BACKGROUND FROM SET THEORY

Nowdyas it is a common approach to build a mathematical theory from a list of axioms. And the theory of sets underlines all mathematical theories. Most mathematicians use set theory in a "naive" way based on the intuitive understanding of the notion of a set. However, sometimes a deeper understanding of the world of sets is necessary and then our intuition may be not sufficient. This is why we need strong fundaments for set theory which are provided by the axioms. For the list of all axioms and a nice introduction to the methods of modern set theory we recommend the book by K. Ciesielski [1].

Among the axioms of set theory there is one, called the **Axiom of Choice**, which caused a lot of controversy during the developement of set theory but nowdays it is accepted by the majority of mathematicians:

Axiom of Choice. *If the sets X_i, $i \in I$ are non-empty then there exists a function $f : I \longrightarrow \bigcup_{i \in I} X_i$ such that $f(i) \in X_i$ for all $i \in I$.*

Intuitively, the axiom says that given any family of non-empty sets we can choose at once one element from each set. While this is trivial for finite families, the case of infinite families is much more subtle than our finite world experience may suggest.

Equivalently, the Axiom of Choice says that the product $\Pi_{i \in I} X_i$ of nonempty sets is nonempty.

One of the reasons for the controversy around the Axiom of Choice is the following cute and counterintuitive result:

The Banach-Tarski Paradox. *A solid ball B in $\mathbb{R}^3$ can be decomposed into five pairwise disjoint subsets S_1, S_2, S_3, S_4, S_5 for which there exist isometries t_i of $\mathbb{R}^3$, $i = 1, ..., 5$ such that both $t_1(S_1) \cup t_2(S_2)$ and $t_3(S_3) \cup t_4(S_4) \cup t_5(S_5)$ are isometric to B.*

In other words, we can duplicate a ball! There is even more surprising result, called the **Strong Banach-Tarski Paradox**:

Strong Banach-Tarski Paradox. *Given any two bounded sets A and B in $\mathbb{R}^3$, each having nonempty interior, one can decompose A into a finite number of pieces and rearrange them by rigit motions to form B.*

More about such paradoxical decompositions and their connections with group theory can be found in the beautifull book by Stan Wagon [2]

Let us note that A. Tarski proved in 1924 that the Axiom of Choice is equivalent to the following nowdays commonly used statement:

For any infinite set X there is a bijection from X onto $X \times X$.

There are two other statements equivalent to the Axiom of Choice, which have numerous important applications. Before we formulate them we need to recall some basic concepts about ordered sets.

Definition 1. An **ordering** (or **partial ordering**) of a set X is a binary relation $\leq$ on X such that:

- $x \leq x$ for all $x \in X$;
- if $x \leq y$ and $y \leq x$ then $x = y$;
- if $x \leq y$ and $y \leq z$ then $x \leq z$.

We write $x < y$ if $x \leq y$ and $x \neq y$.

Definition 2. An ordering $\leq$ on a set X is called **linear** (or **total**) if for any two elements $x, y \in X$, either $x \leq y$ or $y \leq x$.

For example, the usual order on positive integers is linear, but if we define $n \leq m$ iff n divides m, then the resulting order is not linear.

Let $(X, \leq)$ be an ordered set and Y a subeset of X. We say that $a \in Y$ is a **maximal** (**minimal**) element of Y if for every element $y \in Y$ such that $a \leq y$ $(y \leq a)$ we have $y = a$. We say that $a \in Y$ is the **largest** (**smallest**) element of Y if for every $y \in Y$ we have $y \leq a$ $(a \leq y)$. It is clear that the largest (smallest) element is unique, if exists, and that it is also a maximal (minimal) element. If Y is linearly ordered by $\leq$ then a maximal (minimal) element is automatically largest (smallest). For example, in the set of all integers larger than 1 ordered by divisibility all prime numbers are minimal elements. An **upper bound** (**lower bound**) of Y is any element $b \in X$ such that $y \leq b$ $(b \leq y)$ for all $y \in Y$.

Definition 3. We say that an ordered set $(X, \leq)$ is **inductively ordered** if every non-empty, linearly ordered by $\leq$ subset of X has an upper bound.

A linearly ordered subset of an ordered set is often called a **chain**.

The next definition introduces a very important class of orderings.

Definition 4. A **well-ordering** of a set X is a linear ordering $\leq$ such that every nonempty subset of X conatains a smallest element.

For example, the set of positive integers with its natural ordering is well ordered, but the set of all integers is not.

A well-ordering $\leq$ of a set X allows us to prove facts about elements of X by the so-called **transfinite induction**. This means that if for each $x \in X$ we have some statement $P(x)$ such that

— $P(a)$ is true for the smallest element a of X;

— for any $b \in X$, if $P(x)$ is true for all $x < b$ then $P(b)$ is true

then $P(x)$ is true for all $x \in X$.

Note also that for each $x \in X$ which is not the largest element of X there is unique element x^* which is "the next element after x", i.e. which is smallest among all elements larger than x. In general, not every element of X is of the form x^* for some $x \in X$. The elements which are not of this form are called **limit** elements.

The following notion is quite useful when dealing with well ordered sets.

Definition 5. Let $(X, \leq)$ be a well ordered set. A subset Y of X is called an **initial** segment of X if either $X = Y$ or there is $a \in X$ such that $Y = \{x \in X : x < a\}$.

Now we can state the results equivalent to the Axiom of Choice.

Zermelo's Lemma. *Every non-empty set can be well ordered.*

Kuratowski-Zorn Lemma. *Every inductively ordered set has a maximal element.*

Theorem 1. *The following statements are equivelent:*

a) the Kuratowski-Zorn Lemma;

b) Zermelo's Lemma;

c) the Axiom of Choice.

Proof: $a) \Longrightarrow b)$: Let X be a non-empty set. Consider the set $WP(X)$ which consists of all pairs $(Y, \leq_Y)$ such that Y is a nonempty subset of X and $\leq_Y$ is a well-ordering of Y. We define an ordering $\preceq$ on $WP(X)$ by $(Y, \leq_Y) \preceq (Z, \leq_Z)$ iff Y is an initial segment of Z and the restriction of $\leq_Z$ to Y coincides with $\leq_Y$. Let M be a linearly ordered subset of $WP(X)$. Define a subset Y of X by $a \in Y$ iff there is $(Z, \leq_Z) \in M$ such that $a \in Z$. For $a, b \in Y$ set $a \leq_Y b$ iff there is $(Z, \leq_Z) \in M$ such that $a, b \in Z$ and $a \leq_Z b$ (note that this does not depend on the choice of $(Z, \leq_Z)$.

The fact that M is linearly ordered implies that $\leq_Y$ is a well-ordering of Y. In fact, let A be a nonempty subset of Y. There is $(Z, \leq_Z) \in M$ such that $A \cap Z \neq \emptyset$. Set m for the smallest element of $A \cap Z$ (it exists since Z is well ordered). We claim that m is the smallest element in A. In fact, if $a \in A$ then there is $(Z', \leq_{Z'}) \in M$ such that $a \in Z'$. Now, either Z is an initial segment of Z' or Z' is an initial segment of Z. If $a \in Z$ then $a \in Z \cap A$, so $m \leq_Z a$ by definition of m and therefore $m \leq_Y a$. If $a \notin Z$ then Z is an initial segment of Z' and all elements in Z are smaller than a. In particular, $m \leq_{Z'} a$ so $m \leq_Y a$.

Note now that $(Z, \leq_Z) \preceq (Y, \leq_Y)$ for all $(Z, \leq_Z) \in M$. Thus $(Y, \leq_Y)$ is an upper bound of M. Consequently, $WP(X)$ is inductively ordered by $\preceq$ so it has a maximal element $(T, \leq_T)$. If $T \neq X$, then there is an $a \in X$, $a \notin T$. Define $T' = T \cup \{a\}$ and extend the order $\leq_T$ to T' by setting $t \leq_{T'} a$ for all $t \in T$. The resulting ordering $\leq_{T'}$ of T' is a well-ordering and $(T, \leq_T) \prec (T', \leq_{T'})$, a contradiction. Thus $T = X$, i.e. X can be well ordered.

$b) \Longrightarrow c)$: Let X_i, $i \in I$ be nonempty sets. The set $X = \bigcup_{i \in I} X_i$ can be well ordered by $\leq$. Define $f(i)$ to be the smallest element of X_i.

$c) \Longrightarrow a)$: Let $(X, \leq)$ be an inductively ordered set and suppose that X does not have maximal elements. Let I be the set of all subsets of X which are well ordered by $\leq$ (so the empty set belongs to I). For every $i \in I$ define U_i to be the set of all upper bounds for i which do not belong to i. Since X has no maximal elements, each U_i is a nonempty subset of X. By the Axiom of Choice, there is a function $g : I \longrightarrow X$ such that $g(i) \in U_i$ for all $i \in I$. Note that if $i \in I$ and $c \in i$ then the set $i_c = \{x \in i : x < c\}$ also belongs to I. We say that $i \in I$ is **good** if $g(i_c) = c$ for every $c \in i$. For example, for any $a \in X$ the sets $\emptyset$, $\{a\}$ and $\{a, g(\{a\})\}$ are good. More generally, for any good set i the set $i \cup \{g(i)\}$ is also good.

The main observation is that if i, j are good sets then either $i \subseteq j$ or $j \subseteq i$. In fact, if for every $x \in i$ we have $i_x = j_y$ for some $y \in j$, then $x = g(i_x) = g(j_y) = y$ so $i \subseteq j$. Otherwise, there is smallest $x \in i$ such that $i_x \neq j_y$ for all $y \in j$. If $a \in i_x$ then $i_a = j_b$ for some $b \in j$. Applying g yields $a = b$. Thus $i_x \subseteq j$ and $i_a = j_a$ for all $a \in i_x$. We will show that $j = i_x$. If not, let $u \in j$ be smallest such that $u \notin i_x$, so $j_u \subseteq i_x$. If $a \in i_x$ and $u < a$ then $u \in j_a = i_a \subseteq i_x$, a contradiction. Thus $a < u$ for all $a \in i_x$, i.e $i_x \subseteq j_u$. We see that $i_x = j_u$, a contradiction.

Thus we proved that the good sets are linearly ordered by inclusion. Let j be the union of all good sets. It is an easy exercise to see that j itself is a good set. But then $j \cup g(j)$ is also a good set, which is larger than j, a contradiction. □

It is evident from our proof that the hardest implication is $c) \implies a)$. This suggests that in a sense the Kuratowski-Zorn Lemma is "strongest" among all three results, and in fact it seems to be the most efficient tool for applications. A typical application of the Kuratowski-Zorn Lemma is in showing that every vector space over a field has a basis:

Theorem. *Any vector space V over a field F has a basis.*

Proof: Recall that a basis of V is a subset of V which is linearly idependent and spans V (we consider the empty subset of V to be linearly independent) Moreover, for any linearly independent subset A of V, the span $S(A)$ of A is the set of all linear combinations of vectors in A and if $v \notin S(A)$ then $A \cup \{v\}$ is also linearly independent (in particular, $S(\emptyset) = \{0\}$). Thus if A is a maximal linearly independent subset of V then it is a basis (here maximal is with respect to inclusion). Hence we just need to show that a maximal linearly independent subsets exist and this is where the Kuratowski-Zorn Lemma comes to rescue. Consider the set $I(V)$ of all linearly independent subsets of V. It is a nonempty set ordered by inclusion. It is easy to see that if M is a linearly ordered subset of $I(V)$ then the union of all members of M is a linearly independent subset of V which is an upper bound for M. Thus $I(V)$ is inductively ordered by inclusion, so maximal elements exist in $I(V)$. □

References

[1] K. Ciesielski, **Set Theory for the Working Mathematician**, London Mathematical Society Student Texts 39, Cambridge University Press 1997.

[2] S. Wagon, **The Banach-Tarski Paradox**, Encyclopedia of Mathematics and its Application Vol. 24, Cambridge University Press 1985.