

The Axiom of Choice:

What the beginning grad student needs to
know

Reference:

K. Ross, Informal Introduction to Set Theory,
<http://www.uoregon.edu/~ross1/SetTheory.pdf>

What is it?

Let I be any set and suppose that

$$\{A_\lambda \mid \lambda \in I\}$$

is a family of non-empty sets indexed by I . Then there exists a *choice* function f from I into $\cup_{\lambda \in I} A_\lambda$ with the property that

$$f(\lambda) \in A_\lambda \quad \text{for each } \lambda \in I.$$

Alternate Statement

Let S be any set. Let $\mathcal{P}(S)$ be the power set of S , that is the set of all subsets of S .

Then there is a function $g : \mathcal{P}(S) - \{\emptyset\} \rightarrow S$ such that

$$g(A) \in A \quad \text{for every } A \subset S.$$

Why the fuss?

- A rigorous foundation for mathematics requires a theory of sets, founded on a set of axioms for construction and manipulation of sets, based on the intuitive notion of a set as a collection of objects. Most mathematicians (if they think about it at all!) use the Zermelo-Fränkel (ZF) axioms.
- The choice function is a useful tool in manipulation of sets and its existence for **finite index sets** follows from ZF.
- For general sets, AC must be added as a separate axiom. While it is intuitively reasonable, it differs qualitatively from theory built on ZF by asserting the existence of a sophisticated object without giving an explicit construction. Therefore, acceptance

of AC has been (and still is to some degree) an issue of contention. It is now assumed by most “working mathematicians.”

- AC is equivalent to strong set theory principles that are profound and not intuitive. AC and its consequences are at the heart of basic theorems in analysis, topology, and algebra.

The talk

- Existence of non-measurable sets
- Equivalents to AC:
 - The Hausdorff Maximality Principle
 - Zorn's Lemma
 - Well ordering Principle
- Applications:
 - Existence of maximal ideals
 - Existence of Hamel bases
 - Derivation of well ordering and Hausdorff maximality from Zorn
 - Inductive definitions

Math To Know

- Sets, unions, intersections, etc. (naive set theory basics)
- Equivalence relations
- Vector spaces:
Example to keep in mind: $C[a, b]$, the continuous, real-valued function functions defined on $[a, b]$, as a vector space over the reals:

$$(\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x)$$

- Rings and Ideals. Definitions

Non-measurable sets

Does there exist a function μ (called a measure),

$$\mu : \mathcal{P}(\mathbb{R}) \rightarrow [0, \infty],$$

that assigns to each subset A of \mathbb{R} a "length" $\mu(A)$ and that satisfies:

- Non-triviality: $0 < \mu([0, 1)) < \infty$.
- Translation invariance:

$$\forall x \in \mathbb{R}, \forall A \subset \mathbb{R}, \quad \mu(A + x) = \mu(A);$$

- Countable additivity: If A_1, A_2, \dots are pairwise disjoint subsets of \mathbb{R}

$$\mu \left(\bigcup_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} \mu(A_n) ?$$

ANSWER: NO! (assuming the axiom of choice!)

A partition of $[0, 1]$ into translates of a set

- We will work on $[0, 1)$. Addition will be interpreted modulo 1: for $0 \leq x, y < 1$,

$$x + y \triangleq (x + y) \mid \text{mod } 1.$$

- Q will denote the rational numbers in $[0, 1)$.

Note: If μ is translation invariant, it is translation invariant also with respect to addition modulo 1.

LEMMA There is a set $C \subset [0, 1)$, such that $C + r$ and $C + q$ are disjoint for any two distinct (modulo 1) rational numbers r and q , and

$$[0, 1) = \bigcup_{r \in Q} C + r.$$

Theorem If μ is translation-invariant and countably additive as a function on the power set of \mathbb{R} , then either $\mu([0, 1)) = 0$ or $\mu([0, 1)) = \infty$.

Proof: $\mu(C + r) = \mu(C)$ for every r . By countable additivity (Q is countable!),

$$\begin{aligned}\mu([0, 1)) &= \mu\left(\bigcup_{r \in Q} C + r\right) \\ &= \sum_{r \in Q} \mu(C + r) \\ &= \sum_{r \in Q} \mu(C).\end{aligned}$$

This can only happen if either $\mu([0, 1)) = 0$ or $\mu([0, 1)) = \infty$.

Proof of Lemma

Define $x \sim y$ iff $x - y$ is rational.

\sim is an equivalence relation.

Let I be the collection of equivalence classes. By AC, $\exists f : I \rightarrow \mathbb{R}$ such that $f(\lambda) \in \lambda$ for each equivalence class λ . Let

$$C \triangleq \bigcup_{\lambda \in I} \{f(\lambda)\}.$$

If w and v are two distinct points in C , they are not equivalent— $w - v \notin Q$. If $x \in (C + r) \cap (C + q)$, $r \neq q$, and $r, q \in Q$, $\exists y, z \in C$, s.t. $x = y + r$ and $x = z + q$. But then $z - y = r - q \in Q$, which is a contradiction. So $(C + r) \cap (C + q) = \emptyset$.

Let $x \in [0, 1)$. There is a unique $\lambda \in I$ such that $x \in \lambda$. Hence $x - f(\lambda) \in Q$. Thus,
 $[0, 1) = \bigcup_{r \in Q} (C + r)$.

Axiom of Choice in Set Theory

Dramatis Personae

- Georg Cantor (1845-1918)
- Ernst Zermelo (1871-1953)
- Felix Hausdorff (1888-1942)
- Max Zorn (1906-1993) (also K. Kuratowski, 1896-1980)

Orderings of sets

A *partial order* on a set A is a relation \preceq which is

- Reflexive: $x \preceq x$ for every $x \in A$.
- Anti-symmetric: $x \preceq y$ and $y \preceq x$ imply $x = y$.
- Transitive: $x \preceq y$ and $y \preceq z$ imply $x \preceq z$.

A *linear order* on a set A is a partial order for which every two elements are comparable.

A *well ordering* of a set A is a linear ordering with the following property: every subset B of A has a least element: if $B \subset A$, $\exists b \in B$ such that $b \preceq x$ for all $x \in B$.

Set theoretically a partial ordering of A is a subset of $A \times A$.

Examples

- (Partial order of set inclusion) Let U be a set. Let $A \triangleq \mathcal{P}(U) = \{x \mid x \text{ is a subset of } U\}$, and let $x \preceq y$ iff $x \subset y$.
- The positive integers with the usual ordering form a well ordered set. The reals with the usual ordering form a linearly ordered set.
- The set of rational numbers of the form $n - (1/m)$ where n and m are positive integers is well ordered using the usual ordering.
- Let H be a vector space over a field K . (Example: all continuous real valued functions on an interval $[a, b]$ as a vector space over the reals.) Let A be the set of all vector subspaces of H . The relation of inclusion defines a partial order.

- Let R be a commutative ring with identity. Let A be the set of all proper ideals. Inclusion again defines a partial order on A .

- Fix sets U and V . Let

$$A \triangleq \{f \mid f : S \rightarrow V, S \subset V\}.$$

Define:

$$f \preceq g \text{ iff } \text{dom}(f) \subset \text{dom}(g), g|_{\text{dom}(f)} = f$$

CHAINS AND MAXIMALITY

Let (A, \preceq) be a partially ordered set.

An element $x \in A$ is *maximal* if no element in A is larger: for any $y \in A$, $x \preceq y$ implies $x = y$.

A *chain* in A is a subset $B \subset A$ which is linearly ordered by \preceq .

A chain is *maximal* if it is not contained properly in another chain.

Example: Take v_1, v_2, \dots a sequence of vectors in vector space H . Let S_n be the subspace spanned by v_1, \dots, v_n : $S_n = \left\{ \sum_1^n \alpha_i v_i \mid \alpha_i \in K \right\}$. Then $S_1 \subset S_2 \subset S_3 \subset \dots$, so the set $\{S_n\}$ is a chain for the the partial order of inclusion in the set of subspaces of H .

H (considered as a subset of itself) is a maximal element.

If $\cup_n S_n = H$, then $\{S_n\}$ is a maximal chain.

Hausdorff Maximality Principle

Every nonempty partially ordered set contains a maximal chain.

Zorn's Lemma

A nonempty partially ordered set for which every chain has an upper bound contains a maximal element.

THEOREM AC \iff Zorn's Lemma \iff
Hausdorff Maximality Principle

Applications of Zorn's Lemma

A nonempty partially ordered set for which every chain has an upper bound contains a maximal element.

1. Zorn \implies Hausdorff.

- A is any nonempty partially ordered set.
- B is the set of chains in A . Partially order B by inclusion.
- Let $\mathcal{C} (\neq \emptyset)$ be any chain in B (a chain of chains!). Check: $\bigcup_{C \in \mathcal{C}} C$ is a chain, i.e. $\bigcup_{C \in \mathcal{C}} C \in B$.
- Since each $c \in \mathcal{C}$ satisfies $c \subset \bigcup_{C \in \mathcal{C}} C$, $\bigcup_{C \in \mathcal{C}} C$ is an upper bound of \mathcal{C} .

- Zorn \implies B has a maximal element, which is a maximal chain in A . Have proved: Every nonempty partially ordered set has a maximal chain.

2. Zorn \implies existence of Hamel bases

Theorem Any vector space H contains a set U such that

- i) Any finite subset of U is a linearly independent set of vectors.
- ii) Any vector $v \in H$ can be expressed in the form $v = \alpha_1 u_1 + \cdots + \alpha_n u_n$, where $u_i \in U$, $1 \leq i \leq n$.

Proof: Let A be the set of subsets of H satisfying condition (i). Partially order A by inclusion. A is non-empty. A maximal element U of A will satisfy (ii).

Take any nonempty chain \mathcal{C} in A .

Check that $\bigcup_{C \in \mathcal{C}} C$ is in A .

\mathcal{C} is thus an upper bound of \mathcal{C} .

Zorn \implies A has a maximal element U .

3. Zorn \implies Existence of Maximal Ideals

(I is an ideal in a commutative ring R if I is closed under addition and if $rs \in I$ whenever $r \in R$ and $s \in I$.)

Theorem: Let R be a commutative ring with identity. If I is an ideal in R and $I \neq R$, then there is a maximal proper ideal containing I .

Proof: Let A be the set of all proper ideals containing I . Note $I \in A$. Order A by inclusion. Show A contains a maximal element. Let \mathcal{C} be a chain in A , that is a chain of proper ideals.

Show that $\bigcup_{J \in \mathcal{C}} J$ is a proper ideal:

- $1 \notin \bigcup_{J \in \mathcal{C}} J$, since $1 \notin J, \forall J \in \mathcal{C}$.
- $r(\bigcup_{J \in \mathcal{C}} J) = \bigcup_{J \in \mathcal{C}} rJ = \bigcup_{J \in \mathcal{C}} J, \forall r \in R$
- if $s, t \in \bigcup_{J \in \mathcal{C}} J, \exists J' \in \mathcal{C}$ so that $s, t \in J'$. So $s + t \in \bigcup_{J \in \mathcal{C}} J$.

$\bigcup_{J \in \mathcal{C}} J$ is an upper bound in A to \mathcal{C} . Apply Zorn!

4. Zorn \implies Well Ordering Principle

Well Ordering Principle Every set can be well ordered.

Derivation from Zorn:

B is an arbitrary, non-empty set.

$$A \triangleq \{(U, \leq_U) \mid U \subset B, \leq_U \text{ well orders } U\}$$

(E.g., A contains all singleton subsets.)

$(U, \leq_U) \preceq (V, \leq_V)$ iff $U \subset V$ and \leq_V coincides with \leq_U on U .

If U is a maximal element then $U = B$. So it suffices to show a maximal element exists.

If \mathcal{C} is any nonempty chain in A , show $\bigcup \mathcal{C}$ is in A and is an upper bound of \mathcal{C} .

Now apply Zorn!

Other prominent applications

- If H is an abelian subgroup of a group G , there is a maximal abelian subgroup containing H . (exercise!)
- Let h be a group homomorphism between a subgroup H of an abelian group G and the reals (with addition). There is a homomorphism of G extending H .
- The **Hahn-Banach Theorem**. It extends linear functionals from subspaces of a vector space to the whole space.
- **Tychonoff's Theorem**. Let $\{X_\lambda \mid \lambda \in I\}$ be an arbitrary family of compact topological spaces. Then the product $\prod_{\lambda \in I} X_\lambda$ is compact in the product topology.

Theorem: The following are equivalent!

- The Axiom of Choice
- The Hausdorff Maximality Principle
- Zorn's Lemma
- The Well Ordering Principle
- Tychonoff's Theorem

Inductive Definitions; an example

Here is a problem of a type common in proving covering lemmas.

Let \mathcal{F} be a collection of balls in \mathbb{R}^n , all of radius less than or equal to 1. We want to extract from \mathcal{F} a countable sequence of balls B_1, B_2, B_3, \dots with the properties: for every n

- (i) the center of B_n is not contained in any of the balls B_1, B_2, \dots, B_{n-1} ;
- (ii) the radius of B_n is at least $1/2$ of the supremum of all radii of balls in \mathcal{F} whose centers do not lie in B_1, B_2, \dots, B_{n-1} .

How do we know such a sequence exists?

Proof by induction will not work, but Zorn's lemma will!

Let \mathcal{C} consist of all functions b such that the domain of b is either $\{1, 2, \dots, n\}$ or all positive integers, each $b(i)$ is a ball in \mathcal{F} and $b(1), \dots, b(n)$ satisfy (i), (ii). Partially order such functions by $b \preceq c$ if c extends b .

Any chain \mathcal{B} in \mathcal{C} has an upper bound u :

$$\text{if } n \in \bigcup_{b \in \mathcal{B}} \text{dom}(b), \quad u(n) \triangleq \tilde{b}(n),$$

where \tilde{b} is any element of \mathcal{B} whose domain contains n .

Zorn implies there is maximal element. The domain of this maximal element must be \mathbb{N} , the set of all positive integers.

AC the implies Banach-Tarski Paradox:

Let U and V be two closed sets in \mathbb{R}^n , where $n \geq 3$. Then there exist a finite n and disjoint partitions $U = E_1 \cup \dots \cup E_n$, $V = F_1 \cup \dots \cup F_n$, such that E_i and F_i are congruent for each i .

Reference:

K. Stromberg, The Banach-Tarski Paradox, American math. Monthly **86**(1979), 151-161.

Some quotes for your amusement

“The Axiom of Choice is obviously true; the Well Ordering Principle is obviously false; and who can tell about Zorn’s Lemma?”

—Jerry Bona

“To choose one sock from each of infinitely many pairs of socks requires the Axiom of Choice, but for shoes the Axiom is not needed.”

—Bertrand Russell