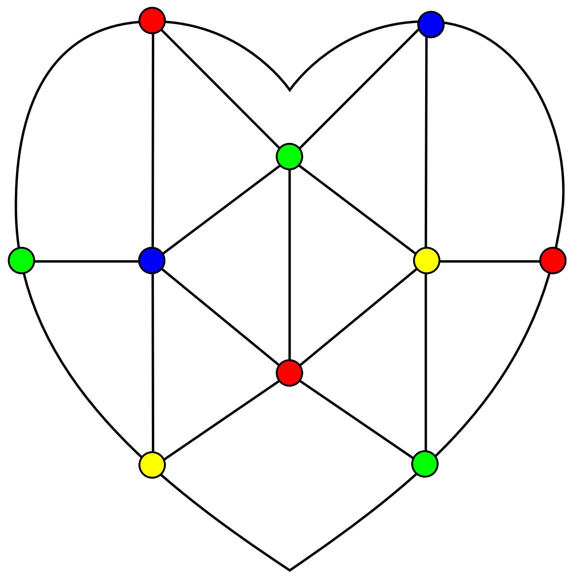


Heart of the Four Color Theorem

Timothy van der Valk

January 2025, Tohoku University

The four color theorem is a famous result from graph theory that has resisted proof for well over a 100 years. An application of this theorem states that any world map can be colored with four colors in such a way, that two neighboring regions receive different colors. When tasked with coloring a map, the solution lies in the idea of breaking the map into smaller pieces that are easy to color. This idea is called *reducibility*. This paper gives a more intuitive explanation of the three forms of reducibility used in the proof of the four color theorem.

We have explained how maps with regions arranged in a *ring* can broken up into smaller maps. This idea was first introduced by Birkhoff in 1913 [1]. We have rewritten his proofs for the reducing of rings of 4 and 5 regions. For rings of 6 regions and above, we have introduced D-reducibility with the Birkhoff Diamond as example. This is where the use of Kempe-chains is streamlined. We improved upon the problems of D-reducibility by introducing C-reducibility with the Bernhart Diamond as example. This form uses reducers to avoid bad colorings of the configuration.

We built up all this theory by leaning on the five color theorem for inspiration. We have put the problem of proving the four color theorem in perspective of several simple concepts, such that the intuition behind the proof as a whole can be better understood. Many figures and examples have been given to this end.

Title image - The Birkhoff diamond shaped like a heart (original).

Summary

The four color theorem states that every map can be colored with four colors in such a way, that two neighboring regions receive different colors. Such a coloring is desired for a world map, because it becomes easy to tell two neighboring regions apart. It has been observed by many map makers that four colors suffice. The problem was first formulated by Francis Guthrie in 1852 while coloring the map of England. He brought the problem to his brother Frederick Guthrie, who in turn brought the problem to his mathematics lecturer Augustus De Morgan.

At the heart of the famous proof of the four color theorem by Appel and Haken in 1976 [4] is shown that any map contains an arrangment of regions that can be removed and recolored later. These are called *reducible configurations*. The four and five color theorem both used the theory of reducible configurations. There are three forms of reducibility of a configuration used in these proofs.

1. *k-reducibility* shows that configurations with a boundary ring on less than 6 regions can be reduced under certain conditions. This is done by proving that the interior and exterior can be independently colored such that the colors on the bordering ring are the same. This justifies removing the configuration from the map and adding it back after coloring the smaller map.

2. *D-reducibility* builds upon k-reducibility for configurations on rings of 6 regions and more. This technique excels at proving reducibility for individual configurations. A configuration is D-reducible if every ring coloring can be fixed to become a valid ring coloring of the configuration. *Kempe-chains* are used to reconfigure invalid ring colorings to valid ones. The Birkhoff Diamond is the most famous example for D-reducibility.

3. *C-reducibility* improves upon D-reducibility in case there are unfixable ring colorings. It avoids these unfixable colorings by replacing the configuration with a smaller map called a *reducer*. The ring colorings of the reducer put constraints on the ring colorings of the configuration. These contraints are selected such that the constrained colorings are all fixable. This avoids the unfixable colorings, making the configuration reducible.

The Bernhart Diamond is C-reducible with a valid reducer. However, two of the constrained colorings from the reducer are unfixable. This turns out to be a flaw in either D-reducibility, or the implementation of D-reducibility by John. P. Steinberger [7]. These two colorings depend on another problem coloring called a *symmetry fault*. A symmetry fault is a coloring that is unfixable, but whose symmetry *is* fixable. Therefore, both symmetries must be fixable. We have not uncovered the cause of these faults, but Bernhart [2] proved that in case of the Bernhart Diamond, they are not a problem.

The first proof of the four color theorem had a set of 1478 configurations to check. An improvement was made by Neil Robertson et al. [6] in 1996 who reduced the set to only 633 configurations that are either C or D-reducible. Lastly, an improvement was made by John P. Steinberger in 2009 who used only D-reducibility at the cost of having to check 2822 configurations. It seems that using more advanced forms of reducibility reduces the amount of configurations on ring 6 and above. This hints that there might be an all-encompassing form of reducibility that uses the least number of configurations, which would be *The Heart of the Four Color Theorem*.

Contents

1 Introduction

The four color theorem was initially formulated from a problem in coloring world maps. A map consists of regions that can border other regions on a flat surface. When we talk about a *coloring* of a map, we mean a way to color its regions such that any two neighbors are colored differently.

The actual shape of the regions in our map is not of importance here. The key information that is needed from a map, is the connectivity between regions. Such information can be represented in a *graph* where vertices (circles) correspond to regions. An edge between two vertices then indicates that the two corresponding regions are neighbors.

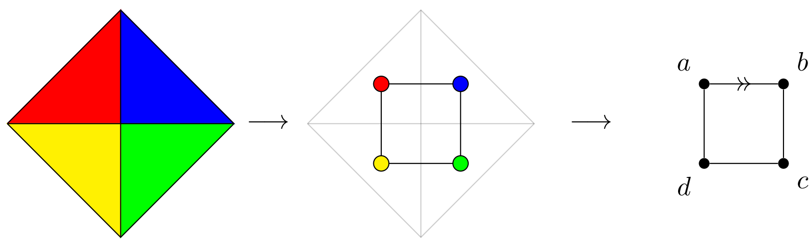


Figure 1: The translation of a map coloring to a graph coloring. In the last step we replace colors by the letters a,b,c and d for convenience. We obtain the coloring called *abcd* on a *planar graph*.

We will be working exclusively with *planar graphs*. See Figure 2 for examples. The order of colors such as in *abcd* is indicated by the $\gg$ edge, which is the first edge between the first two vertices. Two colorings are considered *equal* if they differ only by a renaming of colors.

Definition 1. *A graph is* planar *if it has a* planar embedding *where no two edges cross each other.*

Definition 2. *Two colorings x and y are* equal *if they differ only by a renaming of the colors a, b, c and d. I.e $abab = acac$.*

Theorem 1. *Every planar graph can be colored in at most four colors.*

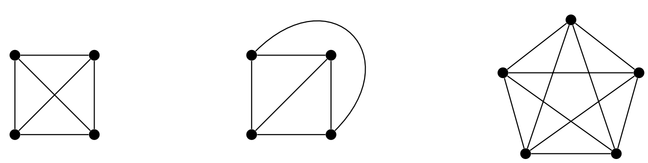


Figure 2: A non-planar embedding of the full graph K_4 on 4 vertices (left). A planar embedding of K_4 (middle), therefore K_4 is a planar graph. A non-planar graph K_5 that has no planar embeddings (right).

1.1 The five color theorem

The proof of this four color theorem required over a 100 years to complete, despite its simple statement. What would you do to prove such a statement? A first step to the proof is to work with an easier problem instead. This is the five color theorem.

Theorem 2. *Every planar graph can be colored in at most five colors.*

Proof. Given a planar graph G. Because G is planar, a known result in graph theory is that G has a vertex with at most five edges. That is, there is a v such that $deg(v) \leq 5$. If we can always free up a color for v regardless of the colors of its neighbors, then we may simply ignore v for now and color the smaller graph $G-v$ first. By repeating the same argument on $G-v$, we will eventually be left with just a single vertex. From there we can build up the 5-coloring of our graph. Therefore, let us show that we can always free up a color for v. We consider two cases for $\deg(v) \leq 5$.

- $\deg(v) \leq 4$. In this case, our vertex has at most four neighbors. These four neighbors have at most four different colors. This means that one color is free to be used for v.
- $\deg(v) = 5$. In this case, we have exactly five neighbors. Should our neighbors require only four colors, we can simply use our fifth color here. However, it might occur that all five neighbors use all five colors. Now we must try to free up a color in these neighbors.

Indeed, to treat the case $\deg(v) = 5$ we should make a sketch of the situation first.

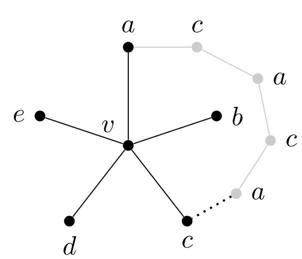


Figure 3: The vertex v when having five differently colored neighbors. In gray, an example ac-chain starting from the neighbor a.

Now we must ask ourselves, do these neighbors really need to use all five colors? We may flip all the vertices colored a or c that are connected to the top neighbor a to change its color to c. This way we freed up the color a. We say that we flipped the *ac-chain* of neighbor a. Such a chain can be seen in Figure 1.1.

However, if the neighbor c is part of this ac-chain (represented by a dotted line), then it will get flipped to a. So in this case we have not freed up the color a. We did however, isolate the vertex b with the ac-chain. It is now impossible for a bd-chain from b to d to exist. Therefore we can flip the bd-chain of b to change its color into d without affecting the neighbor d. This frees up the color b.

Therefore, all cases show that a color can be freed up for v. By our earlier argument the graph is 5-colorable. □

1.2 Fundaments of the four color theorem

The proof of the four color theorem follows the same structure as the five color theorem. We show that every planar graph contains a subgraph that allows us to reduce the coloring problem to a smaller graph. With *smaller* we mean less vertices.

Definition 3. *The size $|G|$ of a graph is the number of vertices of G.*

Definition 4. *A graph H is smaller than G or $H < G$ if $|H| < |G|$.*

This notion of a *subgraph* of a graph requires some extra attention, since we will need a stronger notion of being a subgraph called *containment*. See Figure 4. The graphs $\mathcal{C}$ contained in others that we use for reducibility purposes, are called *configurations*.

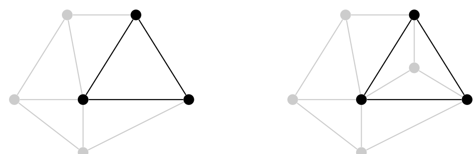


Figure 4: A configuration contained in a graph (left). No containment (right).

Definition 5. *A planar graph $\mathcal{C}$ is* contained *in a graph G, if $G \setminus \mathcal{C}$ (removed $\mathcal{C}$) is connected.*

Definition 6. *A planar graph $\mathcal{C}$ is* reducible *in a graph G, if $\mathcal{C}$ being contained in G implies that there exists graphs $H < G$ called* reductions *whose 4-colorings can be used to 4-color G.*

Therefore, in a reducibility proof, we may assume that any graph H smaller than G is 4-colorable. Then, we show that a 4-coloring of H can be extended to a 4-coloring of G. Using these two definitions we can formulate the key theorem of the four color theorem. We proved the same result for the five color theorem.

Theorem 3. *Every planar graph G contains a configuration $\mathcal{C}$ that is either k-reducible, D-reducible or C-reducible in G.*

From this theorem, a 4-coloring of a planar graph G_0 can be found as follows. This is the same procedure that we used for the vertex of $\deg(v) \leq 5$ in the five color theorem.

1. Find a reducible configuration C_n in G_n.
2. Reduce the graph G_n to the smaller graph G_{n+1}.
3. If G_{n+1} is the empty graph, color all the intermediate graphs starting from G_n all the way until G_0. Else, repeat Step 1 on G_{n+1}.

There are many ways to prove that a configuration is reducible. We will be treating the three central forms of reducibility that are used in the four color theorem. Each of them allows us to test the reducibility of certain configurations. These forms of reducibility are by no means perfect, each has its flaws and uses. A single, most general definition of reducibility is still not found. If such a single form would exist, then it would capture the heart of the four color problem.

2 k-Reducibility

We have seen in the proof of the five color theorem that a vertex surrounded by five or less neighbors can always be colored using one of five colors, even if all of its neighbors initially use all five colors. If we have only four colors available, we will find that it is no longer guaranteed that we can free a color. This is what Alfred Kempe tried to do when he gave the first false proof of the four color theorem.

If we look at the key idea, we see that if one half of a graph is isolated from another half by a group of boundary vertices, we can color this isolated part regardless of the colors on the boundary. Naturally, the fundamental shape that seperates a graph in two halves is a *ring.*

Definition 7. *A* ring *of n vertices R_n in a planar graph G is an induced cycle of G.*

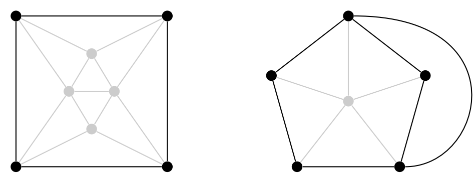


Figure 5: An example of the ring R_4 surrounding a graph on four vertices (left). An example of an invalid ring R_5 (right).

The vertex of ring R_1 can not have a single edge to itself, instead, it acts as a bridge from one part of a graph to the other. For rings present in graphs, we are interested in which colorings are possible on them, called *ring colorings.* The set of these colorings for a graph G we shall define. It turns out that the rings $n \geq 4$ are already reducible configurations. Therefore, all future configurations will only contain triangles (the ring R_3), they will be *triangulations.*

Definition 8. *The set of all* valid *4-colorings of a ring R in a planar graph G is given by $\Phi(R \subset G)$ or $\Phi(G)$ if R is clear from the context.*

Definition 9. *The set of all valid ring colorings of R_n is given by $\Phi(n) = \Phi(R_n)$.*

Theorem 4. *The ring R_n with $n \geq 4$ is reducible in every planar graph G.*

Proof. Let R_n be contained in G. Since the interior of R_n is empty and $n \geq 4$, we may contract the two non-neighboring ring vertices v_1, v_3 to a new vertex u. As a result, we obtain the graph G' on one less vertex. See Figure 6.

Given a 4-coloring of G'. Because R_n is a ring, there will be no edges between the ring vertices v_1 and v_3. Therefore, we may give v_1 and v_3 the same color as u without issue. Let the other vertices of G be given the same color as their G' counterparts. Then we have obtained a 4-coloring of G.

□

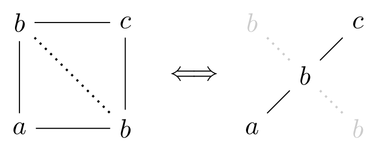


Figure 6: The ring R_4 being contracted to a smaller graph G'. The coloring of G' can be reversed to a coloring of G.

2.1 Ring configurations

We have now seen that rings with nothing on their interior are already reducible. The natural next step would be to consider rings with an interior $\mathcal{K}$ called a *core*. These will be called *ring configurations*, since they are configurations with a ring as their boundary. All configurations used in the four color theorem are ring configurations!

Definition 10. *A planar graph* $\mathcal{C} = R_n + \mathcal{K}$ *consisting of a ring* R_n *and a* core $\mathcal{K}$ *is called a* ring configuration *on* R_n.

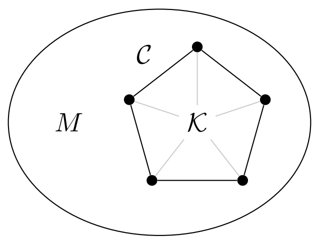


Figure 7: A graph $M + \mathcal{C}$ with a ring configuration $\mathcal{C}$ on R_5.

If a ring configuration $\mathcal{C}$ on R is contained in a graph G, then the ring R acts as border between the interior $\mathcal{K}$ and exterior M. Suppose that we removed from G the interior $\mathcal{K}$ of $\mathcal{C}$ to obtain the graph $M + R$. If we colored $M + R$ and $\mathcal{C}$ individually, then we could add the interior $\mathcal{K}$ back if the two colorings have the same colors on the ring R. This is the essence of k-reducibility. We prove that such a *common ring coloring* can always be found. Mathematically, we phrase this as

$$\Phi(\mathcal{C}) \cap \Phi(M + R) \quad \neq \quad \varnothing. \tag{1}$$

I.e, the sets of possible ring colorings of R in $\mathcal{C}$ and $M + R$ have a common element. Given a configuration $\mathcal{C}$, we can question whether such a common element exists for all graphs $M + \mathcal{C}$. To answer this question, we must know to *some* degree which colorings exist in $\Phi(M + R)$. Such colorings we call *guaranteed colorings*. They are all we can work with if the graph M is arbitrary.

2.2 Kempe-chains

In the five color theorem we worked with a vertex that had 5 neighbors. We used chains of 2 colors between its neighbors to *flip* the colors of these neighbors, creating a new coloring. If we return to the problem of finding a common ring coloring for a ring configuration $\mathcal{C}$ on R and the graph $M + R$, we do not know much of the ring colorings in $\Phi(M + R)$. All we know, are the guaranteed colorings from contracting the ring R. To obtain more guaranteed colorings, we can use the same trick of flipping the colors of vertices on the same chain. Therefore, let us define *Kempe-chains*.

Definition 11. *Let $G_{ab}(x)$ be the subgraph consisting of all the vertices colored ab in the coloring x of G. Then the* Kempe-chain $\kappa_{ab}(v)$ *or* ab-chain *of the vertex v is the component of $G_{ab}(x)$ that contains v.*

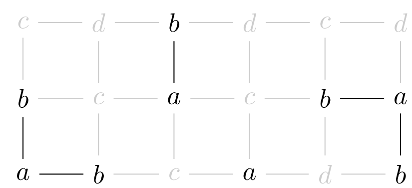


Figure 8: The components of $G_{ab}(x)$ for a planar graph G and its coloring x are highlighted. We write $u \overset{ab}{\frown} v$ or $u \in \kappa_{ab}(v)$ if u and v are on the same component.

We can flip the colors of a chain $\kappa_{ab}(v)$ without breaking the current coloring. Imagine in your head how you can swap the chains in Figure 8 for example. This is the trick to creating new guaranteed ring colorings in $\Phi(M + R)$.

Now, suppose that we are guaranteed that the coloring *abab* is in $\Phi(M + R)$. If we want to flip any Kempe-chains on *abab*, then it is necessary to have information on them. This information is not obvious from the coloring *abab* itself. To add this information, we may consider two cases, one in which a chain is present, and another in which it is not. To visualise the Kempe-chains that are present on a coloring like *abab*, we may draw lines between vertices like

$$\overset{d}{\widehat{a\ b\ a}}\ b\,, \qquad a\ \overset{c}{\widehat{b\ a\ b}}\,, \qquad \overset{\not c}{\widehat{a\ b\ a}}\ b\,. \tag{2}$$

From this notation, it is visible at glance what the structure is of the ring coloring *abab*. A ring coloring together with knowledge of Kempe-chains we call a *scheme*. In the left-most scheme for example, we see $v_1 \overset{ad}{\frown} v_3$. In the middle scheme, we see $v_2 \overset{bc}{\frown} v_4$ and in the right-most scheme, we see that v_1 and v_3 are not connected by an ac-chain.

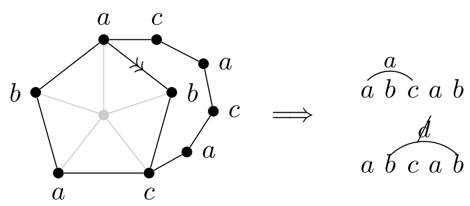
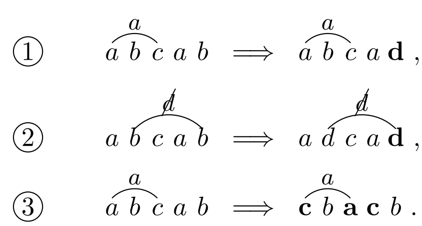


Figure 9: Two ring schemes that are derived from a graph coloring.

Definition 12. *Given a coloring x of a planar graph G and the colors on its ring $x(R)$. The* scheme *on R of x consists of $x(R)$ with knowledge whether $u \in \kappa_{ab}(v)$ for two ring vertices $u, v \in R$ and colors ab.*

We will often make a case dinstinction on the Kempe-chains in the scheme of a ring coloring. We consider a case in which the scheme has the chain, and a case in which it has not. We rarely ever know *all* the Kempe-chains of a scheme, but this is also never necessary. Given the information of just one Kempe-chain in a scheme, we can start *reconfiguring* the colors of the ring by flipping Kempe-chains. This results in a new scheme with different ring colors. We say that one scheme *implies* another scheme.

Definition 13. *Given two schemes x and y. We say that $x \implies y$ if $x = y$ or y can be obtained from x by flipping a Kempe-chain.*



For the two schemes from Figure 9, we have given two implied schemes above (change is highlighted). To deduce these implied colorings, we followed a few simple rules.

1. The two vertices colored b are seperated by the ac-chain $\kappa_{ac}(v_1)$. Therefore, the bd-chain $\kappa_{bd}(v_5)$ can not possibly connect with v_2. This results in only v_5 being recolored to d if we flip this chain.

2. It is now directly given that $\kappa_{bd}(v_5)$ does not connect to v_1, therefore we may flip v_5 to d in the same way as in the first case.

3. Now we know v_1 and v_3 are in the same ac-chain. Because v_4 neighbors v_3 on the ring, it will also be included in this chain. Since all three are on $\kappa_{ac}(v_1)$, flipping this chain results in all their colors to be flipped.

Refer to these examples whenever a step of implying colorings in a proof is confusing. It will be a key concept for the remainder of the paper.

2.3 Reducibility of configurations on R_4

Using just the power of schemes and Kempe-chains, we can already tackle the reducibility of any configuration on R_4. This proof serves as an excellent introduction into applying schemes and Kempe-chains, just like in the five color theorem.

Theorem 5. *Every configuration $\mathcal{C}$ on R_4 is reducible in every planar graph.*

Proof. Let the planar graph $M + \mathcal{C}$ and configuration $\mathcal{C} = \mathcal{K} + R_4$ be arbitrary. Let us denote the sets of ring colorings of the reductions $M + R$ and $\mathcal{C}$ by I and II respectively. We will refer to a coloring in one of these sets by prepending the set name, i.e I($abab$) means coloring $abab$ in set I. The situation is sketched in Figure 10.

$$\mathrm{I} = \Phi(M + R_4) \quad \text{and} \quad \mathrm{II} = \Phi(\mathcal{C}). \tag{3}$$

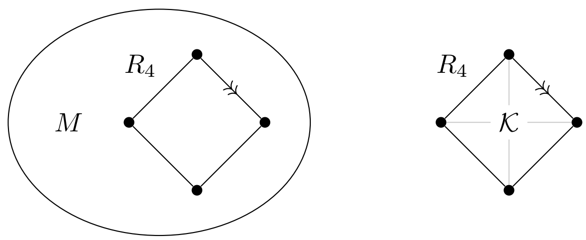


Figure 10: The reductions $M + R_4$ and $\mathcal{K} + R_4$ ($\mathcal{C}$).

Both reductions contain the ring R_4. Therefore, by Theorem 4, we may further reduce these graphs by contracting any two opposing vertices of R_4. Because there are only 2 ways to contract non-neighboring vertices of R_4, we will obtain two guaranteed colorings in I and II. In each of them, the contracted vertices are colored the same.

$$\begin{Bmatrix} abab \text{ or } abac, \\ baba \text{ or } baca \end{Bmatrix} \subset I, II. \tag{4}$$

First note that $abab = baba$ are the same coloring. The possibilities result in a total of 4 different sets of guaranteed colorings for I and II.

$$①= \{abab\}, \quad ② = \begin{Bmatrix} abab \\ baca \end{Bmatrix}, \quad ③ = \begin{Bmatrix} abac \\ baba \end{Bmatrix}, \quad ④ = \begin{Bmatrix} abac \\ baca \end{Bmatrix}. \tag{5}$$

All pairs of sets except ① and ④ already have a common coloring. Therefore, let $\{abab\} \subset$ I and $\{abac, baca\} \subset$ II. Now, we make a case distinction whether the chain $v_1 \overset{ad}{\frown} v_3$ exists in I($abab$) or not.

$$
\begin{aligned}
\mathrm{I}(abab) &= \overset{d}{\overbrace{a\ b\ a}}\ b \implies \mathrm{II}(abac) \\
\mathrm{I}(abab) &= \overset{\not d}{\overbrace{a\ b\ a}}\ b \implies \mathrm{I}(abdb) = \mathrm{II}(baca).
\end{aligned}
\tag{6}
$$

In any case, we obtain a common ring coloring between I and II. Therefore the ring R_4 is 0-reducible. □

2.4 Reducers to constrain ring colorings

As we have seen in the reducibility proof of R_4, we had enough guaranteed colorings to find a common ring coloring through the reconfiguring of schemes. However, as we will see in the next section, we will need guaranteed 3-colorings like *cabab* and *abcab* to prove that configurations on R_5 are reducible. A 3-coloring is not guaranteed from simply contracting ring vertices from Theorem 4. Therefore, we seek a technique to obtain more guaranteed colorings such as 3-colorings.

However, such a technique will come at cost. The cost is that we will no longer be able to reduce to $M + R$, and must instead reduce to a slightly bigger graph $M + S$ where $R \subset S$. The graph S is called a *reducer*. It is chosen in such a way that it guarantees the colorings we need in $\Phi(M + S)$. The reducer that we will use to guarantee 3-colorings of R_5 is shown in Figure 11.

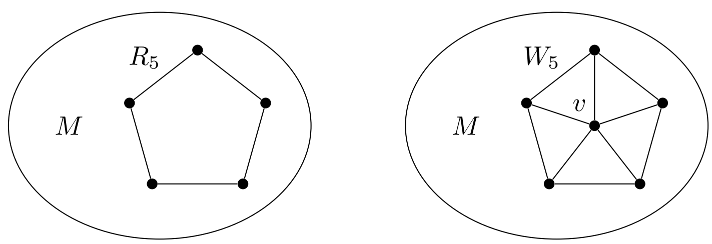


Figure 11: The graph $M + R_5$ (left). The graph $M + W_5$ that guarantees 3-colorings of the R_5 (right).

The reduction we used for R_4 is similar to the left-most graph in Figure 11. This graph allows us to contract two vertices of the R_5 to force them to be the same color. Since there are 5 ways to contract two non-neighboring vertices of R_5, we are guaranteed all of the following 5 colorings.

$$\Phi^\star(5) = \{a{*}{*}a{*}, \quad {*}a{*}{*}a, \quad a{*}a{*}{*}, \quad {*}a{*}a{*}, \quad {*}{*}a{*}a\}. \tag{7}$$

The $*$-colors in these colorings are still unknown. This set $\Phi^\star(5)$ contains all the guaranteed colorings from contracting vertices of the ring R_5. Every ring has such a set. For the ring reducibility proof of R_4 for example, we used the following set of 2 colorings.

$$\Phi^\star(4) = \{a{*}a{*},\ {*}a{*}a\}. \tag{8}$$

Definition 14. *The set $\Phi^\star(n)$ consists of all guaranteed $*$-colorings of R_n obtained from successive contraction of non-neighboring vertices.*

Imagine how larger rings like R_6 are able to contract two opposing vertices twice in a row. This guarantees that 2 pairs of vertices are colored the same. We wont be using these sets for rings beyond R_5, but it should not be difficult to define exactly all the $*$-colorings for every ring R_n.

To continue with our introduction of reducers, let us consider now the colorings of the right-most graph $M + W_5$ in Figure 11. The graph $S = W_5$ is called the *wheel graph* on R_5. It is clear that this graph is 1 vertex bigger than $M + R_5$, hence making it a weaker reduction. However, in exchange we are guaranteed *only one* of the following five 3-colorings.

$$\Phi^{W_5} = \{\underline{c}abab \;\vee\; a\underline{c}bab \;\vee\; ab\underline{c}ab \;\vee\; aba\underline{c}b \;\vee\; abab\underline{c}\}. \tag{9}$$

We have underlined the uniquely colored vertex of each 3-coloring. From the usage of a logical-OR symbol ($\vee$), we will have in fact multiple options for the set Φ^{W_5}. In general, we will use the notation Φ^S to indicate *some* set of guaranteed ring colorings of a reducer S, even if there are multiple options for this set.

Definition 15. *The set Φ^S is* some *set of guaranteed ring colorings of a reducer S.*

From the examples of the sets $\Phi^\star(5)$ and Φ^{W_5}, it is clear that the use of a reducer guarantees more types of ring colorings. These guarantees can then be used to find common ring colorings of two graphs, through the reconfiguration of Kempe-chains. Therefore, let the reduction for M be given by

$$M + S \text{ with } R_n \subset S \text{ and } |S| \leq k + n. \tag{10}$$

The k in the maximum size of our reducer $|S| \leq k + n$ indicates how many other vertices we have in S besides the ring R_n. This parameter determines the amount of *control* we have over the ring colorings.

- If $k = 0$ we can take R_4 as example. In this case we had $S = R_4$ and we reduced to $M + R_4$. This is the smallest possible reduction. We could only use guaranteed colorings in $\Phi^\star(4)$.
- If $k = 1$, we can take R_5 as example. Here we will set $S = W_5$ such that we reduce to $M + W_5$ instead. This gives us a guaranteed 3-coloring from Φ^{W_5}. However, the graph $M + W_5$ would only be smaller than $M + \mathcal{C}$ if $\mathcal{C} > W_5$. This means that there must be at least 2 vertices in the interior of $\mathcal{C}$. This limits which configurations on R_5 are reducible.
- If $k \geq |\mathcal{C}|$, then there is never a point in reducing, since we always obtain the same graph $M + \mathcal{C}$ by setting $S = \mathcal{C}$ or a larger one.

Therefore, we must require that $k < \mathcal{C}$ in order for $\mathcal{C}$ to be reducible. With this, all the components to define k-reducibility of a configuration $\mathcal{C}$ are in place.

Definition 16. *A configuration $\mathcal{C}$ on R_n is k-*reducible *for $k < |\mathcal{C}|$, if for all planar graphs M there exists a reducer with $|S| \leq k$ such that*

$$\Phi(M+S) \cap \Phi(\mathcal{C}) \quad \neq \quad \varnothing. \tag{11}$$

We have already proven that configurations on R_4 are 0-reducible. Configurations on R_1, R_2 and R_3 are also trivial examples of 0-reducibility. Their only colorings are of the form a, ab and abc respectively, therefore, we have $\Phi(M + R_n) = \Phi(\mathcal{C})$ for each of them. We have hinted a non-trivial example in our motivation for reducers. This is where we prove the 1-reducibility of configurations on R_5.

2.5 Reducibility of configurations on R_5

Recall that every planar graph has a vertex with $\deg(v) \leq 5$. If we add edges between the neighbors of this vertex, then we obtain the rings R_1 thru R_5. We have seen the 0-reducibility of configurations on the first four. Therefore, if we could prove that every configuration on R_5 is 0-reducible, any planar graph would be reducible and the four color theorem would follow. Many people have tried to show this and failed, so let this serve as an warning as to why we should prove 1-reducibility instead.

Theorem 6. *A configuration $\mathcal{C}$ on R_5 is 1-reducible in all planar graphs M if it has a 3-coloring, or all planar graphs with $|M| > 1$ if it does not.*

Proof. We may consider a configuration with $|\mathcal{C}| > 1$. Let the planar graph M be arbitary. We will again use the convention of the sets I and II for the ring colorings.

$$\mathrm{I} = \Phi(M+S) \quad \text{and} \quad \mathrm{II} = \Phi(\mathcal{C}). \tag{12}$$

The heart of this proof depends on a guaranteed 3-coloring in both I and II. Because we may set our reducer $S = W_5$ (See Figure 11), we are guaranteed of the following colorings.

$$\Phi^{\star}(5) \subset \mathrm{I}, \mathrm{II} \quad \text{and} \quad \Phi^{W_5} \subset \mathrm{I}, \mathrm{II}. \tag{13}$$

To justify that $\Phi^{W_5} \subset \mathrm{II}$ in case a 3-coloring is not given by $\mathcal{C}$ from the start, let us add a single vertex v to the configuration $\mathcal{C}$ as illustrated in Figure 12. This results in the graph $\mathcal{C}'$ on 1 more vertex. We may assume that $\mathcal{C}'$ has a 4-coloring because $\mathcal{C}' < M + \mathcal{C}$ from our assumption that $|M| > 1$. In particular, $\mathcal{C}'$ will have a 3-coloring on the ring from W_5. We may remove v in this 3-coloring to obtain a 3-coloring of $\mathcal{C}$. Therefore $\Phi^{W_5} \subset \Phi(\mathcal{C})$.

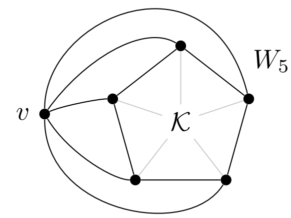


Figure 12: The modified configuration $\mathcal{C}'$ used to guarantee a 3-coloring of $\mathcal{C}$.

A special property of 3-colorings on R_5 is that there will always be a single vertex colored uniquely, we call this the *marked vertex*. This vertex acts as a kind of 'pivot' to tell if two 3-colorings are *adjacent*. These two concepts are key to the proof.

Definition 17. *The uniquely-colored vertex of a 3-coloring of R_5 is called the* marked vertex*, indicated by an underline such as in $\underline{c}abab$.*

Definition 18. *Two 3-colorings of R_5 are called* adjacent *if they have adjacent marked vertices, such as in $\underline{c}abab$ and $a\underline{c}bab$.*

Since we have guaranteed 3-colorings for both I and II, we can split up our proof into 3 cases.

1. I and II have an adjacent coloring ($\underline{c}abab$ and $a\underline{c}bab$).
2. I and II have a non-adjacent coloring ($\underline{c}abab$ and $ab\underline{c}ab$).
3. I and II have a coloring with the same marked vertex ($\underline{c}abab$ and $\underline{d}cbcb$). These are already equal, so we are done.

Therefore, we only need to consider the cases ① and ②. We will prove two lemma's to this end. Their relation is illustrated below.

$$\begin{array}{llll} \text{②} \implies \text{①} & \text{or} \quad \text{common coloring} & & \text{(Lemma 1)} \\ \qquad\quad \downarrow & & & \\ \qquad\quad \text{①} \implies & \text{common coloring} & & \text{(Lemma 2)} \end{array}$$

Lemma 1. *If I and II have a non-adjacent coloring, then they either have an adjacent coloring or a common coloring.*

Proof. Let two non-adjacent colorings $\mathrm{I}(\underline{c}abab)$ and $\mathrm{II}(ab\underline{c}ab)$ be given. We make a case dinstinction whether the chain $v_3 \overset{bc}{\frown} v_5$ exists in $\mathrm{II}(ab\underline{c}ab)$ or not.

$$
\begin{aligned}
\mathrm{II}(ab\underline{c}ab) &= \ a\ b\ \overset{b}{\widehat{c\ a\ b}} \implies \mathrm{II}(abcdb),\\
\mathrm{II}(ab\underline{c}ab) &= \ a\ b\ \overset{\not b}{\widehat{c\ a\ b}} \implies \mathrm{II}(a\underline{c}bab).
\end{aligned}
\tag{14}
$$

The second case results in a coloring adjacent to $\mathrm{I}(\underline{c}abab)$. For the first case, we consider the coloring $\mathrm{I}(*b{*}{*}b)$. The two adjacent $*$-colors must be different from each other and b, therefore we may assume that we have $\mathrm{I}(*bcdb)$. The last $*$-color reveals 3 possibilities.

$$
\begin{array}{lcl}
\mathrm{I}(abcdb) & = & \mathrm{II}(abcdb) \text{ from (14)},\\
\mathrm{I}(cbc\underline{d}b) & \text{adjacent to} & \mathrm{II}(ab\underline{c}ab),\\
\mathrm{I}(db\underline{c}db) & = & \mathrm{II}(ab\underline{c}ab).
\end{array}
\tag{15}
$$

Therefore we obtain either a common coloring or an adjacent coloring. Note that the pair of non-adjacent colorings we chose is the only unique one on R_5 (up to rotational symmetry). □

Lemma 2. *If I and II have an adjacent coloring, then they have a common coloring.*

Proof. Let two adjacent colorings $\mathrm{I}(\underline{c}abab)$ and $\mathrm{II}(a\underline{c}bab)$ be given. We make a case dinstinction whether the chain $v_3 \overset{bd}{\frown} v_5$ exists in $\mathrm{II}(a\underline{c}bab)$ or not.

$$
\begin{aligned}
\mathrm{II}(a\underline{c}bab) &= \ a\ c\ \overset{d}{\widehat{b\ a\ b}} \implies \mathrm{I}(\underline{c}abab).\\
\mathrm{II}(a\underline{c}bab) &= \ a\ c\ \overset{\not d}{\widehat{b\ a\ b}} \implies \mathrm{II}(acdab).
\end{aligned}
\tag{16}
$$

The first case leads to a common coloring. For the second case, we consider the coloring $\mathrm{I}(a{*}{*}a{*})$. We may again assume to have $\mathrm{I}(acda*)$. Then the 3 remaining possibilities for the $*$-color are

$$
\begin{array}{lcll}
\mathrm{I}(acdab) & = & & \mathrm{II}(acdab), \text{ from (16)},\\
\mathrm{I}(ac\underline{d}ac) & = & \text{shifted } +2 & \mathrm{I}(\underline{c}abab),\\
\mathrm{I}(a\underline{c}dad) & = & & \mathrm{II}(a\underline{c}bab).
\end{array}
$$

If we do not obtain common colorings, we may repeat this procedure with $\mathrm{II}(a\underline{c}bab)$ and $\mathrm{I}(ac\underline{d}ac)$ to continously shift the marked vertex two to the right. The pattern that arises is illustrated in Figure 13.

	v_1	v_2	v_3	v_4	v_5	v_1
1	I	II				I
2		II	I			
3			I	II		
4				II	I	
5					I	II

Figure 13: Marked vertices of I and II obtained by repetition of the earlier procedure. Every step shifts the leftmost marked vertex two to the right.

At iteration 5, we obtain that II has the same marked vertex v_1 as I. Therefore, by repetition we obtain that

$$\mathrm{II}(a\underline{c}bab) \to \mathrm{II}(aba\underline{c}b) \to \mathrm{II}(\underline{c}abab) = \mathrm{I}(\underline{c}abab). \tag{17}$$

□

Therefore, adjacent colorings imply a common coloring for I and II. Combining these two lemmas yields a guaranteed common ring coloring for I and II regardless of the case.

□

3 D-Reducibility

We have seen the reducibility of configurations on R_4 and R_5. Naturally, we might try to investigate the reducibility of configurations on R_6 and beyond. However, as we have seen in the increase in complexity for the 1-reducibility on R_5, this will likely be a very tough problem. The difficulty of this problem lies in the fact the we try to prove the reducibility of many configurations on R_n *at once*. However, it is much easier to analyze individual configurations. This why special forms of reducibility were created to analyse such individual configurations. Among which D-reducibility, which is inspired by k-reducibility.

3.1 Definition with the Birkhoff Diamond

We will be working with an example, the *Birkhoff Diamond* (Bir$\Diamond$), which is a configuration on R_6 with 4 vertices in the core. See Figure 14.

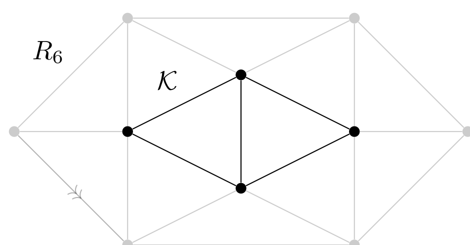


Figure 14: The Birkhoff Diamond $\mathcal{C}$ = Bir$\Diamond$ with the core highlighted.

An individual configuration on a ring is fully defined by its core $\mathcal{K}$ and the amount of edges each vertex of $\mathcal{K}$ has with the ring R. The Birkhoff Diamond is uniquely determined by the four vertices in the middle (its core) and the requirement that each vertex of $\mathcal{K}$ has a total of 5 edges in $\mathcal{C}$. For vertices in the core, the distinction between various heads of a vertex indicates this degree in $\mathcal{C}$. This is a convention set by Heesch [3] in 1969. Four such vertex-heads are given below.

$$\begin{array}{ccc} \bullet & \longrightarrow & \deg(v) = 5 \\ \cdot & \longrightarrow & \deg(v) = 6 \\ \circ & \longrightarrow & \deg(v) = 7 \\ \square & \longrightarrow & \deg(v) = 8 \end{array}$$

We will show that Bir$\Diamond$ is 0-reducible by a careful investigation of all the ring colorings of Bir$\Diamond$ and R_6. Since we know exactly what the graph is of Bir$\Diamond$, we can write down all its valid ring colorings. We have provided all the ring colorings of R_6 and Φ(Bir$\Diamond$) in Figure 15.

Φ(6)			
ababab	abacbd	abcadc	abcdab
ababac	abacdb	abcbab	abcdac
ababcb	abacdc	abcbac	abcdad
ababcd	abcabc	abcbad	abcdbc
abacab	abcabd	abcbcb	abcdbd
abacac	abcacb	abcbcd	abcdcb
abacad	abcacd	abcbdb	abcdcd
abacbc	abcadb	abcbdc	
31			

Figure 15: The set Φ(6) and the valid colorings Φ(Bir♢) highlighted.

To show that Bir♢ is 0-reducible, we must show that all those invalid ring colorings can be fixed using Kempe-chains to become valid. If we let $M + R_6$ be arbitrary, then we can expect any ring coloring of R_6. Let us consider the coloring *ababab* for example. Suppose that $v_4 \overset{bd}{\frown} v_6$. This implies the following colorings.

$$\begin{aligned} a\ b\ a\ \overset{d}{\widehat{b\ a\ b}} &\implies ababcb \\ a\ b\ a\ \overset{\not d}{\widehat{b\ a\ b}} &\implies ababad = ababac. \end{aligned} \tag{18}$$

Because the invalid coloring *ababab* can be reconfigured into a valid coloring with only one Kempe-chain flip, we say that the coloring *ababab* *implies* the set of colorings

$$ababab \implies \{ababcb, ababac\}. \tag{19}$$

Definition 19. *For a coloring x, we have $x \implies II$ if for every scheme $x^\star$ of x, we have $x^\star \implies y \in II$.*

Definition 20. *For two set of colorings I and II, we $I \implies II$ if for every $x \in I$, we have $x \implies II$.*

The set of all colorings of R_6 that require one Kempe-chain flip to be reconfigured in the same way as *ababab* is called the 1-implying set of Bir♢.

$$\Phi_1(\text{Bir}\lozenge) = \ \Phi_0(\text{Bir}\lozenge) \cup \{ababab,\ ababcd,\ abacab,\ abcbcb,\ abcdad\}. \tag{20}$$

This set is the largest that satisfies $\Phi_1(\text{Bir}\lozenge) \implies \Phi_0(\text{Bir}\lozenge) = \Phi(\text{Bir}\lozenge)$. We may repeat what we did for $\Phi_1(\text{Bir}\lozenge)$ to obtain sets of colorings that require 2, 3 and more Kempe-chain flips to become a valid coloring. This higher-order implication between sets of colorings is called *n-implication.*

$$\Phi_5(\text{Bir}\lozenge) \implies \Phi_4(\text{Bir}\lozenge) \implies \Phi_3(\text{Bir}\lozenge) \implies \ldots \implies \Phi(\text{Bir}\lozenge). \tag{21}$$

Definition 21. *For two sets of colorings I and II, we have $I \overset{n}{\Longrightarrow} II$ if there exist sets B_i for $0 < i < n$ such that*

$$I \Longrightarrow B_{n-1}, \quad B_i \Longrightarrow B_{i-1} \quad \text{and} \quad B_1 \Longrightarrow II. \tag{22}$$

Definition 22. *The n-implying set $\Phi_n(\mathcal{C})$ of a configuration $\mathcal{C}$ is the largest set of ring colorings such that $\Phi_n(C) \overset{n}{\Longrightarrow} \Phi_0(\mathcal{C}) = \Phi(\mathcal{C})$.*

The set $\Phi_5(\text{Bir}\Diamond)$ satisfies $\Phi_5(\text{Bir}\Diamond) \overset{5}{\Longrightarrow} \Phi_0(\text{Bir}\Diamond)$. Let us find all the n-implying sets of Bir$\Diamond$. We say that colorings in these sets are *fixable*, because they can always be reconfigured to a valid coloring with Kempe-chains. For D-reducibility, we want all ring colorings to be fixable. All n-implying sets together must cover all ring colorings. This is precisely the case with the Birkhoff Diamond.

$\Phi_0(\text{Bir}\Diamond)$		Φ_1	Φ_2	Φ_3	Φ_4	Φ_5
ababac	abcadb	ababab	abacad	abacbd	abcabd	abcabc
ababcb	abcbab	ababcd	abcbdb	abcbdc	abcadc	
abacac	abcbac	abacab		abcdac	abcdbc	
abacbc	abcbad	abcbcb		abcdbd		
abacdb	abcbcd	abcdad				
abacdc	abcdab					
abcacb	abcdcb					
abcacd	abcdcd					
16		5	2	4	3	1

Figure 16: All n-implying sets of Bir$\Diamond$. The largest covers all 31 ring colorings of R_6.

Definition 23. *The* max-implying *set $\overline{\Phi}(\mathcal{C})$ of a configuration $\mathcal{C}$ is the largest n-implying set $\Phi_n(\mathcal{C})$.*

Definition 24. *A coloring $x \in \overline{\Phi}(\mathcal{C})$ is called a* fixable *ring coloring of $\mathcal{C}$.*

Definition 25. *A configuration $\mathcal{C}$ on R_n is D-reducible if $\overline{\Phi}(\mathcal{C}) = \Phi(n)$.*

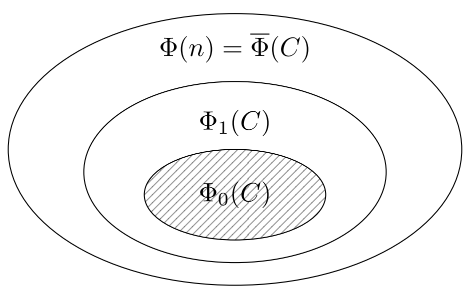


Figure 17: Every ring coloring is fixable for a D-reducible configuration.

4 C-Reducibility

The definition of D-reducibility required that every ring coloring was reconfigurable to a valid coloring of $\overline{\Phi}(\mathcal{C})$. If this is not possible, we must avoid the unfixable colorings. In Section 2.4 we introduced reducers to guarantee certain ring colorings for use in proofs. These reducers constrain which ring colorings can be encountered on the ring of the original configuration. Therefore, we may apply them in the same manner for individual configurations to avoid unfixable colorings in D-reducibility. This results in what is known as C-reducibility.

4.1 Definition with Bernhart's Diamond

We will be working with an example as we develop this theory. This example is a relative of the Birkhoff Diamond (Bir$\Diamond$), called the Bernhart Diamond (Ber$\Diamond$). First proven to be reducible by Arthur Frederick Bernhart in 1947 [2].

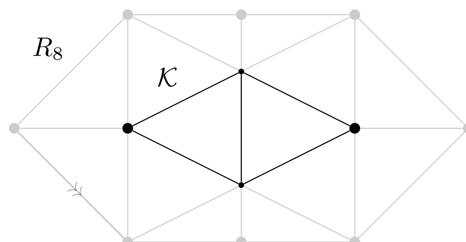


Figure 18: The Bernhart Diamond $\mathcal{C} = \text{Bir}\Diamond$ with the core highlighted.

The core of Ber$\Diamond$ as highlighted in Figure 18 can be found as a configuration in a list by Frank Allaire et al. [5]. The two vertices with a small dot ($\cdot$) require 6 edges in $\mathcal{C}$, as opposed to 5 in Bir$\Diamond$. This is their only difference. We will follow the same procedure as we did for Bir$\Diamond$ by analyzing all compatible colorings. The ring of size 8 has 274 unique colorings, in addition, the max-implying set $\overline{\Phi}(\text{Ber}\Diamond)$ has 92 colorings. Therefore, Ber$\Diamond$ is not D-reducible. See the green and blue colored cells in Figure 19.

$\Phi(8)$					
abababab	abacadac	abcababc	abcadbab	abcbcbdb	abcdadbc
abababac	abacadad	abcababd	abcadbac	abcbcbdc	abcdadbd
abababcb	abacadbc	abcabacb	abcadbad	abcbcdab	abcdadcb
abababcd	abacadbd	abcabacd	abcadbcb	abcbcdac	abcdadcd
ababacab	abacadcb	abcabadb	abcadbcd	abcbcdad	abcdbabc
ababacac	abacadcd	abcabadc	abcadbdb	abcbcdbc	abcdbabd
ababacad	abacbabc	abcabcab	abcadbdc	abcbcdbd	abcdbacb
ababacbc	abacbabd	abcabcac	abcadcab	abcbcdcb	abcdbacd
ababacbd	abacbacb	abcabcad	abcadcac	abcbcdcd	abcdbadb
ababacdb	abacbacd	abcabcbc	abcadcad	abcbdabc	abcdbadc
ababacdc	abacbadb	abcabcbd	abcadcbc	abcbdabd	abcdbcab
ababcabc	abacbadc	abcabcdb	abcadcbd	abcbdacb	abcdbcac
ababcabd	abacbcab	abcabcdc	abcadcdb	abcbdacd	abcdbcad
ababcacb	abacbcac	abcabdab	abcadcdc	abcbdadb	abcdbcbc
ababcacd	abacbcad	abcabdac	abcbabab	abcbdadc	abcdbcbd
ababcadb	abacbcbc	abcabdad	abcbabac	abcbdbab	abcdbcdb
ababcadc	abacbcbd	abcabdbc	abcbabad	abcbdbac	abcdbcdc
ababcbab	abacbcdb	abcabdbd	abcbabcb	abcbdbad	abcdbdab
ababcbac	abacbcdc	abcabdcb	abcbabcd	abcbdbcb	abcdbdac
ababcbad	abacbdab	abcabdcd	abcbabdb	abcbdbcd	abcdbdad
ababcbcb	abacbdac	abcacabc	abcbabdc	abcbdbdb	abcdbdbc
ababcbcd	abacbdad	abcacabd	abcbacab	abcbdbdc	abcdbdbd
ababcbdb	abacbdbc	abcacacb	abcbacac	abcbdcab	abcdbdcb
ababcbdc	abacbdbd	abcacacd	abcbacad	abcbdcac	abcdbdcd
ababcdab	abacbdcb	abcacadb	abcbacbc	abcbdcad	abcdcabc
ababcdac	abacbdcd	abcacadc	abcbacbd	abcbdcbc	abcdcabd
ababcdad	abacdabc	abcacbab	abcbacdb	abcbdcbd	abcdcacb
ababcdbc	abacdabd	abcacbac	abcbacdc	abcbdcdb	abcdcacd
ababcdbd	abacdacb	abcacbad	abcbadab	abcbdcdc	abcdcadb
ababcdcb	abacdacd	abcacbcb	abcbadac	abcdabab	abcdcadc
ababcdcd	abacdadb	abcacbcd	abcbadad	abcdabac	abcdcbab
abacabab	abacdadc	abcacbdb	abcbadbc	abcdabad	abcdcbac
abacabac	abacdbab	abcacbdc	abcbadbd	abcdabcb	abcdcbad
abacabad	abacdbac	abcacdab	abcbadcb	abcdabcd	abcdcbcb
abacabcb	abacdbad	abcacdac	abcbadcd	abcdabdb	abcdcbcd
abacabcd	abacdbcb	abcacdad	abcbcabc	abcdabdc	abcdcbdb
abacabdb	abacdbcd	abcacdbc	abcbcabd	abcdacab	abcdcbdc
abacabdc	abacdbdb	abcacdbd	abcbcacb	abcdacac	abcdcdab
abacacab	abacdbdc	abcacdcb	abcbcacd	abcdacad	abcdcdac
abacacac	abacdcab	abcacdcd	abcbcadb	abcdacbc	abcdcdad
abacacad	abacdcac	abcadabc	abcbcadc	abcdacbd	abcdcdbc
abacacbc	abacdcad	abcadabd	abcbcbab	abcdacdb	abcdcdbd
abacacbd	abacdcbc	abcadacb	abcbcbac	abcdacdc	abcdcdcb
abacacdb	abacdcbd	abcadacd	abcbcbad	abcdadab	abcdcdcd
abacacdc	abacdcdb	abcadadb	abcbcbcb	abcdadac	
abacadab	abacdcdc	abcadadc	abcbcbcd	abcdadad	
274					

Figure 19: Ring colorings of R_8. Green: The max-implying set of Ber◊. Underlined green: Valid ring colorings of Ber◊. Blue: Fixable reducer-constrained colorings. Red: Unfixable reducer-constrained colorings. Black: Symmetry faults that *should* be fixable.

The blue cells indicate reducer-constrained colorings of Ber$\Diamond$. All of them (except for the red ones) are blue and therefore fixable. The way that these constraints are created is a through a process we can control. We created a reducer as follows.

1. Contract using a mapping σ, any non-neighboring vertices of the ring R to a single point. Two contracted vertices will be colored the same. The result is the contracted ring $\sigma(R)$.
2. Turn the contracted ring $\sigma(R)$ into the graph S by including extra edges and vertices on the interior.

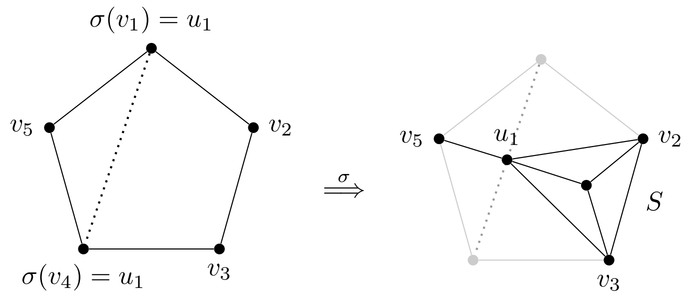


Figure 20: First, the two vertices v_1 and v_4 are contracted to the same vertex y_1, then we turn the contracted graph into S.

A contraction maps the vertices of R to the vertices of $\sigma(R)$. Those vertices that are mapped to the same vertex are contracted. This mapping allows us to directly convert a ring coloring $x(u)$ of S to a ring coloring $x(\sigma(v))$ of R. Note that the reducers we used for k-reducibility used the identity contraction $\sigma(v) = v$, which does not contract any edges.

Definition 26. *A ring contraction $\sigma(v)$ on R is a map from the vertices of $R \mapsto \sigma(R)$. Neighboring vertices of R may not be mapped to the same vertex.*

Definition 27. *$\sigma_{id}(v) = v$ is the* identity contraction.

Definition 28. *A reducer (S, σ) of a configuration $\mathcal{C}$ consists of a contraction $\sigma(v)$ on R and a graph $S < \mathcal{C}$ whose boundary is the contracted ring $\sigma(R)$.*

Definition 29. *The set of reducer-constrained ring colorings $\Phi(S, \sigma)$ consists of all the colorings $x(\sigma(v))$ with $x(u)$ a boundary coloring of S.*

The reducer for Ber$\Diamond$ is given in Figure 21. It consists of two contractions and a single extra edge added by S in the middle. The mapping of colorings from S to R_8 are given below. It is helpful to think of a reducer solely as a set of constraints.

$$v_1\,\mathbf{u_2}\,v_3\,v_5\,\mathbf{u_1}\,v_7 \quad \mapsto \quad v_1\,\mathbf{u_2}\,v_3\,\mathbf{u_2}\,v_5\,\mathbf{u_1}\,v_7\,\mathbf{u_1}. \tag{23}$$

In the next section we will show that the red generated colorings in Figure 19 are actually in $\overline{\Phi}(\text{Ber}\Diamond)$, therefore $\Phi(S, \sigma) \subset \overline{\Phi}(\text{Ber}\Diamond)$. As a result, we may replace Ber$\Diamond$ by its reducer S and be guaranteed that any coloring for S can be reconfigured ot a valid coloring. This is the essence of C-reducibility.

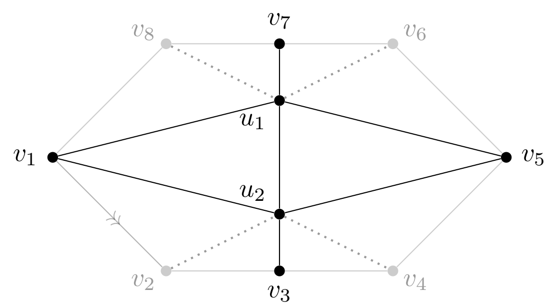


Figure 21: The reducer S of the Bernhart Diamond Bir$\Diamond$. It forces that $v_2 = v_4 \neq v_6 = v_8$. The contractions force the same colors on pairs of vertices, while the added edges force different colors.

Definition 30. *A configuration $\mathcal{C}$ is C-reducible if $\Phi(S, \sigma) \subset \overline{\Phi}(\mathcal{C})$ for some reducer (S, σ).*

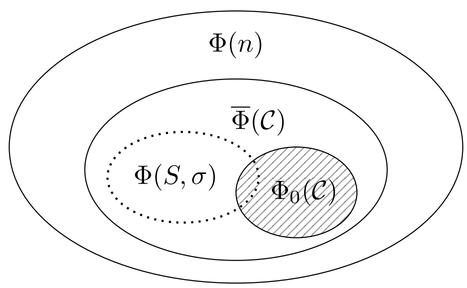


Figure 22: The max-implying set includes all generated colorings of $\Phi(S, \sigma)$.

For any given configuration, there are only a finite number of reducers (S, σ) possible. Therefore, it is feasible to simply test all possible reducers for the condition $\Phi(S, \sigma) \subset \overline{\Phi}(\mathcal{C})$.

4.2 Symmetry faults

If we come back to the red colorings from Figure 19, they are the two colorings we call *faults* R1 and R2. We will be working with the following colorings.

$$\begin{aligned} &abcbacbc && (R1) \\ &abcbdcbc && (R2) \\ &abcbadbc && (F1) \\ &abcbacbd && (F1\star) \in \overline{\Phi}(\text{Ber}\Diamond) \\ &abcbacac && (G1) \in \overline{\Phi}(\text{Ber}\Diamond) \\ &abcbdbcb && (G2) \in \overline{\Phi}(\text{Ber}\Diamond) \end{aligned} \tag{24}$$

It was proven in 1947 by Bernhart [2] that the colorings $R1$ and $R2$ are in fact fixable. This is because their fixability depends on the broken coloring $F1$ that should be fixable, this is one of the black colorings from Figure 19. We have the following implications.

$$\begin{aligned} R2 \implies \{&G2, R1\} \\ &\downarrow \\ &R1 \implies \{G1, F1\} \end{aligned} \tag{25}$$

Therefore, if $F1 \in \overline{\Phi}(\text{Ber}\Diamond)$ then $R1, R2 \in \overline{\Phi}(\text{Ber}\Diamond)$. The argument that Bernhart used was that the ring coloring $F1$ is symmetric to the coloring $F1\star \in \overline{\Phi}(\text{Ber}\Diamond)$, therefore it should also be fixable. The situation is sketched in Figure 23.

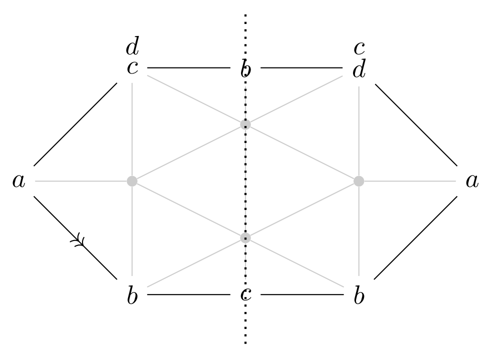


Figure 23: Horizontal symmetry of the colorings *abcbadbc* ($F1$) and *abcbacbd* ($F1\star$) above the other.

We have not delved deeper into this problem, and do not know whether this is a fundamental flaw of implied colorings or an error in our code from [7]. This proves that the Bernhart Diamond is in fact C-reducible, because the red colorings are incorrectly flagged so. The root of the problem being the colorings $F1$ and $F1\star$, we will call them *symmetry faults.* In Figure 23, the symmetry is a bijection between the vertices left and right of the symmetry line.

Definition 31. *A graph symmetry of G is a bijection $f(v)$ on the vertices of G that preserves the neighbors of every vertex.*

Definition 32. *A* symmetry fault *of a configuration $\mathcal{C}$ is a coloring x that is not in $\overline{\Phi}(\mathcal{C})$, but whose symmetry $x\star = f(x)$ is.*

A future study on the symmetry fault in the Bernhart Diamond could lead to insights to improve the definitions of C and D-reducibility. One would first verify if it is a computational error, and then try to find more examples. In Bernhart's 1947 paper [2], the four colorings $R1, R2, F1$ and $F1\star$ were also the only problem case, therefore it is unlikely to be a coincide.

4.3 Generalising D and k-reducibility

We have now seen all three forms of reducibility used in proofs of the four color theorem. You might have noticed many similarities between the different forms. For example, the key use of Kempe-chains, the reconfiguring of ring colorings and the use of reducers. Due to the way we built up this theory, we extended upon each of the shortcomings of one form to motivate the next. It might not come as a surprise then that C-reducibility is the most general of all forms. Let us first consider the relation between D-reducibility and k-reducibility.

Theorem 7. *D-reducibility implies 0-reducibility.*

Proof. Suppose that we have D-reducibility $\mathcal{C}$ on R_n. Clearly, we have that $\Phi(M + R_n) \subset \Phi(n) = \overline{\Phi}(\mathcal{C})$. Therefore, every coloring of $\Phi(M + R_n)$ can be reconfigured to a coloring of $\Phi(\mathcal{C})$. This reconfiguring means that at least one coloring of $\Phi(\mathcal{C})$ must exist in $\Phi(M + R_n)$. Therefore

$$\Phi(M + R_n) \quad \cap \quad \Phi(\mathcal{C}) \quad \neq \quad \varnothing. \tag{26}$$

□

Theorem 8. *C-reducibility with reducer (R_n, σ_{id}) is equivalent to D-reducibility.*

Proof. $\Longrightarrow$ Suppose that we have C-reducibility of $\mathcal{C}$ with reducer (R_n, σ_{id}). Then

$$\Phi(R_n, \sigma_{id}) = \Phi(n) \subset \overline{\Phi}(\mathcal{C}) \quad \Longleftrightarrow \quad \Phi(n) = \overline{\Phi}(\mathcal{C}). \tag{27}$$

$\Longleftarrow$ Now suppose that we have D-reducibility. Then we may set $S = R_n$ and $\sigma = \sigma_{id}$ to obtain the same result as in the above equation. □

Theorem 9. *C-reducibility with reducer* (S, σ_{id}) *on* R_n *implies* k*-reducibility with* $k = |S| - n$*.*

Proof. Suppose that we C-reducibility of $\mathcal{C}$ with (S, σ_{id}). Because $\Phi(M+S) \subset \Phi(S, \sigma_{id}) \subset \overline{\Phi}(\mathcal{C})$, at least some coloring of $\Phi(M+S)$ can be reconfigured to become valid. Therefore

$$\Phi(M+S) \ \cap \Phi(\mathcal{C}) \quad \neq \quad \varnothing. \tag{28}$$

In addition, our reducer S has $k = |S| - n$.

□

Note, that there reverse direction of k-reducibility implying D and C-reducibility is also possible, but has some problems and technicalities that are not covered by the definitions. To go from k-reducibility to D-reducibility for example, one must show that every ring coloring is reconfigurable to $\Phi(\mathcal{C})$. With only the information that there is a common coloring in $\Phi(M+R)$ and $\Phi(\mathcal{C})$, it is only possible to guarantee this reconfigurability if the set $\Phi(M+R)$ contains only this one coloring and all its Kempe-chain reconfigurations. However, do there exist graphs whose colorings are all derived from a single one? It is not known.

5 Conclusion

We have seen how the rings and k-reducibility came forth from the five color theorem. D-reducibility expanded upon k-reducibility for rings 6 and above. Then C-reducibility expanded upon D-reducibility by avoiding unfixable colorings. We have worked through these forms of reducibility with the Birkhoff Diamond and Bernhart Diamond as example. For the Bernhart Diamond, we encountered symmetry faults that require further attention.

The ideas of the four color theorem are generalisations of what we did for the five color theorem. Unavoidability of a reducible configuration plays a key role in both proofs. The theory of Kempe-chains came forth from the need to obtain more colorings from a single guaranteed coloring. For this, we had introduced some forms of notation to clarify our future arguments. The below Figure illustrates how using more advanced forms of reducibility results in less configurations of ring size 6 and above.

Author	Reducibility forms	Configurations
	Perfect reducibility	1
Neil Robertson et al.	C, D	633
Appel and Haken	A,B,C,D	1478
John P. Steinberger	D	2822
	No reducibility	∞

Figure 24: Number of reducible configurations on rings R_6 and higher in proofs of the four color theorem

.

It is expected that there should exist more advanced forms of reducibility that essential hone in on the "heart of the four color theorem". Just like how the vertices of $\deg(v) \leq 5$ were at the heart of the five color theorem. The computer part of the four color theorem was only needed to evaluate the trivial calculations of finding implying sets of configurations and reducers. The essence of the four color theorem does not lie in these calculations. In the end, reducibility lies at the heart of the four color theorem and the various implementation and calculation issues caused the proof to become so long and infamous.

References

[1] George David Birkhoff. “The Reducibility of Maps”. In: *American Journal of Mathematics* Vol. 35.2 (1913).

[2] Arthur Frederick Bernhart. “Another reducible edge configuration”. In: *American Mathematical Journal of Mathematics* (1947).

[3] H. Heesch. “Untersuchungen zum Vierfarbenproblem”. In: *Hochschulskriptum 810/a/b Bilbiographisches Institüt Mannheim* (1969).

[4] Kenneth Appel and Wolfgang Haken. “Every Planar Map is Four Colorable”. In: *Bulletin of the American Mathematical Society* Vol. 82.5 (1976).

[5] Edward Reinier Swart Frank Allaire. “A Systematic Approach to the Determination of Reducible Configurations in the Four-Color Conjecture”. In: *Journal of Combinatorial Theory* Series B 25, 339-362 (1976).

[6] Neil Robertson et al. “A New Proof of the Four-Color Theorem”. In: *Electronic Research Announcements American Mathematical Society* 2.1 (1996).

[7] John P. Steinberger. “An Unavoidable Set of D-reducible Configurations”. In: *arXiv* (2009).