CCS Discrete II	Professor: Padraic Bartlett

Lecture 4: Graph Theory and the Four-Color Theorem

Week 4 | *UCSB 2015*

Through the rest of this class, we're going to refer frequently to things called **graphs**! If you haven't seen graphs before, we define them here:

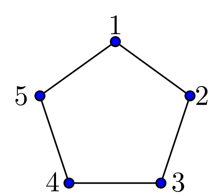
Definition. A **graph** G with n vertices and m edges consists of the following two objects:

1. a set $V = \{v_1, \ldots v_n\}$, the members of which we call G's **vertices**, and
2. a set $E = \{e_1, \ldots e_m\}$, the members of which we call G's **edges**, where each edge e_i is an unordered pair of distinct elements in V, and no unordered pair is repeated. For a given edge $e = \{v, w\}$, we will often refer to the two vertices v, w contained by e as its endpoints.

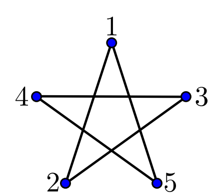
Example. The following pair (V, E) defines a graph G on five vertices and five edges:

- $V = \{1, 2, 3, 4, 5\}$,
- $E = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}, \{5, 1\}\}$.

Something mathematicians like to do to quickly represent graphs is **draw** them, which we can do by taking each vertex and assigning it a point in the plane, and taking each edge and drawing a curve between the two vertices represented by that edge. For example, one way to draw our graph G is the following:



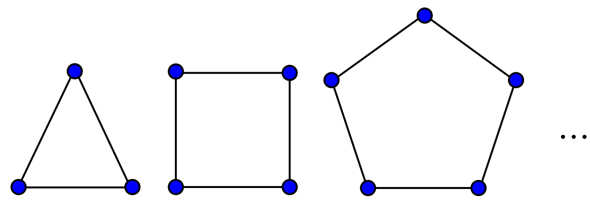
We could also draw our graph like this:



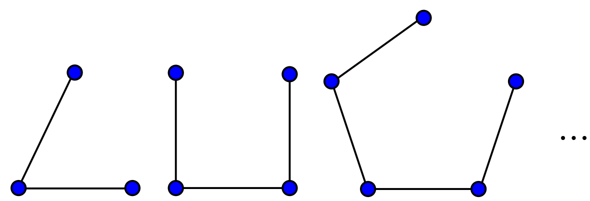
In general, all we care about for our graphs is their vertices and their edges; we don't usually care about how they are drawn, so long as they consist of the same vertices connected via the same edges. Also, we usually will not care about how we "label" the vertices of a graph: i.e. we will usually skip the labelings on our graphs, and just draw them as vertices connected by edges.

Some graphs get special names:

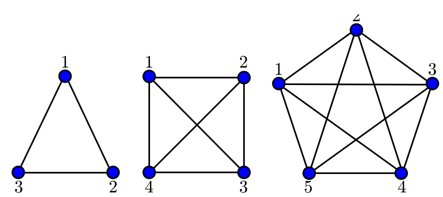
Definition. The **cycle graph** on n vertices, C_n, is the graph on the vertex set $\{v_1, v_2, \ldots v_n\}$ with edge set $E(C_n) = \{\{v_1, v_2\}, \{v_2, v_3\}, \ldots \{v_{n-1}, v_n\}, \{v_n, v_1\}\}$. The cycle graphs C_n can be drawn as n-gons, as depicted below:



Definition. The **path graph** on n vertices, P_n, is the graph on the vertex set $\{v_1, v_2, \ldots v_n\}$ with edge set $E(C_n) = \{\{v_1, v_2\}, \{v_2, v_3\}, \ldots \{v_{n-1}, v_n\}\}$. The path graphs P_n can be drawn as paths of length n, as depicted below:



Definition. The **complete graph** K_n**.** The complete graph on n vertices, K_n, is the simple graph on the vertex set $\{v_1, v_2, \ldots v_n\}$ that has every possible edge: in other words, $E(K_n) = \{\{v_i, v_j\} : i \neq j\}$. We draw several of these graphs below:



.

Every vertex in a K_n has degree $n-1$, as it has an edge connecting it to each of the other $n-1$ vertices; as well, a K_n has $n(n-1)/2$ edges in total in it, by the degree-sum formula. (Explicitly: every vertex has degree $n-1$ and there are n vertices, therefore the sum of the degrees of K_n's vertices is $n(n-1)$. We've shown that this quantity is twice the number of edges in the graph; dividing by 2 then tells us that the number of edges in K_n is $n(n-1)/2$, as claimed.)

Definition. The **complete bipartite graph** $K_{n,m}$**.** The complete bipartite graph on $n+m$ vertices with part sizes n and m, $K_{n,m}$, is the following graph:

- $V(K_{n,m}) = \{v_1, v_2, \ldots v_n, w_1, w_2, \ldots w_m\}$.
- $E(K_{n,m})$ consists of all of the edges between the n-part and the m-part; in other words, $E(K_{n,m}) = \{(v_i, w_j) : 1 \leq i \leq n, 1 \leq j \leq m\}$.

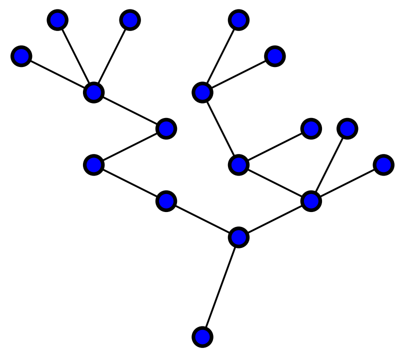
The vertices v_i all have degree m, as they have precisely m edges leaving them (one to every vertex w_j); similarly, the vertices w_j all have degree n. By either the degree-sum formula or just counting, we can see that there are nm edges in $K_{n,m}$.

Definition. Given a graph G and another graph H, we say that H is a **subgraph** of G if and only if $V(H) \subset V(G)$ and $E(H) \subset E(G)$.

Definition. Given a graph G, we call G **connected** if for any two vertices $x, y \in V(G)$, there is a path that starts at x and ends at y in our graph G.

Definition. If a graph G has no subgraphs that are cycle graphs, we call G **acyclic**. A **tree** T is a graph that's both connected and acyclic. In a tree, a **leaf** is a vertex whose degree is 1.

Example. The following graph is a tree:



1 The Four-Color Theorem

Graph theory got its start in 1736, when Euler studied the Seven Bridges of Königsberg problem. However, I claim that it first blossomed in earnest in 1852 when Guthrie came up with the Four-Color Problem.

Theorem. Take any map, which for our purposes is a way to partition the plane $\mathbb{R}^2$ into a collection of connected regions $R_1, \ldots R_n$ with continuous boundaries. There is some way to assign each region R_i to a color in the set $\{R, G, B, Y\}$, such that if two regions R_i, R_j are "touching" (i.e. they share some nonzero length of boundary between them,) then those two regions must receive different colors.

We're going to prove this theorem in this class! First, some background:

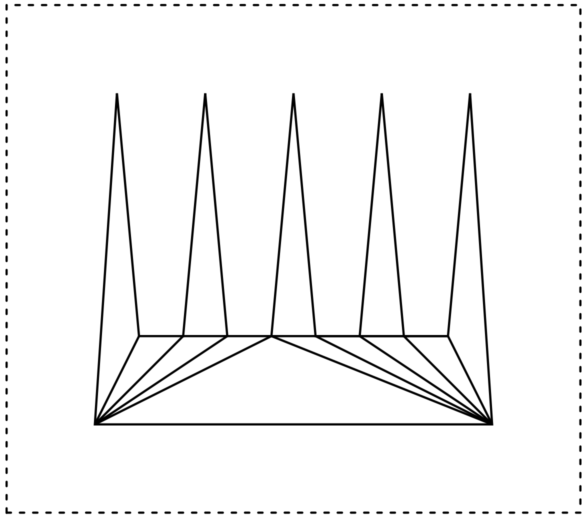
Definition. Take any map M. We can turn this into a graph as follows:

- Assign to each region R_i a vertex v_i.
- Connect v_i to v_j with an edge if the regions R_i, R_j are touching.

We call this graph the **dual graph** to M.

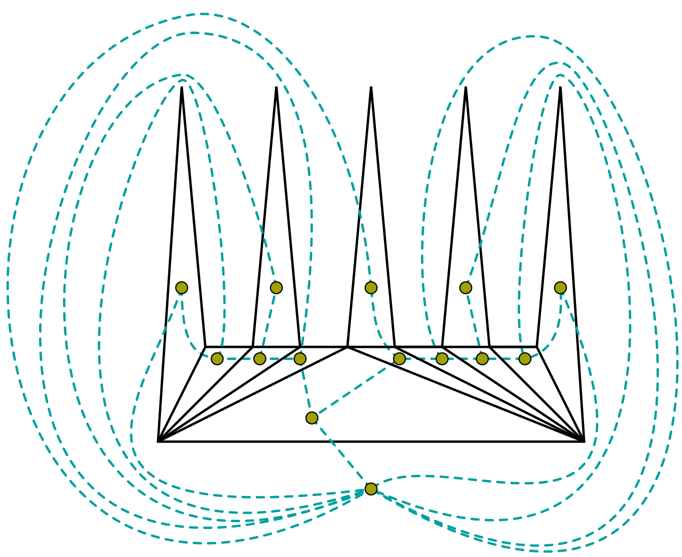
We give an example here:

Example. Consider the following map:



This map consists of 14 regions. If you count, you can see that the figure drawn consists thirteen triangles; as well, we have the "outer" region consisting of everything else left over, which forms a very strange 15-gon.

Now, take each region, and assign to it a vertex. As well, connect two regions sharing a border with an edge: this will give you the following graph, with edges given by the dashed teal lines and vertices given by the yellow dots:



Note how we have drawn the edges so that they connect two adjacent countries by traveling through the border that they share! This observation is useful to recall when thinking about our second definition:

Definition. We say that a graph G is **planar** if we can draw it in the plane so that none of its edges intersect.

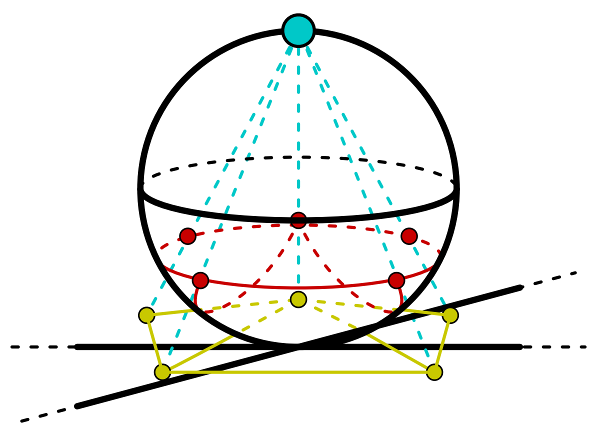
Sometimes, it will help to think of planarity in the following way:

Definition. We call a connected graph G **planar** if we can draw it on the sphere S^2 in the following fashion:

- Each vertex of G is represented by a point on the sphere.
- Each edge in G is represented by a continuous path drawn on the sphere connecting the points corresponding to its vertices.
- These paths do not intersect each other, except for the trivial situation where two paths share a common endpoint.

We call such a drawing a **planar embedding** of G on the sphere.

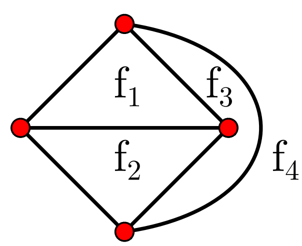
It is not hard to see that this definition is equivalent to our earlier definition of planarity. Simply use the stereographic projection map (drawn below) to translate any graph on the plane to a graph on the sphere:



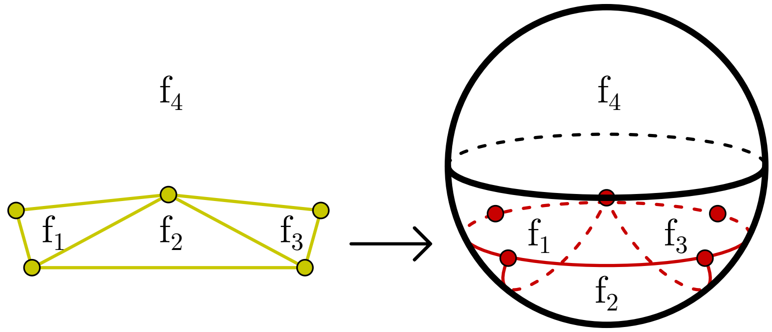
By drawing lines from the "north pole" (0,0,1) through points either in the xy-plane or on the surface of the sphere, we can translate graphs drawn on the sphere (in red) to graphs drawn in the plane (in yellow.)

Definition. For any connected planar graph G, we can define a **face** of G to be a connected region of $\mathbb{R}$ whose boundary is given by the edges of G.

For example, the following graph has four faces, as labeled:



Notice that we always have the "outside" face in these drawings, which can be easy to forget about when drawing our graphs on the plane. This is one reason why I like to think about these graphs as drawn on the sphere; in this setting, there is no "outside" face, as all of the faces are equally natural to work with.



This observation has a nice accompanying lemma:

Lemma. Take any connected planar graph G, and any face F of G. Then G can be drawn on the plane in such a way that F is the outside face of G.

Proof. Take a planar embedding of G on the unit sphere. Rotate this "drawn-upon" sphere so that the face F contains the north pole $(0, 0, 1)$ of the sphere. Now, perform stereographic projection to create a planar embedding of G in $\mathbb{R}^2$. By construction, the face F is now the outside face, which proves our claim. $\square$

It bears noting that not all graphs are planar:

Proposition. The graph K_5 is not planar.

Proof. Draw a 5-cycle on the sphere. If the edges of this 5-cycle do not intersect each other, then the resulting pentagon partitions the sphere into two parts, each part of which is bounded by this pentagon. Take either one of these parts; notice that within that part, we can draw at most two nonintersecting edges connecting nonadjacent vertices in that part. Consequently, it is impossible to draw the additional 5 edges required to create K_5 without using overlapping edges. Therefore it is impossible to find a planar embedding of K_5 on the sphere, as claimed! $\square$

For maps, something you may have noticed is the following:

Observation. The dual graph to any map M is connected and planar.

Proof. On the homework! $\square$

The reason we care about this is that it gives us the following more graph-theoretic way to describe the four-color theorem:

Theorem. Take any connected planar graph on finitely many vertices. There is a way to assign each of its vertices one of the four colors $\{R, G, B, Y\}$ such that no edge in this graph has both endpoints colored the same color.

In general, this concept of **coloring** comes up all the time in graph theory! We give it a name here:

Definition. A graph G is called k**-colorable** if there is a collection of k distinct colors that we can map the vertices of G to, so that no edge in G has both endpoints colored the same color. Given a graph G, we define the **chromatic number** of G, $\chi(G)$, as the smallest number k such that G is k-colorable.

This gives us one last rephrasing of the four-color theorem:

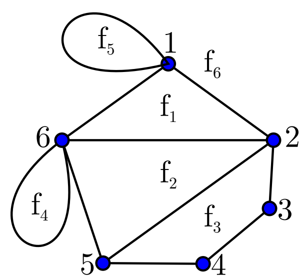
Theorem. If G is a connected planar graph on finitely many vertices, then $\chi(G) \leq 4$.

...So. Before we can start Kempe's proof, we need one last bit of background, which is the concept of Euler characteristic:

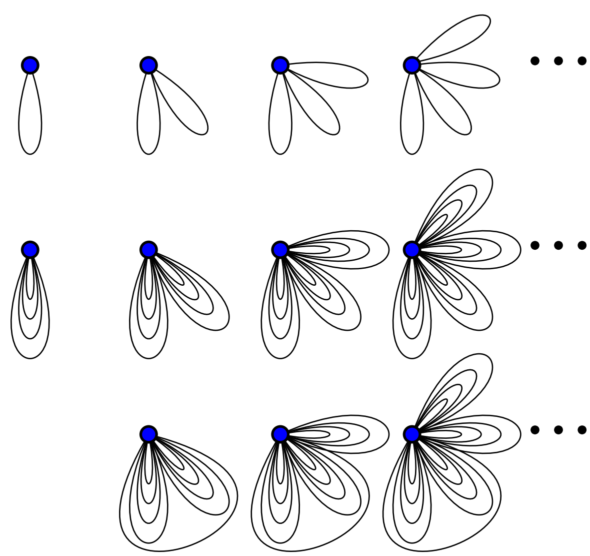
Theorem. (Euler characteristic.) Take any connected graph that has been drawn in $\mathbb{R}^2$ as a planar graph. Then, if V is the number of vertices, E is the number of edges, and F is the number of faces in this graph, we have the following relation:

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

Proof. We will actually prove a stronger claim: we will show that any planar **multigraph** (a graph, but where we allow multiple edges between vertices, and also edges that start and end at the same vertex) satisfies the $V - E + F = 2$ formula. For the rest of this proof, we will assume that graph and multigraph are synonymous; once we are done with this proof, though, we will stop assuming this.

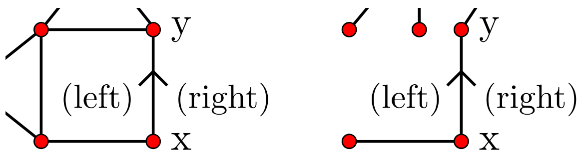


We proceed by induction on the number of vertices. Suppose that $V = 1$. Then our graph looks like something of the following form:



I claim that $V - E + F = 2$ for any of these graphs, and prove it by a second induction on the number of edges. For a zero-edge graph, this is easy; we have one vertex, no edges and one face, we have $V - E + F = 1 - 0 + 1 = 2$. Now, assume via induction that every one-vertex multigraph on n edges has $V - E + F = 2$. Take any graph on one vertex with $n + 1$ edges. Pick one of these edges, and look at it.

I claim that this edge borders exactly two faces. To see why, take any edge, and assign an orientation to it (i.e. if our edge is $\{x, y\}$, then orient the edge so that we travel from x to y.) If you do this, then our edge has two "sides," the left- and right-hand sides, if we travel along it via this orientation.

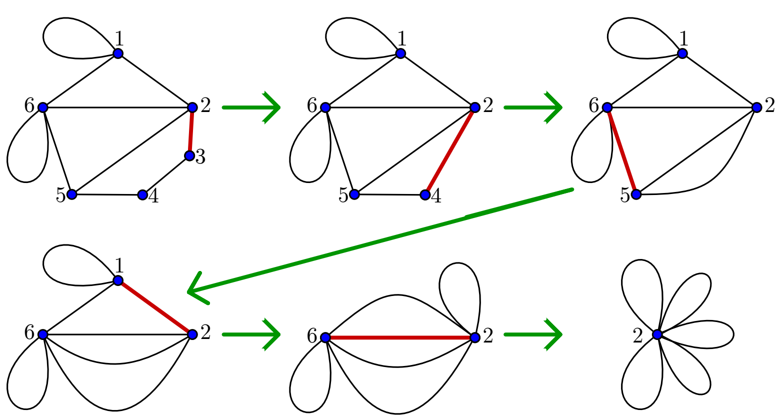


There are two possibilities, as drawn above: either the left- and right-hand sides are different, or they are the same. This tells us that our edge either borders one or two faces! To see that we have exactly two, we now recall that our edge (because our graph has exactly one vertex) must start and end at the same vertex. In other words, it is a **closed** loop: i.e. its outside is different from its inside! In other words, our left- and right-hand sides are different, and our edge separates two distinct faces.

Therefore, deleting this edge does the following things to the graph: it decreases our edge count by 1, and also decreases our face count by 1 (as we merge two faces when we delete this edge.) In other words, deleting this edge does not change $V - E + F$! But by induction we know that $V - E + F = 2$ for all 1-vertex graphs on n edges, which is what we get if we delete this edge from a $n + 1$-edge graph. So we're done!

This settles our base case for our larger induction on V, the number of vertices. We now go to the second phase of an inductive proof: we show how to reduce larger cases to smaller cases!

To do this, consider the following operation, called **edge contraction**. Take any edge with two distinct endpoints. Delete this edge, and combine its two endpoints together: this gives us a new graph! We draw examples of this process below: we start with a graph on six vertices, and contract one by one the edges labeled in red at each step.



Contracting an edge decreases the number of vertices by 1 at each step, as it "squishes together" two adjacent vertices into one vertex. It also decreases the number of edges by 1 at each step, as we are contracting an edge to a point! Finally, it never changes the number of faces; if two faces were distinct before this process happens, they stay distinct, as we're not making any cuts in any of our boundaries (and instead are just shrinking them partially a bit!)

But this means that $V-E+F$ is still constant! Therefore, by induction, if $V-E+F$ holds for every n-vertex multigraph, it holds for any $n+1$-vertex multigraph by just contracting an edge! This finishes our induction, and thus our proof. $\square$

Using this theorem, we can prove the following useful lemma, which is the only part of the Euler characteristic property that we need for our graph:

Lemma. Take any connected planar graph G. Then there is some vertex v in our graph with degree at most 5.

Proof. We proceed by contradiction. Assume that every vertex has at least degree 6; we will create a contradiction to the claim that $V - E + F = 2$.

First, consider the sum $\sum_{v \in G} \deg(v)$. On one hand, this is twice the number of edges in G: this is because each edge shows up twice in this sum (once for each endpoint v when we're calculating $\deg(v)$.) On the other hand, if each vertex has degree at least 6, we have

$$\sum_{v \in G} \deg(v) \geq \sum_{v \in G} 6 = 6V.$$

Consequently, we have $2E \geq 6V$, and therefore $E/3 \geq V$.

Similarly: notice that every face F of our planar graph must have at least three edges bounding it, because our faces are made out of edges in our graph. Also, if we sum over all faces the number of edges in each face, we get again at most twice the number of edges; this is because each edge is in at most two faces (as discussed earlier!) Therefore, we have

$$2E \geq \sum_{f \in G} \text{facedeg}(f) \geq \sum_{f \in G} 3 = 3F,$$

and therefore that $2E/3 \geq F$.

Therefore, we have

$$2 = V - E + F \leq E/3 - E + 2E/3 = 0,$$

which is clearly impossible. Therefore, we have a contradiction, and can conclude that our initial assumption — that all vertices have degree at least 6 — is false! $\square$

2 Kempe's Proof

With this notation set up, Kempe's proof of the four-color theorem is actually fairly straightforward! We give it here.

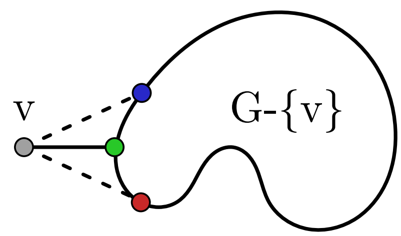
Proof. We proceed by contradiction. Assume not: that there are connected planar graphs on finitely many vertices that need at least 5 colors to be colored properly. Consequently, there must be some **smallest** connected planar graph G, in terms of the number of its vertices, that needs at least five colors to color its vertices! Pick such a graph G. Notice that if we remove any vertex v from G, we have a graph on a smaller number of vertices than G. Consequently, the graph $G \setminus \{v\}$ **can** be colored with four colors!

Let v be the vertex in G with degree at most 5. Delete v from G: this leaves us a graph that we can four-color. Do so.

Our goal is now the following: to add v back in and (by possibly changing the coloring of $G \setminus \{v\}$) give v one of our four colors, so that we have a four-coloring of G! This will prove that our initial assumption — that a G can exist that needs five colors — is false, and therefore prove our theorem.

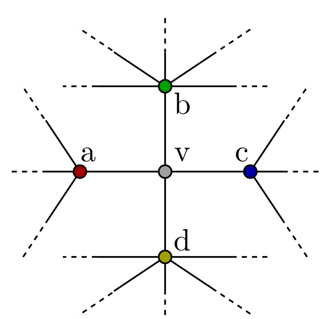
We proceed by cases, considering v's possible degrees.

1. v has degree 1, 2 or 3. In these cases, note that when we add v back in, it is adjacent to at most three colors! So there is some fourth color left over. Give v that color.



2. v has degree 4. In this case, there are two possibilities:

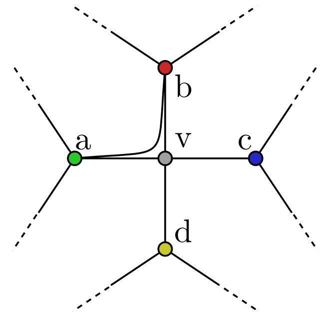
   - In the four neighbors a, b, c, d of v, some color is not used. In this case, we are in the same kind of situation as above: just color v with the color that doesn't show up in its neighbors?
   - In the four neighbors a, b, c, d of v, each color is used exactly once. So, up to the names of the colors, we are in the following situation:

   First, notice that without losing any generality we may assume that there are edges $a \leftrightarrow b \leftrightarrow c \leftrightarrow d \leftrightarrow a$. To see why, notice the following:

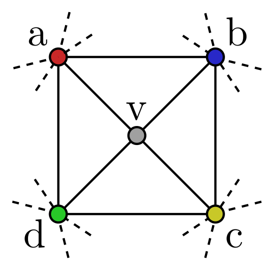
   (a) Adding these edges does not change the assumed property that $\chi(G) \geq 5$, as extra edges only makes it "harder" for us to color a graph.

(b) We can always draw in these edges if they do not exist: for example, if the edge $\{a, b\}$ did not exist, we could add it in without breaking planarity by simply drawing a path that is "very close" to the two edges $\{a, v\}$ and $\{v, b\}$. This path will not cross other edges, as there are no other edges leaving v.



(c) Therefore, because we have preserved $\chi(G) \geq 5$ and G 's planarity, we can put these edges into G without changing any of our arguments thus far!

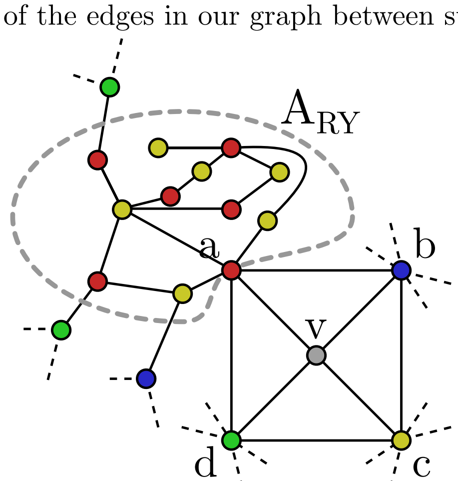
By the argument above, we can now assume our graph looks like the following:



Now, do the following: for any two colors C_1, C_2, let G_{C_1,C_2} denote the subgraph of G given by taking all of the vertices in G that are colored either C_1 or C_2, along with all of the edges that connect C_1 vertices to C_2 vertices.

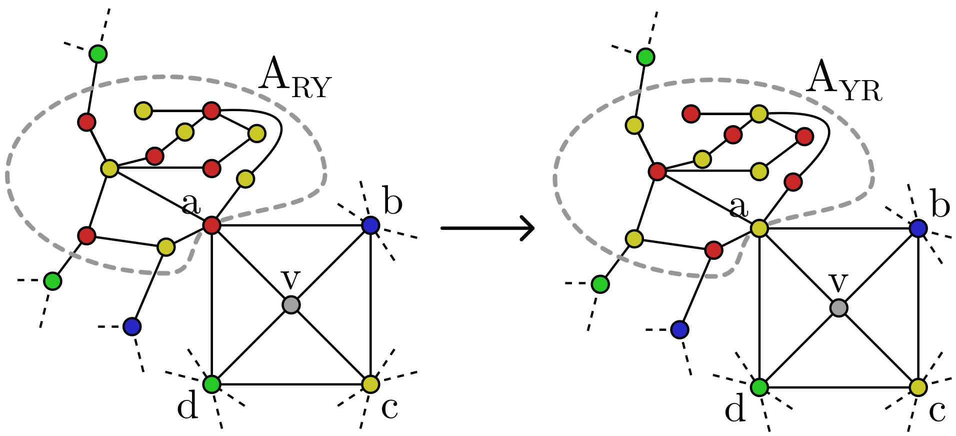
Look at the red-yellow subgraph G_{RY}. In this graph, there are two possibilities:

(a) There is no path from a to c in this graph. In other words, define A_{RY} as the subgraph of G_{RY} given by taking all of the G_{RY} vertices that have paths to a, along with all of the edges in our graph between such vertices.



Suppose that we "switch" the colors red and yellow in the subgraph A_{RY}. Does it create any issues with our coloring?

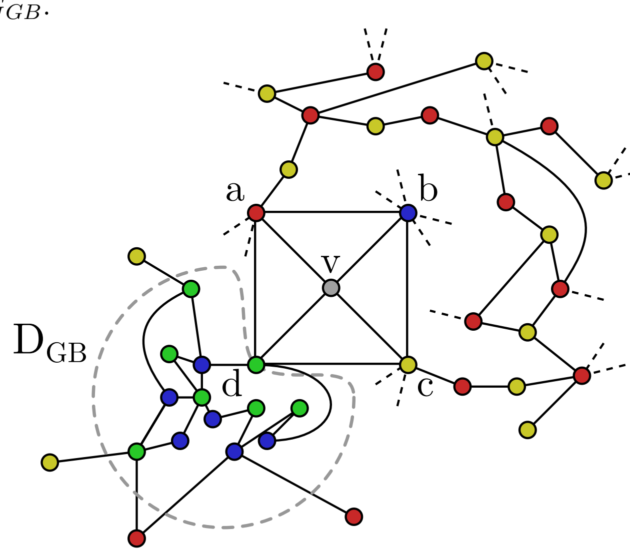
Let's check. No edge between two vertices in A_{RY} is broken (i.e. has both endpoints made the same color) by this process; before it had one red and one yellow endpoint, and now it has one yellow and one red endpoint. As well, no edge that involves no vertices in A_{RY} is broken by this process, as we did not change the colors of either of their endpoints!



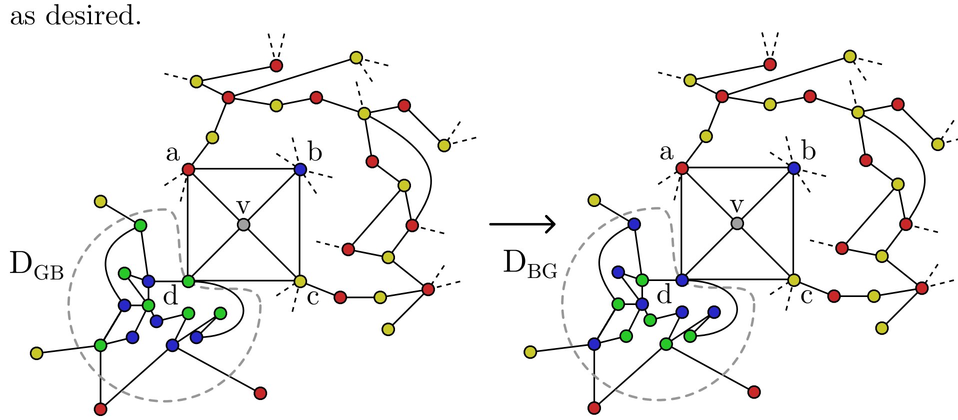
Finally, consider any edge with one endpoint in A_{RY} and another endpoint not in A_{RY}. In order for this edge to have one endpoint in A_{RY} and another not in A_{RY}, one endpoint must be red or yellow (the endpoint in our set) and the other must be green or blue (the endpoint not in our set!) So if we switch red and yellow in A_{RY}, this edge is also not broken!

No edges are broken by this swap; therefore we still have a valid coloring. Furthermore, in this coloring, v has no neighbors that are red; so we can color v red and have a four-coloring of our entire graph G!

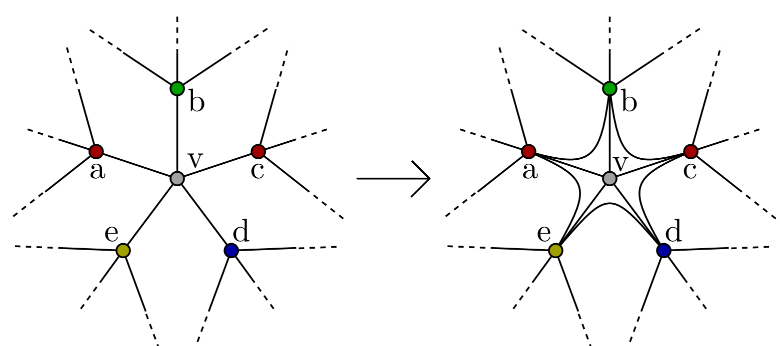
(b) Alternately, (a) does not happen. In this case, there is a path from a to c made entirely of red-yellow vertices linked by edges. In this case: look at the graph G_{GB}.



In particular, notice that there cannot be a path from b to d along green-blue edges, because our graph is planar and any such path would have to cross our red-yellow edges! Therefore, we can define D_{GB} to be the collection of all of the G_{GB} vertices that have paths to d, along with the edges in our graph between such vertices. As noted above, $d \notin B_{GB}$.
Now, switch the colors G and B in D_{GB}! This causes no conflicts, by exactly the same argument as above, and yields a graph where v has no green neighbor; therefore, we can give v the color green, and have a proper four-coloring as desired.



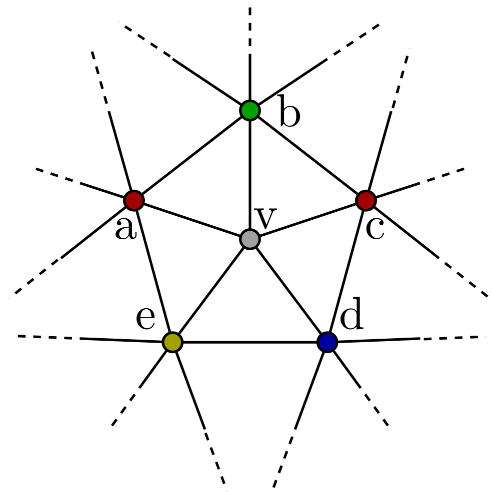
3. v has degree 5. Again, as before, we can assume that all four of the colors in our graph occur on G's neighbors, because if they do not we can simply give v whichever color is missing. Again, as before, we can assume that the neighbors of v are connected by the following pentagonal structure.:



This is because of the following:

- Adding edges to our graph will never make it easier to color a graph: all they do is give us more conditions on what vertices have to have different colors, which only makes coloring **harder**.
- Furthermore we can add these edges without breaking planarity by simply drawing them arbitrarily close to the v-edges.

Up to symmetry and colorings, then, we are in the following situation:



This is because we have to repeat one color (so it might as well be red,) we have to use all of the other colors (so we have green, blue and yellow in some order,) red cannot occur on two adjacent vertices (because there are edges between adjacent vertices,) and therefore up to rotation and flipping we have the above.

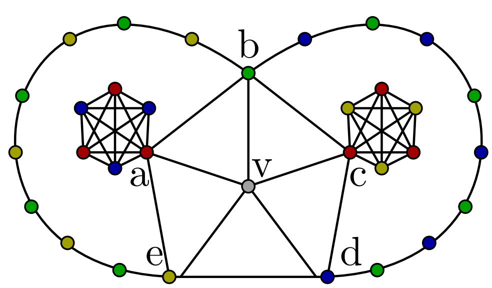
Do the following:

(a) First, look at the G_{GB} subgraph. Either the vertex b is not connected to d in this subgraph, in which case we can do the switching-trick that we discussed earlier. Otherwise, b **is** connected to d in G_{GB}, and we have a green-blue chain from b to d.

(b) Now, look at the G_{GY} subgraph. Similarly, either the vertex b is not connected to e in this subgraph, in which case we can do the switching-trick that we discussed earlier, or it is, and we have a green-yellow chain from b to e.

If we were able to switch in either of the two cases above, then v has only three colors amongst its neighbors, and we can color it with whatever color remains.

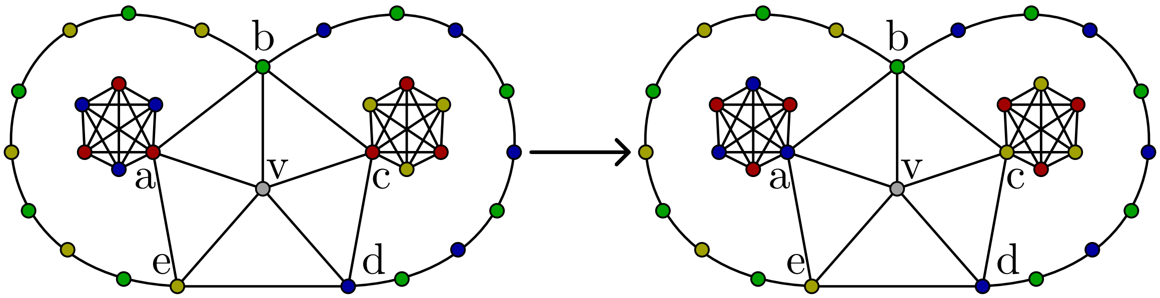
Otherwise, we are in the following case:



Do the following:

(a) First, look at the G_{RB} subgraph. Because of the green-yellow chain, the vertices a and d are not connected to each other. Therefore, we can switch red and blue in the a-connected part of this subgraph!

(b) Now, look at the G_{RY} subgraph. Because of the green-blue chain, the vertices c and e are not connected to each other. Therefore, we can switch red and yellow in the c-connected part of this subgraph!



This yields a graph where v has no red neighbor: consequently, we can color v red, which gives us a proper four-coloring! This proves our claim.

□

3 Plot Twist

This proof . . . is in fact Kempe's famous flawed proof of the four-color theorem, which stood for 11+ years before being disproven! In particular, it was **disproven**: i.e. the proof you've read above is **false**!

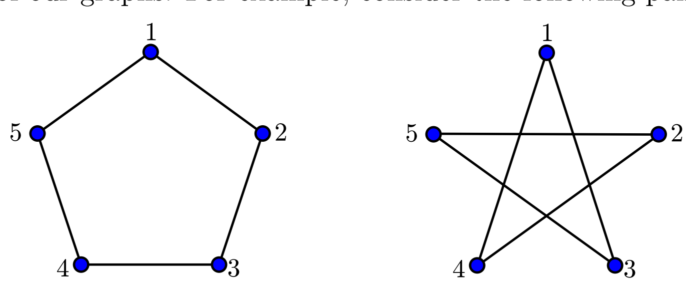
Guess what your HW is for today?

4 Graphs: Terminology

We transition from the four-color theorem to a more careful discussion of the definitions and terms we need to discuss graphs. We start by discussing what it means for two graphs to be the "same," in the same sense that we talked about groups or fields or vector spaces or sets being the same: **isomorphisms**!

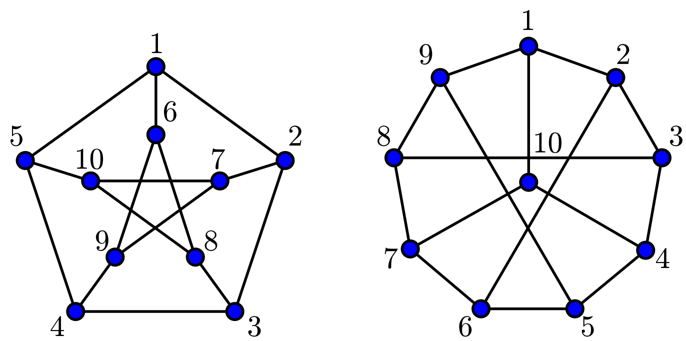
Which is to say the following: In our first examples of graphs, we labeled all of our vertices because this is part of the definition of what a graph *is* – a collection of labeled vertices and edges between them.

However, when we had to actually do tricky mathematics — the four-color theorem — we stopped labeling our graphs! And in general, I claim that we don't really care about the labelings of most of our graphs. For example, consider the following pair of graphs:



These graphs are, in one sense, different; the first graph has an edge connecting 1 to 2, where the second graph does not. However, in another sense, these graphs are representing the same situation: they're both depicting the graph sketched out by a pentagon!

For graphs like the ones in our menagerie, we don't care so much about the labeling of the vertices; rather, the interesting features of these graphs are the intersections of their edges and vertices. In other words, we want to say that both of the graphs below are "the" Petersen graph[1]: even though they initially look rather different, there is a way of "relabeling" the vertices on the second graph so that (i, j) is an edge in the first graph iff it's an edge in the relabeled second graph.



How can we do this? What notion can we introduce that will allow us to regard such graphs as being the "same," in a well-defined sense? Well, consider the following:

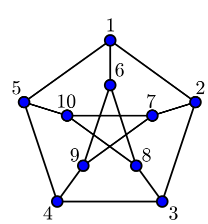
Definition. We say that two graphs G_1, G_2 are **isomorphic** if and only if there is a map $\sigma : V(G_1) \to V(G_2)$ such that

- σ matches each element of $V(G_1)$ to a unique element of $V(G_2)$, and vice-versa: in other words, σ is a way of relabeling G_1's vertices with G_2's labels, and vice-versa.
- $\{v_i, v_j\}$ is an edge in G_1 if and only if $\{\sigma(v_1), \sigma(v_2)\}$ is an edge in G_2.

We will often regard two isomorphic graphs as being the "same," and therefore refer to graphs like K_n or the Petersen graph without specifying or worrying about what the vertices are labeled.

A concept that's much more interesting (given the idea of isomorphism) is the concept of a **subgraph**, which we define below:

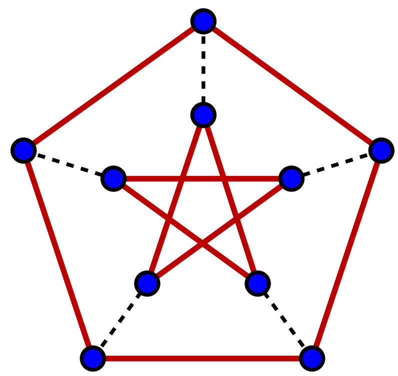
[1]**The Petersen graph** P The Petersen graph P is a graph on ten vertices, drawn below:



The vertices in P all have degree three; by counting or the degree-sum formula, P has 15 edges. It is notorious as being the counterexample or example to a number of conjectures in graph theory!

Definition. Given a graph G and another graph H, we say that H is a **subgraph** of G if and only if $V(H) \subset V(G)$ and $E(H) \subset E(G)$.

Example. The Petersen graph has the disjoint union of two pentagons $C_5 \sqcup C_5$ as a subgraph, which we shade in red below:



In general, when we ask if a graph H is a subgraph of a graph G, we won't mention a labeling of H's vertices; in this situation, we're actually asking whether there is *any* subgraph of G that is **isomorphic** to H.

For example, one question we could ask is the following: what kinds of graphs contain an triangle (i.e. a C_3) as a subgraph? Or, more generally, what kinds of graphs contain an odd cycle (i.e. a C_{2k+1}) as a subgraph?

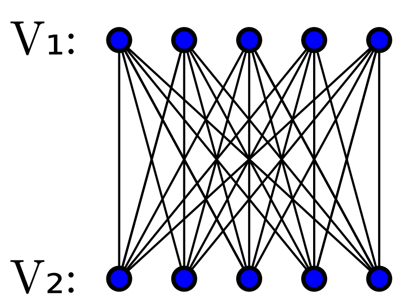
We answer this in the next section:

5 Classifying Bipartite Graphs

Definition. We call a graph G **bipartite** if and only if we can break the set $V(G)$ up into two parts $V_1(G)$ and $V_2(G)$, such that every edge $e \in E(G)$ has one endpoint in $V_1(G)$ and one endpoint in $V_2(G)$.

Alternately, we say that a graph is bipartite iff there is some way to color G's vertices red and blue – i.e. to take every vertex in G and assign it either the color blue or color red, but not both or neither – so that every edge has one blue endpoint and one red endpoint.

Example. The following graph is bipartite, with indicated partition (V_1, V_2):



However, there are graphs that are not bipartite; for example, C_3, the triangle, is not bipartite! This is not very hard to see: in any partition of C_3's vertices into two sets V_1 and V_2, one of the two sets V_1 or V_2 has to contain two vertices of our triangle. Therefore, there is an edge in C_3 with both endpoints in one of our partitions; so this partition does not make C_3 bipartite. Because this holds for every possible partition, we can conclude that no such partition exists – i.e. C_3 is not bipartite!

In general, we can say much more:

Proposition 1. *C_{2k+1} is not bipartite.*

Proof. We will prove this proposition with a proof by contradiction. In other words, we will assume that C_n is bipartite, and from there we'll deduce something we know to be false; from there, we can conclude that our assumption must not have been true in the first place (as it led us to something false,) and therefore that C_n is not bipartite.

To do this: as stated, we'll suppose for contradiction that C_{2k+1} is bipartite. Then, there must be some way of coloring the vertices $\{v_1, \ldots v_{2k+1}\}$ of C_n red and blue, so that no edge is monochrome (i.e. has two red endpoints or two blue endpoints.)

How do we do this? Well: look at v_{k+1}. v_{k+1} has to be either red or blue: without any loss of generality[2], we can assume that it's red. Then, because no edge in C_1 is monochrome, we specifically know that none of v_{k+1}'s neighbors can be red: in other words, they both have to be blue! So both v_k and v_{k+2} are blue.

Similarly, we know that neither of v_k or v_{k+2}'s neighbors can be blue: so both v_{k-1} and v_{k+3} have to be red! Repeating this process, we can see that

- v_{k+1} being red forces
- v_k, v_{k+2} to be blue, which forces
- v_{k-1}, v_{k+3} to be red, which forces
- v_{k-2}, v_{k+4} to be blue, which forces
- ...
- which forces v_1, v_{2k+1} to both be the same color.

But there is an edge between v_1 and v_n in C_{2k+1}! This contradicts the definition of bipartite: therefore, we've reached a contradiction. Consequently, C_n cannot be bipartite. □

This allows us to actually classify a large number of graphs as not being bipartite:

Proposition 2. *If a graph G has a subgraph isomorphic to C_{2k+1}, then G is not bipartite.*

[2] The phrase "without loss of generality" is something mathematicians are overly fond of. In general, it's used in situations where there is some sort of symmetry to the situation that allows you to assume that a certain situation holds: for example, in this use, we're assuming that v_1 is red because it has to be either red or blue, and if it was blue we could just switch the colors "red" and "blue" through the entire proof.

Proof. Suppose that G contains a subgraph H that's not bipartite. Then, for any coloring of H's vertices, there is some edge in H that's monochrome. Therefore, because any coloring of G's vertices into two parts will also color H's vertices, we know that any coloring of G's vertices with the colors red and blue will create a monochrome edge; therefore, G cannot be bipartite. $\square$

Is this it? Or are there other ways in which a graph can fail to be bipartite? Surprisingly, as it turns out, there isn't:

Proposition 3. *A graph G on n vertices is bipartite if and only if none of its subgraphs are isomorphic to an odd cycle.*

Proof. Our earlier proposition proved the "if" direction of this claim: i.e. if a graph is bipartite, it doesn't have any odd cycles as subgraphs. We focus now on the "only if" direction: i.e. given a graph that doesn't contain any odd cycles, we seek to show that it is bipartite.

First, note the following definitions:

Definition. A graph G is called **connected** iff for any two vertices $v, w \in V(G)$, there is a path connecting v and w.

Definition. Given a graph G, divide it into subgraphs $H_1, \ldots H_k$ such that each of the subgraphs H_i are connected, and for any two H_i, H_j's there aren't any edges with one endpoint in H_i and one endpoint in H_j. These parts H_i are called the **connected components** of G; a graph G is connected if it has only one connected component.

Definition. For a graph G and two vertices v, w we define the **distance** $d(v, w)$ between v and w as the number of edges of the smallest path connecting v and w. For a connected graph, this quantity is always defined, $d(v, v) = 0$, and $d(v, w) > 0$ for any $v \neq w$.

Take our graph G, and divide it into its connected components $H_1, \ldots H_k$. If we can find a red-blue coloring of each connected component H_i that shows it's bipartite, we can combine all of these colorings to get a coloring of all of G; because there are no edges between the connected components, this combined coloring would show that G itself is bipartite!

Therefore, it suffices to just show that any connected graph H on n vertices without any odd cycles in it is bipartite. To do this, take any vertex $y \in V(H)$, and construct the following sets:

- $N_0 = \{w : d(v, y) = 0\}$
- $N_1 = \{w : d(v, y) = 1\}$
- $N_2 = \{w : d(v, y) = 2\}$
- $\ldots$
- $N_n = \{w : d(v, y) = n\}$

First, notice that every vertex v shows up in at least one of these sets, as H is connected and has n vertices (and thus, any path in H has length $\leq n$.) Furthermore, no vertex shows up in more than one of these sets, because distance is well-defined. Finally, notice that for any $x \in N_k$ and any path P given by $y = v_0 e_{01} v_1 e_{12} \ldots e_{k-1,k} v_k = x$, each of the vertices v_j lies in N_j. This is because each of these has a path of length j from y to v_j (just take our path and cut it off at v_j), and has no shorter path (because if there was a shorter path, we could use it to get from y to x in less than k steps, and therefore $d(y, x)$ would not be k.)

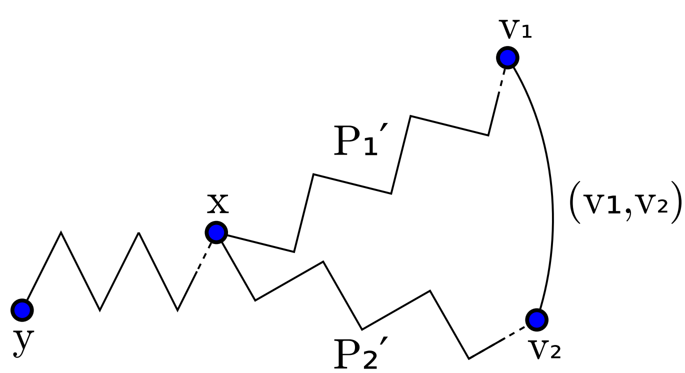
Now, color all of the vertices in the even N-sets red, and all of the vertices in the odd N-sets blue. We claim that there are no monochromatic edges.

To see this, take any edge $\{v_1, v_2\}$ in our graph H. Let $d(y, v_1) = k$ and $d(y, v_2) = l$, P_1 be a path of length k connecting v_1 with y, and P_2 be a path of length l connecting v_2 with y. These paths may intersect repeatedly, so take x to be the furthest-away vertex from y that's in both of these paths. Let P_1' be the path that we get by starting P_1 at x and proceeding to v_1, and P_2' be the path that we get by starting P_2 at x and proceeding to v_2.

There are two possiblities. Either x is one of v_1 or v_2, in which case (because there's an edge from v_1 to v_2) the distance from y to v_1 is either one greater or one less than the distance from y to v_2. In either case, v_1 and v_2 have different colors (because our colors alternated between red and blue as our distance increased,) so this edge is not monochrome.

Otherwise, x is neither v_1 or v_2. In this case, look at the cycle formed by doing the following:

- Start at x, and proceed along P_1'.
- Once we get to v_1, travel along the edge $\{v_1, v_2\}$.
- Now, go backwards along P_2' back to x.



This is a cycle, because P_1' and P_2' don't share any vertices in common apart from x. What is its length? Well, the length of P_1' is just $d(y, v_1) - d(y, x)$, the length of P_2' is just $d(y, v_2) - d(y, x)$, and the length of a single edge is just 1; so, the total length of this path is

$$d(y, v_1) - d(y, x) + d(y, v_2) - d(y, x) + 1 = d(y, v_1) + d(y, v_2) - (2d(y, x) + 1).$$

We know that this cannot be odd, because our graph has no odd cycles; so the number above is even! Because $(2d(y, x) + 1)$ is odd, this means that $d(y, v_1) + d(y, v_2)$ must also be odd; in other words, exactly one of $d(y, v_1), d(y, v_2)$ can be odd ,and exactly one can be even. But this means specifically that exactly one must be blue and one must be red (under our coloring scheme,) so our edge must not be monochromatic.

Therefore, our graph has no monochromatic edges; so it's bipartite! $\square$