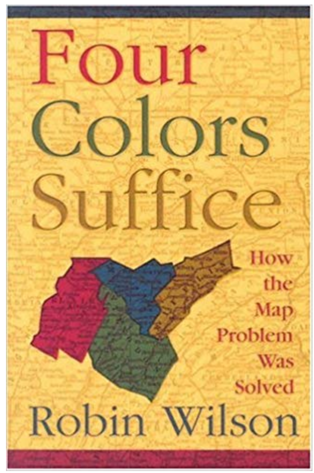
Supplement. The Four-Color Theorem: A History, Part 2

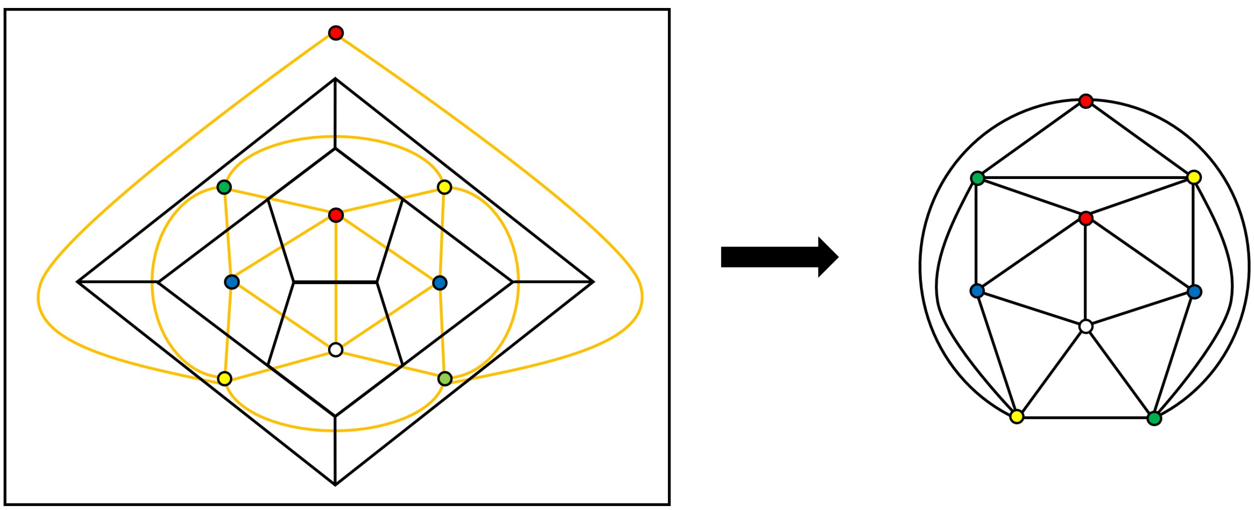
Note. We continue our story of the history of the Four Color Theorem. Our primary source remains Robin Wilson's *Four Colors Suffice*, Princeton University Press (2002).



Note. On the last page of Alfred B. Kempe's paper, "On the Geographical Problem of the Four Colours," *American Journal of Mathematics*, **2**(3), 193–200 (1879), available online at the JSTOR website (accessed 12/23/2022), he states:

> "If we lay a sheet of tracing paper over a map and mark a point on it over each district and connect the points corresponding to districts which have a common boundary, we have on the tracing paper a diagram of a 'linkage', and we have the exact analogue of the question we have been considering, that of lettering the points in the linkage with as few letters as possible, so that no two directly connected points shall be lettered with the same letter."

Of course the marked "points" are the vertices of a graph, and the connections of the points which correspond to districts with a common boundary are the edges of a graph. It is at this early stage that the Four Color Theorem enters the realm of graph theory. If we treat the map itself as a graph, then Kempe's points and connections represent the dual of the graph. This is formalized in Bondy and Murty's graduate text in Section 10.2. Duality. In Supplement. The Four-Color Theorem: A History, Part 1, we gave an example of a small map (below left) and a partial coloring that illustrated the failure of Kempe's "proof" of the Four Color Theorem.



On the left, we have the small map (with black borders) with the dual map superimposed on it (with orange edges and vertices colored according to the colors of the corresponding countries in the partial coloring). On the right, we have the dual by itself (as given in Exercise 15.2.2 of Bondy and Murty's graduate text).

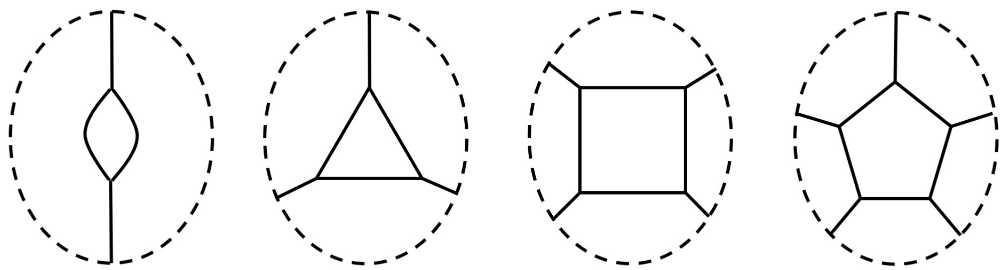
Note FCT.K. Now four coloring a map is equivalent to four coloring the vertices of the dual graph such that any two adjacent vertices have a different color (that is, finding a *proper* vertex four coloring of the dual graph). Approaches to the Four Color Theorem in the 20th century largely are based on consideration of the dual graph. We saw in Section 10.2. Duality that the dual of a plane graph is itself a plane graph (Lemma 10.2.A) and a dual of a plane graph is connected (Proposition 10.9). We also have in Proposition 10.11 of that section that a simple connected plane graph is a triangulation (that is, all countries or "faces" are bounded by three edges) if and only if its dual is cubic (that is, three regular). In Section 15.2. The Four-Colour Theorem we see that a map which is a minimal counterexample to the Four Color Theorem, the dual graph is a triangulation (Proposition 15.2(ii)) and has no vertex of degree less than four (Proposition 15.2(iii)). We then have by Proposition 10.11 that a minimal map counterexample to the Four Color Theorem has exactly three borders meeting at a single point (i.e., the map as a graph is cubic). Notice that these properties are satisfied by the map and its dual from Exercise 15.2.2, given above. The fact that the dual of a minimal counterexample has no vertex of degree less than four corresponds to the fact that a minimal counterexample cannot contain a two-sided country (by Note FCT.C), a three-sided country (by Note FCT.D), nor a four-sided country (by Note FCT.H; this is the result of Kempe's *correct* argument using Kempe chains and Kempe interchange).

Note. By 1900, the flaws in Kempe's attempted proof of the Four Color Theorem was known. The idea that a correct proof had not been presented because sufficiently talented mathematicians had not worked on the problem started to circulate. One story is that Hermann Minkowski, while lecturing on topology at Göttingen University, mentioned the four color problem and stated "This theorem has not yet been proved, but that is because only mathematicians of the third rank have occupied themselves with it." (See Wilson's page 143.) You might know

Minkowski's name from Real Analysis 2 (MATH 5220), where the Triangle Inequality in the classical L^p spaces is called Minkowski's Inequality (see my online notes for this class on Section 7.2. The Inequalities of Young, Höolder, and Minkowski), or from special relativity theory where flat spacetime is known as Minkowski space (see my online notes for Differential Geometry [MATH 5510] on the supplemental topic Section 1.1. The Minkowski Vector Space $\mathbb{V}_4$). Minkowski was one of Albert Einstein's mathematics instructors at the Federal Institute of Technology in Zurich, Switzerland and once referred to Einstein as a "lazy dog." (Upon learning of Einstein's 1905 work on special relativity, Minkowski expressed surprise and admiration!)

Note. Before the turn of the century, the four color problem had mostly been of interest in Britain: De Morgan of the University College London, his student Frederick Guthrie (and his brother Francis), Arthur Cayley of Trinity College Cambridge, Alfred Kempe of London, and Percy Heawood of Durham College England. After the turn of the century, several American mathematicians took an interest in the problem. Wilson mentions George Birkhoff, Oswald Veblen, Philip Franklin, Hassler Whitney by name (see his page 144). Two ideas emerged in work on the problem, the concepts of an *unavoidable set* and a *reducible configuration*. Both ideas were implicit in Kempe's 1879 paper.

Note. To illustrate the concept of an *unavoidable set*, recall that the Only Five Neighbors Theorem states that every map has at least one country with five or fewer neighbors. If we consider the collection of cubic maps (those where each point of intersection of boundaries involve the intersection of exactly three boundaries), we then have that that each cubic map must include one of the following:



That is, these four arrangements of countries form an unavoidable set in the collection of cubic maps.

Note FCT.L. A *reducible configuration* is any arrangement of countries that cannot occur in a minimal counterexample to the Four Color Theorem. If a map contains a reducible configuration, then any coloring of the rest of the map with four colors can be extended, after any necessary recoloring, to a four coloring of the entire map. We have seen that two, three, and four sided countries are examples of reducible configurations in Kempe's work (see Notes FCT.C, FCT.D, and FCT.H). If a five sided country could be shown to be a reducible configuration, then we would have a proof of the Four Color Theorem (in modern terminology, this is what Kempe unsuccessfully attempted to do).

Note FCT.M. The strategy to prove the Four Color Theorem now becomes to show that no counterexample exists. This can be done by **finding an unavoidable set of configurations, each of which is reducible**. We would then have a counterexample that must contain a reducible configuration (which, of course, it cannot contain). This contradiction proves that no counterexample can exist, and hence the Four Color Theorem holds. Ultimately, this is the technique that leads to a successful proof.

Note. Paul Wernicke in "Über den kartographischen Vierfarbensatz," *Math. Ann.*, **58**(3), 413–426 (1904) showed that a cubic map that contains no 2-sided, 3-sided, or 4-sided country must contain either two adjacent pentagons or a pentagon adjacent to a hexagon. We illustrate the technique of *discharging* by showing that the set of configurations consisting of the digon, triangle, square, pair of adjacent pentagons, and pentagon adjacent to a hexagon is an unavoidable set. We will first assume that we have a cubic map that contains none of these, and then find a contradiction. The contradiction then implies that a pentagon can only be adjacent to countries with at least seven edges. Discharging was first published by Heinrich Heesch (pronounced 'haish', June 25, 1906–July 26, 1995) in "Untersuchungen zum Vierfarbenproblem," *Hochschulskriptum* **810ab**, Bibliographisches Institut, Mannheim, 1969.

Note FCT.N. ASSUME that a cubic map contains no 2-sided, 3-sided, or 4-sided country, no two adjacent pentagons, and no pentagon adjacent to a hexagon. We start by assigning a *charge* of $6 - k$ to each country with k boundary lines.

Notice that when $k < 5$ (such as for a digon, triangle, and square), a country gets a positive charge, a hexagon gets a 0 charge, and for $k > 6$ a country gets a positive charge. Let C_k denote the number of countries with k sides. Then we have $C_2 = C_3 = C_4 = 0$ by assumption. The total charge on the map is then

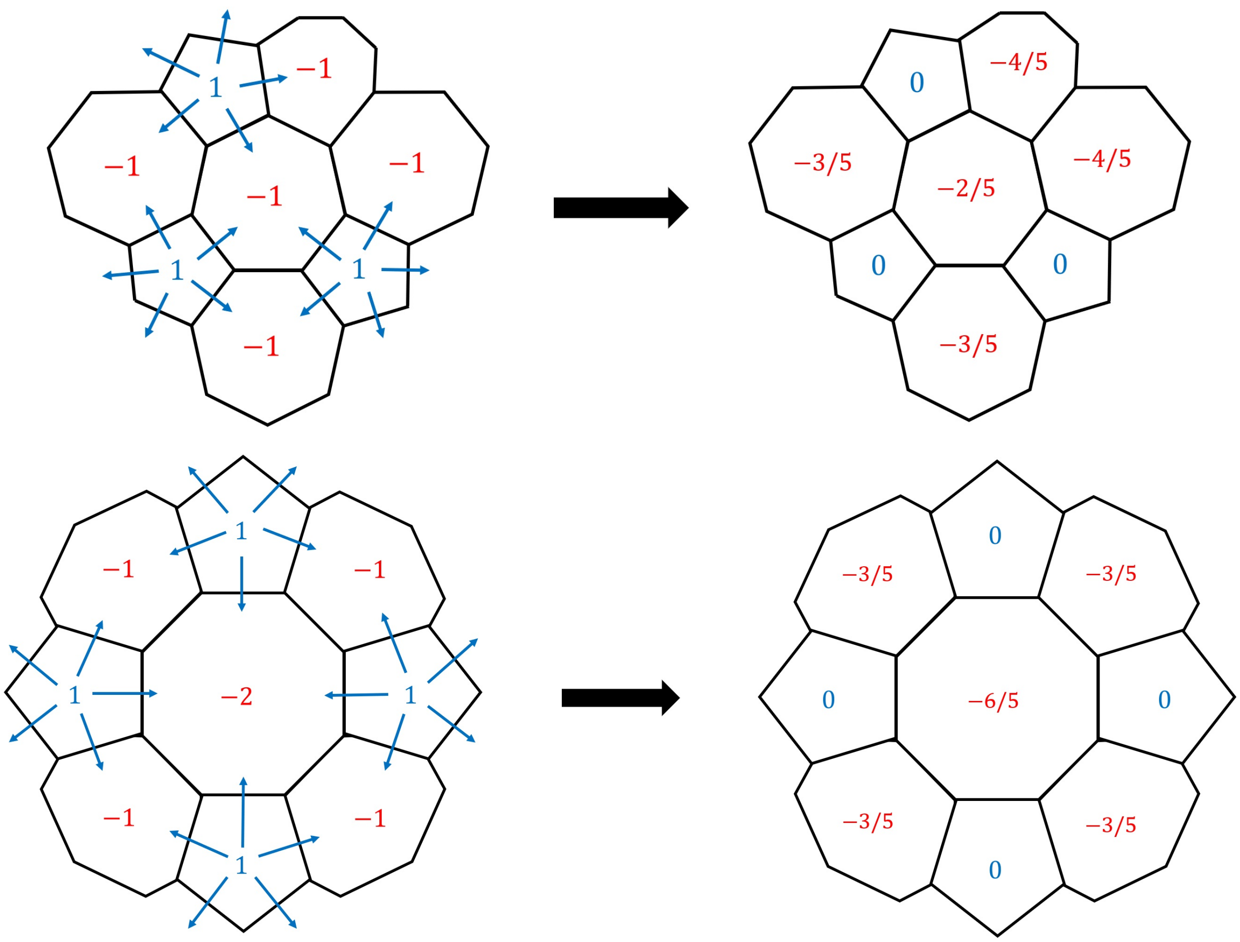
$$(1 \times C_5) + (0 \times C_6) + (-1 \times C_7) + (-2 \times C_8) + \cdots$$

$$= C_5 - C_7 - 2C_8 - 3C_9 - 4C_{10} - \cdots - (k-6)C_k - \cdots .$$

By the Counting Formula for Cubic Maps from Supplement. The Four-Color Theorem: A History, Part 1, we have

$$4C_2 + 3C_3 + 2C_4 + C_5 - C_7 - 2C_8 - 3C_9 - \cdots - (k-6)C_6 - \cdots = 12,$$

or since we have $C_2 = C_3 = C_4 = 0$, $C_5 - C_7 - 2C_8 - 3C_9 - \cdots - (k-6)C_6 - \cdots = 12$. Therefore, the total charge on a cubic map is 12. We now move charges around in such a way that the total charge is preserved (this is called *discharging the map*). We are doing so still under the assumption that the map contains no 2-sided, 3-sided, or 4-sided country, no two adjacent pentagons, and no pentagon adjacent to a hexagon. For each pentagon, we transfer 1/5 of its charge to each of its neighbors. By our assumption, the neighbors have at least seven boundary lines and so have an initial charge of at most $6 - (7) = -1$. After discharging a single pentagon, each neighbor of the pentagon has a new charge of at most $(-1) + 1/5 = -4/5 < 0$. Now we explore the charge on each country after all pentagons are discharged. Since no two pentagons are adjacent, then the number of pentagons neighboring a heptagon (a 7-sided country) is at most 3. So after discharging all pentagon, each heptagon will have a charge of at most $(-1) + 3(1/5) = -2/5 < 0$. The number of pentagon neighboring an octagon is at most 4 (and similarly for a nonagon), so after discharging the pentagons each octagon will have a charge of at most $(-2) + 4(1/5) = -6/5 < 0$ (for a nonagon the charge is at most $(-3) + 4(1/5) = -11/5 < 0$). In general, after discharging a $2n$-gon will have a charge of $(6 - 2n) + n(1/5) = 6 - 9n/5 < 0$ and a $2n + 1$-gon will have a charge of $(6 - (2n + 1)) + n(1/5) = 5 - 9n/5 < 0$. This is illustrated for a heptagon and an octagon below.



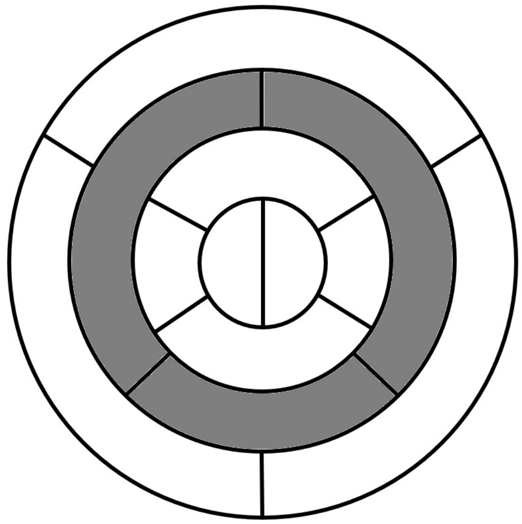
However, this means that, after discharging, all pentagons have charge 0 (they were discharged), all hexagons have charge 0 (which did not change), and all other countries have negative charge. But the total charge is unchanged and should be 12, a CONTRADICTION. So the assumption that a cubic map contains no 2-sided, 3-sided, or 4-sided country, no two adjacent pentagons, and no pentagon adjacent to a hexagon is false. That is, every cubic map must contain at least one of these so that these form an unavoidable set. This same argument is presented in Graph Theory 2 (MATH 5450) in Section 15.2. The Four-Colour Theorem, but applied to the dual of the map so that the charges are on vertices instead of countries.

Note. In the 1920s, Philip Franklin showed that a related unavoidable set contains a digon, triangle, square, and: a pentagon adjacent to two other pentagons, a pentagon adjacent to a pentagon and a hexagon, and a pentagon adjacent to two hexagons. This result is part of Franklin's Ph.D. dissertation at Princeton University. Another contributor to unavoidable sets was Henri Lebesgue (June 28, 1875–July 26, 1941) who, in 1940, wrote a paper in which he used Euler's formula and the Counting Formula for Cubic Maps to construct a number of new unavoidable sets. You know Lebesgue, of course, from Lebesgue measure and Lebesgue integration which are the main topics of Real Analysis 1 (MATH 5210; see my online notes for Real Analysis 1). The method of discharging can be used to show that many sets of configurations are unavoidable. We illustrated this in one case in Note FCT.N; but the discharging algorithm used there is specific to the given set of configurations. Different configurations may require different discharging algorithms. As the 20th century progressed, huge unavoidable sets (of thousands of configurations) were constructed. For each unavoidable set, a discharging algorithm is needed that can address all of the configurations. Ultimately, this is how the Four Color Theorem was proved.

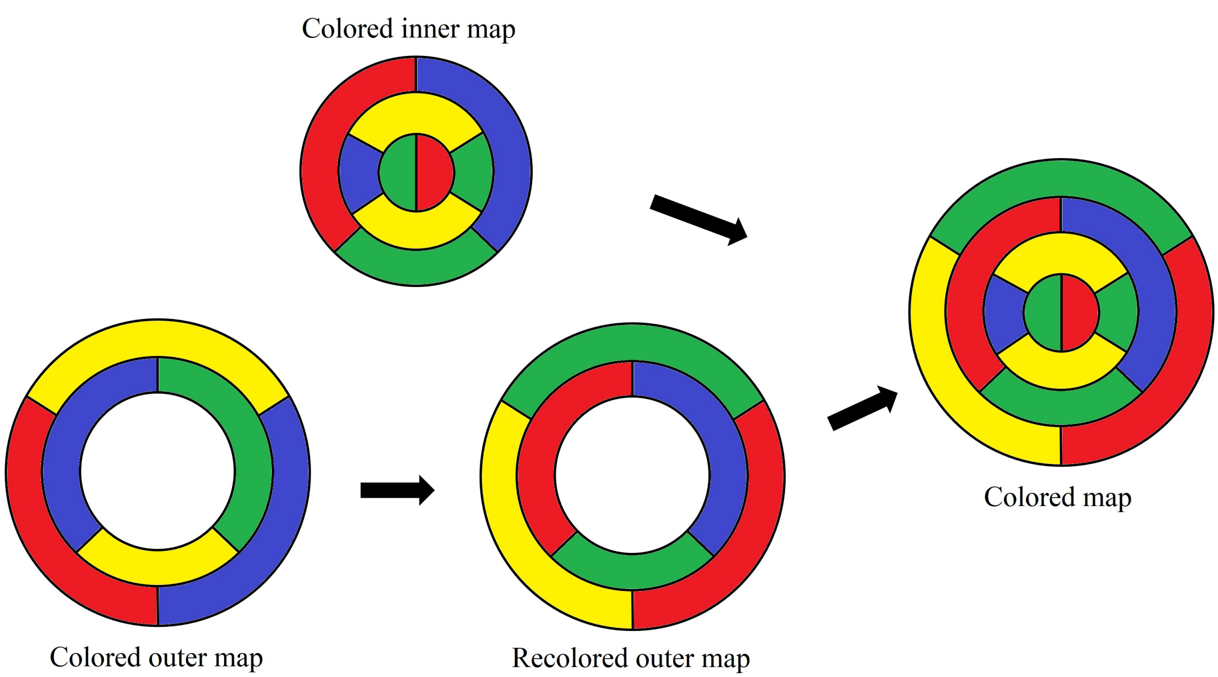
Note FCT.O. George D. Birkhoff published "The Reducibility of Maps," *American Journal of Mathematics*, **35**(2), 115–128 (1913); it available online on JSTOR (accessed 1/5/2023). In the paper, Birkhoff gave an exploration of Kempe chains that would lead to later developments on reducibility. Birkhoff considered rings of countries in a hypothesized minimal counterexample of the Four Color Theorem. Whereas Kempe's approach was to remove one country (and then use the minimality property of the alleged counterexample), Birkhoff's approach removed several countries at a time.

George D. Birkhoff
(March 21, 1884–November 12, 1944)
Image from the MacTutor History of Mathematics
Archive biography of Birkhoff (accessed 1/5/2023)

As an illustration, consider a ring of three countries (in grey) with at least one country inside the ring and at least on country outside the ring:

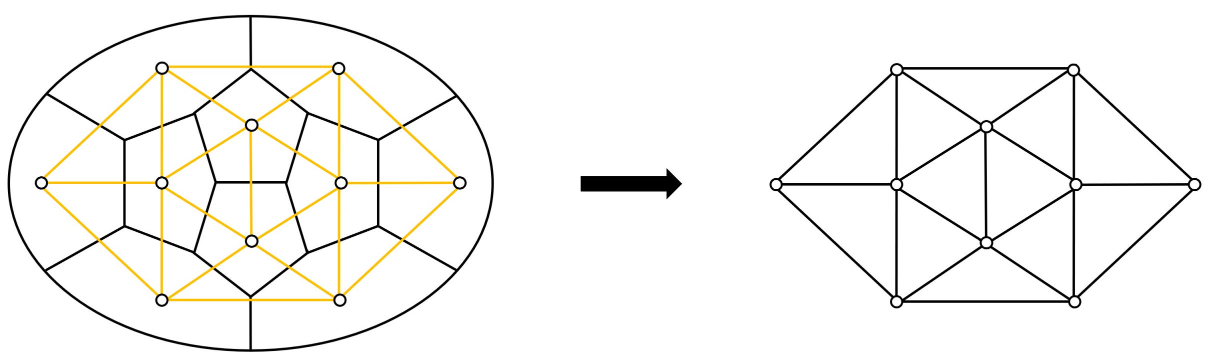


If this map is a minimal counterexample to the Four Color Theorem, then by deleting the countries outside the ring we get a map that is four colorable. Similarly, if we delete the countries inside the ring we also get a map that is four colorable. We can then mesh the two coloring together, permuting the colors in one of the smaller maps if needed, as follows:

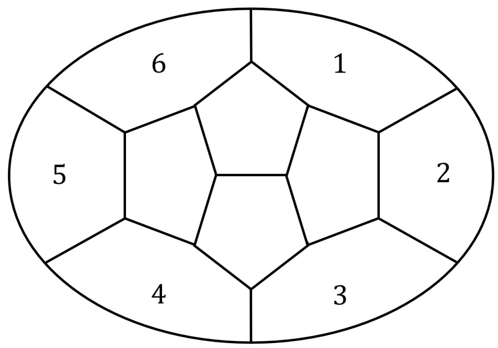


Birkhoff also showed that rings of four countries are reducible (though the construction is more complicated than the case of a ring of three countries). He was mostly, but not entirely, successful in addressing rings of five countries. He could not solve the case of a pentagon surrounded by a ring of five countries; this is the same case that Kempe failed to do correctly. Birkhoff's arguments also applied to a few cases of rings of six countries. Arthur Bernhart was able to complete Birkhoff's work on rings of six countries in "Six-Rings in Minimal Five-Color Maps," *American Journal of Mathematics*, **69**(2), 391–412 (1947), available on JSTOR (accessed 1/6/2023).

Note FCT.P. One well known six-ring of countries is the *Birkhoff diamond*. Here is the map corresponding to the Birkhoff diamond and its dual graph.



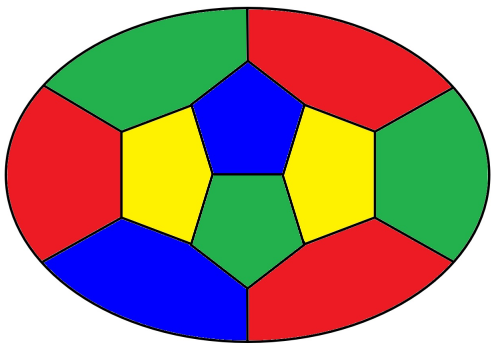
Notice that the dual graph is the Figure 15.7 of Bondy and Murty's Section 15.2. The Four-Colour Theorem (the dual graph is called "Birkhoff's diamond" in Bondy and Murty, since they are addressing the Four Color Theorem in terms of vertex colorings of the dual of a map). Bondy and Murty show that the Birkhoff diamond is reducible in their Theorem 15.7. We now give an of this here, based on the map instead of the dual graph of the map. As usual, we assume that a minimal contradiction to the Four Color Theorem exists which contains the Birkhoff diamond, and then get a contradiction by showing that the map actually is four colorable. We number the countries in the ring of six countries as 1 through 6 as follows:



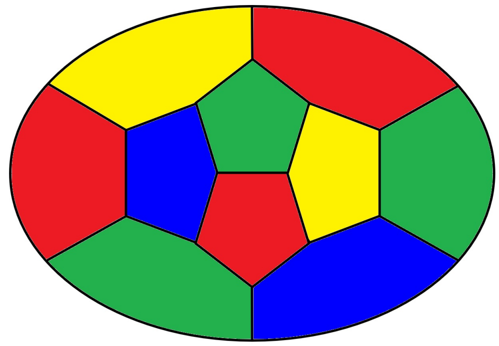
If we use the four colors red, green, blue, and yellow (abbreviated r, g, b, and y, respectively), then we get 31 essentially different colorings. By “essentially different” we mean, in a sense, “up to permutations of colors and of the map”; for example the coloring $rgrgrg$ is the same as the coloring $bybyby$ (we just interchange red and blue, green and yellow). We choose r as the color in country 1 and g is the color in country 2 throughout. The 31 colorings fall into two categories (one of which we indicate with an asterisk) are:

$rgrgrg$	$rgrbrg^*$	$rgrbgy^*$	$rgbrgy$	$rgbryb$	$rgbgbg^*$	$rgbyrg$	$rgbygy^*$
$rgrgrb^*$	$rgrbrb$	$rgrbyg^*$	$rgbrbg^*$	$rgbgrg^*$	$rgbgby$	$rgbyrb$	$rgbyby^*$
$rgrgbg$	$rgrbry$	$rgrbyb^*$	$rgbrby$	$rgbgrb^*$	$gbgyg$	$rgbyry^*$	$rgbyby^*$
$rgrgby^*$	$rgrbgb^*$	$rgbrbg$	$rgbtyg$	$rgbgry^*$	$rgbgyb$	$rgbygb$	

The 16 colorings marked with an asterisk are called (by Wilson) “good colorings.” They are such that the four pentagons can be colored without any adjustment. As examples, we here are the first ($rgrbrg$) and last ($rgbgry$) asterisked cases:

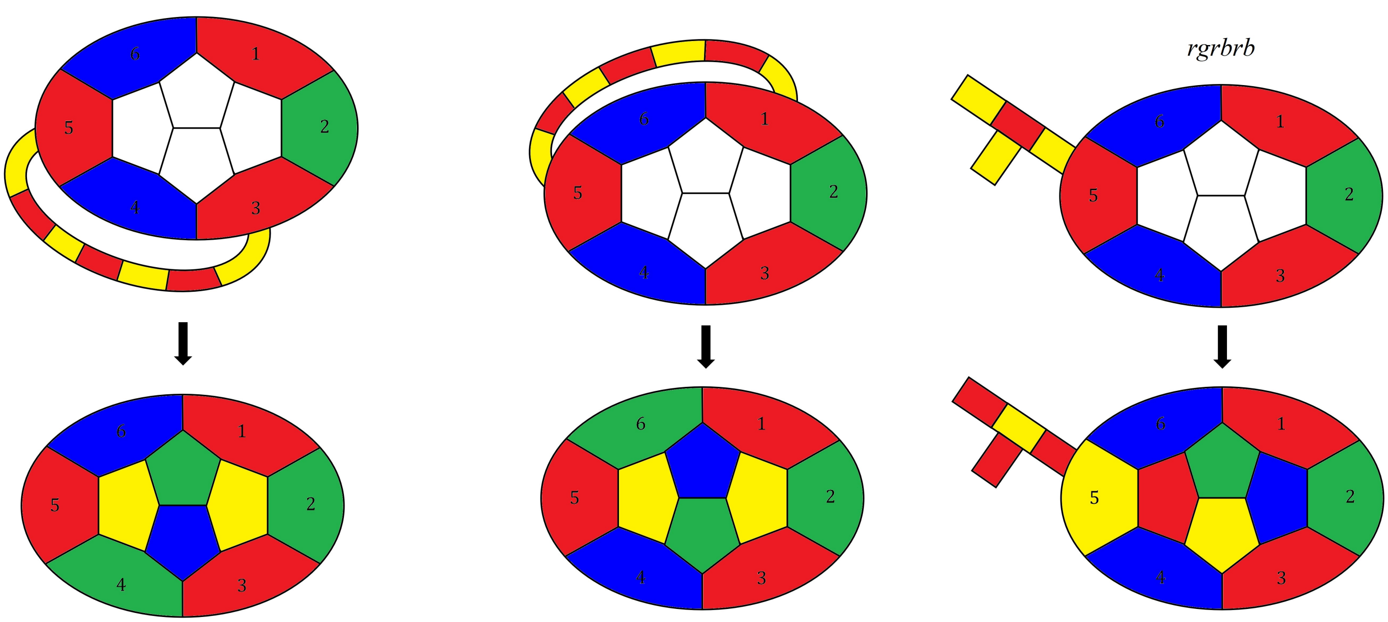


rgrbrg



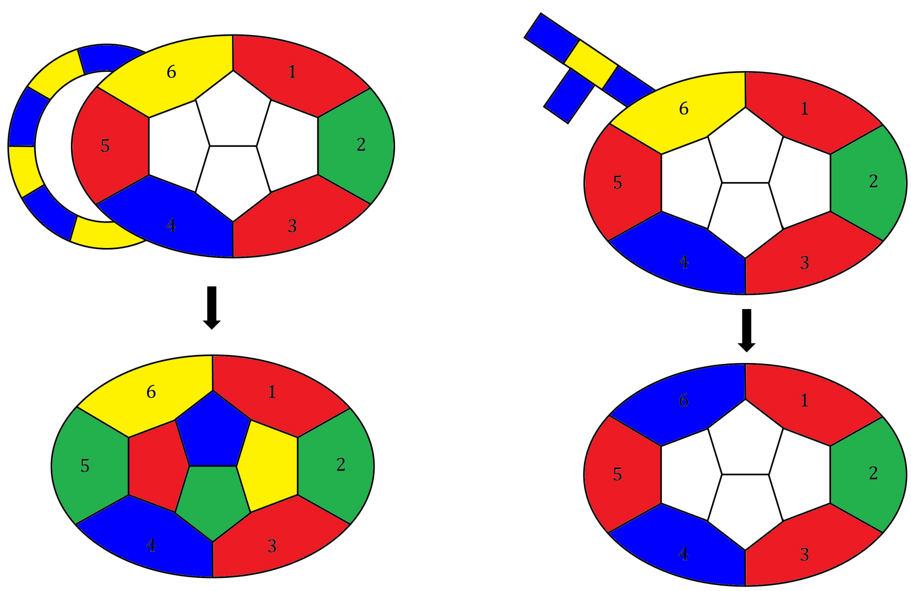
rgbgry

The remaining 15 cases require adjustments to the ring colors before the inner four pentagons can be colored. Consider the coloring of the ring *rgrbrb*. We use Kempe chains to convert it into a good coloring. If there is a red-yellow chain connecting countries 3 and 5, then we can interchange the colors in the blue-green chain starting at country 4 and constrained to be inside the red-yellow chain. This makes all such blue countries (such as country 4) green and all such green countries blue (resulting in no color violations). Then the four pentagons can be colored as given below (left).



If there is a red-yellow chain connecting countries 1 and 5, then we can interchange the colors in the blue-green chain starting at country 6 and constrained to be inside the red-yellow chain. This makes all such blue countries (such as country 6) green and all such green countries blue (resulting in no color violations). Then the four pentagons can be colored as given above (middle). If there is neither a red yellow chain connecting countries 3 and 5, nor one connecting countries 1 and 5, the we can interchange the colors in the red-yellow chain starting at country 5. This makes country 5 yellow and then the four pentagons can be colors as given above (right).

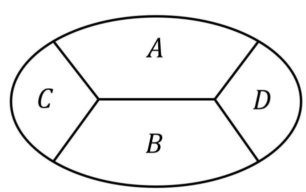
Note FCT.Q. We next consider the coloring $rgrbry$.



rgrbrb

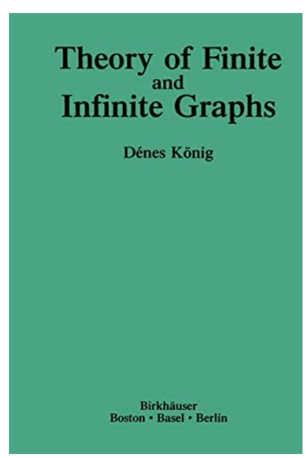
If there is a blue-yellow chain from country 4 to country 6, then the colors in the red-green chain starting at country 5 can be interchanged converting the coloring into $rgrbgy$ which is a good coloring (above left). If no such blue-yellow chain exists, then we can interchange the colors in the blue-yellow chain starting at country 6 (which does not circle back around to the ring of six countries; above right) producing the coloring $rgrbrb$ which is not good, but was dealt with in the previous example. In fact, all 31 of the possible colorings of the ring of six countries are either good or can be modified to be good through Kempe chain changes of color(s).

Note FCT.R. George Birkhoff published another landmark paper on colorings in 1912. He considered the number of λ-colorings of a graph when considering the Four Color Theorem in "A Determinant Formula for the Number of Ways of Coloring a Map," *Annals of Mathematics*, **14**(1/4), 42–46 (1912/13). A copy is available online on the JSTOR (accessed 1/8/2023). With $\lambda \geq 3$ as the number of colors used to color a given map, he showed that the number of ways to λ-color the map is a polynomial in λ called the *chromatic polynomial* of the map. These ideas are covered in Graph Theory 2 (MATH 5450) in Section 14.7. The Chromatic Polynomial where it is discussed in the setting of proper vertex colorings; the fact that the function is a polynomial is shown in those notes in Theorem 14.26. To illustrate this idea, consider the following map:



Country A can be assigned any of the λ colors. Since Country B is a neighbor of country A, then it can be assigned any of the remaining $\lambda - 1$ colors. Since countries C and D are both neighbors of countries A and B but not of each other, they can each be assigned any of the remaining $\lambda - 2$ colors. Therefore the number of possible λ-colorings of the map (by the Fundamental Counting Principle) is given by the polynomial $P(\lambda) = \lambda(\lambda-1)(\lambda-2)^2 = \lambda^4 - 5\lambda^3 + 8\lambda^2 - 4\lambda$. Birkhoff undertook this study in an attempt to prove the Four Color Theorem; notice that if $P(4) \geq 1$ for a given map, then that map can be four colored. In his paper, he proved that for all chromatic polynomials, $P(\lambda) \geq \lambda(\lambda-1)(\lambda-2)(\lambda-3)^{n-3}$ where P is a chromatic polynomial for a map with n countries and this holds for all λ *except* $\lambda = 4$. Had the inequality held for $\lambda = 4$, then the Four Color Theorem would follow since the right-hand-side of the inequality for $\lambda = 4$ is 24. So Birkhoff was close! Following Birkhoff's work, the study of chromatic polynomials and reducible configurations spread widely through the growing graph theory community in the 20th century.

Note. Graph theory was a growing discipline in the middle of the 20th century. It was proving useful in addressing a number of applications. Interest was stimulated by the appearance of several textbooks. Dénes König (September 21, 1884–October 19, 1944) published the first textbook on graph theory, *Theorie der enlichen und unendlichen Graphen*, in 1936. A translation, appearing as *Theory of Finite and Infinite Graphs*, by Richard McCoart with commentary by William Tutte, was published by Birkhäuser Press in 1990. A limited preview of the translation is available on Archive.org (accessed 1/8/2023).



The above images are from the MacTutor History of Mathematics Archive biography of König and Amazon.com (accessed 1/8/2023). The 1960s saw several graph theory books published, including: Claude Berge's *Theory of Graphs and Its Applications* (Wiley, 1961), Oystein Ore's *Theory of Graphs* (American Mathematical Society, 1961), and Frank Harary's *Graph Theory* (Addison-Wesley, 1969). Ore also published the first book devoted to map coloring with *The Four-Color Problem* (Academic Press, 1967). Another followed by Gerhardt Ringel: *Map Color Theorem* (Springer, 1974). By the way, Ringel is the coauthor of the book from which I have developed online notes for Introduction to Graph Theory (MATH 4347/5347). The timing of these early textbooks shows how "young" graph theory is as a discipline. As a quick observation, your instructor (who has a master's degree in a graph theory-related area) entered graduate school in 1984 and had never heard of graph theory at the time!

Note. Wolfgang Haken (June 21, 1932–October 2, 2022) did his Ph.D. research on topology and knot theory at the University of Kiel in (West) Germany. He published his dissertation work in 1961 and its high quality lead to him ultimately getting a job at the University of Illinois a Urbana-Champaign. There he struggled (unsuccessfully) with the Poincaré Conjecture (one of the most famous unsolved mathematical problems of the 20th century, it fell in 2002–2003 when it was proved by the eccentric Grigori Perelman). Following this, he turned his attention to the Four Color Theorem (well, the Four Color "Problem" or "Conjecture," at the time). As mentioned above, Heinrich Heesch published the idea of discharging in 1969. Heesch had come to think that an unavoidable set of reducible configurations did exist (implying the validity of the Four Color Theorem), that these configurations will not be large, but that there is likely to be very many of them in the set (Wilson, page 176). In 1967, Haken contacted Heesch about the problem. At that time, Heesch was working on extending Birkhoff's 1913 ideas for generating reducible configurations.

Heinrich Heesch

Wolfgang Haken

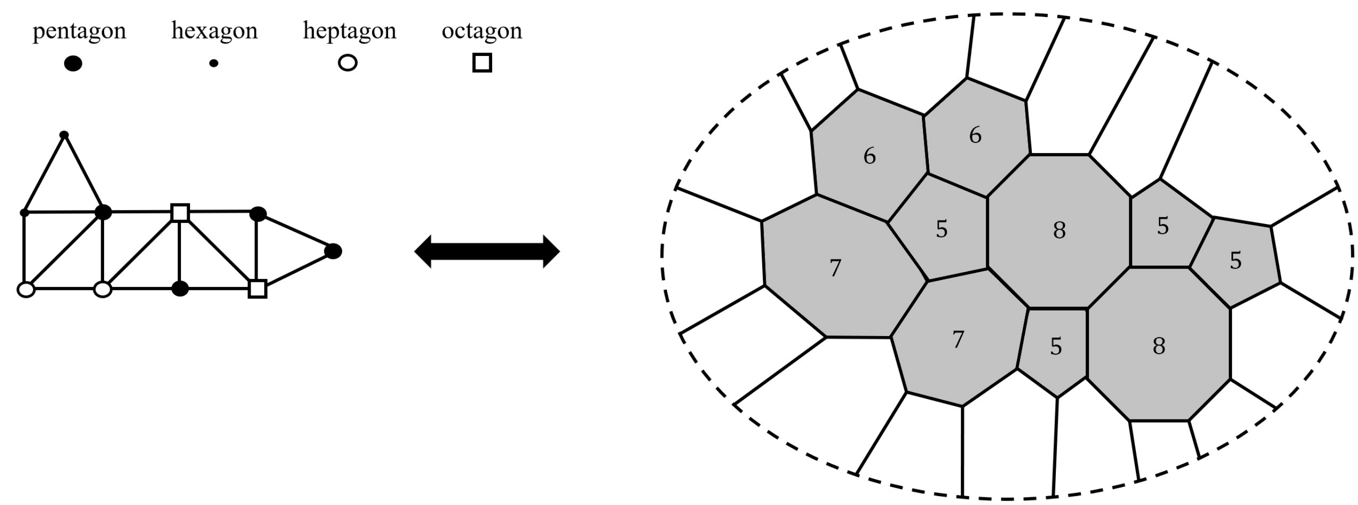
These images are from the Wikipedia page for Heesch and the Wikipedia page for Haken (accessed 1/8/2023).

Heesch defined two categories of reducible configurations. He labeled as *D-reducible* those configurations for which every coloring of the surrounding ring of countries yields an extension of the coloring without modification (we called these "good colorings" in Note FCT.P), or which can be converted into a good coloring by a succession of Kempe-chain color interchanges. The configurations in Notes FCT.P and FTC.Q (including the Birkhoff diamond) are examples of D-reducible configurations. Heesch labeled as *C-reducible* those configurations that can be proved reducible after they have been "modified in some way" (we leave this category vaguely defined like this). Heesch introduced these categories in an attempt to systematize Birkhhoff's approach and to make the testing of configurations algorithmic so that showing reducibility could be done by a computer. In 1965 Heesch, with the help of Karl Dürre, had software (written in "Algol 60") that could show the D-reducibility of many configurations of increasing complexity. The complexity of a configuration is determined by its ring-size; as we claimed in Note FCT.P when the ring has size 6, there are 31 possible colorings to consider. The complexity grows rapidly as:

ring-size	6	7	8	9	10	11	12	13	14
colorings	31	91	271	820	2461	7381	22144	64430	199291

Computational complexity is discussed in more detail (though still a bit informally) in my online notes for Mathematical Modeling Using Graph Theory (MATH 5870); see Section 8.1. Computational Complexity. It was thought at the time that ring-size up to 18 would have to be considered to prove the Four Color Theorem (with ring-size 18, there are over 16 million colorings; Wilson page 181). It turns out that Appel and Haken's 1976 solution showed that only ring-size up to 14 had to be considered (Wilson page 182).

Note FCT.S. Heinrich Heesch came up with a short-hand notation to represent configurations. The publication of the proof of the Four Color Theorem had dozens of pages filled with configurations represented using a similar notation. Notice below that Heesch's notation is related to the dual of part of the map (provided we treat each of the different symbols as representing vertices; the part represented is gray in the figure); see Section 10.2. Duality. Since the maps are all cubic, then the dual graphs all have faces of degree three; that is, the dual graph is a triangulation of the plane. Also, the symbol determines the degree of each vertex in the dual so that, for example, the dots representing pentagons, "•," are each of degree five in the dual graph. This is implied by Heesch's notation, but it is not the case that each such vertex is degree five in Heesch's *notation*; this is because the notation represents a configuration that is *part* of a larger graph.



Note. At Haken's home university, the University of Illinois, no supercomputer was available (though one was under construction) so the software could not be tested there. Heesch and Dürre connected with Yoshio Simamoto (1924–August 27, 2009) of the Atomic Energy Commission's Brookhaven Laboratory who had access to a Cray supercomputer. Shimamoto was a "devotee" of the Four Color Theorem and had access to time on the Cray. Dürre converted the software from Algol to Fortran. Using hours of time on the Cray, Heesch and Dürre were able to confirm the D-reducibility of over 1000 configurations of ring-size 14 or less. Shimamoto, pursuing his on research on the problem, was able to show that if he could fins a single configuration with certain properties and if this configuration were D-reducible, then the Four Color Theorem would follow (Wilson, page 186). He found the following configuration, called the *Shimamoto horseshoe* and proved that if this is D-reducible, then the Four Color Theorem holds.

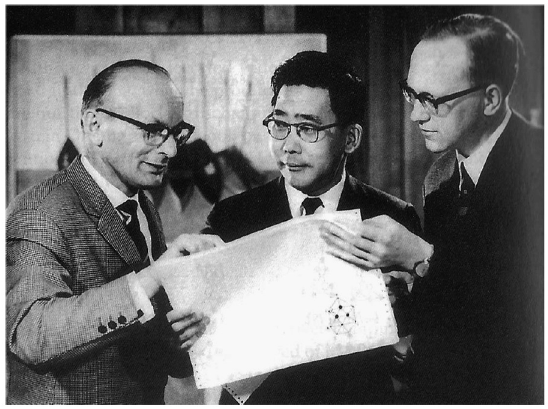
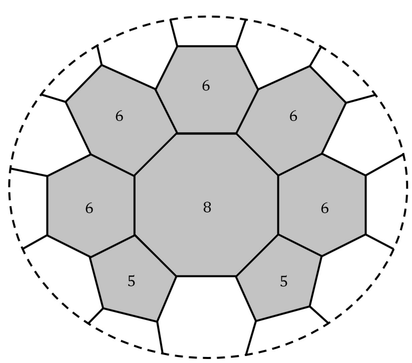


FIGURE 7. Heesch, Shimamoto, and Dürre at Brookhaven National Laboratory.

This image from Robin Wilson's "Wolfgang Haken and the Four-Color Problem," *Illinois Journal of Mathematics*, **60**(1), 149–178 (2016); available on the Celebration Mathematica webpage (accessed 1/10/2023).

Shimamoto, pursuing his on research on the problem, was able to show that if he could fins a single configuration with certain properties and if this configuration were D-reducible, then the Four Color Theorem would follow (Wilson, page 186). He found the following configuration, called the *Shimamoto horseshoe* and proved that if this is D-reducible, then the Four Color Theorem holds.



Ultimately, in late 1971 the Cray showed that the horseshoe is *not* D-reducible after 26 hours of computer time (Wilson page 188). The Four Color "Theorem" remained an unproved conjecture. This would change in five years.

Note. We now return the story to Wolfgang Haken. He had been in communication with Heesch during the time of the Brookhaven computer work, and Heesch had sent him unpublished results on reducible configurations. At this time (the early 1970s), most approaches were to collect reducible configurations by the hundreds and then form an unavoidable set. Haken prioritized the unavoidable sets themselves and looked for configurations likely to be reducible. Any configurations that later proved not to be reducible could be addressed one at a time (Wilson pages 193 and 194). During a lecture Haken gave on the Brookhaven work and the Shimamoto horseshoe, he stated that the "computer experts" told him that a computational approach was not promising. In attendance was Kenneth Appel (October 8, 1932–April 19, 2013), who did Ph.D. work in mathematical logic and algebra, and who had computing experience. Appel approached Haken after the talk and told him that the "experts" were wrong and that computational approach may take time, but that it should work. This lead to the Appel/Haken collaboration starting in 1972.

Note. When their collaboration started in late 1972, they did not have a clear idea of how to completely process the problem. Early computer runs proved informative and hinted at the direction to proceed. However, the computer outputs were huge printouts and many configurations were repeated. They revised their programs and a second run of their software showed improvements, including a reduced printout. They began regularly revising their discharging algorithm and their printed output shrank further. After six months, they were confident that their technique could be used to produce a finite unavoidable set which could be processed by computer in a reasonable amount of time (Wilson page 197). In mid-1974, Appel visited the computer science department at the University of Illinois looking for a graduate student who could, as part of their dissertation work, assist with additional programming work. John Koch was available and became part of the project (he would earn his Ph.D. under the direction of Appel).

This image from Robin Wilson's "Wolfgang Haken and the Four-Color Problem," available on the Celebration Mathematica webpage (accessed 1/10/2023).

John Koch in 2020

This image is from The Archives & Special Collections of Wilkes University webpage (accessed 1/10/2023).

Koch was first assigned the problem of the C-reducibility of configurations of ring-size 11. He found an efficient method for testing for this reducibility, which Appel extended to configurations with ring-sizes 12, 13, and 14. By the end of 1975, the discharging method was causing occasional problems related to pentagons and hexagons. In order to avoid rewriting their software, it was decided to perform the final part of the discharging process by hand. Though this was work-intense, it gave them a level of flexibility that allowed them to restrict all of their configurations to ring-size 14 or less (Wilson pages 201 and 202). During the first half of 1976, Appel and Haken finalized their discharging procedure and ended up with 487 discharging rules. In March 1976, the University of Illinois acquired a powerful new computer for administrative use. Appel and Haken were granted time on the computer. Appel, Haken, and Haken's daughter, Dorothea, spent months working through the 1936 configurations that would eventually form the unavoidable set. By late June of 1976, the unavoidable set was constructed and in two days Appel was able to test it for reducibility. Additional checking of details was needed and the results had to be written up, but on July 22, 1976 Appel and Haken publicly announced their proof of the Four Color Theorem.

Note. Appel and Haken (and Koch) were not the only ones working on the problem. Frank Allaire of the University of Waterloo in Ontario was working on reducibility methods that he had inherited from Jean Mayer. By 1976, Allaire was months ahead of Appel and Haken and was expecting to complete his solution in a few months. He joined with Ted Swart of the University of Rhodesia (now Zimbabwe), who was working on a similar approach to that of Allaire's. Before Appel and Haken made their announcement, Allaire and Swart submitted a paper dscribing their algorithm for determining reducibility and including a list of all reducible configurations with ring-size 10 or less. Their paper appeared as "A Systematic Approach to the Determination of Reducible Configurations in the Four-Color Conjecture," *Journal of Combinatorial Theory B*, **25**(3), 339–362 (1978); that can be viewed online on the ScienceDirect.com webpage (accessed 1/11/2023). Allaire and Swart had a more systematic approach then that of Appel and Haken. Allaire was understandably disappointed when he learned of Appel and Haken's result, but he was professional about it and in fact was the referee of the reducibility part of the Apple-Haken paper (Wilson pages 211 and 212). Others working on the Four Color Theorem at the time were Walter Stromquist, a doctoral student at Harvard University, and Frank Bernhart of the University of Oklahoma who had published previous work on Birkhoff's work on rings of six countries and was creating additional reducibility arguments.

Note. The first publication concerning the proof of the Four Color Theorem is a two page research announcement: Kenneth Appel and Wolfgang Haken, "Every Planar Map is Four Colorable," *Bulletin of the American Mathematical Society*, **82**(5) (September 1976). This can be read online on the Project Euclid webpage (accessed 1/11/2023). Appel and Haken decided to submit their paper to the *Illinois Journal of Mathematics* (since this was a "local" journal for the authors, they were able to suggest those best people to referee that paper). Jean Mayer (mentioned above in connection with Allaire's reducibility methods) refereed the discharging arguments and Frank Allaire refereed the reducibility part. Haken and Appel spent late 1976 and early 1977 refining their paper, conferring with the referees on details, and preparing a microfiche of 450 pages of diagrams and explanations. The paper had 1482 reducible configurations (instead of the 1936 that were considered in an earlier version of the paper) and 487 discharging rules, and appeared in two parts plus the microfiche supplement. The references are:

- K. Appel, W. Haken, "Every planar map is four colorable. Part I: Discharging," *Illinois Journal of Mathematics*, **21**(3), 429–490 (September 1977), and
- K. Appel, W. Haken, J. Koch, "Every planar map is four colorable. Part II: Reducibility," *Illinois Journal of Mathematics*, **21**(3), 491–567 (September 1977).

These papers can be read or downloaded from the Project Euclid webpage (accessed 1/11/2023). The first figure in Part I is a legend that illustrates the presentation of the many configurations; in fact, they use the same notation as introduced by Heesch and illustrated above in Note FCT.S.

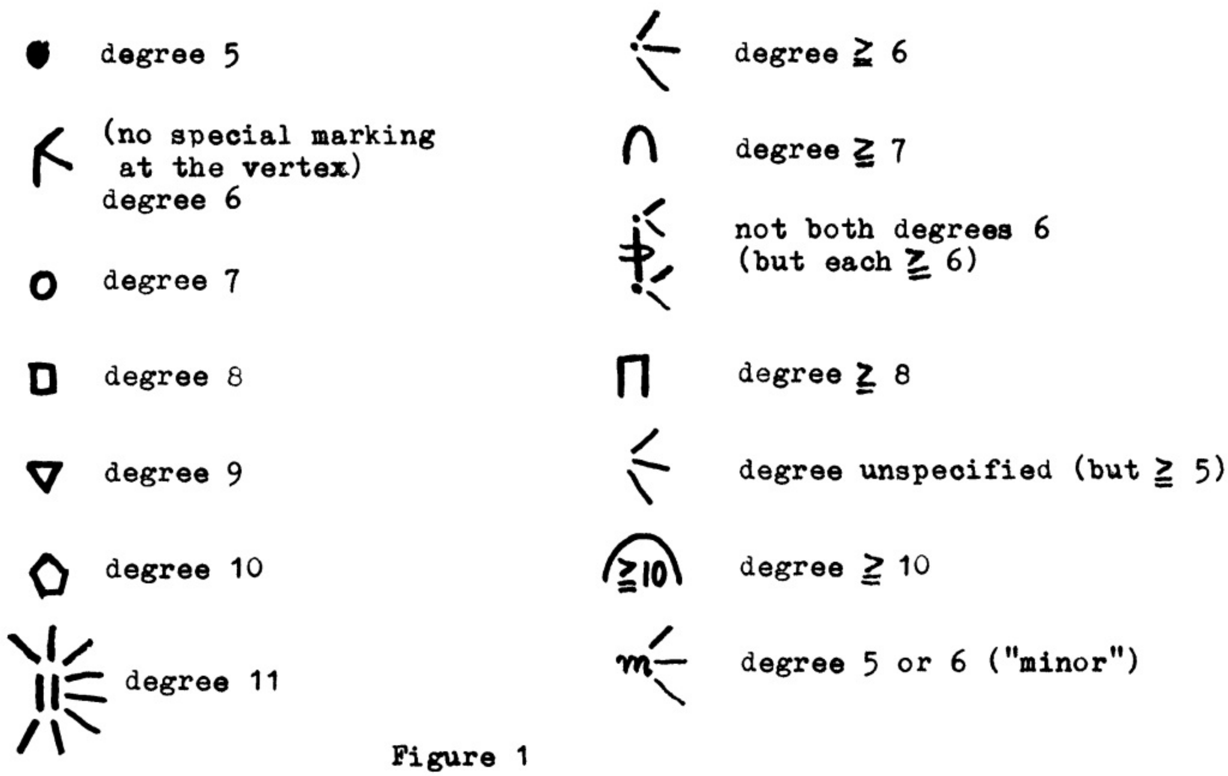


Figure 1 from Part I: Discharging, page 436

Part I: Discharging includes 24 pages of figures (on pages 440–457 and 472–477) related to "situations of small/large dischargings." Part II: Reducibility includes 14 pages of prose interspersed with figures, followed by 63 pages of figures (on pages 505–567) related to reducible configurations. An example of this material is Figure 14 from page 548 of Part II is below. N. Robertson, D. Sanders, P. Seymour, and R. Thomas in "The Four-Colour Theorem," *Journal of Combinatorial Theory, Series B*, **70**(1), 2–44 (1997), used "only" 32 discharging rules to find 633 unavoidable configurations (again, with computer assistance; available online on the Science Direct website; accessed 1/11/2022).

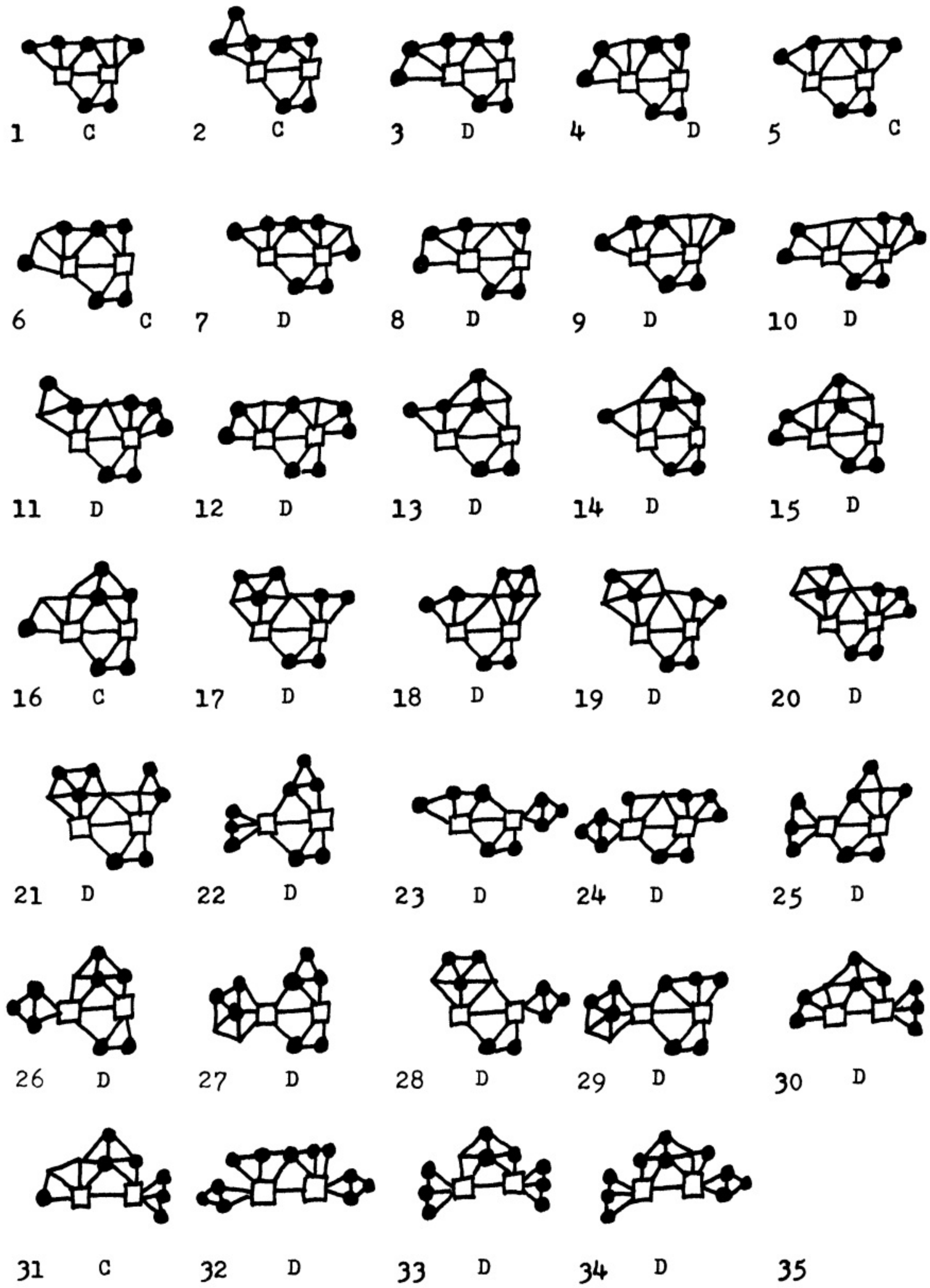


Figure 44

Figure 44 from Part II: Reducibility, page 548

Note. The Appel, Hakin, and Koch 1977 proof drew quick criticism. Since key parts of the proof were dependent on a computer search that could not be checked by hand, the proof itself could not be checked. The controversy receded some with the publication of Appel and Haken's *Every Planar Map is Four Colorable*, Contemporary Mathematics #98, 741 pp., American Mathematical Society (1989). This work is described on the AMS Bookstore webpage as:

> "...the book contains the full text of the supplements and checklists, which originally appeared on microfiche. The thirty-page introduction, intended for nonspecialists, provides some historical background of the theorem and details of the authors' proof. In addition, the authors have added an appendix which treats in much greater detail the argument for situations in which reducible configurations are immersed rather than embedded in triangulations. This result leads to a proof that four coloring can be accomplished in polynomial time." (Accessed 9/4/2022.)

The relatively simple 1997 *JCT-B* paper of N. Robertson, D. Sanders, P. Seymour, and R. Thomas also calmed things. Some additional infomation is given in my online notes on Section 15.2. The Four-Colour Theorem.

Note. The complaints that computer techniques have intruded into pure mathematics proofs still persist. One resolution for the Four Color Theorem would be to introduce a non-computer-based proof. But this would require new ideas and, over the past 45 to 50 years since Appel and Haken's initial success, there seems to be no progress in this direction. The mathematics writer Ian Stewart is paraphrased by Wilson (page 220) as complaining that the Appel-Haken proof "did not explain why the theorem is true—partly because it was too long for anyone to grasp all the details, and partly because it seemed to have no structure." The purpose of a mathematical proof is not so much to learn *what* is true about mathematical structures, but *why* the mathematical structures have the proven properties.

Note. Objections to the proof are all based on the use of computers; it is not the fact that it required so many cases or that the papers were so long. The classification of finite simple groups was a project that took over 30 years, covered between 5,000 and 10,000 journal pages spread over 300 to 500 individual papers. Yet there is no objection to this proof; see my online notes for Introduction to Modern Algebra (MATH 4127/5127) on Supplement. Finite Simple Groups. *Personally*, I have no strong objection to the use of computers in searching cases in a discrete math proof. I have used this approach twice in my research (but only because I could not find a simpler argument) In R. Gardner and T. Holt, "Decompositions of the Complete Symmetric Digraph into Orientations of the 4-Cycle with a Pendant Edge," *Congressus Numerantium*, **190**, 173–182 (2008), to show the nonexistence of a certain structure, I considered 69,120 cases and tested them with a small program written in basic. I would have much preferred that such a structure would have existed, so that I could simply have presented it! Similarly, if a counterexample to the Four Color Theorem existed, then it could simply be given...but we don't always get what we want (only, according to the Rolling Stones, what we need).

Revised: 4/27/2023