The Four-Color Problem: Concept and Solution

Steven G. Krantz

October 14, 2007

Prologue

Modern mathematics is a rich and complex tapestry of ideas that have evolved over thousands of years. Unlike computer science or biology, where the concept of truth is in a constant state of flux, mathematical truth is permanent. Ideas that were discovered 2000 years ago by Pythagoras are still valid today. Proofs (also millenia old) in the style of Euclid are as valid today as when they were first created.

As a result, modern mathematics can be quite complex and technical. It requires someone with considerable advanced training to understand the current problems, much less solve them. So it is particularly charming when we can find problems that *anyone* can understand, but that still resist the best efforts of the world's great experts.

The problem that we wish to discuss today is charming and simple. It is appealing because it is geometric, and it has an interesting and unusual genesis.

In 1852 Francis W. Guthrie, a graduate of University College London, posed the following question to his brother Frederick:

Imagine a geographic map on the earth (i.e., a sphere) consisting of countries only—no oceans, lakes, rivers, or other bodies of water. The only rule is that a country must be a single contiguous mass—in one piece, and with no holes—see Figure 1.

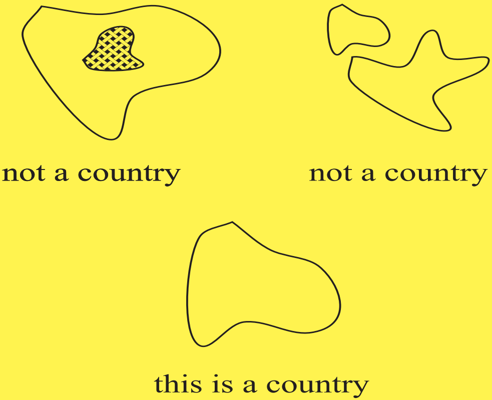

Figure 1. What is a country?

As cartographers, we wish to *color* the map so that no two adjacent countries (countries that share an edge) will be of the same color (Figure 2). How many colors should the map-maker keep in stock so that he can be sure he can color any map that may arise?

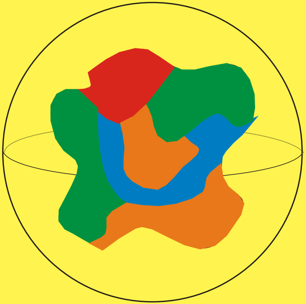

Figure 2. A typical map and its coloring.

It is not difficult to write down an example of a map that surely needs 4 colors. Examine Figure 3. Each of the countries in this figure is adjacent to each of the others. There are four countries, and they all must be of a different color.

Is there a map that will require 5 colors? Mathematicians at the best universities beat their collective heads against this question for many decades.

Figure 3. A map that requires four colors.

The eminent geometer Felix Klein (1849–1925) in Göttingen heard of the problem and declared that the only reason the problem had never been solved is that no capable mathematician had ever worked on it. *He*, Felix Klein, would offer a class, the culmination of which would be a solution of the problem. *He failed.*

In 1879, A. Kempe (1845–1922) published a solution of the four-color problem. That is to say, he showed that any map on the sphere whatever could be colored with four colors. Kempe's proof stood for eleven years. Then a mistake was discovered by P. Heawood (1861–1955). Heawood studied the problem further and came to a number of fascinating conclusions:

In 1879, A. Kempe (1845–1922) published a solution of the four-color problem. That is to say, he showed that any map on the sphere whatever could be colored with four colors. Kempe's proof stood for eleven years. Then a mistake was discovered by P. Heawood (1861–1955). Heawood studied the problem further and came to a number of fascinating conclusions:

- Kempe's proof, particularly his device of "Kempe chains" (a sequence of countries that alternates between just two colors), *does* suffice to show that any map whatever can be colored with at most 5 (*not* 4) colors. We say that the *chromatic number* of the sphere is *at most* 5, but it *could be* 4.

In 1879, A. Kempe (1845–1922) published a solution of the four-color problem. That is to say, he showed that any map on the sphere whatever could be colored with four colors. Kempe's proof stood for eleven years. Then a mistake was discovered by P. Heawood (1861–1955). Heawood studied the problem further and came to a number of fascinating conclusions:

- Kempe's proof, particularly his device of "Kempe chains" (a sequence of countries that alternates between just two colors), *does* suffice to show that any map whatever can be colored with at most 5 (*not* 4) colors. We say that the *chromatic number* of the sphere is *at most* 5, but it *could be* 4.
- Heawood found a formula that gives an estimate for the chromatic number of any surface that is geometrically more complicated than the sphere.

Here is how to understand Heawood's idea. It is known that any surface in space is geometrically equivalent to a sphere with handles attached. See Figure 4. The number of handles is called the *genus*, and we denote it by g. The Greek letter chi ($\chi(g)$) is the chromatic number of the surface—the least number of colors that it will take to color any map on the surface with genus g.

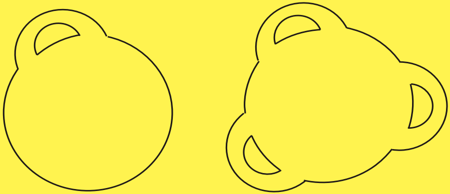

Figure 4. Any surface in space is a sphere with handles attached.

Heawood' formula is

$$\chi(g) \leq \frac{1}{2}\left(7+\sqrt{48g+1}\right)$$

so long as $g \geq 1$.

The torus (see Figure 5) is topologically equivalent to a sphere with one handle. Thus the torus has genus $g = 1$. Then Heawood's formula gives the estimate 7 for the chromatic number:

$$\chi(1) \leq \frac{1}{2}\left(7+\sqrt{48\cdot 1+1}\right) = \frac{1}{2}(7+7) = 7\,.$$

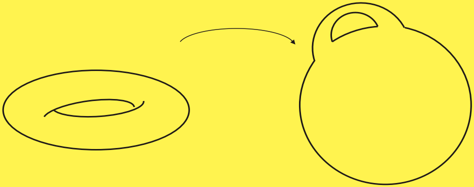

Figure 5. The torus.

We can show that this is the *right* (or best possible) estimate by first performing the trick of cutting the torus apart. See Figure 6. By cutting the torus around the small rotation and then across the large rotation, we render it as a rectangle with identifications.

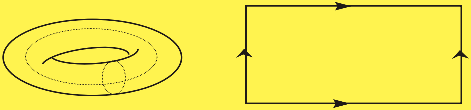

Figure 6. The torus cut apart.

With one cut, the torus becomes a cylinder; with the second cut it becomes a rectangle. The arrows on the edges indicate that the left and right edges are to be identified (with the same orientation), and the upper and lower edges are to be identified (with the same orientation). Now we can see in Figure 7 how to color the torus (rendered as a rectangle). For clarity, we call our colors "1", "2", "3", "4", "5", "6", "7".

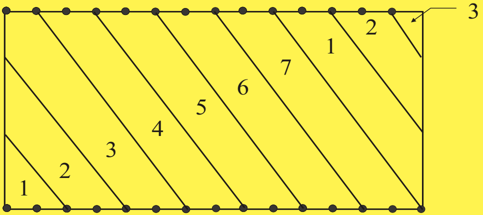

Figure 7. Seven colors on the torus.

We may see that there are seven countries shown in our Figure 7, and every country is adjacent to (i.e., touches) every other. Thus they all must have different colors! This is a map on the torus that *requires* 7 colors; it shows that Heawood's estimate is sharp for this surface.

For the double-torus with two handles (genus 2—see Figure 8), Heawood's estimate gives an estimate of 8. Is that the best number? Is there a map on the double torus that actually *requires* 8 colors? And so forth: we can ask the same question for every surface of every genus. Heawood could not answer these questions. Nor could anyone else.

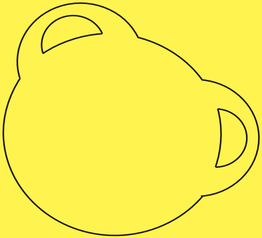

Figure 8. The sphere with two handles.

The late nineteenth century saw more alleged solutions of the four-color problems, many of which stood for as long as eleven years. Eventually errors were found, and the problem remained open on into the twentieth century.

What is particularly striking is that Gerhard Ringel (1919–) and J. W. T. Youngs (1910–1970) were able to prove in 1968 that all of Heawood's estimates, for the chromatic number of any surface of genus at least 1, are sharp. So the chromatic number of a torus is indeed 7. The chromatic number of a "double-torus" with two holes is 8. And so forth. But the Ringel/Youngs proof, just like the Heawood formula, does not apply to the sphere. They could not improve on Heawood's result that 5 colors will always suffice. The 4-color problem remained unsolved.

Then in 1974 there was blockbuster news. Using 1200 hours of computer time on the University of Illinois supercomputer, Kenneth Appel and Wolfgang Haken showed that in fact 4 colors will always work to color any map on the sphere. Their technique is to identify 633 fundamental configurations of maps (to which all others can be reduced) and to prove that each of them is reducible to a simpler configuration. But the number of "fundamental configurations" was very large, and the number of reductions required was beyond the ability of any human to count. And the reasoning is extremely intricate and complicated. Enter the computer.

In those days computing time was expensive and not readily available, and Appel and Haken certainly could not get a 1200-hour contiguous time slice for their work. So the calculations were done late at night, "off the record", during various down times. In fact, Appel and Haken did not know for certain whether the calculation would ever cease. Their point of view was this:

In those days computing time was expensive and not readily available, and Appel and Haken certainly could not get a 1200-hour contiguous time slice for their work. So the calculations were done late at night, “off the record”, during various down times. In fact, Appel and Haken did not know for certain whether the calculation would ever cease. Their point of view was this:

- If the computer finally stopped then it will have checked all the cases and the 4-color problem was solved.

In those days computing time was expensive and not readily available, and Appel and Haken certainly could not get a 1200-hour contiguous time slice for their work. So the calculations were done late at night, “off the record”, during various down times. In fact, Appel and Haken did not know for certain whether the calculation would ever cease. Their point of view was this:

- If the computer finally stopped then it will have checked all the cases and the 4-color problem was solved.
- If the computer never stopped then they could draw no conclusion.

Well, the computer stopped. But the level of discussion and gossip and disagreement in the mathematical community did not. Was this really a proof? The computer had performed tens of millions of calculations. Nobody could ever check them all.

But now the plot thickens. Because in 1975 a mistake was found in the proof. Specifically, there was something amiss with the algorithm that Appel and Haken fed into the computer. It was later repaired. The paper was published in 1976. The four-color problem was declared to be solved.

In a 1986 article, Appel and Haken point out that the reader of their seminal 1976 article must face

In a 1986 article, Appel and Haken point out that the reader of their seminal 1976 article must face

- 50 pages containing text and diagrams;

In a 1986 article, Appel and Haken point out that the reader of their seminal 1976 article must face

- 50 pages containing text and diagrams;
- 85 pages filled with almost 2500 additional diagrams;

In a 1986 article, Appel and Haken point out that the reader of their seminal 1976 article must face

- 50 pages containing text and diagrams;
- 85 pages filled with almost 2500 additional diagrams;
- 400 microfiche pages that contain further diagrams and thousands of individual verifications of claims made in the 24 statements in the main section of the text.

But it seems as though there is always trouble in paradise. Errors continued to be discovered in the Appel/Haken proof. Invariably the errors were fixed. But the stream of errors never seemed to cease. So is the Appel/Haken work really a proof?

Well, there is hardly anything more reassuring than another, independent proof. Paul Seymour and his group at Princeton University found another way to attack the problem. In fact they found a new algorithm that seems to be more stable. They also needed to rely on computer assistance. But by the time they did their work computers were *much*, much faster. So they required much less computer time. In any event, this paper appeared in 1994.

It is still the case that mathematicians are most familiar with, and most comfortable with, a traditional, self-contained proof that consists of a sequence of logical steps recorded on a piece of paper. We still hope that some day there will be such a proof of the four-color theorem. After all, it is only a traditional, Euclidean-style proof that offers the understanding, the insight, and the sense of completion that all scholars seek.

And there are new societal needs: theoretical computer science and engineering and even modern applied mathematics require certain pieces of information and certain techniques. The need for a workable device often far exceeds the need to be *certain* that the technique can stand up to the rigorous rules of logic. The result may be that we shall re-evaluate the foundations of our subject. The way that mathematics is practiced in the year 2100 may be quite different from the way that it is practiced today.