Sierpiński's Theorem in Coq

The Generalized Continuum Hypothesis Implies the Axiom of Choice

Felix Rech

June 15, 2019

The *generalized continuum hypothesis* in Zermelo-Fraenkel set theory states that no cardinality lies strictly between that of any infinite set and its power set. The *axiom of choice* states that for every set A of nonempty sets, there is a choice function that maps each $B \in A$ to an element of B. Both are independent of the Zermelo-Fraenkel axioms but Sierpiński's theorem [5] states that the generalized continuum hypothesis implies the axiom of choice. We present a proof of that result in a second-order axiomatization of Zermelo-Fraenkel set theory in the type theory of Coq. Large parts of the proof depend on the axiom of excluded middle. For convenience, we also assume function extensionality. All presented results are formalized.

First, we introduce the axioms of the set theory and some basic constructions (Section 1). We continue with important facts about the cardinality of sets and classes (Section 2). As a key step of our main theorem, we will show that the generalized continuum hypothesis implies the *well-ordering theorem*: Every set has a well-order. We prepare this step with some observations about ordinals as unique representatives of isomorphism classes of well-orders (Section 3). The core of the proof can be found in Sections 4 and 5.

1 Elementary Set Theory

We assume a type $\mathscr{S}$ of objects that we call **sets**. We think of those as collections of other sets and assume a binary relation that tells us if one set is an **element** of another.

General collections of sets that are defined by an arbitrary predicate of type $\mathscr{S} \to \mathbb{P}$ are called **classes**. By a notation of the form $\{x \mid P(x)\}$, we denote the class of all sets x that satisfy the property $P : \mathscr{S} \to \mathbb{P}$. By $x \in A$, we express that a set x is an element of a class A. Furthermore, we use other common set-theoretic notations without explicit introduction. We identify every set with the class of its elements, which is justified by the axiom of extensionality (1.1) below. Moreover, we identify every class A with the refinement type $\sum y.\, y \in A$ of its elements. For example, when we talk about functions from a class A to a class B, we mean functions that take a set that is an element of A as input and return a set that is an element of B as output.

1.1 Basic Axioms

We introduce now our axiomatization of second-order Zermelo-Fraenkel set theory and some basic set constructions. The axiomatization is inspired by that of Kirst and Smolka [4] which in turn is close to that of Barras [1]. Some of the definitions and proofs in this and the following sections are adapted from Smullyan and Fitting [6].

1.1 Axiom of Extensionality. *Two sets are equal if they contain the same elements.*

1.2 Axiom of Empty Set. *The empty class is a set.*

1.3 Definition. For class A, we define the **union** of A as

$$\bigcup A := \{x \mid \exists B \in A, x \in B\}.$$

1.4 Axiom of Union. *The union of every set is a set.*

1.5 Definition. We say that a class A is **subclass** of a class B if all elements of A are also in B. Sometimes, we write this as $A \subseteq B$. If A is a set then we call it a **subset**.

1.6 Definition. We define the power class of a class A as the class of all subsets

$$\mathcal{P}(A) := \{x \mid x \subseteq A\}.$$

1.7 Axiom of Power Set. *If A is a set then the power class $\mathcal{P}(A)$ is also a set. We call it the power set of A.*

1.8 Axiom of Replacement. *Let A be a set and $R : \mathscr{S} \to \mathscr{S} \to \mathbb{P}$ be a functional relation on all sets, that is, for every set x, there is at most one set y that satisfies $R(x,y)$. In this context, the class*

$$\{y \mid \exists x \in A.\, R(x,y)\}$$

is a set.

This is our first axiom that deviates significantly from the standard axiomatization of Zermelo-Fraenkel set theory in first-order logic because we can use all the power of Coq to define the relation R, not just first-order statements. Since the relation does not need to be total, we get the principle of separation as an immediate consequence:

1.9 Fact (Separation)**.** *Given a set A and a predicate $P : A \to \mathbb{P}$, the class*

$$\{x \in A \mid P(x)\}$$

is a set.

The next axiom asserts that the element relation on sets is well founded, which gives us a useful induction principle. First-order axiomatizations typically use the axiom of *regularity* instead that is equivalent under excluded middle but seems less natural. As preparation for the axiom, we define what it means for a relation to be well-founded.

1.10 Definition. Fix a type A and a relation $R : A \to A \to \mathbb{P}$ on A. We define inductively that an element $x : A$ is **accessible** through R if all its *predecessors* are accessible, that is, all $y : A$ with $R(y, x)$. We say that R is **well-founded** if all elements of A are accessible through R.

1.11 Axiom of Foundation. *The element relation on the type of sets is well-founded.*

This implies the principle of well-founded induction:

1.12 Fact (Well-founded induction)**.** *If we want to show a property for all sets, it suffices to show the property for a fixed but arbitrary set under the assumption that it holds on all of its elements.*

Of course, an equivalent statement holds on all types with well-founded relations. We highlight one significant consequence that we need later.

1.13 Fact. *No set contains itself.*

In Coq, there is a difference between the propositional inhabitation of a type and the ability to give a concrete element. In general, the first does not imply the second. When talking about sets, however, if we have the propositional existence of a *unique* set that satisfies a given property then we can also obtain such a set from the axioms:

1.14 Fact (Description)**.** *Fix a class A and a predicate $P : A \to \mathbb{P}$ on A. If there is propositionally a unique element of A that satisfies P, then we can construct this element explicitly from the axioms of sets.*

In addition to the constructions defined so far, we use the common constructions of sets like union, intersection and set difference and write finite sets in the form $\{x_1, \ldots, x_n\}$ by the list of their elements. By $f[A]$, we denote the image of a class A under a function f.

1.2 Numerals

Until now, we have no way to construct infinite sets. The typical axiom of infinity in first-order logic would state that there is a set that contains the empty set and is closed under a suitable successor function. Again, we use the features of Coq for a similar but slightly more natural formulation. We will define an encoding of natural numbers as sets. We call the encoded numbers *numerals* and assume axiomatically that the class of numerals is a set. We encode 0 as the empty set and use the following successor function.

1.15 Definition. We define the **successor function** by $\sigma(x) := x \cup \{x\}$ for all sets x.

We will not only use this for numerals but also for ordinals in Section 3.

1.16 Definition. We define an encoding of natural numbers by

$$\begin{aligned} \mathit{encode}_{\mathbb{N}}(0) &:= \emptyset \\ \mathit{encode}_{\mathbb{N}}(1+n) &:= \sigma(\mathit{encode}_{\mathbb{N}}(n)). \end{aligned}$$

The sets in the image of this encoding are called **numerals**. From here on, we leave the encoding implicit and write the natural number in place of its encoding.

Note that for every natural number n, the corresponding numeral contains exactly n elements. This allows us to use numerals not just as representatives for natural numbers but also as canonical sets of any finite cardinality.

1.17 Axiom of Infinity. *The class of numerals is a set.*

We list the three most important facts to work with numerals:

1.18 Fact. *The encoding of natural numbers is injective.*

This is an important ingredient for the following observations and makes sure that the set of numerals is actually infinite.

1.19 Fact (Recursion)**.** *If we want to define a function $f : \mathbb{N} \to A$ for any class A, it suffices to give a base value $x \in A$ and a step function $g : \mathbb{N} \to A \to A$. The resulting function satisfies the computation rules*

$$\begin{aligned} f(\emptyset) &= x \\ f(\sigma(n)) &= g(n, f(n)). \end{aligned}$$

In the statement of this fact, it is important that A needs to be a class, not an arbitrary type. For arbitrary types, this does not seem to be true since we cannot easily define a function that maps numerals back to their corresponding natural numbers.

1.20 Fact (Induction)**.** *If we want to prove a property P on all numerals, it suffices to prove $P(\emptyset)$ and to show that $P(n)$ implies $P(\sigma(n))$ for all numerals n.*

1.3 Cartesian Product

In this subsection, we introduce the Cartesian product of two sets which is a set that consists of encoded pairs. We start with the definition of the encoding function.

1.21 Definition. We define an encoding of pairs of sets by $\mathit{encode}_{\times}(x, y) \coloneqq \{\{x\}, \{x, y\}\}$. Like the encoding of natural numbers, we usually leave this implicit. We define the **Cartesian product** of two sets A and B as the set of all encoded pairs

$$A \times B \coloneqq \{(x, y) \mid x \in A \wedge y \in B\}.$$

1.22 Fact. *The encoding of pairs is injective.*

1.23 Definition. Fix two sets A and B and an element $z \in A \times B$. We define the **left projection** $\pi_1(z)$ by description (Fact 1.14) as the unique $x \in A$ such that there is a $y \in B$ with $z = (x, y)$. In a similar way, we define the **right projection** $\pi_2(z) \in B$.

We conclude with a remark that may not seem useful at the moment but becomes crucial in Section 4.

1.24 Remark. For all sets A, it holds that $A \times A \subseteq \mathcal{P}^2(A)$.

1.4 Disjoint Union

When working with sets, it is sometimes useful to assume without loss of generality that two sets are disjoint before one builds the union of both. In this subsection, we present one way to make that more precise. Our solution is analogous to sum types in type theory.

1.25 Definition. We define the **left injection** of a set x by $\mathit{inj}_1(x) := (0, x)$ and the **right injection** by $\mathit{inj}_2(x) := (1, x)$. For given sets A and B, the set $A + B := \mathit{inj}_1[A] \cup \mathit{inj}_2[B]$ is called the **disjoint union** of A and B.

The important point of this definition is that inj_1 and inj_2 are injective and their images are always disjoint. Like for numerals, we have an eliminator and a form of induction principle:

1.26 Fact (Match)**.** *If we want to define a function $f : A + B \to C$ for any sets A, B and C then it suffices to give a function $f_A : A \to C$ and a function $f_B : B \to C$. The resulting f satisfies the computation rules*

$$\begin{aligned} f(\mathit{inj}_1(x)) &= f_A(x) \\ f(\mathit{inj}_2(y)) &= f_B(y). \end{aligned}$$

1.27 Fact (Case analysis)**.** *If we want to prove a property $P : A + B \to \mathbb{P}$ on all elements of the disjoint union $A + B$ of two sets A and B then it suffices to prove $P(\mathit{inj}_1(x))$ for all $x \in A$ and $P(\mathit{inj}_2(y))$ for all $y \in B$.*

2 Cardinality

In this section, we define the types of *bijections*, *injections* and *surjections* as a means to compare types by size. In addition, we provide some facts about the size of classes and sets.

2.1 Definition. A **bijection** from a type A to a type B is a function $f : A \to B$ for which we have a function $g : B \to A$ that satisfies $f(g(y)) = y$ for all $y \in B$ and $g(f(x)) = x$ for all $x \in A$.

If the type of bijections from A to B is inhabited then we say that A and B have the same **cardinality** or that they are **equipotent**, abbreviated as $A \sim B$.

2.2 Fact. *Bijections satisfy three basic properties:*

1. *The identity function on any type is a bijection.*

2. *The composition of two bijections is a bijection.*

3. *Every bijection has an inverse bijection.*

This implies that equipotency is an equivalence relation.

2.3 Definition. An **injection** from a type A to a type B is a function $f : A \to B$ such that for all $x, x' \in A$ with $f(x) = f(x')$, it follows that $x = x'$. We denote the type of injections from A to B by $A \hookrightarrow B$.

If the type $A \hookrightarrow B$ is inhabited then we say that A has smaller or equal cardinality or just that it is smaller or equal to B, abbreviated as $A \leq B$.

2.4 Fact. *Basic properties of injections are:*

1. *The identity function on any type is an injection.*
2. *The composition of two injections is an injection.*

2.5 Lemma. *Every bijection is also an injection.*

We continue with an important fact about power sets.

2.6 Fact. $\mathcal{P}(A + B) \sim \mathcal{P}(A) \times \mathcal{P}(B)$.

Proof. We define functions in both directions by

$$\begin{aligned} f &: \mathcal{P}(A + B) \to \mathcal{P}(A) \times \mathcal{P}(B) \\ f(C) &:= (\{x \in A \mid \mathit{inj}_1(x) \in C\}, \{y \in B \mid \mathit{inj}_2(y) \in C\}) \end{aligned}$$

$$\begin{aligned} g &: \mathcal{P}(A) \times \mathcal{P}(B) \to \mathcal{P}(A + B) \\ g(A', B') &:= \mathit{inj}_1[A'] \cup \mathit{inj}_2[B']. \end{aligned}$$

With extensionality, we can show that those are mutually inverse. □

Our next major goal is to describe a condition under which a set A satisfies the equation $A \sim A + A$ (Fact 2.6). Under the axiom of choice, this holds for all infinite sets, but without it, we will need to be more specific. We divide our proof into multiple steps. First however, we have to define what it means for a set to be infinite. The most common definition states that a set is infinite if it is not equipotent to a numeral. For our purpose, the following is more convenient.

2.7 Definition. A set A is **Dedekind-infinite** or just **infinite** if $\mathbb{N} \leq A$.

Under the axiom of choice, both statements are equivalent, but without it, ours is stronger.

Now, we want to show that every infinite set A satisfies $A \sim 1 + A$. First, we show this property for the set of numerals, then we use the fact that every infinite set is equipotent to a disjoint union of the form $\mathbb{N} + \_$ to generalize the result.

2.8 Lemma. $\mathbb{N} \sim 1 + \mathbb{N}$.

Proof. We define functions in both directions by

$$\begin{aligned} f &: \mathbb{N} \to 1 + \mathbb{N} & g &: 1 + \mathbb{N} \to \mathbb{N} \\ f(\emptyset) &:= \mathit{inj}_1(0) & g(\mathit{inj}_1(\_)) &:= \emptyset \\ f(\sigma(n)) &:= \mathit{inj}_2(n) & g(\mathit{inj}_2(n)) &:= \sigma(n). \end{aligned}$$

By induction on the numerals and the disjoint union respectively, we can show that those are mutually inverse. □

2.9 Lemma. *If $B \subseteq A$ then $A = B \cup (A \setminus B)$.*

2.10 Corollary. *If $B \subseteq A$ then $A \sim B + (A \setminus B)$.*

Proof. This follows with $B \cup (A \setminus B) \sim B + (A \setminus B)$ which holds since B and $A \setminus B$ are disjoint. □

2.11 Fact. *If A is an infinite set then $A \sim 1 + A$.*

Proof. Let $B \subseteq A$ be the image of the injection from $\mathbb{N}$ to A. Then $B \sim \mathbb{N}$ which implies $B \sim 1 + B$ (Lemma 2.8). We conclude

$$\begin{aligned} A &\sim B + (A \setminus B) \\ &\sim (1 + B) + (A \setminus B) \\ &\sim 1 + (B + (A \setminus B)) \\ &\sim 1 + A \end{aligned}$$

by Corollary 2.10 and associativity of the disjoint union with respect to cardinality. □

Finally, we get our desired result:

2.12 Fact. *If A is an infinite set then $\mathcal{P}(A) \sim \mathcal{P}(A) + \mathcal{P}(A)$.*

Proof. We deduce

$$\begin{aligned} \mathcal{P}(A) &\sim \mathcal{P}(1 + A) \\ &\sim \mathcal{P}(1) \times \mathcal{P}(A) \\ &= 2 \times \mathcal{P}(A) \\ &= \mathcal{P}(A) + \mathcal{P}(A) \end{aligned}$$

with Fact 2.11 and Fact 2.6. □

We also define the concept of *surjections* with the exclusive goal to formulate a variant of *Cantor's theorem* that we need in Section 5.

2.13 Definition. A function f from a type A to a type B is a **surjection** if for all $y \in B$ there is an $x \in A$ such that $f(x) = y$.

2.14 Cantor's Theorem. *There is no surjection from any set into its power set.*

We end the section with an observation that will be useful to prove that a certain class is actually a set.

2.15 Fact. *Every class that is smaller than a set is itself also a set.*

Proof. Fix an arbitrary class A and a set B that is larger or equal to A. By definition, we have an injection $f : A \hookrightarrow B$. The class A is the image of f^{-1} and thus a set by the axiom of replacement (1.8). □

3 Orderings

We start this section with standard definitions of *orders*, *well-orders* and *order isomorphisms*. In the second half, we introduce *ordinals* as unique representatives for isomorphism classes of well-orders.

3.1 Definition. A **strict total order**, or just **order**, on a type A is a binary relation $\_ < \_ : A \to A \to \mathbb{P}$, that satisfies three properties:

Transitivity: If $x < y$ and $y < z$ then $x < z$ for all $x, y, z : A$.

Irreflexivity: It never holds that $x < x$ for any $x : A$.

Trichotomy: All $x, y : A$ satisfy $x < y$, $x = y$ or $x > y$.

An **ordered type** is a type that has an order.

3.2 Definition. A **well-order** is an order that is well-founded (Definition 1.10).

3.3 Definition. An **order isomorphism** from an ordered type A to an ordered type B is a bijection $f : A \to B$ that preserves the order in both directions:

$$x < y \leftrightarrow f(x) < f(y).$$

If the type of such isomorphisms is inhabited, we say that A as ordered type is **order isomorphic** to B, abbreviated as $A \simeq B$.

3.4 Fact. *There are three basic laws of order isomorphisms:*

1. *The identity function on any ordered type is an order isomorphism.*

2. *The composition of any two order isomorphisms is an order isomorphism.*

3. *Every order isomorphism has an inverse order isomorphism.*

This implies that isomorphism between ordered types is an equivalence relation.

3.1 Ordinals

Next, we define *ordinals* which we use as sets that uniquely represent well orders up to isomorphism. They can also be seen as a generalization of numerals that allows us to count past infinity. Our definition is quite unconventional and analogous to the characterization of the cumulative hierarchy by Kirst and Smolka [4].

3.5 Definition. We define the class of **ordinals** inductively by the rules that

1. the successor $\sigma(\alpha)$ of an ordinal α is an ordinal and

2. the union $\bigcup A$ of a set of ordinals A is an ordinal.

The successor function σ in this case is the same that we used for the definition of numerals (1.15). Note that the empty set is also an ordinal, which implies that all numerals are ordinals. As announced before, ordinals represent well-orders:

3.6 Fact. *Every ordinal is well-ordered by the element relation.*

As representatives, they are unique by the following fact.

3.7 Fact. *Isomorphic ordinals are equal.*

We will also need a characterization of ordinals that differs from the definition. It makes use of the concept of transitive sets:

3.8 Definition. We call a set A **transitive** if for all sets x and y, whenever $x \in A$ and $y \in A$ then $x \in A$ — in other words, if every element is also a subset.

Note that this is similar to transitivity of the element relation. Now we can formulate the characterization of ordinals.

3.9 Fact. *A set is an ordinal if and only if it is transitive and every element is an ordinal.*

4 Hartogs Number

Our goal for this section is to assign to every set an ordinal, the so-called *Hartogs number.* If we assumed stronger axioms, such as the generalized continuum hypothesis (Definition 5.1), then we could show that the Hartogs number of a set is always greater than that set. Without further axioms, however, the closest thing that we can achieve is to show that it is at least not smaller or equal in cardinality. As a by-product of the construction, we will obtain an upper bound on the Hartogs number that will be crucial in the next section.

Hartogs introduced the Hartogs number in 1915 [3], also with the goal to derive the well-ordering theorem. The upper bound was first proved by Sierpiński [5]. We deviate slightly from those works with the intention to make the construction a bit clearer.

4.1 Definition. We define the **Hartogs number** of a given set A as

$$\aleph(A) := \{\alpha \in \text{ordinals} \mid \alpha \leq A\}.$$

If it turns out that the Hartogs number is an ordinal then $\aleph(A) \not\leq A$ will follow immediately from this definition, because otherwise, the Hartogs number would contain itself. We proceed in three steps:

1. We show that $\aleph(A) \leq \mathcal{P}^6(A)$ which implies that the Hartogs number is a set.

2. We show that the Hartogs number is an ordinal.

3. We conclude that $\aleph(A) \not\leq A$.

4.2 Fact. *The Hartogs number of every set A satisfies the upper bound*

$$\aleph(A) \leq \mathcal{P}^6(A).$$

Proof. By the general upper bound on the Cartesian product (Remark 1.24), we have

$$\begin{aligned} \mathcal{P}\left(\mathcal{P}(A) \times \mathcal{P}(A \times A)\right) &\subseteq \mathcal{P}\left(\mathcal{P}(A) \times \mathcal{P}^3(A)\right) \\ &\leq \mathcal{P}\left(\mathcal{P}^3(A) \times \mathcal{P}^3(A)\right) \\ &\subseteq \mathcal{P}^6(A). \end{aligned}$$

With transitivity, it suffices to define an injection

$$\begin{gathered} f : \aleph(A) \hookrightarrow \mathcal{P}(\mathcal{P}(A) \times \mathcal{P}(A \times A)) \\ f(\alpha) \coloneqq \{x \in \mathcal{P}(A) \times \mathcal{P}(A \times A) \mid x \simeq \alpha\}, \end{gathered}$$

where we treat every $x \in \mathcal{P}(A) \times \mathcal{P}(A \times A)$ as a subset of A with a relation on it that can satisfy $x \simeq \alpha$ if the relation is an order. To see that f is injective, fix two ordinals $\alpha, \beta \in \aleph(A)$ with $f(\alpha) = f(\beta)$. By definition of the Hartogs number, there is an injection $\alpha \hookrightarrow A$. We embed the order on α along this injection to obtain an $x \in \mathcal{P}(A) \times \mathcal{P}(A \times A)$. Note that $x \simeq \alpha$. Therefore $x \in f(\alpha) = f(\beta)$ and hence, $x \simeq \beta$ by definition of f. Together, we have $\alpha \simeq x \simeq \beta$ which implies $\alpha = \beta$ since isomorphic ordinals are equal (Fact 3.7). □

As mentioned before, this upper bound is very generous. We could use a different encoding of ordered subsets to get the bound down to $\mathcal{P}^3(A)$. For our purpose, however, this does not matter.

4.3 Corollary. *The Hartogs number of a set is also a set.*

Proof. This holds since a class that is smaller than a set, is itself also a set (Fact 2.15). □

4.4 Fact. *The Hartogs number of a set is an ordinal.*

Proof. Fix a set A. We know that the Hartogs number $\aleph(A)$ contains only ordinals by definition and that it is a set by the previous corollary. With the characterization of ordinals in Fact 3.9 it suffices to show that it is transitive. Fix two ordinals x and y with $y \in x \in \aleph(A)$. Our goal is to prove that $y \in \aleph(A)$. By definition of the Hartogs number, $x \in \aleph(A)$ is an ordinal that satisfies $x \leq A$ and we have to show those properties for y.

1. As element of an ordinal, y is one too (Fact 3.9).

2. As element of an ordinal, y is also a subset (Fact 3.9). Hence, $y \subseteq x \leq \aleph(A)$. □

4.5 Hartogs' Theorem. *For all sets A, we have $\aleph(A) \not\leq A$.*

Proof. Assume that $\aleph(A) \leq A$. By definition, we get that $\aleph(A) \in \aleph(A)$, which is a contradiction (Fact 1.13). □

5 Sierpiński's Theorem

This section focuses on the statement and the proof of our main theorem. Our proof is an adaption of that from Gillman [2] which in turn is a modified form of that from Sierpiński [5]. The theorem relates the following two important statements that are neither provable nor refutable under the axioms that we assumed so far.

5.1 Definition. The **generalized continuum hypothesis** states that there is no set strictly between any *infinite* set A and the power set of A. In other words, for all infinite sets A and B such that $A \leq B \leq \mathcal{P}(A)$, we have either $A \sim B$ or $B \sim \mathcal{P}(A)$.

The *specialized* continuum hypothesis, in contrast, talks only about $\mathbb{N}$ instead of an arbitrary set A.

5.2 Definition. The **axiom of choice** states that for every family $F : I \to S$ of inhabited sets over an index set I, there is propositionally a choice function that maps every $i \in I$ to an element of $F(i)$.

Our goal is to show that the generalized continuum hypothesis implies the axiom of choice. During our proof, we will use the following statement as an intermediate step.

5.3 Definition. The **well-ordering theorem** states that every set has a well-order.

5.4 Theorem. *The well-ordering theorem implies the axiom of choice.*

We continue with a rather technical lemma.

5.5 Lemma. *All sets A and B such that $A \sim A + A$ and $A + B \sim \mathcal{P}(A)$ satisfy the inequality $\mathcal{P}(A) \leq B$.*

Proof. Without loss of generality, we assume that A and B are disjoint. From the assumptions and Fact 2.6, we deduce

$$A \cup B \sim A + B \sim \mathcal{P}(A) \sim \mathcal{P}(A + A) \sim \mathcal{P}(A) \times \mathcal{P}(A).$$

Hence, there is a bijection $f : A \cup B \to \mathcal{P}(A) \times \mathcal{P}(A)$. We compose f with the first projection to obtain a function $\pi_1 \circ f : A \to \mathcal{P}(A)$. By Cantor's theorem (2.14), this cannot be a surjection, that is, there is an $A_1 \in \mathcal{P}(A)$ that is not the first component of any $f(x)$ for $x \in A$. Conversely, for all $A_2 \in \mathcal{P}(A)$, the preimage $f^{-1}(A_1, A_2)$ must come from B. This leads us to the conclusion that the injection

$$\begin{aligned} \mathcal{P}(A) &\hookrightarrow A \cup B \\ A_2 &\mapsto f^{-1}(A_1, A_2) \end{aligned}$$

is actually an injection of type $\mathcal{P}(A) \hookrightarrow B$. □

Under the axiom of choice, we could replace the condition that $A \sim A + A$ by the condition that A has to be infinite. Without the axiom of choice however, that does not seem to be strong enough.

5.6 Sierpiński's Theorem. *The generalized continuum hypothesis implies the axiom of choice.*

Proof. We assume the generalized continuum hypothesis and show the well-ordering theorem which implies the axiom of choice (Theorem 5.4). Let us fix a set A that we want to well-order. It suffices to find some ordinal greater than or equal to A.

Since the Hartogs number of A is at least not smaller, it seems like a reasonable candidate. However, it will be useful to consider instead the Hartogs number of some larger set B with the property that $\mathcal{P}^i(B)$ is infinite and satisfies

$$\mathcal{P}^i(B) \sim \mathcal{P}^i(B) + \mathcal{P}^i(B) \tag{1}$$

for all natural numbers i. By Lemma 2.12 and the definition of infinity (2.7), $B \coloneqq \mathcal{P}(A+\mathbb{N})$ satisfies this property.

We know that $\mathcal{P}^6(B)$ is an upper bound on the Hartogs number of B (Fact 4.2). We show by induction that, in general, for all natural numbers i, the upper bound $\aleph(B) \leq \mathcal{P}^i(B)$ implies the existence of a well ordering on A. The base case is simple since the assumption that $\aleph(B) \leq \mathcal{P}^0(B) = B$, contradicts Hartogs' theorem (4.5).

For the inductive step, we fix a natural number i and assume the upper bound $\aleph(B) \leq \mathcal{P}^{i+1}(B)$. We need an occasion to apply the generalized continuum hypothesis, so we deduce

$$\mathcal{P}^i(B) \leq \aleph(B) + \mathcal{P}^i(B) \leq \mathcal{P}^{i+1}(B),$$

where the first inequality is trivial and the second one follows with Equation 1 as

$$\begin{aligned}\aleph(B) + \mathcal{P}^i(B) &\leq \mathcal{P}^{i+1}(B) + \mathcal{P}^i(B)\\ &\leq \mathcal{P}^{i+1}(B) + \mathcal{P}^{i+1}(B)\\ &\leq \mathcal{P}^{i+1}(B).\end{aligned}$$

By the generalized continuum hypothesis, this leads to two possible cases:

Case 1. If $\aleph(B) + \mathcal{P}^i(B) \sim \mathcal{P}^i(B)$ then we can immediately conclude that $\aleph(B) \leq \mathcal{P}^i(B)$. This is exactly the condition for the inductive hypothesis which proves our goal.

Case 2. If $\aleph(B) + \mathcal{P}^{i(B)} \sim \mathcal{P}^{i+1}(B)$ then $\mathcal{P}^{i+1}(B) \leq \aleph(B)$ by Lemma 5.5 with Equation 1. Hence, $A \leq B \leq \mathcal{P}^{i+1}(B) \leq \aleph(B)$ and since the Hartogs number is an ordinal and therefore has a well-order, we conclude that A has a well-order too. □

References

[1] Bruno Barras. “Sets in Coq, Coq in sets”. In: *Journal of Formalized Reasoning* 3.1 (2010), pp. 29–48.

[2] Leonard Gillman. “Two classical surprises concerning the axiom of choice and the continuum hypothesis”. In: *The American Mathematical Monthly* 109.6 (2002), pp. 544–553.

[3] Friedrich Hartogs. “Über das Problem der Wohlordnung”. In: *Mathematische Annalen* 76.4 (1915), pp. 438–443.

[4] Dominik Kirst and Gert Smolka. “Categoricity results for second-order ZF in dependent type theory”. In: *International Conference on Interactive Theorem Proving*. Springer. 2017, pp. 304–318.

[5] Wacław Sierpiński. “L’hypothèse généralisée du continu et l’axiome du choix”. In: *Fundamenta Mathematicae* 1.34 (1947), pp. 1–5.

[6] Raymond M. Smullyan and Melvin Fitting. *Set theory and the continuum problem*. Dover Publications, 2010.