Steinitz's Theorem and Zorn's Lemma

ADACHI Norio
Waseda University

§1. Preliminaries

It is well-known that the following three statements are equivalent:

1 (**Well-Ordering Theorem**). *Every set can be well ordered.*

2 (**Axiom of Choice**, abbreviated AC). *If for each* $i \in I, S_i \neq \emptyset$, *then* $\prod_{i \in I} S_i \neq \emptyset$.

3 (**Maximal Principle** or **Zorn's Lemma**). *Every inductive set*[1] *has a maximal element.*

AC was first formulated by Zermelo in 1904 to prove the Well-Ordering Theorem which had been earlier proposed by Cantor as a kind of law. The Maximal Principle was formulated by Zorn in 1935 to give a simple proof to Steinitz's Theorem on the existence and uniqueness of algebraic closure. A full account of the history of these axioms or principles can be found in the fascinating book of G. Moore ([Moor]).

Note that the Axiom of Choice is not self-evident even in the countable case. It is tempting to consider the following argument as a proof:

Argument 1. The ascending chain condition holds $\Longleftrightarrow$ the maximal condition holds.[2] In fact, $\Longleftarrow$ is clear. To show $\Longrightarrow$, suppose that mc does not hold in some ordered set S. Let A be a non-empty subset of S which has no maximal element, and let a_0 be any element of A. Since a_0 is not maximal, there is an element $a_1 \in A$ such that $a_0 < a_1$. Since a_1 is not maximal, there is an element $a_2 \in A$ such that $a_1 < a_2$. *Inductively*, we find a_{n+1} such that $a_n < a_{n+1}$. In this way we could construct an infinite ascending chain, which implies that acc does not hold. Thus the proof of the implication $\Longrightarrow$ is complete.

Argument 1 is false, though the equivalence 'acc $\Longleftrightarrow$ mc' itself is true. If you believe that Argument 1 is legitimate, you must also accept the following argument:

[1] An ordered set S is said to be inductive if every totally ordered subset of S has an upper bound in S.

[2] In an ordered set S, the condition that no infinite ascending chain exists in S is called the ascending chain condition, abbreviated acc. The condition that any non-empty subset of S has a maximal element is called the maximal condition, abbreviated mc.

Argument 2. Let $S_n (n \in \mathbb{N})$ be a family of non-empty sets. Then $S_1 \neq \emptyset$. Suppose that $\prod_{i=1}^{n} S_i \neq \emptyset$. Then

$$\prod_{i=1}^{n+1} S_i \simeq \prod_{i=1}^{n} S_i \times S_{n+1} \neq \emptyset.$$

Hence, *by mathematical induction*,

$$\prod_{n \in \mathbb{N}} S_n \neq \emptyset.$$

This argument is equivalent to the assertion that the Countable Axiom of Choice (abbreviated AC_0) is unconditionally true. Most mathematicians today would dismiss Argument 2 as fallacious; what is really proved here is only $\prod_{i=1}^{n} S_i \neq \emptyset$ for any n. Therefore Argument 1 must also be considered as faulty. As Aristotle said, *infinity is that which cannot be passed through.*

As is well-known, Argument 1 can be repaired by invoking the power of AC as follows:

Suppose that S does not satisfy mc. Let A be a non-empty subset of S which has no maximal element. For any a in A put $A_a = \{x \ (\in A) \mid a < x\}$. Then $A_a \neq \emptyset$. Let $\varphi : A \longrightarrow A$ be a function satisfying $\varphi(a) \in A_a$, whose existence is guaranteed by AC. Let a_1 be an arbitrary element of A, putting $a_{n+1} = \varphi(a_n)$ for $\forall n \in \mathbb{N}$. Then $(a_n)_n$ is an infinite ascending chain in S.[3]

The result of this inference has been formulated as a principle:

Principle of Dependent Choices (abbreviated PDC). *If R is a relation on a set A such that for every x in A there exists some y in A with xRy, then there is a sequence $a_1, a_2, \ldots$ such that for every positive integer n, a_n is in A and $a_n R a_{n+1}$ holds.*

Argument 1, thus, is valid if one simply replaces 'inductively' by 'because of PDC'.

PDC was proposed by Bernays in 1942 as a weakened form of AC. Incidentally, the following facts are known: AC implies PDC, but not vice versa: And, in turn, PDC implies AC_0, but not vice versa (cf. [Moor] or [Jech]):

$$\mathrm{AC} \underset{\not\Longleftarrow}{\Longrightarrow} \mathrm{PDC} \underset{\not\Longleftarrow}{\Longrightarrow} \mathrm{AC}_0$$

§2. Critical Study of the Proofs of Steinitz's Theorem

[3]Let f be a function from a set X to X itself, and $a \in X$. Then there exists a unique sequence $(a_n)_n$, that is, a function $a : \mathbb{N} \longrightarrow X$, such that $a_0 = a$, $a_{n+1} = f(a_n)$ for $\forall n \in \mathbb{N}$ (Recursion Theorem).

Steinitz's Theorem (1910, [Stein]). *For any field K, an algebraic closure of K exists and is unique up to isomorphism.*

There are several (plausible) proofs of the theorem:

Argument 3. For simplicity, suppose that K is countable. Let $f_1(X), f_2(X), \cdots$ be an enumeration of the monic polynomials of positive degree. Then we begin with $K_0 = K$ and construct K_{n+1} inductively as the splitting field over K_n. The union Ω of all the K_n is a set (since the union of a family of sets is also a set), and one can prove that Ω is an algebraic closure of K.

The procedure just sketched can be used also in the general case by invoking the power of transfinite induction.

A closer examination of Argument 3 reveals a minor flaw. A splitting field over a given field K is not unique, though it is unique up to isomorphism, and, in general, each contains no more than a subfield isomorphic to K. Even if one considers the totality of all isomorphic fields, trying to apply PDC, it cannot be a 'set'.

B. L. van der Waerden got over the difficulty by providing n symbols $\alpha_1, ..., \alpha_n$ as the roots of an irreducible polynomial $f(X)$ of degree n to uniquely construct a splitting field $L(\alpha_1, ..., \alpha_n)$ of $f(X)$ over a given field L. This may explain the somewhat awkward proof of Steinitz's Theorem in any edition of [Waer].

Historical Note. In his most seminal work ([Stein]) Steinitz stressed the importance of AC, proving far-reaching results, among which are Steinitz's Theorem and the existence of a transcendence base for a field extension. This was the first place where many important theorems in practical mathematics were proved by conscious appeal to AC. After his work, AC gradually came to be regarded as an essential tool.

In the first edition of [Waer], van der Waerden simplified Steinitz's proof. However, he was persuaded by intuitionists to abandon his support for AC, limiting Steinitz's Theorem, for instance, to the countable case ([Waer], 2nd edition, 1937). Nevertheless, in the proof of the uniqueness of algebraic closure, he could not help mentioning the necessity of the Axiom of Choice. Strongly urged by other friends, he reinstated AC, transfinite induction, and Steinitz's theory in the 3rd edition (1950).

Argument 4 (Proof given by Zorn, [Zorn]). This proof is like the construction of $\overline{\mathbb{Q}}$ over $\mathbb{Q}$. In short, one need only adjoin all the roots of all polynomials to the ground field K. For every monic irreducible polynomial $p(X) \in K[X]$ of degree n, consider n indeterminates $y_1^{(p)}, ..., y_n^{(p)}$; Let Y be the set of all such $y_i^{(p)}$. In the ring $K[Y]$, let I be the ideal generated by all the coefficients $z_1^{(p)}, ..., z_n^{(p)}$ of the form

$$p(X) - (X - y_1^{(p)}) \cdots (X - y_n^{(p)}) = z_1^{(p)} X^{n-1} + \cdots + z_n^{(p)}.$$

By Zorn's Lemma, there exists a maximal ideal $J \supseteq I$ in the ring $K[Y]$. The quotient

$\Omega = K[Y]/J$ is a field containing a subfield isomorphic to K, which is identified with K. The field Ω thus obtained is an algebraic closure of the field K.

This argument contains nothing doubtful, since it is done without any recurrence procedure.

Historical Note. Influenced by the 1st edition of van der Waerden's textbook, Zorn published his proof of Steinitz's Theorem based on his type of maximal principle ([Zorn]). Although several kinds of maximal principles had been sporadically studied ([Kurat], for instance), Zorn was the first who recognized the principle not as a theorem but as an axiom. He observed that his maximal principle yields AC, but it was Bourbaki and J. W. Tukey that established their equivalence. Algebraists were soon to apply the maximal principle to a variety of applications. Until then they had used the Well-Ordering Theorem, though with some degree of discomfort. Chevalley introduced the maximal principle to Bourbaki, who became most influential in spreading it to the mathematical world.

Argument 5 (attributed to Artin). Define inductively a sequence $(L_n)_{n\geq 0}$ of fields as follows: Put $L_0 = K$; and if $n \geq 0$ and a field L_n has been defined, construct an algebraic extension L_{n+1} of L_n in which every nonconstant polynomial in $L_n[X]$ admits a zero. Such a construction is possible by a similar (but simpler) method to Zorn's proof stated above. Then the compositum $\bigcup_{n=0}^{\infty} L_n$ is an algebraic closure of K.

This is a typical misuse of 'mathematical induction', though widely known as a simple 'proof'. In fact, the extension L_{n+1} is not uniquely constructed over L_n even up to isomorphism in this stage of discussion. Recall that we use Zorn's Lemma to get a maximal ideal in the construction of L_{n+1} over L_n. In addition, L_n is not contained in but, in reality, is embedded in L_{n+1} as it was in Argument 3. In short, since there are too many fields to be called a set as a whole, we cannot make use of PDC, not to mention mathematical induction.

When one wants to closely examine the validity of proof, one must be careful especially in dealing with infinite sequences: For, *though in algebra we abuse the method of identification by embedding, it is not a formal procedure.*

One way of repairing the fault in Argument 5 is to prove the following remark pointed out by Bourbaki ([Bour], Ch. V, §3) which says that the construction of a sequence of fields in Artin's proof is superfluous:

If Ω is algebraic over K and if every non-constant polynomial of $K[X]$ has a root in Ω, then Ω is an algebraic closure of K.

For the proof, which makes use of Galois Theory, see the excellent textbook of Bastida ([Bast], p.157).

Argument 6. Consider the set S of all algebraic extensions L of K. Order S by writing

$L_1 \leq L_2$ for $L_1, L_2 \in S$ when $L_1 \subseteq L_2$. It is very easy to prove that S is an inductive set. Therefore it follows that there exists a maximal element Ω of S. The existence of a proper algebraic extension of Ω would contradict the maximality of Ω. Thus Ω is found to be an algebraic closure of K.

Even the laziest student of mathematics would find that this argument is false. In short, there exist so many fields to be dealt with that we cannot say S forms a set as a whole.

One way to correct the fallacy in Argument 6 is found in [Zar/Sam] and [Brau]. The proof runs as follows. Let P denote the set of all monic irreducible polynomials $p(X) \in K[X]$. Form the set $P \times \mathbb{N}$ consisting all (p, r), $p \in P; r = 1, 2, \ldots$ and let H be the subset consisting of (p, r) with $r \leq \deg p$. If $p = X - c$ with $c \in K$, then we agree to replace $(p, 1)$ by c of K. Thus $K \subset H$. A field L is called admissible, if the following three conditions are satisfied:

1. $L \subset H$.
2. If $(p, r) \in L$, then (p, r) is a zero of the polynomial $p(X) \in L[X]$.
3. $K \subset L$.

Next form the set S of all admissible fields L. if we work in this set S, the proof goes the same as in Argument 6.

The difficulty is in making everything lie in a 'set'. However, we should say that this device is too complicated and artificial for such a basic theorem as Steinitz's.

§3. Corredting the Arguments by Using a Universal Set

The defect common to all of the preceding naive arguments is reduced to the fact that a totality of too many sets cannot be a set. While, for a formula P, $\{x \mid P(x)\}$ is not a set in general, $\{x \mid P(x)\} \cap S$ is a set for any set S (the *Axiom of Separation*). This leads us to the following correction of Argument 6:

Choose a set D containing the ground field K such that $|D| > |K|$; in case K is finite, choose any uncountable set as D. Let Γ be the set of all algebraic extensions of K contained in D:

$$\Gamma = \{L\ (\subset D) \mid L \text{ is an algebraic extension field of } K\}$$

Order the set Γ by defining $L_1 \leq L_2$ if L_1 is a subfield of L_2. It is easy to see that Γ is a nonempty inductive set. Then a maximal element Ω of Γ is an algebraic closure of K. In fact, if Ω had a proper algebraic extension Ω', then we might suppose that Ω' is a finite extension of Ω. Then, since it is easy to prove that $| \Omega' | = | \Omega | < | D |$, Γ would contain a

proper extension Ω'' of Ω which is isomorphic to Ω', which contradicts the maximality of Ω.

We owe this simple and elegant proof to Jacobson ([Jacob], Ch. IX). The same trick works for Arguments 3 and 5 as well. For example, to repair Argument 5, under the same notation as above, define a relation $\prec$ in Γ by writing $L \prec L'$ if every non-constant polynomial with coefficients in L admits a zero in L'; L, then, is necessarily a subfield of L'. By using the fact that an algebraic (not necessarily finite) extension does not increase the cardinality of an infinite field, we can prove that for every L in Γ there exists some L' in Γ with $L \prec L'$, as stated in Argument 5. Then we use PDC to form an ascending chain

$$L_0 \prec L_1 \prec L_2 \cdots \prec L_n \prec \cdots .$$

The union $\bigcup_{n=0}^{\infty} L_n$ is certainly an algebraic closure of K.

Finally, for the sake of completeness only, we mention other possible proofs with recourse to mathematical logic.

One which is set theoretical is Steinitz's (and hence van der Waerden's simplified) proof using transfinite induction. The following is a kind of model theoretic proof.

Let P be the set of all irreducible monic polynomials with coefficients in K, and let I be the set of all finite subsets of P. For each $F \in I$ take an extension field of K in which any $f \in F$ splits into a product of linear factors; here we need AC. Define $V_f = \{F \mid f \in F \in I\}$. Then since $\{V_f \mid f \in P\}$ has the finite intersection property, there exists an ultrafilter $\mathcal{F}$ such that $V_f \in \mathcal{F}$ $(f \in P)$. The set of all algebraic elements over K in $\prod_{F \in I} K_F/\mathcal{F}$ is an algebraic closure of K. This argument is close to Bourbaki's proof which uses the tensor product of all splitting fields of irreducible polynomials ([Bour], V, §4).

Alternatively, we may proceed this way. Choose an axiom of set theory which is not usually used in algebra, say, the Axiom of Power Sets, and call it P. Next take a *set* X which satisifies all axioms of set theory except P, and develop algebra in X. All those naive arguments are then justified (needless to say, PDC must be used instead of mathematical induction), since though the class of all algeraic extensions of K in X is *not* a set in the model X, it *is* a set in the whole set theory.

As a final remark, it is known that there is a model of ZF in which a field exists which has no algebraic closure ([Jech], Ch. 10). Thus AC is indispensable in the proof of Steinitz's Theorem. For detailed deductive relations, see Appendix 2 of [Moor].

Conclusion. We can justify most of doubtful inferences in algebra related to an infinite sequence of sets by presupposing that we are working in a kind of universal set, say D, with a larger cardinality. However, we must invoke the power of PDC when the recursion procedure to get from one set to the next admits a range of choice. Mathematical induction is of no use in this case.

Which argument is the best as a heuristic proof of Steinitz's Theorem for introductory courses or textbooks of algebra? The author recommends Argument 6. It is not only the simplest in idea but also the easiest to correct. There are other reasons for this preference. Steinitz's Theorem has itself a set-theoretical character. An algebraic closure of a field K is defined as a maximal algebraic extension of K. Hence it is not so artificial to require a little cardinal arithmetic. Moreover, Steinitz's Theorem is so basic that it is desirable to prove it at a stage as early as possible. Certainly not after introducing normality, separability, and so on, because an algebraic closure plays the role of a universe for such objects as equations and algebraic extension fields, for instance. Arguments 4 and 5, even heuristically, need the notion of polynomial ring in infinitely many variables which rarely, if ever, plays an essential role in other places.

The author wishes to express his thanks to K. Eda, his colleague, for valuable suggestions on the discussion related to set theory in this article.

References

[Bast] J. R. Bastida, *Field Extensions and Galois Theory*, Addison-Wesley, 1984

[Bour] N. Bourbaki, *Eléments de mathèmatique, Algèbre II*, Actualités Sci. Indust., Hermann, 1959

[Brau] R. Brauer, *Galois Theory*, Harvard Univ. 1957/8, revised 1963/64

[Jacob] N. Jacobson, *Lectures in Abstract Algebra III-Theory of Fields and Galois Theory*, Van Nostrand, 1964

[Jech] T. H. Jech, *The Axiom of Choice*, North-Holland, 1973

[Kurat] K. Kuratowski, Une méthode d'élimination des nombres transfinis des raisonnements mathématiques, Fund. Math., V (1922), 76-108

[Moor] G. H. Moore, *Zermelo's Axiom of Choice*, Springer-Verlag, 1982

[Stein] E. Steinitz, Algebraische Theorie der Körpern, J. reine und angew. Math. 137(1910), 167-309

[Waer] B. L. van der Waerden, *Moderne Algebra*, Springer-Verlag, 1930, 1937, 1950

[Zar/Sam] O. Zariski, O.-P. Samuel, *Commutative Algebra, Vol. I*, Van Nostrand, 1958

[Zorn] M. Zorn, A Remark on method in transfinite algebra, Bull. Amer. Math. Soc. 41 (1935), 667-670