9. Transfinite induction and recursion

April 16, 2015

In this chapter we continue the discussion of ordinal numbers, generalizing the induction and recursion principles to them. So far we have only used ordinals in order to define cardinals and size of sets. In this chapter we do not use the axiom of choice. Now we consider ordinals in general, as a generalization of the order aspect of the natural numbers. Let us list the main things that we know about ordinals so far. Recall that an ordinal is defined to be a transitive set a well-ordered by the set $\in_a$; this set is $\{(b,c) : b, c \in a$ and $b \in c\}$.

7.1: Every natural number is an ordinal.
7.2: ω is an ordinal.
7.3: The ordering $<$ between ordinals, which is the same as $\in$, is transitive.
7.4: Every member of an ordinal is an ordinal.
7.5: $\alpha \subset \beta$ iff $\alpha \in \beta$, for ordinals α, β.
7.6: trichotomy: any two ordinals are comparable under $\leq$.
7.7: Any nonempty set of ordinals has a least element.
8.3: If α and β are order-isomorphic ordinals, then $\alpha = \beta$.
8.6: Every well-ordered set is order-isomorphic to a unique ordinal.
8.28: If Γ is a set of ordinals, then $\bigcup \Gamma$ is an ordinal, and it is the least upper bound of Γ.

We now develop some further simple properties of ordinals.

Proposition 9.1. *If x is an ordinal, then so is $x \cup \{x\}$.*

Proof. If $z \in y \in x \cup \{x\}$, then either $z \in y \in x$ and so $z \in x \subseteq x \cup \{x\}$ because x is transitive, or $z \in y = x$, so obviously $z \in x \cup \{x\}$. This proves that $x \cup \{x\}$ is transitive. If A is a nonempty subset of $x \cup \{x\}$, there are two possibilities. If $A = \{x\}$, then x is the least element of A, since it is the only element of A. If $A \neq \{x\}$, then $A \cap x \neq \emptyset$, and the least element of $A \cap x$ is also the least element of A. Hence $x \cup \{x\}$ is an ordinal. □

In a later chapter we will introduce addition, multiplication, and exponentiation of ordinals. Then, as for the natural numbers, $\alpha \cup \{\alpha\}$ will turn out to equal $\alpha +_o 1$. We use a subscript for $+_o$ since this is not the same as $\alpha + 1$ for a cardinal α. To aid the intuition we define $\alpha +_o 1 = \alpha \cup \{\alpha\}$ now; and when we introduce addition of ordinals in general, this will be seen to be consistent with that definition.

Corollary 9.2. *For any ordinals α, β we have: $\alpha < \beta$ iff $\alpha +_o 1 \leq \beta$.*

Proof. $\Rightarrow$: Assume that $\alpha < \beta$. If $\beta < \alpha +_o 1$, then $\beta \in \alpha \cup \{\alpha\}$, hence $\beta \leq \alpha < \beta$, contradiction. So $\alpha +_o 1 < \beta$ by 7.6.

$\Leftarrow$: Assume that $\alpha +_o 1 \leq \beta$. Hence $\alpha \cup \{\alpha\} \subseteq \beta$, so $\alpha \in \beta$, hence $\alpha < \beta$. □

There are so many ordinals that the set of all of them is one of the illegal sets mentioned before. This is an important fact philosophically, and is actually technically useful too.

Theorem 9.3. *There does not exist a set which has all ordinals as members.*

Proof. Assume that A is such a set, and let $B = \{x \in A : x \text{ is an ordinal}\}$. Thus B is the set of all ordinals. Since every member of an ordinal is an ordinal, B is a transitive set. By 7.7 (see above), B is well-ordered by $\in$, so B itself is an ordinal. Hence $B \in B$, contradiction. $\square$

We now expand a little on Theorem 7.7.

Theorem 9.4. *If Γ is a nonempty set of ordinals, then $\bigcap \Gamma$ is an ordinal, and is in fact the least element of Γ.*

Proof. We know from 7.7 that Γ has a least element α, so it suffices to show that $\alpha = \bigcap \Gamma$. If $\beta \in \Gamma$, then $\alpha \leq \beta$, which means by 7.5 that $\alpha \subseteq \beta$. Thus $\alpha \subseteq \beta$ for every $\beta \in \Gamma$, so $\alpha \subseteq \bigcap \Gamma$.

We claim that $\bigcap \Gamma$ is an ordinal. To show that it is transitive, suppose that $x \in y \in \bigcap \Gamma$. If $\beta \in \Gamma$, then $y \in \beta$, so also $x \in \beta$ since β is transitive. So $x \in \bigcap \Gamma$. This shows that $\bigcap \Gamma$ is transitive. Now Γ has some member γ, so $\bigcap \Gamma \subseteq \gamma$, hence $\bigcap \Gamma$ is well-ordered by $\in$ because γ is. So $\bigcap \Gamma$ is an ordinal.

Thus our assertion above that $\alpha \subseteq \bigcap \Gamma$ implies by 7.5 that $\alpha \leq \bigcap \Gamma$. But also $\alpha \in \Gamma$, so $\bigcap \Gamma \subseteq \alpha$ and hence $\bigcap \Gamma \leq \alpha$ by 7.5. So $\bigcap \Gamma = \alpha$. $\square$

We now introduce a standard classification of ordinals; ordinals are of three mutually exclusive types:

- The ordinal 0.
- Successor ordinals: ordinals of the form $\alpha +_o 1$ for some ordinal α.
- Limit ordinals: nonzero ordinals which are not successor ordinals.

Examples of successor ordinals are $1, 2, 3, \ldots$ and also $\omega +_o 1$. So far we have encountered only one limit ordinal, namely ω.

Proposition 9.5. *Every infinite cardinal is a limit ordinal.*

Proof. Let κ be an infinite cardinal. Thus $\omega \leq \kappa$, so $\omega \subseteq \kappa$. Suppose that α is an ordinal and $\kappa = \alpha +_o 1$; we want to get a contradiction. We will do this by defining a bijection f from κ to α. We define, for any $\beta < \kappa$,

$$f(\beta) = \begin{cases} \beta +_o 1 & \text{if } \beta < \omega, \\ \beta & \text{if } \omega \leq \beta < \alpha, \\ 0 & \text{if } \beta = \alpha. \end{cases}$$

Clearly f maps κ into α. Now $f \restriction \omega$ is a bijection from ω to $\omega \backslash \{0\}$, $f \restriction (\alpha \backslash \omega)$ is the identity map on $\alpha \backslash \omega$, and f maps α itself to 0. Clearly then f is the indicated bijection, contradiction. $\square$

Proposition 9.6. *The following conditions are equivalent:*

(i) α is a limit ordinal;

(ii) $\alpha \neq 0$, and for every $\beta < \alpha$ there is a γ such that $\beta < \gamma < \alpha$.

(iii) $\alpha = \bigcup \alpha \neq 0$.

Proof. (i)⇒(ii): suppose that α is a limit ordinal. So $\alpha \neq 0$, by definition. Suppose that $\beta < \alpha$. By 9.2 we have $\beta +_o 1 < \alpha$, since α is not a successor ordinal. Thus $\gamma = \beta +_o 1$ works as indicated.

(ii)⇒(iii): if $\beta \in \bigcup \alpha$, choose $\gamma \in \alpha$ such that $\beta \in \gamma$. Then $\beta \in \alpha$ since α is an ordinal. This shows that $\bigcup \alpha \subseteq \alpha$.

Conversely, if $\beta \in \alpha$, choose γ with $\beta < \gamma < \alpha$. Thus $\beta \in \bigcup \alpha$. This proves that $\alpha = \bigcup \alpha$, and $\alpha \neq 0$ is given.

(iii)⇒(i): suppose that $\alpha = \beta +_o 1$. Then $\beta \in \beta \cup \{\beta\} = \beta +_o 1 = \alpha = \bigcup \alpha$, so choose $\gamma \in \alpha$ such that $\beta \in \gamma$. Thus $\beta < \gamma \leq \beta$ by 9.2, so $\beta < \beta$, contradiction. Thus α is not a sucessor ordinal. By (iii) it is also not 0, so it is a limit ordinal. □

Proposition 9.7. *If* $\alpha = \beta +_o 1$, *then* $\bigcup \alpha = \beta$.

Proof. If $\gamma \in \beta$, then $\gamma \in \beta \in \alpha$, so $\gamma \in \bigcup \alpha$. Suppose that $\gamma \in \bigcup \alpha$. Choose $\delta \in \alpha$ such that $\gamma \in \delta$. Then $\gamma < \delta \leq \beta$, so $\gamma < \beta$, i.e., $\gamma \in \beta$. □

Now we turn to the discussion of transfinite induction. This is a generalization of induction on the natural numbers, ω, to induction on other ordinals. The principle of complete transfinite induction generalizes 6.8, the principle of complete induction on ω.

Theorem 9.8. (Complete transfinite induction) *Suppose that* α *is an ordinal,* $\Gamma \subseteq \alpha$, *and*

$(*)$ *For all* $\beta < \alpha$, *if* $\gamma \in \Gamma$ *for all* $\gamma < \beta$, *then* $\beta \in \Gamma$.

Then $\Gamma = \alpha$.

Proof. Suppose not. Then $\alpha \backslash \Gamma$ is nonempty, and we let β be the least element of it. Thus $\gamma \in \Gamma$ for all $\gamma < \beta$, so by the assumption $(*)$, also $\beta \in \Gamma$, contradiction. □

There is also an ordinary principle of transfinite induction, in which the argument goes step-by-step, except for limit ordinals, where we have to do complete induction again.

Theorem 9.9. (Ordinary transfinite induction) *Suppose that* α *is an ordinal,* $\Gamma \subseteq \alpha$, *and the following three conditions hold:*

(i) If $0 < \alpha$, *then* $0 \in \Gamma$.

(ii) If $\beta +_o 1 < \alpha$ *and* $\beta \in \Gamma$, *then* $\beta +_o 1 \in \Gamma$.

(iii) If $\beta < \alpha$ *is a limit ordinal, and if* $\gamma \in \Gamma$ *for all* $\gamma < \beta$, *then* $\beta \in \Gamma$.

Under these assumptions, $\Gamma = \alpha$.

Proof. Again, suppose not, and let β be the least element of $\alpha \backslash \Gamma$. Then $\beta \neq 0$ by (i). Suppose that β is a successor ordinal $\gamma +_o 1$. Then $\gamma \in \Gamma$, and (ii) is contradicted. Finally, if β is a limit ordinal, then (iii) is contradicted. □

There are also transfinite induction principles which involve properties of ordinals rather than sets of ordinals. The intuition is like that for sets, except that properties might be too big to be considered as sets. The situation is similar to that for the comprehension axioms, where we had properties which could be used to define subsets of a given set. Now, though, we are not trying to introduce any new axioms, but just to indicate another valid form of reasoning. When, later, we give concrete applications of these "property" forms of transfinite induction, the notions should become clearer.

Theorem 9.10. (Complete transfinite induction principle, for properties) *Let $A(\alpha)$ be a property of ordinals. Assume also:*

($*$) *For every ordinal β, if $A(\gamma)$ holds for all $\gamma < \beta$, then $A(\beta)$ holds.*

Then $A(\alpha)$ holds for every ordinal α.

Proof. Suppose not. Choose β so that $A(\beta)$ is false. Then the set

$$\Gamma \stackrel{\text{def}}{=} \{\gamma \in \beta +_o 1 : A(\gamma) \text{ fails}\}$$

is nonempty, since β is in it. (Remember once again that $\beta +_o 1 = \beta \cup \{\beta\}$.) Since Γ is a nonempty subset of the ordinal $\beta +_o 1$, it has a least element, γ. Now for every $\delta < \gamma$ we have $\delta \in \beta +_o 1$ because $\beta +_o 1$ is an ordinal. By the choice of γ, then, it must be the case that $\delta \notin \Gamma$, and so $A(\delta)$ holds. This being true for every $\delta < \gamma$, the hypothesis ($*$) implies that $A(\gamma)$ holds too, contradicting the fact that $\gamma \in \Gamma$. $\square$

Theorem 9.11. (Ordinary transfinite induction, for properties) *Let $A(\alpha)$ be a property of ordinals satisfying the following conditions:*

(i) $A(0)$ holds;

(ii) for all α, if $A(\alpha)$ then $A(\alpha +_o 1)$;

(iii) for every limit ordinal α, if $A(\beta)$ for all $\beta < \alpha$, then $A(\alpha)$.

Then $A(\alpha)$ for every ordinal α.

Proof. Suppose not, and let α be any ordinal such that $A(\alpha)$ fails. Let β be the least element of $\{\gamma \in \alpha +_o 1 : A(\gamma) \text{ fails}\}$. By (i), $\beta \neq 0$. By (ii), β is not a successor ordinal. So β is a limit ordinal, contradicting (iii).

Now we turn to transfinite recursion.

Theorem 9.12. (Transfinite recursion principle) *Suppose that α is an ordinal, A is a set, and g is a function mapping $\bigcup_{\beta<\alpha}({}^{\beta}A)$ into A. Then there is a unique function $f : \alpha \to A$ such that, for every $\beta < \alpha$,*

$$f(\beta) = g(f \restriction \beta).$$

Proof. The proof is very similar to that of 6.10, the recursion theorem for ω. Let $B = \bigcup_{\beta\leq\alpha}({}^{\beta}A)$. Define

$$\begin{aligned}\mathscr{F} = \{h :\ &\text{there is a } \beta \leq \alpha \text{ such that } h : \beta \to A\\ &\text{and for all } \gamma < \beta,\ h(\gamma) = g(h \restriction \gamma))\}.\end{aligned}$$

Each such h is in B, so this definition is legal. As in the case of the natural numbers, we are considering here all of the approximations to the function we are trying to find.

We claim

(1) If $h, k \in \mathscr{F}$, $h : \beta \to A$, $k : \gamma \to A$, and $\beta \leq \gamma$, then $h = k \restriction \beta$.

To prove (1), we need to show that $h(\delta) = k(\delta)$ for every $\delta < \beta$. We do this by transfinite induction. Suppose that $\delta < \beta$ and $h(\varepsilon) = k(\varepsilon)$ for all $\varepsilon < \delta$. Then $h \restriction \delta = k \restriction \delta$, and so $h(\delta) = g(h \restriction \delta) = g(k \restriction \delta) = k(\delta)$. This finishes the inductive proof. So (1) holds.

(2) For every $\beta \leq \alpha$ there is an $h \in \mathscr{F}$ such that $\mathrm{dmn}(h) = \beta$.

Again we prove this by transfinite induction. Suppose that $\beta \leq \alpha$, and for every $\gamma < \beta$ there is an $h \in \mathscr{F}$ such that $\mathrm{dmn}(h) = \gamma$. This function is unique by (1), so call it $k(\gamma)$. Thus we have defined a function k with domain β, at least intuitively. Let us see how we get it more rigorously. So, we redefine k:

$$k = \{(\gamma, h) \in \beta \times \mathscr{F} : \mathrm{dmn}(h) = \gamma\}.$$

By (1), k is a function, and by the inductive assumption it has domain β, as desired.

Now we define

$$h = \bigcup_{\gamma<\beta} k(\gamma).$$

By (1), h is still a function. Let δ be its domain.

First suppose that β is a successor ordinal $\gamma +_o 1$. Then clearly $h = k(\gamma)$. Let $s = h \cup \{(\gamma, g(h))\}$. Then s is a function with domain β and range included in A, If $\varepsilon < \gamma$, then

$$s(\varepsilon) = (k(\gamma))(\varepsilon) = g((k(\gamma)) \restriction \varepsilon) = g(s \restriction \varepsilon).$$

And $s(\gamma) = g(h) = g(s \restriction \gamma)$. It follows that $s \in \mathscr{F}$ and $\mathrm{dmn}(s) = \beta$.

Second, suppose that $\beta = 0$. In this case it is trivial that h is in $\mathscr{F}$. Third, suppose that β is a limit ordinal. Then $\beta = \delta$, and clearly $h \in \mathscr{F}$.

This completes the inductive proof of (2). The case $\beta = \alpha$ in (2) gives the existence of the function required in the theorem.

Now we prove uniqueness. Suppose that both f and k satisfy the conditions of the theorem. We prove that $f(\beta) = k(\beta)$ for all $\beta < \alpha$ by transfinite induction. Suppose that this is true for all $\beta < \gamma$, where $\gamma < \alpha$. Then $f \restriction \gamma = k \restriction \gamma$, and so

$$f(\gamma) = g(f \restriction \gamma) = g(k \restriction \gamma) = k(\gamma),$$

finishing the inductive proof. □

There is also a version of transfinite recursion involving class functions. Remember from our treatment of the ordinal representation theorem that this is an intuitive concept of a function that may be too big to actually be a set. Before giving this general recursion principle we need a lemma indicating when a class function can be represented by a set.

Lemma 9.13. *Suppose that F is a class function, defined for all ordinals, and α is an particular ordinal. Then there is a unique function f with domain α such that $f(\xi) = F(\xi)$ for every $\xi < \alpha$.*

Proof. By the axiom of replacement, the set

$$A \stackrel{\text{def}}{=} \{x : F(\xi) = x \text{ for some } \xi < \alpha\}$$

exists. Let

$$f = \{(\xi, \eta) \in \alpha \times A : F(\xi) = \eta\}.$$

Clearly f is as desired, and it is unique. □

The function f asserted to exist in this lemma will be denoted by $F \restriction \alpha$.

Theorem 9.14. (Class version of the transfinite recursion principle) *Suppose that G is a class function whose domain consists of all (set) functions. Then there is a unique class function F defined for all ordinals such that for every ordinal α we have $F(\alpha) = G(F \restriction \alpha)$.*

Proof. This proof is very similar to that of 9.12. We consider the following condition:

(*) f is a function with domain α, and for every $\xi < \alpha$ we have $f(\xi) = G(f \restriction \xi)$.

First we show

(1) If f, α satisfy (*) and g, β satisfy (*) and $\alpha \leq \beta$, then $f = g \restriction \alpha$.

To prove this, we prove by transfinite induction on ξ that if $\xi < \alpha$ then $f(\xi) = g(\xi)$. Suppose that this is true for all $\eta < \xi$, where $\xi < \alpha$. Then $f \restriction \xi = g \restriction \xi$, so $f(\xi) = G(f \restriction \xi) = G(g \restriction \xi) = g(\xi)$, finishing the inductive proof.

(2) For every ordinal α there is a function f such that (*) holds.

We prove this by transfinite induction. Assume that it is true for all $\beta < \alpha$. By (1), for each $\beta < \alpha$ there is a unique f satisfying (1); we denote it by f_β (the replacement axiom is being used). Let $g = \bigcup_{\beta<\alpha} f_\beta$. Then g is a function by (1), and its domain is clearly $\bigcup \alpha$.

(3) For any $\beta < \alpha$ we have $f_\beta = g \restriction \beta$, and $g(\beta) = G(g \restriction \beta)$.

In fact, the first condition is clear. For the second,

$$g(\beta) = f_{\beta +_o 1}(\beta) = G(f_{\beta +_o 1} \restriction \beta) = G(g \restriction \beta).$$

So, (3) holds.

If $\alpha = 0$, clearly (*) holds.

If α is a limit ordinal, then $\bigcup \alpha = \alpha$ and (*) holds for g and α.

Finally, suppose that $\alpha = \beta +_o 1$ for some β. Thus $\bigcup \alpha = \beta$. Let $h = g \cup \{(\beta, G(g))\}$. So, h is a function with domain α. Suppose that $\gamma < \alpha$. If $\gamma < \beta$, then $h(\gamma) = g(\gamma) = G(g \restriction \gamma) = G(h \restriction \gamma)$. If $\gamma = \beta$, then $h(\gamma) = G(g) = G(h \restriction \gamma)$. Thus h and α satisfy (*).

This finishes the inductive proof of (2).

Now for any ordinal α we let $F(\alpha) = f(\alpha)$, where f is chosen so that (*) holds for $\alpha +_o 1$ and f. This definition is unambiguous by (1). Also by (1), we have $F \restriction \alpha = f \restriction \alpha$. Hence $F(\alpha) = f(\alpha) = G(f \restriction \alpha) = G(F \restriction \alpha)$.

This finishes the proof of existence.

For uniqueness, suppose that H also satisfies the conditions of the theorem. We prove that $F(\alpha) = H(\alpha)$ for every ordinal α by induction. Suppose that this is true for all $\beta < \alpha$. Then $F \upharpoonright \alpha = H \upharpoonright \alpha$, and hence $F(\alpha) = G(F \upharpoonright \alpha) = G(H \upharpoonright \alpha) = H(\alpha)$. This finishes the inductive proof. □

As a first application of transfinite recursion we prove some theorems which show that, in a sense, all sets are built up from the emptyset!

Theorem 9.15. *There is a class function V, defined for all ordinals, with the following properties:*

(i) $V_0 = \emptyset$.

(ii) For any ordinal α, $V_{\alpha +_o 1} = \mathscr{P}(V_\alpha)$.

(iii) For any limit ordinal α, $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$.

Proof. We define a class function G, with domain the class of all functions f, as follows:

$$G(f) = \begin{cases} \emptyset & \text{if } f = \emptyset, \\ \mathscr{P}(f(\alpha)) & \text{if } \mathrm{dmn}(f) = \alpha +_o 1 \text{ for some ordinal } \alpha, \\ \bigcup_{\beta<\alpha} f(\beta) & \text{if } \mathrm{dmn}(f) \text{ is a limit ordinal } \alpha, \\ \emptyset & \text{otherwise.} \end{cases}$$

Now we apply the transfinite recursion principle to get a class function V such that for any ordinal α, $V_\alpha = G(V \upharpoonright \alpha)$. Then

$$\begin{aligned} V_0 &= G(V \upharpoonright 0) = G(\emptyset) = \emptyset; \\ V_{\alpha+_o 1} &= G(V\langle(\alpha +_o 1)) = \mathscr{P}(V_\alpha); \\ V_\alpha &= G(V \upharpoonright \alpha) = \bigcup_{\beta<\alpha} V_\beta \quad \text{for limit } \alpha. \end{aligned}$$

□

Proposition 9.16. *V_α is transitive, for every ordinal α.*

Proof. We prove this by ordinary transfinite induction for properties. Since $V_0 = \emptyset$, and $\emptyset$ is trivially transitive, our statement is true for $\alpha = 0$. Assume that V_α is transitive (induction hypothesis), and suppose that $a \in b \in V_{\alpha+_o 1}$. Since $V_{\alpha+_o 1} = \mathscr{P}(V_\alpha)$, we have $b \subseteq V_\alpha$, and hence $a \in V_\alpha$. By the inductive hypothesis, V_α is transitive, and hence $a \subseteq V_\alpha$. So $a \in V_{\alpha+_o 1}$. This finishes this inductive step.

There is another inductive step: assume that α is a limit ordinal, and V_β is transitive for all $\beta < \alpha$; we want to show that V_α is transitive. Suppose that $a \in b \in V_\alpha$. Now $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$, so we can choose $\beta < \alpha$ so that $b \in V_\beta$. By the inductive hypothesis, V_β is transitive, and $a \in b$, so $a \in V_\beta \subseteq V_\alpha$, as desired. □

Proposition 9.17. *If α and β are ordinals and $\alpha \leq \beta$, then $V_\alpha \subseteq V_\beta$.*

Proof. First note that $V_\gamma \subseteq V_{\gamma+_o 1}$ for every ordinal γ. In fact, if $x \in V_\gamma$, then $x \subseteq V_\gamma$ by 9.16, and so $x \in V_{\gamma+_o 1}$.

Now we prove the proposition by induction on β, with α fixed. If $\beta = 0$, then also $\alpha = 0$, and the conclusion is trivial. Assume that $\alpha \le \beta$ implies that $V_\alpha \subseteq V_\beta$, and suppose that $\alpha \le \beta +_o 1$. Again the conclusion is trivial if $\alpha = \beta +_o 1$, so suppose that $\alpha < \beta +_o 1$. Then $\alpha \le \beta$, and so $V_\alpha \subseteq V_\beta$ by the inductive assumption. By our initial remark, $V_\beta \subseteq V_{\beta+_o 1}$, so $V_\alpha \subseteq V_{\beta+_o 1}$. The inductive step to a limit ordinal β is very easy. □

Theorem 9.18. *For any set x there is a transitive set y such that $x \subseteq y$ and for every transitive set z, if $x \subseteq z$ then $y \subseteq z$.*

We call this set y, which is clearly unique, the *transitive closure* of x.

Proof. By recursion on ω let

$$A_0 = x,$$
$$A_{m+_o 1} = A_m \cup \bigcup A_m \quad \text{for any } m \in \omega.$$

Then let $y = \bigcup_{m\in\omega} A_m$. Clearly $x \subseteq y$. To show that y is transitive, suppose that $a \in b \in y$. Choose $m \in \omega$ so that $b \in A_m$. Then $a \in b \in A_m$, so $a \in \bigcup A_m$, and hence $a \in A_{m+_o 1} \subseteq y$, and so $a \in y$. Thus y is transitive.

Now suppose that $x \subseteq z$ and z is transitive. We show by induction on m that $A_m \subseteq z$ for all $m \in \omega$. First, $A_0 = x \subseteq z$, so this is true for $m = 0$. Now suppose that $A_m \subseteq z$ (induction hypothesis). Take any $a \in A_{m+_o 1}$. Recall that $A_{m+_o 1} = A_m \cup \bigcup A_m$. If $a \in A_m$, then $a \in z$ by the inductive hypothesis. Suppose that $a \in \bigcup A_m$. Say $a \in b \in A_m$. Since $A_m \subseteq z$ by the inductive hypothesis, we have $b \in z$. Now $a \in b \in z$ and z is transitive, so $a \in z$. This shows that $A_{m+_o 1} \subseteq z$ and finishes the inductive proof. It now follows that $y = \bigcup_{m\in\omega} A_m \subseteq z$. □

Theorem 9.19. *For any set x there is an ordinal α such that $x \in V_\alpha$.*

Proof. Suppose that this is not true, and let x be a set such that $x \notin V_\alpha$ for every ordinal α. Let y be the transitive closure of $x \cup \{x\}$. We define

$$M = \{a \in y : a \notin V_\alpha \text{ for every ordinal } \alpha\}.$$

Thus $x \in M$, so M is nonempty. Choose $a \in M$ such that $a \cap M = \emptyset$, by the foundation axiom. Thus for all $z \in a$ we have $z \notin M$; but $z \in y$ since y is transitive, so there must exist an ordinal α such that $z \in V_\alpha$; let α_z be the least such ordinal. By the axiom of replacement we have a set $\{\alpha_z : z \in a\}$; call this set Γ. Let $\beta = \bigcup \Gamma$; recall that $\bigcup \Gamma$ is an ordinal. Now for any $z \in a$ we have $z \in V_{\alpha_z} \subseteq V_\beta$ by 9.17. Thus $a \subseteq V_\beta$, so $a \in V_{\beta+_o 1}$, contradiction. □

There will be many uses of transfinite induction and recursion in the remainder of these notes.

Exercises, Chapter 9

The first few exercises give five equivalent definitions of ordinal, counting the official definition.

1. We say that x is an ordinal$_2$ iff x is transitive, and for all $y, z \in x$, either $y \in z$, $y = z$, or $z \in y$. (We say that y and z are *comparable* then.) Show that every ordinal is an ordinal$_2$.

2. We say that x is an ordinal$_3$ iff x is transitive, and for all y, if $y \subset x$ and y is transitive, then $y \in x$. Show that every ordinal$_2$ is an ordinal$_3$. Hint: apply the foundation axiom to $x \backslash y$.

3. We say that x is an ordinal$_4$ iff x is transitive, and every member of x is transitive. Show that every ordinal$_3$ is an ordinal$_4$. Hint: assume that x is an ordinal$_3$, and $y = \{z \in x : z$ is an ordinal$_4\}$, and get a contradiction from assuming that $y \subset x$.

4. Show that every ordinal$_4$ is an ordinal$_2$. Hint: Assume that x is an ordinal$_4$ and it is not an an ordinal$_2$. Find a "minimal" $y \in x$ not comparable with some $z \in x$, and take a "minimal" such z. Prove that $y = z$ (contradiction).

5. We say that x is an ordinal$_5$ iff the following two conditions hold:
   (i) for all $y \in x$, either $y \cup \{y\} = x$ or $y \cup \{y\} \in x$;
   (ii) for all $y \subseteq x$, either $\bigcup y = x$ or $\bigcup y \in x$.
Show that every ordinal$_4$ is an ordinal$_5$.

6. Show that every ordinal$_5$ is an ordinal$_4$. Hint: let $A = \{y \in x : y$ is an ordinal$_4\}$. Show that both $\bigcup A$ and $\bigcup A \cup \{\bigcup A\}$ are ordinal$_4$s. Then apply (ii) followed by (i).

7. Show that every ordinal$_4$ is an ordinal.

8. Suppose that $\alpha < \beta$, and A is a set of ordinals such for all γ with $\alpha \le \gamma < \beta$, if $\delta \in A$ for all δ such that $\alpha \le \delta < \gamma$ then $\gamma \in A$. Show that every ordinal γ such that $\alpha \le \gamma < \beta$ is in A. This is a principle of transfinite induction from α to β.

9. For any set x, let $\rho(x)$ be the least ordinal such that $x \in V_{\alpha +_o 1}$. Show that for any set x, $\rho(x) = \bigcup_{y \in x} (\rho(y) +_o 1)$.

10. (Continuing 9) Show that if $x \subseteq y$, then $\rho(x) \le \rho(y)$.

11. (Continuing 9) Show that $\rho(\alpha) = \alpha$ for any ordinal α. Hint: use complete induction on α.

12. (Continuing 9) Show that for any ordinal α, $\rho(V_\alpha) = \alpha$. Hint: first show that $\rho(\mathscr{P}(x)) = \rho(x) +_o 1$ for any set x.

13. (Continuing 9) Show that for any ordinal α, $V_\alpha = \{x : \rho(x) < \alpha\}$.

14. (Continuing 9) For any sets x, y, determine $\rho(\{x\})$, $\rho(\{x, y\})$, $\rho((x, y))$, $\rho(\bigcup x)$, and $\rho(x \times y)$ in terms of $\rho(x)$ and $\rho(y)$.

By 7.31 there is a cardinal number greater than ω; let ω_1 be the least such cardinal.

15. Show that if $A \subseteq \omega_1$ and A is countable, then $\bigcup A < \omega_1$.

A subset C of ω_1 is *closed* iff for every countable $A \subseteq C$ we have $\bigcup A \in C$.

16. Show that the intersection of any nonempty collection of closed subsets of ω_1 is again closed.

A subset C of ω_1 is *unbounded in* ω_1 iff for every $\alpha < \omega_1$ there is a $\beta \in C$ such that $\alpha < \beta$.

17. Show that C is unbounded in ω_1 iff for every $\alpha < \omega_1$ there is a $\beta \in C$ such that $\alpha \leq \beta$.

18. Let $C \subseteq \omega_1$. Prove that the following conditions are equivalent:
   (i) C is closed and unbounded.
   (ii) There is a function $f : \omega_1 \to \omega_1$ with the following properties:
      (a) For all $\alpha, \beta < \omega_1$, if $\alpha < \beta$ then $f(\alpha) < f(\beta)$.
      (b) For every limit ordinal $\alpha < \omega_1$, $f(\alpha) = \bigcup_{\beta<\alpha} f(\beta)$.
      (c) $\mathrm{rng}(f) = C$.

19. Prove that the intersection of two closed unbounded subsets of ω_1 is again closed and unbounded. Hint: for unbounded, do a back-and-forth recursion on ω.

20. Prove that the intersection of countably many closed unbounded subsets of ω_1 is again closed and unbounded. Hint: generalize the procedure used for exercise 19.

21. Show that every closed and unbounded subset of ω_1 has a limit ordinal as a member.

10. Equivalents of the axiom of choice

In this chapter we prove the equivalence of several versions of the axiom of choice, and indicate a few applications of that axiom. The set of axioms of ZFC with the axiom of choice removed is denoted by ZF; so we work in ZF in this section. We begin with a lemma which will simplify the equivalence proofs.

Lemma 10.1. *For any set A there is an ordinal α such that there is no one-one function mapping α into A.*

Proof. For each well-ordered set $(B, <)$ such that $B \subseteq A$, let $\beta_{B,<}$ be the unique ordinal to which it is isomorphic. Then let $\mathscr{W} = \{(B, <) : B \subseteq A$ and $(B, <)$ is a well-ordering$\}$, and let

$$\alpha = \left(\bigcup_{(B,<)\in\mathscr{W}} \beta_{B,<} \right) +_o 1.$$

Suppose that f is a one-one function from α into A. Let B be the range of f, and define $\prec$ to be the set $\{(b_0, b_1) : b_0, b_1 \in B$ and $f^{-1}(b_0) < f^{-1}(b_1)\}$. So $(B, \prec)$ is a well-ordering and $B \subseteq A$. Hence $\alpha = \beta_{B,\prec}$. It follows by the definition of α that $\alpha < \alpha$, contradiction. □

For ease of reference we now define several statements which we will prove are equvalent to the axiom of choice, starting with our original form of the axiom of choice.

Axiom of Choice. For any sets A, B and any function f mapping A onto B there is a function $g : B \to A$ such that $f \circ g = \mathrm{Id}_B$.

Although we have chosen this as our definition of the axiom of choice, it is not the most commonly used formulation of this axiom. This most common form is as follows.

Axiom of choice, second version. For any set A, there is a function f whose domain is $\mathscr{P}(A)\backslash\{\emptyset\}$ such that $f(X) \in X$ for all $X \in \mathscr{P}(A)\backslash\{\emptyset\}$.

We call the function in this version of the axiom of choice a *choice function for nonempty subsets of A.*

Product choice axiom. For any system $\langle A_i : i \in I\rangle$ of nonempty sets, $\prod_{i\in I} A_i \neq \emptyset$.

Zorn's Lemma. *If $(A, <)$ is a partial ordering such that $A \neq 0$ and every subset of A simply ordered by $<$ has an upper bound, then A has a maximal element under $<$, i.e., an element a such that there is no element $b \in A$ such that $a < b$.*

Here an *upper bound* for a subset X of A is an element $a \in A$ such that $x \leq a$ for all $x \in X$. This is a very popular form of the axiom of choice, particularly in advanced algebra arguments.

Well-ordering principle. *For every set A there is a well-ordering of A, i.e., there is a relation $<$ such that $(A, <)$ is a well-ordering.*

These are all of our equivalents. However, many others are known. A comprehensive list can be found in the following book; some 240 equivalents are given, with proofs of equivalence.

Rubin, H.; Rubin, J. **Equivalents of the axiom of choice.** 1985.

Theorem 10.2. *On the basis of all of our axioms except the axiom of choice, the above statements are equivalent.*

Proof. Axiom of choice $\Rightarrow$ Axiom of choice, second version: Assume the axiom of choice, and suppose that A is any set. Let

$$B = \{(X, a) : X \subseteq A \text{ and } a \in X\}$$

For each $(X, a) \in B$ let $g(X, a) = X$. Thus g maps B onto $\mathscr{P}(X)\backslash\{\emptyset\}$, so by the axiom of choice there is a function $h : \mathscr{P}(X)\backslash\{\emptyset\} \to B$ such that $g \circ h$ is the identity function on $\mathscr{P}(X)\backslash\{\emptyset\}$. Now define $f(X) = 2^{nd}(h(X))$ for every $X \in \mathscr{P}(X)\backslash\{\emptyset\}$. Then for any $X \in \mathscr{P}(X)\backslash\{\emptyset\}$ we have $g(h(X)) = X$, and hence by the definition of g, $h(X)$ must have the form (X, a). So $f(X) = a \in X$, as desired.

Axiom of choice, second version $\Rightarrow$ Product choice axiom: Assume the second version of the axiom of choice, and suppose that $\langle A_i : i \in I\rangle$ is a system of nonempty sets. Let $B = \bigcup_{i\in I} A_i$, and let f be a choice function for nonempty subsets of B. Now define $g(i) = f(A_i)$ for all $i \in I$. Then $g \in \prod_{i\in I} A_i$, as desired.

Product choice axiom $\Rightarrow$ Axiom of choice: Assume the product choice axiom, and suppose that $f : A \to B$ is a surjection; we want to find a function $g : B \to A$ such that $f \circ g = \mathrm{Id}_B$. For each $b \in B$ let $A_b = \{a \in A : f(a) = b\}$. Then $\langle A_b : b \in B\rangle$ is a system of nonempty sets, since f maps onto B. Let $g \in \prod_{b\in B} A_b$. Then for any $b \in B$ we have $g(b) \in A_b$, and hence $f(g(b)) = b$, as desired.

Axiom of choice $\Rightarrow$ well-ordering principle: Let a set A be given. We want to well-order A. The idea is to use transfinite recursion to list out the elements of A, one after the other, thereby producing a well-order of A. Let γ be chosen by Lemma 10.1: it is an ordinal such that there is no one-one function from γ into A. Let f be a choice function for $\mathscr{P}(A)\backslash\{0\}$, i.e., assume that f is a function with domain $\mathscr{P}(A)\backslash\{0\}$ such that $f(X) \in X$ for every $X \in \mathscr{P}(A)\backslash\{0\}$. By the transfinite recursion principle on γ there is a function h such that for every ordinal $\alpha < \gamma$,

$$h(\alpha) = \begin{cases} f(A\backslash h[\alpha]) & \text{if } A\backslash h[\alpha] \neq 0, \\ A & \text{if } A\backslash h[\alpha] = 0. \end{cases}$$

The idea here is that the first clause in the definition cannot always hold, since eventually we run out of elements. Taking the first time that we run out of elements, we get our desired listing of elements of A. Rigorously, we first show

(1) If $\alpha < \beta$ and $h(\alpha) = A$, then $h(\beta) = A$.

In fact, assume that $\alpha < \beta$ and $h(\alpha) = A$. So by the definition, $A\backslash h[\alpha] = 0$, which means that $A \subseteq h[\alpha]$. Now $h[\alpha] \subseteq h[\beta]$, so also $A \subseteq h[\beta]$, and so $A\backslash h[\beta] = 0$; so $h(\beta) = A$, as asserted in (1).

(2) If $\alpha < \beta$ and $h(\beta) \neq A$, then $h(\alpha) \neq h(\beta)$.

In fact, assume that $\alpha < \beta$ and $h(\beta) \neq A$. By (1), also $h(\alpha) \neq A$. Then $h(\beta) = f(A\backslash h[\beta]) \in A\backslash h[\beta]$, and $h(\alpha) \in h[\beta]$, so $h(\beta) \neq h(\alpha)$.

(3) There is an ordinal α such that $h(\alpha) = A$.

For, otherwise by (2) h is a one-one function from γ into A, contradicting the choice of γ.

By (3), let α be minimum such that $h(\alpha) = A$. Let $k = h \restriction \alpha$. Then k is a one-one function mapping α onto A. We define $a_0 < a_1$ iff $k^{-1}(a_0) < k^{-1}(a_1)$ for all $a_0, a_1 \in A$. Hence for any ordinals $\beta, \delta < \alpha$ we have $\alpha < \beta$ iff $k(\alpha) < k(\beta)$. From this it is straightforward to check that $<$, as defined on A, is a well-ordering; see an exercise.

Well-ordering principle $\Rightarrow$ Zorn's lemma. Let $(A, <)$ be a partial ordering such that $A \neq 0$ and every subset of A simply ordered by $<$ has an upper bound. Also, let $\prec$ be a well-ordering of A. Again, choose an ordinal γ such that there is no one-one function from γ into A. For any subset X of A we set $\mathrm{sub}(X) = \{a \in A : b < a \text{ for all } b \in X\}$. ("sub" for "strict upper bound") Then we define by transfinite recursion, for each $\alpha < \gamma$,

$$h(\alpha) = \begin{cases} \prec\text{-least element of } \mathrm{sub}(h[\alpha] \cap A) & \text{if there is such,} \\ A & \text{otherwise.} \end{cases}$$

The idea is similar to the above. We start listing out some elements of A in increasing order according to both $\prec$ and $<$, and eventually we must stop; the stopping place is a maximal element of A. Rigorously the argument is similar to the above:

(1) If $\alpha < \beta$ and $h(\alpha) = A$, then $h(\beta) = A$.

For, assume that $\alpha < \beta$ and $h(\alpha) = A$. This means that $\mathrm{sub}(h[\alpha] \cap A)$ is empty, i.e., there is no element of A which is $>$ each element of $h[\alpha] \cap A$. Since $h[\alpha] \subseteq h[\beta]$, the same is true for β, so $h(\beta) = A$. Thus (1) holds.

(2) If $\alpha < \gamma$ and $h(\alpha) \neq A$, then $h(\alpha) \in A$.

This is clear.

(3) If $\alpha < \beta$ and $h(\beta) \neq A$, then $h(\alpha) < h(\beta)$.

For, by (1), also $h(\alpha) \neq A$, so $h(\alpha) \in h[\beta] \cap A$. Hence $h(\alpha) < h(\beta)$.

(4) There is an ordinal α such that $h(\alpha) = A$.

In fact, otherwise h is a one-one function from γ into A, contradicting the definition of γ.

Again, we take the least α such that $h(\alpha) = A$. Let $f = h \restriction \alpha$. By (3), we have $f(\beta) < f(\gamma)$ whenever $\beta < \gamma < \alpha$. Also note that $\alpha \neq 0$, since, A being nonempty, $h[0] = 0$ has a strict upper bound. Hence by the hypothesis of Zorn's lemma, there is an element $a \in A$ such that $f(\beta) \leq a$ for all $\beta < \alpha$. Since $f[\alpha]$ has no strict upper bound in A, there is no element $b \in A$ such that $a < b$. So a is a maximal element of A.

Zorn's lemma $\Rightarrow$ axiom of choice. Let A be any set. Define

$$\mathscr{A} = \{f : f \text{ is a function and } \operatorname{dmn} f \subseteq \mathscr{P}(A)\backslash\{\emptyset\} \text{ and } f(a) \in a \text{ for all } a \in \operatorname{dmn}(f)\}.$$

Now $\mathscr{A} \neq 0$, since $\emptyset \in \mathscr{A}$. $\mathscr{A}$ is partially ordered by $\subseteq$. If $\mathscr{B}$ is a subset of $\mathscr{A}$ simply ordered by inclusion, then $\bigcup_{f\in\mathscr{B}} f$ is a function. For, if $(x,y),(x,z) \in \bigcup_{f\in\mathscr{B}} f$, choose $f,g \in \mathscr{B}$ such that $(x,y) \in f$ and $(x,z) \in g$. Since $\mathscr{B}$ is simply ordered by $\subseteq$, say $f \subseteq g$. So $(x,y),(x,z) \in g$, hence $y = z$. Clearly, then, $\bigcup_{f\in\mathscr{B}} f \in \mathscr{A}$, and it is an upper bound for $\mathscr{B}$. Therefore, by Zorn's lemma let h be a maximal member of $\mathscr{A}$ under $\subseteq$. Suppose that $\operatorname{dmn} h \neq \mathscr{P}(A)\backslash\{\emptyset\}$. Take any $a \in (\mathscr{P}(A)\backslash\{\emptyset\})\backslash\operatorname{dmn} h$, and take any $x \in a$. Let $f = h \cup \{(a,x)\}$. Then clearly $f \in \mathscr{A}$ and $h \subset f$, contradiction. It follows that $\operatorname{dmn} h = \mathscr{P}(A)\backslash\{\emptyset\}$, and it is the desired choice function. $\square$

Applications of the axiom of choice in analysis are usually straightforward uses of choice functions. We give a few applications of the axiom of choice in algebra. Each application requires some knowledge of other parts of mathematics. We state exactly what knowledge is needed, and then the treatment here will be self-contained based upon that knowledge.

Vector spaces. We consider vector spaces over fields. If the notion of a field is not familiar, little is lost in this discussion if one assumes that the field is the usual system of real numbers with ordinary addition and multiplication.

A *vector space* over a field F is a triple $(V, +, \cdot)$ such that $+$ is a binary operation on V, $\cdot$ maps $F \times V$ into V, and the following conditions hold for all $x, v, w \in V$ and all $a, b \in F$; members of V are called *vectors* and members of F are called *scalars*.

(1) $x + (v + w) = (x + v) + w$.

(2) $x + v = v + x$.

(3) There is a unique vector $z \in V$ such that $z + v = v + z = v$ for all $v \in V$; we denote this vector by 0.

(4) For any $v \in V$ there is a unique vector w such that $v + w = w + v = 0$.

Given a vector space with this notation, a *linear combination* of members of a subset $X \subseteq V$ is a vector of the form

$$a_1v_1 + a_2v_2 + \cdots + a_mv_m \tag{$*$}$$

with $m \in \omega$, each $a_i \in F$, and each $v_i \in X$. Here $m = 0$ is allowed; the expression $(*)$ is then taken to be the 0 vector. We say that X is *linearly independent* iff 0 cannot be written as a linear combination as in $(*)$ with some $a_i \neq 0$. The *span* of a subset $X \subseteq V$ is the collection of all vectors which can be written as linear combinations of members of X, including 0 as a linear combination. A *basis* for V is a linearly independent set which spans the whole space.

The only result we assume from linear algebra is that if X and Y are both finite bases for V, then $|X| = |Y|$.

Theorem 10.3. *If V is a vector space over a field F, then V has a basis.*

Proof. We are going to apply Zorn's lemma.

Let $A = \{X \subseteq V : X \text{ is linearly independent}\}$, partially ordered by $\subseteq$. Then $A \neq \emptyset$, since trivially $\emptyset \in A$. Now suppose that B is a subset of A simply ordered by $\subseteq$. We claim that $\bigcup B \in A$; this will verify the hypothesis of Zorn's lemma. Suppose that $v_1, \ldots, v_n \in \bigcup B$, $a_1, \ldots, a_n \in F$, and $a_1v_1 + \cdots + a_nv_n = 0$; we want to show that all a_i are 0. For each $i = 1, \ldots, n$ choose $X_i \in B$ such that $v_i \in X_i$. Now $\{X_i : i = 1, \ldots, n\}$ has a largest member X_j under $\subseteq$, since B is simply ordered. Here we are using 6.34. Clearly $v_i \in X_j$ for all $i = 1, \ldots, n$. Since X_j is linearly independent, it follows that each $a_i = 0$, as desired.

Now we apply Zorn's lemma to obtain a maximal member Y of $\mathscr{A}$ under $\subseteq$. We claim that Y is a basis for A. Since Y is linearly independent, it suffices to show that Y spans A. Suppose that $w \in A$. If $w \in Y$, then obviously w is in the span of Y. Suppose that $w \notin Y$. Then $Y \subset Y \cup \{w\}$ so by the maximality of Y, $Y \subset Y \cup \{w\}$ is linearly dependent. Hence there is a natural number n, elements $v_1, \ldots, v_n \in Y \cup \{w\}$, and elements $a_1, \ldots, a_n \in F$, not all 0, such that $a_1v_1 + \cdots + a_nv_n = 0$. Since Y is linearly independent, not all v_i are in Y; say that $v_j = w$. Then again because Y is linearly independent, we must have $a_j \neq 0$. So

$$w = \left(-\frac{a_1}{a_j}v_1\right) + \cdots + \left(-\frac{a_{j-1}}{a_j}v_{j-1}\right) + \left(-\frac{a_{j+_o1}}{a_j}v_{j+_o1}\right) + \cdots + \left(-\frac{a_n}{a_j}v_n\right),$$

so that w is in the span of Y, as desired. □

Theorem 10.4. *Let V be a vector space over a field F. Then any two bases for V over F have the same number of elements.*

Proof. Let X and Y be bases for V over F. If both of them are finite, then $|X| = |Y|$ by the assumed result above. Suppose now that at least one of them is infinite, but they have different cardinalities. Say that $|X| < |Y| \geq \omega$. Now each $v \in X$ is a linear combination of members of Y. So we can write

$$v = a_{1,v}w_{1,v} + \cdots + a_{m_v,v}w_{m_v,v} \tag{$*$}$$

with each $a_{i,v} \in F$ and nonzero, and each $w_{i,v} \in Y$. Since the $a_{i,v}$'s and $w_{i,v}$'s are uniquely determined by v, the axiom of choice is not involved here. Now let $W = \{w_{i,v} : v \in X, i = 1, \ldots, m_v\}$. Then W spans V. In fact, given $x \in V$, we can write

$$x = b_1 v_1 + \cdots + b_n v_n$$

with each $b_i \in F$ and each $v_i \in X$, and then by $(*)$ it is clear that x is a linear combination of members of W. However, $|W| < |Y|$. In fact, if X is finite, then so is W, and hence $|W| < |Y|$. If X is infinite, then $|W| = |X| < |Y|$ by exercise 15 in Chapter 8. Thus, indeed, $|W| < |Y|$. Choose $y \in Y \backslash W$. Then y is a linear combination of members of W since W spans V, and this contradicts Y being linearly independent. □

Algebraically closed fields. As mentioned above, a field is a structure like the real numbers—a nonempty set together with operations of addition and multiplication satisfying conditions similar to familiar ones for the reals. Another important example of a field is the field $\mathbb{Q}$ of rational numbers.

An extension G of F is *algebraic over* F iff every element of G is the zero of some polynomial with coefficients from F. A field G is said to be *algebraically closed* iff it has a subfield F such that G is algebraic over F, and any polynomial with coefficients in F splits into linear factors over G. There are various equivalent definitions of this notion which we do not need to enter into. We assume these results about fields:

(1) If $f(x)$ is a non-constant polynomial with coefficients in a field F, then F has an algebraic extension G in which $f(x)$ splits into linear factors, and such that $|G| \le |F| + \omega$.

(2) If K is an algebraic extension of G and G is an algebraic extension of F, then K is an algebraic extension of F.

(3) If k is an isomorphism from F onto G, then k extends to an isomorphism of $F[x]$ onto $G[x]$, namely,

$$\begin{aligned} &a_0 + a_1 x + \cdots + a_m x^m \quad \text{maps to} \\ &k(a_0) + k(a_1)x + \cdots + k(a_m)x^m \end{aligned}$$

for any $a_0, \ldots, a_m \in F$.

Theorem 10.5. *For any field F there is an algebraically closed field G which extends F and has size at most $|F| + \omega$. Moreover, every element of G is algebraic over F.*

Proof. Let $\kappa = |F| + \omega$, and let λ be a cardinal number greater than κ; this cardinal exists by 7.31. By induction, $|F^m| \le \kappa$ for all $m \in \omega$. We can map $\langle a_0, \ldots, a_m\rangle$ to $a_0 + a_1 x + \cdots + a_m x^m$, thus mapping F^m onto the set of all polynomials of degree at most m. So there are at most κ polymomials of degree at most m. Hence the total number of polynomials is at most $\omega \cdot \kappa = \kappa$. Let $\langle f_\alpha(x) : \alpha < \kappa\rangle$ list all polynomials with coefficients in F, possibly with repetitions.

Let g be a bijection from a subset G_0 of λ onto F. We can make G_0 into a field such that g is an isomorphism from G_0 onto F by the following definition. For any $a, b \in G_0$,

$$\begin{aligned} a + b &= g^{-1}(g(a) + g(b)) \quad \text{and} \\ a \cdot b &= f^{-1}(g(a) \cdot g(b)). \end{aligned}$$

Let $\alpha < \kappa$, and let

$$f_\alpha(x) \text{ be } a_{\alpha,0} + a_{\alpha,1}x + a_{\alpha,2}x^2 + \cdots + a_{\alpha,m_\alpha}x^{m_\alpha}.$$

We then let

$$f'_\alpha(x) \text{ be } g^{-1}(a_{\alpha,0}) + g^{-1}(a_{\alpha,1})x + g^{-1}(a_{\alpha,2})x^2 + \cdots + g^{-1}(a_{\alpha,m_\alpha})x^{m_\alpha}.$$

So $f'_\alpha(x)$ is a polynomial with coefficients in G_0.

Next, we define $\mathscr{F}$ to be the collection of all fields $(H, +_H, \cdot_H)$ such that $H \subseteq \lambda$, and we let c be a choice function for nonempty members of $\mathscr{P}(\mathscr{F})$; that is, for any nonempty $X \subseteq \mathscr{F}$, $c(X) \in X$.

We make a construction of fields G_α for $\alpha \leq \kappa$ by transfinite recursion as follows, starting with G_0. Suppose that G_α has been defined as an extension of G_0, with $|G_\alpha| \leq \kappa$. By the result (1) on field extensions which we are assuming, there is a field $(K, +_K, \cdot_K)$ which is an algebraic extension of G_α and in which $f'_\alpha(x)$ splits into linear factors, with $|K| \leq \kappa$. Let

$$\begin{aligned} X = \{(H, +_H, \cdot_H) : H \subseteq \lambda, H &\text{ is an algebraic extension of } G_\alpha, \\ |H| \leq \kappa, \text{ and } f'_\alpha(x) &\text{ splits into liinear factors in } H\}. \end{aligned}$$

Let k be a one-one function mapping K into λ, with k the identity on $K \cap \lambda$; this is possible because $|K| < \lambda$: for then we have $|K \cap \lambda| < \lambda$ and so $|\lambda \backslash K| = \lambda$, and we can map $K \backslash \lambda$ one-one into $\lambda \backslash K$. Let H be the range of k. Then we make H into a field by defining, for any $a, b \in K$,

$$\begin{aligned} a + b &= k(k^{-1}(a) + k^{-1}(b)) \quad \text{and} \\ a \cdot b &= k(k^{-1}(a) \cdot k^{-1}(b)). \end{aligned}$$

Then k is an isomorphism of K onto H. The coefficients of $f'_\alpha(x)$ are in $G_\alpha \subseteq \lambda$, and so they are fixed by k. Now by result (3), k extends to an isomorphism of $K[x]$ onto $H[x]$. It follows that $f'_\alpha(x)$ splits into linear factors over H too. Moreover, H is an algebraic extension of G_α. In fact, let $b \in H$. Then $k^{-1}(b) \in K$. Since K is an algebraic extension of G_α, there is a polynomial $a_0 + a_1x + \cdots + a_mx^m$ with coefficients $a_i \in G_\alpha$ such that $a_0 + a_1k^{-1}(b) + \cdots + a_m(k^{-1}(b))^m = 0$. Since $G_\alpha \subseteq \lambda$ and k is the identity on $K \cap \lambda$, we have $k(a_i) = a_i$ for all i, and hence

$$0 = k(0) = k(a_0 + a_1k^{-1}(b) + \cdots + a_m(k^{-1}(b))^m) = a_0 + a_1b + \cdots + a_mb^m.$$

This shows that b is algebraic over G_α.

We have now shown that $X \neq \emptyset$. We let $G_{\alpha+_o1} = c(X)$.

At limit steps $\alpha \leq \lambda$ we define $G_\alpha = \bigcup_{\beta<\alpha} G_\beta$. Note that then

$$|G_\alpha| \leq \sum_{\beta<\alpha} |G_\beta| \leq \sum_{\beta<\alpha} \lambda \leq \lambda \cdot \lambda = \lambda.$$

Now let l be a one-one function mapping G_κ onto a superset L of F, extending g. We make L into a field by setting, for any $a, b \in L$,

$$\begin{aligned} a + b &= l(l^{-1}(a) + l^{-1}(b)) \quad \text{and} \\ a \cdot b &= l(l^{-1}(a) \cdot l^{-1}(b)). \end{aligned}$$

We claim that L is as desired in the theorem. Since l is a bijection, L has size at most $\lambda = |F| + \omega$. We have $F \subseteq L$. If $a, b \in F$, then

$$\begin{aligned} a +_L b &= l(l^{-1}(a) +_{G_0} l^{-1}(b)) \\ &= l(g^{-1}(a) +_{G_0} g^{-1}(b)) \\ &= l(g^{-1}(g(g^{-1}(a)) +_F g(g^{-1}(b))) \\ &= l(g^{-1}(a +_F b)) \\ &= g(g^{-1}(a +_F b)) \\ &= a +_F b. \end{aligned}$$

Similarly, $a \cdot_L b = a \cdot_F b$. Hence L extends F.

For any $\alpha < \kappa$, the polynomial $f'_\alpha(x)$ splits into linear factors over G_κ; thus

$$\begin{aligned} f'_\alpha(x) \text{ is } & g^{-1}(a_{\alpha,0}) + g^{-1}(a_{\alpha,1})x + g^{-1}(a_{\alpha,2})x^2 + \cdots + g^{-1}(a_{\alpha,m_\alpha})x^{m_\alpha} \\ \text{which is } & (b_0 x - c_0)(b_1 x - c_1) \ldots (b_{m_\alpha} x - c_{m_\alpha}) \end{aligned}$$

for some b_i's and c_i's in G_κ. By (3), l extends to an isomorphism from G_κ onto L. Clearly $l(f'_\alpha(x)) = f_\alpha(x)$, and so we see that $f_\alpha(x)$ splits into linear factors over L, namely, $f_\alpha(x)$ is

$$(l(b_1)x - l(c_1)(l(b_2)x - l(c_2)) \cdots (l(b_{m_\alpha})x - l(c_{m_\alpha})).$$

Finally, we show that L is an algebraic extension of F. To do this, we first prove that each G_α is an algebraic extension of G_0, by transfinite induction. This is obviously true for $\alpha = 0$. If we have shown that G_α is an algebraic extension of G_0, then $G_{\alpha +_o 1}$ is also, by (2). If α is limit $\leq \kappa$ and we know that each G_β for $\beta < \alpha$ is an algebraic extension of G_0, then each element of G_α is in some G_β for $\beta < \alpha$, and hence is the zero of some polynomial with coefficients in G_0, as desired. This finishes the inductive proof. Hence G_κ is an algebraic extension of G_0. Now take any $b \in L$. Then $l^{-1}(b)$ is a zero of some polynomial $g(x)$ with coefficients in G_0, so it is clear that b is a zero of the polynomial $l(g(x))$ which has coefficients in F; a detailed argument in a similar situation was given above. □

It is of some interest to recognize some statements which go beyond the axioms of ZF, are implied by the axiom of choice, but are demonstrably weaker than AC. Here is a partial list of such statements, taken from T. Jech, **The axiom of choice**:

Boolean prime ideal theorem. *Every Boolean algebra has a prime ideal.*

Ordering principle. *Every set can be simply ordered.*

Axiom of choice for families of well-orderable sets. *If $\langle A_i : i \in I\rangle$ is a system of non-empty sets each of which can be well-ordered, then there is a function f with domain I such that $f(i) \in A_i$ for all $i \in I$.*

Axiom of choice for families of two-element sets. *If $\langle A_i : i \in I\rangle$ is a system of non-empty sets, and for each $i \in I$ there exist distinct x, y such that $A_i = \{x, y\}$, then there is a function f with domain I such that $f(i) \in A_i$ for all $i \in I$.*

Axiom of choice for families of finite sets. *If $\langle A_i : i \in I\rangle$ is a system of non-empty finite sets, then there is a function f with domain I such that $f(i) \in A_i$ for all $i \in I$.*

Countable axiom of choice. *If $\langle A_i : i \in \omega\rangle$ is a system of non-empty sets, then there is a function f with domain ω such that $f(i) \in A_i$ for all $i \in \omega$.*

Exercises, Chapter 10

1. Show carefully how the function h defined in the proof of 10.2, Axiom of Choice $\Rightarrow$ Well-ordering principle, is shown to exist by the transfinite recursion principle.

2. Prove that the relation $<$ on A, defined in the proof that the axiom of choice implies the well-ordering principle, really does well-order A.

3. Show carefully how the function h defined in the proof of 10.2, Well-ordering principle $\Rightarrow$ Zorn's lemma, is shown to exist by the transfinite recursion principle.

4. Show by induction on m, without using the axiom of choice, that if $m \in \omega$ and $\langle A_i : i \in m\rangle$ is a system of nonempty sets, then there is a function f with domain m such that $f(i) \in A_i$ for all $i \in m$.

5. Using AC, prove the following, which is called the *Principle of Dependent Choice* (which is also weaker than the axiom of choice, but cannot be proved in ZF). If R is a relation, $R \subseteq A \times A$, and for every $a \in A$ there is a $b \in A$ such that aRb, then there is a function $f : \omega \to A$ such that $f(i)Rf(i+1)$ for all $i \in \omega$.

The next exercises give some more equivalents to the axiom of choice; so each exercise states something provable in ZF. Consider the following statements.

(1) If $<$ is a partial ordering and $\prec$ is a simple ordering which is a subset of $<$, then there is a maximal (under $\subseteq$) simple ordering $\ll$ such that $\prec$ is a subset of $\ll$, which in turn is a subset of $<$.

(2) For any two sets A and B, either there is a one-one function mapping A into B or there is a one-one function mapping B into A.

(3) For any two nonempty sets A and B, either there is a function mapping A onto B or there is a function mapping B onto A.

(4) A family $\mathscr{F}$ of subsets of a set A has *finite character* if for all $X \subseteq A$, $X \in \mathscr{F}$ iff every finite subset of X is in $\mathscr{F}$. Principle (4) says that every family of finite character has a maximal element under $\subseteq$.

(5) For any relation R there is a function $f \subseteq R$ such that $\operatorname{dmn} R = \operatorname{dmn} f$.

6. Show that the axiom of choice implies (1). [Use Zorn's lemma]

7. Prove that (1) implies (2). [Given sets A and B, define $f < g$ iff f and g are one-one functions which are subsets of $A \times B$, and $f \subset g$. Apply (1) to $<$ and the empty simple ordering.]

8. Prove that (2) implies (3). [Easy]

9. Prove that (3) implies the axiom of choice. [Show that any set A can be well-ordered, as follows. Use Lemma 10.1 to find an ordinal which cannot be mapped one-one into $\mathscr{P}(A)$. Show that if $f : A \to \alpha$ maps onto α, then $\langle f^{-1}[\{\beta\}] : \beta < \alpha \rangle$ is a one-one function from α into $\mathscr{P}(A)$.

10. Show that the axiom of choice implies (4). [Use Zorn's lemma.]

11. Show that (4) implies (5). [Given a relation R, let $\mathscr{F}$ consist of all functions contained in R.]

12. Show that (5) implies the axiom of choice. [Given a family $\langle A_i : i \in I \rangle$ of nonempty sets, let $R = \{(i, x) : i \in I \text{ and } x \in A_i\}$.]