A note on a proof of Zorn's Lemma from Axiom of Choice without transfinite induction

Koji Nuida

November 13, 2011 (1st ed.), March 30, 2023 (3rd ed.)

Abstract

In this note, we give a proof of Zorn's Lemma from Axiom of Choice without transfinite induction (whose essential idea is the same as the proof of Theorem 4.19 in the book [H. Rubin and J. E. Rubin, "Equivalents of the Axiom of Choice, II", Second Edition, Studies in Logic and the Foundations of Mathematics vol.116, North-Holland, 1985], as I noticed after writing the first version of this note ...). For the sake of comparison, we also include as an appendix a proof of the claim using transfinite induction.

Throughout this note, let $(X, \leq)$ denote an arbitrary non-empty partially ordered set in which every chain[1] C has an upper bound[2]. Then Zorn's Lemma states that such an X always has a maximal element[3]. In this note, we give a proof of Zorn's Lemma from Axiom of Choice (in Zermelo–Fraenkel set theory), without transfinite induction which would be used in a "natural" proof of the claim.

Assume, for the contrary, that some X satisfying the hypothesis has no maximal elements. We will derive a contradiction from this assumption.

We prepare some definitions and terminology. For any chain C in X, let U_C denote the set of upper bounds for C belonging to $X \setminus C$. For any chain C in X and any $x \in X$, we define $s_C(x) := \{y \in C \mid y < x\}$. Let $\mathcal{C}$ denote the set of non-empty and well-ordered[4] chains of X. We note that when $C \in \mathcal{C}$, any non-empty subset of C also belongs to $\mathcal{C}$. Now we have the following property.

Lemma 1. *We have $U_C \neq \emptyset$ for any $C \in \mathcal{C}$.*

Proof. By the hypothesis of Zorn's Lemma, C has an upper bound $x \in X$. As we have assumed that X has no maximal elements, there is an $y \in X$ with $x < y$. For this y, we have $y \not\leq x$, therefore $y \notin C$ by the choice of x. Moreover, y (as well as x) is an upper bound for C. Hence we have $y \in U_C$, therefore $U_C \neq \emptyset$. □

We define $\mathcal{U} := \{S \subseteq X \mid S = U_C \text{ for some } C \in \mathcal{C}\}$. By Lemma 1, the family $\mathcal{U}$ consists of non-empty sets, therefore Axiom of Choice yields its choice function f. That is, $f(U_C) \in U_C$ for any $C \in \mathcal{C}$. We write the element $f(U_C)$ simply as $f(C)$.

Note that $X \neq \emptyset$ by the hypothesis of Zorn's Lemma. Fix an element $x_0 \in X$. We introduce the following definition:

- We say that a chain $C \in \mathcal{C}$ is *f-consecutive* if $x_0 = \min C$ and for any $c \in C \setminus \{x_0\}$ we have $f(s_C(c)) = c$ (note that now $x_0 < c$ and hence $s_C(c) \neq \emptyset$, therefore $s_C(c) \in \mathcal{C}$). Let $\mathcal{C}_f$ denote the set of f-consecutive elements of $\mathcal{C}$.

We note that, in the definition above, if $c = x_0$, then we have $s_C(c) = \emptyset$. By defining $f(\emptyset) := x_0$, it will hold that for any $C \in \mathcal{C}_f$ and any $c \in C$, we always have $f(s_C(c)) = c$.

Define $C^* := \bigcup_{C \in \mathcal{C}_f} C$. As $\{x_0\} \in \mathcal{C}_f$, it follows that $C^* \neq \emptyset$ and $x_0 = \min C^*$.

[1] totally (or linearly) ordered subset

[2] that is, an element $x \in X$ satisfying that $c \leq x$ for every $c \in C$

[3] that is, an element $x \in X$ satisfying that there exist no elements $y \in X$ with $x < y$

[4] that is, any non-empty subset has the minimum element

Lemma 2. *If $C_1, C_2 \in \mathcal{C}_f$, $x \in C_1$, and $y = \min(C_2 \setminus s_{C_1}(x))$, then $y \in C_1$.*

Proof. By the minimality of y, we have

$$s_{C_2}(y) \subseteq s_{C_1}(x) \ , \tag{1}$$

therefore $x \in C_1 \setminus s_{C_2}(y)$. As C_1 is well-ordered, the element $z := \min(C_1 \setminus s_{C_2}(y))$ exists and satisfies that $z \leq x$. The definition of y implies that $y \notin s_{C_1}(x)$, therefore $y \in C_2 \setminus s_{C_1}(z)$. As C_2 is well-ordered, the element $w := \min(C_2 \setminus s_{C_1}(z))$ exists and satisfies that

$$w \leq y \ . \tag{2}$$

The minimality of w implies that $s_{C_2}(w) \subseteq s_{C_1}(z)$. Conversely, when $u \in s_{C_1}(z)$, the minimality of z implies that

$$u \in s_{C_2}(y) \ . \tag{3}$$

Now if we assume that $u \notin s_{C_2}(w)$, then as C_2 is a chain, we have

$$w \leq u \ , \tag{4}$$

therefore we have $w \in s_{C_2}(y)$ by Eq.(3). By this and Eq.(1), we have $w \in C_1$, while the definition of w implies that $w \notin s_{C_1}(z)$. As C_1 is a chain, it follows that $z \leq w$, therefore we have $z \leq u$ by Eq.(4); while $u \in s_{C_1}(z)$ by the choice of u. This is a contradiction. Hence we have $u \in s_{C_2}(w)$. As a result, we have $s_{C_2}(w) = s_{C_1}(z)$. As both C_1 and C_2 are f-consecutive, we have

$$w = f(s_{C_2}(w)) = f(s_{C_1}(z)) = z \in C_1 \cap C_2 \ .$$

Now the definition of z implies that $z \notin s_{C_2}(y)$. As C_2 is a chain, we have $y \leq z = w$. Combining this and Eq.(2), it follows that $y = w \in C_1$. Hence the claim holds. □

Lemma 3. *If $C_1, C_2 \in \mathcal{C}_f$, then $C_2 \setminus C_1 \subseteq U_{C_1}$. Hence any element of C_1 and any element of C_2 are comparable.*

Proof. Let $x_2 \in C_2 \setminus C_1$ and $x_1 \in C_1$. Then we have $x_2 \in C_2 \setminus s_{C_1}(x_1)$. As C_2 is well-ordered, the element $y := \min(C_2 \setminus s_{C_1}(x_1))$ exists and satisfies that $y \leq x_2$. Now Lemma 2 implies that $y \in C_1$. On the other hand, by the definition of y, we have $y \notin s_{C_1}(x_1)$. As C_1 is a chain, it follows that $x_1 \leq y \leq x_2$. Hence the claim holds. □

By Lemma 3, C^* is a chain.

Lemma 4. *If $C \in \mathcal{C}_f$ and $x \in C$, then $s_{C^*}(x) = s_C(x)$.*

Proof. The definition of C^* implies that $C \subseteq C^*$; therefore it suffices to prove that $s_{C^*}(x) \subseteq C$. For this goal, it suffices to deduce a contradiction by assuming that $y \in s_{C^*}(x) \setminus C$. By the definition of C^*, we have $y \in C'$ for some $C' \in \mathcal{C}_f$. Now Lemma 3 implies that $y \in U_C$, therefore we have $x \leq y$, contradicting the property $y \in s_{C^*}(x)$ in the assumption. Hence the claim holds. □

To prove that C^* is well-ordered, we let S be a non-empty subset of C^* and prove that S has the minimum element. Fix an $x \in S$. The claim already holds when $x = \min S$; we consider the other case from now. As C^* is a chain, we have $y < x$ for some $y \in S$. Hence $s_{C^*}(x) \cap S \neq \emptyset$. By Lemma 4, $s_{C^*}(x) \cap S$ is a non-empty subset of some $C \in \mathcal{C}_f$; as C is well-ordered, $s_{C^*}(x) \cap S$ has the minimum element, say y. Now for any $z \in S$, if $z < x$, then we have $z \in s_{C^*}(x) \cap S$ and therefore $y \leq z$ by the choice of y. On the other hand, if $x \leq z$, then the choice of y implies that $y < x$, therefore $y < z$. Hence we have $y \leq z$ in any case, therefore y is the minimum element of S. Hence C^* is well-ordered, therefore $C^* \in \mathcal{C}$. Moreover, for any $x \in C^* \setminus \{x_0\}$, we take a chain $C \in \mathcal{C}_f$ with $x \in C$; then Lemma 4 implies that $s_{C^*}(x) = s_C(x)$. As C is f-consecutive, we have $f(s_{C^*}(x)) = f(s_C(x)) = x$. Hence C^* is f-consecutive as well. Summarizing, we have $C^* \in \mathcal{C}_f$. Now as $f(C^*) \in U_{C^*}$, the set $C^{**} := C^* \cup \{f(C^*)\}$ is also an element of $\mathcal{C}_f$, while $f(C^*) \notin C^*$ implies that $C^{**} \nsubseteq C^*$. This contradicts the definition of C^*. This completes the proof of Zorn's Lemma.

Appendix: A proof using transfinite induction

In this appendix, for the sake of comparison, we describe a proof of Zorn's Lemma from Axiom of Choice using transfinite induction. First we clarify the statement of the principle for "definition by transfinite recursion" (see e.g., Theorem 9.3 in Chapter I of [K. Kunen, "SET THEORY, An Introduction to Independence Proofs", Elsevier, 1980]):

Theorem 1. *Let $\varphi(x, y)$ be a formula (in Zermelo–Fraenkel set theory) with free variables x, y satisfying $\forall x \exists! y \varphi(x, y)$. Then there exists a formula $\Phi(x, y)$ with free variables x, y satisfying the following two conditions;*

1. $\forall x((x \in \mathbf{ON} \to \exists! y \Phi(x, y)) \wedge (\neg x \in \mathbf{ON} \to \neg \exists y \varphi(x, y)))$;
2. $\forall x(x \in \mathbf{ON} \to \forall y, z(y = \Phi \restriction_x \wedge \varphi(y, z) \to \Phi(x, z)))$,

where "$x \in \mathbf{ON}$" is an abbreviation of "x is an ordinal number" and "$\Phi \restriction_x$" is an abbreviation of the set $\{\langle a, b\rangle \mid a \in x \wedge \Phi(a, b)\}$ (with $\langle a, b\rangle$ denoting the ordered pair of a and b).

Intuitively, the theorem means that, if we would like to define a "function" Φ with domain consisting of all ordinal numbers (the whole of which is never a set) in such a way that the value of Φ at each ordinal number α is determined by a given rule from the values of Φ at ordinal numbers less than α, then there indeed exists such a "function" Φ. Note that this is a theorem of ZF set theory and does not depend on Axiom of Choice.

Now we give a proof of Zorn's Lemma from Axiom of Choice using Theorem 1 (as well as transfinite induction). Let $X \neq 0$ be a partially ordered set appeared in the statement of Zorn's Lemma. Assume, for the contrary, that X has no maximal elements. Then, for each non-empty subset C of X which is isomorphic to an ordinal number (hence a chain), it follows from Axiom of Choice that there exists a distinguished upper bound b_C of C with $b_C \in X \setminus C$.

To apply Theorem 1, first we define a formula $\varphi(x, y)$ in the following manner, where we fix an element $a \in X$ throughout the proof:

- If $x = 0$ $(= \emptyset)$, then let $\varphi(x, y)$ mean that $y = a$.
- If x is a function from an ordinal number $\alpha > 0$ to X which is an isomorphism (between partially ordered sets) onto the image $\mathrm{Im}(x)$ of x, then let $\varphi(x, y)$ mean that $y = b_{\mathrm{Im}(x)}$ (note that $\mathrm{Im}(x)$ is isomorphic to the non-empty ordinal number α, therefore $b_{\mathrm{Im}(x)}$ is indeed defined).
- Otherwise, let $\varphi(x, y)$ mean that $y = 0$.

This formula $\varphi(x, y)$ satisfies the hypothesis of Theorem 1, therefore a formula $\Phi(x, y)$ as in the theorem exists. Now we have the following lemma:

Lemma 5. *Let x be an ordinal number, and let x' be the unique element satisfying $\Phi(x, x')$.*

1. *We have $x' \in X$.*
2. *If $y < x$ and $\Phi(y, y')$, then $y' < x'$ in X.*

Proof. We prove the claim by transfinite induction on x. First, if $x = 0$, then it follows from the definition of the formula φ that $x' = a$, therefore the specified conditions are satisfied. Secondly, suppose that $x > 0$. Then, by the hypothesis of the transfinite induction, the set $\Phi \restriction_x$ in the statement of Theorem 1 is an isomorphism from x to a subset of X, say, C (note that x is totally ordered). Now by the definitions of Φ and φ, it follows that $x' = b_C$, therefore the specified conditions are satisfied for x (the second condition follows from the property that $b_C \in X \setminus C$ is an upper bound of C). Hence the claim holds. $\square$

By the second property shown in Lemma 5, for each element $v \in X$, there exists at most one ordinal number x satisfying $\Phi(x, v)$. Let X' denote the subset of X defined in such a way that $v \in X'$ if and only if $v \in X$ and $\Phi(x, v)$ for some (or equivalently, a unique) ordinal number x. By the Axiom Schema of Replacement applied to the set X' and the formula $\Phi'(x, y) := \Phi(y, x)$, there exists a set Y for which we have $y \in Y$ if y is an ordinal number and the unique element y' satisfying $\Phi(y, y')$ belongs to X'. Now by the first property shown in Lemma 5, the set Y contains every ordinal number. However, this contradicts Burali–Forti Paradox (which states that there exist no sets containing all ordinal numbers). Hence X should have a maximal element, concluding the proof of Zorn's Lemma.