

Research Article

A simple and elementary proof of Zorn's Lemma

Koji Nuida*

Institute of Mathematics for Industry, Kyushu University, Fukuoka, Japan

(Received: 28 August 2023. Received in revised form: 24 April 2024. Accepted: 10 May 2024. Published online: 11 July 2024.)

© 2024 the author. This is an open-access article under the CC BY (International 4.0) license (www.creativecommons.org/licenses/by/4.0/).

Abstract

Zorn's Lemma is a well-known equivalent of the Axiom of Choice. It is usually regarded as a topic in axiomatic set theory, and its historically standard proof (from the Axiom of Choice) relies on transfinite recursion, a non-elementary set-theoretic concept. However, the statement of Zorn's Lemma itself uses only elementary terminology of partially ordered sets. Hence, it is worthy to establish a proof using only such elementary terminology. Following this line of study, a new simple proof of Zorn's Lemma is given that does not even use the notion of a well-ordered set.

Keywords: partially ordered sets; Zorn's Lemma; elementary proof.

2020 Mathematics Subject Classification: 06A06, 03E25, 97E60.

1. Introduction

Zorn's Lemma is one of the most famous equivalents of the Axiom of Choice under the Zermelo–Fraenkel set theory. It is historically regarded as a topic of axiomatic set theory, and the “standard” proof of Zorn's Lemma (from the Axiom of Choice) relies on a non-elementary set-theoretic concept, called the transfinite recursion (for example, see [8, Theorem 6.1], [13, Theorem 2.2], [16, Theorem 5.4], [15, Theorem 2.1]). However, Zorn's Lemma is ubiquitously used in mathematics, not just in set theory, and Zorn's Lemma can be stated with only elementary terminology of partially ordered sets (posets). Therefore, it is worthy to establish a proof of Zorn's Lemma by using such elementary terminology only. There exists a large number of proofs, without transfinite recursion, of either Zorn's Lemma (for example, see [1, 3, 7, 9, 11, 12, 14, 17, 20, 22–24, 28–30]) or some property that easily implies Zorn's Lemma; for instance, Hausdorff Maximal Principle [6, 10, 18, 21, 25, 27] and the existence of fixed points for some mappings on a poset [2, 4, 5, 19, 26, 31]. Following this direction of study, in this article, yet another simple proof of Zorn's Lemma is given using elementary terminology only.

The proof of Zorn's Lemma, given in the present article, is a descendant of a proof due to Lewin [22]. Intuitively speaking, Lewin's proof constructed a “highest” chain in a given poset as the union of some family \mathcal{C} of chains, which was defined by using the notion of a well-ordered set. The nontrivial point of the present work is to extract what is really required in the proof of the properties of well-ordered sets. From this point of view, in the present work, two properties named “(i-C)” and “(ii-C)” are introduced; a trick in the present proof is that the construction of the family \mathcal{C} is now in two steps, where the first step uses the condition (i-C) to construct an auxiliary family \mathcal{C}_0 and the second step constructs \mathcal{C} by using the condition (ii-C) defined in terms of the family \mathcal{C}_0 . As a result, the terminology in the present proof is more elementary, not even using the notion of a well-ordered set.

Notations, terminology, and the statement of Zorn's Lemma

A poset (partially ordered set) is denoted by (P, \leq) , where P is a set and “ \leq ” is a binary relation on P with the following three axioms (for arbitrary $x, y, z \in P$): (I) $x \leq x$; (II) $x \leq y$ and $y \leq x$ imply $x = y$; (III) $x \leq y$ and $y \leq z$ imply $x \leq z$. We write $x < y$ to mean “ $x \leq y$ and $x \neq y$ ”. A subset C of P is called a *chain* in P if any pair (x, y) of elements of C satisfies either $x \leq y$ or $y \leq x$. Note that any subset of a chain in P is also a chain in P . We say that $x \in P$ is an *upper bound* of a chain $C \subseteq P$ if $y \leq x$ holds for any $y \in C$; and $x \in P$ is a *strict upper bound* of C if $y < x$ holds for any $y \in C$, or equivalently, x is an upper bound of C and $x \notin C$. A poset (P, \leq) is said to be *inductively ordered* if any chain in P has an upper bound in P . We say that an element $x \in P$ is *maximal* if there exists no $y \in P$ satisfying $x < y$. On the other hand, we say that an element x of a chain C in P is the *maximum element* of C , denoted by $\max C$, if $y \leq x$ holds for any $y \in C$ (i.e., x is an upper bound of C). With the above terminology, Zorn's Lemma is stated as follows.

*E-mail address: nuida@imi.kyushu-u.ac.jp

Theorem 1.1 (Zorn’s Lemma). *Any inductively ordered poset (P, \leq) has a maximal element.*

2. Proof of Theorem 1.1

Let $\mathcal{T} := \{C \subseteq P : C \text{ is a chain in } P\}$. For each $C \in \mathcal{T}$, define

$$\overline{U}_C := \{x \in P : x \text{ is an upper bound of } C\}, U_C := \{x \in P : x \text{ is a strict upper bound of } C\} = \overline{U}_C \setminus C.$$

Then $\overline{U}_C = U_C \cup \{\max C\}$ if the maximum element $\max C$ of C exists, and $\overline{U}_C = U_C$ otherwise. Therefore,

$$\text{if } x \in \overline{U}_C, y \in P, \text{ and } x < y, \text{ then } y \in U_C \tag{1}$$

(note that if $\max C < y$ then $y \notin C$). Moreover,

$$\text{if } C_1, C_2 \in \mathcal{T} \text{ and } U_{C_1} \not\subseteq U_{C_2}, \text{ then } C_1 \cap \overline{U}_{C_2} = \emptyset \tag{2}$$

(indeed, if $x \in C_1 \cap \overline{U}_{C_2}$, then any $y \in U_{C_1}$ satisfies that $x < y$ by definition of U_{C_1} and hence $y \in U_{C_2}$ by (1), contradicting the assumption $U_{C_1} \not\subseteq U_{C_2}$). Now, the Axiom of Choice yields a choice function f_0 for the family of non-empty subsets X of P ; that is, $f_0(X) \in X$ for any such X . Then we define a function $f : \mathcal{T} \rightarrow P$ satisfying that, for each $C \in \mathcal{T}$,

$$f(C) := f_0(U_C) \in U_C \subseteq P \setminus C \text{ if } U_C \neq \emptyset, f(C) := \max C \in C \cap \overline{U}_C \text{ if } U_C = \emptyset$$

(note that when $U_C = \emptyset$, the assumption on (P, \leq) being inductively ordered implies that C has an upper bound $x \in P$ that is not strict, which satisfies $x \in C$ and must be the maximum element of C). By the construction of f ,

$$\text{if } C_1, C_2 \in \mathcal{T} \text{ and } U_{C_1} = U_{C_2} \neq \emptyset, \text{ then } f(C_1) = f(C_2) \in U_{C_1} (= U_{C_2}). \tag{3}$$

Now, let \mathcal{C}_0 denote the set of all $C \in \mathcal{T}$ satisfying the following condition:

$$\text{(i-}C\text{)} \ S \subseteq C \text{ and } U_S \not\subseteq U_C \text{ imply } f(S) \in C.$$

Then let \mathcal{C} denote the set of all $C \in \mathcal{C}_0$ satisfying the following condition:

$$\text{(ii-}C\text{)} \ C' \in \mathcal{C}_0 \text{ implies } C \subseteq C' \cup U_{C'}.$$

Let $C^* := \bigcup \mathcal{C}$. We prove that $C^* \in \mathcal{C}$, by verifying the defining conditions of \mathcal{C} as follows:

- For (ii- C^*), let $C' \in \mathcal{C}_0$. Then each $C \in \mathcal{C}$ satisfies $C \subseteq C' \cup U_{C'}$ by (ii- C), therefore $C^* = \bigcup \mathcal{C} \subseteq C' \cup U_{C'}$, as desired.
- We show that $C^* \in \mathcal{T}$, i.e., any $x, y \in C^*$ satisfy $x \leq y$ or $y \leq x$. We can take $C, C' \in \mathcal{C}$ with $x \in C$ and $y \in C'$. Now, $C \subseteq C' \cup U_{C'}$ by (ii- C), therefore we have either $x, y \in C'$ or $x \in U_{C'}$ (hence $y < x$), implying the claim in any case.
- For (i- C^*), suppose that $S \subseteq C^*$ and $U_S \not\subseteq U_{C^*} = \bigcap_{C \in \mathcal{C}} U_C$. Then $U_S \not\subseteq U_C$ for some $C \in \mathcal{C}$. Now, $S \cap U_C = \emptyset$ by (2), while $S \subseteq C^* \subseteq C \cup U_C$ by (ii- C^*) applied to $C \in \mathcal{C}_0$, therefore $S \subseteq C$. Also, by applying (i- C) to $S \subseteq C$, we have $f(S) \in C \subseteq C^*$, therefore $f(S) \in C^*$, as desired.

Now, if $U_{C^*} = \emptyset$, then $f(C^*) = \max C^*$ is a maximal element of P (as otherwise $\max C^* < y$ for some $y \in P$ and hence $y \in U_{C^*}$ by (1), a contradiction), as desired. Hence, the proof will be completed once we obtain a contradiction assuming that $U_{C^*} \neq \emptyset$ (hence $f(C^*) \in U_{C^*}$). Let $u := f(C^*) \in U_{C^*}$ and $C^{**} := C^* \cup \{u\}$; therefore we have $u = \max C^{**} \notin C^*$, $C^{**} \not\subseteq C^*$, and $C^{**} \in \mathcal{T}$ as $C^* \in \mathcal{T}$. We prove that $C^{**} \in \mathcal{C}$, by verifying the remaining defining conditions as follows:

- For (i- C^{**}), suppose that $S \subseteq C^{**}$ and $U_S \not\subseteq U_{C^{**}}$. Then $S \cap \overline{U}_{C^{**}} = \emptyset$ by (2) and hence $u \notin S$. This implies that $S \subseteq C^*$ and hence $U_{C^*} \subseteq U_S$. Now, if $U_S \subseteq U_{C^*}$, then $U_S = U_{C^*} \neq \emptyset$ and $f(S) = f(C^*) = u \in C^{**}$ by (3); while if $U_S \not\subseteq U_{C^*}$, then we have $f(S) \in C^* \subseteq C^{**}$ by applying (i- C^*) to $S \subseteq C^*$. Hence $f(S) \in C^{**}$ in any case, as desired.
- For (ii- C^{**}), let $C' \in \mathcal{C}_0$. As $C^* \subseteq C' \cup U_{C'}$ by (ii- C^*), it suffices to show that $u \in C' \cup U_{C'}$, or equivalently, $u \in C'$ if $u \notin U_{C'}$. Now, we have $U_{C^*} \not\subseteq U_{C'}$ as $u \in U_{C^*}$. Hence $C^* \cap U_{C'} = \emptyset$ by (2), while $C^* \subseteq C' \cup U_{C'}$ as above, therefore $C^* \subseteq C'$. Hence $u = f(C^*) \in C'$ by applying (i- C') to $C^* \subseteq C'$, as desired.

However, this fact $C^{**} \in \mathcal{C}$ yields a contradiction, as $C^{**} \not\subseteq C^* = \bigcup \mathcal{C}$. This completes the proof of Theorem 1.1.

3. Remarks

If we change the family \mathcal{T} in our proof to the family of all well-ordered subsets of P , then it yields a proof with weakened assumption on (P, \leq) being inductively ordered where the existence of upper bounds is now assured only for well-ordered subsets of P . Indeed, now to show that the set C^* in the proof is a member of \mathcal{T} , for any non-empty subset $S \subseteq C^*$, fix any $C \in \mathcal{T}$ with $S \cap C \neq \emptyset$, and take $x := \min(S \cap C)$. For $y \in S$, if $y \in C$ then $x = \min(S \cap C) \leq y$; while if $y \notin C$ then $y \in C^* \setminus C \subseteq U_C$ from (ii- C^*) and $x < y$. Hence $x = \min S$, therefore $C^* \in \mathcal{T}$. The proof of $C^{**} \in \mathcal{T}$ is almost the same as the original proof, and the remaining part of the proof is not affected by the change of the definition of \mathcal{T} .

We now explain the difference between our proof and that of Lewin [22], mentioned in the Introduction. The outline of the proof is common to both proofs, i.e., (1) defining some family \mathcal{C} of chains and (2) showing that $C^* := \bigcup \mathcal{C} \in \mathcal{C}$ and that $C^{**} := C^* \cup \{f(C^*)\} \in \mathcal{C}$ where $f(C^*)$ is a strict upper bound of C^* , yielding a contradiction. However, in contrast to our proof where the key property $C^* \in \mathcal{C}$ in Step (2) is derived directly from the defining conditions for \mathcal{C} , Lewin's proof required an intermediate step to show some extra property for \mathcal{C} that is seemingly stronger than the defining conditions for \mathcal{C} . In detail, the following comparability property was shown in Lewin's proof: for any two members of \mathcal{C} , one of them is an initial segment of the other. This is analogous to a property of well-ordered sets, and to ensure this property, \mathcal{C} was defined in a way that each member of \mathcal{C} should be a well-ordered subset of P . Our main idea is that the full comparability property is in fact not necessary in the proof. Our new condition (ii- C) can be seen as a weaker variant of the comparability property (indeed, the condition $C \subseteq C' \cup U_{C'}$ holds when one of C and C' is an initial segment of the other), which (for $C^* = \bigcup \mathcal{C}$ and $C^{**} = C^* \cup \{f(C^*)\}$) can be proved directly without requiring that members of \mathcal{C} are well-ordered subsets.

Acknowledgment

The author deeply thanks Árpád Száz for several detailed comments to a previous version of the manuscript, including a comprehensive list of known proofs of Zorn's Lemma without transfinite recursion (on which the reference list in the current paper is based).

References

- [1] J. An, *A Proof of Zorn's Lemma*, 2019, Available at <https://web.archive.org/web/20240111075954/https://www.math.pku.edu.cn/teachers/anjp/Zorn.pdf>.
- [2] C. Berge, *Topological Spaces*, Oliver and Boyd, Edinburgh, 1963.
- [3] G. Bergman, *The Axiom of Choice, Zorn's Lemma, and all that*, 1997, Available at <https://web.archive.org/web/20240703172930/https://www1.mat.uniroma1.it/people/desole/didattica/2022-algebra1/bergman.pdf>.
- [4] N. Bourbaki, Sur le théorème de Zorn, *Arch. Math.* **2** (1949) 434–437.
- [5] A. Brown, A remark on set theory, *Proc. Amer. Math. Soc.* **25** (1970) 920–921.
- [6] K. Brown, *Zorn's Lemma*, 2010, Available at <https://web.archive.org/web/20240419031710/https://pi.math.cornell.edu/~kbrown/6310/zorn.pdf>.
- [7] B. Bukh, *Math Studies Algebra: Axiom of Choice*, 2021, Available at https://www.borisbukh.org/MathStudiesAlgebra1718/notes_ac.pdf.
- [8] P. J. Cameron, *Sets, Logic and Categories*, Springer, New York, 1998.
- [9] J. Dugundji, *Topology*, Allyn and Bacon, Boston, 1966, 29–36.
- [10] O. Frink, A proof of the maximal chain theorem, *Amer. J. Math.* **74** (1952) 676–678.
- [11] D. R. Grayson, *Zorn's Lemma*, 2007, Available at <https://web.archive.org/web/20240703173322/https://pages.uoregon.edu/apolish/Zorn.pdf>.
- [12] P. R. Halmos, *Naive Set Theory*, Van Nostrand, Princeton, 1960.
- [13] H. Herrlich, *Axiom of Choice*, Springer, New York, 2006.
- [14] A. Jain, Zorn's Lemma: An elementary proof under the Axiom of Choice, *arXiv:1207.6698 [math.LO]*, (2012).
- [15] T. Jech, *The Axiom of Choice*, North-Holland, Amsterdam, 1973.
- [16] T. Jech, *Set Theory*, Springer, New York, 2002.
- [17] H. Kneser, Eine direkte Ableitung des Zornschen Lemmas aus dem Auswahlexiom, *Math. Zeitschrift* **53** (1950) 110–113.
- [18] K. Kuttler, *Modern Analysis*, CRC Press, Boca Raton, 1998, 529–533.
- [19] S. Lang, *Real and Functional Analysis*, Springer, New York, 1993, 10–16.
- [20] S. Lang, *Algebra*, Springer, New York, 2002.
- [21] L. Lektorean, *Zorn's Lemma*, 2010, Available at <https://web.archive.org/web/20190819061255/https://people.math.aau.dk/~cornean/index.html/zornlemma.pdf>.
- [22] J. Lewin, A simple proof of Zorn's Lemma, *Amer. Math. Monthly* **98**(4) (1991) 353–354.
- [23] J. Nagata, *Modern General Topology*, North-Holland, Amsterdam, 1985, 16–23.
- [24] H. Rubin, J. E. Rubin, *Equivalents of the Axiom of Choice, II*, North-Holland, Amsterdam, 1985.
- [25] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, 1987, 395–396.
- [26] J. Shurman, *The Axiom of Choice, Zorn's Lemma, and the Well Ordering Principle*, Available at <https://web.archive.org/web/20240506042828/https://people.reed.edu/~jerry/332/23zorn.pdf>.
- [27] N. Spronk, *Axiom of Choice, Zorn's Lemma and the Well-Ordering Principle*, Available at <https://web.archive.org/web/20240703174716/https://www.math.uwaterloo.ca/~nspronk/math351/AofC.pdf>.
- [28] T. Szele, On Zorn's lemma, *Publ. Math. Debrecen* **1** (1950) 254–256.
- [29] D. V. Tausk, *Proof of Zorn's Lemma*, 2018, Available at <https://web.archive.org/web/20240503001714/https://www.ime.usp.br/~tausk/texts/Zorn.pdf>.
- [30] J. D. Weston, A short proof of Zorn's Lemma, *Arch. Math. (Basel)* **8** (1957) 279–279.
- [31] M. Zarouali-Darkaoui, On the Bourbaki's fixed point theorem and the axiom of choice, *arXiv:1905.09782v2 [math.LO]*, (2019).