Introduction to Set Theory
Third Edition, Revised and Expanded
by Karel Hrbacek and Thomas Jech

Theorems Document

by Dan Whitman

September 17, 2019

Copyright

Chapter 6 Ordinal Numbers

§6.4   Transfinite Induction and Recursion

Theorem 6.4.5

Let $\mathbf{G}$ be an operation; then the property $\mathbf{P}(x, y)$ defined by

1. x is an ordinal number and $y = t(x)$ for some computation of length x (with respect to $\mathbf{G}$), or

2. x is not an ordinal and $y = \varnothing$.

defines an operation $\mathbf{F}$ such that $\mathbf{F}(\alpha) = \mathbf{G}(\mathbf{F} \restriction \alpha)$ for all ordinals α.

Note that, for ordinal α, we say that t is a computation of length α if t is a transfinite sequence whose domain is $\alpha + 1$ and such that, for all $\beta \leq \alpha$, we have $t(\beta) = \mathbf{G}(t \restriction \beta)$.

Solution:

Proof. Suppose that $\mathbf{G}$ is an operation, which is to say that for each x there is unique y such that $y = \mathbf{G}(x)$.

Now we define the property $\mathbf{P}(x, y)$ as described above. We prove that $\mathbf{P}$ defines an operation. Hence we have to show that for any x there is a unique y such that $\mathbf{P}(x, y)$ holds. So consider any set α. If α is not an ordinal than clearly $\mathbf{P}(\alpha, \varnothing)$ holds and $\varnothing$ is unique. So suppose that α is an ordinal. Then it suffices to show that there is a unique computation of length α (with respect to $\mathbf{G}$) since this will make $y = t(\alpha)$ unique. We show this via transfinite induction.

So assume that for all $\beta < \alpha$ that there is a unique computation of length β and we show that there exists a unique computation of length α, which completes the proof that $\mathbf{P}$ defines an operation.

Existence. First define a property $\mathbf{R}(x, y)$ such that $\mathbf{R}(x, y)$ holds if and only if

1. x is an ordinal where $x < \alpha$ and y is a computation of length x (with respect to $\mathbf{G}$), or

2. x is is an ordinal and $x \geq \alpha$ and $y = \varnothing$, or

3. x is not an ordinal and $y = \varnothing$.

Clearly by the induction hypothesis this property has a unique y for every x. Hence we can apply the Axiom Schema of Replacement, according to which there is a set T such that for every $\beta \in \alpha$ (so that $\beta < \alpha$) there is a t in T such that $\mathbf{R}(\beta, t)$ holds. That is

$$T = \{t \mid t \text{ is the unique computation of length } \beta \text{ for all } \beta < \alpha\}$$

Now, T is a system of transfinite sequences (which are functions) so define $\bar{t} = \bigcup T$ and let $\tau = \bar{t} \cup \{(\alpha, \mathbf{G}(\bar{t}))\}$.

Claim 1: $\mathrm{dom}\,(\tau) = \alpha + 1$. So consider any $\beta \in \mathrm{dom}\,(\tau)$. Clearly if $\beta = \alpha$ then $\beta \in \alpha + 1$. On the other hand if $\beta \in \mathrm{dom}\,(\bar{t})$ then there is a $t \in T$ such that $\beta \in \mathrm{dom}\,(t)$. But since t is a computation of length γ and $\gamma < \alpha$ it follows that $\beta \leq \gamma < \alpha < \alpha + 1$ so that $\beta \in \alpha + 1$. Hence since β was arbitrary $\mathrm{dom}\,(\tau) \subseteq \alpha + 1$.

Now consider any $\beta \in \alpha + 1$ so that $\beta \leq \alpha$. If $\beta = \alpha$ then clearly by definition $\beta \in \mathrm{dom}\,(\tau)$. On the other hand if $\beta \neq \alpha$ then $\beta < \alpha$. So consider the $t \in T$ where t is the unique computation of length β (which exists since $\beta < \alpha$). Then clearly $\beta \in \mathrm{dom}\,(t)$ so that $\beta \in \mathrm{dom}\,(\bar{t})$. From this it follows that clearly $\beta \in \mathrm{dom}\,(\tau)$ so that $\alpha + 1 \subseteq \mathrm{dom}\,(\tau)$ since β was arbitrary. This proves the claim.

Claim 2: τ is a function. Consider any $\beta \in \text{dom}\,(\tau) = \alpha + 1$ so that again $\beta \leq \alpha$. If $\beta = \alpha$ then clearly $\tau(\beta) = \tau(\alpha) = \mathbf{G}(\bar{t})$ is unique since $\mathbf{G}$ is an operation. On the other hand if $\beta < \alpha$ then τ is a function so long as $\bar{t}$ is, and this is the case so long as T is a compatible system of functions since $\bar{t} = \bigcup T$. We show this presently.

So consider any arbitrary $t_1, t_2 \in T$ where t_1 is the computation of length β_1 and t_2 is the computation of length β_2. Without loss of generality we can assume that $\beta_1 \leq \beta_2$. We must show that $t_1(\gamma) = t_2(\gamma)$ for all $\gamma \leq \beta_1$. This we show by transfinite induction. So suppose that $t_1(\delta) = t_2(\delta)$ for all $\delta < \gamma \leq \beta_1$. Then clearly $t_1 \restriction \gamma = t_2 \restriction \gamma$ and since $\mathbf{G}$ is an operation we have $t_1(\gamma) = \mathbf{G}(t_1 \restriction \gamma) = \mathbf{G}(t_2 \restriction \gamma) = t_2(\gamma)$. This completes the proof of the claim.

Claim 3: $\tau(\beta) = \mathbf{G}(\tau \restriction \beta)$ for all $\beta \leq \alpha$. So consider any $\beta \leq \alpha$. If $\beta = \alpha$ then then since $\bar{t} = \tau \restriction \alpha$ we clearly have $\tau(\beta) = \tau(\alpha) = \mathbf{G}(\bar{t}) = \mathbf{G}(\tau \restriction \alpha) = \mathbf{G}(\tau \restriction \beta)$ by the definition of τ. On the other hand if $\beta < \alpha$ then let $t \in T$ be the computation of length β (which exists since $\beta < \alpha$). Then $\tau(\beta) = t(\beta) = \mathbf{G}(t \restriction \beta) = \mathbf{G}(\tau \restriction \beta)$ since t is a computation (with respect to $\mathbf{G}$) and clearly $t \subseteq \tau$.

Claims 1 through 3 show that τ is a computation of length α and hence that such a computation exists.

Uniqueness. Now let σ be another computation of length α. We show that $\sigma = \tau$, which proves uniqueness. Since both σ and τ are functions with $\text{dom}\,(\sigma) = \alpha + 1 = \text{dom}\,(\tau)$ it suffices to show that $\sigma(\gamma) = \tau(\gamma)$ for all $\gamma \leq \alpha$. We show this once again by using transfinite induction. So suppose that $\sigma(\delta) = \tau(\delta)$ for all $\delta < \gamma \leq \alpha$. It then follows that $\sigma \restriction \gamma = \tau \restriction \gamma$. Then since σ and τ are computations we have that $\sigma(\gamma) = \mathbf{G}(\sigma \restriction \gamma) = \mathbf{G}(\tau \restriction \gamma) = \tau(\gamma)$, thereby completing the uniqueness proof.

This completes the proof that $\mathbf{P}$ defines an operation.

So let $\mathbf{F}$ be the operation defined by $\mathbf{P}$. The last thing we need to show to complete the proof of the entire theorem is that $\mathbf{F}(\alpha) = \mathbf{G}(\mathbf{F} \restriction \alpha)$ for all ordinals α, noting that we are treating $\mathbf{F}$ as a function even though it is an operation. Thus, for any set X, $\mathbf{F} \restriction X$ denotes the set $\{(x, \mathbf{F}(x)) \mid x \in X\}$, which forms a function with domain X. The range of this function is a set whose existence is guaranteed by the Axiom Schema of Replacement since $\mathbf{F}$ is an operation.

So consider any ordinal α and a computation t of length α. Then clearly for any $\beta \in \alpha$ we have that $t_\beta = t \restriction (\beta + 1)$ is a computation of length β. Since this computation is the unique computation of length β, by the definition of $\mathbf{F}$ as it relates to $\mathbf{P}$ we have $\mathbf{F}(\beta) = t_\beta(\beta) = t(\beta)$. Since β was arbitrary this shows that $\mathbf{F} \restriction \alpha = t \restriction \alpha$.

So for any ordinal α let t be the unique computation of length α. Then clearly we have $\mathbf{F}(\alpha) = t(\alpha) = \mathbf{G}(t \restriction \alpha) = \mathbf{G}(\mathbf{F} \restriction \alpha)$ by what was just shown above. $\qquad\square$

§6.5 Ordinal Arithmetic

Theorem 6.5.3

Let $(W_1, <_1)$ and $(W_2, <_2)$ be well-ordered sets, isomorphic to ordinals α_1 and α_2, respectively, and let $(W, <)$ be the sum of $(W_1, <_1)$ and $(W_2, <_2)$. Then $(W, <)$ is isomorphic to $\alpha_1 + \alpha_2$.

Solution:

Lemma 6.5.3.1. *Any initial segment of an ordinal is also an ordinal.*

Proof. Suppose that α is an ordinal and that β is an initial segment of α.

First we show that β is transitive. So consider any $x \in \beta$ and any $y \in x$. Since clearly $\beta \subseteq \alpha$ we have that $x \in \alpha$ so that x is an ordinal by Lemma 7.2.8. But then since $y \in x$ we have that y is an ordinal too by the same lemma. Hence we have $y < x$ so that $y \in \beta$ since $x \in \beta$ and β is an initial segment. Hence since $y \in x$ was arbitrary it follows that $x \subseteq \beta$ so that β is transitive since x was arbitrary.

Now consider any subset X of β. Then since $\beta \subseteq \alpha$ clearly $X \subseteq \alpha$ as well so that X has a least element since α is well-ordered (since it is an ordinal). Hence since X was arbitrary β is also well-ordered.

This shows by definition that β is an ordinal. $\qquad\square$

Lemma 6.5.3.2. *Suppose that $(W, \prec)$ is a well ordered set isomorphic to ordinal α and $W[a]$ is an initial segment of W isomorphic to ordinal β. If f is the isomorphism from α to W and g the isomorphism from β to $W[a]$ then:*

1. $\beta < \alpha$

2. $f \restriction \beta = g$

3. $f(\beta) = a$

Proof. Consider $h = f^{-1} \circ g : \beta \to \alpha$, which is clearly an isomorphism since f and g are. We also claim that h maps β onto an initial segment of α, i.e. $h[\beta]$ is an initial segment of α. So consider any $x \in h[\beta]$, noting that $x \in \alpha$ since clearly $h[\beta] \subseteq \alpha$. Thus x is an ordinal by Lemma 6.2.8. So consider any ordinal $y < x$. Now since $x \in h[\beta]$ there is a $\gamma \in \beta$ such that $h(\gamma) = x$. Hence we have

$$f^{-1}(g(\gamma)) = x$$
$$g(\gamma) = f(x)$$

and $f(x) = g(\gamma) \in W[a]$ since $\gamma \in \beta$. Also since $y < x$ and f is an isomorphism we have $f(y) \prec f(x)$ so that $f(y) \in W[a]$ since $W[a]$ is an initial segment and $f(x) \in W[a]$. Then since g is onto there is a $\delta \in \beta$ such that

$$g(\delta) = f(y)$$
$$f^{-1}(g(\delta)) = y$$
$$h(\delta) = y \,,$$

which shows that $y \in h[\beta]$. Since $y < x$ was arbitrary this shows that $h[\beta]$ is an initial segment of α by definition.

Thus, since h is an isomorphism, β is isomorphic to an initial segment of α. By Lemma 6.5.3.1 this an initial segment is itself an ordinal. So it has to be that $h[\beta] = \beta$ since if it were any other ordinal then we would have that β is isomorphic to a different ordinal, which is not possible. Then since β is an initial segment of α we have that $\beta \subset \alpha$ so that $\beta < \alpha$ by Lemma 6.2.9, which shows (1).

Moreover since h is an isomorphism from β onto itself it follows from Corollary 6.1.5 part (b) that h is the identity function. Hence for any $x \in \beta$ we have

$$h(x) = f^{-1}(g(x)) = x$$
$$g(x) = f(x)$$

so that clearly (2) follows since x was arbitrary.

Now, let $W' = W - W[a]$, noting that this is not empty since $W[a] \subset W$. Consider any $x \in W'$ so that since $x \notin W[a]$ but $x \in W$ we have that $x \succeq a$. Since x was arbitrary and clearly $a \in W'$ this shows that a is the least element of W'.

Now let $X = \alpha - \beta$, which is also nonempty since $\beta \subset \alpha$. Consider any $x \in X$ so that $x \in \alpha$ but $x \notin \beta$. Thus $x \geq \beta$ so that β is the least element of X since since x was arbitrary and clearly $\beta \in X$.

Now we claim that $f \restriction X$ is an isomorphism from X to W'. So consider any $x \in X$ so that $x \notin \beta$. Suppose that $f(x) \in W[a]$. Then there is a $\gamma \in \beta$ such that $g(\gamma) = f(x)$. But also since $f \restriction \beta = g$ it follows that $f(\gamma) = g(\gamma) = f(x)$. Then since f is injective it follows that $\gamma = x$. But this is a contradiction since we have that $x \notin \beta$ but $\gamma \in \beta$, so it has to be that $f(x) \notin W[a]$. Hence since clearly $f(x) \in W$ we have that $f(x) \in W'$.

Also for any x and y in X where $x < y$ we clearly have $(f \restriction X)(x) = f(x) < f(y) = (f \restriction X)(y)$ since f is an isomorphism. Thus $f \restriction X$ is an isomorphism and therefore also injective. We need only show that $f \restriction X$ is surjective as well. So consider any $b \in W'$ so that $b \in W$ but $b \notin W[a]$. Since f is surjective there is a $\gamma \in \alpha$ such that $f(\gamma) = b$. Suppose that $\gamma \in \beta$ so that since $f \restriction \beta = g$ we have $g(\gamma) = f(\gamma) = b \in W[a]$, a contradiction. Hence $\gamma \notin \beta$ so that $\gamma \in X$. Since we have also shown that $f(\gamma) = b$ and b was arbitrary this shows that $f \restriction X$ is surjective.

Thus we have proven that $f \restriction X$ is an isomorphism between X and W'. Then since β is the least element of X and a is the least element of W' it follows that $f(\beta) = (f \restriction X)(\beta) = a$, which shows (3) and thus the entire lemma. $\square$

Main Theorem.

Proof. We assume that W_1 and W_2 are disjoint and that $(W, <)$ is their sum as defined in Lemma 4.4.5. We prove the result by transfinite induction on α_2, in particular the version of induction in 6.4.3.

So for $\alpha_2 = 0$ we have that $W_2 = \varnothing$ so $W = W_1 \cup \varnothing = W_1$, which is clearly isomorphic to $\alpha_1 = \alpha_1 + 0 = \alpha_1 + \alpha_2$.

Now suppose that $\alpha_2 = \beta + 1$ for some ordinal β so that W_2 has a greatest element a. Noting that a is clearly also the greatest element of W, also assume then that $W[a]$ is isomorphic to $\alpha_1 + \beta$. Then we have clearly that W is isomorphic to $(\alpha_1 + \beta) + 1$ (i.e. the successor of $\alpha_1 + \beta$). But then by Definition 6.5.1b we have $(\alpha_1 + \beta) + 1 = \alpha_1 + (\beta + 1) = \alpha_1 + \alpha_2$.

Lastly, suppose that α_2 is a limit ordinal and that for each $\beta < \alpha_2$ there is an isomorphism f_β from $\alpha_1 + \beta$ to $W[a_\beta]$ where $a_\beta \in W_2$. We note that each f_β is unique. Then let $f = \bigcup_{\beta < \alpha_2} f_\beta$.

First we claim that f is a function, which is the case so long as $\{f_\beta \mid \beta < \alpha_2\}$ is a compatible system of functions. So consider any $\beta_1 < \alpha_2$ and $\beta_2 < \alpha_2$. Without loss of generality we can assume that $\beta_1 \leq \beta_2$. Clearly if $\beta_1 = \beta_2$ then $f_{\beta_1} = f_{\beta_2}$ since the function is unique so suppose that $\beta_1 < \beta_2$. Then by Lemma 6.5.4a $\alpha_1 + \beta_1 < \alpha_1 + \beta_2$. From this it follows that $W[a_{\beta_1}]$ is an initial segment of $W[a_{\beta_2}]$ From Lemma 6.5.3.2 part (2) above it follows that $f_{\beta_2} \restriction (\alpha_1 + \beta_1) = f_{\beta_1}$. Since β_1 and β_2 were arbitrary this shows that the system is compatible and thus that f is a function.

We also show that $\operatorname{dom} f = \alpha_1 + \alpha_2$. So first consider any $\gamma \in \alpha_1 + \alpha_2$ so that we have $\gamma = \alpha_1 + \beta$ for some $\beta < \alpha_2$ and hence that $\gamma \in (\alpha_1 + \beta) + 1 = \alpha_1 + (\beta + 1)$. Thus $\gamma \in \operatorname{dom} f_{\beta+1}$ so that clearly also $\gamma \in \operatorname{dom} f$, noting that clearly $\beta + 1 < \alpha_2$ as well since α_2 is a limit ordinal. Hence since γ was arbitrary $\alpha_1 + \alpha_2 \subseteq \operatorname{dom} f$. Now suppose that $\gamma \in \operatorname{dom} f$ so that there is a $\beta < \alpha_2$ such that $\gamma \in \operatorname{dom} f_\beta$. It then follows by the definition of f_β that $\gamma \in \alpha_1 + \beta$ for some $\beta < \alpha_2$. Then clearly $\alpha_1 + \beta < \alpha_1 + \alpha_2$ so that $\alpha_1 + \beta \in \alpha_1 + \alpha_2$ and $\alpha_1 + \beta \subseteq \alpha_1 + \alpha_2$ since $\alpha_1 + \alpha_2$ is transitive (since it is an ordinal). Thus also $\gamma \in \alpha_1 + \alpha_2$. Hence $\operatorname{dom} f \subseteq \alpha_1 + \alpha_2$ since γ was arbitrary. This shows that $\operatorname{dom} f = \alpha_1 + \alpha_2$.

We also show that $\operatorname{ran} f = W$, which also shows that f is onto. So consider any $b \in \operatorname{ran} f$. Then there is a $\beta < \alpha_2$ where $b \in \operatorname{ran} f_\beta$. Hence $b \in W[a_\beta]$ for some $a_\beta \in W$. Since this is an initial segment of W clearly also $b \in W$. Thus $\operatorname{ran} f \subseteq W$ since b was arbitrary. Now consider any $b \in W$. If $b \in W_1$ then let g be the isomorphism from α_1 to W_1 so that there is a $\gamma \in \alpha_1$ such that $g(\gamma) = b$. Clearly also $\gamma \in a_1 + 1$ and since W_1 is an initial segment of W by Lemma 6.5.3.2, part 2 we have that $f_1 \restriction \alpha_1 = g$ so that $f_1(\gamma) = b$ as well. Thus $b \in \operatorname{ran} f_1$ so that clearly $b \in \operatorname{ran} f$ also. On

the other hand if $b \in W_2$ then let g denote the isomorphism from α_2 to W_2 so that there is a $\gamma \in \alpha_2$ such that $g(\gamma) = b$, which exists since g is onto. Since α_2 is a limit ordinal we have that $\gamma + 1 \in \alpha_2$ so that $\gamma + 1 < \alpha_2$. Then $f_{\gamma+1}$ is an isomorphism from $\alpha_1 + (\gamma + 1)$ to $W[a_{\gamma+1}]$. Thus $f_{\gamma+1}(\alpha_1 + \gamma) = g(\gamma) = b \in W[a_{\gamma+1}]$ since $\alpha_1 + \gamma \in (\alpha_1 + \gamma) + 1 = \alpha_1 + (\gamma + 1)$. Hence $b \in \operatorname{ran} f_{\gamma+1}$ so that also $b \in \operatorname{ran} f$. Hence since b was arbitrary and $b \in \operatorname{ran} f$ in all cases it follows that $W \subseteq \operatorname{ran} f$, which shows that $\operatorname{ran} f = W$ as desired.

Finally we show that f is an isomorphism. So consider any γ_1 and γ_2 in $\alpha_1 + \alpha_2$ where $\gamma_1 < \gamma_2$. Then we have that $\gamma_2 = \alpha_1 + \beta$ for some ordinal β and that $\gamma_1 < \gamma_2 < \gamma_2 + 1 = (\alpha_1 + \beta) + 1 = \alpha_1 + (\beta + 1)$. From this it follows that γ_1 and γ_2 are both in $\operatorname{dom} f_{\beta+1} = \gamma_2 + 1$. Thus since $f_{\beta+1}$ is an isomorphism we clearly have that

$$f(\gamma_1) = f_{\beta+1}(\gamma_1) < f_{\beta+1}(\gamma_2) = f(\gamma_2)$$

so that f is an isomorphism as well since γ_1 and γ_2 were arbitrary. Thus $\alpha_1 + \alpha_2$ is isomorphic to W in this case as well, which proves the entire theorem. $\qquad \square$

§6.6 The Normal Form

Theorem 6.6.1

The ordinal functions $\alpha + \beta$, $\alpha \cdot \beta$, and α^β are *continuous* in the second variable: If γ is a limit ordinal and $\beta = \sup_{\nu < \gamma} \beta_\nu$, then

$$\alpha + \beta = \sup_{\nu < \gamma} (\alpha + \beta_\nu), \qquad \alpha \cdot \beta = \sup_{\nu < \gamma} (\alpha \cdot \beta_\nu), \qquad \alpha^\beta = \sup_{\nu < \gamma} \alpha^{\beta_\nu}$$

Solution:

Lemma 6.6.1.1. *Suppose that A is a nonempty set of ordinals where α is the greatest element of A and that f is a non-decreasing function on A whose range is a set of ordinals. Then*

$$f(\alpha) = \sup_{\beta \in A} f(\beta).$$

Proof. Consider any $\beta \in A$ so that $\beta \leq \alpha$ since α is the greatest element of A. It then follows that $f(\beta) \leq f(\alpha)$ since f is non-decreasing. Since β was arbitrary this shows that $f(\alpha)$ is the greatest element of the set $f(A) = \{f(\gamma) \mid \gamma \in A\}$. Therefore by the remarks after the proof of Theorem 6.2.6 in the text it follows that $f(\alpha) = \sup f(A)$ as desired. $\qquad \square$

Lemma 6.6.1.2. *If A is a set of ordinals then ordinal $\alpha = \sup A$ if and only if α is the least upper bound of A, i.e. α is an upper bound of A and β is not an upper bound of A for any $\beta < \alpha$.*

Proof. ($\rightarrow$) First suppose that $\alpha = \sup A$. Then by the remarks following the proof of Theorem 6.2.6 in the text α is an upper bound of A and if β is an upper bound of A then $\alpha \leq \beta$. This last statement is simply the contrapositive of the statement that $\beta < \alpha$ implies that β is *not* an upper bound of A and hence is logically equivalent.

($\leftarrow$) We show that an ordinal α with the least upper bound property for A is unique, which suffices to show the result since if β has this property then $\beta = \sup A$ since $\sup A$ does as well (by what was just shown above) and the ordinal having this property is unique.

So suppose that ordinals α and β both have the least upper bound property for A but that $\alpha \neq \beta$. Without loss of generality we can assume then that $\alpha < \beta$. But then, since β has the least upper

bound property, α cannot be an upper bound of A, which contradicts the fact that α also has the least upper bound property! Hence it has to be that $\alpha = \beta$, which shows the uniquness. $\square$

Lemma 6.6.1.3. *Suppose that A and B are sets of ordinals and that f is a non-decreasing function defined on $A \cup B$ whose range is a set of ordinals. Also suppose that neither A nor B have greatest elements and that $\sup A = \sup B$. Then*

$$\sup_{\alpha \in A} f(\alpha) = \sup_{\beta \in B} f(\beta).$$

Proof. First we denote $f(A) = \{f(\alpha) \mid \alpha \in A\}$ and $f(B) = \{f(\beta) \mid \beta \in B\}$ so that we must show that $\sup f(A) = \sup f(B)$. We show this by proving that $\sup f(A)$ has the least upper bound property of the set $f(B)$, from which the result follows from Lemma 6.6.1.2.

So first consider any $f(\beta)$ in $f(B)$ so that $\beta \in B$. Then $\beta \leq \sup B = \sup A$. Moreover it has to be that $\beta < \sup B = \sup A$ since B has no greatest element (since if $\beta = \sup B$ then it would be the greatest element of B). Thus β is not an upper bound of A so that there is an $\alpha \in A$ such that $\beta < \alpha$. Then, since f is non-decreasing, we have $f(\beta) \leq f(\alpha) \leq \sup f(A)$. Hence $\sup f(A)$ is an upper bound of $f(B)$ since $f(\beta)$ was arbitrary.

Now consider any $\gamma < \sup f(A)$. Then γ is not an upper bound of $f(A)$ so that there is an $\alpha \in A$ such that $\gamma < f(\alpha)$. Since A has no greatest element it follows that $\alpha < \sup A = \sup B$ so that it is not an upper bound of B. Hence there is a $\beta \in B$ such that $\alpha < \beta$. It then follows that $f(\alpha) \leq f(\beta)$ since f is non-decreasing. Thus $\gamma < f(\alpha) \leq f(\beta)$. This shows that γ is not an upper bound of $f(B)$ and thus the result since $g < \sup f(A)$ was arbitrary. $\square$

Main Theorem.

Suppose that γ is a nonzero limit ordinal, $\{\beta_\nu\}$ is a transfinite sequence of ordinals where $\nu < \gamma$, and $\beta = \sup_{\nu < \gamma} \beta_\nu$.

For what follows let $A = \{\beta_\nu \mid \nu < \gamma\}$, i.e. the range of the sequence, so that clearly then $\beta = \sup_{\nu < \gamma} \beta_\nu = \sup A$. Note also that A is nonempty since the domain of the sequence is a nonzero ordinal.

First we show that $\alpha + \beta = \sup_{\nu < \gamma} (\alpha + \beta_\nu)$ for any ordinal α.

Proof. First note that the function $f(\delta) = \alpha + \delta$ defined for any $\delta \leq \beta$ is non-decreasing by Lemma 6.5.4. With this notation we aim to show that $\alpha + \beta = \sup_{\delta \in A} f(\delta)$.

In the first case suppose that $\beta \in A$ so that clearly β is the greatest element of A (since it is an upper bound of A). The result then follows from Lemma 6.6.1.1.

On the other hand suppose that $\beta \notin A$. It then follows that A has no greatest element and then by Exercise 6.2.7 that β is a limit ordinal. Note that $\beta \neq 0$ since if that were the case then $A = \varnothing$, which is not possible since it is the range of a sequence indexed by a nonzero ordinal. Hence we have by Definition 6.5.1c that

$$\alpha + \beta = \sup_{\delta < \beta} (\alpha + \delta) = \sup_{\delta < \beta} f(\delta).$$

The result then follows from Lemma 6.6.1.3 since f is non-decreasing and by the remarks after Theorem 6.2.10 we have that

$$\sup_{\delta < \beta} \delta = \beta = \sup A$$

since β is a limit ordinal. $\square$

The proof that $\alpha \cdot \beta = \sup_{\nu < \gamma} (\alpha \cdot \beta_\nu)$ is analogous, noting that the function $f(\delta) = \alpha \cdot \delta$ is non-decreasing by Exercise 6.5.7 (even clearly when $\alpha = 0$ since non-decreasing is not strict like it would be if we required that the function be increasing).

Similarly, the proof for the exponential function, i.e. that $\alpha^\beta = \sup_{\nu < \gamma} \alpha^{\beta_\nu}$, is also analogous. However this is only valid for $\alpha > 1$, which is not mentioned in the text. This is because ordinal exponentiation is only non-decreasing for $\alpha > 0$ since $0^1 = 0 < 1 = 0^0$. That it is non-decreasing follows trivially from what is shown in Exercise 6.5.14b. The following is a counterexample that shows that the assertion does not hold for $\alpha = 0$, even though exponentiation not being non-decreasing does not necessarily mean it is false (it just means our proof does not work for $\alpha = 0$).

Suppose that our transfinite sequence is indexed by $\gamma = \omega$ (i.e. it is an ordinary sequence) and that the sequence is $\{\beta_\nu\} = \{1, 0, 0, 0, \ldots\}$. We then clearly have that

$$\beta = \sup_{\nu < \gamma} \beta_\nu = \sup_{\nu < \omega} \beta_\nu = \sup \{1, 0, 0, 0, \ldots\} = 1$$

so that $\alpha^\beta = 0^1 = 0$. However, we also have

$$\sup_{\nu < \gamma} \alpha^{\beta_\nu} = \sup_{\nu < \omega} 0^{\beta_\nu} = \sup \{0^1, 0^0, 0^0, 0^0, \ldots\} = \sup \{0, 1, 1, 1, \ldots\} = 1.$$

Hence the result does not hold since $\alpha^\beta = 0 \neq 1 = \sup_{\nu < \gamma} \alpha^{\beta_\nu}$.

Chapter 7 Alephs

§7.1 Initial Ordinals

Theorem 7.1.7

The Hartogs number of A exists for all A.

Solution:

Proof. For every well-ordered set (W, R) where $W \subseteq A$ there is a unique ordinal α such that $(\alpha, <)$ is isomorphic to (W, R) by Theorem 6.3.1. Then by the Axiom Schema of Replacement there is a set H such that $\alpha \in H$ if and only if it is the ordinal isomorphic to a well-ordering $R \subseteq \mathcal{P}(A \times A)$ of some $W \subseteq A$.

We claim that H contains all ordinals that are equipotent to some subset of A. So consider any such ordinal α and let f be the bijection from α to $W \subseteq A$. Then let

$$R = \{(f(\beta), f(\gamma)) \mid (\beta, \gamma) \in \alpha \times \alpha \wedge \beta < \gamma\}$$

so that clearly f is an isomorphism from $(\alpha, <)$ to (W, R). Hence by definition $\alpha \in H$ as desired.

Thus the set

$$h(A) = \{\alpha \in H \mid \alpha \text{ is equipotent to some subset of } A\}$$

exists by the Axiom Schema of Comprehension. We claim that this is the Hartogs numbers. To justify this we need to show that it is an ordinal and that it is the least ordinal which is not equipotent to a subset of A.

First since H was defined to be a set of ordinals and $h(A) \subseteq H$ it follows that it also is a set of ordinals and thus well-ordered by Theorem 6.2.6d. Now consider any $\alpha \in h(A)$ so that $\alpha \in H$ so that there is a $W \subseteq A$ and well ordering $R \subseteq \mathcal{P}(A \times A)$ such that $(\alpha, <)$ is isomorphic to (W, R). Let f be that isomorphism. Now consider any $\beta \in \alpha$ so that $\beta \subseteq \alpha$ since α is an ordinal (and therefore transitive). Then let $g = f \restriction \beta$ and $X = f[\beta]$. Then clearly g is an isomorphism from $(\beta, <)$ to (X, R) so that $\beta \in H$. Moreover since g is bijective and $X \subseteq W \subseteq A$ clearly β is equipotent to a subset of A. Hence $\beta \in h(A)$. Since $\beta \in \alpha$ was arbitrary this shows that $\alpha \subseteq h(A)$ so that $h(A)$ is transitive. Hence $h(A)$ satisfies the definition of an ordinal number.

Now suppose that $h(A)$ itself is equipotent to some subset of A. Then $h(A) \in H$ since H was shown to contain all such ordinals. Then also by definition $h(A) \in h(A)$, which contradicts Lemma 6.2.7. So it must be that $h(A)$ is in fact not equipotent to any subset of A. Clearly also for any $\alpha < h(A)$ we have $\alpha \in h(A)$ since $h(A)$ is an ordinal. But then by definition α is equipotent to a subset of A. This shows that $h(A)$ is the least ordinal that is not equipotent to a subset of A and hence is the Hartogs number. $\square$

Theorem 7.1.8

We show the remark after Definition 7.1.8 that $|\omega_\alpha| < |\omega_\beta|$ for ordinals α and β where $\alpha < \beta$.

Solution:

Lemma 7.1.8.1. *For well ordered sets A and B either $|A| \leq |B|$ or $|B| \leq |A|$ (or both in which case $|A| = |B|$).*

Proof. By Theorem 6.1.3 we have:

Case: A and B are isomorphic. Let f be the isomorphism from A to B. Then clearly f is a bijection so that $|A| = |B|$. Also since f is injective $|A| \leq |B|$. Clearly also f^{-1} is bijective from B to A so that $|B| \leq |A|$ as well.

Case: A is isomorphic to an initial segment of B. Then if f is the isomorphism clearly f is an injective function from A to B so that $|A| \leq |B|$.

Case: B is isomorphic to an initial segment of A. Then if f is the isomorphism clearly f is an injective function from B to A so that $|B| \leq |A|$.

Since these cases are exhaustive by Theorem 6.1.3 clearly the result has been shown.

Note that this did not require the Axiom of Choice. $\square$

Corollary 7.1.8.2. *If A and B are well ordered sets then $|A| \nleq |B|$ if and only if $|B| < |A|$.*

Proof. ($\rightarrow$) Suppose that $|A| \nleq |B|$. Then it follows from Lemma 7.1.8.1 above that $|B| \leq |A|$. Suppose that $|B| = |A|$. Then there is a bijection f from B to A. But then clearly f^{-1} is also a bijection and therefore injective. Hence by definition $|A| \leq |B|$, a contradiction. So it cannot be that $|B| = |A|$. Hence $|B| < |A|$ by definition as desired.

($\leftarrow$) We show this by proving the contrapositive. So suppose that $|A| \leq |B|$. Also suppose that $|B| \leq |A|$ so that by Lemma 7.1.8.1 above $|A| = |B|$. Thus we have shown that

$$|B| \leq |A| \rightarrow |A| = |B|$$
$$|B| \nleq |A| \vee |A| = |B|$$
$$\neg(|B| \leq |A| \wedge |A| \neq |B|)$$
$$\neg(|B| < |A|),$$

thereby showing the contrapositive. $\square$

Main Theorem.

Proof. We show this by induction. First for α we have by definition that $\omega_{\alpha+1} = h(\omega_\alpha)$. Since it has been shown that $|h(\omega_\alpha)| \not\leq |\omega_\alpha|$ it follows from Corollary 7.1.8.2 that $|\omega_{\alpha+1}| = |h(\omega_\alpha)| > |\omega_\alpha|$ since ω_α and $h(\omega_\alpha)$ are ordinals and therefore well-ordered. Hence $|\omega_\alpha| < |\omega_{\alpha+1}|$.

Now suppose that α is a nonzero limit ordinal and consider any $\gamma < \alpha$. Then $\gamma + 1 < \alpha$ since α is a limit ordinal. Then since ω_α is an upper bound of $\{\omega_\beta \mid \beta < \alpha\}$ it follows that $\omega_{\gamma+1} \leq \omega_\alpha$ since $\gamma + 1 < \alpha$. From this it clearly follows that $|\omega_{\gamma+1}| \leq |\omega_\alpha|$. We also have by what was just shown that $|\omega_\gamma| < |\omega_{\gamma+1}|$. Hence by Exercise 4.1.2a we have $|\omega_\gamma| < |\omega_\alpha|$. Since $\gamma < \alpha$ was arbitrary this completes the induction and the result follows. $\quad\square$

Theorem 7.1.9

We show part (b): If Ω is an infinite initial ordinal number, then $\Omega = \omega_\alpha$ for some α.

Solution:

Lemma 7.1.9.1. *If α is an infinite ordinal then $|\alpha| = |\alpha + 1|$, i.e. α and $\alpha + 1$ are equipotent.*

Proof. First we note that since α is infinite we have $\alpha + 1 > \alpha \geq \omega$. We then construct a bijection from $\alpha + 1$ to α. So define $f : \alpha + 1 \to \alpha$ by

$$
f(\beta) = \begin{cases} \beta + 1 & \beta < \omega \\ \beta & \beta \geq \omega \text{ and } \beta \neq \alpha \\ 0 & \beta = \alpha \end{cases}
$$

for $\beta \in \alpha + 1$.

First we show that f is injective. So consider any β and γ in $\alpha + 1$ where $\beta \neq \gamma$. Without loss of generality we can assume that $\beta < \gamma$. We then have the following:

Case: $\beta < \omega$. Then clearly $f(\beta) = \beta + 1 < \omega$ since $\beta < \omega$ and ω is a limit ordinal, but we also clearly have that $0 < \beta + 1 = f(\beta)$. Now, if also $\gamma < \omega$ then clearly $f(\beta) = \beta + 1 < \gamma + 1 = f(\gamma)$ since $\beta < \gamma$. If $\gamma \geq \omega$ and $\gamma \neq \alpha$ then we have $f(\beta) < \omega \leq \gamma = f(\gamma)$. Lastly if $\gamma = \alpha$ then we have $f(\gamma) = 0 < f(\beta)$.

Case: $\beta \geq \omega$ and $\beta \neq \alpha$. Here since $\beta < \gamma$ we have $\omega \leq \beta < \gamma$. Thus if also $\gamma \neq \alpha$ then clearly we have $f(\beta) = \beta < \gamma = f(\gamma)$. On the other hand if $\gamma = \alpha$ then $f(\gamma) = 0 < \omega \leq \beta = f(\beta)$.

Thus in every case we have $f(\beta) \neq f(\gamma)$, thereby showing that f is injective. We note that the case in which $\beta = \alpha$ is impossible since α is the greatest element of $\alpha + 1$ but $\gamma > \beta$ and $\gamma \in \alpha + 1$.

Next we show that f is surjective. So consider any $\beta \in \alpha$.

Case: $\beta < \omega$. If $\beta = 0$ then clearly $f(\alpha) = 0 = \beta$. On the other hand if $0 < \beta < \omega$ then β is a successor ordinal, say $\beta = \gamma + 1$, so that $\gamma < \beta < \omega$ hence clearly $\gamma \in \alpha + 1$ and $f(\gamma) = \gamma + 1 = \beta$.

Case: $\beta \geq \omega$. Then since $\beta \in \alpha$ we have $\beta < \alpha < \alpha + 1$ so that $\beta \neq \alpha$ but $\beta \in \alpha + 1$. Then clearly $f(\beta) = \beta$.

Hence in all cases there is a $\gamma \in \alpha + 1$ such that $f(\gamma) = \beta$ so that f is injective. Therefore we have shown that f is a bijection so that by definition $\alpha + 1$ and α are equipotent. $\quad\square$

Lemma 7.1.9.2. *Every infinite initial ordinal is a limit ordinal.*

Proof. Suppose that α is an infinite initial ordinal and that it a successor so that $\alpha = \beta + 1$. It was shown in Lemma 7.1.9.1 that $|\beta| = |\beta + 1| = |\alpha|$, but since clearly $\beta < \alpha$ this contradicts the fact that α is an initial ordinal. Hence α must be a limit ordinal. $\square$

Lemma 7.1.9.3. *Suppose that γ is a nonzero limit ordinal and $\{\alpha_\nu\}$ and $\{\beta_\nu\}$ for $\nu < \gamma$ are two transfinite sequences. Also suppose that $\alpha_\nu \leq \beta_\nu$ for every $\nu < \gamma$. Then*

$$\sup_{\nu < \gamma} \alpha_\nu \leq \sup_{\nu < \gamma} \beta_\nu \,.$$

Proof. First let $A = \{\alpha_\nu \mid \nu < \gamma\}$ and $B = \{\beta_\nu \mid \nu < \gamma\}$ be the ranges of the sequences so that we must show that $\sup A \leq \sup B$. Now consider any $\alpha \in A$ so that $\alpha = \alpha_\nu$ for some $\nu < \gamma$. We then have that

$$\alpha = \alpha_\nu \leq \beta_\nu \leq \sup B$$

so that $\sup B$ is an upper bound of A since α was arbitrary. It then follows from the least upper bound property that $\sup A \leq \sup B$ as desired. $\square$

Lemma 7.1.9.4. $\alpha \leq \omega_\alpha$ *for every ordinal α.*

Proof. We show this by induction on α. Clearly for $\alpha = 0$ we have $\alpha = 0 \leq \omega = \omega_0 = \omega_\alpha$. Now suppose that $\alpha \leq \alpha_\omega$. Then by Exercise 6.5.8 we have that $\alpha + 1 \leq \omega_\alpha + 1$. Then, since $h(\omega_\alpha)$ is an initial ordinal by Lemma 7.1.6 and clearly is infinite, it follows that $h(\omega_\alpha)$ is a limit ordinal by Lemma 7.1.9.2. Since clearly $\omega_\alpha < h(\omega_\alpha)$ we then have that $\omega_\alpha + 1 < h(\omega_\alpha)$ as well. Hence we have

$$\alpha + 1 \leq \omega_\alpha + 1 < h(\omega_\alpha) = \omega_{\alpha+1}$$

so that $\alpha + 1 \leq \omega_{\alpha+1}$ is true.

Now suppose that α is a nonzero limit ordinal and that $\gamma \leq \omega_\gamma$ for all $\gamma < \alpha$. Since α is a limit ordinal it follows that $\alpha = \sup\{\gamma \mid \gamma < \alpha\}$ and by definition we have that $\omega_\alpha = \sup\{\omega_\gamma \mid \gamma < \alpha\}$. It then follows from these and the induction hypothesis that

$$\alpha = \sup\{\gamma \mid \gamma < \alpha\} \leq \sup\{\omega_\gamma \mid \gamma < \alpha\} = \omega_\alpha$$

by Lemma 7.1.9.3. This completes the inductive proof. $\square$

Lemma 7.1.9.5. *For initial ordinals α and β, if $|\alpha| \leq |\beta|$, then $\alpha \leq \beta$.*

Proof. Suppose that $|\alpha| \leq |\beta|$ but that $\alpha > \beta$. Then clearly β is isomorphic (and therefore equipotent) to an initial segment of α so that $|\beta| \leq |\alpha|$. Then by the Cantor-Bernstein Theorem we have $|\alpha| = |\beta|$. However since α is an initial ordinal and $\beta < \alpha$ it cannot be that $|\alpha| = |\beta|$. Thus we have a contradiction so that it must be that $\alpha \leq \beta$ as desired. $\square$

Lemma 7.1.9.6. *For any ordinal α and any infinite initial ordinal Ω where $\Omega < \omega_\alpha$, there is a $\gamma < \alpha$ such that $\Omega = \omega_\gamma$.*

Proof. We show this by induction on α. For $\alpha = 0$ we have $\omega_\alpha = \omega_0 = \omega$ so that there is no infinite initial ordinal Ω such that $\Omega < \omega_\alpha = \omega$. Hence the hypothesis is vacuously true. Now suppose that, for every infinite initial ordinal $\Omega < \omega_\alpha$, there is a $\gamma < \alpha$ such that $\Omega = \omega_\gamma$. Consider any infinite initial ordinal $\Omega < \omega_{\alpha+1}$. Then $\Omega < \omega_{\alpha+1} = h(\omega_\alpha)$ so that Ω is equipotent to some subset of ω_α by the definition of the Hartogs number. From this it clearly follows that $|\Omega| \leq |\omega_\alpha|$ and hence $\Omega \leq \omega_\alpha$ by Lemma 7.1.9.5 since both Ω and ω_α are initial ordinals. If $\Omega = \omega_\alpha$ then we are finished but if $\Omega < \omega_\alpha$ then by the induction hypothesis there is a $\gamma < \alpha$ such that $\Omega = \omega_\gamma$ so that we are also finished.

Now suppose that α is a nonzero limit ordinal and that for every $\beta < \alpha$ and infinite initial ordinal $\Omega < \omega_\beta$ there is a $\gamma < \beta$ such that $\Omega = \omega_\gamma$. Consider then any infinite initial ordinal $\Omega < \omega_\alpha$. Then since $\omega_\alpha = \sup \{\omega_\beta \mid \beta < \alpha\}$ it follows that Ω is not an upper bound of $\{\omega_\beta \mid \beta < \alpha\}$ so that there is a $\beta < \alpha$ such that $\Omega < \omega_\beta$. But then by the induction hypothesis there is a $\gamma < \beta$ such that $\Omega = \omega_\gamma$. This completes the transfinite induction. $\qquad\square$

Main Theorem.

Proof. Let Ω be any infinite initial ordinal. Then by Lemma 7.1.9.4 we have

$$\Omega < \Omega + 1 \le \omega_{\Omega+1} \, .$$

Hence there is an α such that $\Omega < \omega_\alpha$. It then follows from Lemma 7.1.9.6 that there is a $\gamma < \alpha$ such that $\Omega = \omega_\gamma$. $\qquad\square$

Chapter 8   The Axiom of Choice

§8.1   The Axiom of Choice and its Equivalents

Theorem 8.1.13

The following statements are equivalent:

(a) (The Axiom of Choice) There exists a choice function for every system of sets.

(b) (The Well-Ordering Principle) Every set can be well-ordered.

(c) (Zorn's Lemma) If every chain in a partially ordered set has an upper bound, then the partially ordered set has a maximal element.

Solution:

Lemma 8.1.13.1. *For any set A, there is a $b \notin A$.*

Proof. Let $X = \{\alpha \in A \mid \alpha$ is an ordinal number$\}$. Then by Theorem 6.2.6e there is an ordinal α such that $\alpha \notin X$. It also has to be that $\alpha \notin A$ since, if it were, then α would be in X since it is an ordinal number, which would be a contradiction. $\qquad\square$

Main Theorem.

Proof. The equivalence of (a) and (b) follows immediately from Theorem 8.1.1; therefore it suffices to show the equivalence of (a) and (c).

(a) implies (c). Suppose that $(A \preccurlyeq)$ is a partially ordered set in which every chain has an upper bound. Then by Lemma 8.1.13.1 there is a $b \notin A$, and by (a) there is a choice function g on $\mathcal{P}(A)$. We then define a transfinite sequence $\langle a_\alpha \mid \alpha < h(A) \rangle$ by transfinite recursion. Given the transfinite sequence $\langle a_\xi \mid \xi < \alpha \rangle$ for an $\alpha < h(A)$ we define the set $A_\alpha = \{a \in A \mid a_\xi \prec a$ for all $\xi < \alpha\}$. We then set

$$a_\alpha = \begin{cases} g(A_\alpha) & \text{if } a_\xi \ne b \text{ for all } \xi < \alpha \text{ and } A_\alpha \ne \varnothing \\ b & \text{otherwise} \, , \end{cases}$$

noting that clearly A_α is in the domain of g since $A_\alpha \subseteq A$.

We can justify the recursive definition of this sequence as follows. First, let $B = A \cup \{b\}$ and $S = \bigcup_{\alpha < h(A)} B^\alpha$ be the set of all transfinite sequences of elements B of length less than $h(A)$. We then define a function $G : S \to B$. For any transfinite sequence $F \in S$ there is an $\alpha < h(A)$ such that $F \in B^\alpha$, and we let $A_F = \{a \in A \mid F(\xi) \prec a \text{ for all } \xi < \alpha\}$. We then simply set

$$G(F) = \begin{cases} g(A_F) & \text{if } F(\xi) \neq b \text{ for all } \xi < \alpha \text{ and } A_F \neq \varnothing \\ b & \text{otherwise} . \end{cases}$$

It then follows from Theorem 6.4.4 that there is a function f such that $f(\alpha) = G(f \restriction \alpha)$ for all $\alpha < h(A)$. Clearly based on our definition of G this f is exactly the transfinite sequence constructed above.

Returning to our main proof, we claim that there is an $\alpha < h(A)$ such that $a_\alpha = b$. To see this, suppose to the contrary that $a_\alpha \neq b$ for all $\alpha < h(A)$ so that it has to be that each $a_\alpha \in A$. Consider now any $\alpha < h(A)$ and $\beta < h(A)$ where $\alpha \neq \beta$. Without loss of generality we can assume that $\alpha < \beta$. Clearly then, by definition, we have that $a_\beta \in A_\beta$ so that $a_\xi \prec a_\beta$ for all $\xi < \beta$. But since $\alpha < \beta$, we have that $a_\alpha \prec a_\beta$ so that $a_\alpha \neq a_\beta$. Since α and β were arbitrary, this shows that the sequence is an injective function from $h(A)$ to A. However, this would mean that $h(A)$ is equipotent to some subset of A, which contradicts the definition of the Hartogs number. Hence it has to be that $a_\alpha = b$ for some $\alpha < h(A)$.

So let $\lambda < h(A)$ be the least ordinal such that $a_\lambda = b$ and let $C = \{a_\xi \mid \xi < \lambda\}$. We claim that C is a chain in $(A, \preccurlyeq)$. So consider any a_α and a_β in C so that $\alpha < \lambda$ and $\beta < \lambda$ Without loss of generality we can assume that $\alpha \leq \beta$. If $\alpha = \beta$ then obviously $a_\alpha = a_\beta$ so that $a_a \preccurlyeq a_\beta$ clearly holds. If $\alpha < \beta$ then, by what was shown above, we have that $a_\alpha \prec a_\beta$ so that $a_\alpha \preccurlyeq a_\beta$ again holds. Hence, in every case, a_α and a_β are comparable in $\preccurlyeq$, which shows that C is a chain since a_α and a_β were arbitrary.

Thus, since C is a chain, it has an upper bound $c \in A$. We claim that c is also a maximal element of A. To show this, suppose that there is an $a \in A$ such that $c \prec a$. Now consider any $\xi < \lambda$. Then, since c is an upper bound of C, we have that $a_\xi \preccurlyeq c \prec a$ so that $a_\xi \prec a$ since orders are transitive. It then follows from the definition of A_λ that $a \in A_\lambda$ so that $A_\lambda \neq \varnothing$. Also note that, by the definition of λ, we have that $a_\xi \neq b$ for any $\xi < \lambda$. Thus, by the recursive definition of the sequence, it follows that $a_\lambda = g(A_\lambda) \neq b$, which contradicts the definition of λ (as the least ordinal such that $a_\lambda = b$). So it has to be that there is no such element a, which shows that c is in fact a maximal element of A.

(c) implies (a). Suppose that T is any system of sets. Then, supposing that $\varnothing \in T$, $S = T - \{\varnothing\}$ is a system of nonempty sets. If we can show that S has a choice function g then $f = g \cup \{(\varnothing, \varnothing)\}$ is clearly a choice function for T. Thus it suffices to show that any system S of nonempty sets has a choice function.

So let S be such a system. Let F be the system of functions f for which $\operatorname{dom} f \subseteq S$ and $f(X) \in X$ holds for any $X \in \operatorname{dom} f$. Note that $F \in \mathcal{P}(\mathcal{P}(S \times \bigcup S))$, which exists by the Axiom of Power Set and the Axiom of Union. We claim that F is ordered by $\subseteq$. Hence we must show that $\subseteq$ is reflexive, asymmetric, and transitive. Since, for any $f \in F$, we obviously have that $f \subseteq f$, clearly $\subseteq$ is reflexive. Also, clearly if we have $f \subseteq g$ and $g \subseteq f$ then $f = g$ so that $\subseteq$ is asymmetric. Lastly, if $f \subseteq g$ and $g \subseteq h$, then obviously $f \subseteq h$ as well so that $\subseteq$ is transitive. Hence $\subseteq$ is a (partial) order on F.

Now suppose that F_0 is a linearly ordered subset of F. We claim that F_0 is a compatible set of functions. Consider any f and g in F_0. Since $\subseteq$ is a linear ordering of F_0 we have that either $f \subseteq g$ or $g \subseteq f$. Without loss of generality, we can assume that $f \subseteq g$. Then clearly $f \cup g = g$. Since this is a function, it follows from Lemma 2.3.11a that f and g are compatible. Since f and g were arbitrary, this shows that F_0 is a compatible system. It then follows from Theorem 2.3.12 that $f_0 = \bigcup F_0$ is a function and that it extends every $g \in F_0$ and hence is compatible with them all.

We also claim that $f_0 \in F$. First consider any $X \in \mathrm{dom}\, f_0$ so that it has to be that $X \in \mathrm{dom}\, g$ for some $g \in F_0$. Then, since also clearly $g \in F$ (since $F_0 \subseteq F$), it follows that $g(X) \in X$. Thus also $f_0(X) = g(X) \in X$ since f_0 and g are compatible. We also have that $\mathrm{dom}\, g \subseteq S$ (since $g \in F$) so that $X \in S$. Since X was arbitrary this shows that $\mathrm{dom}\, f_0 \subseteq S$ and hence that $f_0 \in F$. Lastly, clearly f_0 is an upper bound of F_0 in $(F, \subseteq)$ since $g \subseteq \bigcup F_0 = f_0$ for any $g \in F_0$.

Therefore the conditions of Zorn's Lemma are satisfied so that F_0 has a maximal element $\overline{f}$. We claim that $\mathrm{dom}\, \overline{f} = S$. If this were not the case then it would have to be that $\mathrm{dom}\, \overline{f} \subset S$ since we know that $\mathrm{dom}\, \overline{f} \subseteq S$ since $\overline{f} \in F_0 \subseteq F$. Thus there would be an $X \in S$ where $X \notin \mathrm{dom}\, \overline{f}$. Consider choose $x \in X$ so that clearly the function $\overline{\overline{f}} = \overline{f} \cup \{(X, x)\} \in F$. However, we would also clearly have that $\overline{f} \subseteq \overline{\overline{f}}$ but $\overline{f} \neq \overline{\overline{f}}$, which would contradict the maximality of $\overline{f}$. So it has to be that $\mathrm{dom}\, \overline{f} = S$. Clearly then $\overline{f}$ is a choice function for S since, for any $X \in S$, we have that $\overline{f}(X) \in X$ since $\overline{f} \in F$. $\qquad \square$

§8.2 The Axiom of Choice in Mathematics

Theorem 8.2.10*

Let p be a sublinear functional on the vector space V, and let f_0 be a linear functional defined on a subspace V_0 of V such that $f_0(v) \leq p(v)$ for all $v \in V_0$. Then there is a linear functional f defined on V such that $f \supseteq f_0$ and $f(v) \leq p(v)$ for all $v \in V$.

Solution:

Lemma 8.2.10.1. *If $\mathcal{F}$ is a nonempty system of subspaces of a vector space V over scalar field K that is linearly ordered by $\subseteq$, then $\bigcup \mathcal{F}$ is also a subspace of V.*

Proof. First, since $\mathcal{F}$ is nonempty, there is a subspace $W \in \mathcal{F}$. Since W is a subspace, it follows that the zero vector 0 is in W. Hence clearly $0 \in \bigcup \mathcal{F}$ since $W \in \mathcal{F}$.

Now consider any vectors v_1 and v_2 in $\bigcup \mathcal{F}$. Then there are subspaces (of V) V_1 and V_2 in $\mathcal{F}$ such that $v_1 \in V_1$ and $v_2 \in V_2$. Since $\subseteq$ linearly orders $\mathcal{F}$ we have that either $V_1 \subseteq V_2$ or $V_2 \subseteq V_1$. Without loss of generality we can assume that $V_1 \subseteq V_2$ since the other case is analogous. Thus we have that both v_1 and v_2 are in V_2 since $v_1 \in V_1 \subseteq V_2$. Since V_2 is a subspace it follows that $v_1 + v_2 \in V_2$ so that clearly also $v_1 + v_2 \in \mathcal{F}$ since $V_2 \in \mathcal{F}$.

Lastly, consider a vector $v \in \bigcup \mathcal{F}$ and scalar $c \in K$. Then there is a subspace $W \in \mathcal{F}$ such that $v \in W$. It then follows that $cv \in W$ so that clearly $cv \in \bigcup \mathcal{F}$.

This shows that $\bigcup \mathcal{F}$ is a subspace of V as desired. $\qquad \square$

Main Theorem.

Proof. Let F be the set of all linear functionals g defined on some subspace W of V such that

$$f_0 \subseteq g \text{ and } g(v) \leq p(v) \text{ for all } v \in W.$$

Let $F_0 \subseteq F$ be linearly ordered by $\subseteq$. If $F_0 = \varnothing$ then f_0 is an $\subseteq$-upper bound of F_0 since clearly $f_0 \in F$ (with subspace V_0). So assume that F_0 is nonempty so that $g_0 = \bigcup F_0$ is clearly an $\subseteq$-upper bound of F_0. It is trivial to show that F_0 is a compatible system of functions since it is linearly ordered by $\subseteq$. Hence $g_0 = \bigcup F_0$ is a function with $\mathrm{dom}\,(g_0) = \bigcup_{g \in F_0} \mathrm{dom}\,(g)$ by Theorem 2.3.12 where g_0 extends all $g \in F_0$.

We also show that $\{\mathrm{dom}\,(g) \mid g \in F_0\}$ is a set of subspaces of V that is linearly ordered by $\subseteq$. Consider any g and h in F_0. Then either $g \subseteq h$ or $h \subseteq g$. Without loss of generality we can assume that the former is the case ($g \subseteq h$) since the other case is analogous. Now consider any $v \in \mathrm{dom}\,(g)$ so that clearly $(v, g(v)) \in g$. Then we have $(v, g(v)) \in h$ since $g \subseteq h$. Thus $v \in \mathrm{dom}\,(h)$ as well so that $\mathrm{dom}\,(g) \subseteq \mathrm{dom}\,(h)$ since v was arbitrary. The result then follows since g and h were arbitrary.

Having shown this, it follows from Lemma 8.2.10.1 that $\mathrm{dom}\,(g_0) = \bigcup \{\mathrm{dom}\,(g) \mid g \in F_0\}$ is also a subspace of V. Now we show that g_0 is a linear functional. So consider any u and v in $\mathrm{dom}\,(g_0)$ and α and β in $\boldsymbol{R}$. Then there is a g and g' in F_0 where $u \in \mathrm{dom}\,(g)$ and $v \in \mathrm{dom}\,(g')$. Since we showed that $\{\mathrm{dom}\,(g) \mid g \in F_0\}$ is linearly ordered by $\subseteq$, we have that either $\mathrm{dom}\,(g) \subseteq \mathrm{dom}\,(g')$ or $g' \subseteq g$. If $\mathrm{dom}\,(g) \subseteq \mathrm{dom}\,(g)'$ then both u and v are in $\mathrm{dom}\,(g')$ since $u \in \mathrm{dom}\,(g) \subseteq \mathrm{dom}\,(g')$. Thus $\alpha \cdot u + \beta \cdot v \in \mathrm{dom}\,(g')$ since g' is a linear functional since $g' \in F_0 \subseteq F$. Hence clearly $\alpha \cdot u + \beta \cdot v \in \mathrm{dom}\,(g_0)$ since $g' \in F_0$. We then also have that

$$g_0(\alpha \cdot u + \beta \cdot v) = g'(\alpha \cdot u + \beta \cdot v) = \alpha \cdot g'(u) + \beta \cdot g'(v) = \alpha \cdot g_0(u) + \beta \cdot g_0(v)$$

since g_0 extends g'. The case in which $\mathrm{dom}\,(g') \subseteq \mathrm{dom}\,(g)$ is analogous, thereby showing that g_0 is a linear functional on the subspace $\mathrm{dom}\,(g_0)$.

We also clearly have that $f_0 \subseteq g_0$. To see this, note that, since $F_0 \neq \varnothing$, there is a $g \in F_0$ so that $g \in F$ since $F_0 \subseteq F$. Therefore $f_0 \subseteq g$ by the definition of F. We also clearly have that have that $g \subseteq g_0$ since $g_0 = \bigcup F_0$ and $g \in F_0$. Therefore $f_0 \subseteq g \subseteq g_0$ as desired.

Now we show that $g_0(v) \leq p(v)$ for all $v \in \mathrm{dom}\,(g_0)$. So consider any such v so that there is a $g \in F_0$ such that $v \in \mathrm{dom}\,(g)$ since it was shown above that $\mathrm{dom}\,(g_0) = \bigcup_{g \in F_0} \mathrm{dom}\,(g)$. Thus also $g \in F$ since $F_0 \subseteq F$. It then follows that $g(v) \leq p(v)$ by the definition of F. However, since g_0 extends g, we have that $g_0(v) = g(v) \leq p(v)$. This shows the result since v was arbitrary.

We have therefore shown all of the facts required to conclude that indeed $g_0 \in F$ by the definition of F. Since g_0 is an upper bound of F_0 and F_0 was arbitrary, this shows that the conditions of Zorn's Lemma have been satisfied so that F has a $\subseteq$-maximal element f. We claim that f is our linear functional on V such that $f_0 \subseteq f$ and $f(v) \leq p(v)$ for all $v \in V$.

Clearly all we must show is that $\mathrm{dom}\,(f) = V$ since the rest is true by virtue of f being in F. We know that $\mathrm{dom}\,(f)$ is a subspace of V since $f \in F$ so that $\mathrm{dom}\,(f) \subseteq V$. Assume to the contrary that $\mathrm{dom}\,(f) \neq V$ so that $\mathrm{dom}\,(f) \subset V$ and there is a $u \in V - \mathrm{dom}\,(f)$. Then let W be the subspace of V spanned by $\mathrm{dom}\,(f) \cup \{u\}$. It follows then that every vector $w \in W$ can be uniquely expressed as $w = x + \alpha \cdot u$ for $x \in \mathrm{dom}\,(f)$ and $\alpha \in \boldsymbol{R}$. So define the function

$$f_c(w) = f(x) + \alpha \cdot c$$

for each $w \in W$ (where $w = x + \alpha \cdot u$) and a $c \in \boldsymbol{R}$.

First we show that f_c is a linear functional on W for any choice of c. Clearly $f_c : W \to \boldsymbol{R}$ since $\mathrm{ran}\,(f) \subseteq \boldsymbol{R}$ and α and c are real (so that $f_c(w) = f(x) + \alpha \cdot c$ is clearly real). Now consider any w and v in W and a and b in $\boldsymbol{R}$. Then there are unique $x, y \in \mathrm{dom}\,(f)$ and $\alpha, \beta \in \boldsymbol{R}$ where $w = x + \alpha \cdot u$ and $v = y + \beta \cdot u$. We then have

$$a \cdot w + b \cdot v = a \cdot (x + \alpha \cdot u) + b \cdot (y + \beta \cdot u) = (a \cdot x + b \cdot y) + (a\alpha + b\beta) \cdot u,$$

noting that $a \cdot x + b \cdot y \in \mathrm{dom}\,(f)$ since $\mathrm{dom}\,(f)$ is a subspace and x and y are in $\mathrm{dom}\,(f)$, and $a\alpha + b\beta \in \boldsymbol{R}$. It then follows that

$$f_c(a \cdot w + b \cdot v) = f(a \cdot x + b \cdot y) + (a\alpha + b\beta)c = af(x) + bf(y) + a\alpha c + b\beta c$$
$$= a\,[f(x) + \alpha c] + b\,[f(y) + \beta c] = af_c(w) + bf_c(v)$$

so that f_c is indeed a linear functional on W.

Now consider any $(x, a) \in f$ so that clearly the vector x is in $\mathrm{dom}\,(f)$. Then also $x \in W$ since $W = \mathrm{dom}\,(f) \cup \{u\}$. Also clearly $x = x + 0 \cdot u$ so that $f_c(x) = f(x) + 0 \cdot c = f(x) = a$. Hence $(x, a) \in f_c$ so that $f \subseteq f_c$ since (x, a) was arbitrary. However we also clearly have that $u \in W$ but $u \notin \mathrm{dom}\,(f)$ so that $(u, f_c(u)) \in f_c$ but $(u, f_c) \notin f$. Therefore $f \neq f_c$ so that $f \subset f_c$. This also clearly implies that $f_0 \subseteq f_c$ since $f_0 \subseteq f \subset f_c$.

Now, for all v and t in $\mathrm{dom}\,(f)$ we have

$$f(v) + f(t) = f(v + t) \leq p(v + t) = p([v - u] + [t + u]) \leq p(v - u) + p(t + u)$$

by the linear properties of f and p. Hence

$$f(v) - p(v - u) \leq p(t + u) - f(t)\,.$$

Since v was arbitrary, it follows that the set $\{f(v) - p(v - u) \mid v \in \mathrm{dom}\,(f)\}$ has an upper bound so that $A = \sup\{f(v) - p(v - u) \mid v \in \mathrm{dom}\,(f)\}$ exists. Similarly, since t was arbitrary, the set $\{p(t + u) - f(t) \mid t \in \mathrm{dom}\,(f)\}$ has a lower bound so that $B = \inf\{p(t + u) - f(t) \mid t \in \mathrm{dom}\,(f)\}$ exists. We then also clearly have that $A \leq B$. So let $c = (A + B)/2$ so that $A \leq c \leq B$.

Now consider any $w \in W$ so that $w = x + \alpha \cdot c$ for some $x \in \mathrm{dom}\,(f)$ and $\alpha \in \boldsymbol{R}$.

Case: $\alpha = 0$. Then

$$f_c(w) = f_c(x + \alpha \cdot u) = f(x) + \alpha c = f(x) + 0 = f(x) \leq p(x) = p(x + \alpha \cdot u) = p(w)\,.$$

Case: $\alpha > 0$. Then $1/\alpha \cdot x \in \mathrm{dom}\,(f)$ so that based on the definition of c we have

$$c \leq B \leq p\left(\frac{1}{\alpha} \cdot x + u\right) - f\left(\frac{1}{\alpha} \cdot x\right)$$

Then, since $\alpha > 0$, f is linear functional, and p is a sublinear functional, it follows that

$$\alpha c \leq \alpha \cdot p\left(\frac{1}{\alpha} \cdot x + u\right) - \alpha \cdot f\left(\frac{1}{\alpha} \cdot x\right)$$
$$\alpha c \leq p(x + \alpha \cdot u) - f(x)$$
$$f(x) + \alpha c \leq p(x + \alpha \cdot u)$$
$$f_c(w) \leq p(w)\,.$$

Case: $\alpha < 0$. Then, again $-1/\alpha \cdot x \in \mathrm{dom}\,(f)$ so that based on the definition of c we have

$$f\left(-\frac{1}{\alpha} \cdot x\right) - p\left(-\frac{1}{\alpha} \cdot x - u\right) \leq A \leq c\,.$$

Then, since $-\alpha > 0$, f is a linear functional, and p is a sublinear functional, we have

$$-\alpha \cdot f\left(-\frac{1}{\alpha} \cdot x\right) - (-\alpha) \cdot p\left(-\frac{1}{\alpha} \cdot x - u\right) \leq -\alpha c$$
$$f(x) - p(x + \alpha \cdot u) \leq -\alpha c$$
$$f(x) + \alpha c \leq p(x + \alpha \cdot u)$$
$$f_c(w) \leq p(w)\,.$$

Thus in all cases we have $f_c(w) \leq p(w)$. Since $w \in W$ was arbitrary, this completes the proof that $f_c \in F$. But since $f \subset f_c$, this contradicts the fact that f is a $\subseteq$-maximal element of F. Hence it must be that in fact $\mathrm{dom}\,(f) = V$, thereby completing the overall proof. $\qquad\square$

Theorem 8.2.14*

There is no function $\mu : \mathcal{P}(\mathbf{R}) \to [0, \infty) \cup \{\infty\}$ with the following properties:

0) $\mu([a, b]) = b - a$ for any a and b in $\mathbf{R}$ where $a < b$.

i) $\mu(\varnothing) = 0$ and $\mu(\mathbf{R}) = \infty$.

ii) If $\{A_n\}_{n=0}^{\infty}$ is a collection of mutually disjoint subsets of $\mathbf{R}$, then

$$\mu\left(\bigcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} \mu(A_n).$$

iii) If $a \in \mathbf{R}$, $A \subseteq \mathbf{R}$, and $A + a = \{x + a \mid x \in A\}$, then $\mu(A + a) = \mu(A)$.

iv) If $A \cap B = \varnothing$ then $\mu(A \cup B) = \mu(A) + \mu(B)$.

v) If $A \subseteq B$ then $\mu(A) \leq \mu(B)$.

Solution:

Lemma 8.2.14.1. *If $\{A_i\}_{i \in I}$ is an indexed system of sets and B is another set, then*

$$\bigcup_{i \in I}(B \cap A_i) = B \cap \left(\bigcup_{i \in I} A_i\right)$$

Proof. First, for brevity, let

$$L = \bigcup_{i \in I}(B \cap A_i) \qquad\qquad R = B \cap \left(\bigcup_{i \in I} A_i\right)$$

so that we must show that $L = R$.

($\subseteq$) Consider any $x \in L$ so that there is an $i \in I$ such that $x \in B \cap A_i$. Therefore $x \in B$ and $x \in A_i$. Since $i \in I$ and $x \in A_i$ it follows that $x \in \bigcup_{i \in I} A_i$. Therefore $x \in R$ since also $x \in B$ so that $L \subseteq R$ since x was arbitrary.

($\supseteq$) Now consider any $x \in R$ so that $x \in B$ and $x \in \bigcup_{i \in I} A_i$. Thus there is an $i \in I$ where $x \in A_i$. Then $x \in B \cap A_i$ since also $x \in B$. Since $i \in I$ it follows that $x \in L$ so that $R \subseteq L$ since x was arbitrary.

This shows that $L = R$ as desired. $\qquad\qquad\square$

Main Theorem.

Proof. Suppose to the contrary that there is such a function μ. We define the relation $\approx$ on $\mathbf{R}$ by:

$$x \approx y \text{ if and only if } x - y \text{ is a rational number.}$$

First we show that $\approx$ is an equivalence relation, which is straightforward. First consider any $x \in \mathbf{R}$ so that clearly $x - x = 0 \in \mathbf{Q}$ so that by definition $x \approx x$. Hence $\approx$ is reflexive since x was arbitrary. Now consider any x and y in $\mathbf{R}$ such that $x \approx y$. Then $z = x - y \in \mathbf{Q}$ by definition. Clearly also then $y - x = -(x - y) = -z \in \mathbf{Q}$ as well since $z \in \mathbf{Q}$ so that $y \approx x$. Thus $\approx$ is also symmetric. Lastly, consider x, y, and z in $\mathbf{R}$ where $x \approx y$ and $y \approx z$. Then $a = x - y \in \mathbf{Q}$ and $b = y - z \in \mathbf{Q}$. We then clearly have that

$$x - z = x - y + y - z = (x - y) + (y - z) = a + b \in \mathbf{Q}$$

since a and b are both rational. Thus $x \approx z$ so that $\approx$ is transitive. We have therefore shown that $\approx$ is indeed an equivalence relation.

We then know that $\approx$ partitions $\boldsymbol{R}$ into equivalence classes, and that by the Axiom of Choice and Theorem 8.1.3 the partition has a set of representatives X. Let

$$R = \bigcup \{X + r \mid r \in \boldsymbol{Q}\}$$

and we show that $R = \boldsymbol{R}$. First consider any $y \in R$ so that there is an $r \in \boldsymbol{Q}$ where $y \in X + r$. From this is follows that there is an $x \in X$ such that $y = x + r$. Clearly then, since $x \in \boldsymbol{R}$ (since $X \subseteq R$) and $r \in \boldsymbol{Q}$, we have that $y \in \boldsymbol{R}$ so that $R \subseteq \boldsymbol{R}$ since y was arbitrary. Now consider any real y. Then since X is a set of representatives of a partition of $\boldsymbol{R}$, it follows that there is an $x \in X$ such that $y \in [x]_{\approx}$. Then we have that $y \approx x$ so that $r = y - x \in \boldsymbol{Q}$ by the definition of $\approx$. Thus $y = x + r$ where $x \in X$ so that $y \in X + r$. Since $r \in \boldsymbol{Q}$ it follows that $y \in R$ so that $R \subseteq \boldsymbol{R}$ since y was arbitrary. This shows that $R = \boldsymbol{R}$ as desired.

Now we show that $X + r$ and $X + q$ are disjoint for rationals r and q where $r \neq q$. To see this, suppose that this is not the case and that there is a y where $y \in X + r$ and $y \in X + q$. Then there is an $x_1 \in X$ where $y = x_1 + r$ and an $x_2 \in X$ where $y = x_2 + q$. Now, it cannot be that $x_1 = x_2$ for then we would have that $y = x_1 + r = x_1 + q$, from which it follows that $r = q$, which we know is not the case. So it must be that $x_1 \neq x_2$, which means that x_1 and x_2 are in different equivalent classes since X is a set of representatives of disjoint equivalence classes. In particular this means that $x_1 \not\approx x_2$ so that $x_1 - x_2 \notin \boldsymbol{Q}$. We then have

$$y = x_1 + r = x_2 + q$$
$$x_1 - x_2 = q - r \,,$$

but this is a contradiction since we have shown that the left side is irrational but the right side is clearly rational since q and r are. It therefore must be that $X + r$ and $X + q$ are indeed disjoint.

Now suppose for the moment that $\mu(X) = 0$. It then follows from property iii) that $\mu(X + q) = \mu(X) = 0$ for any rational q. Then, since we showed that $R = \boldsymbol{R}$, we have that

$$\mu(\boldsymbol{R}) = \mu(R) = \mu\left(\bigcup \{X + r \mid r \in \boldsymbol{Q}\}\right) = \sum_{r \in \boldsymbol{Q}} \mu(X + r) = \sum_{r \in \boldsymbol{Q}} 0 = 0$$

by property ii), noting that this applies because $\boldsymbol{Q}$ is countable. This contradicts property i), according to which $\mu(\boldsymbol{R}) = \infty$. So it must be that $\mu(X) > 0$.

Suppose next that $\mu(X \cap [a, b]) = 0$ for every closed interval $[a, b]$. Let $A_n = [n, n+1]$ for $n \in \boldsymbol{Z}$ so that clearly $\bigcup_{n \in \boldsymbol{Z}} A_n = \boldsymbol{R}$. We also have that $\mu(X \cap A_n) = 0$ for any $n \in \boldsymbol{Z}$ since each A_n is a closed interval. It follows from Lemma 8.2.14.1 that

$$\bigcup_{n \in \boldsymbol{Z}} (X \cap A_n) = X \cap \left(\bigcup_{n \in \boldsymbol{Z}} A_n\right) = X \cap \boldsymbol{R} = X$$

since $X \subseteq \boldsymbol{R}$. We then have

$$\mu(X) = \mu\left(\bigcup_{n \in \boldsymbol{Z}} (X \cap A_n)\right) = \sum_{n \in \boldsymbol{Z}} (X + A_n) = \sum_{n \in \boldsymbol{Z}} 0 = 0$$

by property ii) since $\boldsymbol{Z}$ is countable. As before, this contradicts property i) so that it has to be that there is a closed interval $[a, b]$ such that $\mu(X \cap [a, b]) > 0$.

Let $Y = X \cap [a, b]$ and $P = \boldsymbol{Q} \cap [0, 1]$. Next we show that the sets $\{Y + q\}_{q \in P}$ are mutually disjoint. Suppose the contrary so that there are p and q in P where $p \neq q$ and a z such that $z \in Y + p$ and

$y \in Y + q$. Then there is a $y_1 \in Y$ where $z = y_1 + p$ and a $y_2 \in Y$ such that $z = y_2 + q$. We then have that y_1 and y_2 are both in X so that $z = y_1 + p \in X + p$ and $z = y_2 + q \in X + q$. But since $p \neq q$, this is not possible since we have shown above that $X + p$ and $X + q$ are disjoint. So it has to be that $\{Y + q\}_{q \in P}$ are mutually disjoint sets.

We then have that

$$\mu \left(\bigcup_{q \in P} (Y + p) \right) = \sum_{q \in P} \mu(Y + p) = \sum_{q \in P} \mu(Y)$$

by properties ii) and iii), noting that clearly $P = \boldsymbol{Q} \cap [0, 1]$ is countably infinite. Since $\mu(Y) > 0$ this is an infinite sum of a positive constant so that the result diverges and is thus ∞.

Now consider any $z \in \bigcup_{q \in P}(Y + p)$ so that there is a $q \in P$ such that $z \in Y + q$. Then there is an $y \in Y$ such that $z = y + q$. It follows that $y \in [a, b]$ so that $a \leq y \leq b$. We also have that $0 \leq q \leq 1$ since $q \in P = \boldsymbol{Q} \cap [0, 1]$. We then have that $z = y + q \leq b + q \leq b + 1$. Also clearly $z = y + q \geq a + q \geq a + 0 = a$. Therefore $a \leq z \leq b + 1$ so that $z \in [a, b + 1]$, which shows that $\bigcup_{q \in P}(Y + p) \subseteq [a, b + 1]$ since z was arbitrary. However, this implies that

$$\infty = \mu \left(\bigcup_{q \in P} (Y + p) \right) \leq \mu([a, b + 1]) = b + 1 - a$$

by properties v) and 0). This clearly a contradiction though since $b + 1 - a$ is finite. Thus our initial assumption that there exists such a function μ must be incorrect. $\qquad \square$

Chapter 9   Arithmetic of Cardinal Numbers

§9.1   Infinite Sums and Products of Cardinal Numbers

Theorem 9.1.3

Let λ be an infinite cardinal, let κ_α $(\alpha < \lambda)$ be nonzero cardinal numbers, and let $\kappa = \sup \{\kappa_\alpha \mid \alpha < \lambda\}$. Then

$$\sum_{\alpha < \lambda} \kappa_\alpha = \lambda \cdot \kappa = \lambda \cdot \sup \{\kappa_\alpha \mid \alpha < \lambda\} .$$

Solution:

Lemma 9.1.3.1. *If κ and λ are cardinal numbers and $\kappa \leq \lambda$ in the ordinal ordering (since κ and λ are also ordinals), then also $\kappa \leq \lambda$ in the cardinal sense of ordering.*

Proof. A proof of this is quite simple. Since $\kappa \leq \lambda$ in the ordinal sense, either $\kappa < \lambda$ (in the ordinal sense) or $\kappa = \lambda$. In the former case $\kappa \in \lambda$ by the definition of ordinal ordering so that also $\kappa \subseteq \lambda$ since λ is transitive (since it is an ordinal). In the case where $\kappa = \lambda$ then obviously $\kappa \subseteq \lambda$. Hence in either case $\kappa \subseteq \lambda$ so that $\kappa \leq \lambda$ in the cardinal sense since clearly the identity function on κ is an injection from κ to λ. $\qquad \square$

Main Theorem.

Proof. First note that $\sup\{\kappa_\alpha \mid \alpha < \lambda\}$ refers to the ordinal supremum, which is well defined since every cardinal is also an ordinal (though the converse is not true). In particular the cardinal numbers are exactly the initial ordinals. In order to avoid ambiguity in what follows we shall set $\beta = \kappa$ and $\beta_\alpha = \kappa_\alpha$ for each $\alpha < \lambda$. When referring to ordinal ordering we use β and β_α and, when referring to cardinal ordering, use κ and κ_α.

We show first that $\kappa = \beta = \sup\{\beta_\alpha \mid \alpha < \lambda\}$ is actually a cardinal number, for which it suffices to show that β is an initial ordinal. So consider any ordinal $\gamma < \beta$. Then it follows that γ is not an upper bound of $\{\beta_\alpha\}_{\alpha<\lambda}$ so that there is an $\alpha < \lambda$ where $\gamma < \beta_\alpha$. Since β_α is an initial ordinal, we have $|\gamma| < |\beta_\alpha| = \kappa_\alpha$. We also have that $\beta_\alpha \leq \beta$ (since β is an upper bound of $\{\beta_\alpha\}_{\alpha<\lambda}$) so that $\kappa_\alpha = |\beta_\alpha| \leq |\beta|$ by Lemma 9.1.3.1 Therefore we have $|\gamma| < \kappa_\alpha \leq |\beta|$, which shows that β is an initial ordinal since $\gamma < \beta$ was arbitrary. Hence κ is indeed a cardinal number.

Proceeding to prove the actual result, clearly we have that $\kappa_\alpha \leq \kappa$ for each $\alpha < \lambda$ by Lemma 9.1.3.1 since $\beta_\alpha \leq \beta$ (since β is an upper bound of $\{\beta_\alpha\}_{\alpha<\lambda}$). It then follows from Exercises 9.1.2 and 9.1.4 that

$$\sum_{\alpha<\lambda} \kappa_\alpha \leq \sum_{\alpha<\lambda} \kappa = \lambda \cdot \kappa\,.$$

Since each κ_α is nonzero, we have that $1 \leq \kappa_\alpha$ for all $\alpha < \lambda$ and hence

$$\lambda = \lambda \cdot 1 = \sum_{\alpha<\lambda} 1 \leq \sum_{\alpha<\lambda} \kappa_\alpha\,,$$

where we again have used Exercises 9.1.2 and 9.1.4.

Now, let $\beta_s = \kappa_s = \sum_{\alpha<\lambda} \kappa_\alpha$. Also let $\langle A_\alpha \mid \alpha < \lambda \rangle$ be mutually disjoint sets where $|A_\alpha| = \kappa_\alpha$ for each $\alpha < \lambda$, then it follows by definition that $\kappa_s = \sum_{\alpha<\lambda} \kappa_\alpha = \left|\bigcup_{\alpha<\lambda} A_\alpha\right|$. Then, for any $\gamma < \lambda$, we clearly have $A_\gamma \subseteq \bigcup_{\alpha<\lambda} A_\alpha$ so that $\kappa_\gamma = |A_\gamma| \leq \left|\bigcup_{\alpha<\lambda} A_\alpha\right| = \kappa_s$. It then follows from Lemma 7.1.9.5 that $\beta_\gamma \leq \beta_s$. Since γ was arbitrary this shows that κ_s is an upper bound of $\{\beta_\alpha\}_{\alpha<\lambda}$. Then, since β is the *least* upper bound of $\{\beta_\alpha\}_{\alpha<\lambda}$, we have that $\beta \leq \beta_s$ so that $\kappa \leq \kappa_s$ by Lemma 9.1.3.1.

Since λ is infinite, it follows from Corollary 7.2.2 that $\lambda \cdot \kappa$ is the greater of λ and κ. It then follows that $\lambda \cdot \kappa \leq \sum_{\alpha<\lambda} \kappa_\alpha$ since we have shown that both $\lambda \leq \sum_{\alpha<\lambda} \kappa_\alpha$ and $\kappa \leq \kappa_s = \sum_{\alpha<\lambda} \kappa_\alpha$. Hence, since we have shown that both $\sum_{\alpha<\lambda} \kappa_\alpha \leq \lambda \cdot \kappa$ and $\lambda \cdot \kappa \leq \sum_{\alpha<\lambda} \kappa_\alpha$, it follows from the Cantor-Bernstein Theorem that $\sum_{\alpha<\lambda} \kappa_\alpha = \lambda \cdot \kappa$ as desired. $\square$

Theorem 9.1.7

(König's Theorem) If κ_i and λ_i $(i \in I)$ are cardinal numbers, and if $\kappa_i < \lambda_i$ for all $i \in I$, then

$$\sum_{i \in I} \kappa_i < \prod_{i \in I} \lambda_i\,.$$

Solution:

Proof. First, we show that $\sum_{i \in I} \kappa_i \leq \prod_{i \in I} \lambda_i$. So let $\langle A_i \mid i \in I \rangle$ and $\langle B_i \mid i \in I \rangle$ be systems of sets where $|A_i| = \kappa_i$ and $|B_i| = \lambda_i$ for all $i \in I$. Also suppose that $\langle A_i \mid i \in I \rangle$ are mutually disjoint.

Since, for any $i \in I$, we have that $|A_i| = \kappa_i < \lambda_i = |B_i|$, we can clearly assume that $A_i \subset B_i$ (since, if not, then each A_i will still be equipotent to some proper subset of B_i). Hence we have

$$\sum_{i \in I} \kappa_i = \left| \bigcup_{i \in I} A_i \right| \qquad\qquad \prod_{i \in I} \lambda_i = \left| \prod_{i \in I} B_i \right|$$

by the definition of the cardinal sum and product. We now construct an injective function f from $\bigcup_{i \in I} A_i$ into $\prod_{i \in I} B_i$, the existence of which then shows that

$$\sum_{i \in I} \kappa_i = \left| \bigcup_{i \in I} A_i \right| \leq \left| \prod_{i \in I} B_i \right| = \prod_{i \in I} \lambda_i \,.$$

So do this, we first note that $B_i - A_i \neq \varnothing$ (since $A_i \subset B_i$) for all $i \in I$, so that we can choose any $d_i \in B_i - A_i$ for each $i \in I$ (which requires the Axiom of Choice). Now, let x be any element of $\bigcup_{i \in I} A_i$ so that there is a unique $i_x \in I$ (unique because $\langle A_i \mid i \in I \rangle$ are mutually disjoint) such that $x \in A_{i_x}$. We then set $f(x) = \langle a_i \mid i \in I \rangle$ where

$$a_i = \begin{cases} x & i = i_x \\ d_i & i \neq i_x \end{cases}$$

for any $i \in I$. Clearly, when $i = i_x$, we have $a_i = x \in A_{i_x}$ so that $a_i \in B_{i_x} = B_i$ since $A_{i_x} \subset B_{i_x}$. If $i \neq i_x$ then we have that $a_i = d_i \in B_i - A_i$ so that obviously $a_i \in B_i$ as well. Thus clearly $f(x) = \langle a_i \mid i \in I \rangle \in \prod_{i \in I} B_i$ so that f is a function from $\bigcup_{i \in I} A_i$ to $\prod_{i \in I} B_i$.

Now consider any x and y in $\bigcup_{i \in I} A_i$ where $x \neq y$, and let $f(x) = \langle a_i \mid i \in I \rangle$ and $f(y) = \langle b_i \mid i \in I \rangle$. If $i_x = i_y$ then we have $a_{i_x} = x \neq y = b_{i_y} = b_{i_x}$. On the other hand, if $i_x \neq i_y$, then $a_{i_x} = x \in A_{i_x}$ while $b_{i_x} = d_{i_x} \in B_{i_x} - A_{i_x}$ so that $b_{i_x} \notin A_{i_x}$. Hence it has to be that $a_{i_x} \neq b_{i_x}$ here to. This, in both cases, we clearly have that $f(x) = \langle a_i \mid i \in I \rangle \neq \langle b_i \mid i \in I \rangle = f(y)$ since $a_{i_x} \neq b_{i_x}$ and $i_x \in I$. This clearly shows that f is injective since x and y were arbitrary, which of course shows that indeed $\sum_{i \in I} \kappa_i \leq \prod_{i \in I} \lambda_i$.

Now let us show that $\sum_{i \in I} \kappa_i < \prod_{i \in I} \lambda_i$. To this end, suppose that $\langle B_i \mid i \in I \rangle$ are sets such that $|B_i| = \lambda_i$ for all $i \in I$. Suppose for the moment that $\sum_{i \in I} \kappa_i$ and $\prod_{i \in I} \lambda_i$ are actually equal, and let $\langle A_i \mid i \in I \rangle$ be mutually disjoint sets where $|A_i| = \kappa_i$ for each $i \in I$. It would then follow that there is a bijection f from $\bigcup_{i \in I} A_i$ to $\prod_{i \in I} B_i$. We then have that $X_i = f[A_i]$ is a subset of $\prod_{i \in I} B_i$ for each $i \in I$. Furthermore, suppose i and j in I where $i \neq j$ and that $b \in X_i \cap X_j$. Since f is a bijection, it follows that f^{-1} is also a bijective function so that $f^{-1}[X_i] = f^{-1}[f[A_i]] = A_i$ and $f^{-1}[X_j] = f^{-1}[f[A_j]] = A_j$. Hence we would have that $\inf f(b) \in A_i$ since $b \in X_i$ and $f^{-1}(b) \in A_j$ since $b \in X_j$, but this violates the supposition that $\langle A_i \mid i \in I \rangle$ are mutually disjoint since $i \neq j$. So it must be that $\langle X_i \mid i \in I \rangle$ are also mutually disjoint.

Since each X_i is a subset of $\prod_{i \in I} B_i$, clearly we have that $\bigcup_{i \in I} X_i \subseteq \prod_{i \in I} B_i$. Then, for any $b \in \prod_{i \in I} B_i$ we have that $f^{-1}(b) \in \bigcup_{i \in I} A_i$ so that there is an $i_0 \in I$ where $f^{-1}(b) \in A_{i_0}$. Then clearly we have that $b = f(f^{-1}(b)) \in X_{i_0}$ since $X_{i_0} = f[A_{i_0}]$. Hence clearly $b \in \bigcup_{i \in I} X_i$ since $i_0 \in I$. There for $\prod_{i \in I} B_i \subseteq \bigcup_{i \in I} X_i$ also so that in fact $\bigcup_{i \in I} X_i = \prod_{i \in I} B_i$. Moreover, clearly since f is a bijection, we have that $|X_i| = |f[A_i]| = |A_i| = \kappa_i$ for all $i \in I$. Therefore we have shown that $\sum_{i \in I} \kappa_i = \prod_{i \in I} \lambda_i$ implies that there are mutually disjoint subsets $\langle X_i \mid i \in I \rangle$ of $\prod_{i \in I} B_i$ such that $|X_i| = \kappa_i$ for all $i \in I$ and $\bigcup_{i \in I} X_i = \prod_{i \in I} B_i$. We show by contradiction that this situation is impossible, from which it follows that $\sum_{i \in I} \kappa_i \neq \prod_{i \in I} \lambda_i$ (and therefore the desired result follows since it has already been shown that $\sum_{i \in I} \kappa_i \leq \prod_{i \in I} \lambda_i$).

So suppose that there are such mutually disjoint subsets $\langle X_i \mid i \in I \rangle$ of $\prod_{i \in I} B_i$. For each $i \in I$, we let $A_i = \{a_i \mid \langle a_j \mid j \in I \rangle \in X_i\}$. Then, for any $x \in A_i$, there is an $a \in X_i$ where $x = a(i)$.

It then follows that $x \in B_i$ since $a \in X_i \subseteq \prod_{i \in I} B_i$. Thus $A_i \subseteq B_i$. We also clearly have $|A_i| \leq |X_i| = \kappa_i < \lambda_i = |B_i|$ so that $A_i \subset B_i$. Thus there is a b_i such that $b_i \in B_i$ but $b_i \notin A_i$. So we define $b = \langle b_i \mid i \in I \rangle$ so that clearly $b \in \prod_{i \in I} B_i$. However, for any $j \in I$, we have that $b_j \notin A_j$. If it were the case that $b \in X_j$ then we would have that $b(j) = b_j \in A_j$ by the definition of A_j. So it has to be that $b \notin X_j$ since we know that $b_j \notin A_j$. Since j was arbitrary, it follows that $b \notin X_i$ for all $i \in I$, and hence $b \notin \bigcup_{i \in I} X_i$. But, since $b \in \prod_{i \in I} B_i$, this contradicts the fact that $\bigcup_{i \in I} X_i = \prod_{i \in I} B_i$. The result follows from this contradiction as described above. $\qquad \square$