§11 - Axiom of Choice

1 Motivation

What constitutes a rigourous mathematical argument? Normally we think of a proof as being airtight if it starts from a true premise, then uses logical deductions to arrive at a new statement. For example, in our proof that every second countable, regular space was normal (§9.6) we started with the assumption that a space was second countable and regular, we unwound those definitions, used some facts about topology and set manipulation, then we arrived at the conclusion that the space was normal.

Our proof was airtight but something we did requires a bit of explanation. Take a second to look through that proof, and try to find what part of the proof was "non-canonical"; that is, which part of the proof did we use the existence of something, without *really* writing it down? This proof (as written) has the peculiar property that two different people going through this proof could choose different U_a and different V_a for the proof, and it would still work. Is this okay? Is this a valid part of a mathematical argument?

Non-Canonical Exercise: "Fix" that proof that second countable, regular spaces are normal by forcing each reader of the proof to choose the same U_a and the same V_a. One way to do this is to specify a *canonical* choice, so that the reader does not have any freedom.

Phew! If you completed that exercise, then you have *really, really* made that proof airtight. We won't be able to "fix" every one of our proofs in this course, but that shouldn't bother us *too* much. We will briefly look at whether or not these types of arguments are valid or not, and we will also look at the type of theorems that rely on non-canonical choices. One key tool will be Zorn's Lemma.

2 Axiom of Choice

Our main goal here is to introduce the Axiom of Choice.

Definition. *Let $\mathcal{A}$ be a (non-empty) collection of non-empty sets. A function $f : \mathcal{A} \longrightarrow \bigcup \mathcal{A}$ such that $f(A) \in A$ $(\forall A \in \mathcal{A})$ is called a* ***choice function*** *(for $\mathcal{A}$).*

In English, "**a choice function picks an element out of each set**". Choice functions are used often in math, although usually we avoid *explicitly* mentioning them. They tend to make the notation a bit messy.

Some examples:

- Suppose $\mathcal{A} := \{ \{a, b\}, \{\text{Mike}, \text{Janet}, \text{Ali}\}, \mathbb{N} \}$. Then the function f defined by $f(\{a, b\}) = a$, $f(\{\text{Mike}, \text{Janet}, \text{Ali}\}) = \text{Mike}$ and $f(\mathbb{N}) = 10$ is a choice function. It is really just picking an element from each set. To be very clear here, the codomain of the choice function is:

$$\bigcup \mathcal{A} = \{\, a, b, \text{Mike}, \text{Janet}, \text{Ali}, 1, 2, 3, \dots \} = \{a, b\} \cup \{\text{Mike}, \text{Janet}, \text{Ali}\} \cup \mathbb{N}$$

- Suppose $\mathcal{A} := \{\, A : \emptyset \neq A \subseteq \mathbb{N} \,\} = \mathcal{P}(\mathbb{N}) \setminus \{\emptyset\}$. Then the function g defined by $g(A) = \min(A)$ is a choice function. It is a very important choice function!

Empty Set Exercise: Why did we add the condition that each element of $\mathcal{A}$ must be a non-empty set?

Now we can state the Axiom of Choice!

Axiom (The Axiom of Choice)**.** *Suppose $\mathcal{A}$ is a (non-empty) collection of non-empty sets. There exists a choice function for $\mathcal{A}$.*

In English, the Axiom of Choice says "**if there are bunch of decisions to make, we can make each of them**". The strength of this axiom comes from the fact that "a bunch" could be infinite.

Finite Choice: Convince yourself that the "finite Axiom of Choice" is true. That is, if $\mathcal{A} = \{\, A_i : 1 \leq i \leq N \,\}$ is a (non-empty) collection of non-empty sets then there exists a choice function for $\mathcal{A}$. (Does your argument extend to $\mathcal{A}$ countable? If it does then you've just proved the "countable Axiom of Choice" is true, and you will win a Fields medal.)

Example. *Here is a classic example used to illustrate the power of the Axiom of Choice. Suppose you have an infinite collection of pairs of shoes. Is there a way to choose exactly one shoe from every pair? Of course, always take the left shoe. Now, suppose you have an infinite collection of pairs of socks. Is there a way to choose exactly one sock from every pair? It seems like the answer should be yes, but now we can't say "just take the left ones" because socks don't have a left sock and a right sock. Here we would need the Axiom of Choice to pick out a sock from each pair.*

Why do we call the Axiom of Choice an "axiom"? It is because we cannot prove it, and it seems like a reasonable assertion. We now know that the Axiom of Choice is independent from the other axioms of mathematics, and we are free to assume that "it is true". In this class some of our results will depend on this axiom. In the early 20th century there were bitter debates about this axiom. Friends were lost and hurtful words were said. Nowadays, the general consensus is that the Axiom of Choice is a legitimate mathematical tool, although its use should be avoided if possible.

Why avoid the Axiom of Choice? The Axiom of Choice provides us with a "non-constructive" construction in proofs when we use it. In the socks example, the Axiom of Choice would tell us that we could pick a sock from each pair, but we wouldn't know anything about the socks that were chosen. The Axiom of Choice does not give us a "method for choosing the socks", it just presents us with the socks. In mathematics that deals with the real world this can be problematic. How can you say that you have a proper method for solving a problem if it relies on the Axiom of Choice?

3 Equivalences to the Axiom of Choice

As it turns out, we will rarely use the Axiom of Choice as stated. We will instead use various corollaries of it, and some things that are actually equivalent to it. The most important one is called "Zorn's Lemma" (which we will discuss in a moment) and we have already seen the well-ordering principle in §10 when we constructed ω_1.

Theorem. *The following are equivalent:*

1. *The Axiom of Choice (AC);*
2. *The well-ordering principle;*
3. *Zorn's Lemma.*

This is a theorem of set theory and logic that we will not prove. If you are interested in the proof you may consult Jech and Hrbacek's book "Introduction to Set Theory", or you may wish to take the class MAT409.

Before we introduce Zorn's Lemma, there is a classic line about this theorem: "The Axiom of Choice is clearly true, the well-ordering Principle is clearly false, and as for Zorn's lemma, who knows?". The idea is that even though we know they are equivalent statements, the Axiom of Choice seems very harmless, the well-ordering principle seems crazy, and Zorn's Lemma will be a statement that seems completely alien.

4 Zorn's Lemma

Zorn's Lemma is a statement about the existence of maximal elements in partial orders. We already know some of the language about partial orders, but we will need a little bit more.

Definition. *Let $(\mathbb{P}, \leq)$ be a partial order. A set $C \subseteq \mathbb{P}$ is said to be **bounded above** if there is an element $r \in \mathbb{P}$ such that $c \leq r$ for all $c \in C$. A set $C \subseteq \mathbb{P}$ is said to be **bounded below** if there is an element $l \in \mathbb{P}$ such that $l \leq c$ for all $c \in C$. A set $C \subseteq \mathbb{P}$ is said to be **bounded** if it is both bounded above and bounded below.*

We have already seen this notion. When we talked about ω_1 we noted that every countable subset of ω_1 is bounded above. Now we need the notion of a maximal element.

Definition. *Let $(\mathbb{P}, \leq)$ be a partial order. An element $m \in \mathbb{P}$ is said to be **maximal** if there is no $p \in \mathbb{P}$, different from m such that $m \leq p$.*

Some examples:

- The natural numbers with their usual ordering is a partial order with no maximal element.
- The partial order $(\mathcal{P}(\mathbb{R}), \subseteq)$ has a maximal element: $\mathbb{R}$. (In fact, prove that any chain $\mathcal{C} \subseteq \mathcal{P}(\mathbb{R})$ has $\bigcup \mathcal{C}$ as an upper bound of $\mathcal{C}$.)

Now we are ready to state Zorn's Lemma:

Theorem (Zorn's Lemma)**.** *Suppose that $(\mathbb{P}, \leq)$ is a (non-empty) partial order with the property that "every chain has an upper bound" then $\mathcal{P}$ has a maximal element.*

Upside Down Exercise: Using Zorn's Lemma, prove that "If $(\mathbb{P}, \leq)$ is a (non-empty) partial order with the property that "every chain has an lower bound" then $\mathcal{P}$ has a minimal element."

What will make this a very powerful theorem is by making clever choices for which partial order we use. We will look at a couple of examples and hopefully it will make sense. (Truth be told, it is really easy to forget the meaning of the words in Zorn's Lemma. Try to remember a couple of examples of how Zorn's Lemma works and everything should be fine!)

Example 1. *Let us show, using Zorn's Lemma, that ω_1 has an uncountable collection of pairwise disjoint open sets.*

Proof. The idea is to design a partial order where the maximal element is an uncountable collection of pairwise disjoint open sets in ω_1. Well, let

$$\mathbb{P} := \{\, \mathcal{A} : \mathcal{A} \text{ is a collection of pairwise disjoint open, (countable) subsets of } \omega_1 \,\}$$

and order the collections "by inclusion", that is $\mathcal{A} \leq \mathcal{B}$ iff $\mathcal{A} \subseteq \mathcal{B}$.

Claim: If $\mathcal{M} \in \mathbb{P}$ is a maximal element, then $\mathcal{M}$ is an *uncountable* collection of pairwise disjoint open sets.

Since $\mathcal{M} \in \mathbb{P}$ we already know that $\mathcal{M}$ is a collection of pairwise disjoint open sets, we only need to check that it is uncountable. Well, suppose that it was countable. Then $\bigcup \mathcal{M}$ is a countable union of countable sets, so it is countable. Finding an upper bound α for $\bigcup \mathcal{M}$, we see that there is a (countable) open set U that lives above α (and hence above $\mathcal{M}$), and we see that U is disjoint from each element of $\mathcal{M}$. Thus $\mathcal{M} < \mathcal{M} \cup \{U\}$, so $\mathcal{M}$

is not a maximal element.

Claim: In $\mathbb{P}$ every non-empty chain has an upper bound.

Let $\mathcal{C} \subseteq \mathbb{P}$ be a non-empty chain. I claim that $\bigcup \mathcal{C} = \bigcup \{ \mathcal{A} : \mathcal{A} \in \mathcal{C} \}$ is an upper bound for $\mathcal{C}$. (This is a very standard choice for an upper bound. Usually in Zorn's Lemma-type proofs the upper bound in question is the union.) Here though, it is obvious that if $\mathcal{A} \in \mathcal{C}$ then $\mathcal{A} \subseteq \bigcup \mathcal{C}$. What isn't clear is if $\bigcup \mathcal{C}$ is actually an element of $\mathbb{P}$. That is, is $\bigcup \mathcal{C}$ "a collection of pairwise disjoint open, (countable) subsets of" ω_1? The only thing that could go wrong is if there are open sets are not pairwise disjoint. Well let $A, B \in \bigcup \mathcal{C}$. Then they came from elements of $\mathcal{C}$, so there are $\mathcal{A}, \mathcal{B} \in \mathcal{C}$ such that $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Since $\mathcal{C}$ was a chain, either $\mathcal{A} \subseteq \mathcal{B}$ or $\mathcal{B} \subseteq \mathcal{A}$. Without loss of generality, say $\mathcal{A} \subseteq \mathcal{B}$. So then $A \in \mathcal{B}$, and so A and B must be disjoint.

So now we have checked all the hypotheses for Zorn's Lemma, so it applies, and our argument shows that a maximal element is the desired uncountable collection of pairwise disjoint open sets. □

It Always Gets Me Exercise!: "we have checked all the hypotheses for Zorn's Lemma" was a lie, (sorry!). Find the one condition of Zorn's Lemma that we didn't check and check it for yourself.

It's Collections All The Way Down Exercise: That proof had a lot of collections of collections of chains of collections of collections of .... Go through it with a picture, and use remember the exercise from §10 that showed that ω_1 is not ccc.

More CCC Exercise: This proof gave a general method for proving that a space is not ccc. It used a little bit about ω_1 but not very much. See if you can find a general topological property that would allow this proof to go through. Alternatively, find a topological space that isn't ccc, and prove it using a method similar to this one.

You will get more experience with Zorn's Lemma on Assignment 6. We will need it to prove Tychonoff's Theorem later in the course.

5 Using Zorn's Lemma

Let me give you some intuition for using Zorn's Lemma. This technique is useful for creating objects that contain "a lot of information". In our previous example we wanted to create an uncountable collection of mutually disjoint open sets in ω_1. Here an uncountable collection contains "more information" than a small (i.e. countable) collection. Here our notion of information is similar to that of a computer; a text file does not contain a lot of information (it is a small file), but a video tends to contains a lot of information (it is a big file).

It is usually a good idea to try to use Zorn's Lemma when you need to create a highly complicated object, that contains "a lot of information". In this case, try to construct a partial order such that a larger element of the partial order contains more information. That way a maximal element contains a maximal amount of information. In our previous example, the partial order contained collections of mutually disjoint open sets, and the bigger elements were ones with more sets.

Some examples:

- "Every vector space has a basis". The object we want to create is a basis, which is a (spanning) set of linearly independent vectors. So a good guess for a partial order would be collections of linearly independent vectors. That way a maximal element (if it exists) will be a spanning set of linearly independent vectors.

- "Every collection $\mathcal{A}$ of non-empty sets has a choice function". The object we want to create is a choice function, which picks an element from *every* set $A \in \mathcal{A}$. Here, a good guess for the partial order would be *partial* choice functions of $\mathcal{A}$. That is, a partial choice function only picks out elements from *some of* the $A \in \mathcal{A}$. The ordering would be that larger elements (in the order) make choices for more elements of $\mathcal{A}$. This way a maximal element (if it exists) makes a choice for *every* set $A \in \mathcal{A}$.

In an actual application of Zorn's Lemma, when you are trying to show that "every chain has an upper bound" there is a standard approach for certain nice partial orders. If your partial order contains only collections of objects (like in all of our examples), you can *usually* take the union of the chain to be your upper bound. This doesn't always work, but it is the standard approach which often does work.

6 More Equivalences and some Corollaries

Before we leave, lets give some more equivalences to the Axiom of Choice that are interesting, but not particularly useful for this course.

Theorem 2. *The following are equivalent:*

1. *The Axiom of Choice;*
2. *Every vector space has a basis;*
3. *Every equivalence relation has a set of representatives;*

4. *If* $\{A_\alpha : \alpha \in I\}$ *is a collection of non-empty sets, then* $\emptyset \neq \prod_{\alpha \in I} A_\alpha$. *(We will talk about this later.);*

5. *Tychonoff's Theorem (which we will see later in this course);*

6. *(The Compactness Theorem in Logic) Every finitely consistent set of sentences is consistent.*

Additive Exercise: If you did New Ideas question NI.1 on Assignment 4 (and that was *everyone* except 3 people), then construct a (non-continuous) additive function $f : \mathbb{R} \longrightarrow \mathbb{R}$ that isn't just a line. Use the fact that $\mathbb{R}$ (as a vector space over $\mathbb{Q}$!) has a basis.

And finally let's state some results from various branches of math that are corollaries of the Axiom of Choice.

Theorem 3. *The following are corollaries of the Axiom of Choice:*

1. *(Topology) In a first countable space, "a function* f *is continuous" is equivalent to "*$x_n \to x$ *implies* $f(x_n) \to f(x)$*".*

2. *(Topology) Various forms of compactness are equivalant.*

3. *(Set Theory) Every set has a (well defined) cardinality.*

4. *(Set Theory) A countable union of countable sets is countable.*

5. *(Algebra) Every ideal in a ring is contained in a maximal ideal;*

6. *(Algebra)* $\mathbb{Q}$ *has a (minimal) algebraic completion.*

7. *(Real Analysis) The Hahn-Banach Theorem;*

8. *(Real Analysis) Vitali's Theorem that there is a subset of* $\mathbb{R}$ *that is not Lebesgue Measurable;*

9. *(Real Analysis) The Banach-Tarski "Paradox".*

Wait. Did I just say that the fact that "a countable union of countable sets is countable" uses the Axiom of Choice? That seems weird.

Countability and Choice Exercise: Go through the proof of that "a countable union of countable sets is countable" (§4) and identify where we use the Axiom of Choice. A different theorem in that section also uses the Axiom of Choice. Which one?

7 Summary of Exercises

These exercises aren't for submission, but are useful for understanding.

Non-Canonical : "Fix" that proof that second countable, regular spaces are normal by forcing each reader of the proof to choose the same U_a and the same V_a. One way to do this is to specify a *canonical* choice, so that the reader does not have any freedom.

non-$\emptyset$: In the definition of a choice function, why did we add the condition that each element of $\mathcal{A}$ must be a non-empty set?

Finite Choice : Convince yourself that the "finite Axiom of Choice" is true. That is, if $\mathcal{A} = \{ A_i : 1 \le i \le N \}$ is a (non-empty) collection of non-empty sets then there exists a choice function for $\mathcal{A}$. (Does your argument extend to $\mathcal{A}$ countable? If it does then you've just proved the "countable Axiom of Choice" is true, and you will win a Fields medal.)

IAGM : Find the one condition of Zorn's Lemma that we didn't check in the proof that ω_1 is not ccc and check it for yourself.

All The Way Down : That proof had a lot of collections of collections of chains of collections of collections of .... Go through it with a picture, and use remember the exercise from §10 that showed that ω_1 is not ccc.

More CCC : That Zorn's Lemma proof gave a general method for proving that a space is not ccc. It used a little bit about ω_1 but not very much. See if you can find a general topological property that would allow this proof to go through. Alternatively, find a topological space that isn't ccc, and prove it using a method similar to this one.

NI1 : Using some form of AC, construct a non-additive function that isn't a line.

Ctble + Choice : Go through the proof of that "a countable union of countable sets is countable" (§4) and identify where we use the Axiom of Choice. A different theorem in that section also uses the Axiom of Choice. Which one?