The Four Colour Theorem

Ewan Wakelin-Smith

Teddy Rocks Maths 2025

When you first look at the problem, it seems like a riddle: how many colours do you need to shade a map so that no two touching regions share the same colour? The answer (four for those of you who don't want to wait until the next paragraph) may seem simple, but proving it was certainly not. The theorem left mathematicians scratching their heads for a little over a hundred years and was the first major problem to be proven with computer assistance. What started out as a fun puzzle had turned into a cornerstone of graph theory.

What is the Theorem?

The theorem states that no more than four colours are required to colour the regions of any map so that no two adjacent regions share a common boundary of non-zero length (not merely a corner where three or more regions meet – this would mean that a hexagon connecting each of its vertices to the centre would have to have seven colours).

(https://commons.wikimedia.org/wiki/File:France-four-color-map.svg)

In graph terminology, this means that using at most four colours a planar graph (any graph that can be drawn without any of its edges crossing) can have its nodes (points) coloured so that no two adjacent ones have the same colour.

The History

The conjecture that any map could be coloured with four colours first appeared in a letter from Augustus De Morgan to his friend William Rowan Hamilton. The problem had been suggested to De Morgan in 1852 by one of his students – Frederick Guthrie, whose elder brother Francis (a botanist) had come up with the idea whilst trying to colour in counties on a map of England. By 1878, the question was widely known throughout the mathematical community and mistaken proofs were offered to this problem. There were many attempts to prove the theorem, such as the one by Sir Alfred Kempe – who studied mathematics at Cambridge and was taught by Cayley. He became a barrister and kept mathematics as a hobby which he devoted a lot of time to, he came up with the idea of Kempe chains but failed to prove the theorem. Finally in 1976, Kenneth Appel and Wolfgang Haken managed to prove the theorem using the assistance of a computer.

De Morgan's Attempt (around 1860)

De Morgan used the fact that in a map with four regions, each region touching the other three, one of them is completely enclosed by the others. However, he could not find a way to prove this and thus used it as an axiom (a statement that is taken to be true, to serve as a premise for further reasoning and arguments). He believed, incorrectly, that this idea lay at the heart of the problem which soon became his obsession

De Morgan came up with a proof by induction (which means proving for an initial case, like one, and then proving that if the theory is true for n, it is also true for n+1) that says if you have a map (M) with n+1 regions and you know the result is true for any map with 4 or less regions, then we suppose it is true for a map with n regions. He came up with the following points:

(i) If M has four mutually adjacent regions, one of them is surrounded by the others. Removing this region results in a map M' with n regions, which can be four-coloured by the induction hypothesis. When the lost region is reinstated it is adjacent to at most three colours, and so there is a colour available for it.

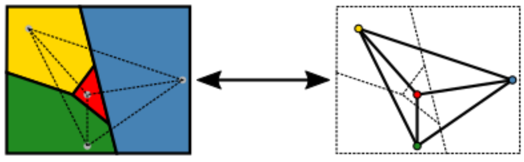
(ii) If M does not have four mutually adjacent regions, four colours are never needed in a 'neighbourhood' and so the induction step proceeds as in (i). Thus the theorem holds for each n.

However, part (ii) of De Morgan's proof is where it falls down. If you have four regions mutually adjacent in the map, then you are never forced to use four different colours in a "local neighbourhood". In other words, the proof assumes every region's immediate neighbours can be coloured in at most three colours.

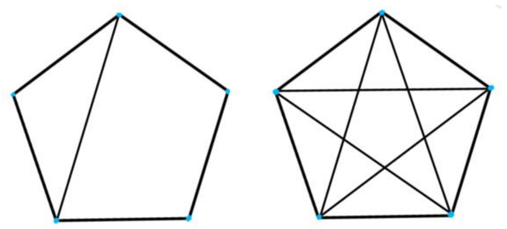
Even if no set of four regions are mutually adjacent, it is still possible to arrange regions so that a region as adjacent to four regions (or to regions where you have to use four different colours). In that case, when you try to add back a region you removed into a four-coloured map, it might be adjacent to regions in all four colours – so you are not left with a free colour to colour it.

Displaying Maps as Graphs

With any map, you can represent the middle of the regions as vertexes and any edges connecting as vertexes between points. A graph is planar if it can be drawn on a flat surface without any of the edges crossing except at their vertices.



(From Wikipedia https://en.m.wikipedia.org/wiki/Four_color_theorem)



A planar graph on the left and a non planar graph on the right.

Kempe's Proof

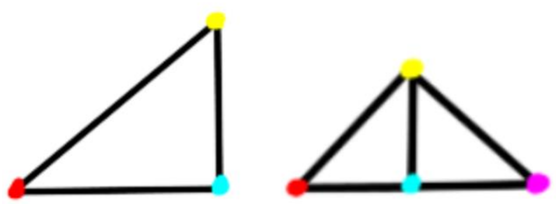
Kempe's "proof" uses strong induction on the number of regions in the map, meaning he attempts to show that a graph with n nodes can be coloured using only four colours (four coloured) and assumes that all graphs with less than n nodes can be four coloured.

Before he gets to the inductive idea, Kempe begins with a clever way to categorize the possible graphs. He first states that we might as well consider graphs all of whose faces are triangles. These maximal planar graphs include the maximum number of edges possible, and if they can be four-coloured, then so can other graphs with fewer edges. He then proves a lemma (a helper theorem) that any maximal planar graph must have at least one node with degree 5 (five nodes connected to the node) or less (By Euler's theorem, if all the nodes had degree six or higher, then the graph would not be planar).

Kempe then iterates through the four cases where degrees equal 2, 3, 4 or 5; if a graph had a degree of one, it would not be triangular.

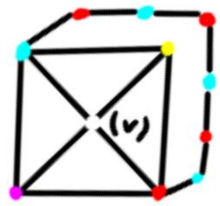
When the degree is two or three:

With degrees of two and three, the proof is easy as for four colours, you can just assign them as needed.



When the degree is four:

Kempe now needed to invent the idea of what is now called a Kempe chain (A connected chain of vertices on a graph with alternating colours).



(A Kempe chain)

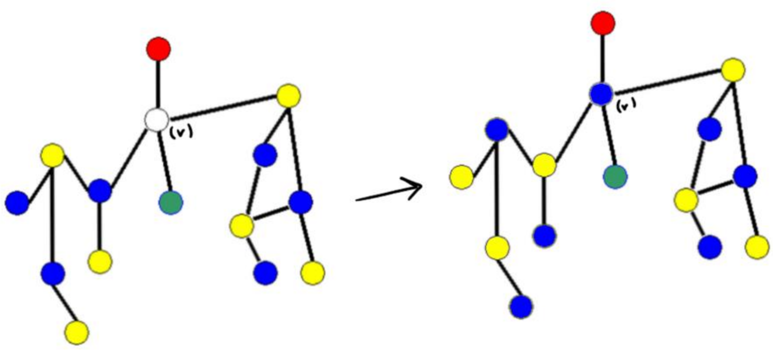
The process is as follows, we can remove v from the middle of the graph and then colour the points around it by induction. If the nodes around v are four different colours, then we would not be able to assign a different colour to v and thus we must recolour the graph, if there are less than four colours surrounding v, then we can colour v the one of the missing colour(s). Please note that the edges between the nodes in the subgraph and the original graph connect blue to yellow and green to red (and vice versa)

For this worst case scenario, we can start from blue node and then create a subgraph by following this graph through all nodes coloured blue or yellow. There are now two cases, either the subgraph that we have creates does not include the yellow node (case one) or it does (case 2).

Case 1:

This is the easier of the two cases, as we can toggle the blue and yellow nodes on the left hand side of the graph so that we now have an extra colour to use in node v.

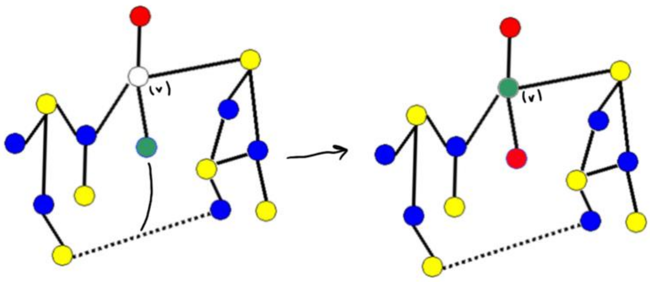
This is the easier case, as we can now toggle yellow to blue and blue to yellow, and this frees up the v for a colour.



(Image from https://sites.google.com/site/shaisimonson/professional-page/mathexp/lab4, slightly modified)

Case 2:

This subgraph now links the blue node next to v to the yellow node next to v. This is called a cyclical Kempe chain. Kempe's idea of colour switching does not work here because the the blue and yellow nodes adjacent to v switch colour. Kempe then fixes this problem by pointing out that if there is path between the nodes adjacent to v coloured blue and yellow, then there cannot be a path between the green and red nodes. This connected Kempe chain thus acts as a physical barrier and stops the red and green nodes from connecting.



(Image from https://sites.google.com/site/shaisimonson/professional-page/mathexp/lab4, slightly modified)

Now, If there is no path between the red or green nodes, we can simply toggle one of the red and green nodes connecting to v (and the respected nodes above or below it) and have a colour free.

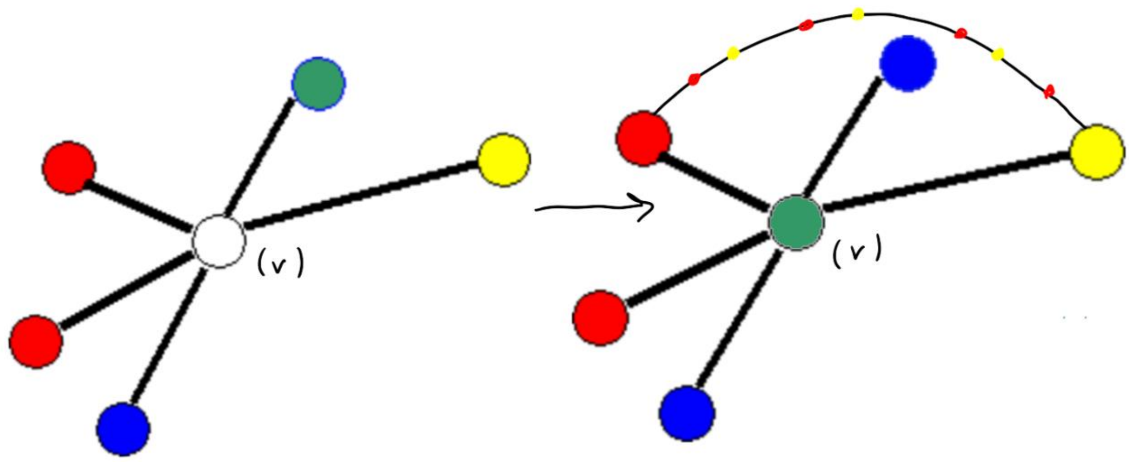
When the degree is five

Now we have five nodes connecting to x. We can follow the usual process and use induction to colour the smaller graph with four colours by induction. If the five nodes use fewer than four colours, we can simply colour v. However if the five adjacent nodes use four different colours then we must consider the possible cases.

Case one:

Case 1 is where we have two of the same colour next to each other (In this case red).

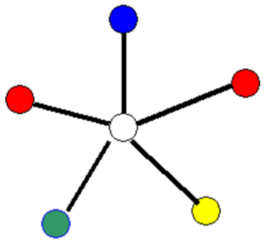
(Image from https://sites.google.com/site/shaisimonson/professional-page/mathexp/lab4, slightly



modified)

We can handle it much the same way as the fourth degree situation by using Kempe chains and toggling.

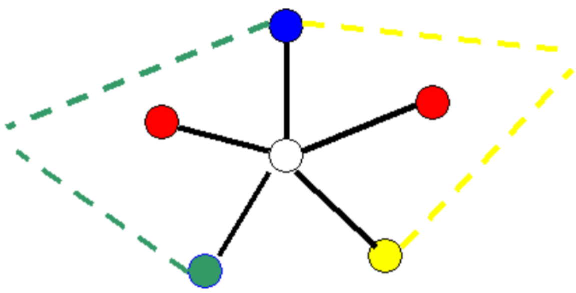
Now, if we take case 2, which is where the same coloured nodes are separated.



(Image from https://sites.google.com/site/shaisimonson/professional-page/mathexp/lab4)

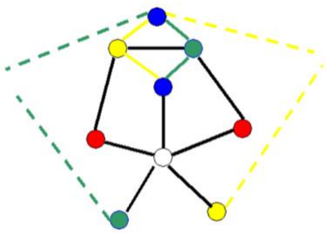
We create a Kempe chain from the blue node by following all edges that connect to nodes coloured blue or yellow. As before, if this chain does not contain the node adjacent to v that is coloured yellow, we toggle the colours of the chain, resulting in blue being able to be coloured yellow and leaving a colour free. However if the chain does contain a yellow node, then we give up on that chain and create a new Kempe chain starting from the blue node but this time by following all edges that connect to nodes coloured blue or green. If this chain does not contain the node adjacent to v that is coloured green, then we can toggle the colours of this chain, once again allowing us to colour v blue. If this second chain does contain the node coloured green, then we must reject this chain as well, and we are left with the hardest case. We cannot use any Kempe chain starting from blue.

Kempe's solution in this case is to start at each of the two red nodes and create two Kempe chains, one with red and green and the other with red and yellow. From the red node that is surrounded by the blue green Kempe chain, he creates a red and yellow Kempe chain; this chain cannot reach the yellow node adjacent to v, so red becomes yellow. This means the red node on the right is coloured green and the red node on the left is coloured yellow. This leaves red free for the middle to now use.



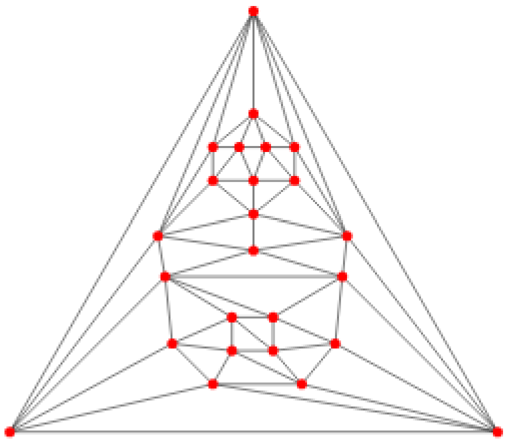
(Image from https://sites.google.com/site/shaisimonson/professional-page/mathexp/lab4)

However, this final step is where the flaw in Kempe's proof, found by Heawood, creeps in. Duplication is in the last option of case 2 as Kempe explains the two Kempe chains going away from the two red nodes. The problem is with the last step of the process where the two reds get changed to different colours. The claim is correct in why we can colour the nodes adjacent to v but a problem forms as it is possible to have edges between yellow nodes in the red yellow Kempe chain and green in the red green Kempe chain. If we change the colours according to Kempe's method, then yellow and green get changed to red. Without loss of generality (meaning we are picking one case, but the same logic works for the rest of the cases too), assume we first change yellow to red and then when green is changed to red later, the yellow which is now red gets changed to green and that green may have an edge to another green in the blue green Kempe subgraph.



(Image from https://sites.google.com/site/shaisimonson/professional-page/mathexp/lab4)

We can get misled into thinking the green nodes surrounded by the blue yellow chain cannot be adjacent to any yellow nodes surrounding the blue green chain, but they can. This is what is called Kempe Chain entanglement.



This is the shape that Heawood produced which leads to unavoidable entanglement, which stops the proof from working.

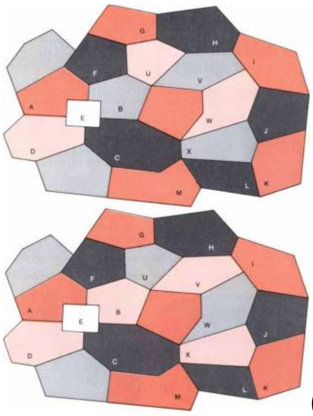
(picture from wolfram mathworld https://mathworld.wolfram.com/HeawoodFour-ColorGraph.html)

The Computer Proof

In the 1960s and 1970s, Heinrich Heesch developed methods of using searching for a proof with computers. He was the first to use discharging for the theorem and also expanded on reducibility and developed a test for it along with Ken Durre (concepts explained below). Unfortunately, he was unable to proceed due to lack of supercomputer time to finish his work.

Others took up his methods and Kenneth Appel and Wolfgang Haken at the University of Illinois were the first to prove the theorem. They built their proof using ideas like Kempe chains that they got from Kempe's failed proof.

They came up with the following key steps in their method to prove the theorem:

1. **Reducing to a fixed set of configurations**
   They proved that any minimal counterexample (the smallest map that cannot be four-coloured) must contain at least one member of a specific finite set of configurations—called an “unavoidable set.” Initially, they identified 1,936 such configurations, but this number was able to be brought down later.
   
   (an example of a configuration https://celebratio.org/Haken_W/article/794/)
2. **Showing reducibility for each map**
   For each configuration in this unavoidable set, they used computers to show that if a configuration appears, it can be replaced by a smaller map which you can colour with four colours (reducible). This step gets rid of the error in Kempe’s proof by avoiding his recolouring steps.
3. **The Discharging Method:**
   To prove that every map must contain one of these configurations, they employed a “discharging” technique—where you redistribute charge (which is a value assigned to each vertex) throughout the map keeping the overall sum the same. The discharging rules made sure that at least one member of the unavoidable set remained after redistribution.

Later work by Robertson, Sanders, Seymour, and Thomas streamlined the process further by reducing the unavoidable set from 1,936 configurations to 633, making the proof more efficient and the computer verification more manageable. The original proof by Appel and Haken took over 1200 hours of computer time to verify.

However, this proof was controversial at the time. Lots of mathematicians raised questions like “How can we trust a proof we can’t fully verify ourselves” and “What if there is a bug in the hardware or software”. People also objected to the way that the computer’s proof was less about insight and elegant reasoning but rather brute force verification.

Why is this Proof Important Today?

Whilst being an interesting challenge the four colour problem also has some practical applications that you probably haven’t considered. One use is with phone towers, as you only need at most four frequencies to fill the entire map and provide coverage. Another use is when colouring chemical bonds, as it helps represent molecules where different bonds need to be distinct.

References

https://nrich.maths.org/articles/four-colour-theorem

https://sites.google.com/site/shaisimonson/professional-page/mathexp/lab4

https://projecteuclid.org/journals/involve-a-journal-of-mathematics/volume-2/issue-3/How-false-is-Kempes-proof-of-the-Four-Color-Theorem/10.2140/involve.2009.2.249.full

https://mathworld.wolfram.com/HeawoodFour-ColorGraph.html

https://math.stackexchange.com/questions/3650823/what-does-pairwise-mean-in-graph-theory

https://celebratio.org/Haken_W/article/794/

https://mathworld.wolfram.com/HeawoodFour-ColorGraph.html

https://en.citizendium.org/wiki/Four_color_theorem

https://mathshistory.st-andrews.ac.uk/Biographies/Kempe/

https://en.wikipedia.org/wiki/Kempe_chain