Four Colour Theorem

Sebastian Wheeler

June 19, 2018

Abstract

This paper gives a brief overview of the Four Colour Theorem and a proof thereof. In 2005, Gonthier managed to use Coq to prove the theorem. This proof will be outlined here, explaining some steps in detail and also pointing out where it differs from the original incorrect proof by Kempe in 1879.

Contents

1 The Four Colour Theorem

In 1852, Francis Guthrie conjectured the Four Colour Theorem. The main topic of this paper is the Four Colour Theorem and the formal proof of the theorem done by Gonthier explained in [4]. First of all, recall the theorem:

Theorem (Four Colour Theorem) [4], p. 2
The regions of any simple planar map can be coloured with only four colours, in such a way that any two adjacent regions have different colours.

In order to understand the theorem, some more definitions are required:

Definition (Planar map, regions, simple map) [4], p. 3
A planar map *is a set of pairwise disjoint subsets of the plane, called* regions. *A* simple *map is one whose regions are connected open sets.*

Definition (Adjacent regions) [4], p. 3
Two regions of a map are adjacent *if their respective closures have a common point that is not a corner of the map.*

Definition (Corner) [4], p. 3
A point is a corner of a map if and only if it belongs to the closures of at least three regions.

With the help of these definitions the Four Colour Theorem can be understood. This easy to understand theorem is quite complicated to prove and it took several failed attempts and over 100 years until it was finally proven by Appel and Haken in 1976. The first (incorrect) proof was found by Kempe in 1879 and although it contains some mistakes, the main ideas of the proof were used in most of the subsequent proofs, also in the one presented here. (cf. [1], p.4 and [4], p. 5)

2 The proof

The main goal of Gonthier was to make sure that every detail of the proof is correct. Therefore, he used Coq v7.3.1 to check every part of the proof. In order to be able to use a computer to check the proof, the theorem has to be formulated in such a way that a computer can work with the underlying structures. Although most of the literature speaks about graphs when considering the Four Colour Theorem, Gonthier decides to stick to a combinatorial formulation of the problem using hypermaps. This is supposed to make it easier to implement in Coq. Consequently, the following structure is used:

Definition (polygonal outline, polygonal map, face) [4], p. 5
A polygonal outline *is the pairwise disjoint union of a finite number of open line segments, called edges, with a set of nodes that contains the endpoints of the edges. The regions of a* finite polygonal planar map*, called* faces*, are the connected components of the complement of a polygonal outline.*

A polygonal map is too general, as it allows structures that are not necessary for the Four Colour Theorem.

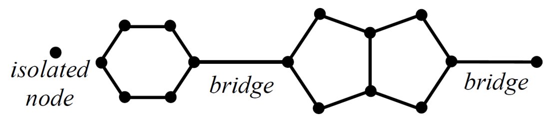

Figure 1: A general polygonal map ([4], p. 6)

When colouring a map, both isolated nodes and bridges do not make a difference, so they need not be considered. The exact definitions of an isolated node and a bridge are omitted, Figure 1 gives a good intuition. This fact gives rise to the next definition:

Definition (polyhedral map) [4], p. 6
A polyhedral map *is a finite bridgeless connected polygonal map.*

In his proof, Gonthier only considers polyhedral maps as this does not imply a loss of generality.

Another important theorem that is used in the proof is the well-known Euler formula:

Theorem (Euler formula) [6], p. 75
The Euler polyhedron formula (Euler 1752)

$$V - E + F = 2 \tag{1}$$

where V,E,F are the number of vertices, edges and faces, is valid for any schema which represents the sphere.

With all of this known, a brief overview of the proof can be given.

2.1 Structure of the proof

There are several parts of the idea of the proof. A brief overview will be given here. Most of these steps were already performed by Kempe in his incorrect proof. Only a few were added later on, so most of these steps can be found in [1] and [4].

2.1.1 Simplify the map

The first step is to simplify the map, i.e. identifying a special case that can be used but does not entail a loss of generality. Kempe and Cayley already knew that it suffices to consider only cubic maps, that is maps that only have vertices of degree 3. There is a simple way to see that if every cubic map is 4-colourable also every other map is 4-colourable:

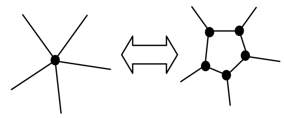

Figure 2: Reducing to cubic map ([4], p. 6)

If the cubic map is 4-colourable, the center face that was added can be removed to get back to the original map. The result is the map with a colouring with at most 4 colours. By removing the face in the middle, it cannot happen that a valid colouring becomes invalid.

Now that only cubic maps have to be considered for the rest of the proof, the Euler formula can be looked at again. Obviously, for cubic maps it holds that $3V = 2E$. Putting this in to (1) yields

$$2E = 6F - 12. \tag{2}$$

Every edge belongs to exactly two faces, so the total number of sides of all the faces is exactly $2E$. Dividing (2) by F leads to

$$\frac{2E}{F} = 6 - \frac{12}{F} \Leftrightarrow \frac{\#sides}{F} = 6 - \frac{12}{F}.$$

The left-hand side can be read as number of sides per face, showing that, on average, every face has just under 6 sides.

2.1.2 Consider a minimal counterexample

The next step in the proof is to consider a minimal counterexample, so a cubic map that is not 4-colourable and with a minimal number of faces. If a map that needs at least 5 colours exists, then there also must be one with a minimal number of faces.

This counterexample is a cubic map, so as shown in 2.1.1 the faces have less than six sides on average, so there must be at least one face that has five sides or less. Figure 3 shows the possible shape and neighbourhood of that face.

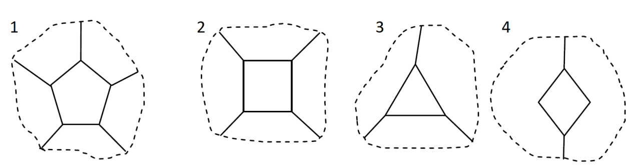

Figure 3: Possible shape and neighbourhood of the smaller face ([4], p. 7)

In the next section, it will be shown that all of these maps are 4-colourable and that therefore there cannot be a counterexample, so the Four Colour Theorem must be valid. It is important to remember that a *minimal* counterexample was considered, i.e. any map with less faces is 4-colourable.

2.1.3 Prove that the "counterexample" is 4-colourable

This section sets out to prove that the minimal counterexample is actually 4-colourable by removing edges from the central face and getting a smaller, thus 4-colourable map. The smaller 4 -colouring will be used to construct a 4-colouring of the original map. One of the cases of Figure 3 will occur in the example. The hope is that for every case, a 4-colouring can be found:

Cases 3 and 4 For both of these cases, remove any side of the central face. As this removes a face, the resulting map has to be 4-colourable (recall that a *minimal* counterexample was considered). This means that the central face is bordered by 3 (case 3) or 2 (case 4) colours, both cases leaving a fourth colour for the central face. This proves that the counterexample cannot have either of these structures as a part of it.

Case 2: square In the case of the square, the same can be done when choosing two appropriate sides to remove. When removing two opposite sides of the square, the faces touching those sides can either be the same, be adjacent or neither. If they are adjacent or the same, then the faces touching the other two sides of the square can be neither the same nor adjacent, as one of those faces will be surrounded by the two previous faces. So it is always possible to remove two edges such that the touching faces are neither the same nor adjacent. Figure 4 illustrates this fact.

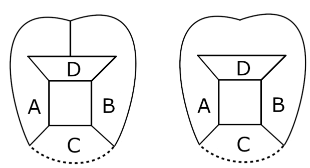

Figure 4: Possible situations of the square

A and B can either be adjacent (left side of Figure 4), the same (right side of Figure 4) or neither. In the first two cases, it can easily be seen that C and D are not adjacent, in the third case A and B are not adjacent, so there are always two faces that are not adjacent. The edges of the square touching these faces are the edges that are removed. This smaller map is definitely 4-colourable (as a *minimal* counterexample was considered). As two sides were removed, there are only three faces in the smaller map, so there is a fourth colour that can be used to colour the square.

Case 1: pentagon The last remaining case is the pentagon. As with the square, two sides are removed such that the faces touching those sides are neither the same nor adjacent. The reduced map is obviously 4-colourable again. The problem that remains is to prove that this colouring can be adapted to the original map.

Kempe thought he found a way to do this by using "Kempe chains" that give rise to a recolouring that leads to a valid 4-colouring. Kempe chains are chains of faces with alternating colours. Obviously, the colouring stays valid, if all of the colours of a longest possible chain are switched. An example of a Kempe chain is shown in Figure 5:

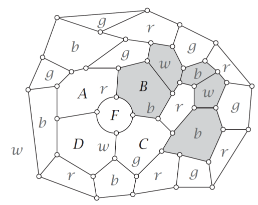

Figure 5: Kempe chain coloured in grey ([1], p. 11)

Kempe's idea was to find two Kempe chains that would split the colours of the map appropriately to allow a recolouring. The following image shows the general situation:

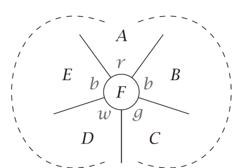

Figure 6: Kempe chains splitting the colours ([1], p. 11)

In Figure 6, let the dashed lines be two Kempe chains, one connecting A and D with a red-white chain and the other connecting A and C with a red-green chain. If one of these chains does not exist, simply taking the longest part of that chain starting in A and switching the colours will remove the colour red from the direct neighbourhood of the face F, so a colour for F has been found. So let both of these chains exist. Then the blue-white chain starting in B cannot possibly contain D, as it would have to go through the red-green chain from A to C. Similarly, the blue-green chain starting in E cannot contain C. Switching the colours of both of these chains then allows F to be coloured blue. This was Kempe's idea which at first sight seems to be correct.

However, he did not bear in mind that after switching the first set of colours, the second chain might change. This mistake was discovered in 1890 by Heawood when he provided a counterexample to Kempe's proof.

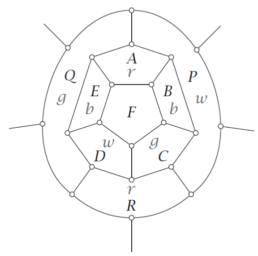

Figure 7: Counterexample to Kempe's proof ([1], p. 17)

If the same steps as above are performed on the example in Figure 7, the colouring is no longer valid. The reason is that as soon as the blue-white chain starting in B switches colours, the red-white chain connection A and D no longer exists and therefore C and E can be contained in the same blue-green chain. Although this ruins Kempe's proof, the main idea still can be used for a correct proof.

2.1.4 Correct Kempe's mistake

Kempe's problem was that he did not consider all possible neighbourhoods of the central pentagon in question. The solution is to consider even bigger neighbourhoods of the pentagon, enumerate all relevant possibilities and prove that all of them are 4-colourable.

The task is to find a certain set of configurations that are:

- *reducible*, meaning that it can be shown that they are 4-colourable and
- *unavoidable*, meaning that one of these configurations has to appear in every counterexample considered.

In his paper from 1913 ([2]), Birkhoff provided some good insights on how to compute reducibility. Basically, it was applying Kempe's idea to more general configurations. The bigger problem was the unavoidability.

In [5], it was proved that every possible minimal counterexample would have to contain one of 633 listed configurations. As a second step, they proved that each of these configurations is reducible, meaning that there also is a map that disproves the Four Colour Theorem but is smaller. This is a contradiction to the assumption that it was a minimal counterexample so in total, there cannot be a counterexample. However, a computer was used to check the 633 configurations which leads to scepticism among mathematicians and is a reason why some do not accept the proof.

Gonthier used the same 633 configurations to complete his computer-checked proof. The difference is that he uses Coq, which checks the proof for validity. Mathematicians therefore no longer have to trust the computer program, only Coq and as Coq is used more generally for all sorts of theorems, this is not as much of a problem (although some people surely still will not believe it). The last section of this paper will focus on some methods used to prove unavoidability.

2.2 Unavoidability

The problem of unavoidability was quite difficult to solve. So far, this paper has shown that the pentagon is unavoidable, but this does not help as it has not yet been proven that it is reducible. The task is to create a larger, more complicated set of configurations that is unavoidable but such that all of the configurations are reducible. In 1969, Heesch published a systematic method for determining unavoidability, called *discharging*. Following [1], this method is explained now:

At the beginning, every face gets an initial charge

$$i(f) = 6 - \#sides(f).$$

This means that every pentagon gets a charge of 1, every hexagon a charge 0, every heptagon a charge -1 and so on. Using (2) it can easily be seen that

$$\sum_{f \in F} i(f) = 6|F| - 2|E| = 6|F| - 6|F| + 12 = 12$$

where F is the set of all faces and E is the set of all edges.

As the total charge is always > 0, there must always be local areas with a positive charge. These local areas are unavoidable configurations. If none of these configurations were present, then all local charges would be ≤ 0 leading to a total charge ≤ 0 which is a contradiction. So a method has been found to find unavoidable configurations.

Next, a set of *discharge rules* has to be created that transfer charges between faces but keep the total global charge the same. Choosing appropriate discharge rules is the key to finding an unavoidable and reducible set. The simplest possible discharge rule is:

Discharge Rule 1 *Do nothing.*

This rule leaves all charges as they are, so the only areas with positive charge are the pentagons, so all that is known is that a pentagon is an unavoidable configuration. Another possible discharge rule is the following:

Discharge Rule 2 *Transfer* $\frac{1}{5}$ *from every pentagon to every adjacent face with more than six sides.*

The resulting charge r can now be inspected. If a pentagon f_5 still has a charge $r(f_5) > 0$ then it cannot have five adjacent faces with each more than six sides as each of them would reduce the pentagon's charge by $\frac{1}{5}$. If this is the case, then one of the neighbours of the pentagon either has to be another pentagon or a hexagon.

Next, have a look at the other faces. As hexagons neither receive nor lose charge, they will not have charge > 0. A heptagon f_7 has initial charge $i(f_7) = -1$, so the resulting charge is $r(f_7) = -1 + \frac{m}{5}$ where m is the number of adjacent pentagons. For $r(f_7) > 0$, m has to be at least 6. If a heptagon has six adjacent pentagons, two of those pentagons will be adjacent, so the pentagon-pentagon configurations appears again.

For any face with $k \geq 8$ sides, $i(f_k) = 6 - k$, so

$$r(f_k) = (6-k) + \frac{m}{5} \leq (6-k) + \frac{k}{5} = \frac{30 - 5k + k}{5} \leq \frac{-2}{5} < 0$$

where m again is the number of adjacent pentagons. This means that the only two cases that can possibly give positive charge have been analysed and in both cases, certain structures have been found. These configurations are therefore an unavoidable set:

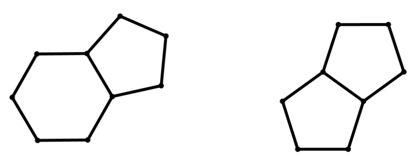

Figure 8: First unavoidable set

Obviously, it is good that an unavoidable set has been found. However, these configurations are not known to be reducible, so it does not solve the problem. Therefore, other steps with additional discharge rules can be performed in the hope to find a better set.

Discharge Rule 3 *Transfer $\frac{1}{4}$ from every pentagon to up to 4 adjacent faces with more than six sides.*

Again, positive resulting charge r will lead to an unavoidable set. For $r(f_5) > 0$, there must be at most three neighbours with more than six sides. This also means that there must be two neighbours with five or six sides.

For the heptagons, the formula is $r(f_7) = -1 + \frac{m}{4}$ where m again is the number of adjacent pentagons.

$$r(f_7) > 0 \Leftrightarrow m \geq 5$$

This means that there are five pentagons around the heptagon which implies that three of the pentagons must be adjacent (chain of three pentagons).

For $k \geq 8$ sides, the formula reads

$$r(f_k) = (6 - k) + \frac{m}{4} \leq \frac{24 - 3k}{4} \leq 0.$$

So the only configurations with positive charge are the ones previously found, leading to another, larger unavoidable set:

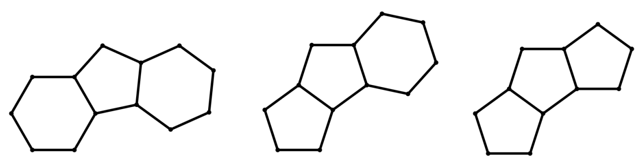

Figure 9: Second, larger unavoidable set

As demonstrated, certain discharge rules lead to certain unavoidable sets.

Appel and Haken managed to create an unavoidable and reducible set containing 1936 configurations. For reducibility, they only used Kempe's original idea simply applied to these larger configurations. Obviously, these computations cannot be done by hand but have to be done by computers which leads to disbelief of the proof. In [5], the unavoidable set was reduced to 633 configurations by using 32 special discharge rules. In his proof, Gonthier used the same unavoidable set.

2.3 Conclusion of the proof

The last step of the proof is to verify that all of the 633 configurations of the unavoidable set are actually reducible. Reducibility is the most computation-intensive part of the proof. All of the 633 configurations have to be proved to be 4-colourable. The idea of reducing is to remove some edges of the configuration leading to a smaller graph. This smaller graph has to be 4-colourable as a *minimal* counterexample was considered. The goal is to find a colouring of the original graph by using the colouring of the smaller graph.

As this part of the proof is a lot of enumeration, it would be better if there were less cases to consider. It turns out that it is possible to solve a different problem containing less cases. Instead of colouring the faces using four colours, it is sufficient to colour the edges using three colours. This obviously reduces the necessary computations/enumerations significantly.

Unfortunately, proving this fact and explaining how this can be done would go beyond the scope of this paper. Further information can be found in [4] and [1]. Here, it is sufficient to know that it can be done and formally verified.

Finding such a reducible and unavoidable set was the last step to concluding the proof. As Gonthier performed the entire proof and all of the computations within Coq, every single step was formally verified by the computer. And although some people might still mistrust the proof, it is widely accepted as correct.

All in all, it has taken quite a long time until finally a proof has been created that is accepted by most people. It seems unlikely that a proof will be found that doesn't rely on the help of computers for the enumeration parts. Of course, those steps could also be done by hand but humans are more likely to make mistakes and more importantly, it would take a person several years to perform all the checks. It also took Gonthier and his team several years until they had all of their code written, but Coq can verify this formal proof within a few days, possibly even a few hours with today's computing power.

References

[1] Aigner, Martin (2015): Graphentheorie. eine Einführung aus dem 4-Farben Problem. Springer Spektrum Wiesbaden.

[2] Birkhoff, George David (1913): The Reducibility of Maps (American Journal of Mathematics 35, pp. 115-128). The John Hopkins University Press.

[3] Fritsch, Rudolf and Fritsch, Gerda (1994): Der Vierfarbensatz. Geschichte, topologische Grundlagen und Beweisidee. Mannheim: B.I. - Wissenschaftsverlag.

[4] Gonthier, Georges (2005): A Computer-Checked Proof of the Four Colour Theorem.

[5] Robertson, Neil and Sanders, Daniel and Seymour, Paul and Thomas, Robin (1997): The Four-Colour Theorem (Journal of Combinatorial Theory, Series B 70).

[6] Stillwell, John (1980): Classical Topology and Combinatorial Group Theory (Graduate Texts in Mathematics 72). Springer-Verlag New York.