SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

The Logic behind Kőnig's Lemma

av

Cordelia Mangia Jansson

2021 - No K19

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

The Logic behind Kőnig's Lemma

Cordelia Mangia Jansson

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Sven Raum

2021

Abstract

The goal of this paper is to uncover the complexity of Kőnig's Lemma and find a deeper understanding of how axioms and theorems are connected. In order to get an insight on the logic behind this lemma, different versions of the lemma are introduced, the relative strengths of them are compared and some applications are discovered. The role of mathematical logic is presented together with the rise of axiomatic systems such as Peano Axioms and two different set theories, namely "Zermelo-Fraenkel set theory" (ZF) and "Zermelo-Fraenkel set theory with Axiom of Choice" (ZFC). A comparison between the two set theories ZF and ZFC is given with a particular focus on one of the most discussed axioms in mathematics, namely the "Axiom of Choice". Some weaker versions of this axiom are then introduced and compared, in particular "Weak Axiom of choice for countable families of finite sets". Mathematical statements equivalent to Weak Kőnig's Lemma as well as Kőnig's Lemma are found and the axioms necessary to prove these equivalence relations are investigated. Finally, the axiomatic systems for which weak Kőnig's Lemma and Kőnig's Lemma exist are defined.

Keywords: *Kőnig's Lemma, weak Kőnig's Lemma, trees, sequences, sequential compactness, planar graphs, graph coloring, axiomatic systems, Axiom of Choice, weak Axiom of choice for countable families of finite sets, reverse mathematics.*

Acknowledgement

I would like to express my gratitude to my advisor, Sven Raum, for guiding me with contagious entusiasm through the process of this thesis. I am thankful for introducing me to a new perspective of mathematics.

I would also like to thank my parents and my grandmother, Armanda Gandini, for encouraging me along the road of becoming a mathematician.
Last, my gratitude goes to my partner, Emil Abrahamsson, for always being there for me and providing all of my study hours with a warm cups of tea.

Contents

Introduction

Imagine yourself underneath an infinitely big oak tree. Standing at it's roots, you can only see how the branches get thinner and thinner, but you can't see its end. This is the closest you've ever been with infinity.

What if you could reach it and find the answer to the questions that have made mathematicians and philosophers scratch their heads for centuries?

Fearless, you decide to defy infinity, and you start climbing up the tree. You reach the first branch and, while dangling your feet in the air, you're reminded of a theorem about infinite trees.

Theorem 1 (Kőnig's Lemma)**.** *Let T be an infinite (k-ary) tree. Then T contains an infinite path p.*

From this, you realize you have to choose wisely what way to climb up if you want to encounter infinity. You have to choose the right path, the infinite path. But how?

The core of this paper is Kőnig's Lemma, which helps us solve this puzzle.

Through formal mathematical analysis, we are going to uncover the complexity of Kőnig's Lemma, which at first glance seems simple.

In order to get an insight on the logic behind this lemma, we are going to examine diverse versions of the lemma, namely Weak Kőnig's Lemma and Kőnig's Lemma. We will compare the relative strengths and discover some applications.

We will then learn about logic within mathematics, the rise of axiomatic systems (such as Peano Axioms and the Zermelo-Fraenkel set theories with and without the axiom of choice) and more about which axiomatic systems are necessary to prove Kőnig's lemma and its applications. We are going to slice down the proofs until we find the axioms that they built upon. In particular, we are going to put focus on the theorems relation with one of the most discussed axioms in mathematics: the Axiom of Choice.

At the end of this paper, we will conclude by finding that some theorems which have been introduced as implications of Kőnig's Lemma are in fact equivalent to the lemma. We will define for which axiomatic systems these equivalences hold.

This paper aims not only to give a deep understanding of Kőnig's Lemma, but also to be a good guideline for examining any mathematical theorem.

Structure

In Chapter 1 we are going to give a general introduction to binary rooted trees which we will define as a set of sequences. We will also present some fundamental definitions about these mathematical structures in order to understand Kőnig's Lemma.

In Chapter 2 we are going to prove two versions of Weak Kőnig's Lemma. The first one will treat infinite binary rooted trees and the second one will treat infinite k-ary rooted trees. We then prove that the sequential compactness of $[0, 1]$ can be deduced from Weak Kőnig's Lemma.

At the end of the chapter we will introduce some graph theory and theory about graph coloring. We will see that an interesting property of infinite k-colourable planar graphs is nothing more than a consequence of Kőnig's Lemma.

In Chapter 3 we will introduce the foundations of logic, what axiomatic systems are, in particular Peano Axioms and two versions of Zermelo-Fraenkel set theory, namely ZF and ZFC. We will also examine the relationship of these different axiomatic systems. In order to do this, we will introduce the Axiom of Choice which is one of the most discussed axioms in mathematics, and some weaker forms of AC. At the end of this chapter we will prove the statement "the countable union of finite sets is countable" is equivalent to a weak version of the axiom of choice.

In Chapter 4 we will introduce reverse mathematics. Starting with a theorem, we will go backwards in order to discover which axiomatic systems are needed. We are also going to prove an equivalence between several statements and Kőnig's Lemma in the context of variate axiomatic systems. Also, we will summarize the results of the paper.

1 Theory of trees

In order to get a deep understanding of Kőnig's Lemma, we have to start by going through some fundamental definitions of trees. Generally, a tree is a mathematical structure that can be viewed as either a graph or a data structure. These mathematical structures are used in a lot of fields, such as: mathematics, data science, philosophy, operations research, chemistry, transport systems, logistics and more.

A universal idea of a what a rooted tree is, is given by Richard Kate in the book "Mathematics of logic".

"A tree is a diagram (often called a graph) with a special point or node called the root, and lines or edges leaving this node downwards to other nodes. These again may have edges leading to further nodes. The thing that makes this a tree (rather than a more general kind of graph) is that the edges all go downwards from the root, and that means the tree cannot have any loops or cycles."[1, page 1].

In this paper we want to compare different formulations of Kőnig's Lemma. In order to do this, we are going to give different definitions of trees. In particular, we are first going to look at a formalism of rooted trees given by numerical sequences. A consequence of this type of formalism is that the rooted trees can be coded and are therefore countable. Later on, in Section 2.3, we are going to look at a formalism of trees (without roots) and graphs which can not be coded as the ones mentioned above.

Since we first are going to define rooted trees as collections of sequences, we have to define what a sequence is.

Definition 1 (Sequence)**.** *A sequence is a function whose domain is either the set $\mathbb{N}$ of natural numbers or a subset of it in the form $\{x \in \mathbb{N} : x < n\}$, for some $n \in \mathbb{N}$. The length of a sequence is the number of elements in the domain of the function.*

Remark. *The set of natural numbers, $\mathbb{N}$, will be defined as $\mathbb{N} = \{0, 1, 2, 3...\}$ in this thesis.*

To write a sequence, we list its digits in order, for example 013019, which length is 6.

For the following definitions and examples of this chapter, we are only going to use binary sequences. Therefore every sequence will only be composed by at most two different digits (0 or 1).

Definition 2. *If s is a sequence of length $l \in \mathbb{N}$ and $n \in \mathbb{N}$ is at most l, then $s \restriction n$ denotes the initial part of s of length n.*

For example, if $s = 011011$ then $s \restriction 3 = 011$.

Definition 3. *If s is a sequence of length l and x is 0 or 1, then sx is the sequence of length $l + 1$ whose last element is x and all other elements agree with those of s.*

For example, if $s = 01011$ then $s0 = 010110$ and $s1 = 010111$.

With this basic background about sequences, we can now give a definition of rooted trees.

Definition 4 (Rooted tree)**.** *A rooted tree is a non-empty set of finite numerical (i.e with values in* $\mathbb{N}$*) sequences* T *such that for any* $s \in T$ *of length* n *and for any* $l < n$ *then* $s \upharpoonright l \in T$*.*

To better comprehend this definition, let's look at the rooted tree in Figure 1.

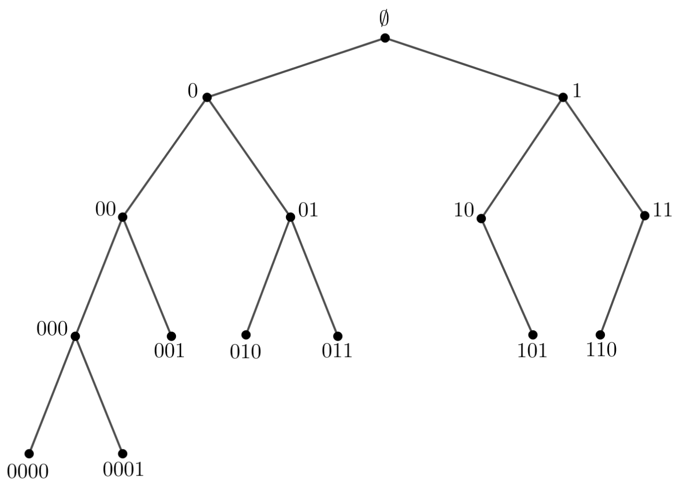
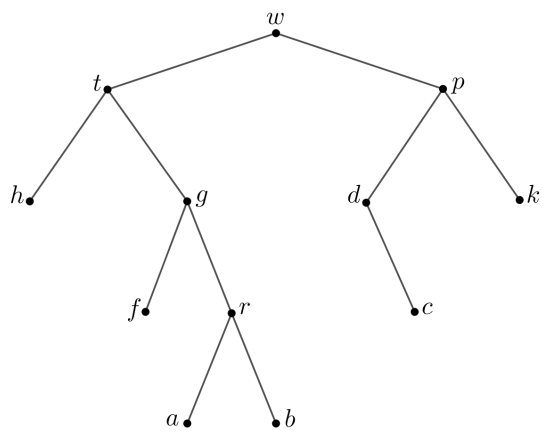


Figure 1: *Example of a finite binary rooted tree.*

As we can see, each node corresponds to a unique sequence and, in particular, the root corresponds to the empty sequence $\emptyset$.

The sequences illustrated at each node tells us something about the position of each node, in particular about the path from the root to the node (see Definition 6). We could think about each digit in the sequence as a representation of a choice. Imagine standing at the root and walking down the tree. Each time you choose to go right, you add a 0, and each time you choose left, you add a 1.

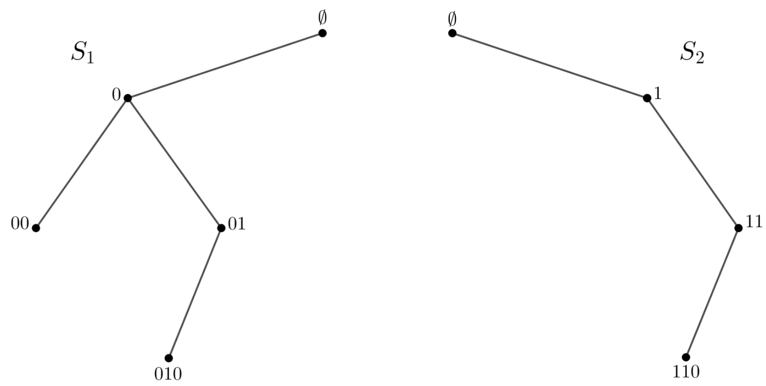
Example. *Given the tree illustrated below, try to find the node identified by the sequence* $s = 0110$*.*



We have to trace the path described by the digits of the sequence. We start at the root. Since first digit of s is 0, we choose the left edge and arrive at node t. The second digit of s is 1, so now we choose the right edge and arrive at node g. The node we are now at corresponds to the sequence 01. We continue in this way until we discover that a is the node we were looking for.

Definition 5 (Subtree)**.** *A subtree of a tree T is a subset S of T that is a tree in its own right.*

In Figure 2 we can see some examples of subtrees of the rooted tree in Figure 1.



It's important to keep in mind that every tree has the empty set $\emptyset$ as an element. The next definitions will tell us more about a particular type of subsets of a tree.

Definition 6 (Path with a root)**.** *A path with a root, p, in a rooted tree T is a subtree of T such that for any sequence $s, t \in p$ with lengths n, k respectively $n \leq k$, we have $s = t \restriction n$.*

Therefore the length of a path from the root to a node n is the same thing as the length of the sequence n.

A first introduction to Kőnig's Lemma will be given by introducing a weaker form of the lemma, namely Weak Kőnig's Lemma, which concerns rooted trees such that every node is connected to at most two lower nodes. These rooted trees are called binary trees.

Definition 7 (Binary tree)**.** *A binary rooted tree is a rooted tree T where all the sequences in it are functions from some $\{n \in \mathbb{N} : n < k\}$ to $\{0, 1\}$.*

In other words, every sequence of a binary rooted tree will at most be composed by two different digits: 0 or 1. We can also think of a binary rooted tree as a tree which, from every node, has at most 2 edges.

As mentioned above, Kőnig's Lemma treats infinite trees.

Definition 8 (Infinite tree)**.** *A tree is infinite if it contains infinitely many sequences.*

By the concepts of graphs which we will see in Section 2.3, the sequences will become vertices and edges will be given by two sequences $s \sim r$ if $|l(s) - l(r)| = 1$ and

$$\begin{cases} \text{s} \restriction l(r) = r, \text{ or} \\ \text{r} \restriction l(s) = s. \end{cases}$$

In this framework, we could define an infinite tree as:

- a tree with infinitely many nodes,
- a tree with infinitely many edges.

With these definitions in mind, we can take a close look at Weak Kőnig's Lemma and some of its applications. Note that later in this paper (Section 2.3), new definitions such as *graphs* and *trees* will substitute the definition of *rooted tree* in order to give a new formulation of Kőnig's Lemma.

2 (Weak) Kőnig's Lemma and applications

In this chapter we will introduce two versions of Weak Kőnig's Lemma. The first version treats rooted binary trees, the second version treats rooted k-ary trees. The proofs of both versions are very similar, but we will point out the main differences.

There are many mathematical constructions that can be thought of as finding paths on infinite binary trees [2, page 10]. In Section 2.2 and 2.3 we are going to see some of these constructions when introducing two applications of Kőnig's Lemma. We will also introduce some basic concepts of graph theory since Section 2.3 is an application within this field.

There is more than one way to prove Kőnig's Lemma. This is not an unusual quality for mathematical theorems. The existence of two or more proofs for a theorem could be an opportunity to gain more information about the theorem itself, such as its value within different fields of mathematics. On the other hand, it could also be an invitation to shift our focus to the methods used in the proof and analyze how these different methods might be connected, and maybe even find some common properties.

In the proof below, we are going to deal with infinity. Mathematical induction is a universal method used to reduce infinite problems to finite problems.

Remark. *The induction technique consists of three steps:*

1. *Base case. Show that the mathematical statement (induction hypothesis) holds for a first value.*

2. *Induction step. Assume that the mathematical statement holds for an arbitrary value n and then show that it also holds for the value $n+1$.*

3. *Conclusion. Since the induction hypothesis holds for both the base case and the inductive step, then it holds for any value.*

However, it is good to know that induction is not a universally functional mechanism, but an axiom (actually, it is a whole family of axioms) in Peano Arithmetics (more about this in Chapter 3).

Theorem 2 (Weak Kőnig's Lemma for binary trees)**.** *Let T be an infinite binary tree. Then T contains an infinite rooted path p.*

Proof. Assume that T is an infinite binary tree, as the one represented in Figure 2.
Take the sequence s of length n,where $n \in \mathbb{N}$, and let T_s be the subtree of T below s, such that

$$T_s = \{r \in T : r \upharpoonright n = s\} \cup \{s \upharpoonright k : k < n\}.$$

Graphically we can represent T_s as the union of the two subtrees given in Figure 2.

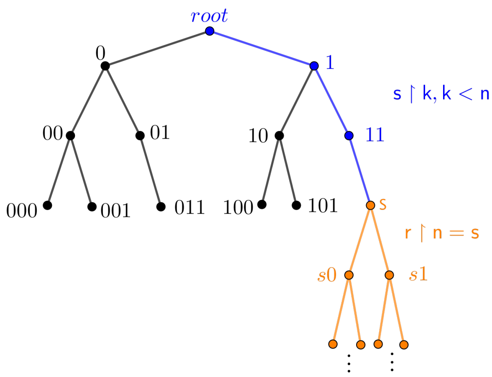


Figure 2: *The union of the blue and the orange subtrees gives us T_s.*

To check that T_s is a tree, it is sufficient to note that that for any $s \in T_s$ of length n, and for any $l < n$ then $s \restriction l \in T$.

To prove Kőnig's Lemma, we have to find a sequence $s(n)$ such that:

1. $s(n)$ has length n,
2. $s(n) = s(n+1) \restriction n$
3. the tree $T_{s(n)}$ below $s(n)$ is infinite

The set $\{s(n) : n \in \mathbb{N}\}$ will be our infinite path when the proof is completed.

Since we're proving this lemma with induction technique, we will start by choosing the third property above as our induction hypothesis.

Suppose we've chosen a sequence $s = s(n)$ of length n and that T_s is infinite. Since the tree is binary, we can then define T_s as:

$$T_s = \{r \in T : r \restriction (n+1) = s0\} \cup \{r \in T : r \restriction (n+1) = s1\} \cup \{s \restriction k : k \leq n\}.$$

Since $T_{s(n)}$ is infinite, at least one of the sets above has to be infinite. Let's have a graphical look at this infinite tree and each set.

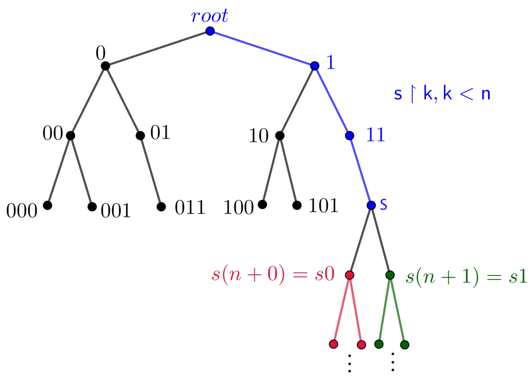


Figure 3: *Define T_s as the union of the blue, green and red subtree.*

The last set is obviously finite, so at least one of the first two subsets has to be infinite.
Now we have two cases.

1. If the first set $\{r \in T : r \upharpoonright (n+1) = s0\}$ is infinite, then we can set $s(n+1) = s0$ and in this case

$$T_{s(n+1)} = \{r \in T : r \upharpoonright (n+1) = s0\} \cup \{s0\} \cup \{s \upharpoonright k : k \leq n\},$$

   which is infinite.

2. If the first set isn't infinite, then the second set $\{r \in T : r \upharpoonright (n+1) = s1\}$ has to be infinite. In this case we set $s(n+1) = s1$, and just as above, $T_{s(n+1)}$ would as well be infinite.

Either way, we have defined a sequence $s(n+1)$ and proved the induction hypothesis for $n+1$, so we have proved Kőnig's Lemma. □

Note that since we have coded binary trees as sequences, then they are countable. In Chapter 4 (Theorem 10), we will see a version of Kőnig's Lemma for which there is no assumption of countability.

2.1 Presentation of Weak Kőnig's Lemma (for infinite subtrees of a k-rooted tree)

The above described Lemma is a weak variant of Kőnig's Lemma. We will now prove a more common variant, which we still call Weak Kőnig's Lemma, but for infinite k-ary rooted trees. Actually, this version is equivalent to Theorem 2, since it is possible to code k-ary rooted trees to binary trees. (The equivalence between Theorem 2 and Theorem 3 will not be proven in this paper.)

For k-ary rooted trees, we will now not have only two choices at each node (as we had with binary rooted trees), but, we will have up to (at most) k choices, with $k \in \mathbb{Z}^+$. In other words, we are going to extend the lemma to a more general case, i.e rooted k-ary trees.

Definition 9 (k-ary tree). *A rooted k-ary tree is a tree whose elements are sequence with values in* $\{0, ..., k-1\}$.

The following proof is very similar to the proof of Kőnig's Lemma for binary rooted trees, so we will only give a sketch of the idea of the proof for rooted $k-$ary trees.

Theorem 3 (Weak Kőnig's Lemma for rooted k-ary trees). *Let T be an infinite rooted k-ary tree. Then T contains an infinite path p.*

Proof. Take an infinite rooted k-ary tree T. Let T_s be the subtree of T as defined with root in s. Define the subtree T_s as

$$T_s = \{s \upharpoonright t : t \leq n)\} \cup \{r \in T : r \upharpoonright (n+1) = si\}$$

for $n \in \mathbb{Z}^+$, $t \in \{0, ..., n\}$ and $i \in \{0, 1, ..., k-1\}$. As before, we want to find a sequence $s(n)$ of elements of T such that:

1. $s(n)$ has length n,
2. $s(n) = s(n+1) \upharpoonright n$,
3. the tree $T_{s(n)}$ below $s(n)$ is infinite.

The third statement is our induction hypothesis. When the proof is completed, the set $\{s(n) : n \in \mathbb{N}\}$ will be our infinite path.

The the last set can be interpreted as the set of every subtree $T_{s(n+1)}$ of T_s with root in a node si. Since there are at most k such nodes and since we know by hypothesis that T_s is infinite, then at least one of the subtrees $T_{s(n+1)}$ has to be infinite.

Therefore there has to be a subtree $T_{s(n+1)}$ with root in the node si. Define the infinite subtree as $T_{s(n+1)} = \{s \upharpoonright t : t \leq n\} \cup \{r \in T : r \upharpoonright (n+1) = si\} \cup \{si\}$. In this way we have proved the induction hypothesis for $(n+1)$ and defined $s(n+1)$. □

Just as with the case of rooted binary trees, rooted $k-$ary trees are also countable since they are coded as sequences. In Chapter 4 we are going to introduce the stronger version of this lemma, namely Kőnig's Lemma for which we will formalize the lemma with *graphs*, and no longer sequences. Before we introduce graph theory, we are going to see an application for Weak Kőnig's Lemma within metric spaces.

2.2 Sequential compactness of $[0, 1]$

A first example of application of (Weak) Kőnig's Lemma is the proof of sequential compactness for a set $X = [0, 1]$.

We are going to introduce this with some definitions that will help us through the proof.

Definition 10 (Metric space). *A metric space is an ordered pair (X,d) where X is a set of points and d is a metric on X i.e, $d : X \times X \to \mathbb{R}$ such that for any points x,y,z the following holds:*

1. *d(x,y)=0 $\iff$ x=y, for arbitrary $x, y \in X$*

2. *d(x,y)=d(y,x), for $x, y \in X$*

3. *d(x,z) $\leq$ d(x,y) + d(y,z), for $x, y, z \in X$*

In other words, a metric is simply a function that defines the distance between each pair of elements in a set.

Remark. *A sequence $(x_n)_{n\in\mathbb{N}}$ in a metric space is a function $\mathbb{N} \to X$.*

Remark. *A subsequence of a given sequence $(x_n)_{n\in\mathbb{N}}$ is any other sequence $(y_n)_{n\in\mathbb{N}}$ that is of the form $(y_n)_{n\in\mathbb{N}} = (x_{n_k})_{k\in\mathbb{N}}$ where $(n_k)_{k\in\mathbb{N}}$ is an increasing sequence of natural numbers.*

Definition 11. *A sequence $(x_n)_{n\in\mathbb{N}}$ in a metric space (X, d) is said to converge to a point $p \in X$, if for every $\epsilon > 0$ there exists an $M \in \mathbb{N}$ such that $d(x_n, p) < \epsilon$ for all $n \geq M$.*

Note that the point p is the *limit* of $(x_n)_{n\in\mathbb{N}}$.

Definition 12. *(Sequential compactness) A metric space (X,d) is sequentially compact if for each sequence $(x_k)_{k\in\mathbb{N}}$ in X there is a subsequence $(x_{k_n})_{k\in\mathbb{N}}$ that is convergent.*

In order to better understand the concepts of convergent sequence and convergent subsequence, we are going to look at an the following example.

Example. *Fix $X = [0, 1]$ with standard metric. Take the sequence defined as:*

$$x_k = \begin{cases} \frac{1}{k}, \textit{if } k \textit{ is odd} \\ 1, \textit{else.} \end{cases}$$

We see that the sequence x_k then has two subsequences with different limits. In fact, the subsequence $(x_{2k+1})_k$ converges to 0 and that the subsequence $(x_{2k})_k$ converges to 1. The sequence x_k does not converge.

In other words, if a subsequence of a sequence converges, then the sequence does not have to converge itself. But if a sequence converges to q, then every subsequence of the sequence also converges to q.

We will prove that Weak Kőnig's Lemma implies that the metric space $X = [0, 1]$ is sequentially compact. In the proof we will construct an infinite tree dividing the interval $[0, 1]$. Since this construction is a bit complicated, we are going to give an example of it before we head on into the proof.

Example. *In the proof of Theorem 4 we are going to construct a binary tree whose nodes at depth k are identified with the intervals $I_k = [\frac{r-1}{2^k}, \frac{r}{2^k}]$ for $r \in \{0, ..., 2^k\}$. Take the following example given in Figure 4.*

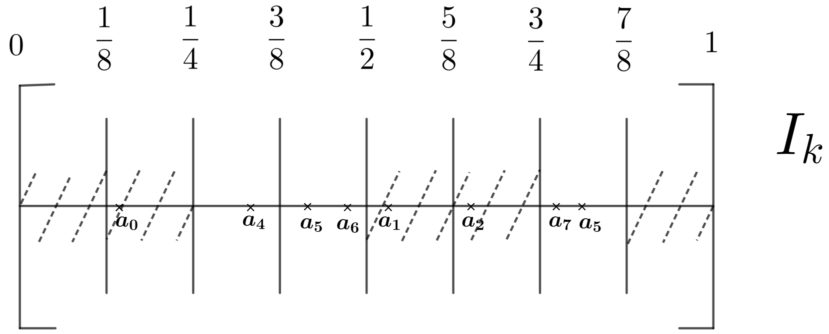


Figure 4: *This figure illustrates an example of the construction used in the proof of Theorem 4.*

In this figure we only have finite number of elements a_n *but the method is the same as for infinitely many elements* a_n.

The first step of the construction is to divide the interval, beginning with $[0,1]$ *in half and inspect the new intervals. At stage* k *of the construction we throw away any interval that does not contain any element* a_n *such that* $n \geq k$.

We will now apply this construction on the interval shown in Figure 1 with the elements $a_0, a_1, ..., a_6$ *in it.*

First, put $n_0 = 0$ *and note that in the interval* $[0,1]$ *there are elements* a_n, $n > 0$. *Then take* $k = 1$ *and divide the interval* $[0,1]$ *in two. Note that there are elements* a_n *such that* $n \geq 1$ *in both* $[0, \frac{1}{2}]$ *and* $[\frac{1}{2}, 1]$. *Now take* $k = 2$, *and divide* $[0, \frac{1}{2}]$ *and* $[\frac{1}{2}, 1]$ *in two. Since there is no element* a_n *such that* $n \geq 2$ *in the interval* $[0, \frac{1}{4}]$ *and* $[\frac{1}{2}, \frac{3}{4}]$, *we throw away these intervals. In the next step, for* $k = 3$, *notice that the interval* $[\frac{7}{8}, 1]$ *will be thrown away, and so on.*

This construction could also be illustrated by the binary tree given in Figure 5.

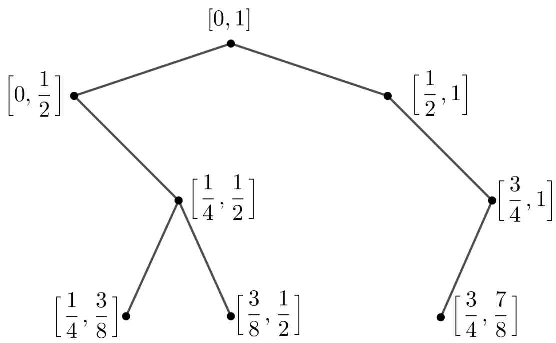


Figure 5: *This figure illustrates an example of the construction used in the proof of Theorem 4.*

Every time a new interval (which isn't thrown away due to the element a_n *it contains) is constructed, a new node is added to the tree.*

Imagine that for any stage k of the construction there will always be at least one interval containing an element a_n such that $n < k$. If so, the three will have an infinite path.

Another easier example could be given by the sequence defined as $a_k = 1$ for all k in $\mathbb{N}$. If we apply the same construction as above, we will get a tree with an infinite path such that every node has valency 2.

With this example in mind, we will prove the following theorem by constructing an infinite tree and finding a path.

Theorem 4. *Every sequence (a_n) of elements of $X = [0, 1]$ has a subsequence converging to some element in X.*

Proof. Fix (a_n) in $[0, 1]$. We will construct a binary rooted tree to find an infinite path.

A graphical example of this construction was given in the example above.

We now want to apply Kőnig's Lemma so we have to construct a binary tree. This is, a tree which nodes at depth k are identified with the intervals $I_k = [\frac{r-1}{2^k}, \frac{r}{2^k}]$ for $r \in \{0, ..., 2^k\}$ which have infinitely many elements a_n. The root of this tree will then be identified as $[0, 1]$. In particular this tree is by construction infinite, since there are infinitely many elements a_n (i.e, infinitely many nodes). Kőnig's Lemma tells us that there is an infinite path through this tree.

So there are closed intervals

$$I_0 = [0, 1] \supseteq I_1 \supseteq I_2 ... \supseteq I_k \supseteq I_{k+1} \supseteq ...$$

such that:

1. $|I_k| \to 0$, and
2. for all $k \in \mathbb{N}$ there is some $n \in \mathbb{N}$ such that $a_n \in I_k$.

From statement 2 and the fact that the $(I_k)_k$ are descending, we can conclude that for all $k \in \mathbb{N}$, the set $\{n \in \mathbb{N} : a_n \in I_k\}$ is infinite. Let us now recursively construct an increasingly function $k \to n_k$ such that $a_{n_k} \in I_k$ for all $k \in \mathbb{N}$.

Put $n_0 = 0$ then $a_{n_0} = a_0 \in [0, 1] = I_0$. If $n_0, ..., n_k$ have been defined, consider the set

$$S = \{n \in \mathbb{N} : a_n \in I_{k+1}\} \setminus \{n \in \mathbb{N} : \exists 0 \leq i \leq k : n_i = n\}.$$

Let $n_{k+1} = min(S)$. Then $n_{k+1} > n_i$ for all $0 \leq i \leq n$ and $a_{n_{k+1}} \in I_{k+1}$. So we have defined $a_{n_{k+1}}$. After constructing $(a_{n_k})_{k \in \mathbb{N}}$, we have to show that it converges in $[0, 1]$. To this end, observe that for all $i \geq k$ we have $a_{n_i} \in I_k$. Since the length of I_k is 2^{-k}, then for $i, j \geq k$ we have $d(a_{n_i}, a_{n_j}) \leq \frac{1}{2^k}$, where $d(a_{n_i}, a_{n_j})$ is the distance from a_{n_i} to a_{n_j}.

Recall that a sequence of real numbers $x_1, x_2, x_3, ...$ in a metric space (X, d) is called a Cauchy sequence if for every positive real number ϵ, there is an integer N such that $|x_m - x_n| < \epsilon$ for all natural numbers $m, n > N$.

Therefore $(a_{n_k})_{k \in \mathbb{N}}$ is a Cauchy sequence. Since $[0, 1]$ is a complete metric space, this shows that $(a_{n_k})_k$ converges.

□

We have shown that Weak Kőnig's Lemma implies sequential compactness of the interval $[0, 1]$. We observe that the proof does not use any induction, although the definition of the tree we used might appear inductive. In fact, it can be alternatively described directly as a set of binary sequences

$$T = \{s| \; \forall N > l(s) \; \exists n > N : a_n \in [c_s, c_s + 2^{-l(s)}],$$

where $c_s = \sum_{i=1}^{l(s)} s_i 2^{-i}$ denotes the starting point intervall corresponding to s, as illustrated in Figure 5. In the last chapter of this paper we are going to show the implication the other way around, also avoiding induction in our proof, hence we'll see that sequential compactness of $[0, 1]$ and Weak Kőnig's Lemma are equivalent, and hint at the possibility to express this equivalence in a weaker axiom system than ZF set theory, which will be introduced in Chapter 3.

2.3 k-colourable infinite graphs

Kőnig's Lemma does also have applications in the theory of graphs, in particular graph coloring. Following there will be an introduction of some basic concepts within graph theory. Even though Kőnig himself wrote the first textbook on the field of grah theory, most of the following definitions have been taken from a book written by Distel R [3].

Definition 13. *A graph is a pair (V, E) such that $E \subseteq V \times V$ and E is symmetric and irreflexive.*

Remark. *Given a graph G, we refer to the vertices of G as $V = V(G)$ and the edges of G as $E = E(G)$*

Remark. *Every rooted k-ary tree defines a graph in the sense of Definition 12.*

In other words, a graph consists of two data, namely vertices and edges. An example of a graph is given in Figure 5. From this figure we see that an edge is an unordered pair of vertices and that it is not important how the edges are drawn.

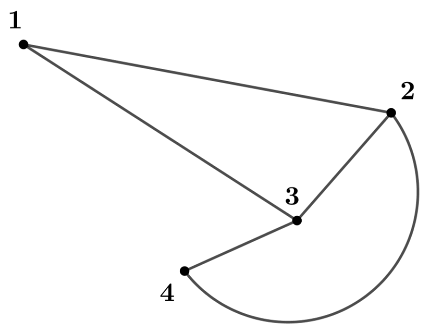


Figure 6: *Example of a graph. The vertices of this graph are given by $V = \{1, 2, 3, 4\}$, and the edges are given by $E = \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{2, 4\}, \{3, 4\}\}$.*

To be precise, Definition 13 is commonly known as the definition of an undirected graph. The symmetrical property $(v, w) \in G \iff (w, v) \in G$ is what distinguishes directed and undirected graphs. In this thesis we are only interested in undirected (and unweighted) graphs, therefore our definition of graph given above.

Definition 14 (Adjacent vertices)**.** *Two vertices connected by an edge are called adjacent.*

In Figure 5, for example, the vertices 1 and 2 are adjacent.

We will now introduce some definitions concerning graph coloring.

Historically, graph coloring was introduced by the mathematician Francis Guthrie who noticed that with four colors it was possible to color the map of the counties of England such that no region with common borders shared colors. Graph coloring is still commonly applied in practical areas.

A graph coloring problem is to assign colors to certain elements of a graph (usually vertices or edges) subjects to restrictions. The most common graph coloring problem is vertex coloring.

Definition 15 (Vertex coloring)**.** *A vertex coloring of a graph G=(V,E) is a map $c : V \to S$ such that $c(v) \neq c(w)$ whenever v and w are adjacent. The elements of the set S are called the available colors.*

Figure 7 illustrates examples of valid colorings of graphs.

If we look at graph G_1, it is composed of four vertices and no edges, hence the graph can be colored with only one color. If we connect two of the vertices, we get graph G_2. For G_2 we have to have at least 2 colors since the two adjacent vertices has to have different colorings.

The figure does also illustrate three different colorings of the graph G, all of them valid, and there are only three different ways of coloring it.

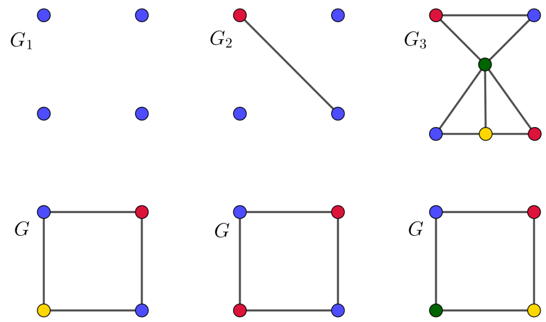


Figure 7: *Examples of graph coloring.*

In particular, the coloring of G_1, G_2, G_3 gives us the chromatic number of these graphs.

Definition 16 (Chromatic number)**.** *The chromatic number $X(G)$ of a graph G is the minimal k such that there is a coloring $V(G) \to \{1, ..., k\}$.*

A graph G with $X(G) = k$ is called $k - chromatic$; if $X(G) \leq k$ then it's called $k - colourable$.

The chromatic number of G_1 is $X(G_1) = 1$, hence the smallest numbers of colors needed to color the graph is one. Also, we have that $X(G_2) = 2$, $X(G_3) = 4$. From graph G we see that $X(G) = 2$ G is 4-colourable.

In the next theorem we will mention countable graphs. A countable set is a set that has the same cardinality (number of elements) as some subset of the set of the natural numbers.

With these definitions in mind, we can now prove a theorem which is implied by Kőnig's Lemma.

Theorem 5. *Let G be a countable graph such that every finite subgraph of G is k-colourable. Then G is k-colourable.*

Proof. Assume G is a countable graph, then we are able to enumerate it's vertices as v_0, v_1, v_2, v_3, etc. Let G_n be a subgraph of G induced by the vertices v_0 through v_n (for example $G_3 = v_0, v_1, v_2, v_3$). Each vertices v_n corresponds to a subset G_n, therefore there are countably many G_n. Since $G_n \subset G_{n+1}$ by construction, it follows that $\bigcup_{n=0}^{\infty} G_n = G$.

For each G_n there is a set C_n of $k - colorings$ of G_n.

Take a coloring of G_{n+1} and remove the vertex v_{n+1}, we now have a coloring of G_n that is an element of C_n [1]. Therefore, for every coloring $c_{n+1} \in C_{n+1}$ of G_{n+1}, there is some induced coloring $c_{ind} \in C_n$ such that $c_{ind} \prec c_{n+1}$ [2].

The next step is to construct a tree and use Kőnig's Lemma. Given the enumeration $v_0, v_1, v_2, ...$ of the graphs vertices, and the subgraphs $G_n = \{v_0, v_1, ..., v_n\}$, we construct a k-ary tree. Observe that every coloring of G_n can be encoded as a sequence of elements in $\{1, ..., k\}$ of length n. Indeed, at the i-th position we write down the color of v_i. Encoded in this way, consider all k-colorings of all G_n at the same time.

The resulting set of sequences is a k-ary tree. This tree is actually infinite since, by assumption, every G_n admits at least some k-coloring. Hence there is an infinite path. This path defines a k-coloring of G by assigning to v_i the colori in $\{1, ..., k\}$ that is written at the i-th position of the path. Therefore, we can state that there exists a k-coloring for G. □

(The idea for this proof was taken from a post published in Stack Exchage [4].) This theorem is very powerful in graph theory, but is important to note that this theorem only holds for countable graphs. In Chapter 4 we are going to see a theorem which also holds for non-countable graphs. The theorem in Chapter 4 will therefore be more applicable, but it relies on more axioms. More about this in the following chapter.

[1] Note that removing vertices does not invalidate a coloring.

[2] If we have a coloring $c_{n+1} : G_n \to \{1, ..., k\}$ and a coloring $c_n : G_n \to \{1, ..., k\}$ with $G_n \subseteq G_{n+1}$, then $c_n \prec c_{n+1}$.

3 Foundations of set theory

In order to find the foundations of set theory, we are going introduce logic and it's role within mathematics.

Logic is the study of forms of reasoning. The aim of logicians is to define a well-structured system of reasoning that, if applied on *true* [3] assumptions, guarantees correct conclusions.

Logic has its deepest roots in philosophy since ancient times and was initially used as a method for studies of argument, meaning and existence. Traditionally, logic was built upon only three laws which were theorised by the great philosopher Aristotle: the law of identity, the law of non-contradiction and the law of the excluded middle. The law of identity states that 'whatever is, is' and can be formalized as $\forall p(p = p)$. The law of non-contradiction states that 'nothing can both be and not be', and is formalized as $\forall p \neg(p \wedge \neg p)$. At last, the law of excluded middle states that 'everything must either be or not be', i.e $\forall p(p \vee \neg p)$. These laws are also at the foundation of mathematics. (The laws given above are expressed in the formal language of mathematical logic, but we will not introduce this topic in this paper.)

During the late 19th century, the study of the foundations of mathematics became of great interest. Logic became a useful tool to study the formal systems of mathematical models and the deductive processes of formal proof systems. This was the time when mathematicians and philosophers began to analyze the underlying pillars of mathematics, aiming to give mathematics a logical foundation.

Today logic is an interdisciplinary area that covers philosophy, mathematics, linguistics, computer science, artificial intelligence, sociology and more. In the context of mathematics, logic is broadly used in different fields, such as set theory, model theory, recursion theory and proof theory.

In the next chapters we are going to represent Zermelo Fraenkels set theory and Peanos axioms, which are some of the most notorious cases of mathematics explained with logic.We are also going use logic principles to seek which axioms seem to underlie Kőnig's Lemma and if all of them are necessary.

While writing this paper I found it very interesting to learn more about how aspects of logic interact with theorems. In fact, depending on how we formulate a theorem, the number of axioms that underlies the theorem might variate. For example, depending on the formulation of a theorem, the proof used to proof the theorem might vary and, as a consequence, we might recall to different axioms in the different proofs. A theorem is nothing more than a logical consequence of axioms. The proof of a theorem is a logical argument which establishes its truth through the inference of rules of a deductive system. Finding exactly which axioms are necessary for the proof of a theorem is not an easy game. Also, in some proofs we might implicitly use some axioms without even noticing that we are applying them. Learning about which axioms theorems are built on is crucial both to understand the theorem itself and to understand it's consistency. If we find out that a theorem (or even a axiomatic system) is based on axioms from which contradictions rise, the theorem won't seem as convincing anymore.

[3] Note that the truth value of a sentence is not defined in the field of logic. Logic is the study about what follows from what, i.e what conclusion follows from some premises. Whether the premises are true or false, is to be considered in other disciplines.

This is not different from when we deduce something from false premises. In Section 3.2 we are going to discuss a paradox arised from the Axiom of Choice.

3.1 Short summary of ZF and Peano Axioms

One attempt of formalizing all mathematical reasoning, using logic, was given was given by Ernst Zermelo and Abraham Fraenkel. They constructed the ZF-set theory, one of the most well known axiomatic systems of mathematics, trying to describe all mathematics in terms of sets and operation of sets.

Set theory had been developed before Zermelo and Fraenkel, by the german mathematician Georg Cantor in the mid-late 19th century. His theory is a 'naive' theory of sets, since it is defined informally, with no use of formal logic nor any rigorous axiomatic structure. After the discovery of some paradoxes in the 'naive' set theory, such as Russell's paradox, Zermelo and Fraenkel decided to re-construct set theory with a more systematical approach. The ZF-set theory was initially composed by eight axioms (extensionality, regularity, specification, pairing, union, infinity, replacement, power set), each of them expressed in first-order logic, but we will not define them in this paper. Later on, the axiom of choice was added to ZF, creating an extension of ZF, namely ZFC (see next chapter). We can today reduce most known maths to those nine axioms, even though some of the ZF-axioms can not be considered pure logic. For example, the axiom of Infinity and the axiom of Chioce.

Also, as an effect of the results from Gödel's incompleteness theorems, the dream of finding a complete [4] and consistent [5] set of axioms for all mathematics vanished. Gödel showed that a formal deductive system can't show it's own consistency. As a consequence, it is not possible to show that a formal deductive system is consistent, but it is possible to show that it isn't consistent by finding a contradiction in the system. This discovery includes all formal deductive systems which can express Peano Axioms. We are going to see that ZFC is one of them. Since ZFC is identified with ordinary mathematics, the consistency of ZFC can not be demonstrated in ordinary mathematics. However, even though we can not prove the consistency of ZF and therfore if it really is the ferm foundation of mathematics, it is still today the most common axiomatic structure for set theories. And until now, ZF has been immune to paradoxes.

In the next chapter we will see that some of the theorems mentioned in this paper are based upon Zermelo-Fraenkel set theory and some are based on a weaker axiomatic system.

Another attempt of formalization of a part of mathematics, was given by the italian mathematician Giuseppe Peano. Peano created a system of axioms which describes the essence of natural numbers without evoking any numerical concepts. During his time, the language of mathematical logic was only just starting to develop. In order to formalize arithmetic, Peano created new logical notations to present his axioms, such as the symbol for set membership $\in$ and the symbol of implication $\subset$.

The Peano Axioms (PA) are five and the can be derived from the ZF axioms. In this paper, we are going to show how we can create a model of PA in ZF.

[4] A theory is called complete with respect to a particular propertyif every formula having the property can be derived using the theory.

[5] A theory is consistent if it does not entail any contradiction.

A model of second-order Peano arithmetic is a set N with a function S defined on N such that:

(PA1) There is a distinguished element $0 \in N$. In particular N is a non-empty set.

(PA2) $S(n) \in N$ for all $n \in N$.

(PA3) For all $n \in N$ and $m \in N$, then $n = m$ if and only if $S(n) = S(m)$. That is, S is an injection.

(PA4) There is no $n \in N$ such that $S(n) = 0$.

(PA5) If X is a set such that $0 \in X$ and if $n \in N \cap X$ implies $S(n) \in N \cap X$, then $N \subseteq X$.

A model of second order Peano arithmetic is defined as the triple $(N, 0, S)$ for which the axioms above holds. Note that there is no mention about what actually a $N, 0, S$ is, indeed these axioms do only describe the behaviour of the set of natural numbers and the function S (which behaves like the successor function). Even though the second-order Peano arithmetic is what Peano originally wrote, there is a similar model which is better-known today, namely Peano arithmetic. The difference between these two Peano models lies in the fifth axiom. In fact, for second-order Peano arithmetic PA5 states that induction holds for all subsets of the model, while in Peano arithmetic the fifth axiom states that induction holds for sets defined by a formula of first order logic. For this reason, Peano arithmetic is also known as first-order Peano arithmetic. The reason why we gave a description of the second-order Peano arithmetic is because its models admit a more concise self-contained definition. Yet, every model of the second-order version is also a model of the firs-order version.

Now we are going to show how we can derive the model of second-order Peano arithmetics from ZF. This is the so called Zermelo's Construction.

Theorem 6. *A model for Peano arithmetics exists, if Zermelo-Fraenkel set theory holds.*

Sketch of proof. *This model is given by the triple* $(N, 0, S)$. *This is, a set* N *such that* $N = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, ...\}$, *a distinguished element* $0 = \emptyset \in N$ *and a function* $S : N \to N$ *on sets such that* $S(A) = \{A\}$.

If we apply S *on the empty set, we then have* $S(\emptyset) = \{\emptyset\}$. *Note that if we apply the successor operation* S *again, we get* $S(\{\emptyset\}) = \{\{\emptyset\}\}$, *and so on. We could think of the set* N *as the aggregate of all these nested sets. The set* N *together with the function* S *and element* $0 = \emptyset$ *satisfies Peano axioms. It is possible to show that this models holds for the Peano axioms, but we will not prove it in this paper.*

To understand that this model gives us a construction of the natural numbers, we can visualize the sets in the following way.
We indeed define
$0 = \emptyset \in N$,
$1 = \{\emptyset\} \in N$,
$2 = \{\{\emptyset\}\} \in N$,
$3 = \{\{\{\emptyset\}\}\} \in N$,

$\ldots,$
and so on.

$\square$

The underlying idea of this proof is that natural numbers can be synthesized from sets.

We have now shown that we can construct the natural numbers in the ZF set theory, which is a great first step towards the goal of describing mathematics in terms of sets and operation of sets. In the next chapter we will look deeper into some axioms of ZFC and some complications arised from the Axiom of Choice.

3.2 Different flavours of the Axiom of Choice and related statements

Among the most discussed axioms of mathematics, the Axiom of Choice (AC) is the dominant one after Euclid's Axiom of Parallels [5]. Even though the axiom appears self-evident, paradoxes has arised from it, such as the Banach–Tarski paradox. In short, this paradox states that given a sphere S^3 it is possible to cut it into a finite number of pieces and then re-assest them into two spheres identical to S^3. As a consequence, not every set is measurable. In some frameworks, for example physics, it can be interesting to be able to measure every set, but in this case the Axiom of Choice can not be used. In general, not being able to measure all sets, is not a problem and therefore the Axiom of Choice very used.

Because of the nature of the Axiom of Choice, the Zermelo-Fraenkel set theory has been divided in two camps. There is the Zermelo–Fraenkel set theory without the axiom of choice (ZF) and the Zermelo–Fraenkel set theory with the axiom of choice (ZFC). Even though ZF one gets rid of the complications arising from the paradox, ZFC has become the standard form of axiomatization of set theory and a lot of mathematical ideas have developed from it.

In order to gain a deep understanding of AC, let's first define what a choice function is. Intuitively, a choice could for example be: what movie to see, what bicycle to buy, what mountain to climb. However, within mathematics the axiom of choice needs rigorous definition which would make this idea less intuitive. To explain this mathematical idea we will define a function and then a choice function.

Definition 17. *A function from X to Y is a subset $A \subseteq X \times Y$ such that:*

- $\forall x \in X\ \exists y \in Y : (x, y) \in A$
- $\forall x \in Y\ \forall y, y' \in Y : (x, y), (x, y') \in A \iff y = y'$

We write $f : X \to Y$ for a function $A \subseteq X \times Y$ from X to Y and $f(x) = y$ if $(x, y) \in A$.

There are different ways of presenting the axiom of choice. We will begin by explaining it with choice functions.

The *axiom of choice* states that for each family of pairwise disjoint non-empty sets, there is a *choice function* [6] defined on that set (or a 'choice set').

[6] A choice function is a function that operates on a collection X of non-empty sets and assigns to each set S some element f(S) of S.

A choice function matches our intuition behind choice. A daily-life interpretation of a choice function could be choosing what to wear. Look at Figure 8 for this example.

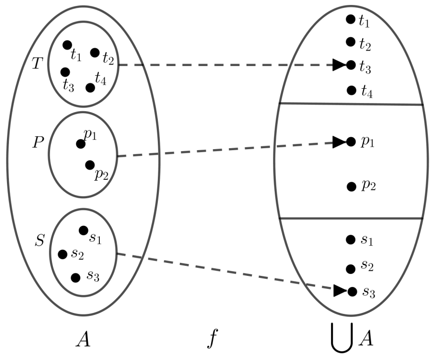


Figure 8: *Example of a choice function.*

Assume that T is a set of t-shirts, P is set of pants, and S a set of socks. Let A be the collection of the sets P, T, S. Then there is a function $f : A \to \bigcup A$, where $\bigcup A$ is the set of all elements of the sets in A. We could then say that choosing what to wear is the same as taking each of those sets and using a function f to send it to one of its elements.

A formal definition of choice function is given below.

Definition 18. *Given a family $(A_i)_{i\in\mathbb{I}}$ of non-empty sets, a choice function for $(A_i)_{i\in\mathbb{I}}$ is a function $f : \mathbb{I} \to \bigcup_{i\in\mathbb{I}} A_i$ such that $f(i) \in A_i$.*

We are now going to introduce two formal definition of the Axiom of choice.

Definition 19 (Axiom of choice - *AC*)**.** *The axiom of choice requires that for every family of non-empty sets there is a choice function.*

This first definition is presented in order to be compared with Definition 22 and will be used for Theorem 7.

Definition 20 (Axiom of choice - *AC*)**.** *Given a collection $Y = \{A_i : i \in \mathbb{I}\}$, such that A_i are pairwise disjoint non-empty sets, and a set X such that $A_i \in X$, there exists a function $f : X \to Y$ such that $f(A_i) \in A_i$.*

In set theory it's more common to describe the axiom of choice with sets. For example, Schindler [6] defines AC as following:

$$\forall x(\forall y(y \in x \to y \neq \emptyset) \wedge \forall y\forall y'(y \in x \wedge y' \in x \wedge y \neq y' \to y \cap y' = \emptyset)$$

$$\to \exists z\forall y(y \in x \to \exists u\forall u'(u' = u \leftrightarrow u' \in z \cap y))).$$

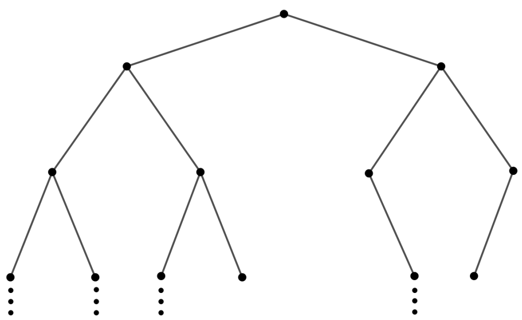
If we take a close look at these two definitions of AC, we can se how they relate. Note that for Definition 18 we use the definition of *function*, but for the definition given from Schindler, we don't use the definition of function. The pairwise disjoint non-empty sets A_i (from the first definition) correspond to variable y defined in Schindler's formalization. Also, X (from the first definition) corresponds to x. Schindler's interpretation states that for every family (namely x) of non-empty and pairwise disjoint sets (namely y), there is a "choice set". We will now take a look at a well-known variant of the Axiom of Choice which is sufficient to develop a lot of areas in mathematics, even though it is not as powerful as the Axiom of Choice. In order to understand the next axiom, we will list some definitions.

Definition 21. *A binary relation over a set X and Y is a subset over the cartesian product $X \times Y$. This set has the ordered pairs (x,y) as elements such that $x \in X$ and $y \in Y$.*

A special case of Axiom of Choice is the *Axiom of Dependent Choice* (DC). This axiom is a weak variant of the Axiom of Choice and we will see how we can prove DC from AC.

Definition 22 (Axiom of dependent choice - *DC*)**.** *Let R be a binary relation on a set $X \neq \emptyset$, then there is a sequence (x_n), $n \in \mathbb{N}$, such that $(x_n, x_{n+1}) \in \mathbb{R}$ for all $n \in \mathbb{N}$.*

To better understand this, let's look at the graph below.



Take the set X of all finite sequences, then $(x_n, x_{n+1}) \in R$ if x_{n+1} is the continuation of x with one step only.

We will now show how AC and DC are related by proving DC from AC.

Theorem 7. *The axiom of choice implies the axiom of dependent choice.*

Proof. Take an entire binary relation R on $X \times X$. Let $P(X \times X)$ be the set of all subsets of $X \times X$, namely the power set of $X \times X$. We now define the non-empty sets A_x as

$$A_x = \{(x, y) \in X \times X : (x, y) \in R\} = R \cap (\{x\} \times X).$$

Then $A_x \in P(X \times X)$. From the axiom of choice we know that there is a choice function $f : X \to X \times X$ such that $f(x) \in A_x$, hence $f(x) = (x, y_x)$ for an element $y_x \in X$ such that $(x, y_x) \in R$. Choose some $x_0 \in X$ and define inductively x_{n+1} as the unique element satisfying $f(x_n) = (x_n, x_{n+1})$. Then the sequence $(x_n)_{n \in \mathbb{N}}$ satisfies $(x_n, x_{n+1}) \in R$ for all $n \in \mathbb{N}$, by construction of f. □

3.2.1 Weak Axiom of Choice for finite sets

In the paragraph above we introduced the Axiom of Choice and the Axiom of Dependent Choice, which we saw is a weaker form of AC.

In this section we will introduce Weak Axiom of Choice for finite sets, which in turn is weaker than Axiom of Dependent Choice, and describe some consequences that comes with it. We will first recall the notion of a *countable* set.

Definition 23. *A countable set is a set with the same cardinality as some subset of the set of natural numbers.*

A countable set can either be a finite set or an infinite countable set. A set S is countable if there is an injective function from the set S to the set $\mathbb{N}$ of natural numbers. If we can show that such function is bijective, then S is called a countably infinite set.

Definition 24 (Weak axiom of choice for finite sets)**.** *Every countable family of non-empty finite sets must have a choice function.*

If we compare Definition 20 with Definition 24, we can note some similarity.

Next, we are going to show the equivalence between the weak axiom of countable choice for finite sets and the statement that says that 'a countable union of finite sets is countable'. We begin with proving that a countable union of finite sets is countable if we assume the Weak Axiom of Choice for finite sets.

Theorem 8. *A countable union of finite sets is countable if we assume that countable families of finite sets have choice functions.*

Proof. Let $(R_i)_{i\in\mathbb{I}}$ be a countable family of finite subsets of a set M. We have to prove that $\bigcup_{i\in\mathbb{I}} R_i \subseteq M$ is countable. In order to do this we shall produce a surjective function from $\mathbb{N}$ to $\bigcup_{i\in\mathbb{I}} S_i$. By countability of $\mathbb{I}$ there is an injective function $f : \mathbb{I} \hookrightarrow \mathbb{N}$. We define

$$S_n = \begin{cases} R_i, & \text{if } n = f(i), i \in \mathbb{I} \\ \emptyset, & \text{otherwise.} \end{cases}$$

Then $\bigcup_{n\in\mathbb{N}} S_n = \bigcup_{i\in\mathbb{I}} R_i$. Also, since we know that all $(R_i)_i$ are finite, then all $(S_n)_n$ are finite.

For all $n \in \mathbb{N}$, denote the cardinality of S_n as $k_n = |S_n| \in \mathbb{N}$. Since S_n is finite, then there is some function from S_n to $\{0, ..., k-1\}$ with k uniquely determined as above.

For $n \in \mathbb{N}$ define $A_n = \{\gamma : S_n \to \{0, ..., k-1\} | \gamma \text{ bijection .}\}$ By the choice of k_n, we know $A_n \neq \emptyset$ for all $n \in \mathbb{N}$. Actually, A_n is finite (of cardinality $k_n!$) since $\{0, ..., k-1\}$ is finite. Take the (finite) permutation group

$$Sym(k_n) = \{\gamma : \{0, ..., k-1\} \to \{0, ..., k-1\} | \gamma \text{ bijection } \}.$$

There is a bijection

$$\Omega : Sym(k_n) \to A_n$$

given by $\gamma \mapsto \gamma \circ \gamma_0$ for a fixed chosen element $\gamma_0 \in A_n$ and for $\gamma \in Sym(k_n)$.

An inverse to this map is given by the assignment

$$\Sigma : \gamma \longmapsto \gamma \circ \gamma_0^{-1}.$$

Indeed, we see that

$$\Omega(\Sigma(\gamma)) = (\gamma \circ \gamma_0^{-1}) \circ \gamma_0 = \gamma \circ (\gamma_0^{-1} \circ \gamma_0) = \gamma = id(\gamma)$$

for all $\gamma \in A_n$. Also,

$$\Sigma(\Omega(\gamma)) = (\gamma \circ \gamma_0) \circ \gamma_0^{-1} = \gamma \circ (\gamma_0 \circ \gamma_0^{-1} = \gamma = id(\gamma))$$

for all $\gamma \in Sym(k_n)$.

Let g be a choice function for $(A_n)_{n\in\mathbb{N}}$, which exists by the theorems assumption. Then

$$g(n) : S_n \to \{0, ..., k_n - 1\},$$

is a bijection for all $n \in \mathbb{N}$. A diagonal argument shows that

$$T = \bigcup_{n\in\mathbb{N}} \{n\} \times \{0, ..., k_n - 1\} \subseteq \mathbb{N} \times \mathbb{N} \hookrightarrow \mathbb{N}$$

is countable, since T is a subset of the countable set $\mathbb{N} \times \mathbb{N}$.

Let $\psi : \mathbb{N} \to T$ be a surjective function. We now construct a surjection $\nu : T \to \bigcup_{n\in\mathbb{N}} S_n$. Then $\nu \circ \psi : \mathbb{N} \to \bigcup_{n\in\mathbb{N}} S_n$ is also surjective, proving that $\bigcup_{n\in\mathbb{N}} S_n$ is countable.

Set $\nu(n, i) = [g(n)^{-1}](i)$ for $n \in \mathbb{N}$ and $i \in \{0, ..., k_n - 1\}$. So we get that

$$g(n)^{-1} : \{0, ..., k_n - 1\} \to S_n,$$

which is the map $\nu : T \to \bigcup_{n\in\mathbb{N}} S_n$ we were looking for. □

Now we will prove the implication the other way around, i.e the countable union condition for finite sets implies the axiom of countable choice.

Theorem 9. *Countable union condition for finite sets implies the axiom of countable choice for finite sets.*

Proof. Assume that we have a set A and let all S_i for $i \in \mathbb{N}$ be a subset of A. Define $(S_i)_{i\in\mathbb{N}}$ as a countable family of finite non-empty sets, then $\bigcup S_i$ is countable.

We want to show that there is a function $g : \mathbb{N} \to \bigcup S_i$, such that $g(i) \in S_i$ for all $i \in \mathbb{N}$. If we can reduce the situation to $A = \mathbb{N}$, then we can take the choice function as the function $g(i) = \min S_i \subseteq \mathbb{N}$.

Since the countable union of the finite set $(S_i)_i$ is countable, then there is an injection

$$\iota : \bigcup S_i \hookrightarrow \mathbb{N}.$$

Therfore we can define the choice function as

$$g(i) = \iota^{-1}(\min(\iota S_i))) \in \bigcup_{i\in\mathbb{N}} S_i.$$

Note that $\iota(S_i) \subseteq \mathbb{N}$ and that $\min(\iota(S_i)) \in \iota(S_i) \subseteq \iota(\bigcup_{i\in\mathbb{N}} S_i)$. In other words, we found that $\iota^{-1}(\min(\iota(S_i)))$ is well defined and it's also an element of S_i. Therfore g is a choice function for $(S_i)_{i\in\mathbb{N}}$.

□

We have now proven the equivalence. Note that we have proven this equivalence in the ZF axiomatic system, and not in ZFC. We can therefore sum up this result in the following corollary.

Corollary 1. *In ZF the following statements are equivalent:*

1. *Every countable family of finite sets has a choice function.*
2. *The union of countable many finite sets is countable.*

In the next chapter we will see a new interpretation of Kőnig's Lemma, which is stronger then the previous interpretations. We will also see that within the Zermelo-Fraenkel axiomatic system, the strong version of Kőnig's Lemma is equivalent to the two statements given in Corollary 1.

4 Reverse mathematics: equivalent descriptions of Kőnig's Lemma

In this chapter we will be going backwards, from theorems to axioms. Reverse mathematics is a program used in the aera of mathematical logic to find out which axioms are necessary to prove certain theorems.

Recall that in Chapter 2 we formalized trees with sequences, which made it possible to codify the trees and, therefore, made the trees countable. Thanks to this property, we have been able to prove Weak Kőnig's Lemma without using any version of the axiom of choice. We proved Weak Kőnig's Lemma in the Zermelo-Fraenkel axiomatic system. It is possible, but more complicated, to prove Weak Kőnig's Lemma in an axiomatic system which is weaker than ZF.

In this chapter we are going to introduce the stronger version of the lemma. Instead of sequences, we will formalize trees with the use of abstract sets. As a consequence of this formalization, the trees will no longer have the property of countability. By loosing this property, a stronger axiomatic system will be necessary for proving Kőnig's Lemma.

We will now introduce some definitions within graph theory.

Definition 25. *A graph G is infinite if* $|V| + |E| = \infty$.

Definition 26. *The valency of a graph G is given by*

$$val(v) = |\{w \in V(G) | (v, w) \in E(G)\}|.$$

Definition 27. *A graph G is locally finite if* $val(v) < \infty$ *for all* $v \in V(G)$.

To clarify, the valency of a vertex of a graph is the number of edges incident to the vertex.

Observation. *If G is an infinite, locally finite, graph, then* $|V(G)| = \infty$.

Definition 28. *A path in G is a sequence of vertices* $v_0, v_1, ...$ *such that* $v_i \sim v_{i+1}$ *for all* i. *If the sequence is finite, sat* $v_0, ..., v_n$ *we call this a finite path from* v_0 *to* v_n. *Otherwise it is called an infinite path.*

Definition 29. *A graph G is connected if for all* $v, w \in V(G)$ *there is a path from* v *to* w.

A path such that the first vertex and the last vertex are the same, for example $v_0, ..., v_0$ is a called a cyclic path.

Definition 30. *A graph is acyclic if it has no graph cycles.*

With these definitions in mind we are going to prove the following theorem.

Theorem 10 (Kőnig's lemma)**.** *If we assume that every countable family of finite sets has a choice function, then the following statement holds.*

For every infinite, locally finite and connected graph, there exists an infinite path $v_0, v_1, v_2, ...$ *such that* $v_i \sim v_{i+1}$.

Proof. Take an infinite, locally finite and connected graph G. Assume that the graph G is acyclic [7]. Even though this assumption is not necessary, it makes

[7] A connected acyclic graph is also known as a tree.

the proof a bit smoother. If the graph was cyclic, we could have picked a cycle and loop through it infinitely many times.

If we take a fix node $g_0 \in G$, we get a tree rooted in g_0. Now define T_n as the set of the nodes of G such that the unique path from g_0 to the elements of T_n has exactly n nodes.

First we are going to show that **for each n, T_n is finite.** The set T_0 is finite, since $T_0 = \{g_0\}$ by definition. Also, if we assume that T_n is finite, then the elements of T_{n+1} are all neighbours to the elements of T_n. But since G is locally finite, then T_{n+1} is finite, as it is a subset of a finite union of finite sets. It follows that for each n, T_n is finite.

Now, we want to prove that **for each n,T_n contains at least one node which is not ultimately terminal.** Define a vertex $g \in G$ as *ultimately terminal* if there is some $l_g \in \mathbb{N}$ such that all paths going from g_0 through g are at most of length l_g. We note that if all the elements of T_n were terminal, we could write l_g as the maximal length of a path from g_0 going through g, then $l = max\{l_g | g \in T_n\}$. This would implicate that $G = \bigcup_{k \leq l} T_k$, i.e G would be finite. But since G is infinite by assumption, then for each n,T_n contains at least one node which is not terminal.

Next step is to prove that **if the node $g \in T_n$ is not terminal, then there is some neighbour g' to g, such that $g' \in T_{n+1}$ such that g' is not terminal.** If g is connected only to terminal nodes in T_{n+1}, let $l = max\{l_{g'} | g' \in T_{n+1}$and g' is neighbour of g$\}$. Then every path that goes from g_0 through g must be at most of length l, and therefore g is terminal too. So, if the node $g \in T_n$ is not terminal, then there is some neighbour g' to g, such that $g' \in T_{n+1}$ such that g' is not terminal.

Define T'_n to be the subset of those nodes which are not ultimately terminal, as in Figure 9.

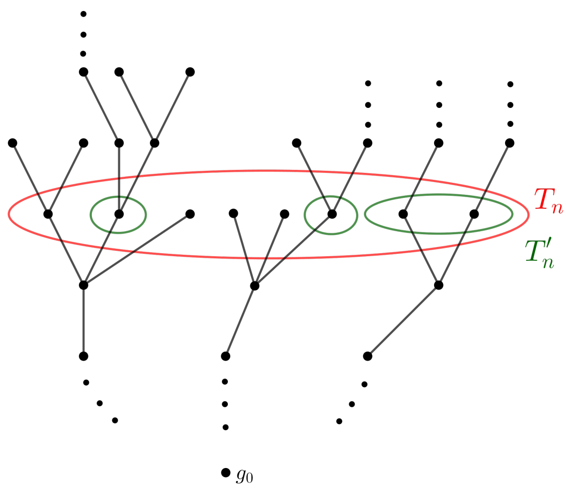


Figure 9: *This is an example of representation of the set T_n and its subset T'_n for a infinite, locally finite and connected graph.*

Since T_n is finite and non-empty, then T'_n is finite. Say that T'_n has k_n elements. Then the set S_n of bijective functions from $\{0, ..., k_{n-1}\}$ to T'_n is finite and non-empty [8]. We now want to find a way to enumerate T'_n in a way that allows us to always choose the "least element" of it's subsets. In order to do this, let F be a choice function for $(S_n)_{n\in\mathbb{N}}$. Let $T' = \bigcup_{n\in\mathbb{N}} T'_n$. Define a function $t : T'_n \to \{0, ..., k_n - 1\}$ on the graph as the following function:

$$t(g) = \begin{cases} F(n+1)(i), & g \in T'_n, i = min\{j < k_{n+1} | F(n+1)(j) \text{is a neighbour of} g\} \\ g, & g \in G \setminus T' \end{cases}$$

In other words, the function $t(g)$ picks the next node in an infinite path that goes through g, if there is such a node. If there is not, then the function "stops".

We are now going to use the recursion theorem which guarantees that recursively defined functions exist. If we start at a vertex g_0 and apply the function t recursively, we will produce the sequence

$$g_0, t(g_0), t^{\circ 2}(g_0), t^{\circ 3}(g_0), ...$$

We now want to show that **this sequence is a path**, and in order to do this, it suffices to observe that $t^{\circ n+1}(g_0) = t(t^{\circ n}(g_0))$ is adjacent to $t^{\circ n}(g_0)$, for $n \in \mathbb{N}$ by construction of t.

We therefore have found an infinite path in the graph G.

□

(The idea for this proof was inspired from a post published in Stack Exchage [4].)

With this proof we have shown that the Axiom of Choice is significantly stronger than what is necessary to prove Kőnig's Lemma. We have defined this lemma in ZF set theory and, therefore, ZFC set theory is not necessary. Although it is possible to prove the lemma with the axiom of choice, and therefore using the ZFC axiomatic system. Despite this, ZF alone is not enough for Kőnig's Lemma, in fact we have also used a weaker version of the chioce axiom, namely the weak axiom of choice for finite sets (Definition 24).

In the next section we are going to see a familiar theorem which is equivalent to Weak Kőnig's Lemma.

4.1 Sequential compactness of $[0, 1]$ and Weak Kőnig's Lemma

In chapter 2 we proved that Weak Kőnig's Lemma implied sequential compactness of $[0, 1]$, in this section we will prove the implication the other way around.

Theorem 11. *Sequential compactness of* $[0, 1]$ *implies Weak Kőnig's Lemma for binary trees.*

Proof. Assume that $[0, 1]$ is compact and let T be an infinite binary tree. We want to show that T has an infinite path.

Let $\sigma : \bigcup_{k\in N}\{0, 1\}^k \to [0, 1]$ be the map [9] identified by

$$\sigma(r) = \sum_{n=0}^{k} \frac{3}{4^{n+1}} r(n)$$

[8]Note that this is a reference to Section 3, Theorem 7

[9]Note that $\{0, 1\}^k$ is the set of k-tuples comprised of '0' digits and '1' digits. For example, if $k = 2$, then $\{0, 1\}^2 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$.

for $r \in \{0,1\}^k$.

Note that

$$0 \leq \sum_{n=0}^{k} \frac{3}{4^{n+1}} r(n) \leq \frac{3}{4} \sum_{n=0}^{\infty} \frac{1}{4^n} = \frac{3}{4} \frac{1}{1-\frac{1}{4}} = 1,$$

since $r(n)$ either has the value 0 or 1. So σ is well defined.

By sequential compactness, the countable set $\sigma(T) \subseteq [0,1]$ contains a convergent sequence, say $(\sigma(r_k))_{k \in \mathbb{N}}$ with $r_k \in T$.

Write $x = \lim \sigma(r_k)$. We will show that for every $N \in \mathbb{N}$ there is $K \in \mathbb{N}$ such that for all $k \geq K$ the initial parts satisfy $r_k \restriction N = r_K \restriction N$. This will prove the existence of an infinite path in T, since the length of $(r_k)_k$ must go to infinity. (There are only finitely many paths of a given length in T).

Fix $N \in \mathbb{N}$ and let $k \in \mathbb{N}$ be chosen such that $|\sigma(r_k) - x| < \frac{1}{2} - \frac{1}{4^{n+1}}$ for all $k \geq K$. Observe that for all $k \geq K$, we have

$$|\sigma(r_n \restriction N) - \sigma(r_k)| = \sum_{n=N+1}^{l(r_k)} \frac{3}{4^{n+1}} r_k(n) \leq \frac{3}{4^{N+2}} \sum_{n=0}^{\infty} \frac{1}{4^k}$$

$$= \frac{3}{4^{N+2}} \frac{4}{3} = \frac{1}{4^{N+1}},$$

where $l(r_k)$ denotes the length of the sequence r_k.

We find that

$$\begin{aligned} &|\sigma(r_k \restriction N) - \sigma(r_K \restriction N)| \\ &\quad \leq |\sigma(r_k \restriction N) - \sigma(r_k)| + |\sigma(r_k) - x| + |x - \sigma(r_K)| + |\sigma(r_k) - \sigma(r_K \restriction N)| \\ &\qquad < \frac{1}{4^{N+1}} + \frac{1}{2}\frac{1}{4^{N+1}} + \frac{1}{2}\frac{1}{4^{N+1}} + \frac{1}{4^{N+1}} = \frac{3}{4^{N+1}}. \end{aligned}$$

Since at most every third interval of the form $[\frac{n}{4^{N+1}}, \frac{n+1}{4^{N+1}}]$ contains one of the number $\sigma(r_k \restriction N), \sigma(r_K \restriction N)$, this shows their equality.

□

Inspection of proof shows that no induction has been used. In other words, we have found that we can prove the equivalence of Weak Kőnig's Lemma and sequential compactness of $[0,1]$ in an axiomatic system which is less than the ZF axiomatic system. In Chapter 2 we have shown that we can prove Weak Kőnig's Lemma in ZF alone, but it is possible to prove it in an axiomatic system which is even weaker than ZF, but this won't be proven in this paper.

4.2 Weak axiom of choice for countable families of finite sets and Kőnig's Lemma

In this section we are going encapsulate some results from Chapter 3 and Chapter 4 into a new corollary, which will tell us more about the relation between Kőnig's Lemma and the Weak Axiom of Choice for countable families of finite sets.

We are going to expand Corollary 1 by proving that the implication in Theorem 10 is truly an equivalence.

Recall Theorem 10, which states that the weak axiom of choice for countable families of finite sets implies that every locally finite, infinite graph has an infinite path. We will now prove we implication the other way around.

Theorem 12. *Assume that every locally finite, infinite graph has an infinite path. Then every countable family of finite sets has a choice function.*

Proof. Let $(S_n)_{n\geq 1}$ be a countable family of finite, non-empty sets.

Define $V(T) = \bigcup_{n\geq 0} S_0 \times S_1 \times ... \times S_n$ and declare two vertices $(s_1, ..., s_n)$ and $(t_1, ..., t_m)$ adjacent in T if $|n - m| = 1$ and

$$\begin{cases} (s_1, ..., s_{n-1}) = (t_1, ..., t_m), & \text{if } n-1 = m \\ (t_1, ..., t_{m-1}) = (s_1, ..., s_n), & \text{if } m-1 = n. \end{cases}$$

In this way we construct a tree as in Figure 6.

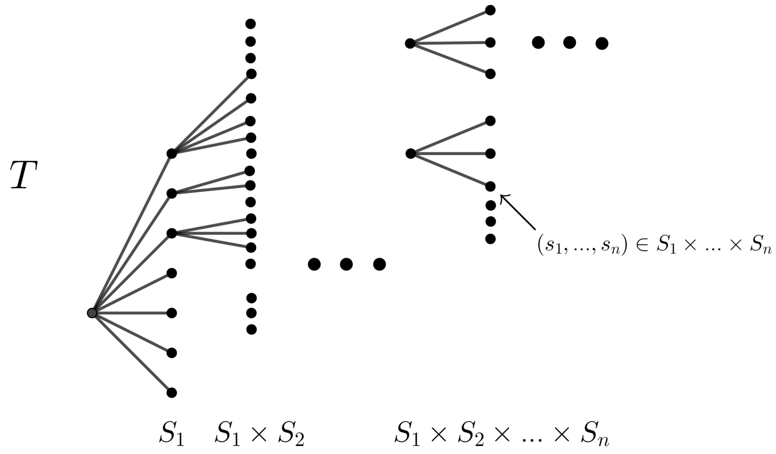


Figure 10: *This is a graphical illustration of the tree construction. Note that we want to construct the tree in an non-trivial way. If we had loops in the tree, then the existence of an infinite path would be trivial, therefore the tree is loop-free.*

We want to find a function $f : \mathbb{N} \to \bigcup_{n\in N} S_n$ such that $f(n) \in S_n$.

Since $S_n \neq \emptyset$ for all $n \geq 1$, the graph T is infinite. Note that T is connected since its node $(s_1, ..., s_n)$ is connected to the empty sequence by

$$\varnothing \sim (s_1) \sim (s_1, s_2) \sim ... \sim (s_1, ..., s_{n-1}) \sim (s_1, ..., s_n).$$

Also, T is locally finite, since there are precisely $|S_{n+1}| + 1$ many neighbours of $(s_1, ..., s_n)$.

So there is an infinite path $(v_0, v_1, v_2, ...)$ in T.

We may assume that $v_0 = \varnothing$. We define $f(n) \in S_n$ by the formula

$$v_n = (v_{n-1}, f(n)) = (f(1), f(2), ..., f(n))$$

for $n \geq 1$.

This is well defined, since $v_n \in S_1 \times ... \times S_n$. □

Remark. *Note that the vertices of the tree constructed in the proof of Theorem 12 are not numerical sequences, as used in Definition 4 of Section 1. This is a crucial difference.*

We have then shown that the following statements are equivalent:

1. Every locally finite, infinite graph has an infinite path.
2. Every countable family of finite non-empty sets has a choice function.

Now recall that we have already seen the second statement in Corollary 1. We can now summarize the result in the following corollary.

Corollary 2. *In ZF the following statements are equivalent.*

1. *Every countable family of non-empty finite sets has a choice function.*
2. *Every infinite, locally finite and connected graph has an infinite path.*
3. *The union of countable many finite sets is countable.*

Note that for this corollary to be valid, the axiomatic system ZF is necessary, and therefore ZFC is not necessary. In fact, the Axiom of Choice is not necessary to prove these equivalences.

We have then proved that Kőnig's Lemma is truly equivalent to the Weak Axiom of Choice for finite sets.

Results

We have now uncovered the complexity hidden behind the different versions of Kőnig's Lemma and unlike our oak tree, we have come to an end.

In the first part of this paper we have proven Weak Kőnig's Lemma for binary trees and Weak Kőnig's Lemma for k-ary trees in ZF. We have also proved that Weak Kőnig's Lemma and the sequential compactness of $[0, 1]$ are equivalent. This equivalence is implied by an axiomatic system which is weaker than ZF, since it has been proved without induction nor Axiom of Choice.

In the second part of this paper we have learnt about the Axiom of Choice and seen that it is not necessary to prove the stronger version of Kőnig's Lemma either. In fact, ZF together with a weak version of the Axiom of Choice, namely the Weak Axiom of Choice for countable families of finite sets, implies Kőnig's Lemma. Also, we have seen that this version of Kőnig's Lemma is equivalent to axiom of choice for countable families of finite sets. At last, we have shown that Kőnig's Lemma is also equivalent to the theorem which states that the countable union of finite sets is countable. Both these equivalences with Kőnig's Lemma have been proved in ZF.

References

[1] Richard Kaye. *The Mathematics of Logic.* (2007)
ISBN 9780521708777.

[2] Denis R. Hirschfeldt. *Slicing the truth: On Computability Theoretic and Reverse Mathematical Analysis of Combinatorial Principles.* (2014)
ISBN 9789814612616.

[3] Reinhard Diestel. *Graph Theory.* (2000)
ISBN 0387989765.

[4] Asaf Karagila. *(Long) Detailed Proof of Kőnig's Lemma (Explicit, Down to Axiom of Choice).* Math Stack Exchange answer. (April 2020, retrieved May 2021)
`https://math.stackexchange.com/a/3615484/2439`

[5] John L. Bell. *The Axiom of Choice.* The Stanford Encyclopedia of Philosophy. (2021)
`https://plato.stanford.edu/entries/axiom-choice/`

[6] Ralf Schindler. *Set Theory.* (2014)
ISBN 9783319067254

Wikipedia and WolframAlpha are general sources used for definitions (note that these have been adjusted to the papers terminology).