

The Axiom of Choice, the Well Ordering Principle and Zorn's Lemma

Dag Normann *

January 9, 2012

Abstract

In this note we prove the equivalence between the axiom of choice, the well ordering principle and Zorn's lemma, and discuss to some extent how large fragment of ZF we need in order to prove the individual implications.

1 Introduction

This note is a supplementary text to the curriculum of MAT4640 Axiomatic Set Theory. The main textbook is Kunen [1]. We assume that the reader is familiar with the basic results on well orderings from Kunen. Nevertheless, the note should be readable for anyone familiar with some naive set theory, as we will state all properties of well orderings that we will use.

We will discuss which axioms from *Zermelo-Fraenkel Set Theory*, ZF, we need for the proofs we give. Readers only interested in the equivalence results may ignore this.

We use one joint enumeration of lemmas, theorems, definitions and exercises, and another enumeration of the claims. Claims cannot be read out of the context, as they will be integrated parts of proofs of lemmas or theorems.

2 Preliminaries

We will use the concepts *partial ordering* and *total ordering* in the meaning of *less than*, i.e. they are irreflexive. Following Kunen [1], we let a partial or total ordering be a pair consisting of a set and a binary relation, but the domain of this relation does not need to be a subset of the set. The advantage is that we can talk about *initial segments* without changing the relation, only the set. In one proof this convention will be a disadvantage. This is reflected in our definition of an *f*-string. All our relations will be binary, i.e. sets of ordered pairs.

*Department of Mathematics, The University of Oslo, P.O. Box 1053, Blindern N-0316 Oslo, Norway, email dnormann@math.uio.no

Definition 1 Let A be a set and R a relation.

a) R is a *partial ordering* on A if

- $\forall a \in A \neg(aRa)$
- $\forall a, b, c \in A (aRb \wedge bRc \Rightarrow aRc)$.

b) A partial ordering R on A is *total* if we in addition have

$$\forall a, b \in A (aRb \vee a = b \vee bRa).$$

c) If we replace $\forall a \in A \neg(aRa)$ with

$$\forall a, b \in A (aRb \wedge bRa \Leftrightarrow a = b)$$

we say that R is a *reflexive ordering* on A .

d) If R is a partial ordering on a set A , $B \subseteq A$ and $a \in A$ we say that a is an *upper bound* for B if $\forall b \in B (bRa \vee b = a)$.

Definition 2 Let R be a total ordering on a set A .

We say that R is a *well ordering* of A if every nonempty subset $X \subseteq A$ has an R -least element.

We will prove the equivalence, modulo ZF of the following three statements:

AC *The Axiom of Choice*

Let F be a function defined on a set I such that $F(i)$ is a nonempty set for all $i \in I$.

Then there is a *choice function* f defined on I such that $f(i) \in F(i)$ for all $i \in I$.

WO *The Well Ordering Principle*

Let A be a set.

Then there is a wellordering R of A .

ZORN *Zorn's Lemma*

Let A be a set and R a partial ordering on A .

Assume that whenever $B \subseteq A$ is totally ordered by R , then B has an upper bound in A .

Then there is an R -maximal element in A .

Remark 3 Kunen [1] call the wellordering principle for the axiom of choice. This is convenient, since it is often easier to prove the well ordering principle than the axiom of choice for models of set theory as constructed in his book. When we have established the equivalence, we may prove the version we prefer, depending on what is easiest to establish.

Quite often **AC** is formulated as the existence of choice functions for indexed families $\{A_i\}_{i \in I}$ of nonempty sets. This is a matter of notation only, as the statements mathematically mean the same.

3 Zorn's lemma implies the axiom of choice

Let F be a function mapping each I in a set I to a nonempty set $F(i)$. We will use Zorn's lemma, and prove that there is a choice function f for F defined on I .

We will let X be the set of partly successful attempts to construct a choice function. More precisely, a pair $\langle J, g \rangle \in X$ if $J \subseteq I$, g is a function with domain J and $g(j) \in F(j)$ for all $j \in J$.

We let \sqsubseteq be the reflexive ordering on X defined by

$$\langle J_1, g_1 \rangle \sqsubseteq \langle J_2, g_2 \rangle$$

when $J_1 \subseteq J_2$ and g_2 is an extension of g_1 .

We let \sqsubset be the corresponding irreflexive ordering.

Claim 1 *The assumption of Zorn's lemma will hold for $\langle X, \sqsubset \rangle$.*

Proof

Let $\{\langle J_y, g_y \rangle \mid y \in Y\}$ be a totally ordered subset of X and let

$$J = \bigcup \{J_y \mid y \in Y\}.$$

For $j \in J$, let $g(j) = g_y(j)$ whenever $j \in J_y$. The choice of y does not matter, this is where we use that the set is totally ordered.

Clearly, $\langle J, g \rangle$ is an upper bound for $\{\langle J_y, g_y \rangle \mid y \in Y\}$, and this is exactly what we need in order to verify that the claim holds.

End of proof of Claim

Using Zorn's lemma, we let $\langle J, g \rangle$ be a maximal element in $\langle X, \sqsubset \rangle$.

Claim 2 *$J = I$ and g is a choice function for F .*

Proof

Since g is a choice function for F on its domain, it suffices to show that $J = I$. Assume that this is not the case, and that $i \in I \setminus J$. Let $a \in F(i)$ and let f defined on $J \cup \{i\}$ extend g by letting $f(i) = a$.

Then $\langle J \cup \{i\}, f \rangle$ properly extends $\langle J, g \rangle$, contradicting the maximality of the latter. Thus the claim follows by contradiction.

End of proof of claim

As a consequence of the two claims we obtain

Theorem 4 *Zorn's lemma implies the axiom of choice.*

In this argument, we have used the power set axiom, the replacement axiom and the comprehension axiom. We used the power set axiom and the comprehension axiom in singling out those $J \subset I$ with a choice function. We then used the power set axiom to select all choice functions for each J and the replacement axiom again to form the set X .

4 The well ordering principle implies Zorn's lemma

In this proof we will use *transfinite recursion*, a method of construction discussed intimately in MAT4640, and we assume that the reader is familiar with this kind of construction.

We need the axiom of replacement to justify transfinite recursion. The principle idea is that we may define a function by transfinite induction over a well ordering, because if the recursion does not make sense, it will be a least element where the function we aim to construct is undefined, or multiply defined, and this will contradict the uniqueness of recursion.

Let A with the partial ordering R satisfy the assumption in Zorn's lemma, and let \prec be a well ordering of A . The intuition is that we use transfinite recursion over the well ordering \prec to construct a totally ordered subset $B \subseteq A$ securing that no $a \in A$ can be a strict upper bound of B . Then any upper bound for B will be maximal. The precise argument is given below.

Claim 3 *There is a unique subset $B \subseteq A$ satisfying*

$$a \in B \Leftrightarrow \forall b \in A (b \in B \wedge b \prec a \Rightarrow bRa)$$

for all $a \in A$.

Proof

Uniqueness follows from the fact that \prec is a well ordering, there is no minimal element where we may choose if $a \in B$ or not. For the existence, we need the replacement axiom, showing by transfinite induction that for each $c \in A$, there is a B_c satisfying the definition of B up to $\{a \in A \mid a \prec c\}$.

End of proof of claim.

The set B constructed in the above Claim will be totally ordered by R , since for a and b in B , we will have that

$$a \prec b \Leftrightarrow aRb.$$

By the assumption in Zorn's lemma, there will be an upper bound b for B . If b is not maximal, there is an a strictly R -above b . By the construction of B , we will have that $a \in B$ since in particular, a is an upper bound for

$$\{c \in B \mid b \prec a\}.$$

This contradicts that a is a strict upper bound for B , something derived from the assumption that b is not R -maximal. Thus b is R -maximal, and the conclusion of Zorn's lemma is proved.

We then have

Theorem 5 *The well ordering principle implies Zorn's lemma.*

There is no significant use of the power set axiom in this argument, and indeed, this theorem can be proved without the use of the power set axiom.

5 The axiom of choice implies the well ordering principle

In this section we will make use of the following facts about well orderings:

Lemma 6 a) Let $\langle A, R \rangle$ and $\langle B, S \rangle$ be isomorphic well orderings. Then the isomorphism is unique.

b) Let $\langle A, R \rangle$ and $\langle B, S \rangle$ be two well orderings. Then exactly one of the three statements below will hold:

1. $\langle A, R \rangle$ and $\langle B, S \rangle$ are isomorphic.
2. $\langle A, R \rangle$ is isomorphic to a proper initial segment of $\langle B, S \rangle$.
3. $\langle B, S \rangle$ is isomorphic to a proper initial segment of $\langle A, R \rangle$.

In all cases, the isomorphism (and initial segment) is unique.

This Lemma is proved in Kunen [1]. a) is trivial, and b) is fairly easy.

Now, let X be an arbitrary set. We want to use the axiom of choice to show that there must be a well ordering of X .

If you have a bucket full of blueberries, and have the patience to pick one blueberry at random out of the bucket as long as there are blueberries left, you may pick out all the blueberries and lay them in a total ordered row.

Our intuition behind the proof is that if we replace the bucket with X , the blueberries with the elements of X and the random picking with a choice function on the set of nonempty subsets of X , the result of a transfinite picking should be a well ordering of X . We of course have to support this intuition with a mathematically sound argument.

By the axiom of choice, let f be a function such that $f(Y) \in X \setminus Y$ whenever $Y \subset X$ has a non empty complement.

The rest of our construction and proof will be given relative to this chosen f .

Definition 7 An f -string is a pair $\langle A, R \rangle$ where $A \subseteq X$ and R is a well ordering of A such that

$$\forall a \in A (a = f(\{b \in A \mid bRa\}))$$

and R consists only of ordered pairs from A^2 .

Remark 8 If $\langle A, R \rangle$ is a nonempty f -string and a_0 is the R -least element of A , then

$$a_0 = f(\{b \in A \mid bRa_0\}) = f(\emptyset).$$

Thus the least element will be common for all nonempty f -strings. For all f -strings with more elements, the next object will also be common. In fact, we will see that it is only the length of f -strings that may distinguish them, not the order in which the elements come.

Lemma 9 An initial segment of an f -string is an f -string.

Proof

If $\langle B, S \rangle$ is an initial segment of an f -string $\langle A, R \rangle$, and $a \in B$, we have that

$$a = f(\{b \in A \mid bRa\}) = f(\{b \in B \mid bSa\}).$$

The claim follows.

Claim 4 *Let $\langle A, R \rangle$ and $\langle B, S \rangle$ be two f -strings, and let g be an isomorphism between initial segments (not necessarily proper) of the strings. Then $g(a) = a$ for all a in the domain of g .*

Proof

Assume not. Let a be the R -least element in A such that $g(a) \neq a$ and let $b = g(a)$. Then $g(c) = c$ whenever $c \in A$ and cRa .

Since g is an isomorphism, we must have that

$$\{d \in B \mid dSb\} = \{g(c) \mid cRa\} = \{c \in A \mid cRa\}.$$

Since both $\langle A, R \rangle$ and $\langle B, S \rangle$ are f -strings, it follows that

$$a = f(\{c \in A \mid cRa\}) = f(\{d \in B \mid dSb\}) = b.$$

This is in conflict with the assumption.

End of proof of Claim

Claim 5 *Let $\langle A, R \rangle$ and $\langle B, S \rangle$ be two f -strings. Then they are either equal, or one is a proper initial segment of the other.*

Proof

By Lemma 6b), they are either isomorphic or one is isomorphic to a proper initial segment of the other. By claims 3 and 4, the isomorphism will in all cases be an identity function, and the claim follows.

End of proof of Claim

As a consequence of Claim 5 we see that if we order the f -strings by the initial segment ordering, this ordering is total.

Let A be the union of all B such that $\langle B, S \rangle$ is an f -string, and let R be the union of all S such that $\langle B, S \rangle$ is an f -string.

Claim 6 *$\langle A, R \rangle$ is an f -string.*

Proof

We must use that $S \subset B^2$ in order to prove that $\langle A, R \rangle$ is a total ordering. This is left for the reader.

We then see that $\langle A, R \rangle$ is a well ordering: Let $C \subseteq A$ be nonempty.

Then there is an f -string $\langle B, S \rangle$ with $B \cap C \neq \emptyset$.

Let a be the S -least element of $C \cap B$.

Since, by construction, $\langle B, S \rangle$ is an initial segment of $\langle A, R \rangle$, we have that A must be the R -least element of C .

This argument also shows that each proper initial segment of $\langle A, R \rangle$ is an f -string, and it follows that $\langle A, R \rangle$ itself must be an f -string.

End of proof of Claim

Clearly the f -string $\langle A, R \rangle$ is the largest f -string

Claim 7 *The set A constructed above will be equal to X , and thus R is a well ordering of X .*

Proof

Assume that A is a proper subset of X , and let $b = f(A)$. Let $B = A \cup \{b\}$ and let $S = R \cup \{\langle a, b \rangle \mid a \in A\}$.

Then $\langle B, S \rangle$ is also an f -string, so $b \in A$ by construction of A .

This is clearly a contradiction.

End of proof of Claim

We have now given all steps in the proof of

Theorem 10 *The axiom of choice implies the well ordering principle.*

Remark 11 In this proof we have used the power set axiom and the replacement axiom, and the reader will also see a direct application of the union axiom. The union axiom is also used to justify transfinite recursion, and of course there are hidden uses of the union axiom and the comprehension axiom in our proofs. We may view the final f -string as the least fixed point of an inductive definition.

6 Final comments

We have proved that, relative to ZF the three statements **AC**, **WO** and **ZORN** are equivalent. We have proved three of the implications between them, and of course, by transitivity of implication, the three other equivalences will hold.

The proof of **ZORN** from **WO** did not use the power set axiom.

Exercise 12 Give a proof of **AC** from **WO** where you do not use the power set axiom.

In the literature you may find proofs of **ZORN** directly from **AC**. Often, a careful reading reveals that the proof of **WO** from **AC** is implicit in the argument. You may also observe, reading mathematics in general, that applications of **ZORN** normally may be replaced by a combined use of **WO** and transfinite recursion. Examples are the existence of ultrafilters extending a filter, every vector space has a base, and less transparently, Tychonov's theorem and Hahn-Banach.

This discussion is of course not proper mathematics, but as a future scholar in set theory, you may wonder why mathematicians avoid the idea of transfinite recursion.

References

- [1] K. Kunen, Set Theory An introduction to Independence Proofs, Elsevier 1980.