigraph) of a digraph G is the graph (or multigraph) such that $\{u, v\}$ is an edge whenever $(u, v) \in G$.

Theorem:

1. A graph has an Eulerian tour if and only if it is connected and Eulerian.
2. A digraph has an Eulerian tour if and only if it has a connected underlying graph and if it is Eulerian.

Proof of statement 1:

($\Longrightarrow$) If G has an Eulerian tour $v_0e_0v_1e_1v_2e_2\ldots v_ke_kv_0$, then for any v_i, the tour has to use an edge into v_i and an edge out of v_i each time it visits v_i. So, $d(v_i)$ is even for all i.

($\Longleftarrow$) Now suppose G is connected and all vertices have even degree. Then let $T = v_0e_0v_1e_1v_2e_2\ldots v_le_lv_{l+1}$ be a longest trail in G.

If $v_{l+1} \neq v_0$, then we know that v_{l+1} has an odd degree in T as the trail goes into v_{l+1} and doesn't leave. However, because we assumed that all vertices in G have an even degree, we know there must be an even number of edges coming out of v_{l+1}. So, we can add another edge to our trail to get a longer trail. But this contradicts our assumption that T is a longest trail in G. Hence, we conclude that $v_{l+1} = v_0$, meaning T is a tour.

Now consider if T is not an Eulerian tour. In that case, there is an edge of G not in T. Additionally, because G is connected, we know that that edge will have the form $e = \{v_i, w\}$. So now consider a new trail: T' defined as $v_ie_iv_{i+1}e_{i+1}\ldots v_0e_0v_1e_1\ldots v_iew$. Importantly, T' is a longer trail than T. So we have a contradiction as T is not a longest trail.

Therefore, the longest trail T in G must be an Eulerian tour.

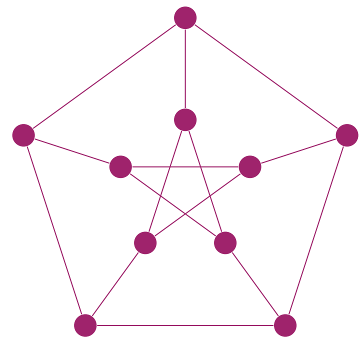
The proof of statement 2 is nearly identical.

Note that this proof can be interpretted as giving an algorithm for finding an Eulerian tour.

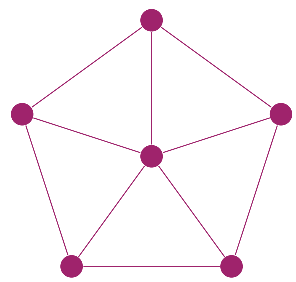
1. Make a trail T.
2. Add edges to T until you get stuck at a vertex. Then you know that your trail forms a tour.
3. If T is not an Eulerian tour, then going by the steps in the proof above, define T'. Then do step 2. on T'.
4. If T is an Eulerian tour, you're done.

A harder problem is whether a graph has a Hamiltonian (spanning) cycle or not.

This graph does not have a Hamiltonian cycle.



This graph does have a Hamiltonian cycle.



Dirac's Theorem: Let $n \geq 3$ and let G be an n-vertex graph of minimum degree at least $\frac{n}{2}$. Then G is hamiltonian.

Proof:
Suppose the theorem is false. Let G be a counter-example with as many edges as possible (a maximal counter-example). Then, we know there exists an edge $\{u, v\} \notin E(G)$ as G cannot equal K_n since K_n is hamiltonian. Furthermore, we know that $G + \{u, v\}$ is hamiltonian since G was maximal. So, there is a hamiltonian cycle $C \subseteq G + \{u, v\}$ using the edge $\{u, v\}$. This in turn means that $C - \{u, v\}$ is a hamiltonian path belonging to G.

Let $P = v_1 e_1 v_2 e_2 \ldots e_n v_n$ be the hamiltonian path in G from u to v and let $N^+(w)$ denote the set of vertices immediately following a neighbor of w on the path P. In other words: $N^+(w) = \{v_{i+1} \mid v_i \in N_G(w)\}$.

By the theorem's assumption about the minimum degree of the graph, we know that $|N_G(u)| \geq \frac{n}{2}$. Meanwhile on the other end of P, since $v = v_n \notin N_G(v)$, we know that every neighbor of v has an element following it on the path. So $|N^+(v)| \geq \frac{n}{2}$. Thus $|N_G(u)| + |N^+(v)| \geq n$. But now note that u does not belong to either of the above sets. So $|N_G(u) \cup N^+(v)| \leq n - 1$. As a consequence of this, $|N_G(u) \cap N^+(v)| > 0$.

Let v_i be a vertex belonging to $(N_G(v) \cap N^+(u))$. Then we can draw a cycle in G visiting all the vertices of G in the following order:

$$v_1, v_2, \ldots, v_{i-1}, v_n, v_{n-1}, \ldots, v_i, v_1$$

This contradicts our assumption that G would be a counter example and thus not Hamiltonian. So, we assume no such counter example exists.

We can also show that Dirac's Theorem gives the best possible minimum degree for a graph to be guarenteably Hamiltonian. Consider a graph G containing two copies of K_m sharing exactly one vertex. In that case, $n = |V(G)| = 2m - 1$ and $\delta(G) = m - 1 = \frac{n-1}{2}$. However, this graph does not have a spanning subcycle as any spanning cycle would have to cross that shared vertex twice.

Let P be a longest path in G going from a vertex u to a vertex v. Additionally, for any $w \in V(P)$, let w^+ be the vertex following w as one travels from u to v along P. Importantly, since P is a longest path, we know that $(N_G(u) \cup N_G(v)) \subseteq V(P)$. Furthermore, for each $w \in N_G(v)$, we can define $Q = P - \{w, w^+\} + \{v, w\}$ where Q is a longest path of G going from u to w^+ instead of u to v. We call Q a rotation of P at v.

Pósa's Rotation Lemma: Suppose G is a graph and for every $S \subseteq V(G)$ with $|S| \leq t$, $|N(S)| > 2|S|$. Then G contains a path of length $3t + 1$.

$N(S)$ is referring to the union of the neighborhoods of each $v \in S$ minus any vertices in S.

Proof:
Let P be a longest path ending at a vertex v. Also, let S be the set of end vertices of all possible longest paths that could be obtained through any number of rotations starting with P. Finally, let S^+ and S^- denoted the vertices of P immediately after and immediately before vertices in S respectively.

Obviously, $|S^+| \leq |S|$ and $|S^-| \leq |S|$ as all vertices except the first and last vertex of P have exactly one vertex before and after them in P. Also note that $N(S) \subseteq S^+ \cup S^-$. This is because if there did exist $w \in N(S)$ such that $w \notin S^+ \cup S^-$, then we would know that no rotation of P made it so that w was not proceeded by w^- and followed by w^+. So, doing a rotation with the vertex w, we would show that either w^+ or w^- belonged to S, thus reaching a contradiction.

Overall, this means that $|N(S)| \leq |S^+ \cup S^-| \leq |S^+| + |S^-| \leq 2|S|$. However, note that by the theorem's assumption about G, we know that $|S| \geq t$ because otherwise we'd have that $2\,|S| < |N(S)|$. So, let T be a subset of S such that $|T| = t$. Then, because T and $N(T)$ are disjoint subsets of $(S \cup N(S))$ which itself is a subset of $V(P)$, we know that $|V(P)| \geq |T| + |N(T)|$. And since $|N(T)| > 2\,|T| = 2t$ by the theorem's assumption about G, we thus can say that $|V(P)| > 3t$. ■

Lecture 5: 1/23/2024

Theorem: If for every set S of vertices in a graph G, we have that $|N(S)| \geq \min\{2\,|S| + 1,\ |V(G) \setminus S|\}$, then G has a hamiltonian path.

Proof:
Once again let us define P as a longest path of G, as well as S as the set of end vertices of all possible rotations of P. Then by the same reasoning as before, we know that $|N(S)| \leq 2\,|S|$. Therefore, since $|N(S)|$ isn't greater than or equal to $2\,|S| + 1$, we know by the assumption of the theorem that $N(S)$ is greater than or equal to $|V(G) \setminus S|$.

Now $S \cup (V(G) \setminus S) = V(G)$ and $(S \cup N(S)) \subseteq V(P)$. Additionally, S and $(V(G) \setminus S)$ are disjoint to each other, as is S and $N(S)$. Thus, we can say that

$$|V(G)| = |S| + |(V(G) \setminus S)| \geq |S| + |N(S)| \leq |V(P)|$$

Therefore the longest path P must cover every vertex of G, meaning that it is a Hamiltonian path. ■

This is what is used to find Hamiltonian paths in random graphs.

A graph is uniquely Hamiltonian if it has exactly one Hamiltonian cycle.

Theorem: If all vertices in a graph G have odd degree, then every edge is in an even number of hamiltonian cycles.

In other words such a graph is not uniquely Hamiltonian.

Proof:
If the graph G in the theorem has no hamiltonian cycles, then we're done. Every edge is in 0 hamiltonian cycles.

Now pick an edge and suppose that there is a Hamiltonian cycle C containing it. We'll call that edge $\{u, v\}$ and let w be the vertex coming before u on C.

Then, let us define a new graph H whose vertices are Hamiltonian paths in G which start with the edge $\{u, v\}$. For example, $(C - \{u, w\}) \in V(H)$. Additionally, let $\{P, Q\}$ be an edge of H if P and Q are rotations of each other.

If $P \in V(H)$ is a hamiltonian path in G ending in a vertex $x \in V(G)$, then:

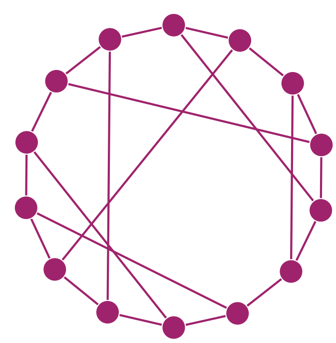
$$d_H(P) = \begin{cases} d_G(x) - 1 & x \notin N_G(u) \\ d_G(x) - 2 & x \in N_G(u) \end{cases}$$

Essentially, for every edge connecting to x except for the one already used by P, there is a rotation of P. We can then say that that rotation is a vertex of H if it includes the edge $\{u, v\}$ (In other words, a rotation including the edge $\{u, x\}$ would not be included in H).

If x is adjacent to u, then $P + \{\{u, x\}\}$ is a hamiltonian cycle containing $\{u, v\}$.

Now here is the clever part: since $d_G(x)$ is assumed to be odd for all $v \in V(G)$, we have that $d_H(P)$ is even if x is not adjacent to u and odd if x is adjacent to u. But now note that every graph has an even number of vertices of odd degree because of the handshaking lemma. So, there must be an even number of paths including the edge $\{u, v\}$ and ending in a vertex x such that x is adjacent to u. Or in other words, G has an even number of hamiltonian cyles containing the edge $\{u, v\}$. ■

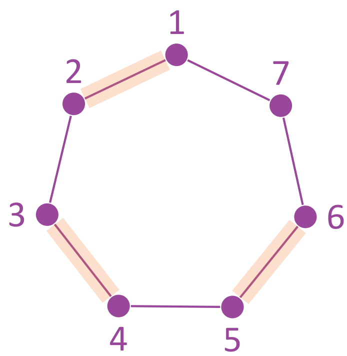
For example, by the above theorem we know that this graph is not uniquely Hamiltonian.



One note about the above theorem is that it can be interpretted as giving an algorithm for finding a second hamiltonian cycle given one hamiltonian cycle in a graph where all degrees are odd.

A matching in a graph is a set of vertex disjoint edges in the graph.

A vertex is saturated by a matching if one of the edges of the matching contains the vertex. Otherwise, we say the vertex is exposed by the matching.



The matching shown to the left is:
$M = \{\{1,2\},\{3,4\},\{5,6\}\}$

In the matching to the left, 7 is exposed and all other vertices are saturated.

A maximum matching in a graph is a matching with the maximum number of edges.

A perfect matching (or 1-factor) is a matching which saturates all the vertices of a graph.

Proposition: for a graph to have a perfect matching, it must have an even number of vertices.

Lecture 6: 1/30/2024

Hall's Theorem: Let G be a bipartite graph with parts A and B. Then G has a matching saturating A if and only if for every set $X \subseteq A$,

$$|N_G(X)| \geq |X|.$$

Note: we refer to the below statement as Hall's condition:
$$\forall X \subseteq A,\ \ |N_G(X)| \geq |X|.$$

Proof:
($\Longrightarrow$) If G has a matching M containing A, then for any $X \subseteq A$, we trivially have that $|N_G(X)| \geq |N_m(X)| = |X|$.

($\Longleftarrow$) To prove the other way, we proceed by induction on A while assuming Hall's condition is true.

Base case: assume that $|A| = 1$. In that case, because Hall's condition is assumed to be true, we know that $|N_G(A)| \geq |A| = 1$. Thus, for some $a \in A$ and $b \in B$, we know there exists an edge $\{a,b\} \in E(G)$. The matching of just that edge saturates A.

Induction step: Assume that Hall's theorem works if $1 \le |A| < n$. Then assume that $|A| = n$. Since we are assuming Hall's condition to be true, we know that for all proper subsets $X \subset A$, either $|N_G(X)| > |X|$ or $|N_G(X)| = |X|$ is true.

- Case 1: For all proper subsets X of A, $|N_G(X)| > |X|$...
  Pick an edge $\{a, b\} \in E(G)$ and consider $H = G - \{a\} - \{b\}$. We know that H will be a bipartite graph with two sets of vertices: $A' = A \setminus \{a\}$, and $B' = B \setminus \{b\}$. Additionally, given any $X \subseteq A'$, we know that $|N_H(X)| \ge |N_G(X)| - 1 \ge |X|$. So, by our inductive hypothesis, H has a matching covering A'. Now add the edge $\{a, b\}$ to this matching to get a matching in G covering A.

- Case 2: There exists a proper subset X of A such that $|N_G(X)| = |X|$...
  In this second case our reasoning for case 1 breaks down because given $X \subseteq A'$, we can no longer guarentee that $|N_G(X)| - 1 \ge |X|$. So, using that same set X, consider G_1 the induced graph of $N_G(X) \cup X$ and G_2 equal to $G - V(G_1)$.

  For any $Y \subseteq X$, we have that $|N_{G_1}(Y)| = |N_G(Y)| \ge |Y|$ due to Hall's condition. Thus, by our inductive hypothesis there exists a matching M_1 saturating X in G_1.

  Additionally, for any $Y \subseteq A \setminus X$, consider $N_G(Y \cup X)$. By assuming Hall's condition, we know that $|N_G(Y \cup X)| \ge |Y \cup X|$. And, because X and Y are disjoint, $|Y \cup X| = |Y| + |X|$. On the other hand, we also know that $N_G(Y \cup X) = N_{G_2}(Y) + N_G(X)$. So, $|N_{G_2}(Y)| + |N_G(X)| \ge |Y| + |X|$. Finally, because $|N_G(X)| = |X|$, we can cancel terms to get that: $|N_{G_2}(Y)| \ge |Y|$. Hence by our inductive hypothesis, we know there exists a matching M_2 saturating $A \setminus X$ in G_2.

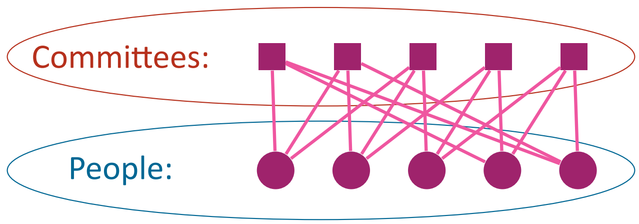
  We now can combine M_1 and M_2 in order to get a matching in G covering A.

Side proposition: if in additional to Hall's condition holding, $|A| = |B|$, then the matching for G generated by the above proof is perfect. Meanwhile, if $|A| \ne |B|$, then it is impossible to make a perfect matching. To prove this, assume without loss of generality that $|A| < |B|$. Then, because $|N_G(B)|$ is a most $|A|$, we know that $|N_G(B)| < |B|$. So no matching can exist covering B.

One application of Hall's Theorem is the "systems of distinct representatives" problem.

Let S_1, S_2, ..., S_n be committees and let $s_i \in S_j$ refer to some person in the jth committee. Then, is it possible to select a representative s_i in each committee S_j so that each s_i is distinct?

To answer this, note that we can represent the above situation as a bipartite graph where vertex set A contains people and vertex set B contains committees. Then for a person $a \in A$ and committee $b \in B$, we add the edge $\{a, b\}$ if person a is in committee b.



Theorem: Committees S_1, S_2, ..., S_n have a system of distinct representatives if and only if for all $X \subseteq \{1, 2, \ldots, n\}$:

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

A 1-factorization of a graph is a collection of pairwise edge-disjoint 1-factors M_1, M_2, ..., M_r such that $E(G) = M_1 \cup M_2 \cup \ldots \cup M_r$.

Theorem: Every r-regular bipartite graph has a 1-factorization if $r \geq 1$.

Proof:

Let $X \subseteq A$. Then the number of edges going out from X to B, written as $e(X, B)$, is equal to $r\,|X|$. Meanwhile, the number of edges going from $N(X)$ back to A is equal to $e(N(X), A) = r\,|N(X)|$.

Now note that every edge counted in $e(X, B)$ is also counted in $e(N(X), A)$. Therefore, $r\,|X| = e(X, B) \leq e(N(X), A) = r\,|N(X)|$. And finally as $r > 0$, we have that $|X| \leq |N(X)|$. Thus Hall's condition is satisfied.

Because Hall's condition is satisfied, we know by Hall's theorem that there is a matching M_1 saturating A. Additionally, because the total number of edges in the graph is equal to $e(A, B) = r\,|A| = r\,|B|$, we know that $|A| = |B|$. So the matching we found that saturates A also saturates B.

Now, that we've a found a 1-factor M_1, let's now subtract M_1 from our bipartite graph. Then, the resulting subgraph is $(r - 1)$ regular. So, we can apply the above reasoning again and again to get more 1-factors. And importantly, all of these 1-factors must be disjoint.

After r iterations, our bipartite graph will have no edges left and we will have created a collection of r many 1-factors whose union is the original edge set. ■

Lecture 7: 2/1/2024

For a graph G, let $\mathrm{odd}(G)$ denote the number of components of G with an odd number of vertices.

If G has a perfect matching, then $\mathrm{odd}(G) = 0$. Furthermore, for any set $S \subseteq V(G)$:

$$\mathrm{odd}(G - S) \leq |S|$$

To understand why, let ϵ be the minimum number of exposed vertices in $G - S$ and consider that:

- We know $\epsilon \leq |S|$ because if M is a perfect matching of G, then the number of exposed edges by the matching: $M \cap E(G - S)$ is at most equal to $|S|$.
- Meanwhile, because no odd component of $G - S$ can have a perfect matching, we know that each odd component of $G - S$ must contribute at least one exposed vertex to an optimal matching of $G - S$. So $\mathrm{odd}(G - S) \leq \epsilon$.

Combining the two bounds, we have $\mathrm{odd}(G - S) \leq \epsilon \leq |S|$. Therefore, $\mathrm{odd}(G - S) \leq |S|$.

I do not understand how the professor saw this as an obviously apparent truth that didn't need a proof but whatever.

This condition is called Tutte's condition. As shown above, it is a necessary condition for a graph to have a perfect matching. However, it happens to also be a sufficient condition.

Theorem: A graph G has a perfect matching if and only if $\mathrm{odd}(G - S) \leq S$ for every set $S \subseteq V(G)$.

Proof:
We already showed that Tutte's condition was necessary for a graph to have a perfect matching. So, what we need to show now is that assuming Tutte's condition holds, you can construct a perfect matching.

We shall proceed by induction. Firstly note that if $S = \emptyset$, then there must be zero odd components. Therefore, Tutte's condition already limits us to focuing on graphs with an even total number of vertices. So, let our base case be when $|V(G)| = 2$. In that case $G = K_2$ and so G has a perfect matching.

Now suppose $V(G) > 2$ and that our theorem holds for any graph with less than $V(G)$ vertices.

Let S be the largest subset of $V(G)$ such that $|S| = \text{odd}(G - S)$. We know that S is nonempty because if $|S|$ equals 1, the equality must hold.

> $|V(G)|$ having to be even implies that $|V(G)| - 1$ is odd. So $G - S$ must have an odd number of odd component. But by Tutte's condition, we know the number of odd components is at most 1. Hence, equality must hold when $|S| = 1$.

This guarentees that $V(G - S) \leq V(G)$.

Let's now consider any even component of $G - S$ which we shall denote H. We claim that H has a perfect matching.

> Consider any $R \subseteq V(H)$. Because R only contains vertices from an even component of $G - S$, we have that every odd component in $G - S$ is also in $G - (R \cup S)$. This combined with Tutte's condition tells us that:
>
> $$\text{odd}(H - R) + \text{odd}(G - S) = \text{odd}(G - (R \cup S)) \leq |S| + |R|$$
>
> However, since we specified that $|S| = \text{odd}(G - S)$, we can cancel out terms to get that $\text{odd}(H - R) \leq |R|$. So Tutte's condition holds for H. And since $V(H) < V(G)$, we can thus conclude by our inductive hypothesis that H has a perfect matching.

Meanwhile, consider any odd component of $G - S$ which we shall denote F. We claim that for any $v \in V(F)$, we know that $F' = F - \{v\}$ has a perfect matching.

> Because $V(F') < V(G)$, we know by our inductive hypothesis that if F' does not have a perfect matching, then Tutte's condition must not hold. Or in other words, there exists $Q \subseteq V(F')$ such that $\text{odd}(F' - Q) > |Q|$.
>
> Now note that for any $R \subseteq V(F)$, because $|V(F)|$ is odd ($\equiv 1 \mod 2$):
>
> - $|R| \overset{2}{\equiv} 0 \implies |V(F - R)| \overset{2}{\equiv} 1 \implies \text{odd}(F - R) \overset{2}{\equiv} 1$
> - $|R| \overset{2}{\equiv} 1 \implies |V(F - R)| \overset{2}{\equiv} 0 \implies \text{odd}(F - R) \overset{2}{\equiv} 0$
>
> So $\text{odd}(F - R) + |R| \overset{2}{\equiv} |V(F)| \overset{2}{\equiv} 1$.
>
> This means that $\text{odd}(F - (Q \cup \{v\})) + |Q| + 1 \overset{2}{\equiv} 1$, which in turn says that $\text{odd}(F' - Q) + |Q| \overset{2}{\equiv} 0$. So both $\text{odd}(F' - Q)$ and $|Q|$ must have the same parity, meaning that $\text{odd}(F' - Q) > |Q| \Rightarrow \text{odd}(F' - Q) \geq |Q| + 2$.
>
> Also observe that $\text{odd}(G - S \cup \{v\} \cup Q) = \text{odd}(G - S) - 1 + \text{odd}(F' - Q)$. This is because every odd component of $G - S$ except for F is also an odd component of $G - S \cup \{v\} \cup Q$.
>
> Therefore we can say by Tutte's condition that:
> $|S| + |Q| + 1 \geq \text{odd}(G - S) - 1 + \text{odd}(F' - Q) \geq \text{odd}(G - S) + |Q| + 1$.

Therefore we've shown that $\text{odd}(G - S \cup \{v\} \cup Q) = |S \cup \{v\} \cup Q|$. However, this contradicts our previous assertion that S is the largest subset of $V(G)$ such that $\text{odd}(G - S) = |S|$. So we have shown that Q cannot exist, meaning that F' has a perfect matching.

Finally, let C be the set of vertices consisting of one vertex from each odd component of $G - S$ such that each vertex in C has an edge connecting to S. Then consider the bipartite graph $G(S, C)$ consisting of every edge in G going between C and S. We claim that this bipartite graph has a perfect matching.

For every $X \subseteq C$, we know that removing $N(X)$ will make the components which each $x \in X = C$ belong to odd. So $|X| = \text{odd}(G - N(X))$. And by Tutte's condition: $\text{odd}(G - N(X)) \leq N(X)$. So $|X| \leq |N(X)|$, meaning that by Hall's theorem, $G(S, C)$ has a perfect matching.

Combining all the perfect matchings we made for $G(S, C)$ and each H and F', we get a perfect matching for all G. ■

A bridge is an edge which when removed increases the number of components of a graph. Meanwhile, a cubic graph is a graph that is 3-regular.

Petersen's Theorem: Every bridgeless cubic graph G has a perfect matching.

We check that for all $S \subseteq V(G)$, we have that $\text{odd}(G - S) \leq |S|$.

By handshaking lemma, each component of G must have an even number of vertices. Thus, if $S = \emptyset$, then $\text{odd}(G - S) = \text{odd}(G) = 0$.

Now suppose if $S \neq \emptyset$. Then each component of $G - S$ sends at least two edges to S since G is bridgeless. However, consider a hypothetical odd component H of $G - S$. Because the sum of degrees of H must be even, we know that H must send an odd number of edges to S. Therefore if every component of $G - S$ was odd, then the number of edges coming into S must be at least $3(\text{odd}(G - S))$. But, because G is 3-regular, the max number of edges that S could possibly accept is $3|S|$. So, $\text{odd}(G - S) \leq |S|$.

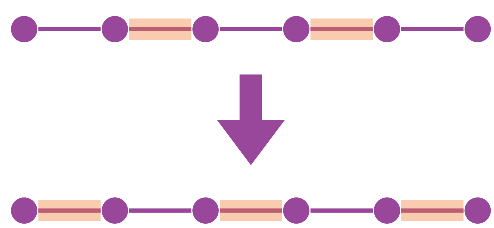
Matching Algorithms:

Let G be a graph and M a matching in G.

- An M-alternating path is a path whose edges are alternatly in and not in M.
- An M-augmenting path is an alternating path whose first and last edge are not in M.

Theorem: If G is a graph and M is a matching in G, then M is a maximum matching if and only there is no augmenting path between two vertices not covered by M.

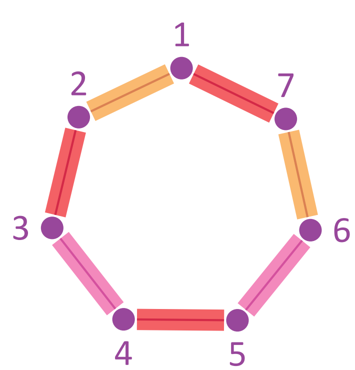
Proof:
($\Longrightarrow$) the contrapositive of this is easy to show. If such an augmenting path does exist, then we can draw a larger matching by doing this:



($\Longleftarrow$) we're not rigorously showing this direction for some reason in this class...

Edge Coloring:

A proper edge coloring of a graph G is a map $c : E(G) \longrightarrow S$, where S is a set such that $c(e) \neq c(f)$ whenever $e \cap f \neq \emptyset$.



This is an edge coloring of C_7.
$S = \{$■, ■, ■$\}$ and c maps each each edge of C_7 to an element of S as shown to the side.

The edge-chromatic number $\chi'(G)$ is the minimum k for which G has a proper k-edge-coloring.

Observe that $\chi'(G) \geq \Delta(G)$ (the maximum degree of G). After all, each edge coming off the vertex with the max degree have to have a different color.

König's Theorem: Let G be a bipartite multigraph. Then $\chi'(G) = \Delta(G)$.

Proof:
By taking two copies of G and adding multiple edges between copies of the same vertex, we can get a $\Delta(G)$-regular-multigraph. Then, by the theorem on page 20 (which I realize we only explicitely proved for normal graphs but all the logic we used should also work for multigraphs...), we know that this new graph has $\Delta(G)$ disjoint perfect matches. Then assign colors to the original edges of G based on which perfect matching that edge wound up in.

A different proof is in the textbook...

Lecture 8: 2/6/2024

We define $\mu(G)$ to be the size of a maximum matching of a graph G. Meanwhile we define $\mathrm{ex}(G)$ to be the number of vertices exposed by a maximum matching of G.

König-Ore Theorem: Let G be a bipartite graph with parts A and B. Additionally, define $\mathrm{ex}(G, A) = |A| - \mu(G)$. In other words, this is the number of vertices in A exposed by a maximum matching. Then:

$$\mathrm{ex}(G, A) = \max_{S \subseteq A} \{|S| - |N(S)|\}$$

Proof:
Denote the right hand side of the above equation by d. Then consider adding d vertices to B which are all adjacent to all vertices of A. Hall's condition is trivially true in this new graph, meaning that this new graph has a matching covering all vertices of A. Therefore, when we remove those added vertices and edges from our matching over A, we will still have a matching covering at least $|A| - d$ vertices in A. So, $\mathrm{ex}(G, A) \leq d$.

On the other hand, if M is a matching of size: $|A| - \mathrm{ex}(G, A)$, meaning that M is a maximum matching, then each $S \subseteq A$ has at least $|S| - \mathrm{ex}(G, A)$ neighbors in B. Or in other words, $|N(S)| \geq |S| - \mathrm{ex}(G, A)$ We know this because Hall's condition applise to A if we remove the exposed vertices in A. Then, by subtracting $\mathrm{ex}(G, A)$ from $|S|$, we are effectively canceling the influence of those exposed points which cause Hall's Theorem to fail. We can rewrite the above formula as $\mathrm{ex}(G, A) \geq |S| - |N(S)|$. And since this is true for all S, we thus know that $d = \max_{S \subseteq A} \{|S| - |N(S)|\} \leq \mathrm{ex}(G, A)$

So in conlusion: $d \leq \mathrm{ex}(G, A) \leq d \Longrightarrow d = \mathrm{ex}(G, A)$.

Here are two other theorems given without proof:

Tutte-Berge Formula: For any multigraph G:

$$\text{ex}(G) = \max_{S \subseteq V(G)} \{\text{odd}(G - S) - |S|\}$$

Vizing's Theorem: For every graph G of maximum degree Δ, either $\chi'(G) = \Delta$ or $\chi'(G) = \Delta + 1$.

If $\chi'(G) = \Delta$, then G is referred to as class 1. Meanwhile, if $\chi'(G) = \Delta + 1$, then G is referred to as class 2.

That said, it is an NP-complete problem to generally determine if a graph is class 1 or not.

An embedding of a graph $G = (V, E)$ is a function $f : V \cap E \longrightarrow \mathbb{R}^2 \cap \mathcal{C}$ (where $\mathcal{C}$ is the set of continuous curves in $\mathbb{R}^2$) such that f is injective, $f(v)$ is a point in $\mathbb{R}^2$ for all $v \in V$, and $f(\{u, v\})$ is a continuous curve in $\mathbb{R}^2$ with ends at $f(u)$ and $f(v)$ for all $\{u, v\} \in \mathbb{R}^2$.

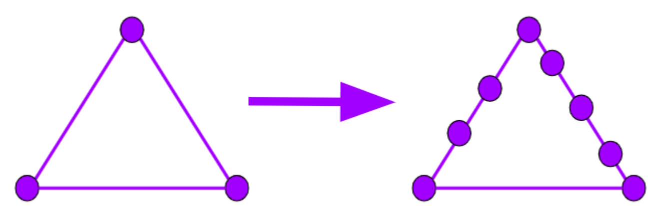
A graph is a planar if we can choose f such that for all distinct $e, e' \in E$, $f(e)$ only intersects $f(e')$ at its end points if at all.

In other words, G is planar if we can draw it without crossings (and we won't get into the topology of what that means...)

A plane graph is a graph which we have embedded in the plane without crossings.

A subdivision of a graph H is a graph obtained by replacing every edge $e = \{u, v\}$ with a path P_e which ends at u and v such that $V(P_e) \cap V(P_f) = e \cap f$ for all $e, f \in E(H)$.

Here is an example of a subdivision:

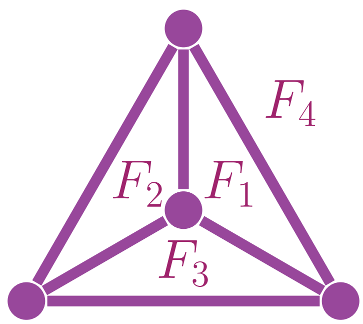


Kuratowski's Theorem: A graph is planar if and only if it does not contain a subdivision of K_5 or $K_{3,3}$.

The proof for this is beyond the scope of this class.

Let G be a plane graph. Then the faces of G are the maximal connected regions of $\mathbb{R}^2 \setminus G$. We refer to the set of faces of G as $F(G)$.

The plane graph below has four faces:

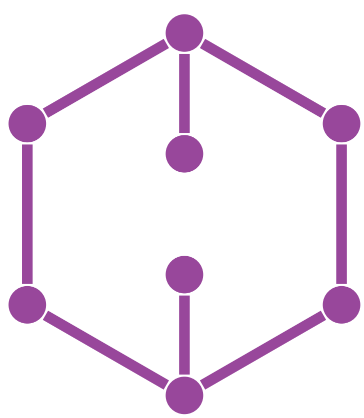


Beware that the same graph can have multiple plane drawings.

If G is connected, then the boundary walk of a face of G is the shortest closed walk which uses every edge on the topological boundary of any face.

The degree of a face is the length of its boundary walk.

The middle face the graph below has degree 10...



The two vertices protruding inside the face have to be traversed twice by any walk going around the boundary of the internal face.

Proposition: The degree of the only face in a plane graph of a tree is twice the tree's number of edges.

Handshaking Lemma: $\sum_{f \in F(G)} \text{degrees}(f) = 2|E(G)|$

Intution why:
Each edge has a face on its left hand and right hand side (it might be the same face on both side). Thus the boundary walk of the face on the edge's left hand side will have to traverse that edge once, whereas the face on the edge's right hand side will also have to traverse that edge once.

Euler's formula: Let G be a connected plane graph. Then:

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

Proof: (we shall prove by induction on the number of edges)

If $|E(G)| = |V(G)| - 1$, then G is a tree and $|F(G)| = 1$. So equality holds.

Now assume that the formula holds if $|E(G)| = |V(G)| + n - 1$ for some integer $n \geq 0$. Then consider a graph G such that $|E(G)| = |V(G)| + n$.
We know that G cannot be tree and so G must contain a cycle C. If $e \in E(C)$, then $G - \{e\}$ is still connected. So we can now apply our inductive hypothesis to say that $|V(G - \{e\})| - |E(G - \{e\})| + |F(G - \{e\})| = 2$.

Note that $F(G - \{e\}) = F(G) - 1$ because by removing e, we removed the boundary seperating two different faces, thus merging them. Additionally, $V(G) = V(G - \{e\})$ and $E(G) = E(G - \{e\}) - 1$. So we can thus conclude that: $|V(G)| - (|E(G)| - 1) + (|F(G) - 1) = 2$. Or in other words:

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

Theorem: Let G be a connected planar graph containing at least one cycle but containing no cycles of length less than g. Then:

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

Proof: We know that the degree of each face $f \in F(G)$ is at least g. Thus, by the handshaking lemma, we get that $g|F(G)| \leq 2|E(G)|$. Now, inserting this into Euler's formula, we get that:

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

This can then be manipulated to get the above formula.

This upper bound is maximized when $g = 3$, thus giving a an upper bound not dependent on g of:

$$|E(G)| \leq 3|V(G)| - 6$$

All in all, this is a neat necessary condition for a graph G to be planar. Beware though that this is not a sufficient condition. Some nonplanar graphs satisfy this inequality.

Some example calculations:

- K_5: The minimum cycle in K_5 is of length 3. So note:
$$\frac{3}{3-2}(|V(G)|-2) = 9 < 20 = |E(K_5)|$$
Thus, we have shown that K_5 can't be planar.

- Petersen Graph: The minimum cycle length in the Petersen graph is 5. So note:
$$\frac{5}{5-2}(|V(G)|-2) = 13\frac{1}{3} < 15 = |E(K_5)|$$
Thus, we have shown that the Petersen graph can't be planar.

Lecture 9: 2/8/2024

A proper (vertex) coloring of G is a map $c : V(G) \longrightarrow S$ where S is a set such that $c(u) \neq c(v)$ if $\{u, v\} \in E(G)$.

The chromatic number $\chi(G)$ is the minimum k for which G has a proper k-coloring.

A graph G is d-degenerate if every induced subgraph of G has a vertex of degree at most d.

Lemma: If G is d-degenerate, then $\chi(G) \leq d + 1$.

Proof:
Firstly, observe that if a graph G is d-degenerate, then any subgraph resulting by removing vertices from G will also be d-degenerate.

Also, observe that for any induced subgraph H of G, we clearly have that $\delta(H) \leq d$. Therefore, assuming H is nonempty, we know there exists a vertex $v \in V(H)$ such that $d(v) \leq d$.

Using the above two observation, we can do induction over a d-degenerate graph as follows:

For each integer $i \in \{1, \ldots, |V(G)|\}$, remove a vertex v_i from $H_i = G - \{v_1, \ldots, v_{i-1}\}$ satisfying the property that $d_{H_i}(v_i) \leq d$.

Doing this, we will eventually be left with an empty graph. This graph obviously has a proper coloring using at most $d + 1$ colors.

Next, consider adding vertices back in. For any $i \in \{1, \ldots, |V(G)|\}$, we shall inductively assume that $G - \{v_1, \ldots, v_{i-1}, v_i\}$ has a proper coloring using less than $d + 1$ colors. Then, because we specified that v_i has at most a degree of d in $G - \{v_1, \ldots, v_{i-1}\}$, we know there is at least one available color to assign v_i. So $G - \{v_1, \ldots, v_{i-1}\}$ has a proper coloring using at most $d + 1$ colors. ■

Brook's Theorem: If G is a connected graph, then $\chi(G) \leq \Delta G$ unless G is an odd cycle or a complete graph.

Proof: (we shall proceed by induction on the number of vertices in G)

To start, consider that $\chi(K_n) = \Delta(K_n) + 1$. Similarly for an odd cycle C, we have that $\chi(C) = \Delta(C) + 1$. Hence, this is why Brook's theorem specifies that G is not one of the above graphs. Additionally, note that if $\Delta G < 3$, then the theorem is obviously true. So, we only need to focus on when $\Delta G \geq 3$.

Now, note that we may assume for all $v \in V(G)$ that $G - \{v\}$ is connected.

After all, if $G - \{v\}$ was not connected, then we could write G as the union of two smaller connected graphs: G_1 and G_2 such that $V(G_1) \cap V(G_2) = \{v\}$. Now if G_i is an odd cycle or a complete graph, then we know $\chi(G_i) = \Delta(G_i) + 1$. However, then $\Delta(G_i) \leq \Delta(G) - 1$ because the degree of the vertex connecting G_i to the rest of G is equal to $\Delta(G_i)$. Meanwhile if G_i is not an odd cycle or complete graph, then we can inductively prove Brook's Theorem for G_i to get that $\chi(G_i) \leq \Delta(G_i) \leq \Delta(G)$. In either case, we will get a coloring over G_i using at most ΔG colors.

So G_1 and G_2 both have colorings with at most $\Delta(G)$ colors. By setting v in both graphs to have the same color, we can then merge the colorings of G_1 and G_2 to get a suitable coloring of G.

Thus, assume that $G - \{v\}$ is connected for all $v \in V(G)$. Then, if $G \neq K_{\Delta(G)+1}$, we can pick vertices v_1, v_{n-1} and v_n such that $\{v_1, v_{n-1}\} \in E(G)$, $\{v_1, v_n\} \in E(G)$, and $\{v_{n-1}, v_n\} \notin E(G)$.

Case 1: Suppose $H = G - \{v_n\} - \{v_{n-1}\}$ is connected. Then, we can order the vertices of H: $v_1, v_2, \ldots, v_{n-2}$ so that for $i \geq 2$, v_i always has at least one neighbor v_j where $j < i$.

This was proven in a homework exercise...

Assign v_n and v_{n-1} the color 1. Then, considering our graph G, since v_i has at most $\Delta(G) - 1$ neighbors v_j where $i < j$, we can iterate from v_{n-2} to v_2, assigning whatever color is available to each vertex. After all, it is guarenteed that we will not have assigned all $\Delta(G)$ colors to the neighborhood of v_i yet since we could have only assigned $\Delta(G) - 1$ colors at most.

If v_i is neighbors with v_n or v_{n+1} then that just means it isn't neighbors with one or two more v_j in our ordering. So, this doesn't affect our conclusion.

Also, the number of colors assigned in $N(v_1)$ is at most $\Delta(G) - 1$ since two of v_1's neighbors are assigned the same color. So we can assign it a remaining color afterwards. Hence G has a proper coloring.

Case 2: Now suppose $H = G - \{v_n\} - \{v_{n+1}\}$ is disconnected. Then we may view G as the union of two smaller connected graphs G_1 and G_2 such that $V(G_1) \cap V(G_2) = \{v_n, v_{n+1}\}$. By the same type of induction that we did earlier, we can prove that G_1 and G_2 can be covered by at most $\Delta(G)$ colors. So, choose a coloring such that v_n and v_{n+1} are assigned the same colors in both G_1 and G_2. Then, we can merge the two colorings to get a proper coloring of G. ■

Note that $\Delta(G)$ can differ greatly from $\chi(G)$. For instance, bipartite graphs are always 2-colorable no matter how many edges they have.

A maximal planar graph is a graph that is planar but the addition of any edge results in a non-planar graph. Similarly, a maximal plane graph is a plane drawing of a maximal planar graph.

Theorem: Every planar graph has a vertex v of degree at most 5.

Proof:

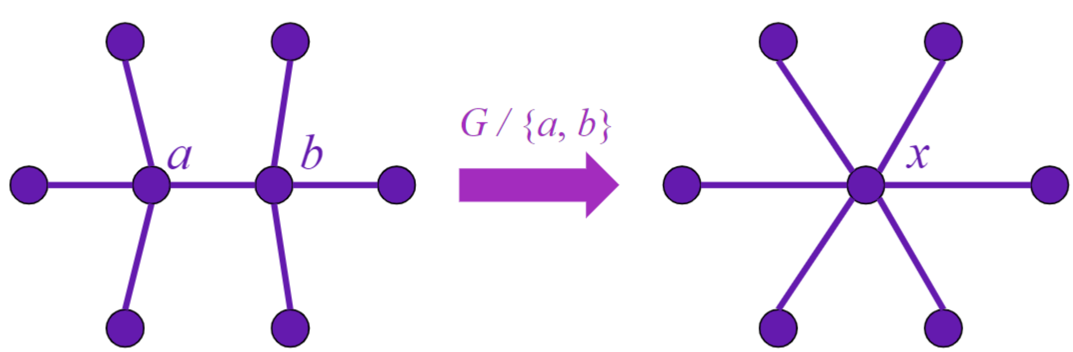
If G is planar, then $|E(G)| \leq 3|V(G)| - 6$. However,

$$\sum_{v \in V(G)} d_G(v) = 2|E(G)| \leq 6|V(G)| - 12 < 6|V(G)|.$$

So not every vertex in G can have a degree of 6 or more.

A contraction of an edge $\{a, b\}$ in a graph G, notated as $G/\{a, b\}$, is the graph obtained by adding a vertex x to $G - \{a\} - \{b\}$ so that x is adjacent to all vertices in $N_G(a) \cup N_G(b)$. An important property of contractions is that if G is planar, then $G/\{a, b\}$ is also planar.

What a contraction looks like:



The 5-color Theorem: If G is a planar graph, then $\chi(G) \leq 5$.

Proof:
If $|V(G)| \leq 5$, then we may simply assign all vertices different colors to show that $\chi(G) \leq 5$. So, consider when $|V(G)| \geq 6$.

Then if $\delta(G) \leq 4$, we know there exists $v \in V(G)$ such that $d_G(v) \leq 4$. So, by induction we can find a proper 5-coloring of $G - \{v\}$. And because v has at most 4 neighbors, there is at least one color left over to assign v afterwards. Meanwhile, because of the previous theorem, we know that any induced subgraph of G will have a minimum degree of at most 5. So, the only non-trivial case we need to prove is when $\delta(G) = 5$.

Thus, assume there exists a vertex $v \in V(G)$ such that $d(v) = 5$. Since K_5 is not planar, there exists some pair of vertices: $a, b \in N_G(v)$ which are not adjacent. Define $H = G/\{a, b\}$ with a new vertex w adjacent to $N_G(v) \cup N_g(a)$.

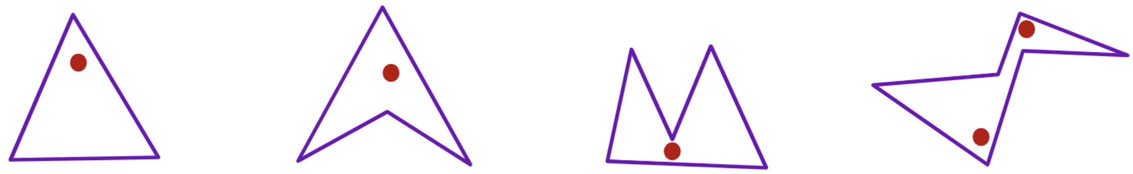
Next contract H again, defining $I = H/\{b, w\}$ with a new vertex x adjacent to $N_H(w) \cup N_H(b)$. Note that I and H are planar. Thus, since I is a planar graph with less vertices than G, we can do induction to find a proper 5-coloring of I. We'll call this coloring c.

For each $y \in V(I) \cap V(G)$, define $c'(y) = c(y)$. This effectively gives a proper coloring of $G - a - b - v$. Then assign $c'(a) = c'(b) = c(x)$. We can do this because we chose a and b to not be adjacent and because none of a and b's neighbors could have been assigned $c(x)$. And finally, because a and b in $N(v)$ have been assigned same color, there is at least one color which we can assign to v. ■

Lecture 10: 2/13/2024

The Art Gallery Problem:
Let R be a region in the plane bounded by a polygon. Also, define that two points are "*mutually visible*" if there exists a straight line between them that is contained entirely in R. Then, what is the smallest size of a set of points S in R such that every point in R is mutually visible to a point in S.

In terms of the art gallery analogy, R is the floor plan of an art gallery and $|S|$ is the number of guards you've hired to guard the gallery. Here are some example galleries with guards:



3 edges, 1 guard | 4 edges, 1 guard | 5 edges, 1 guard | 6 edges, 2 guard

Art Gallery Theorem: For an n-sided polygon, one needs to include at most $\lfloor \frac{n}{3} \rfloor$ many points in S in order to have that every point in R be mutually visible with a point in S.

Proof:

Consider triangulating R. In other words, add straight edges between vertices of R in order to partition R into a bunch of triangular regions.

Having done that, now consider the graph G whose vertices are the vertices of R and whose edges are the boundaries of the triangles of our triangulation of R. We can prove that $\chi(G) = 3$.

When $n = 3$, we have that G consists of a single triangle. So $\chi(G) = 3$.

Now we proceed by induction. Assume that for $n > 3$, we know that $\chi(G) = 3$ if R has less than n sides. Therefore, we now consider a graph G made from an n-sided polygon. Let $\{u, v\}$ be any edge of G not on the boundary of the infinite face. Importantly, we know such a side will exist because $n > 3$. Then, we can partition G into two subgraphs G_1 and G_2 such that $V(G_1) \cap V(G_2) = \{u, v\}$ and $E(G_1) \cup E(G_2) = E(G)$.

G_1 and G_2 are the triangulations of some polygon with less than n sides. Therefore by strong induction, there exists the proper vertex 3-colorings: $c_1 : V(G_1) \to \{1, 2, 3\}$, and $c_2 : V(G_2) \to \{1, 2, 3\}$. Shifting around the colors so that $c_1(u) = c_2(u)$ and $c_1(v) = c_2(v)$, we can then define a proper 3-coloring of all G. So $\chi(G) \leq 3$.

Since G obviously can't be colored by 2 or less colors since the triangles in G need three colors, we thus conclude that $\chi(G) = 3$.

Having shown that G has a 3-coloring, now note that in that 3-coloring, the three vertices enclosing any triangular face of G must all have a different color assigned to them. Additionally, every point in a triangle is mutually visible to another point in that triangle.

Letting the least assigned color be called i, make S the set of locations of vertices which were assigned the color i. Then any point in any of the triangles partioning R is mutually visible to a point in S. Also, $|S|$ is at most $\lfloor \frac{n}{3} \rfloor$. ■

Some other problems related to the art gallery problem are:

- The rectilinear art gallery problem.

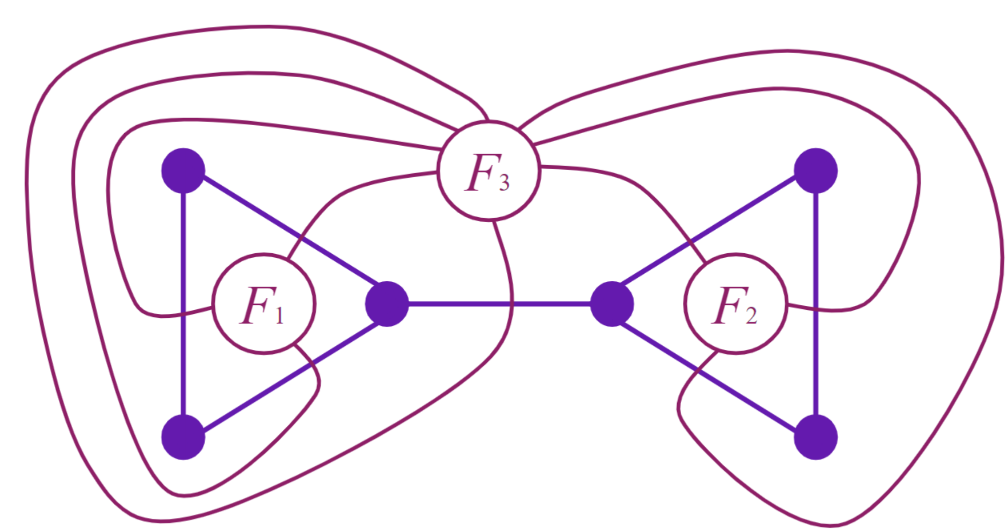
  In this version of the problem, all angles between sides must be right angles. As it turns out, with this restriction on R we only need $\lfloor \frac{n}{4} \rfloor$ points in S so that every point in R is mutually visible with a point in S. However, we aren't proving that in this class.

- The 3d art gallery problem
- The non-polygon art gallery problem

The 4-color Theorem: If G is a planar graph, then $\chi(G) \leq 4$.

Unfortunately, no short proof of this is known. The quickest proof considers 633 different cases.

Let G be a plane graph. To get the dual of G, denoted G^*, represent each face $f \in F(G)$ as a vertex in G^* and draw it inside its corresponding face in G. Then, for each edge $e \in E(G)$, draw an edge $\{f_1, f_2\} \in E(G^*)$ such that e is between the faces f_1 and f_2 of G and $\{f_1, f_2\}$ is drawn to only intercept e at one point.



Note that G^* is a planar multigraph.

Proposition: If G is connected, then $(G^*)^* \cong G$. In other words, while $(G^*)^*$ might be drawn in the plane differently from G, both graphs will have the same vertex, edge, and face sets.

The proof for this is beyond the scope of this class (and requires math I don't know yet).

The significance of the above proposition is that by taking the dual of a connected plane graph G, we can convert the faces of G into vertices and the vertices of G into faces.

Application: giving a proper coloring of the faces of G is equivalent to giving a proper coloring of the vertices of G^*.

Here is one notable approach to the four color theorem:

A graph G is k-edge-connected if for every set X of less than k edges, $G - X$ is still connected.

Proposition: G^* is simple if and only if G is 3-edge-connected.

Proof: (by contrapositive)
($\Longrightarrow$) If G is 1-edge-connected, then we know there is a bridge $\{u, v\}$ in G. That bridge must have the same face on either side of it or else there'd be an alternate path going from u to v, thus contradicting that $\{u, v\}$ is a bridge. This then means that G^* must have a loop. So G^* is not simple.

Now consider if G is 2-edge-connected. In that case there exists edges e_1 and e_2 which when both removed split G into two components. Let e_1 be on the boundary of f_1 and f_2. The effect of subtracting e_1 from G is that we merged only those two faces in the graph $G - e_1$. Meanwhile, we know that e_2 must be a bridge in $G - e_1$. Therefore it must have the same face on either side of it. However, e_2 wasn't a bridge in G. This tells us that e_2 must also be on the boundary of f_1 and f_2. So f_1 and f_2 must share two edges on their boundaries, which in turn means that G^* is not simple.

($\Longleftarrow$) If G^* has a loop, then that means there is an edge in G not on the boundary of two distinct faces. Then that edge is a bridge, which means that G is 1-edge-connected.

If G^* has multiple edges going between two vertices, then two faces of G share more than one edge. By removing one of those shared edges, all the remaining shared edges become bridges. Hence, removing two shared edges disconnects the graph. So G is 2-edge connected.

Now the 4-color theorem is obviously true if $|V(G)| < 4$. Thus, to prove the 4-color theorem, we can focus exclusively on graphs with at least 4 vertices. Additionally, it is enough to prove the 4-color theorem for maximal planar graphs (i.e. a planar graph with as many edges as possible). After all, removing edges does not make a coloring invalid.

Observation 1: G is a maximal planar graph if and only if every face of G is a triangle.

Observation 2: The dual of any maximal planar graph with at least 4 vertices is a simple planar 3-edge-connected cubic graph.

Let G be a maximal planar graph. Then as every face of G is a triangle, every face of G has three neighbors. Hence G^* is cubic.

Additionally, as every face of G must be a triangle, if two faces share two edges $\{u, v\}$ and $\{v, w\}$, then they must also both have an edge $\{u, w\}$. Since the graph G has at least four vertices, we know that G is not just a single triangle. Hence, the only way for both faces to have an edge $\{u, w\}$ would be if G was a multigraph (something we haven't allowed for). So we conclude that G^* has at most one edge going between any two vertices.

Since G and thus $(G^*)^*$ is assumed to be simple since we did not indicate otherwise, we know that G is 3-edge-connected.

Observation 3: The dual of any planar 3-connected cubic graph is a maximal planar graph.

Let G be a planar 3-connected cubic graph. Then G^* being 3-connected means that G^* is simple (so it makes sense to even ask if G^* is a maximal planar graph). Additionally, G being cubic means that every face of G^* is a triangle. So G^* is a maximal planar graph.

This then leads to the following theorem:

Theorem: Every planar graph G is 4-colorable if and only if every 3-connected cubic planar graph is 3-edge colorable.

Proof:

($\Longrightarrow$) Let G be a 3-connected cubic planar graph. Then we know that G^* is a maximal planar graph. If G^* is 4-colorable, then we can properly color the vertices of G^*, which is equivalent to properly coloring the faces of G.

Let $c : F(G) \to \{1, 2, 3, 4\}$ be a proper face coloring of G. Then because G is 3-connected, we know that no edge in G is a bridge. So, every edge in G is between two faces which are colored differently.

We now define the following 3-edge-coloring of G:

> If e is between face colors 1 and 2, or 3 and 4, define $c'(e) = 1$.
> If e is between face colors 1 and 3, or 2 and 4, define $c'(e) = 2$.
> If e is between face colors 1 and 4, or 2 and 3, define $c'(e) = 3$.

Hence $\chi'(G) = 3$.

($\Longleftarrow$) Now let G be a maximal plane graph. If G^* is 3-edge-colorable, then there exists an edge coloring $c' : E(G^*) \to \{1, 2, 3\}$.

Consider the set $H_1 = \{e \in E(G^*) \mid c'(e) \in \{1, 2\}\}$. H_1 must consist of some number of disjoint cyles which span G^*. Similarly, $H_2 = \{e \in E(G^*) \mid c'(e) \in \{1, 3\}\}$ must also consist of some number of disjoint cycles which span G^*. Note that $H_1 \cup H_2$ span G^*. So every edge is on the boundary of a cyle of either H_1 or H_2.

Define $c_1(f)$ such that $c_1(f) = 1$ if f is in the interior of a cycle of H_1, whereas $c_1(f) = 2$ if f is not in the interior of a cycle of H_1. Similarly define $c_2(f)$ but using H_2.

Now we define the following face coloring of G^*:

> $c(f) = 1$ if $(c_1(f), c_2(f)) = (1, 1)$
> $c(f) = 2$ if $(c_1(f), c_2(f)) = (1, 2)$
> $c(f) = 3$ if $(c_1(f), c_2(f)) = (2, 1)$
> $c(f) = 4$ if $(c_1(f), c_2(f)) = (2, 2)$

Then c is a proper 4-face-coloring of G^*. So, we in turn know that G has a proper 4-vertex-coloring.

Here's a quick useful calculation for the homework that the professor gave us:

> Let G be a maximal planar graph with n vertices. By the handshaking lemma, we know that $3|F(G)| = 2|E(G)|$. And since G is obviously connected, we can thus say that: $|V(G)| - |E(G)| + |F(G)| = n - \frac{1}{2}|F(G)| = 2$. So, $|F(G)| = 2n - 4$. Plugging this into Euler's formula again, we then get that $|E(G)| = 3n - 6$.
>
> As a result, we know that G^* has $2n - 4$ vertices and $3n - 6$ edges.

Lecture 11: 2/15/2024

Let $\mathcal{F}$ be any family of graphs.
The external numbers / Turán numbers for $\mathcal{F}$ are the quantities $\mathrm{ex}(n, \mathcal{F})$ which denote the maximum number of edges in an n vertex graph not containing any graph in $\mathcal{F}$.

We call a graph that does not contain any member of $\mathcal{F}$ an $\mathcal{F}$-free graph.
A k-core of a graph G is the graph obtained through the following algorithm:

1. Let X be the set of vertices removed from G in previous steps.

2. If there is no vertex $v \in V(G - X)$ with degree less than k in $G - X$, then you are done as $G - X$ is a k-core of G.

3. Otherwise, add v to the set X and go back to step 2.

If not empty, a k-core is the largest subgraph of G with minimum degree at least k.

The order in which vertices are removed does not matter. The k-core of a graph is unique.

Lemma: If G is an n-vertex graph with more than $(k-1)n - \binom{k}{2}$ edges, then G has a nonempty k-core.

Proof:
If $\delta(G) \geq k$, we're done. So, assume that there is a vertex $v_1 \in V(G)$ with $d(v_1) \leq k - 1$. Then:

$$|E(G - \{v\})| > (k-1)n - \binom{k}{2} - (k-1) = (k-1)(n-1) - \binom{k}{2}$$

Now if $n = k$, then $|E(G)| > (k-1)k - \binom{k}{2} = \binom{k}{2}$. But that would mean that G has more edges than K_n, which is impossible. So, we know that $n > k$. Therefore, it makes sense to consider repeating the above step $n - k$ times. Having done that, we will once again get that there are more than $(k-1)(k) - \binom{k}{2}$ edges remaining. However, that's impossible as a k vertex graph can't have more than $\binom{k}{2}$ edges. Hence, we conclude that the algorithm must have terminated before all $n - k$ steps could be done.

Corollary: If G is a graph with n vertices and more than $(k-1)n - \binom{k}{2}$ edges, then G contains a cycle of length at least $k+1$ unless $G = K_k$.

Proof:
Let $H \subseteq G$ be a non-empty k-core in G. Since $\delta(H) \geq k$, H contains a cycle of length at least $k+1$. Thus so does G. ■

A cut: (A, B), of a graph G is a spanning subgraph of G such that such that $V(G)$ is partitioned into two sets A and B and an edge $\{a, b\} \in E(G)$ is only included in the cut if $a \in A$ and $b \in B$.

A max cut is a cut with as many edges as possible.

Theorem: Every graph with m edges has a cut with at least $m/2$ edges.

Proof:

Let G have m edges. On each vertex of G, flip a fair coin. Then, let A be the set of vertices which have heads and let B be the set of vertices which have tails. That way, (A, B) is a cut.

Since the assignment of each vertex is independent from the assignment of any other vertex, we have that there is a $\frac{1}{2}$ probability of any edge being included in the cut. Thus, the mean number of edges between A and B is:

$$\sum_{e \in E(G)} \mathbb{P}(e \text{ is in the cut}) = \sum_{i=1}^{m} \frac{1}{2} = \frac{1}{2}m$$

This implies that some cut (A, B) of G must have at least $m/2$ edges. (It also implies some cut has at most $m/2$ edges.)

Erdős-Gallai Theorem: Let $k \geq 1$, $n \geq 1$, and let G be an n-vertex P_k-free graph. Then $|E(G)| \leq (k-1)\frac{n}{2}$ with equality if and only if k divides n and every component of G is K_k.

Proof: (we proceed by induction)

Base Case:

If $k = 1$, then $\text{ex}(n, P_k) = 0 = (k-1)\frac{n}{2}$. Thus, the theorem is trivially true.

If $n \leq k$, then $\text{ex}(n, P_k) = E(K_n) = (n-1)\frac{n}{2} \leq (k-1)\frac{n}{2}$. Specifically, equality holds if $n = k$.

Inductive step: (assume the theorem holds if $|V(G)| < n$)

Suppose G is an n vertex P_k-free graph.

If G is disconnected, then there are two disjoint subgraphs G_1 and G_2 such that $G = G_1 + G_2$. Furthermore, G is P_k-free if and only if G_1 and G_2 are both P_k-free. Let $|V(G_1)| = n_1$ and $|V(G_2)| = n_2$. Because $n_1, n_2 < n$, we know by induction that for G_1 and G_2 to be P_k-free, we must have that $E(G_1) \leq (k-1)\frac{n_1}{2}$ and that $E(G_2) \leq (k-1)\frac{n_2}{2}$. Therefore:

$$E(G) \leq (k-1)\frac{n_1}{2} + (k-1)\frac{n_2}{2} = (k-1)\frac{n}{2}$$

Now we suppose G is connected.

If G has a vertex v of degree less than $\frac{k}{2}$, then consider that because $G - \{v\}$ must be P_k free, we know by induction that $|E(G - \{v\})| \leq (k-1)\frac{n-1}{2}$. This in turn means that:

$$|E(G)| \leq (k-1)\tfrac{n-1}{2} + \tfrac{k-1}{2} = (k-1)\tfrac{n}{2}.$$

Also, note that by induction, $|E(G - \{v\})| = (k-1)\frac{n-1}{2}$ if and only if $|E(G - \{v\})|$ is a bunch of disjoint K_k components. But in that case, if $d(v) \geq 1$, then G cannot be P_k-free. So, we can actually conclude that $|E(G)| < (k-1)\frac{n}{2}$.

Meanwhile, if $\delta(G) \geq \frac{k}{2}$, then assume towards a contradiction that $|E(G)| > (k-2)\frac{n}{2}$. That way, by induction on k, we can conclude that G has a path P of length $k-1$

This was why we included the base case of when $k = 1$.

If P can be extended, then we have a contradiction as G is not P_k pathless. So assume G can't be extended. Then, using the same trick as we used previously to prove Dirac's Theorem (pages 14 and 15), we can turn P into a cycle of length $k-1$. But this gives another contradiction because G is connected. So, there must be a vertex adjacent to our cycle of length k, which means that we can remove an edge of the cycle and add an edge to that adjacent vertex to get a path of length k in G.

So, when $\delta(G) \geq \frac{k}{2}$, we have shown that $|E(G)| \leq (k-2)\frac{n}{2} < (k-1)\frac{n}{2}$.

Notably, $|E(G)| = (k-1)\frac{n}{2}$ if and only if the only induction we ever did was considering if G was disconnected, and if the only base case we ever reached was for when a subgraph equals K_k. So equality holds if and only if k divides n and every component of G is K_k.

Lecture 12: 2/20/2024

The Turán graph $T_r(n)$ has n vertices which are partitioned into r sets of size $\lfloor \frac{n}{r} \rfloor$ or $\lceil \frac{n}{r} \rceil$ such that $E(T_r(n))$ consists of every possible edge going between different partitions.

Observe: $K_{r+1} \not\subseteq T_r(n)$ because $\chi(K_{r+1}) = r+1 > r = \chi(T_r(n))$ for any n.

If $r \mid n$, then by handshaking lemma:

$$e(T_r(n)) = \left(\tfrac{n}{r}(r-1)\right)\tfrac{n}{2} = \binom{r}{2}\tfrac{n^2}{r^2} = \tfrac{n^2}{2}(1 - \tfrac{1}{r})$$

Turán's Theorem: For all $n \geq 1$ and $r \geq 2$, $\mathrm{ex}(n, K_{r+1}) = e(T_r(n))$. Also, if G is a K_{r+1}-free graph on n vertices and $e(G) = e(T_r(n))$, then $G = T_r(n)$.

In other words, $T_r(n)$ is the unique maximal graph.

Proof: (we proceed by induction)

Base Case:

When $n \leq r$, we clearly have that $\mathrm{ex}(n, K_{r+1}) = e(K_n)$. However, we also have $T_r(n) = K_n$ when $n \leq r$. Thus, Turán's theorem is clearly true when $n \leq r$.

Inductive Step: (assume the theorem holds when $r \leq |V(G)| < n$)

Assume G is a K_{r+1}-free graph such that $e(G) \geq e(T_r(n))$.

Remove edges from G until we get a graph G' with $e(G') = e(T_r(n))$. Then let v be a vertex of G' of minimum degree. We know that $d_{G'}(v) \leq \delta(T_r(b))$ because if this weren't the case, then G' would have to have more edges than $T_r(n)$ by the handshaking lemma. Therefore:

$$e(G' - \{v\}) \geq e(T_r(n)) - \delta(T_r(n)) = e(T_r(n-1)).$$

But now by induction, we know that $G' - \{v\}$ is K_{r+1}-free only if $e(G' - \{v\}) \leq e(T_r(n-1))$. Therefore $e(G' - \{v\}) = e(T_r(n-1))$. Furthermore, this tells us by induction that $G' - \{v\} = T_r(n-1)$.

Next, we try to fit v back into $G' - \{v\} = T_r(n-1)$. Let $W_1, \ldots, W_r$ be the partitions of $G' - \{v\}$. If $N_{G'}(v) \cap W_i \neq \emptyset$ for all $i \in \{1, \ldots, r\}$, then by picking one v_i from each $W_i \cap N_{G'}(v)$, we get that $\{v, v_1, \ldots, v_r\}$ induces a K_{r+1} subgraph inside G'. This is clearly a contradiction. So we conclude that there is a partition W_i such that $N_{G'}(v) \cap W_i = \emptyset$.

Hence, G' is an r-partite graph. And since $e(G') = e(T_r(n))$ and $T_r(n)$ has more edges than every other n vertex r-partite graph, we know that $G' = T_r(n)$. Finally, notice that $T_r(n)$ is a maximal K_{r+1}-free graph. Therefore, we must have that $e(G) = e(G')$ if G is K_{r+1}-free.

Here is a theorem given without proof:

Erdős-Simonovitz-Stone theorem: If F is any graph of chromatic number $r+1$, then:

$$\lim_{n \to \infty} \left(\frac{\mathrm{ex}(n, F)}{\binom{n}{2}} \right) = 1 - \frac{1}{r}.$$

Jensen's inequality: If a function $f(x)$ is convex and $x_1, x_2, \ldots x_t$ are real numbers, then:

$$\frac{1}{t}\sum_{i=1}^{t} f(x_i) \geq f\left(\frac{1}{t}\sum_{i=1}^{t} x_i\right)$$

I will probably learn much more about this inequality in a later course.

Kövari-Sós-Turán theorem: For $s \geq r \geq 2$,

$$\mathrm{ex}(n, K_{r,s}) \leq \tfrac{1}{2}\left((s-1)^{\frac{1}{r}} n^{2-\frac{1}{r}} + (r-1)n\right)$$

Proof:

If G is an n-vertex $K_{r,s}$-free graph with average degree d and $d \leq r-1$, then $e(G) \leq \frac{1}{2}(r-1)n$. So the inequality of the theorem is true automatically.

So, suppose $d > r-1$. Since G is $K_{r,s}$-free, we know that no set of r vertices in G has more than $s-1$ common neighbors. This leads us to the inequality that:

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

Explanation of this inequality:

Let X be the average number of common neighbors of any set of r vertices in G. Meanwhile, let Y be the total sum of the numbers of common neighbors of every set of r vertices in G. Then:

$$X = \frac{Y}{\text{\# of vertex sets of size } r} = \frac{Y}{\binom{n}{r}}$$

If $X > (s-1)$, then it is guarenteed that at least one set of r vertices will have more than $s-1$ common neighbors. So, we must have that $X \leq (s-1)$.

Meanwhile, $Y = \displaystyle\sum_{v \in V(G)} \binom{d(v)}{r}$.

So one get's the above inequality by multiplying both sides of $X \leq (s-1)$ by $\binom{n}{r}$.

Fact: $f(x) = \begin{cases} 0 & \text{if } x < r \\ \binom{x}{r} & \text{if } x \geq r \end{cases}$ is a convex function for any r.

Thus, applying Jensen's inequality, we can say that:

$$\sum_{v \in V(G)} \binom{d(v)}{r} \geq \sum_{v \in V(G)} \binom{d}{r} = n\binom{d}{r}$$

Now since $d \geq r$, we can safely cancel $(d-r)!$ to get:

$$\binom{d}{r} = \frac{d!}{r!(d-r)!} = \frac{d\cdots(d-r+1)}{r!} \geq \frac{(d-r+1)^r}{r!}$$

Meanwhile, since $n \geq r$, we can safely cancel $(n-r)!$ to get:

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{n\cdots(n-r+1)}{r!} \leq \frac{n^r}{r!}$$

Combining inequalities, we get that: $n\frac{(d-r+1)^r}{r!} \leq (s-1)\frac{(n)^r}{r!}$. This can then be reshuffled as follows:

$$\begin{aligned} n\tfrac{(d-r+1)^r}{r!} \leq (s-1)\tfrac{n^r}{r!} \quad &\implies \quad (d-r+1)^r \leq (s-1)n^{r-1} \\ &\implies \quad d-r+1 \leq \left((s-1)n^{r-1}\right)^{\frac{1}{r}} \\ &\implies \quad d \leq (s-1)^{\frac{1}{r}}n^{1-\frac{1}{r}} + r - 1 \\ &\implies \quad \tfrac{dn}{2} \leq \tfrac{n}{2}\left((s-1)^{\frac{1}{r}}n^{1-\frac{1}{r}} + r - 1\right) \\ &\implies \quad e(G) \leq \tfrac{1}{2}\left((s-1)^{\frac{1}{r}}n^{2-\frac{1}{r}} + (r-1)n\right) \blacksquare \end{aligned}$$

Lecture 13: 2/22/2024

A Sidon set in an abelian group Γ (with the group operation $+$) is a set $A \subseteq \Gamma$ such that if $a+b=c+d$ with $a,b,c,d \in A$, then $\{a,b\} = \{c,d\}$.

> For those who haven't taken algebra yet:
>
> A group G is a set endowed with a binary operation $\bullet$ satisfying:
> - Associativity: $\forall g,h,k \in G,\ \ (g\bullet h)\bullet k = g\bullet(h\bullet k)$
> - Identity: $\exists e_G \in G\ \ s.t.\ \ \forall g \in G,\ \ g\bullet e_G = g = e_G\bullet g$
> - Inverse: $\forall g \in G,\ \ \exists h \in G\ \ s.t.\ \ g\bullet h = e_G = h\bullet g$
>
> An abelian group is a group where $\bullet$ is also commutative: $\forall h,g \in G,\ \ h\bullet g = g\bullet h$.

Proposition: An upper bound to $|A|$ is given by $|A|(|A|-1) \leq |\Gamma|-1$.

Proof:

Consider any two distinct ordered pairs (a,b) and (c,d) of elements from A such that both pairs have two distinct elements (i.e. $a \neq b$ and $c \neq d$). Letting $-b$ and $-d$ be the inverses of b and d respectively, we know that if $a-b=c-d$, then $a+d=b+c$.

Now note that because A is a sidon set, we must have that $\{a,d\} = \{b,c\}$. Therefore, since we assumed above that $a \neq b$ and $c \neq d$, we must have that $a=c$ and $b=d$. However, this contradicts that (a,b) and (c,d) are distinct pairs.

This tells us that there must be a one-to-one mapping of ordered pairs of distinct elements of A to Γ. The number of such ordered pairs is given by $\frac{|A|!}{(|A|-2)!} = |A|(|A|-1)$. Therefore, $|A|(|A|-1) \leq |\Gamma|$.

Finally, note that no pair (a,b) could be mapped to 0 (the identity element of Γ) in the above process. After all, if $a-b=0$, then $a=b$ and so a and b are not pairwise distinct. Hence, we get the bound of the proposition:

$$|A|(|A|-1) \leq |\Gamma|-1$$

We can connect Sidon sets to graph theory as follows:...

Given any set $A \subseteq \Gamma$, we define the Cayley sum graph $G(A)$ such that $V(G(A)) = \Gamma$ and $\{x,y\} \in E(G(A))$ whenever $x+y \in A$.

Note that $G(A)$ may be a pseudograph because if $x+x$ belongs A, then $G(A)$ has a loop. Also, $G(A)$ is $|A|$-regular because $x+(-x+a)=a$ for all $x \in \Gamma$ and $a \in A$.

If $G(A)$ contains a quadrilateral (x,y,z,w,x), then there exists distinct elements $a,b,c,d \in A$ such that:

$$x+y=a \quad y+z=b \quad z+w=c \quad w+x=d$$

But then A cannot be a Sidon set because:

$$\begin{matrix} x+y-y-z=a-b \Longrightarrow x-z=a-b \\ x+w-w-z=d-c \Longrightarrow x-z=d-c \end{matrix} \Longrightarrow a-b=d-c$$

So if A is a Sidon set, then $G(A)$ is C_4-free.

Unfortunately, because $G(A)$ may not be simple, we can not apply the Kövari-Sós-Turán theorem on $G(A)$ to get an equivalent or better upper bound on $|A|$. However, what we have done is found a way of generating C_4-free graphs. After all, $G(A)$ minus its loops is a simple C_4-free graph.

Let q be an odd prime and $\mathbb{Z}_q$ denote the cyclic group of integers modulo q.

In math 109, we referred to these as the set of congruence classes modulo q.
Since q is a prime number, $\mathbb{Z}_q$ satisfies all the properties of a field and thus is a commutative group.

Let $\Gamma_q = \mathbb{Z}_q \times \mathbb{Z}_q$ with addition as the group operation.

i.e. $(a, b) + (c, d) = (a + c \mod (q),\ \ b + d \mod (q))$

Also define n as $|\Gamma_q| = q^2$.

Then, consider $A = \{(x, x^2) \in \Gamma_q \mid x \in \mathbb{Z}_q\}$.

$|A| = q = \sqrt{n}$.

Also, A is a Sidon set.

Consider if $(a, a^2) + (b, b^2) = (c, c^2) + (d, d^2)$. Then we'd have that:

$$\begin{matrix} a + b \overset{q}{\equiv} c + d \\ a^2 + b^2 \overset{q}{\equiv} c^2 + d^2 \end{matrix} \quad \xLeftrightarrow{\text{which is equivalent to saying that}} \quad \begin{matrix} a - c \overset{q}{\equiv} d - b \\ a^2 - c^2 \overset{q}{\equiv} d^2 - b^2 \end{matrix}$$

Now note that $a^2 - c^2 = (a - c)(a + c)$. And since $\mathbb{Z}_q$ is a field, we can divide both sides of $a^2 - c^2 \overset{q}{\equiv} d^2 - b^2$ by $a - c \overset{q}{\equiv} d - b$ to get: $a + c \overset{q}{\equiv} d + b$.

Then combining $a + c \overset{q}{\equiv} b + d$ and $a - c \overset{q}{\equiv} d - b$ tells us that $2a \overset{q}{\equiv} 2d$. And since q is odd (coprime with 2), we can cancel out the 2s to get that $a \overset{q}{\equiv} d$.

Therefore $(a, b) = (c, d)$.

Now consider the Cayley sum graph $G(A)$.

$G(A)$ has q^2 vertices, $\frac{q^3}{2}$ edges, and is C_4-free.

Also, $G(A)$ has a loop for every solution to $2(x, y) = (a, a^2)$. Since the solution to $2(x, y) = (a, a^2)$ is unique for each $a \in \mathbb{Z}_q$, we thus have that there are q loops.

So, after removing loops, we have shown that for each odd prime q:

$$\mathrm{ex}(q^2, C_4) \geq \frac{q^3}{2} - q$$

Fact from number theory: for any $n \geq 2$, there exists a prime number q between n and $n + n^{\frac{2}{3}}$.

Using the above fact, we can strategically pick a prime number q such that:

$$\mathrm{ex}(n, C_4) \geq \frac{q^3}{2} - q \geq \frac{n^{\frac{3}{2}}}{2}(1 - o(1))$$

(where $o(1)$ is some unknown function $f(n)$ such that $\lim\limits_{n\to\infty}(f(n)) = 0$)

This is a taste of what researching external combinatorics is like when no exact answer for the external numbers of a graph can be derived.

As for an upper bound for a graph without C_4, note that $C_4 = K_{2,2}$. So we can use the Kövari-Sós-Turán theorem in order to say that: $\text{ex}(n, C_4) \leq \frac{1}{2}(n^{\frac{3}{2}} + n)$. Thus, the asymptotic behavior of $\text{ex}(n, C_4)$ is that $\text{ex}(n, C_4) = (\frac{1}{2} + o(1))n^{\frac{3}{2}}$.

If G is a graph containing a cycle, then the girth of G is the length of a shortest cycle in G.

Theorem (Moore Bound): Let G be a d-regular graph of girth at least g with n vertices. Then:

- $1 + d \displaystyle\sum_{i=0}^{\lfloor \frac{g-1}{2} \rfloor} (d-1)^{i-1} \leq n$ (if g is odd)
- $2 \displaystyle\sum_{i=0}^{\frac{g}{2}-1} (d-1)^{i} \leq n$ (if g is even)

Proof:

(g is odd):

Imagine drawing a BFS-tree inside G starting at a vertex v and name each layer of the tree N_i such that i is the distance from v to any element of N_i. Because there can be no cycles of length $g-1$ or less in G, we must have that for $i < \lfloor \frac{g-1}{2} \rfloor$, all edges with an end in N_i not going to N_{i-1} must be going to a unique element in N_{i+1}. Therefore:

- $N_0 = \{v\}$ obviously has 1 vertex.
- N_1 obviously will have d vertices.
- N_2 must have $d(d-1)$ vertices.
- N_3 must have $d(d-1)^2$ vertices.

$\vdots$

- $N_{\lfloor \frac{g-1}{2} \rfloor}$ must have $d(d-1)^{\lfloor \frac{g-1}{2} \rfloor - 1}$ vertices.

Adding up the number of vertices in each of those layers, we get a lower bound for n.

(g is even):

Our previous reasoning can be improved when g is even because assuming that equality holds in the previously proven inequality, the shortest lenth of a cycle in G is $2\lfloor \frac{g-1}{2} \rfloor + 1$. But that will be an odd number which we are not allowing. So clearly, the greatest lower bound we can give is higher when g is even.

Consider two adjacent vertices u and v in G. Then for each one of them, draw a BFS tree spanning $G - \{u, v\}$. Let $N_i(u)$ and $N_i(v)$ refer to the ith layer of tree coming out of u and v respectively.

Like before, we must have that for $i < \frac{g}{2} - 1$, all edges with an end in $N_i(u)$ not going to $N_{i-1}(u)$ must be going to a unique vertex in $N_{i+1}(u)$, while all edges with an end in $N_i(v)$ not going to $N_{i-1}(v)$ must be going to a unique vertex in $N_{i+1}(v)$. Additionally, each $N_i(u)$ must be disjoint to every $N_i(v)$ when $i < \frac{g}{2} - 1$. If these facts weren't true, we would be able to draw a shorter cycle than length g. So:

- $N_0(u) \cup N_0(v) = \{u\} \cup \{v\}$ obviously has 2 vertices.
- $N_1(u) \cup N_1(v)$ must have $2(d-1)$ vertices.
- $N_2(u) \cup N_2(v)$ must have $2(d-1)^2$ vertices.

$$\vdots$$

- $N_{\frac{g}{2}}(u) \cup N_{\frac{g}{2}}(v)$ must have $2(d-1)^{\frac{g}{2}-1}$ vertices.

Adding up the number of vertices in each of those layers, we get a lower bound for n.

A Moore graph is a graph which acheives equality in the Moore Bound.

For $g \in \{3, 4\}$, the Moore graphs are obvious. However, finding Moore graphs when $d \geq 3$ and $g \geq 5$ is an active area of research.

Also, the Petersen graph is a 3-regular Moore graph with girth $g = 5$.

Lecture 14: 3/5/2024

Even Cycle Theorem:
For some constant c_k (which we prove is at most $8k$): $\mathrm{ex}(n, C_{2k}) \leq c_k n^{1+\frac{1}{k}}$.

The proof in class for this was very long and glossed over a lot of details because most of us in the class don't know enough math to do the entire proof yet. So, I'm just going to skip writing it down.

Lecture 15: 3/7/2024

The Ramsey number $r(s,t)$ for $s, t \in \mathbb{Z}^+$ is the minimum N such that whenever the edges of K_N are colored red and blue, either a red K_s occurs or a blue K_t occurs.

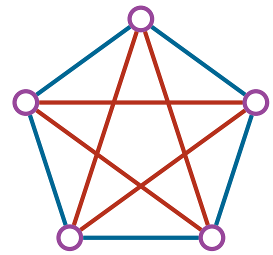
Equivalently, $r(s,t)$ is the minimum N such that every K_s-free graph on N vertices contains an independent set of size t.

To quote from homework 3 of the class:
An **independent set** in a graph G is a set $X \subseteq V(G)$ such that $e(X) = 0$, and the **independence number** $\alpha(G)$ is the largest size of an independent set in G.

A K_s-free graph with $r(s,t)-1$ vertices and no independent set of size t is called a Ramsey graph, and a coloring of K_n with $n = r(s,t)-1$ vertices and no monochromatic K_s or K_t is called a Ramsey coloring.

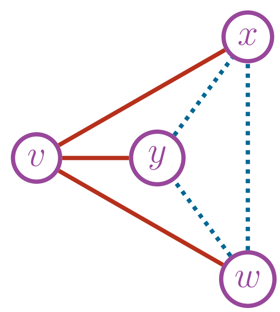
Example:

We can show with the following construction that $r(3,3) > 5$



Meanwhile for $N > 6$, note that given any vertex v there must be three edges $\{v,x\}$, $\{v,y\}$, and $\{v,w\}$ with the same color. This is a result of the pigeonhole principle. Without loss of generality, assume that they are colored red.

Now consider the edges $\{x,y\}$, $\{x,w\}$, and $\{y,w\}$. If any one of them are red, we have a red triangle. Meanwhile, if all of them are not red, then we have a blue triangle. Thus, for $N \geq 6$, we must have a red or blue triangle. Hence, $r(3,3) \leq 6$.



Here's a hopefully obvious observation:
If all red-blue colorings of K_N contain a forbidden graph, then all red-blue colorings of K_{N+1} will contain the forbidden graph because K_N is a subgraph of K_{N+1}.

Erdős-Szekeres Theorem:
For $s \geq 2$ and $t \geq 2$, $r(s,t) \leq r(s-1,t) + r(s,t-1)$. In particular:

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

Proof:
$r(1,t) = r(s,1) = 1$ trivially. Also trivially, if $s = 2$, then $r(s,t) = t$. Meanwhile if $t = 2$, then $r(s,t) = s$. Indeed, both of these satisfy the inequalities of the theorem.

Now we proceed by induction, assuming $s \geq 3$, $t \geq 3$, and that the theorem is true when s or t is smaller.

Let $N = r(s,t)$ and suppose we are given a red-blue coloring of $E(K_N)$. Picking $v \in V(K_N)$, we know $d(v) = N-1$. Then if $N \geq r(s-1,t) + r(s,t-1) + 1$, we get at least either $r(s-1,t)$ red edges on v or $r(s,t-1)$ blue edges on v.

If there are $r(s-1,t)$ red edges on v, then the set $N_{\text{red}}(v)$ of vertices joined to v by a red edge contains either a red K_{s-1} or a blue K_t by induction. This is because $|N_{\text{red}}(v)| \geq r(s-1,t)$. Now if there is a blue K_t contained in the vertices of $N_{\text{red}}(v)$, we're done. Otherwise, the red K_{s-1} contained in the vertices of $N_{\text{red}}(v)$ combined with v itself form K_s.

As for if there are $r(s,t-1)$ blue edges on v, then we can do similar reasoning to show that K_n contains either a red K_s or blue K_t.

So $r(s,t) \leq r(s-1,t) + r(s,t-1)$ for $s \geq 2$ and $t \geq 2$. Now let's get the non-recursively defined inequality.

We already said that $r(2,t), r(s,2) \leq \binom{s+t-2}{s-1}$. Now once again proceed by induction, assuming $s \geq 3$, $t \geq 3$, and that the theorem is true when s or t is smaller.

$$\begin{aligned} r(s,t) &\leq r(s-1,t) + r(s,t-1) \\ &\leq \binom{s+t-3}{s-2} + \binom{s+t-3}{s-1} \end{aligned}$$

Now use Pascal's Triangle Identity: $\displaystyle\binom{a}{b} = \binom{a-1}{b} + \binom{a-1}{b-1}$.

Here is a quick algebraic proof of this identity:

$$\begin{aligned} \tbinom{a-1}{b} + \tbinom{a-1}{b-1} &= (a-1)!\left(\tfrac{1}{b!(a-b-1)!} + \tfrac{1}{(b-1)!(a-b)!}\right) \\ &= (a-1)!\left(\tfrac{(a-b)}{b!(a-b)!} + \tfrac{b}{b!(a-b)!}\right) \\ &= \tfrac{a!}{b!(a-b)!} = \tbinom{a}{b} \end{aligned}$$

Letting $a = s+t-2$ and $b = s-1$, we can say that:

$$\binom{s+t-3}{s-2} + \binom{s+t-3}{s-1} = \binom{s+t-2}{s-1}$$

As a consequence of this, we've shown that Ramsey numbers are well-defined for all $s, t \in \mathbb{Z}^+$.

Here's how to get an easier-to-write upper bound for $r(s,t)$:

$$\binom{s+t-2}{s-1} = \frac{(s+t-2)!}{(s-1)!(t-1)!} = \frac{(s+t-2)(s+t-3)\cdots(t)}{(s-1)!}$$
$$\leq \frac{(s+t-2)^{s-1}}{(s-1)!} \leq (s+t-2)^{s-1}$$

In particular, this means that $r(3,t) \leq (t+1)^2$. So $r(3,t)$ has a quadratically growing upper bound. That said, this is not the best upper bound that has been found. Recently, it was found that:

$$\frac{(t+1)^2}{4\log(t)} \leq r(3,t) \leq \frac{(t+1)^2}{\log(t)}$$

Also, professor Verstraete coauthored a paper very recently (two weeks before this lecture) showing that:

$$\frac{(t+1)^3}{10000(\log(t))^4} \leq r(4,t) \leq \frac{(t+1)^3}{\log(t)}$$

And there is an unsolved conjecture stating that $r(s,t)$ is of order $\frac{(t)^{s-1}}{(\log(t))^{c_s}}$ where c_s is a constant depending on s.

Now a reason we care about bounds is that trying to find exact Ramsey numbers is incredibly difficult. Here are all the currently known Ramsey numbers:

- $r(2,t) = t$

- $r(3,3) = 6$
- $r(3,4) = 9$
- $r(3,5) = 14$
- $r(3,6) = 18$
- $r(3,7) = 23$

- $r(3,8) = 28$
- $r(3,9) = 36$

- $r(4,4) = 18$
- $r(4,5) = 25$

Theorem (Erdős, 1947): $r(t,t) \geq \sqrt{2}^t$ for $t \geq 3$.

Proof:
Let $N = \lfloor \sqrt{2}^t \rfloor$. Color $E(K_N)$ randomly such that each edge independently has $\frac{1}{2}$ probability to be red rather than blue.

The average number of red K_t is: $\mathbb{E}(\text{\# of red } K_t) = \binom{N}{t}\left(\frac{1}{2}\right)^{\binom{t}{2}}$

Explanation:
There are $\binom{N}{t}$ possible groupings of t vertices in K_N. Then $\left(\frac{1}{2}\right)^{\binom{t}{2}}$ is the probability that every edge in a grouping of t vertices will be colored red.

Similarly, $\mathbb{E}(\text{\# of blue } K_t) = \binom{N}{t}\left(\frac{1}{2}\right)^{\binom{t}{2}}$

Thus the average number of red or blue K_t is:

$$2\binom{N}{t}\left(\frac{1}{2}\right)^{\binom{t}{2}} \leq 2\frac{N^t}{t!}\cdot 2^{\frac{-t(t-1)}{2}} \leq 2\frac{(\sqrt{2})^{t^2}}{t!}\cdot 2^{\frac{-t^2}{2}+\frac{t}{2}} = \frac{2}{t!}\cdot 2^{\frac{t}{2}}$$

When $t \geq 3$, $\frac{2}{t!}\cdot 2^{\frac{t}{2}} < 1$ (you can visually check that on Desmos). This tells us that there must be a red-blue coloring of K_N with no red K_t or blue K_t.

Side note: this was the first proof ever done using probabilistic combinatorics.

Using the Erdős-Szekeres theorem, we can also get the following upper bound:

$$r(t,t) \leq \binom{2t-2}{t-1} \leq \frac{(2t-2)!}{(t-1)!(t-1)!} \leq \frac{(2(t-1))^{2t-2}}{((t-1)^{t-1})^2} \quad \text{(because } n! \leq n^n\text{)}$$

$$\leq \frac{2^{2t-2}(t-1)^{2t-2}}{(t-1)^{2t-2}}$$

$$\leq 2^{2t-2} = 4^{t-1}$$

Thus $r(t,t)$ grows exponentially with t.

Discussion Section Additional Notes

Generalization of Ramsey numbers:

Let H and F be graphs. Then $r(H,F)$ is the minimum N such that any red-blue coloring of $E(K_N)$ will have either a red H or blue F. Note that $r(s,t) = r(K_s, K_t)$.

An equivalent way of thinking about Ramsey numbers is through the complements of graphs. Let G be an n-vertex graph. Then the complement of G is $\overline{G}$ where $\overline{G}$ has the same vertices as G and all the edges that G doesn't have.

Then $r(H,F)$ is the minimum N such that for any graph G with $|V(G)| \geq N$, $H \nsubseteq G \Longrightarrow F \subseteq \overline{G}$.

Some tips for solving Ramsey number problems:

- Restate the problem.
- Think of basic lower bound constructions such as making one of the colors form a collection of cliques or a bipartite graph.
- Remember the pigeonhole principle.

Example 1: Prove that every red-blue coloring of K_n contains a monochromatic spanning tree.

Let $G = \mathrm{red}(K_n)$. We showed earlier that any connected graph with finite vertices has a spanning tree. So if G is connected, we're done. Otherwise, we know that $\overline{G} = \mathrm{blue}(K_n)$ is connected. So $\overline{G}$ has a spanning tree. ■

Example 2: Prove that $\left\lceil \frac{3(k-1)}{2} \right\rceil \leq r(P_k, P_k) \leq 2k$.

Upper Bound:

By the Erdős-Gallai Theorem: $\mathrm{ex}(n, P_k) \leq (k-1)\frac{n}{2}$. Let G have n vertices. If G is P_k-free, then $|E(G)| \leq (k-1)\frac{n}{2}$, which means that:

$$|E(\overline{G})| \geq \frac{n(n-1)}{2} - \frac{n(k-1)}{2} = \frac{n}{2}(n-k).$$

If $n \geq 2k$, then $\frac{n}{2}(n-k) \geq \frac{n}{2}(k) > \frac{n}{2}(k-1)$. So $\overline{G}$ contains P_k. Hence, $R(P_k, P_k) \leq 2k$.

Lower Bound:

Now consider a graph G consisting of two cliques of size $k-1$ and l. If $l \leq k-1$, then G does not contain P_k. Meanwhile note that $\overline{G} = K_{k-1,l}$. If $l < \frac{k-1}{2}$ then the longest path in $\overline{G}$ is $2l+1 < k$.

If k is odd, then letting $l = \frac{k-1}{2} - 1$, we have that $|V(G)| = k + \frac{k-1}{2} - 2$. Thus:

$$r(P_k, P_k) \geq k + \frac{k-1}{2} - 1 = \left\lceil \frac{3(k-1)}{2} \right\rceil.$$

If k is even, then letting $l = \frac{k}{2} - 1$, we have that $|V(G)| = k - 2 + \frac{k}{2}$. Thus:

$$r(P_k, P_k) \geq k - 1 + \frac{k}{2} = \left\lceil \frac{3(k-1)}{2} \right\rceil. \blacksquare$$

Lecture 16: 3/12/2024

Example 3: Prove that $r(K_{1,2}, K_t) = 2t - 1$.

If we red-blue color K_N and there is no red $K_{1,2}$, then $\mathrm{red}(K_N)$ forms a matching M. Letting $N = 2t-1$, we can find t vertices with no red edges between them. Specifically, pick one vertex from each match in M plus some number of vertices outside the matching. Thus $\mathrm{blue}(K_n)$ contains k_t, meaning $r(K_{1,2}, K_t) \leq 2t - 1$

Now color $E(K_{2t-2})$ with a red perfect matching and all other edges blue. Then the largest blue clique is K_{t-1}. So $r(F,G) > 2t-2$. ■

Example 4: Prove that $r(C_4, K_t) < 4t^2$ for $t \geq 1$.

In a red-blue coloring of $E(K_N)$, if there is no red C_4, then by the Kövari-Sós-Turán theorem: $|E(\text{red})| \leq \frac{1}{2}(N^{\frac{3}{2}} + N)$. Meanwhile, by Turán's Theorem, if there is no blue K_t, then $|E(\text{blue})| \leq \binom{t-1}{2} \cdot \frac{N^2}{(t-1)^2} = \frac{t-2}{t-1}\left(\frac{N^2}{2}\right)$.

We did not prove this at the time, but even when $r \not| \, n$:

$$|E(T_r(n))| \leq \left(\frac{n}{r}(r-1)\right)\frac{n}{2} = \binom{r}{2}\frac{n^2}{r^2} = \frac{n^2}{2}(1-\frac{1}{r}).$$

Putting those together, we have that: $E(K_N) = \binom{N}{2} \leq \frac{1}{2}(N^{\frac{3}{2}} + N) + \frac{t-2}{t-1}\left(\frac{N^2}{2}\right)$.

$$\binom{N}{2} \leq \frac{1}{2}(N^{\frac{3}{2}} + N) + \frac{t-2}{t-1}\left(\frac{N^2}{2}\right) \implies \frac{N^2-2N-N^{\frac{3}{2}}}{2} \leq \frac{t-2}{t-1} \cdot \frac{N^2}{2}$$
$$\implies -(\tfrac{1}{2}N^{\frac{3}{2}} + N) \leq \frac{-1}{t-1} \cdot \frac{N^2}{2}$$
$$\implies \tfrac{1}{2}N^{\frac{3}{2}} + N \geq \frac{1}{t-1} \cdot \frac{N^2}{2}$$

For $N \geq 4$, we apparently have that $\frac{1}{2}N^{\frac{3}{2}} \geq N$. Now we can easily check that:

- $r(C_4, K_1) = 0 < (t+1)^2$ when $t = 1$
- $r(C_4, K_2) = 4 < (t+1)^2$ when $t = 2$
- $r(C_4, K_t) > 4$ when $t \geq 3$.

So since we now have permission to focus on when $t \geq 3$, we can assume $N \geq 4$ without contradicting our previous assumptions.

$\frac{1}{t-1} \cdot \frac{N^2}{2} \leq \frac{1}{2}N^{\frac{3}{2}} + N \leq 2 \cdot \frac{1}{2}N^{\frac{3}{2}}$. Therefore $\sqrt{N} \leq 2(t-1)$, which means that $N \leq 4(t-1)^2 < 4t^2$. So $r(C_4, K_t) \leq 4t^2$. ■

The best lower bound known is $r(C_r, K_t) \geq ct^{\frac{3}{2}}\sqrt{\log(t)}$ where $c > 0$ is a constant.

Here is an application of Ramsey theory:

Schur's Theorem: If $n \geq 2$ and $p > \lfloor e \cdot n! \rfloor + 1$ is prime, then $x^n + y^n \overset{p}{\equiv} z^n$ has a solution modulo p.

Proof:

Start by defining $r_k(3)$ to be the minimum N such that if we color $E(K_N)$ with k colors, we will get a monochromatic triangle.

$r_2(3) = r(3,3) = 6$ $\qquad$ $r_3(3) = 17$ (this is the highest exact value known)

Theorem: For $k \geq 2, \;\; r_k(3) \leq \lfloor e \cdot k! \rfloor + 1$

Proof:
Let us k-color $E(K_N)$ and proceed by induction on k.

Assume that K_N has no monochromatic triangles. Then pick any $u \in V(K_N)$. By the pigeonhole principle, there must be some color with at least $\frac{N-1}{k}$ edges touching u. Letting U be the set of neighbors of u in that color, we have that $|U| \geq \frac{N-1}{k}$.

Now if for any two vertices $x, y \in U$, we had that $\{x, y\}$ is colored the same as $\{x, u\}$ and $\{y, u\}$, then we would have a monochromatic triangle between x, y, and u. So the edges in U must have at most $k-1$ colors. Since, there must not be a monochromatic triangle in any of the other colors, we know $|U| < r_{k-1}(3)$.

Thus, we know that $\frac{N-1}{k} < r_{k-1}(3)$. Or in other words, $\frac{N-1}{k} \leq r_{k-1}(3) - 1$. Rewriting this, we get that $N \leq kr_{k-1}(3) - k + 1$. So:

$$r_k(3) - 1 \leq k(r_{k-1}(3) - 1) + 1.$$

Let's define $a_k = r_k(3) - 1$ and rewrite the expression as $a_k \leq \frac{k!}{k!} + \frac{k!}{(k-1)!}a_{k-1}$. Then magically:

$$a_k \leq \frac{k!}{k!} + \frac{k!}{(k-1)!}a_{k-1} \leq \frac{k!}{k!} + \frac{k!}{(k-1)!} + \frac{k!}{(k-2)!}a_{k-2} \leq \cdots \leq \sum_{i=2}^{k} \frac{k!}{i!} + k! \cdot a_1$$

$a_1 = r_1(3) - 1 = 2$. So, we can rewrite $k! \cdot a_1$ as $\frac{k!}{1!} + \frac{k!}{0!}$. Finally:

$$a_k \leq \sum_{i=0}^{k} \frac{k!}{i!} = k! \cdot \sum_{i=0}^{k} \frac{1}{i!} < k! \cdot e$$

Thus, $r_k(3) \leq \lfloor e \cdot k! \rfloor + 1$.

For a lower bound on $r_k(3)$, consider recursively splitting K_N into a colored complete bipartite graph colored and 2 uncolored remaining cliques. Then repeat on the remaining cliques.

If $N \leq 2^k - 1$, we can color all the edges of K_N with k different colors. And because all the colors are bipartite, there is no monochromatic odd cycle including a triangle. So $r_k(3) \geq 2^k$.

Now, we want to show that if we color the integers in $S = \{1, 2, \ldots, N\}$ with k colors, there must exist $x, y, z \in S$ such that all three have the same color and $x + y = z$ if N is large enough.

So for $N = |S|$, give each vertex in $V(K_N)$ a distinct integer from S. Then assign any coloring $c : \{1, 2, \ldots, N\} \to \{1, 2, \ldots, k\}$ to S and finally color each edge $\{v_i, v_j\} \in E(K_N)$ the color $c(|i - j|)$. Having done that, we have k-colored $E(K_N)$.

If $N \geq r_k(3)$, then there must be a monochromatic triangle $\{v_{i_1}, v_{i_2}, v_{i_3}\}$. Or in other words, $c(|i_1 - i_2|) = c(|i_2 - i_3|) = c(|i_3 - i_1|)$.

Let $\{x, y, z\} = \{|i_1 - i_2|, |i_2 - i_3|, |i_3 - i_1|\}$. Then x, y, and z all have the same color. Additionally, without loss of generality, if we assume that $i_1 < i_2 < i_3$, then $|i_1 - i_2| + |i_2 - i_3| = |i_3 - i_1|$. So we can have that $x + y = z$.

Now let θ be a multiplicative generator of $\mathbb{Z}_p \setminus \{0\}$ where p is a prime number.

The definition of a "cyclic group" G with the operation $\bullet$ is that $\exists g \in G$ such that letting $g^k = g \bullet g \bullet \cdots \bullet g$ (a.k.a g '$\bullet$ed' with itself k times), for every $h \in G$ there exists $n \in \mathbb{N}$ such that $h = g^n$.

While we did not talk about this in math 109, hopefully it is trivial to see that $\mathbb{Z}_p$ would be a cyclic group when equipped with the operation $+$. Less obviously, $\mathbb{Z}_p \setminus \{0\}$ is also a cyclic group when equipped with the operation $\times$ (multiplication).

Each element $x \in \mathbb{Z}_p \setminus \{0\}$ can be written as θ^{a_x} where $1 \leq a_x \leq p - 1$.

$\theta^p \overset{p}{\equiv} \theta$ by Fermat's little theorem. So there is no reason to assign a_x to any integer greater than $p - 1$.

Thus, color each element of $\{1, 2, \ldots, p-1\}$ with the color $a_x \bmod n$ where $x = \theta^{a_x}$. This is an n-coloring of $\{1, 2, \ldots, p-1\}$. And as we showed above, if $p - 1 \geq \lfloor e \cdot n! \rfloor + 1 \geq r_k(3)$, then there exists $x, y, z \in \{1, 2, \ldots, p-1\}$ with the same color that also satisfy $x + y = z$.

So we've found $\theta^{a_x} + \theta^{a_y} \overset{p}{\equiv} \theta^{a_z}$ where for constants $\alpha, \alpha', \alpha''$, and β, we have that $a_x = \alpha n + \beta,\ \ a_y = \alpha' n + \beta$, and $a_z = \alpha'' n + \beta$. Thus canceling out θ^β from both sides of our equation, we get that:

$$(\theta^\alpha)^n + (\theta^{\alpha'})^n \overset{p}{\equiv} (\theta^{\alpha''})^n$$

Hence we have found a solution to $x^n + y^n \overset{p}{\equiv} z^n$. ■

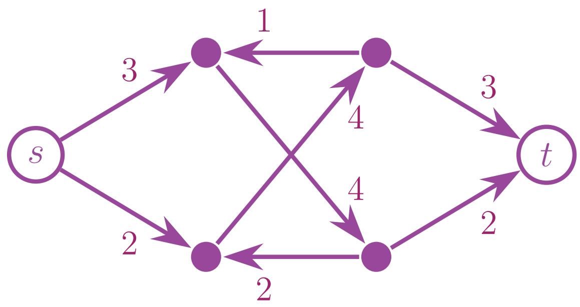
A flow in a digraph G is a map $f : E(G) \to \mathbb{R}^{\geq 0}$ such that for every $v \in V(G)$:

$$\text{flow out } = \sum_{(v,w) \in E(G)} f(v, w) = \sum_{(w,v) \in E(G)} f(w, v) = \text{ flow in}$$

(This is called Kirchoff's Law of Conservation of Flows.)

To make the flow more interesting, we consider G to have a source $s \in V(G)$ and sink $t \in V(G)$ where the conservation law doesn't need to hold.

Here's an example of flow in a digraph:



Note that repeat edges [i.e. $\{(x,y),(x,y)\} \subseteq E(G)$] are completely redundant. There is no reason to ever have them. However, loops [i.e. (x,x)] can be used to "store" flow at a vertex. Also antiparallel edges [i.e. $(x,y),(y,x)$] let a flow go either way.

If s is the source of a flow f, then the value of the flow f is:

$$v(f) = \sum_{v \in N^+(s)} f(s,v) \; - \sum_{v \in N^-(s)} f(v,s) = \text{ net flow leaving } s.$$

A capacity function is a map $c : E(G) \to \mathbb{R}^{\geq 0}$ such that if $f : E(G) \to \mathbb{R}$ is a flow, then $f(v,w) \leq c(v,w)$ for all $(v,w) \in E(G)$.

Note that if you want to have more than one source, you can treat several sources as all being connected to one supersource. A similar trick also works if you want multiple sinks.

Lecture 17: 3/14/2024

Let G be a digraph and f a flow with a source s and sink t. If we have sets S and T that partition $V(G)$ such that $s \in S$ and $t \in T$, then we can define the s-t cut: (S,T), to be the spanning subgraph of G containing all the edges going out of S into T.

Theorem: For any flow f and s-t cut (S,T), $v(f) = \sum_{\substack{(x,y)\in E(G) \\ x \in S \\ y \in T}} f(x,y) \; - \sum_{\substack{(x,y)\in E(G) \\ x \in S \\ y \in T}} f(y,x)$.

Proof:
By Kirchoff's law, for all $x \in S \setminus \{s,t\}$, $\sum_{y \in N^+(x)} f(x,y) = \sum_{y \in N^-(x)} f(y,x)$.

Thus, $\sum_{x \in S \setminus \{s\}} \left(\sum_{y \in N^+(x)} f(x,y) \right) = \sum_{x \in S \setminus \{s\}} \left(\sum_{y \in N^-(x)} f(y,x) \right)$.

Meanwhile, by the definition of $v(f)$: $\sum_{y\in N^+(s)} f(s,y) = \sum_{y\in N^-(s)} f(y,s) + v(f)$.

So, $\sum_{x\in S}\left(\sum_{y\in N^+(x)} f(x,y)\right) = \sum_{x\in S}\left(\sum_{y\in N^-(x)} f(y,x)\right) + v(f)$.

Now for each $x \in S$ and $y \in N^+(x)$, if we also have that $y \in S$, then $f(x,y)$ will be added on both sides of the equation. So we can cancel out those edges, leaving us with the equation of the theorem.

Corollary: If t is the sink of a flow f, then $-v(f) = \sum_{v\in N^+(t)} f(t,v) - \sum_{v\in N^-(t)} f(v,t)$.

Proof:
Take $S = V(G) \setminus \{t\}$ and $T = \{t\}$. Then by the above theorem,
$\sum_{v\in N^-(t)} f(v,t) = \sum_{v\in N^+(t)} f(t,v) + v(f)$.

Given a capacity function c and s-t cut (S,T), we define the capacity of an s-t cut as:

$$c(S,T) = \sum_{e\in(S,T)} c(v,w).$$

Max Flow Min Cut Theorem: The maximum value of any flow equals the minimum capacity of any s-t cut.

Proof:
Let c be the capacity function for a digraph G with a source s and sink t.
For any flow f and s-t cut (S,T), we must have that $v(f) \leq c(S,T)$.

To see why, first note that:

$$v(f) = \sum_{\substack{(x,y)\in E(G)\\ x\in S\\ y\in T}} f(x,y) - \sum_{\substack{(x,y)\in E(G)\\ x\in S\\ y\in T}} f(y,x) \leq \sum_{\substack{(x,y)\in E(G)\\ x\in S\\ y\in T}} f(x,y) \leq \sum_{\substack{(x,y)\in E(G)\\ x\in S\\ y\in T}} c(x,y)$$

Now note that $\{(x,y) \mid (x,y) \in E(G), x \in S, y \in T\} = E((S,T))$.
So we can rewrite our above inequality as $v(f) \leq c(S,T)$.

This tells us that $v(f)$ is at most the minimum value of $c(S,T)$. So, what's left to show is that for some flow f, $v(f)$ equals the minimum value of $c(S,T)$.

Let f be a maximal flow in G. Then define a set $S \subseteq V(G)$ as follows:

- Start by putting s into S.
- Then repeatedly until there are no more satisfactory y:
    - If $(x,y) \in E(G)$, $x \in S$, and $f(x,y) < c(x,y)$, then put y into S.
    - If $(y,x) \in E(G)$, $x \in S$, and $f(y,x) > 0$, then put y into S.

Suppose $t \in S$. Then there exists a path P with vertices: $v_1, v_2, \ldots, v_k$, going from s to t (potentially including edges directed in opposing directions) such that $v_i \in S$ for each i and exactly one of the following inequalities is defined and true for every edge in the path:

$$f(v_i, v_{i+1}) < c(v_i, v_{i+1}) \text{ or } f(v_{i+1}, v_i) > 0,$$

Let ε be the smallest of the defined $c(v_i, v_{i+1}) - f(v_i, v_{i+1})$ and $f(v_{i+1}, v_i)$. Then we can define a valid flow $g : E(G) \to \mathbb{R}^{\geq 0}$ as:

$$g(x,y) = \begin{cases} f(x,y) + \varepsilon & \text{if } (x,y) \in P \;\; s.t. \;\; x = v_i \text{ and } y = v_{i+1} \\ f(x,y) - \varepsilon & \text{if } (x,y) \in P \;\; s.t. \;\; x = v_{i+1} \text{ and } y = v_i \\ f(x,y) & \text{otherwise} \end{cases}$$

Then $v(g) = v(f) + \varepsilon$, contradicting that f was a maximal flow.

So we conclude that $t \notin S$, which means that letting $T = V(G) \setminus S$, we know:

$$v(f) = \sum_{\substack{(x,y) \in E(G) \\ x \in S \\ y \in T}} f(x,y) \; - \sum_{\substack{(x,y) \in E(G) \\ x \in S \\ y \in T}} f(y,x)$$

Finally, by how we constructed S, we know that for $(x,y) \in E(G)$ such that $x \in S$ and $y \in T$, we must have that $f(x,y) = c(x,y)$. Meanwhile, for $(x,y) \in E(G)$ such that $x \in T$ and $y \in S$, we must have that $f(x,y) = 0$. Therefore:

$$v(f) = c(S,T) - 0 = c(S,T). \blacksquare$$

Note that this proof can also be interpretted as an iterative algorithm for finding a maximal flow and minimum cut.

- Start with any valid flow f_1.
- Now inductively for $n \geq 1$, create a set S using the flow f_n as seen in the proof. If while creating S you ever put t into S, then define a new flow f_{n+1} via the formula seen above.
- If you finish making S without ever putting in t, then you're done because you've found a minimum cut and maximal flow.

I would like to caution that we have not showed that this algorithm terminates for a general capacity function $c : E(G) \to \mathbb{R}^{\geq 0}$. In fact, as it is written out now, it will not terminate for some c. However, due to the next corollary, this is not a problem in this class.

Corollary: Given a capacity function c, if $c(e)$ is an integer for all $e \in E(G)$, then there exists a maximal flow f such that $f(e) \in \mathbb{Z}$ for all $e \in E(G)$. Furthermore, by making f_1 map every edge of G to 0, the algorithm described on the previous page will terminate with such an integral maximal flow in a finite number of steps.

Proof:
Let $c : E(G) \to \mathbb{Z}^{\geq 0}$ and define $f_1 : E(G) \to \{0\}$. Then do the algorithm described on the previous page.

Assuming c is not also a zero function, we know that the algorithm will do at least two iterations. So we can consider that when $n \geq 2$, we have that $v(f_n) = v(f_{n-1}) + \varepsilon$ where $\varepsilon > 0$. This tells us that the sequence $v(f_n)$ is strictly increasing.

Also, note that since either $\varepsilon = f_n(e)$ or $\varepsilon = c(e) - f_n(e)$ for each iteration that that the algorithm doesn't terminate on, if both f_n and c_n are functions mapping $E(G)$ to the integers, then ε will be an integer. This tells us two things.

1. By induction, every f_n will map $E(G)$ to the integers if f_1 and c both map $E(G)$ to the integers. Thus, if this algorithm does terminate, it will terminate on an integral maximal flow.
2. $(v(f_n))$ is a strictly increasing sequence of integers.

Finally note that $v(f_n)$ is both bounded below by 0 and bounded above by the capacity of the minimum s-t cut which also must be an integer because the sum of many integers is still an integer. Thus, there are finitely many values that $v(f_n)$ could be. This combined with the fact that $(v(f_n))$ is strictly increasing means that there must be finitely many f_n. So the algorithm must terminate in a finite number of steps, giving us in the process a maximal flow $f : E(G) \to \mathbb{Z}^{\geq 0}$. ■

Interestingly, the max flow min cut theorem is very closely related to Stokes theorem, although showing that is outside the scope of this class.

To conclude the class, let's revist matching theory and Hall's Theorem.

Here is an algorithm for finding an optimal matching in a bipartite graph which uses the max flow min cut theorem.

If G be a bipartite graph with parts A and B. Then define the digraph G^* such that its vertex set is $V(G) \cup \{s, t\}$. Additionally, make $E(G^*)$ consist of all the edges of G directed from A to B in addition to adding every possible edge directed from s to A and every possible edge directed from B to t.

Next, define a capacity function c assigning each edge of G a capacity of 1, and make s and t the source and sink respectively of G.

Having done all that, define $f_1 : E(G^*) \to \{0\}$ and then do the algorithm from page 58 using f_1 as your starting flow. We know that this will eventually converge onto a maximal flow $f : E(G^*) \to \mathbb{Z}^{\geq 0}$.

Claim: The set of edges: $\{\{a, b\} \in E(G) \mid a \in A, b \in B,$ and $f(a, b) = 1\}$, is an optimal matching.

Proof:

Firstly, note that if the optimal matching M of G has size N, then we can define the flow:

$$g(x, y) = \begin{cases} 1 & \text{if } \{x, y\} \in M \\ 1 & \text{if } x = s \text{ and } \{y, z\} \in M \\ 1 & \text{if } y = t \text{ and } \{z, y\} \in M \\ 0 & \text{otherwise} \end{cases}$$

The value of this flow is N. So we know $v(f) \geq N$.

Because $0 \leq f(x, y) \leq 1$ and $f(x, y) \in \mathbb{Z}$ for every $(x, y) \in E(G^*)$, we know that f maps every edge to either 0 or 1. Thus if $v(f) = M$, then there must be exactly M edges going from s to A which f maps to 1.

In turn, this implies that there are M edges going from A to B which f maps to 1. Because flow can not go backwards from B to A, these are the only edges with nonzero flow between A and B. Now note that each of these edges must be associated with a unique vertex of A since each vertex of A has only one incoming edge. Also, each of these edges must be associated with a unique vertex of B because each vertex of B only has one outgoing edge.

So, each of the M edges going from A to B that have a flow of 1 are disjoint. Hence, they form a matching of size M over G.

If $M > N$, then this contradicts that N is the size of the optimal matching of G. Thus, as we already established that $M \geq N$, we must have that $M = N$. So, an optimal matching in G is formed by the M edges in the set:

$$\{\{a, b\} \in E(G) \mid a \in A, b \in B, \text{ and } f(a, b) = 1\},$$

Finally, here is a second proof for Hall's Theorem which states that given a bipartite graph G with parts A and B, G has a matching saturating A if and only if for every set $X \subseteq A,\ |N_G(X)| \geq |X|$.

Proof:
($\Longrightarrow$) This direction is trivial to prove. So we don't care about it here.

($\Longleftarrow$) Assume G is a bipartite graph with parts A and B satisfying Hall's condition. Then construct the digraph G^* just like in the algorithm covered right before this. If we can show that for any maximal flow f in G^*, we have that $v(f) = |A|$, then we know that our algorithm for finding an optimal matching will give us a matching saturating A.

Define (S, T) to be any s-t cut of G^*. Then $c(S, T) = |E((S, T))|$ and $E((S, T))$ consists of all edges going from s to $(A \setminus S)$, $(A \cap S)$ to $(B \setminus S)$, and $(B \cap S)$ to t. Thus if k is the number of edges going between $(A \cap S)$ and $(B \setminus S)$, then $c(S, T) = |A \setminus S| + k + |B \cap S|$.

Consider if $|B \cap S| > |A \cap S|$. Then we're already done because:

$$c(S,T) \geq |A \setminus S| + k + |B \cap S| \geq |A \setminus S| + |B \cap S| > |A \setminus S| + |A \cap S| = |A|$$

So it is safe to assume that $|B \cap S| \leq |A \cap S|$. In that case, note that by Hall's condition, $|B \cap N(A \cap S)| \geq A \cap S$. Furthermore, in the worst case that $B \cap S \subseteq B \cap N(A \cap S)$, we still have that:

$$|(B \setminus S) \cap N(A \cap S)| \geq |A \cap S| - |B \cap S|.$$

k is minimized if there are as few edges going to $(B \setminus S) \cap N(A \cap S)$ from $A \cap S$ as possible. So $k \geq |(B \setminus S) \cap N(A \cap S)| \geq |A \cap S| - |B \cap S|$. Hence, we have once again shown that $c(S, T) \geq |A \setminus S| + |A \cap S| = |A|$.

So $c(S, T) \geq |A|$. Now finally to show that there exists (S, T) such that $c(S, T) = |A|$, consider $S = \{s\}$ and T containing everything else. For that (S, T), we have that $c(S, T) = |A|$.

Hence the minimum capacity of a cut of G^* is $|A|$, which means that the max value of a flow in G^* is $|A|$. ■

The textbook for this class was the professor's course notes.
Also, the T.A. Nicholas Sieger was immensely helpful.