

A gentle and friendly guide to the Bourbaki-Witt theorem and the Zorn's lemma

Mohamad Bakri Zubir

Institut Sains Matematik, Universiti Malaya 50603 Kuala Lumpur
mbakri@um.edu.my

ABSTRACT

In this article we discuss partially ordered sets and the Bourbaki-Witt Theorem. The theorem is then used to prove Zorn's Lemma without using the Axiom of Choice. The purpose of the write-up is to provide a gentle introduction to some topics in Set Theory and dissect the proof of Bourbaki-Witt theorem.

Keywords: poset, inflationary function, Bourbaki-Witt, Zorn's lemma

PRELIMINARIES

We introduce some basic definitions to be used throughout the article. Where possible visual representations of concepts related to partially ordered sets are provided to aid understanding especially in the proof of the Bourbaki Witt theorem.

Definition 1. A partially ordered set or poset is a pair (X, \leq) , where X is a set and \leq is a relation on X that satisfies

- (i) for all $x \in X, x \leq x$ (reflexivity)
- (ii) for all $x, y \in X$, if $x \leq y$ and $y \leq x$ then $x = y$ (antisymmetry)
- (iii) for all $x, y, z \in X$ if $x \leq y$ and $y \leq z$ then $x \leq z$ (transitivity).

Remark 1. If $a \leq b$ and $a \neq b$, we write $a < b$.

Definition 2 (Comparability). The elements x, y in a poset is said to be *comparable* if either $x \leq y$ or $y \leq x$.

Definition 3 (Chain and antichain). In a poset, a subset S is a *chain* if it is totally ordered, that is, for all $x, y \in S$, the elements x and y are comparable. An *antichain* is a subset in which any two distinct element are incomparable.

Remark 2. The word chain is used interchangeably with the phrase totally ordered set in this article.

Definition 4 (Hasse diagram). A Hasse diagram is a pictorial representation of a finite partial order on a set. Each element is shown as a vertex. Two elements are connected by a line if and only if they are related.

Definition 5 (Cover). An element y of a poset *covers* an element x provided that there is no element z in the poset for which $x \leq z \leq y$.

Definition 6 (Maximum element). In a poset X , an element $M \in X$ is *maximum* if for any element $x \in X, x \leq M$.

Definition 7 (Maximal element). In a subset S of a poset X , an element $m \in X$ is *maximal* if there is no element $s \in S$ such that $m \leq s$ and $s \neq m$.

By using symbolic logic, the definition is

- for all $s \in S$, it is false that $(m \leq s)$ and $(s \neq m)$,
- or for all $s \in S$, $\neg[(m \leq s) \wedge (s \neq m)]$,
- or for all $s \in S$, $\neg(m \leq s) \vee \neg(s \neq m)$,
- or for all $s \in S$, $\neg(m \leq s) \vee (s = m)$,
- or for all $s \in S$, $(m \leq s) \rightarrow (s = m)$.

Therefore the alternative definition for being maximal is for any $s \in S$, if $m \leq s$ then $s = m$. This definition is usually used to prove the maximality of a given element.

The differences between a maximum and a maximal element are:

- M is maximum: no element above M and all elements are below it
- m is maximal: no element above m but not all elements are below it.

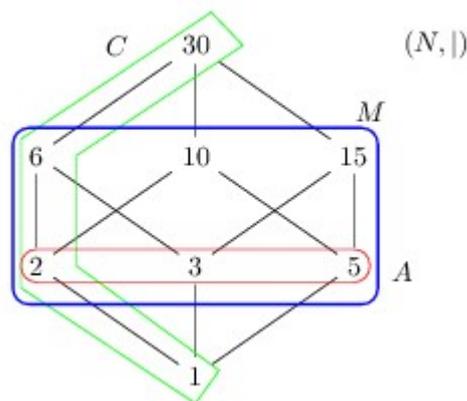
In a poset if the maximum exists, it is unique, whereas we can have more than one maximal element.

Minimum and minimal elements are defined similarly.

Example 1. Let $N = \{1, 2, 3, 5, 6, 10, 15, 30\}$ with the “divides” relation on N : For all $a, b \in N$,

$$a \mid b \Leftrightarrow b = ka \text{ for some integer } k.$$

The Hasse diagram for the poset is



The set N is a poset under the relation. Here

- $C = \{1, 2, 6, 30\}$ is a chain
- $A = \{2, 3, 5\}$ is an antichain
- $M = \{2, 3, 5, 6, 10, 15\}$ has maximal elements 6, 10, 15, minimal element 2, 3, 5, maximum element is 30 and minimum element is 1
- 6 covers 2 but 30 does not cover 2

□

Definition 8 (Upper bounds and supremum). The *upper bound* for any subset S of X is an $x \in X$ such that for all $s \in S$ we have $s \leq x$.

The element $u \in X$ is a *least upper bound*, *supremum* or *join* of S , written $u = \sup S$ or $u = \vee S$ if

- (i) u is an upper bound for S ,
- (ii) for all $x \in X$, if x is an upper bound for S , then $u \leq x$.

That is, to be a least upper bound, an element must be an upper bound first, and for all upper bounds it is the smallest.

Likewise we have lower bounds and infimum.

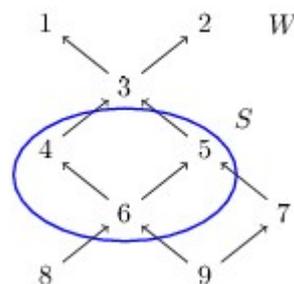
Definition 9 (Lower bounds and infimum). The *lower bound* for any subset S of X is an $x \in X$ such that for all $s \in S$ we have $x \leq s$.

The element $l \in X$ is a *greatest lower bound*, *infimum* or *meet* of S , written $l = \inf S$ or $l = \wedge S$ if

- (i) l is a lower bound for S ,
- (ii) for all $x \in X$, if x is a lower bound for S , then $x \leq l$.

To be a greatest lower bound, an element must be a lower bound first, and for all lower bounds it is the largest.

Example 2. Let $W = \{1, 2, 3, \dots, 8, 9\}$ be ordered as follows:



Consider the subset $S = \{4, 5, 6\}$. The set of upper bounds for S is $V = \{1, 2, 3\}$. Here $\sup S = 3$ because it is the smallest element in the set V .

The set of lower bounds for S is $U = \{8, 9\}$. However S has no infimum because U has no maximum since 8 and 9 are not comparable. The element 7 is not a lower bound for S because we cannot say 7 is below 4.

□

Our aim is to give the proof of Zorn's Lemma by using the Bourbaki-Witt theorem. Otherwise the Axiom of Choice has to be used to prove the Zorn's Lemma.

THE BOURBAKI-WITT THEOREM

Given a function f of a set X into X (or into some set containing X), one of the few questions we can ask is whether some point is mapped into itself; that is, does the equation

$$f(x) = x$$

have a solution? If so x is called a *fixed point* of f . The Bourbaki-Witt theorem is a basic fixed point theorem for partially ordered sets.

Definition 10. A function $f: X \rightarrow X$ is *inflationary* if $x \leq f(x)$ according to the given order \leq . That is, $f(x)$ is always as big as x .

Remark 3. An inflationary function is not the same as a monotonic function. A function g is monotonic (increasing) if for any $x \leq y$ implies that $g(x) \leq g(y)$ for a given order \leq . Not every monotonic function is inflationary.

Example 3. The function $f(x) = |x|$, $x \in \mathbb{R}$ is inflationary because $x \leq |x|$. However it is not monotonic because for example although $-3 \leq 1$ but $|-3| \not\leq |1|$.

□

Definition 11. A *fixed point* of a function $f: X \rightarrow X$ is an x such that $f(x) = x$.

Definition 12. A set X is *inductively* ordered set if every non-empty totally ordered subset of X has an *upper bound*. The set X is *strictly* inductively ordered if every non-empty totally ordered subset has a *least upper bound*.

Theorem 1 (Bourbaki-Witt). Let X be a (non-empty) strictly inductively ordered set. Suppose $f: X \rightarrow X$ be such that $x \leq f(x)$ for all $x \in X$. Then there exists an $s \in X$ such that $s = f(s)$.

The proof is deferred until we have proven two important lemmas.

The intuitive idea is that given an element $a \in X$, we can produce an ascending sequence

$$a \leq f(a) \leq f^2(a) \leq \dots \leq f^n(a) \leq \dots \quad (1)$$

If the theorem were false, then f did not have a fixed point above, then all inequalities in (1) would be strict. If X is finite, then we would obtained an infinite set of elements $f^n(a)$, $n \in \mathbb{N}$, reaching a contradiction.

If X were totally ordered then by assumption it would have a least upper bound $s \in X$ so

$$\begin{aligned} s &\leq f(s), && \text{because } f \text{ is inflationary} \\ &\leq s, && \text{because } s \text{ is the least upper bound} \end{aligned} \quad (2)$$

therefore $s = f(s)$.

Suppose X is not totally ordered. What we would like to do is capture a subset M of X that is totally ordered with two properties, the least upper bound s of M is in M , and for all $x \in M$, $f(x) \in M$. With these properties, then by (2), the element s is the fixed element.

Definition 13. A subset A of X is *admissible* if it has the following properties

- (a) it contains $a \in X$,
- (b) it is closed under f , that is for all $x \in A$, $f(x) \in A$,
- (c) if T is any totally ordered subset of A , then the least upper bound of T is in A .

The set X itself is admissible and so we can define

$$M = \bigcap \{A \mid A \text{ is an admissible subset of } X\}.$$

The set M is admissible because

- (a) if $a \in A$ for each admissible set A then a belongs to the intersection M ,
- (b) if $x \in M$ then x is in every admissible set A , so $f(x)$ is in every admissible set. Whence $f(x) \in M$,
- (c) if T is a totally ordered subset of M , and b is the least upper bound of T in X , then T lies in every admissible subset A . By definition of admissible set, b belongs to every admissible subset A and hence lies in M .

Therefore under order by inclusion, M is contained in every admissible subset of X , so it is the **smallest** admissible subset of X .

We show that the element a is the minimum element of M . Let $A_0 = \{x \in M \mid a \leq x\}$. The set A_0 is admissible because

- (a) the element $a \in A_0$ by definition,
- (b) for any $x \in A_0$, by definition $a \leq x$, and $x \leq f(x)$ because f is inflationary. This implies $a \leq x \leq f(x)$, whence $f(x) \in A_0$,
- (c) if T is any chain in A_0 with least upper bound u in X , then for all $t \in T$, we have $a \leq t \leq u$. Then by definition, $u \in A_0$.

By definition of A_0 we have $A_0 \subseteq M$, and since by order of inclusion $M \subseteq A_0$, therefore $A_0 = M$ and hence a is the minimum element.

Observe that the chain

$$a \leq f(a) \leq f^2(a) \leq \dots \leq f^n(a) \leq \dots$$

is in M .

Our task is to show that M is a chain. That is, given any $x, c \in M$,

$$x \leq c \text{ or } c \leq x,$$

and once we achieve that, by (2), M contains a fixed element.

There are two things that we want elements in the set M to satisfy:

- (i) for any $x \in M$ and **some** $c \in M$, either $x \leq c$ or $c \leq x$,
- (ii) for any $x \in M$ and **any** $c \in M$, either $x \leq c$ or $c \leq x$.

We do it in two stages.

Stage 1

For every $x \in M$ we would like to show that f jumps directly from x to $f(x)$, and in particular no element is in between. Here we need to introduce extreme element of M .

Definition 14. An element $c \in M$ is *extreme* iff for every $x \in M$ whenever $x < c$ then $f(x) \leq c$.

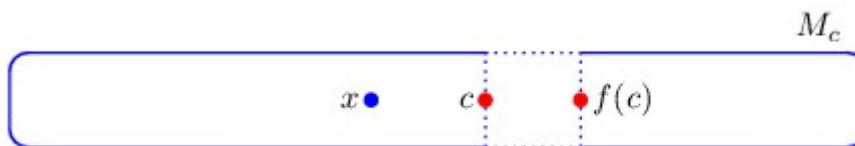
This begs the question, does extreme element exist? Consider the minimum element a . If it was not extreme, then there would be an $x \in M$ such that $x < a$ but $a < f(x)$. But this contradicts the fact that there is no such x below a . So the minimum element a is an extreme element.

Let $a \leq \dots \leq x$ be a chain in M and $x < c$. By hypothesis the chain has least upper bound u . If we take $u = f(x)$ then $x \leq f(x)$ and $x \leq f(x) \leq c$. If $x < c$ but $c < f(x)$, this would contradict $f(x)$ being the least upper bound for the chain. Therefore an extreme element, apart from a , must exist in M , and there is no element between c and $f(c)$.

Then for any chain $a \leq \dots \leq z$, if we define the function f to be such that $f(z)$ is the least upper bound for the chain, then f is inflationary.

With this extreme element c we form the set

$$M_c = \{x \in M \mid x \leq c \text{ or } f(c) \leq x\}.$$



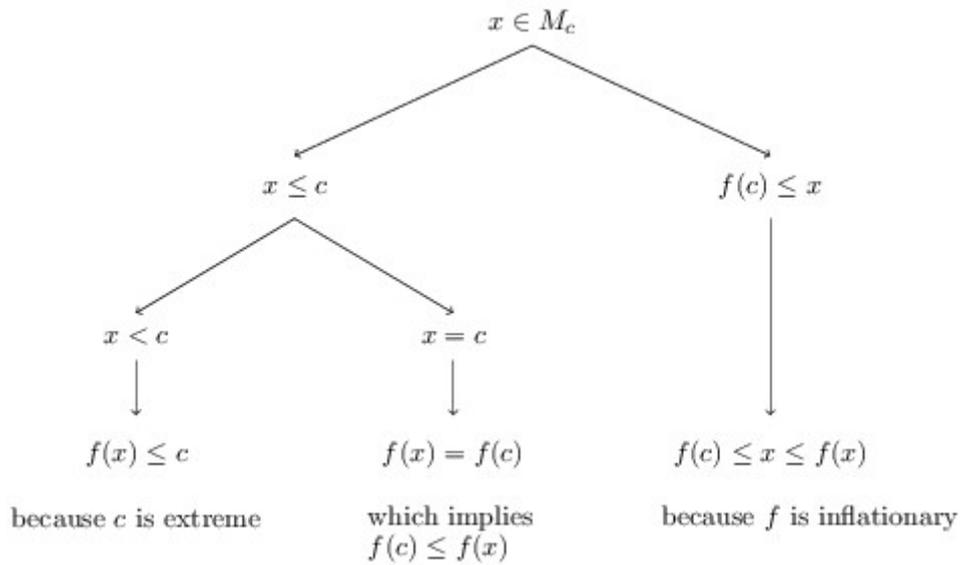
We want to show for any $x \in M$, either $x \leq c$ or $c \leq x$.

Lemma 1. For each extreme point c , $M_c = M$.

Proof. To show $M_c = M$ we show M_c is admissible. By definition $M_c \subseteq M$, and since under order by inclusion, M is the smallest admissible set, $M \subseteq M_c$, whence $M_c = M$.

(a) The element $a \in M_c$ because a is the minimum element.

(b) Let $x \in M_c$. We want to show $f(x) \in M_c$, that is $f(x) \leq c$ or $f(c) \leq f(x)$. Now $x \leq c$ or $f(c) \leq x$ by definition of M_c .



There are three cases to consider. Suppose $x \leq c$.

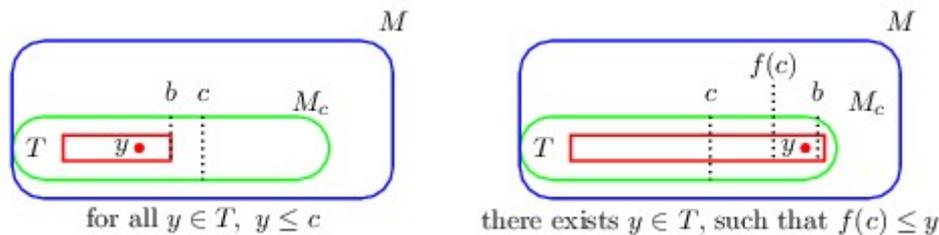
If $x < c$ then $f(x) \leq c$ because c is extreme.

If $x = c$ we have $f(x) = f(c)$ which implies $f(c) \leq f(x)$.

If $f(c) \leq x$ then $f(c) \leq x \leq f(x)$ because f is inflationary.

So if $x \in M_c$ we have either $f(x) \leq c$ or $f(c) \leq f(x)$. By definition $f(x) \in M_c$.

(c) Let T be a totally ordered subset of M_c with least upper bound $b \in X$. We want show $b \in M_c$, that is $b \leq c$ or $f(c) \leq b$. Since $T \subseteq M_c \subseteq M$ and M is admissible then $b \in M$. This is because by definition, any totally ordered subset of M has a least upper bound in M .



If for all $y \in T, y \leq c$ then $b \leq c$ because b is the least upper bound. Otherwise there exists a $y \in T$ such that $f(c) \leq y$, as we have shown earlier that there is no element between c and $f(c)$. This implies $f(c) \leq y \leq b$.

So we have shown $b \leq c$ or $f(c) \leq b$. By definition of M_c we have $b \in M_c$.

Therefore M_c is admissible and $M \subseteq M_c$, whence $M = M_c$. □

Stage 2

Our next task is to show that every element of M is actually extreme. We collect all extreme elements in the set E .

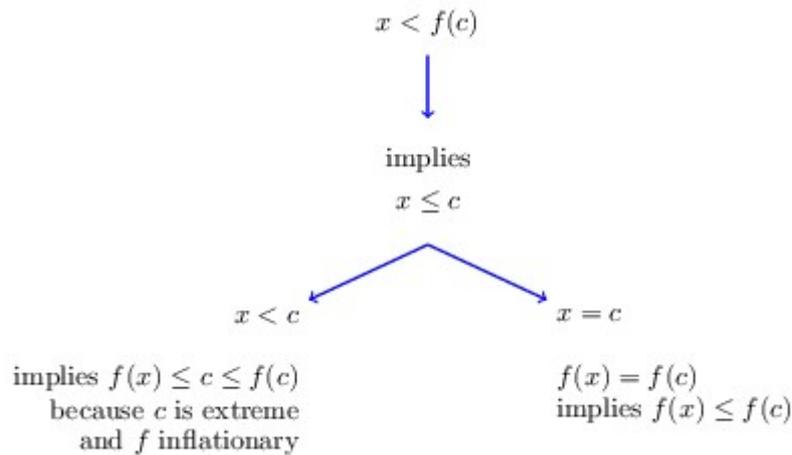
Lemma 2. Let $E = \{c \in M \mid \text{for all } x \in M \text{ if } x < c \text{ then } f(x) \leq c\}$, that is E is the set of all extreme points of M . Then $E = M$.

By definition $E \subseteq M$. We show E is admissible, and therefore $M \subseteq E$, whence $E = M$.

Proof.

(a) The element a belongs to E vacuously. Suppose $a \notin E$. Then there exists an $x \in M$ such that $x < a$ but $f(x) \not\leq a$ or $f(x) > a$. But this is a contradiction because there is no such x since a is the minimum element.

(b) Given any $c \in E$ we want to show $f(c) \in E$, that is for any $x \in M$, if $x < f(c)$ then $f(x) \leq f(c)$.



If $x < f(c)$ then $x \leq c$. If $x < f(c)$ but $x > c$ we have $c < x < f(c)$, a contradiction because there is no element between c and $f(c)$.

There are two cases:

If $x < c$ then $f(x) \leq c \leq f(c)$, because c is extreme and f inflationary.

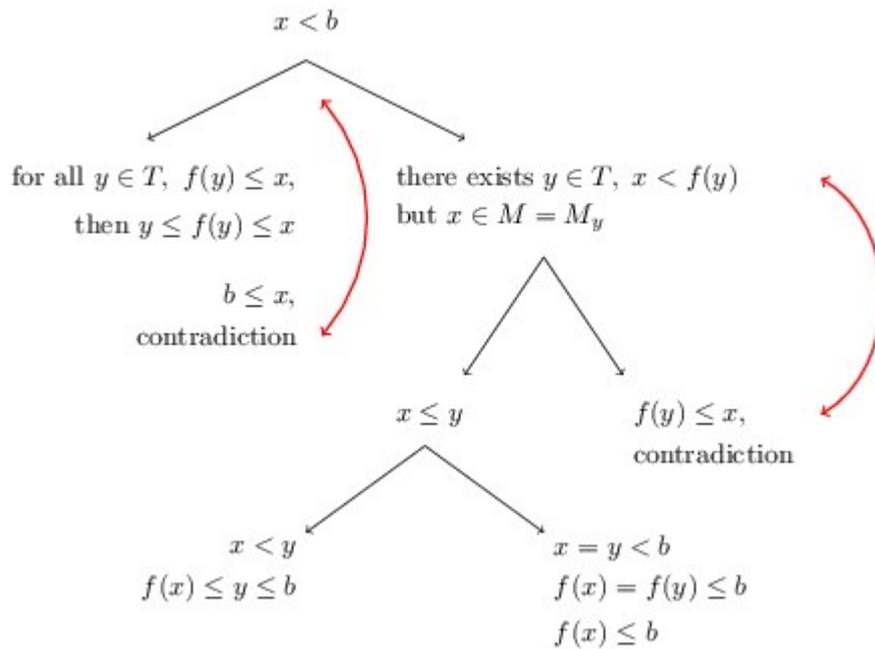
If $x = c$ then $f(x) = f(c)$ which implies $f(x) \leq f(c)$.

So we have shown if $x < f(c)$ then $f(x) \leq f(c)$, and by definition, $f(c)$ is extreme and therefore is in E .

(c) We want to show that if T is any totally totally order subset of E , the least upper bound b of T in X belongs to E . Since $T \subseteq E \subseteq M$ and M is admissible, $b \in M$.

That is we show b is an extreme element. By definition, we have to show for all $x \in M$ if $x < b$ then $f(x) \leq b$.

Since b is the supremum of T we have for all $y \in T$, $y \leq b$. By our assumption $x < b$.



If for all $y \in T$ we have $f(y) \leq x$, then $y \leq f(y) \leq x$. This implies x is an upper bound for T and therefore $b \leq x$, a contradiction to our assumption.

Therefore there must be a $y \in T$ such that $x < f(y)$.

Since $x \in M = M_y$, we must have $x \leq y$. We cannot have $f(y) \leq x$ because it would contradict $x < f(y)$.

There are two cases:

If $x < y$ then $f(x) \leq y \leq b$ because y is extreme and b is the least upper bound for each $y \in T$.

If $x = y$ then $y = x < b$ and because y is extreme, $f(y) \leq b$, and therefore $f(x) = f(c) \leq b$.

□

We can now prove the Bourbaki-Witt theorem because

Theorem 2. The subset M of X is totally ordered.

Proof. Let $x < y \in M$. Then x is an extreme point of M and $y \in M_x$ so

$$y \leq x \text{ or } x \leq f(x) \leq y,$$

proving that M is totally ordered. Then by (2) f has a fixed point.

□

THE ZORN'S LEMMA

Zorn's Lemma is a fundamental axiom of logic. It guarantees the existence of maximal elements in partially ordered sets. Zorn's lemma is used most often in the situation when \mathcal{A} is a family of subsets of a set X ordered by inclusion and such that for every chain $T \subseteq \mathcal{A}$, its union $\bigcup T$ is also in \mathcal{A} .

In most literature, Zorn's Lemma is proved by using the Axiom of Choice. Here we use the Bourbaki-Witt theorem instead.

As a corollary to the Bourbaki-Witt theorem we obtain the weaker form of Zorn's lemma. In the weaker form, every totally ordered subset has a least upper bound. We then prove Zorn's lemma.

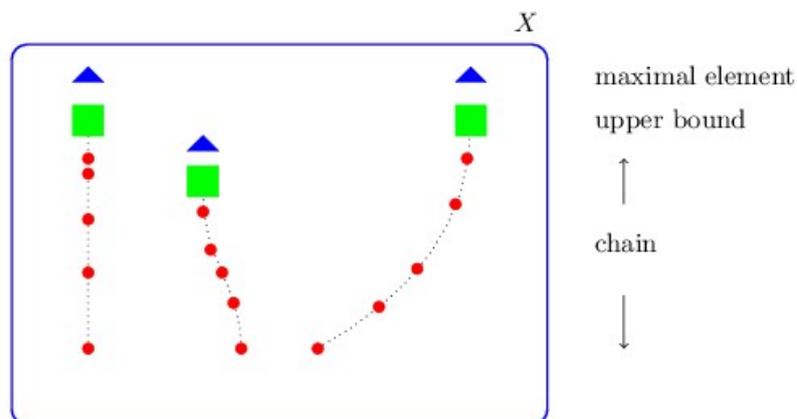
Corollary 1 (Weak form of Zorn's lemma). Let X be a non-empty strictly inductive ordered set, that is every totally ordered subset has a least upper bound. Then X has a maximal element.

Proof. We prove by contradiction. Suppose X does not have a maximal element. Then for any $x \in A$, there exists an element $y_x \in A$ such that $x < y_x$. Let $f: X \rightarrow X$ be the map such that $f(x) = y_x$ for all $x \in X$. Then f is inflationary. Now X and f satisfy the hypothesis of the Bourbaki-Witt theorem, and applying the theorem yields a contradiction. \square

We can now state and prove the Zorn's lemma.

Theorem 3 (Zorn's Lemma). Let X be a (non-empty) posets in which every chain has an upper bound in X . Then it has a maximal element (in X).

It must be emphasized that the upper bound of any given *must* be in X . Otherwise we cannot conclude it has a maximal element. Also the lemma does not tell what the maximal element is, only its existence.



Proof. There are two steps in the proof.

Step 1

The set X being non-empty, has subsets. Some of these subsets contain elements which are totally ordered and some not.

We collect all non-empty totally ordered subsets of X in the set \mathcal{A} . For all $C, D \in \mathcal{A}$, we define $C \leq D$ to mean $C \subseteq D$. Then \mathcal{A} , under order of inclusion, is a partially ordered set. In fact it is strictly inductively ordered.

Let $T = \{C_i\}_{i \in I}$ be a totally ordered subset of \mathcal{A} . We reiterate that $C_1 \subseteq C_2 \subseteq C_3 \subseteq \dots \subseteq C_n \subseteq \dots$ where the elements in each C_i are themselves totally ordered. (Figure 1)

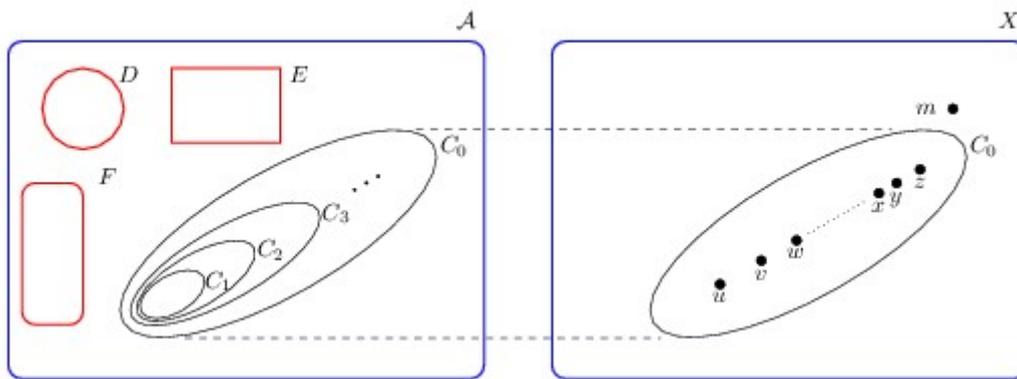


Figure 1 Totally ordered subsets whose elements are totally ordered

So what is the upper bound for the chain T ? Well, take the union of all C_i s, that is, let $Z = \bigcup_{i \in I} C_i$. Then for all $i \in I$, we have $C_i \subseteq Z$. Is Z is also in \mathcal{A} ?

Let $x, y \in Z$. Then $x \in C_i$ and $y \in C_j$ for some $i, j \in I$. Since T is totally ordered, then $C_i \subseteq C_j$, say. Then $x, y \in C_j$. Because C_j is itself totally ordered, $x \leq y$ or $y \leq x$. Thus Z is totally ordered and therefore belongs to \mathcal{A} .

The subset Z is in fact a least upper bound for T in \mathcal{A} . This is because if Ω is an upper bound for T then Ω contains C_i s and therefore contains Z , the union of all C_i s.

By the weaker form of Zorn's Lemma, Corollary (1), \mathcal{A} has a maximal element C_0 . This means C_0 is a maximal, totally ordered subset of S .

Step 2

The maximal element C_0 is a totally ordered subset of S . Let m be an upper bound C_0 in S . Then m is the desired maximal element of X . (Figure 1)

To show m is maximal we must show for each $x \in X$ if $m \leq x$ then $m = x$. Consider the subset $C_0 \cup \{x\}$ of X . This set is totally ordered, and therefore must be equal to C_0 by the maximality of C_0 . Whence $x \in C_0$ and since m is the upper bound for x , we have $x \leq m$. Hence $m = x$, which was what we wanted to prove.

□

Example 4. Let

\mathbb{N} = the set of natural numbers

$\mathcal{Q} = \mathcal{P}(\mathbb{N})$, the power set of \mathbb{N}

\mathcal{S} the family of finite subset of \mathbb{N}

$\mathcal{C} = \{\{1\}, \{1, 3\}, \{1, 3, 5\}, \dots, \}$

Then \mathcal{C} is a chain in each of \mathcal{Q} and \mathcal{S} . The upper bound of \mathcal{C} is $\bigcup_{C \in \mathcal{C}} C = \{1, 3, 5, \dots\}$. Then $\bigcup_{C \in \mathcal{C}} C \in \mathcal{Q}$ but $\bigcup_{C \in \mathcal{C}} C \notin \mathcal{S}$. Observe that \mathcal{Q} has a maximal element, which is \mathcal{N} . However \mathcal{S} has no maximal element, if there was, the maximal element would be a finite subset which would be a proper subset of many other finite subsets of \mathbb{N} . \square

APPLICATION OF ZORN'S LEMMA

One of the widely discussed application of the Zorn's Lemma is to prove

Corollary 2. *Every vector space V has a basis.*

Proof. A subset B of a vector space V is a basis if it linearly independent and spans V .

We collect all subsets B of V which are linearly independent in the set \mathcal{J} . Under order of inclusion, \mathcal{J} is a poset. There are two steps involved. The first step is showing that any chain in \mathcal{J} has an upper bound and therefore by Zorn's lemma has linearly independent maximal subset. The second step is showing the maximal subset spans the vector space.

Step 1

Let $T = \{B_i\}_{i \in I}$ be a chain in \mathcal{J} . Let $Z = \bigcup_{i \in I} B_i$. Then $Z \supseteq B_i$ for all $i \in I$, so Z is an upper bound for the chain T . If we can show Z is linearly independent, then $Z \in \mathcal{J}$ and so by Zorn's Lemma, \mathcal{J} it has a maximal linearly independent subset.

Suppose $\lambda_1 b_1 + \lambda_2 b_2 + \dots + \lambda_n b_n = 0$ for some $b_1, b_2, \dots, b_n \in Z$ and scalars $\lambda_1, \lambda_2, \dots, \lambda_n$. Each element $b_i \in B_i$ and since T is a chain, there is an m such that for $i = 1, \dots, n$, all $b_i \in B_m$. Since B_m is linearly independent all λ_i s are zero. Therefore Z is linearly independent and therefore in \mathcal{J} .

By Zorn's lemma, \mathcal{J} has a maximal linearly independent subset B_0 . If we can show B_0 spans V then we are done.

Step 2

Suppose B_0 did not span V . Choose any nonzero $w \in V \setminus \text{span}(B_0)$. We show the set $B_0 \cup \{w\}$ is linearly independent. This would contradict the maximality of B_0 .

We note that for any nonzero scalar a and nonzero vector w , the product $aw \neq 0$. If not, $w = a^{-1}aw = a^{-1}0 = 0$, a contradiction.

BIBLIOGRAPHY

- Forster, T. (2003), *Logic, induction and sets*. Vol. 56. Cambridge University Press.
- Goubault-Larrecq, J. (2013). *Non-Hausdorff topology and domain theory: Selected topics in point-set topology* (Vol. 22). Cambridge University Press.
- Kuper, L. (2015, August 31). "Composition.al." *What's the Difference between Inflationary and Monotonic Functions?* Retrieved from <http://composition.al/blog/2015/08/31/whats-the-difference-between-inflationary-and-monotonic-functions/>.
- Lang, S. (1993). *Real and functional analysis*, Volume 142 of Graduate Texts in Mathematics. Springer-Verlag, New York.
- Lipschutz, S (1964). *Theory and problems of set theory and related topics*. McGraw-Hill.
- Väth, M. (2012). *Topological analysis: from the basics to the triple degree for nonlinear Fredholm inclusions* (Vol. 16). Walter de Gruyter.