Hausdorff's Maximality Theorem

Jae Won Choi

June 29, 2021

Hausdorff's maximality theorem (or Zorn's lemma) is used often in graduate level real analysis course. Because the proof is a bit convoluted, most courses use it without proving it and professors usually direct us to the proof in Rudin's real and complex analysis [1]. There is a two page appendix that proves the Hausdorff's maximality theorem. Just like most of the proofs in the book, the proof is very dense and it took me a long time to understand it. Below I provide an easy to follow proof in multiple steps.

First we prove the following lemma:

Lemma. *Suppose $\mathcal{F}$ is a nonempty collection of subsets of a set X such that the union of every subchain of $\mathcal{F}$ belongs to $\mathcal{F}$. Supppose g is a function which associates to each $A \in \mathcal{F}$ a set $g(A) \in \mathcal{F}$ such that $A \subset g(A)$ and $g(A) - A$ consists of at most one element. Then there exists an $A \in \mathcal{F}$ for which $g(A) = A$.*

I always find it nice to write out an example before diving into proofs. Let's say that

$$\begin{aligned}
\mathcal{F} &= \{\{\phi_1, \phi_2, \phi_3\}, \{\phi_1, \phi_2, \phi_3, \phi_4\}, \{\psi_1, \psi_2, \psi_3\} \dots\} \\
\Phi &= \{\{\phi_1, \phi_2, \phi_3\}, \{\phi_1, \phi_2, \phi_3, \phi_4\}, \dots\} \\
\Psi &= \{\{\psi_1, \psi_2, \psi_3\}, \{\psi_1, \psi_2, \psi_3, \psi_4\}, \dots\}
\end{aligned}$$

Then Φ is a subchain of $\mathcal{F}$, and since Φ is totally ordered, if $A, B \in \mathcal{F}$, then either $A \subset B$ or $B \subset A$. With this in mind, let's prove it:

Proof. Fix $A_0 \in \mathcal{F}$. Call a subcollection $\mathcal{F}'$ of $\mathcal{F}$ a tower if $\mathcal{F}'$ has the following three properties:

- $A_0 \in \mathcal{F}'$
- The union of every subchain of $\mathcal{F}'$ belongs to $\mathcal{F}'$
- If $A \in \mathcal{F}'$, then also $g(A) \in \mathcal{F}'$

The family of all towers is nonempty because if $\mathcal{F}_1$ is the collection of all $A \in \mathcal{F}$ such that $A_0 \subset A$, then $\mathcal{F}_1$ is a tower. Let $\mathcal{F}_0$ be the intersection of all towers.

Step 1: $\mathcal{F}_0$ is a tower

We want to show that $\mathcal{F}_0$ satisfies all three properties listed above.

(a) Suppose $\{\mathcal{F}'_n\}_n$ is a set of all towers in $\mathcal{F}$, then $\mathcal{F}_0 = \bigcap \mathcal{F}'_n$. Since $A_0 \in \mathcal{F}'_n$ for all n, $A_0 \in \mathcal{F}_0$.

(b) Suppose Φ is a subchain in $\mathcal{F}_0$ then Φ is a subchain in $\mathcal{F}'_n$ for all n, and since the union of subchain is in all $\mathcal{F}'_n$, $\mathcal{F}_0$ also contains the union of subchain.

(c) If $A \in \mathcal{F}_0$, then $A \in \mathcal{F}'_n \quad \forall n$, and since $g(A) \in \mathcal{F}'_n$ for all n, $g(A) \in \mathcal{F}_0$.

Step 2: $\mathcal{F}_0$ is a subchain of $\mathcal{F}$

Let Γ be the collection of all $C \in \mathcal{F}_0$ such that every $A \in \mathcal{F}_0$ satisfies either $A \subset C$ or $C \subset A$:

$$\Gamma = \{C \in \mathcal{F}_0 : A \subset C \text{ or } C \subset A \text{ for every } A \in \mathcal{F}_0\}.$$

For each $C \in \Gamma$, let $\Phi(C)$ be the collection of all $A \in \mathcal{F}_0$ such that either $A \subset C$ or $g(C) \subset A$:

$$\Phi(C) = \{A \in \mathcal{F}_0 : A \subset C \text{ or } g(C) \subset A\}.$$

Step 2.1: Γ and $\Phi(C)$ satisfy tower properties (a) and (b)

Both Γ and $\Phi(C)$ satisfy condition (a) since $A_0 \subset \Gamma$ and $A_0 \subset \Phi(C)$ for all $C \in \Gamma$. In addition, by construction, if a subchain is in Γ, then it is a subchain in $\mathcal{F}_0$ which contains the union of the subchain. Since every element of the subchain is either contained or contains every set in $\mathcal{F}_0$, the union of all elements must also be either contained or contains every set in $\mathcal{F}_0$. Hence the union of the subchain is in Γ. Similarly, if a subchain exists in $\Phi(C)$, then the union of subchain exists in $\mathcal{F}_0$. Since each element in $\Phi(C)$ is either contained by C or contains $g(C)$, there are three possibilities:

i all elements are contained by C

ii all elements contain $g(C)$

iii some elements are contained by C, some contain $g(C)$

Clearly for (i) and (ii), the union of the subchain must also either be contained by C or contains $g(C)$, so it is in $\Phi(C)$. For (iii), since $C \subset g(C)$, the union of subchain must contain $g(C)$, hence it is in $\Phi(C)$.

Step 2.2: $\Phi(C)$ satisfies tower property (c), hence $\Phi(C)$ is a tower

Fix $C \in \Gamma$ and suppose $A \in \Phi(C)$. We want to show that $g(A) \in \Phi(C)$. If $A \in \Phi(C)$ there are three possibilities:

i $A \subset C$ and $A \neq C$: If A is a proper subset of C (i.e. $C \setminus A$ has at least one element), then C cannot be a proper subset of $g(A)$ (i.e. $\#(g(A) \setminus C) \geq 1$) otherwise $\#(g(A) \setminus A) \geq 1 + \#(C \setminus A) = 2$ which contradicts the assumption that $g(A) \setminus A$ has at most one element. Since $C \in \Gamma$ (i.e. $C \subset A$ or $A \subset C\ \forall A \in \mathcal{F}_0$) and $C \not\subset g(A)$, $g(A) \subset C$. Then by definition of $\Phi(C)$, $g(A) \in \Phi(C)$.

ii $A = C$: If $A = C$, then $g(A) = g(C)$ and $g(C) \subset g(C) = g(A)$ so $g(A) \in \Phi(C)$.

iii $g(C) \subset A$: If $g(C) \subset A$, then $g(C) \subset g(A)$ since $A \subset g(A)$. Thus, $g(A) \in \Phi(C)$.

In all three cases, $g(A) \in \Phi(C)$ so $\Phi(C)$ is a tower.

Step 2.3: $\Phi(C) = \mathcal{F}_0$ for all $C \in \Gamma$

By construction, $\Phi(C) \subset \mathcal{F}_0$. Additionally, by definition, $\mathcal{F}_0$ is the smallest tower, and since $\Phi(C)$ is a tower, $\mathcal{F}_0 \subset \Phi(C)$. Therefore, $\mathcal{F}_0 = \Phi(C)$ for all $C \in \Gamma$.

Step 2.4: Γ satisfies tower property (c), hence Γ is a tower

Since $\Phi(C) = \mathcal{F}_0$, it implies that for all $A \in \mathcal{F}_0$, by definition of $\Phi(C)$, either $A \subset C$ or $g(C) \subset A$. Since $C \subset g(C)$, we can rewerite it as $A \subset g(C)$ or $g(C) \subset A$. This is the definition of Γ, hence $g(C) \in \Gamma$. Since $C \in \Gamma \implies g(C) \in \Gamma$, Γ satisfies property (c) and it is a tower.

Step 2.5: $\Gamma = \mathcal{F}_0$

Since Γ is a tower and $\mathcal{F}_0$ is the smallest tower, $\Gamma \supset \mathcal{F}_0$. By construction, $\Gamma \subset \mathcal{F}_0$. Hence $\Gamma = \mathcal{F}_0$.

Step 2.6: $\mathcal{F}_0$ is totally ordered

By construction, Γ is totally ordered: let C_1, C_2 be elements in Γ. Then for all $A \in \mathcal{F}_0$, either $A \subset C_1$ or $C_1 \subset A$, and $A \subset C_2$ or $C_2 \subset A$. So we have four possibilities:

i $A \subset C_1$ and $A \subset C_2$ for all A: $C_1, C_2 \in \mathcal{F}_0$ and $\mathcal{F}_0$ is a tower, so C_1, C_2 are the union of $\mathcal{F}_0$ hence $C_1 = C_2$.

ii $A \subset C_1$ and $C_2 \subset A$ for some $A \implies C_2 \subset C_1$

iii $C_1 \subset A$ and $A \subset C_2$ for some $A \implies C_1 \subset C_2$

iv $C_1 \subset A$ and $C_2 \subset A$ for all A: Since $\mathcal{F}_0$ is a tower, $A_0 \in \mathcal{F}_0$, hence if $C_1 \subset A$ and $C_2 \subset A$ for all A, then $C_1 = C_2 = A_0$.

Since this is true for any pair in Γ, Γ is totally ordered, hence $\mathcal{F}_0$ is totally ordered.

Step 2.7: There exists $A \in \mathcal{F}_0$ for which $g(A) = A$

Let A be the union of $\mathcal{F}_0$. Since $\mathcal{F}_0$ is a tower, $A \in \mathcal{F}_0$. By property (c), $g(A) \in \mathcal{F}_0$. Since A is the largest member of $\mathcal{F}_0$, $A \supset g(A)$. Since $A \subset g(A)$ by definiton of g, we get $A = g(A)$.

□

Recall that

Definition. A choice function for a set X is a function f which associates to each nonempty subset E of X an element of $E : f(E) \in E$. In more informal terminology, f chooses an element out of each nonempty subset of X.

The Axiom of Choice. *For every set there is a choice function.*

Now we prove the main result:

Theorem. ***(Hausdorff's Maximality Theorem)*** *Every nonempty partially ordered set P contains a maximal totally ordered subset.*

(Before the proof, here is a quick example of what a collection of totally ordered set looks like:

$$\begin{aligned} a_n &= \{\phi_1, \phi_2, \phi_3, \dots, \phi_n\} \\ b_n &= \{\psi_1, \psi_2, \psi_3, \dots, \psi_n\} \\ P &= \{a_n | n \in \mathbb{N}\} \cup \{b_n | n \in \mathbb{N}\} \end{aligned}$$

If $\mathcal{F}$ is the collection of all totally ordered subsets of P, then $\{a_{12}, a_1, a_{2748}, a_{52}\} \in \mathcal{F}$ since $a_1 \preceq a_{12} \preceq a_{52} \preceq a_{2748}$ but $\{a_4, b_{26}, a_5\} \notin \mathcal{F}$.

Example of $\subseteq$-chain: $\{\{a_1\}, \{a_1, a_2\}, \{a_1, a_2, a_3\}, \dots\})$

Proof. **Step 1: Construct $\mathcal{F}$, a non-emtpy collection of subsets of a set P**

Let $\mathcal{F}$ be the collection of all totally ordered subsets of P. Since every subset of P which consists of a single element in totally ordered $\mathcal{F}$ is not empty.

Step 2: Union of any chain of totally ordered sets is totally ordered

Let $(I, \leq)$ be some totally ordered set of indices, and let $\{X_i : i \in I\} \subseteq \mathcal{F}$ be a $\subseteq$-chain of members in $\mathcal{F}$. In other words, each $X_i \subset P$ is totally ordered by $\subseteq$, and for any $i, j \in I$ we have $X_i \subseteq X_j$ iff $i \leq j$. Now we want to show that $\bigcup_{i \in I} X_i$ is totally ordered.

Let $x, x' \in \bigcup_{i \in I} X_i$, then there are $i, j \in I$ such that $x \in X_i$ and $x' \in X_j$. Without loss of generality, let $i \leq j$ which implies $X_i \subseteq X_j$ and thus $x, x' \in X_j$. Since X_j is totally ordered by $\subseteq$, we have that $x \subseteq x'$ or $x' \subseteq x$. Since x, x' are arbitrary, we see that $\bigcup_{i \in I} X_i$ is totally ordered.

(Example of $\subseteq$-chain: $\{\{a_1\}, \{a_1, a_2\}, \{a_1, a_2, a_3\}, \dots\} = \{X_1, X_2, X_3, \dots\}$ where $X_i \subseteq X_j$ iff $i \leq j$. Then for any $a_k, a_\ell \in \bigcup_{i \in I} X_i$, either $a_k \subseteq a_\ell$ or $a_\ell \subseteq a_k$)

Step 3: Construct a function g (defined in previous lemma)

Let f be a choice function for P. If $A \in \mathcal{F}$, let A^* be the set of all x in the complement of A such that $A \cup \{x\} \in \mathcal{F}$. If $A^* = \emptyset$, put $g(A) = A \cup \{f(A^*)\}$. If $A^* \neq \emptyset$, put $g(A) = A$.

Step 4: Apply previous lemma

By the lemma, $A^* = \emptyset$ for at least one $A \in \mathcal{F}$, and any such A is a maximal element of $\mathcal{F}$.

□

References

[1] Walter Rudin. *Real and Complex Analysis, 3rd Ed.* McGraw-Hill, Inc., USA, 1987.