Sur le Théorème de Zorn

par Nicolas Bourbaki à Nancago, translated by Paul Taylor

It is well known that it is more convenient to use Zorn's Lemma than Zermelo's [well-ordering] theorem [Zer08] in most arguments that use "transfinite induction". In fact the two results are equivalent, and both are equivalent to the Axiom of Choice.

At the request of many readers of my *Elements of Mathematics* [Bou57], I will show briefly how these two theorems can be derived from the Axiom of Choice and from a general result about ordered sets that itself does not require that Axiom and whose proof essentially copies the one that Zermelo gave for his own theorem.

Theorem 0.1 Let E be an ordered set, a an element of E, f a function from E to E such that, for all $x \in E$, we have $f(x) \geq x$. Let $\mathcal{F}$ be the set of subsets X of E with the following properties:

- $a \in X$;
- if $x \in X$ then $f(x) \in X$;
- if a non-empty subset $Y \subset X$ has a least upper bound in E, this least upper bound belongs to X.

Then the set $\mathcal{F}$ is not empty, the intersection A of the sets $X \in \mathcal{F}$ belongs to $\mathcal{F}$ and for every pair of elements $x, y \in A$, we have

$$y \leq x \quad \vee \quad y \geq f(x),$$

from which it follows that A is totally ordered.

Proof It is immediate that $\mathcal{F}$ is not empty, because the set of elements of E that are are $\geq a$ belongs to $\mathcal{F}$; we also see immediately that $A \in \mathcal{F}$.

Write B for the subset of A formed by those elements $x \in A$ which have the following property:

$$\text{(P):} \qquad \text{if } y \in A \text{ and } y \leq x \text{ then } y = x \text{ or } f(y) \leq x.$$

The subset B is not empty, because clearly $a \in B$.

We will show first that if $x \in B$ and $y \in A$ then the following holds:

$$\text{(Q):} \qquad y \leq x \text{ or } y \geq f(x).$$

Indeed, for any element $x \in B$, let C be the subset of A formed of elements y for which the property (Q) is true; we will show that $C \in \mathcal{F}$; since $C \subset A$, it will follow from the definition of A that $C = A$.

Indeed:

- Since $a \leq x$, we have $a \in C$;
- If $y \in C$ and $y \geq f(x)$, we have $f(y) \geq y \geq f(x)$ and so $f(y) \in C$ by definition. In the other case, $y \leq x$, so from (P) we deduce $y = x$ (and so $f(y) = f(x)$), or that $f(y) \leq x$. In either case we have $f(x) \in C$.
- Let Y be a subset of C having a least upper bound b in E; then $b \in A$ since $A \in \mathcal{F}$. If,

for all $y \in Y$, we have $y \leq x$, then $b \leq x$. If, otherwise, there is at least one $y \in Y$ such that $y \geq f(x)$ then $b \geq f(x)$. So $b \in C$ in both cases.

Hence we have shown that $C \in \mathcal{F}$ and so $C = A$, and it follows that the property (Q) holds for all $x \in B$ and $y \in A$. We will now show that $B = A$, from which the theorem follows: since $B \subset A$ it will be enough to show that $B \in \mathcal{F}$.

Indeed,

- We have already seen that $a \in B$.
- If $x \in B$ and if $y \in A$ is such that $y < f(x)$ then $y \leq x$ by the property (Q). Then by (P) we have $y = x$ or $f(y) \leq x$. In both cases we have $f(y) \leq f(x)$, since $x \leq f(x)$. We see that in both cases $f(x) \in B$.
- Let Y be a subset of B with a least upper bound b in E, which must belong to A. Let $y \in A$ be such that $y < b$; we can't have $y \geq x$, let alone $y \geq f(x) \geq x$ for *all* $x \in Y$, because then we would have $y \geq b$, contrary to the hypothesis. hence by (Q) there is at least one $x \in Y$ such that $y < x$, and it follows from (P) that $f(y) \leq x \leq b$, so we have $b \in B$. This is enough to prove that $B \in \mathcal{F}$, so the theorem has been proved. Observe that this actually shows that a is the *least* element of A. □

Corollary 0.2 If A has a least upper bound b in E then $b \in A$ and $f(b) = b$. Conversely, of there is an element $b \in A$ such that $f(b) = b$, then b is the *largest* element of A.

Proof If b is a least upper bound of A in E then $b \in A$ by the definition of $\mathcal{F}$. Then b is the largest element of A. Since $f(b) \in A$ and $f(b) \in b$, necessarily $f(b) = b$.

Conversely, if $b \in A$ is such that $f(b) = b$, let X be the subset of A consisting of those elements $x \leq b$. We will see that $X \in \mathcal{F}$, from which it will follow that $X = A$ and then that b is the largest element of A.

Indeed:

- Clearly, $a \in X$.
- If $x \in X$ then $x \leq b$. If $x = b$ then $f(x) = f(b) = b$ belongs to A. If otherwise $x < b$ then $f(x) \leq b$ by Theorem 0.1. So in both cases $f(x) \in X$.
- If Y is a subset of A having a least upper bound c in E, we have $c \in A$, whilst $c \leq b$ by definition of X, which completes the proof. □

We see that that if $x \in A$ is not the greatest element of A then $x < f(x)$. Moreover, there is no element $y \in A$ such that $x < y < f(x)$.

Corollary 0.3 The subset A is well ordered.

Proof Let B be a non-empty subset of A and C the set of lower bounds of B in A. To see that B has a smallest element, it is enough to show that C has a least upper bound c in E. Indeed, c then belongs to A and so to C by definition; it comes down to showing that $c \in B$. Otherwise, since B is not empty, c would not be the greatest element of A, so one would have $c < f(c)$, and $x \geq f(x)$ for all $x \in B$, which would contradict the definition of c.

So we show that C has a least upper bound. We will see that otherwise one would have $C = A$, contrary to the hypothesis that B is non-empty. It is enough, as always, to show that, under these assumptions, $C \in \mathcal{F}$.

Indeed,

- We have $a \in C$ by definition.
- If $x \in C$, by hypothesis $x \in B$ is impossible; then for all $y \in B$ we have $x < y$, which

implies $f(x) \le y$ by the Theorem 0.1, and so $f(x) \in C$.

- If Y is a subset of C with a least upper bound in E, that must belong to A, so since all elements of B lie above it, it must be an element of C.

The hypothesis would therefore imply $C = A$, which is impossible, and so the result has been proved. □

It is then easy to prove Zorn's Lemma:

Theorem 0.4 Let E be an inductively ordered set, that is, such that every totally ordered subset of E has a least upper bound in E. Then E has a maximal element.

Proof By the Axiom of Choice, there is a function $f : E \to E$ such that $f(x) = x$ if x is maximal and $f(x) > x$ otherwise. Let A be the subset defined from an element $a \in E$ and the function f according to the method of Theorem 0.1. Since A is totally ordered (by Theorem 0.1), it has a least upper bound b in E by hypothesis. It then follows by Corollary that $b \in A$ and $f(b) = b$, which is to say that b is maximal. □

Recall that if E is a well ordered set then an [initial] *segment* S of E is a subset of R such that if $x \in S$ and $y \le x$ then $y \in S$. It is easily shown that the segments of E are E itself and, for each $a \in E$, the set of $x \in E$ such that $x < a$.

In fact we can weaken the hypothesis of Theorem 0.4 and show

Theorem 0.5 If E is an ordered set such that every well ordered subset of E is bounded in E, then E has a maximal element.

Proof Write $\mathcal{B}$ for the set of subsets B of E that are *well ordered*, and consider the relation on $\mathcal{B}$ that "X is a segment of Y". It is immediate that this is an order relation, that we write $X \le Y$. Moreover, equipped with this relation, $\mathcal{B}$ is *inductive*: this amounts to showing that, if (B_α) is a totally ordered set of elements of $\mathcal{B}$ then the union B of B_α belongs to $\mathcal{B}$, in other words it is well ordered: indeed, if Z is a non-empty subset of B, $Z \cap B_\alpha$ is non-empty for at least one α, so has a smallest element in B_α, and we see immediately that it's also the smallest element of Z in B.

Since the set $\mathcal{B}$ is inductive, it has a *maximal* element X_0. By the hypothesis, X_0 has a bound b in E. Then $b \in X_0$; otherwise, the set $X_1 = X_0 \cup \{b\}$ would be well ordered and X_0 would be a segment of X_1 distinct from it, contrary to the definition of X_0. The element b is therefor the greatest element of X_0 and by a similar argument b is *maxima* in E. □

We are not saying that Zorn's Lemma *must* be deduced directly from Theorem 0.1, since it is in fact very argument of Zermelo that we are repeating. We recall simple that one can apply Theorem 0.1 to the set $\mathcal{P}(E)$ of subsets of any set E, ordered by inclusion. By the Axiom of Choice, for every subset $X \subset E$ such that $X \ne E$, one defines an element $g(X)$ of E such that $g(x) \notin X$. Then for all $X \subset E$, we put $f(X) = X \cup \{g(X)\}$ if $X \ne E$ and $f(E) = E$. It is then enough to apply Theorem 0.1 to the function f and the element a equal to the empty subset of E. Then the set A to one obtains in this way has, it is easy to see, E as its greatest element, and the set of elements of A distinct from E is in bijective correspondence with E, from which we have Zermelo's theorem.

It seems to me that it is more instructive to deduce Zermelo's theorem directly from Zorn's Lemma. Consider the set $\mathcal{S}$ of order structures on the subsets of E (a set that is in bijective correspondence with a certain subset of $\mathcal{P}(E \times E)$), and let $\mathcal{B}$ be the set of those order structures that are *well orderings*. For every structure $s \in \mathcal{B}$, let A be the subset of E

where s is defined. We define an order relation $s \leq s'$ on $\mathcal{B}$ that is equivalent to $A_s \subset A_{s'}$. The structure induced by s' on A_s is identical to s, and A_s is a segment $A_{s'}$ for the structure s'. We easily see, by the same argument as in Theorem 0.1, that $\mathcal{B}$ is not just ordered but *inductive*, so let s_0 be a maximal element of $\mathcal{B}$. It remains to show that the set A_{s_0} is equal to E; indeed, otherwise on the set composed of A_{s_0} and an element that doesn't belong to A_{s_0} we may define a well ordered set s_1 such that $s_1 > s_0$. This proves the result.

Archiv der Mathematik **2** (1949) 434–7
doi 10.1007/bf0203694.
Received on 15 November 1949.
Nancago is a name for Nancy invented by the Bourbaki group: Nancy+Chicago.

References

[Bou49] Nicolas Bourbaki. Sur le théorème de Zorn. *Archiv der Mathematik*, 2:434–7, 1949.

[Bou57] Nicolas Bourbaki. *Eléments de Mathématique XXII: Théories des Ensembles, Livre I, Structures*. Number 1258 in Actualités scientifiques et industrielles. Hermann, 1957. English translation, "Theory of Sets" 1968.

[Cam78] Paul Campbell. The origin of Zorn's lemma. *Historia Mathematica*, 5:77–89, 1978.

[Fel62] Walter Felscher. Doppelte Hülleninduktion und ein Satz von Hessenberg und Bourbaki. *Archiv der Mathematik*, 13:160–5, 1962.

[Lan65] Serge Lang. *Algebra*. Addison-Wesley, 1965. Third edition published by Springer-Verlag in 2002.

[Tay23] Paul Taylor. Well founded coalgebras and recursion. `www.paultaylor.eu/ordinals`, 2023.

[vH67] Jan van Heijenoort. *From Frege to Gödel: a Source Book in Mathematical Logic, 1879–1931*. Harvard University Press, 1967.

[Wil01] Todd Wilson. An intuitionistic version of Zermelo's proof that every choice set can be well-ordered. *Journal of Symbolic Logic*, 66:1121–6, 2001.

[Wit51] Ernst Witt. Beweisstudien zum Satz von M. Zorn. *Mathematische Nachrichten*, 4:434–8, 1951.

[Zer08] Ernst Zermelo. Neuer Beweis für die Möglichkeit einer Wohlordnung. *Mathematisches Annalen*, 65:107–128, 1908. English translation in [vH67], pages 183–198.

[Zor35] Max Zorn. A remark on transfinite algebra. *Bull. Amer. Math. Soc.*, 41(10), 1935.

Note that this list consists of apparently relevant works mainly from later than Bourbaki's paper; very few were actually cited there.