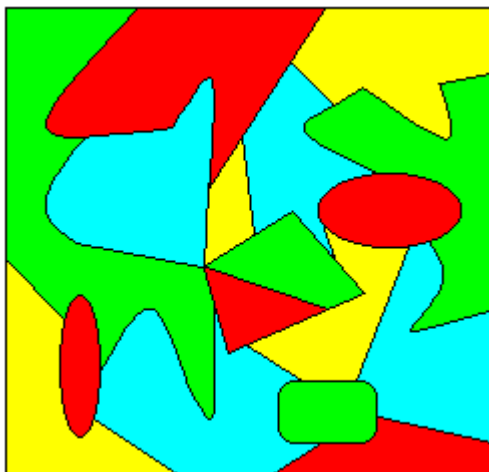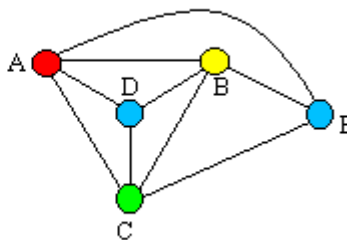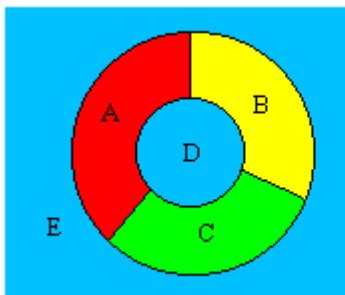# The Four Color Theorem

How many different colors are sufficient to color the countries on a map in such a way that no two adjacent countries have the same color? The figure below shows a typical arrangement of colored regions.

Notice that we define adjacent regions as those that share a common boundary of non-zero length. Regions which meet at a single point are not considered to be "adjacent".

After examining a wide variety of different planar graphs, one discovers the apparent fact that every graph, regardless of size or complexity, can be colored with just four distinct colors. This "four-color conjecture" was first noted by August Ferdinand Mobius in 1840. In 1852 a young man named Francis Guthrie wrote about the problem in a letter to his brother Frederick, then a student at University College in London. Neither of the brothers was able to prove the conjecture, so Frederick asked one of his professors, Augustus DeMorgan (1806-1871). DeMorgan too was unable to prove the conjecture, and after recognizing the difficulty of the problem, he wrote to Sir William Rowan Hamilton (1805-1865) to ask for help. It might seem that this problem would be irresistible to Hamilton, since anything having to do with the number "four" immediately suggests a connection with his beloved quaternions. Also, Hamilton made contributions to graph theory (such as the idea of a Hamiltonian circuit, i.e., a path along the edges of a graph that visits each vertex exactly once), a subject that was developed largely through efforts to prove the four color conjecture. Nevertheless, Hamilton immediately wrote back that he did not believe he would solve DeMorgan's "quaternion of colors" any time soon.

The coloring of geographical maps is essentially a topological problem, in the sense that it depends only on the connectivities between the countries, not on their specific shapes, sizes, or positions. We can just as well represent each country by a single point (vertex), and the adjacency between two bordering countries can be represented by a line (edge) connecting those two points. It's understood that connecting lines cannot cross each other. A drawing of this kind is called a planar graph. A simple map (with just five "countries") and the corresponding graph are shown below.

A graph is said to be n-colorable if it's possible to assign one of n colors to each vertex in such a way that no two connected vertices have the same color. Obviously the above graph is not 3-colorable, but it is 4-colorable. The Four Color Theorem asserts that every planar graph - and therefore every "map" on the plane or sphere - no matter how large or complex, is 4-colorable. Despite the seeming simplicity of this proposition, it was only proven in 1976, and then only with the aid of computers.

Notice that the above graph is "complete" in the sense that no more connections can be added (without crossing lines). The edges of a complete graph partition the graph plane into three-sided regions, i.e., every region (including the infinite exterior) is bounded by three edges of the graph. Every graph can be constructed by first constructing a complete graph and then deleting some connections (edges). Clearly the deletion of connections cannot cause an n-colorable graph to require any additional colors, so in order to prove the Four Color Theorem it would be sufficient to consider only complete graphs.

Although DeMorgan was unable to prove the four-color conjecture, he did observe that no more than four regions on the plane can all be in mutual contact with each other. The graph of a set of three mutually adjoining regions is simply a topological triangle, and if we add a fourth region, it is represented by a fourth vertex in the graph, which must be located either inside or outside the triangle formed by the graph of the original three vertices. In either case, we must then connect this fourth vertex with each of the three original vertices so that all four of the regions are mutually adjoining. Having done this, we can slide the vertices around (if necessary) to bring the graph into the form shown below.
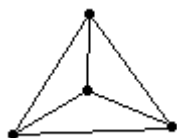
This is the unique graph of four mutually adjoining plane regions (V = 4, E = 6), and also of the tetrahedron. It clearly divides the graph plane into four isolated regions, one region exterior to the big triangle, and the three interior regions. A fifth vertex added to this graph will be able to have edges that reach only three of the four existing vertices, because each of the four regions is completely bounded by three edges that block access to one of the existing vertices. Hence there does not exist a plane graph of five mutually connected vertices, so there does not exist a set of five mutually adjoining regions on the plane.

Of course, the non-existence of a graph with five mutually adjoining vertices does not automatically imply the non-existence of a graph requiring five distinct colors. For example, there exist graphs in which the largest subset of mutually adjoining vertices is two, and yet for which three colors are required. This is the case with any simple loop with an odd number of vertices, as illustrated in the left-hand figure below. Similarly there

are graphs containing no subset of four mutually adjoining vertices that nevertheless require four colors, such as the graph shown on the right.

Therefore, we must consider the possibility that five colors might be required for some graph, even though it's not possible for a graph to contain any set of five mutually connected vertices.

On the other hand, it could be argued that the two example shown above can each be reduced, clarifying the causal links between different parts of the graphs. For example, in the left figure we have an alternating sequence of red, green, red, green, and if these were the only two available colors, it's clear that the chain must continue to alternate, no matter how long - or how short. Thus we could just as well replace any chain of the form rgrg...rg with the simple pair rg, in which case the left-hand figure above consists of three mutually connected vertices. Likewise if we replace the circumferential chain rgrg around the central yellow vertex in the right hand figure with just the pair rg, we get a set of four mutually connected vertices.

In view of this, it's tempting to think that DeMorgan's observation somehow, perhaps indirectly, does ultimately account for the truth of the four-color theorem. Indeed if we construct a graph by adding vertices, one at a time, and make the maximum number of connections at each stage, we will always find the graph plane divided into "triangular" regions, each of which has access to only three vertices. Thus it follows that four colors suffice for any graph that can be constructed in this way. For example, consider the two graphs shown below.
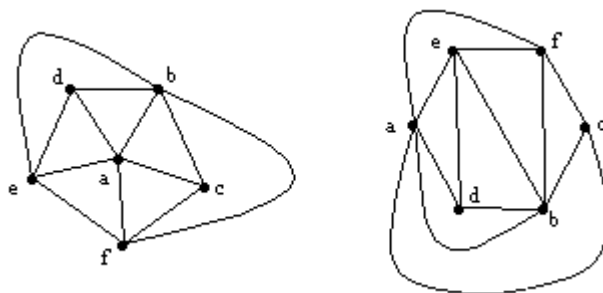

These two graphs each have V = 6 vertices and E = 12 edges, and in fact the two graphs are topologically identical. The plane regions represented by the vertices "a" and "b" each have five adjoining neighbors, whereas the vertices "e" and "f" each have four, and the vertices "c" and "d" each have three. These are complete graphs and, as noted above, a complete graph divides the entire graph plane into "triangular" regions, i.e., regions bounded by three edges connecting three vertices. Nevertheless, depending on how we add the points, it is possible that when a vertex is added to a graph it has more than three neighbors, so we cannot say automatically that four colors would suffice. For example, if vertex "a" is the last to be added, it would have five pre-existing neighbors, and if four colors have already been used in those five vertices, the vertex "a" would require a fifth color.

However, we need not add vertex "a" last. Another topologically equivalent way of drawing the above graph is shown below
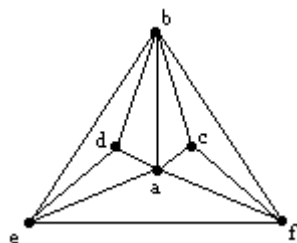
This shows that we could first assign three distinct colors to the vertices e,b,f, and then place the vertex "a" in this triangle, connect it to each of the three surrounding vertices, and give it a fourth color. Then we can place vertex d inside the triangle abe and give it the same color as f. Then we can place vertex c inside the triangle abf and give it the color of e. Hence the graph is 4-colorable. Moreover, any graph, or portion of a graph bounded by a triangle such as ebf , and having this hierarchical pattern of nested triangles, is 4-colorable. This is the case when the graph contains some vertices with only three edges, and when those vertices and edges are removed, the remaining graph has some vertices with only three edges, which can be removed, and so on, until finally all that remains is a single triangle.

Unfortunately (for the prospect of a simple proof), not every graph is of this hierarchical form. For example, consider the complete graph shown below.
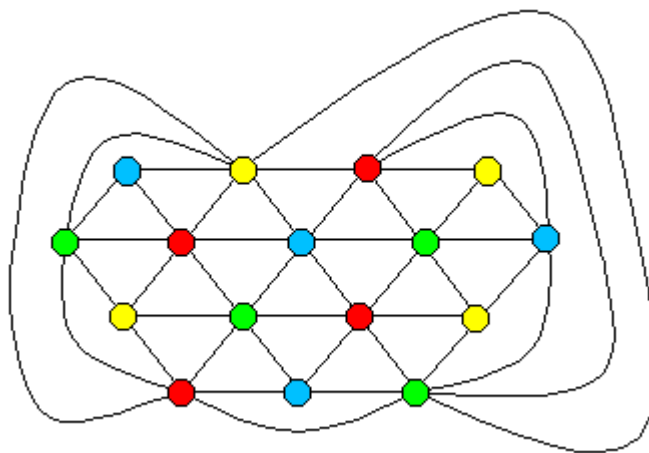
This graph has V = 16 vertices and E = 42 edges. It is not hierarchical, because after removing the two vertices having only three edges each, the remaining graph has four or more edges attached to each vertex. (We can also infer that this graph is not hierarchical from the fact that no three mutually connected vertices are directly connected to a fourth vertex.) Nevertheless, this graph is 4-colorable, as shown in the figure.

The above graph contains a "flat" patch consisting of a uniform hexagonal grid. An *infinite* hexagonal grid is obviously 4-colorable by means of the alternating layered pattern shown below.
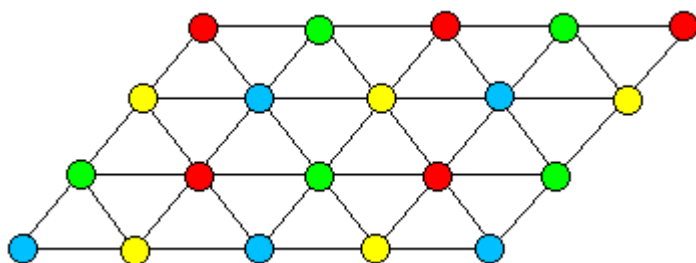
In a sense, two extreme types of graph structures are the purely hierarchical graphs and the perfectly "flat" graphs, both of which are easily seen to be 4-colorable.

Notice that we need not consider graphs containing any configurations of four mutually connected vertices, because, as discussed previously, the edges of such a configuration necessarily split the graph plane into four separate regions, each of which has access to only three of the four vertices. Hence if we can 4-color the interior of such a triangle, we can 4-color the tetrahedral "frame" as well. If we could reduce all the causal links in a graph to their bare minimums, it might be possible to reduce every n-colorable graph to a graph with just n mutually connected vertices, and hence n could never exceed four. However, the reduction of graphs with more than two colors is not a trivial task.

If we denote the number of vertices, edges, and faces (i.e., the bounded regions) of a planar graph by V, E, and F respectively, then Euler's formula for a plane (or a sphere) is $V - E + F = 2$. Furthermore, each face of a complete graph is bounded by three edges, and each edge is on the boundary of two faces, so we have $F = 2E/3$, and Euler's formula for a complete planar graph is simply $E = 3V - 6$. Now, each edge is connected to two vertices, so the total number of attachments (in a complete graph) is $2E = 6V - 12$, and hence the average number of attachments per vertex is $6 - 12/V$. For any incomplete graph, the total number of attachments is less. Consequently the average number of attachments per vertex for any graph (with a finite number of vertices) is less than 6, which implies that at least one vertex has only five or fewer attachments.

If we have six available colors, a vertex with only five neighbors obviously imposes no constraint on the coloring of the other vertices, because, regardless of the colors of its five (or fewer) neighbors, we can assign it a color without exceeding the six available colors. Therefore if we delete this vertex and all its connections from the graph, creating a graph with one fewer vertices, it's clear that if the resulting graph is 6-colorable, then so was the original graph. Moreover, Euler's formula assures us that this reduced graph also contains at least one vertex with five or fewer neighbors, so we can apply this procedure repeatedly, reducing the graph eventually to one with just 6 vertices, which is obviously 6-colorable. Hence the original graph is 6-colorable.

So, we've seen that Euler's formula immediately implies that no graph can require more than six colors. Furthermore, with just a little more work, we can also show that no graph can require more than *five* colors. (Ultimately we will see that no graph can require more than four colors, but it's worthwhile to begin with the proof of the 5-colorability of every planar graph.) Obviously every graph with five or fewer vertices is 5-colorable, so if there exists a finite graph that requires more than five colors, it must have more than five vertices. Let the positive integer V_6 denote the smallest number of vertices on which there exists a graph that requires six (or more) colors. Conversely, every graph with fewer than V_6 vertices is 5-colorable.

Now, assuming the existence of a graph that requires more than five colors, we can consider one that has exactly V_6 vertices. By Euler's formula, this graph must contain at least one vertex with five or fewer connections. However, it cannot contain any vertex with just four (or fewer) connections, because if it did, we could delete such a vertex and leave a graph with just $V_6 - 1$ vertices, which is 5-colorable by definition. Re-inserting the deleted vertex would clearly have no effect on the 5-colorability of the graph, because the vertex has only four (or fewer) neighbors, so the original graph must be 5-colorable, contradicting our assumption. Therefore, a graph with V_6 vertices that requires more than five colors cannot contain any vertex with just four or fewer neighbors.

Since Euler's formula implies that the graph contains at least one vertex with five or fewer

connections, the only remaining possibility is that the graph contains a vertex with exactly five connections. However, this too leads to a contradiction. To show this, it's helpful to introduce the notion of a *k-cluster*, which is specified by a set of k distinct colors and one particular vertex that has one of those colors. The original vertex is included in the cluster, and, in addition, every vertex with one of the k specified colors that neighbors a vertex in the cluster is also in the cluster. By definition the only vertices outside a cluster that are directly connected to vertices inside the cluster have colors that are not in the specified set of k colors. Therefore, we can apply any permutation of the k colors to the vertices in a cluster without invalidating the coloration. In particular, a 2-cluster is a contiguous set of vertices, each with one of two specified colors, and we can transpose these two colors without upsetting the coloration of a graph.

Now consider a graph containing a vertex with exactly five immediate neighbors, of five distinct colors, as illustrated below.
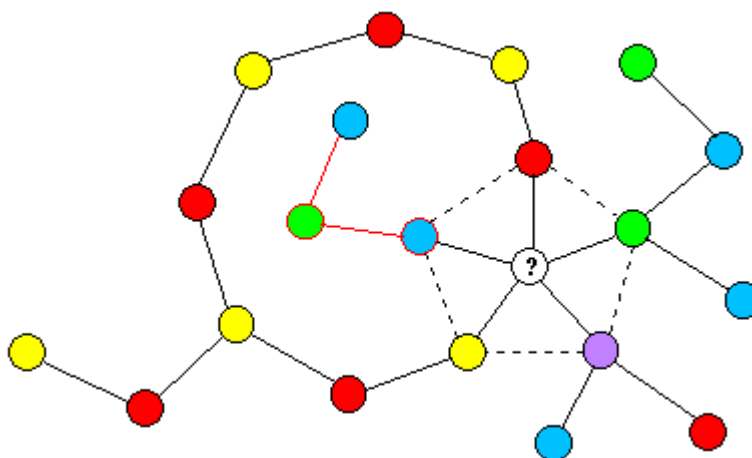
Since the uncolored vertex in the center has neighbors of five distinct colors, it might seem that a sixth color is required. However, notice that we can transpose the blue and green colors in the blue/green 2-cluster attached to the upper left vertex of the central pentagon, so that the upper left vertex is green instead of blue. Once we have done this, the uncolored vertex in the center has neighbors of only four distinct colors. If we delete the central vertex, the overall graph has V_6 - 1 vertices, so it is 5-colorable, but re-inserting this vertex requires no sixth color (once we have transposed the blue and green in the small 2-cluster as described), so the original graph is 5-colorable.

The only possible objection to the transposing of the colors in a cluster as a means of reducing the number of distinct colors around the perimeter of the pentagon is that we cannot necessarily change the color of any vertex to the color of one of the opposite vertices of the pentagon, because two opposite vertices of a pentagon might be in the same 2-cluster. This is the case, for instance, with the yellow and red vertices of the pentagon in the figure above. If we transpose the red and yellow colors in this 2-cluster, there would still be five distinct colors around the perimeter of the pentagon. However, if such a cluster exists, connecting two opposite vertices of the pentagon, we are guaranteed that at least one other pair of vertices are cut off from each other, i.e., they cannot be in the same 2-cluster. For example, the blue/green 2-cluster outlined in red in the above figure cannot be part of the blue/green 2-cluster attached to the upper right of the pentagon. In general, for any vertex connected to exactly five other vertices, we can always (without invalidating the coloration) change the color of at least one of the five neighbors so that there are only four distinct colors. It follows, using the reduction argument described in the preceding paragraph, that no graph can require more than five colors.

This still leaves open the question of whether five colors are ever actually required, or whether four colors will always suffice. If there exist planar graphs requiring five distinct colors, there must be a positive integer V_5 that is the smallest number of vertices on which such a graph can be constructed. Let's call a graph with V_5 vertices that requires five colors a minimal 5-color graph. By the reduction argument described previously, such a graph obviously cannot contain any vertex with three or fewer connections. Moreover, by considering 2-clusters again, we can prove that such a graph cannot contain any vertex with just four connections. To see why, consider the portion of a graph illustrated below.
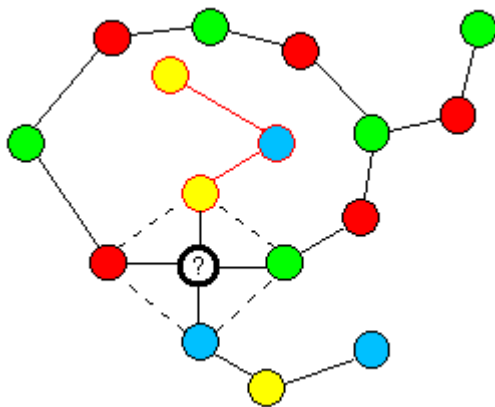
The uncolored vertex in the center has just four neighbors, which we've colored red, yellow, green, and blue. This might seem to require a fifth color for the central vertex, but in fact by permuting the yellow and blue colors in the 2-cluster (outlined in red) above the central vertex we can transform this coloration so that the central uncolored vertex has neighbors of only three distinct colors (red, green, and blue). Since the overall graph from which this configuration was taken is assumed to have exactly V_5 vertices, it follows that if we delete the central vertex, the resulting graph has only V_5 - 1 vertices, so it is 4-colorable, but we can clearly add the central vertex back without requiring a fifth color, so the whole graph must be 4-colorable, contradicting our original assumption. This shows that it's always possible to modify a graph containing a vertex with only four connections in such a way that the vertex is connected to only three distinct colors. At most, only one of the two pairs of opposite vertices can be linked by a common 2-cluster, because if one pair is so linked, the other cannot be linked.

To summarize the argument up to this point, we've shown that no graph can require more than five colors, and we've also shown that if there exists a graph requiring five colors, there must exist a minimal 5-color graph, and such a graph cannot contain any vertex with fewer than five connections. Also, by adding connections, we can "complete" any graph on V vertices so that all the faces are triangular and there are a total of exactly 6V-12 attachments. This graph must still be a minimal 5-color graph because we haven't increased the number of vertices, and adding connections cannot reduce the number of colors required, whereas no graph requires more than five colors. Therefore, letting a_5, a_6, a_7,... denote the number of vertices with precisely 5, 6, 7,... attachments respectively, we must have a minimal 5-color graph such that

$$6V - 12 \;=\; 5a_5 + 6a_6 + 7a_7 + 8a_8 + 9a_9 \ldots$$

where $V = a_5 + a_6 + a_7 + a_8 + a_9 + \ldots$ Substituting this expression for V into the above formula and re-arranging, we have the condition

$$12 \;=\; a_5 - a_7 - 2a_8 - 3a_9 - \ldots$$

This places fairly severe constraints on any putative complete minimal 5-color graph. For example, if we consider only such graphs with no more than six attachments at any single vertex, then this formula implies $a_5 = 12$. In other words, a complete minimal 5-graph with no more than six attachments per vertex must have exactly 12 vertices with five attachments each. This suggests that these 12 vertices are arranged globally in the pattern of an icosahedron, and the remaining vertices with six attachments per vertex are in a regular hexagonal pattern, filling in the "faces" of the icosahedron. (This pattern is the basis for "geodesic domes"). The fundamental graph of this type, with $a_6 = 0$, is shown below.
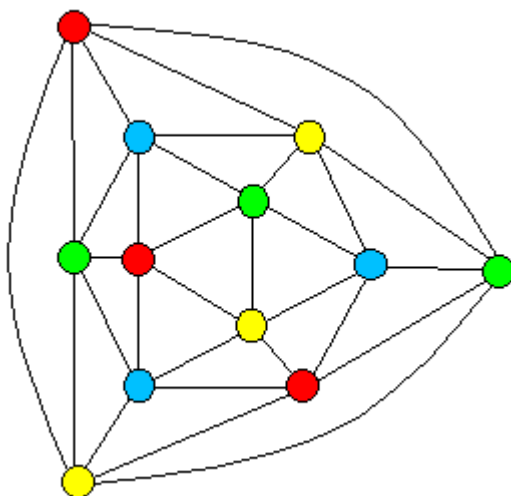
We can give explicit 4-colorings of every such pattern, so such graphs can be shown explicitly to require no more than four colors. Therefore, if a complete minimal 5-color graph exists, it must contain at least one vertex with seven or more attachments.

It was not until 1976 that, with the help of modern computers, the 4-color conjecture was finally proven to be true. The proof, developed by Appel and Haken based on ideas of several earlier people (Kempe, Heawood, Birkhoff, etc) consisted of finding a set of subgraphs, at least one of which must be contained in any normal planar graph (i.e., the set is unavoidable), and such that if a graph containing one of those sub-graphs is not 4-colorable, then neither is a graph with one fewer vertices. This leads immediately to a contradiction, because it implies that, if we hypothesize the existence of a graph that is not 4-colorable, we can always construct another graph, with fewer vertices, that is also not 4-colorable, eventually arriving at a graph with only four vertices, but this graph obviously *is* 4-colorable. Hence we have a contradiction, so we can conclude that the original hypothesis was false, i.e., there does *not* exist a graph that is not 4-colorable. The unavoidable set found by Appel and Haken consisted of nearly 1500 subgraphs, and many of these required considerable analysis to prove that they were "reducible". The development of this unavoidable set, and the proofs of reducibility for its members, was carried out largely on computers, so the proof is notable (and slightly controversial) as an early example of a mathematical proposition whose proof can only be carried out on a computer, and is, in some sense, beyond human mental capabilities to verify.

It's interesting to consider the problem of determining a 4-coloring for any given graph from a purely algebraic standpoint (harking back to Hamilton's quaternions). We might stipulate that each vertex is to be assigned one of four possible values, and the conditions imposed by the connections would be represented by algebraic equations, one equation for each connection. Part of the difficulty of this approach is that the condition of adjacency is not an equality, it is an *in*equality, i.e., we require that two vertices have unequal values (from among the four possible values). Also, there is necessarily a great deal of ambiguity

in the solution, because any permutation of the colors leaves a solution intact. In addition, there is even more ambiguity, since in general there can be more than one distinct 4-coloration of a graph.

Given the four allowable (distinct) values a,b,c,d, we could algebraically impose the requirement for every vertex to have one of these values by requiring that the value u assigned to any vertex satisfies $f(u) = 0$, where the polynomial f is defined as

$$f(x) = (x - a)(x - b)(x - c)(x - d)$$

Then we could algebraically impose the required inequality on the values u,v assigned to two connected vertices by stipulating that $g(uv) = 0$, where the polynomial g is defined as

$$g(x) = (x - ab)(x - ac)(x - ad)(x - bc)(x - bd)(x - cd)$$

The equation $g(uv) = 0$ is satisfied if and only if u and v have distinct values from the set {a,b,c,d}. For example, the set {-2,-1,1,2} gives

$$f(x) = \left(x^2 - 1\right)\left(x^2 - 4\right) \qquad g(x) = (x + 1)(x + 4)\left(x^2 - 4\right)$$

To illustrate, consider again the octahedral graph (which is 3-colorable):
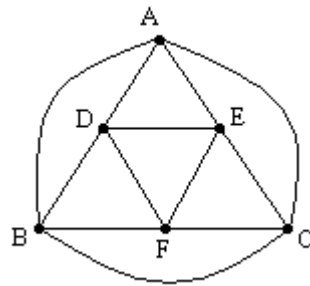
The algebraic conditions for the individual "color" values assigned to the six variables A,B,C,D,E,F are

$$f(A) = f(B) = f(C) = f(D) = f(E) = f(F) = 0$$

and the algebraic conditions imposed by the twelve connections are

$$g(AB) = g(AC) = g(AD) = g(AE)$$
$$= g(BC) = g(BD) = g(BF) = g(CE)$$
$$= g(CF) = g(DE) = g(DF) = g(EF) = 0$$

This amounts to 18 equations in the 6 variables. We can, however, without loss of generality stipulate the "colors" of three of the variables, say, A = 1, B = -1, C = 2, since these three must have mutually distinct values. This leaves us with the following twelve equations in the three unknowns D,E,F

$$f(D) = f(E) = f(F) = 0$$

and

$$g(D) = g(E) = g(-D) = g(-F) = g(2E)$$
$$= g(2F) = g(DE) = g(DF) = g(EF) = 0$$
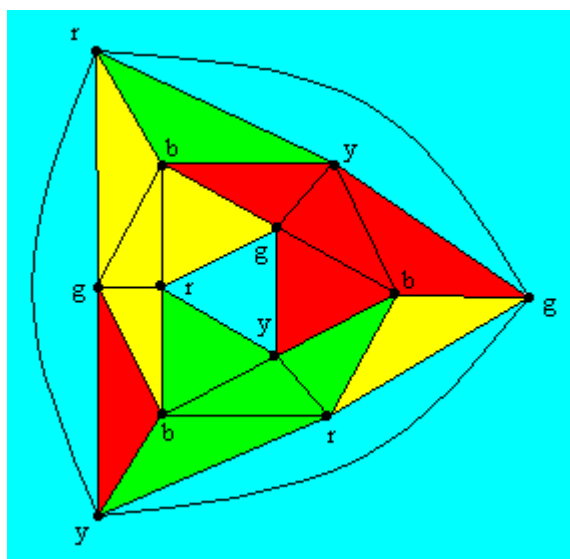
The three equations $f(D) = 0$, $g(D) = 0$, and $g(-D) = 0$ jointly have only two solutions, namely, $D = 2$ and $D = -2$. Therefore those three equations can be replaced with the single equation $D^2 - 4 = 0$. Likewise the three equations involving only E can be reduced to $E^2 + 3E + 2 = 0$, and the three equations involving only F can be reduced to $F^2 + F - 2 = 0$. Thus we have six equations in the three unknowns.

A different approach would be to assign a four-dimensional vector to each vertex, and then the connection between two given vertices would be represented by requiring the dot product of those two vectors to vanish, i.e., requiring that the two vectors be mutually orthogonal. The three vectors assigned to the vertices of any triangle would then be an orthogonal triad with some orientation in 4-dimensional space. Likewise the four vectors assigned to the vertices of a tetrahedral graph would be a complete set of four mutually orthogonal vectors (a tetrad), but still with arbitrary orientation. For example, the conditions on the six vectors A,B,C,D,E,F assigned to the vertices of the octahedral graph shown above would be simply

$$A \cdot B = A \cdot C = A \cdot D = A \cdot E$$
$$= B \cdot C = B \cdot D = B \cdot F = C \cdot E$$
$$= C \cdot F = D \cdot E = D \cdot F = E \cdot F = 0$$

In this context the four color theorem tells us that a space of four dimensions is sufficient to enable us to assign one of the four basis vectors to each vertex of a planar graph in such a way that the vectors of every pair of adjacent vertices are orthogonal. We can assume each vector is of unit length, so it has three independent components. Therefore, we have 18 independent components in all, constrained by 12 equations. Naturally the system is underspecified, because we can apply an arbitrary permutation to the vectors.

Since the edges of a complete graph partition the graph plane into three-sided regions whose vertices are three mutually connected points, we can uniquely assign the fourth color to each of these regions. The result is useful for visualizing the symmetries of the coloring. For example, the 4-colored icosahedral graph discussed above looks like this:

[Return to MathPages Main Menu](#)