Zorns Lemma, or Why Every Vector Space Has a Basis

Notes by Michael Fochler, Department of Mathematical Sciences, Binghamton University,
for a talk given to the Binghamton University Undergraduate Math Club on Nov.29, 2016

0.0.1 Sets

- **a.** Sets X are collections of stuff (elements); $x \in X$: x is an element of X
- • Duplicates and order of elements are ignored: $X = \{-1, 1, -1, 1, \dots\} = \{-1, 1\} = \{1, -1\}$
- **b.** Sets can contain sets, e.g., $\mathfrak{U} = \{\,]a, b[\,: a < b\}$; Powerset $2^X = \{A : A \subseteq X\}$ (all subsets of X)
- • We assume for collections of sets $\mathfrak{U}$ that $\mathfrak{U} \subseteq 2^\Omega$ for some suitable "universal set" Ω
- **c.** Intersections: $x \in A_1 \cap A_2 \;\Leftrightarrow\; x \in A_1$ **and** $x \in A_2$
- • $x \in A_1 \cap \dots \cap A_n = \bigcap_{j=1}^n A_j \;\Leftrightarrow\; x \in A_j$ for **all** $1 \leqq j \leqq n$ —
- • Collection of sets $\mathfrak{U}$; $x \in \bigcap \mathfrak{U} = \bigcap \big[U : U \in \mathfrak{U}\big] \;\Leftrightarrow\; x \in U$ for **all** $U \in \mathfrak{U}$
- **d.** Unions: $x \in A_1 \cup A_2 \;\Leftrightarrow\; x \in$ **at least one** of A_1, A_2
- • $x \in A_1 \cup \dots \cup A_n = \bigcup_{j=1}^n A_j \;\Leftrightarrow\; x \in A_j$ for **at least one** $1 \leqq j \leqq n$ —
- • Collection of sets $\mathfrak{U}$; $x \in \bigcup \mathfrak{U} = \bigcup \big[U : U \in \mathfrak{U}\big] \;\Leftrightarrow\; x \in U$ for **at least one** $U \in \mathfrak{U}$
- **e.** Set difference $A \setminus B = \{x \in A : x \notin B\}$
- • Complement $A^\complement$ of $A \subseteq \Omega$: $A^\complement = \Omega \setminus A$; $A, B \subseteq \Omega \;\Rightarrow\; A \setminus B = A \cap B^\complement$

0.0.2 Types of numbers

- **a.** Natural numbers: $\mathbb{N} = \{1, 2, 3, \dots\}$; Integers: $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$
- **b.** Rational #s: $\mathbb{Q} = \{\frac{n}{d} : n, d \in \mathbb{Z}, d \neq 0\}$; $5/8 \in \mathbb{Q}$, $-7 = \frac{-7}{1} \in \mathbb{Q}$, $1.25 = \frac{5}{4} \in \mathbb{Q}$, $0.33\bar{3} = \frac{1}{3} \in \mathbb{Q}$
- **c.** Real #s: $\mathbb{R} = \{$ all decimals $\} = \{m + \sum_{j=1}^\infty d_j 10^{-j} : m \in \mathbb{Z}$ and $d_j = 0, 1, 2, \dots, 9$ (digits) $\}$
  $\pi, \sqrt{2} \in \mathbb{R}$ but $\pi, \sqrt{2} \notin \mathbb{Q}$ Intervals $[a, b[= \{x \in \mathbb{R} : a \leqq x < b\}$, $]a, b[= \{x \in \mathbb{R} : a < x < b\}, \dots$

0.0.3 Functions $f : X \to Y$, $x \mapsto f(x)$

- **a.** Domain $X \neq \emptyset$ (source of arguments), Codomain $Y \neq \emptyset$ (target contains function values $f(x)$,
  assignment $x \mapsto f(x)$; can write $X \xrightarrow{f} Y$ instead of $f : X \to Y$
- **b.** Example 1: $f : [10, \infty[\,\to\,]-20, \infty[\,$, $x \mapsto f(x) = \sqrt{x - 1}$
- – Example 2: $g : [10, \infty[\,\to\,]3, \infty[\,$, $x \mapsto g(x) = \sqrt{x - 1}$
- – Example 3: $h : [1, 101] \to [0, 10]$, $x \mapsto h(x) = \sqrt{x - 1}$
- • f, g, h are **different** because domains and/or codomains to not match
- **c.** Function $f : X \to Y$, $x \mapsto f(x)$; $\emptyset \neq X' \subseteq X$;
  $f\big|_{X'} : X_1 \to Y$, $x \mapsto f'(x) := f(x)$ is the restriction of f to X' and f is an extension of f' to X

0.0.4 Cardinality

- **a.** finite set X: card$(X) = \#$ of elements in X; empty set $\emptyset$ is finite (no elements) $\Rightarrow$ card$(\emptyset) = 0$
- • X is countably infinite if not finite but can be enumerated (sequenced): $X = \{x_1, x_2, x_3, \dots\}$
- • X is countable if finite or countably infinite
- • X is uncountable if it cannot be sequenced
- **b.** B countable, $A \subseteq B \ \Rightarrow\ A$ countable
- – Proof: discard b_j from $B = \{b_1, b_2, \dots\}$ if $b_j \notin A$
- **c.** A countable union $\bigcup_{n\in\mathbb{N}} A_n$ of countable sets A_n is countable.
- – Proof: $A_n = \{a_{n,1}, a_{n,2}, \dots\}$; traverse the finite diagonals $D_k = \{a_{i,j} : i + j = k\}$ in order, starting with $D_2 = \{a_{1,1}\}$. Skip duplicates and empty slots.
- **d.** Fractions (rational #s) $\mathbb{Q}$ is countable: $\mathbb{Q} = Q_1 \cup Q_2 \cup \cdots =$ countable union of finite sets $Q_n =$ all fractions with denominator n between $-n$ and n: $Q_n = \{-\frac{n^2}{n}, -\frac{n^2-1}{n}, -\frac{n^2-2}{n}, \dots, \frac{n^2-2}{n}, \frac{n^2-1}{n}, \frac{n^2}{n}\}$
- **e.** Decimals $\mathbb{R}$ is uncountable because even the subset $A = \{\sum_{j=1}^{\infty} d_j 10^{-j} : d_j = 0, 1, 2, \dots, 8\}$ (digit 9 is excluded) is uncountable.
- – Proof: Write $m.d_1 d_2 \dots$ for $m + \sum_{j=1}^{\infty} d_j 10^{-j}$. Assume A is countable: $A = \{x_1, x_2, \dots\}$.

$$\begin{aligned}
x_1 &= 0.d_{1,1}d_{1,2}d_{1,3}\dots \\
x_2 &= 0.d_{2,1}d_{2,2}d_{2,3}\dots \\
x_3 &= 0.d_{3,1}d_{3,2}d_{3,3}\dots \\
&\dots\dots\dots\dots\dots\dots \\
x_n &= 0.d_{n,1}d_{n,2}d_{n,3}\dots d_{n,n}\dots \\
&\dots\dots\dots\dots\dots\dots
\end{aligned}$$

Construct $x = 0.d_1 d_2 d_3 \dots d_n \dots$ as follows:

$d_1 = 4$ if $d_{1,1} = 3$ and 3 if $d_{1,1} \neq 3$, hence $x \neq x_1$;
$d_2 = 4$ if $d_{2,2} = 3$ and 3 if $d_{2,2} \neq 3$, hence $x \neq x_2$;
$d_3 = 4$ if $d_{3,3} = 3$ and 3 if $d_{3,3} \neq 3$, hence $x \neq x_3$;
$\dots\dots\dots\dots\dots\dots$
$d_n = 4$ if $d_{n,n} = 3$ and 3 if $d_{n,n} \neq 3$, hence $x \neq x_n$;

Result: $x \in A$ although $A = \{x_1, x_2, \dots\}$ and $x \neq x_j$ for all j. **Contradiction!**

0.0.5 Vector spaces

(linear spaces)

- **a.** vector space (VS) V: Let $x, y, z \in V$, $\alpha, \beta, \gamma \in \mathbb{R}$;
- • addition $(x, y) \mapsto x + y$: commutativity: $x + y = y + x$; associativity: $(x + y) + z = x + (y + z)$ zero vector $0 \in V$: $x + 0 = x$ for all $x \in V$; Negative $-x$ of x: $x + (-x) = 0$; $x - y := x + (-y)$
- • scalar multiplication $(\alpha, x) \mapsto \alpha \cdot x = \alpha x$: $\alpha(\beta x) = (\alpha\beta)x$; $1x = x$;
- • distributivity: $(\alpha + \beta)x = \alpha x + \beta x$; $\alpha(x + y) = \alpha x + \alpha y$
- **b.** Linear combinations are sums $\sum_{j=0}^{n} \alpha_j x_j = \alpha_1 x_1 + \ldots + \alpha_n x_n$ of scalar multiples of vectors $x_1, \dots, x_n \in V$, scalars $\alpha_1, \dots, \alpha_n \in \mathbb{R}$;
- **c.** Nonempty $U \subseteq V$ is sub(–vector)space if $a, b \in A$ and $\alpha \in \mathbb{R} \ \Rightarrow\ a + b \in A$ and $\alpha\, a \in A$;
- • nullspace $\{0\}$ and V are subspaces of V; subspaces are VS
- • $U \subseteq V$ is subspace $\Leftrightarrow$ any lin. comb. of vectors in A belongs to A.
- • Any intersection of subspaces (arbitrarily many) is a subspace

d. $A \subseteq V$, $A \neq \emptyset$; (linear) span $span(A) = \{ \sum_{j=1}^{k} \alpha_j x_j : k \in \mathbb{N}, \alpha_j \in \mathbb{R}, x_j \in A \ (1 \leqq j \leqq k) \}$
$= \{$ all lin. combs of vectors in $A\} = \bigcap \left[W : W \text{ is subspace and } W \supseteq A \right]$ = subspc generated by A

e. $A \subseteq V$, $A \neq \emptyset$ is linearly dependent (LD) if there $(k \in \mathbb{N})$ and scalars $\alpha_1, \alpha_2, \ldots \alpha_k \in \mathbb{R}$ and distinct $x_1, x_2, \ldots x_k \in A$ such that not all scalars α_j are zero $(1 \leqq j \leqq k)$ and $\sum_{j=1}^{k} \alpha_j x_j = 0$.

- Note that if $\alpha_{j_0} \neq 0$ then $x_{j_0} = \sum_{j \neq j_0} \frac{-\alpha_j}{\alpha_{j_0}} \cdot x_j$ is a lin.comb. of the other x_j.

f. $A \subseteq V$, $A \neq \emptyset$ is linearly independent (LI) if A is not LD: Let $k \in \mathbb{N}$, distinct $x_1, x_2, \ldots x_k \in A$ and $\alpha_1, \alpha_2, \ldots \alpha_k \in \mathbb{R}$. If $\sum_{j=1}^{k} \alpha_j x_j = 0$ then $\alpha_j = 0$ for all $1 \leqq j \leqq k$.

- Let $A \subseteq V$ be LI and also $span(A) \neq V$ and $y \in span(A)^{\complement}$. Then $A \cup \{y\}$ is LI.

g. $B \subseteq V$, $B \neq \emptyset$ is a basis for V if **a.** B is LI and **b.** $span(B) = V$

0.0.6 Examples of vector spaces

a. $\mathbb{R}$ is a VS (scalar product = ordinary product);

b. $\mathbb{R}^n$ is a VS: for $\vec{x} = (x_1, .., x_n)^T = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ (the transpose of $(x_1, .., x_n)$), $\vec{y} = (y_1, .., y_n)^T$ and $\alpha \in \mathbb{R}$ define $\vec{z} = \vec{z} + \vec{z}$ and $\vec{w} = \alpha\vec{x}$ as $z_j = x_j + y_j$, $w_j = \alpha x_j$

- $\mathbb{R}^n$ has Basis $\vec{e}_1 = (1, 0, 0, .., 0)^T$, $\vec{e}_2 = (0, 1, 0, .., 0)^T$, $\ldots$ $\vec{e}_n = (0, 0, .., 0, 1)^T$:
  $\vec{x} = (x_1, .., x_n)^T = \sum_{j=1}^{n} x_j \cdot \vec{e}_j$

c. For any set $X \neq \emptyset$: $\mathscr{F}(X) := \{$ all functions $f : X \to \mathbb{R}\}$; for $f, g \in \mathscr{F}(X), \alpha \in \mathbb{R}$:
define sum $f + g$, scalar product $\alpha \cdot f$ as $(f + g)(x) := f(x) + g(x)$ and $(\alpha f)(x) := \alpha f(x)$
$\mathscr{B}(X) := \{ f \in \mathscr{F}(X) : f \text{ is bounded } \}$, $\mathscr{B}(X)$ is a subspace of $\mathscr{F}(X)$
(f bounded means: there is some $\alpha \in \mathbb{R}$ such that $|f(x)| \leqq \alpha$ for all $x \in X$)

- **What is a basis for $\mathscr{F}(X)$? for $\mathscr{B}(X)$?**

0.0.7 Partially ordered sets (PO sets)

a. Equivalence relation $x \sim y$ on a set X:
a. reflexive: $x \sim x$; **b.** symmetric: $x \sim y \Rightarrow y \sim x$; **c.** transitive: $x \sim y$ and $y \sim z \Rightarrow x \sim z$;

–

– Example 2: Function $f : A \to B$; define $x \sim y$ on A: $x \sim y \Leftrightarrow f(x) = f(y)$

b. Partial ordering (PO) $x \preceq y$ on a set X ("x before y" or "y after x"): **a.** reflexive: $x \preceq x$;
b. antisymmetric: $x \preceq y$ and $y \preceq x \Rightarrow x = y$; **c.** transitive: $x \preceq y$ and $y \preceq z \Rightarrow x \preceq z$;

- "PO set" $(X, \preceq)$, $A \subseteq X$. $a \preceq b$ for $a, b \in A$ makes $(A, \preceq)$ a "PO subset" of $(X, \preceq)$.
- Example 1: $X \subseteq \mathbb{R} : x \preceq y \Leftrightarrow x \leqq y$

– Example 2: $X \subseteq \mathbb{R} : x \preceq y \Leftrightarrow x \geqq y$ (!!)

– Example 3: $X \subseteq 2^{\Omega}$: $A \preceq B \Leftrightarrow A \subseteq B \subseteq \Omega$.

– Example 4: $X, Y \neq \emptyset$; $\mathscr{X} \subseteq \{ (A, f) : A \subseteq X$ and f is a function $A \xrightarrow{f} Y\}$. For (A, f), $(B, g) \in \mathscr{X}$
define $(A, f) \preceq (B, g) \Leftrightarrow$ **a.** $A \subseteq B$; **b.** $f = g\big|_A$ (g extends f from A to B)

c. A PO "$\preceq$" on X is a total (linear) order on X if for any $x, y \in X$ $x \preceq y$ or $y \preceq x$ (or both):
Any two items can be compared.

- PO set $(X, \preceq)$; $C \subseteq X$ is a chain if $c \preceq d$ is a linear order on C

– Example 5: Any subset X of $(\mathbb{R}, \leqq)$ is a chain
– Example 6: Given is $(\mathscr{X}, \preceq)$ from example 4. Let $\mathscr{C}$ be an indexed collection of pairs $\big((C_i, f_i)\big)_{i \in I}$ such that $C_i \subseteq X$ and $f_i : C_i \to Y$. Assume there is an index i_0 such that $C_{i_0} \subseteq C_i$ for all i.
Then $\mathscr{C}$ is a chain $\Leftrightarrow$ for any two $i, j \in I$ **a.** $C_i \subseteq C_j$ or $C_j \subseteq C_i$
b. There is a unique extension of f_{i_0} to any of the supersets $C_i \in \mathscr{C}$.

d. PO set $(X, \preceq)$; $m, m' \in X$; m is maximal in X if it does not have a successor: If $x \in X$ such that $m \preceq x$ then $m = x$. m' is the maximum of X if $m' \geqq x$ for all $x \in X$.
Maxima are unique. Write $m' = \max(X)$. $\max(X)$ is maximal in X.

– Example 7: If $(X, \preceq)$ is totally ordered then $x \in X$ is maximal $\Leftrightarrow$ $x = \max(X)$.
But $\max(X)$ may not exist: $([\,0, 1[, \leqq)$ does not have a max even though it is linearly ordered.
– Example 8: For any X let $x \preceq y \Leftrightarrow x = y$. Then each x is maximal but X has no max unless it only has one element.
– Example 9: Let $\mathscr{X} := \{[a, b] \in \mathbb{R} : \; b - a \leq 1\}$. Define $[a, b] \preceq [a', b'] \Leftrightarrow [a, b] \subseteq [a', b']$. Then any interval of length 1 is maximal. $\max(\mathscr{X})$ DNE.
– Example 10: Given is $(\mathscr{X}, \preceq)$ from examples 4 and 6. (M, f) is maximal in $\mathscr{X} \Leftrightarrow f$ cannot be extended to a function g on a larger set B such that $(B, g) \in \mathscr{X}$.

0.0.8 Zorn's Lemma

a. The **ZL** property of a PO set $(X, \preceq)$:

Every chain $C \subseteq X$, possesses an upper bound $u \in X$, i.e., $x \preceq u$ for all $x \in C$. **(ZL)**

- Zorn's Lemma: If a PO set $(X, \preceq)$ is **ZL** then it possesses a maximal element.

b. Zorn's Lemma is equivalent to the Axiom of Choice: Let $X \neq \emptyset$. Then there is a "choice function" $\psi : 2^X \setminus \emptyset \to X$ such that $\psi(A) \in A$ for each $A \in 2^X \setminus \emptyset$:
In other words, it is possible for an arbitrary nonempty set X to specify a mechanism (the choice function) that allows one to choose some $a \in A$ from any non-empty $A \subseteq X$.

c. Accepting (rejecting) Zorn's lemma as a mathematical tool is equivalent to accepting (rejecting) the Axiom of Choice.

0.0.9 Every vector space has a basis

a. VS (vector space V, $A \subseteq V$ such that A is LI (lin. independent); $\mathfrak{B} := \{B \subseteq V : B \supseteq A$ and B is LI $\}$.
Then the PO set $(\mathfrak{B}, \subseteq)$ is **ZL**.

b. V has a basis which contains the set A.

– Proof: Zorn's Lemma $\Rightarrow$ $\mathfrak{B}$ possesses a maximal element $B^\star$ which is LI because $B^\star \in \mathfrak{B}$.
Must show that $\mathit{span}(B^\star) = V$. But otherwise there is $y \in \mathit{span}(B^\star)^c$. From Ch.0.0.5: $B' := B^\star \cup \{y\}$ is LI, hence $B' \in \mathfrak{B}$. But $B^\star \subseteq B'$ together with $B^\star \neq B'$ contradicts maximality of $B^\star$. $\blacksquare$