# SOLUTIONS TO EXERCISES FOR

# MATHEMATICS 144 — Part 6

**Fall 2006**

## VI. Infinite constructions in set theory

### VI.1 : Indexed families and set - theoretic operations

*Exercises to work*

**1.** We first verify the statement about unions. Suppose that $x \in \cup_{j\in J} A_j$, and choose $j(0) \in J$ such that $x \in A_{j(0)}$. Then the inclusion hypothesis implies that $x \in C_{j(0)}$, which in turn implies that $x \in \cup_{j\in J} C_j$. Therefore $\cup_{j\in J} A_j$ is a subset of $\cup_{j\in J} C_j$.

We now verify the statement about intersections. Suppose that $x \in \cap_{j\in J} A_j$, so that $x \in A_j$ for every $j \in J$. The inclusion hypothesis now implies that $x \in C_j$ for every $j \in J$, and therefore we must have $x \in \cap_{j\in J} C_j$. Therefore $\cap_{j\in J} A_j$ is a subset of $\cap_{j\in J} C_j$.■

**2.** We prove the assertions in order. Suppose that $x \in S - \cup_{j\in J} A_j$. Then $x \notin \cup_{j\in J} A_j$, or equivalently there is no $j$ such that $x \in A_j$. Therefore we have $x \notin A_j$ for all $j$, and since $x \in S$ this means $x \in S - A_j$ for all $j$. The latter in turn implies that $x \in \cap_{j\in J} S - A_j$, and therefore we have $S - \cup_{j\in J} A_j \subset \cap_{j\in J} S - A_j$.

Conversely, if $x \in \cap_{j\in J} S - A_j$, then $x \notin A_j$ for each $j$, so that there is no $j$ satisfying $x \in A_j$ and hence $x \notin \cup_{j\in J} A_j$. Since $x \in S$, it follows that $x \in S - \cup_{j\in J} A_j$, and this plus the conclusion of the previous paragraph establishes one of the De Morgan laws.

We now turn to the other De Morgan law. Suppose that $x \in S - \cap_{j\in J} A_j$. Then there is some $j(0) \in J$ such that $x \notin A_{j(0)}$, and accordingly we have $x \in S - A_{j(0)}$. Now the latter set is a subset of the union $\cup_{j\in J} S - A_j$ by the definition of this union, and therefore we have $S - (\cap_{j\in J} A_j) \subset \cup_{j\in J} S - A_j$.

Conversely, if $x \in \cup_{j\in J} S - A_j$, then there is some $j(0)$ such that $x \in S - A_{j(0)}$, so that $x \notin A_{j(0)}$. The last statement implies that $x \notin \cap_{j\in J} A_j$, and since $x \in S$ it follows that $x \in S - (\cap_{j\in J} A_j)$. As in the discussion of the first De Morgan law, this plus the conclusion of the previous paragraph establishes the second De Morgan law.■

**3.** $(a)$ Suppose first that $x \in (\cup_i A_i) \cap (\cup_j B_j)$. Then one can find indices $i(0)$ and $j(0)$ such that $x \in A_{i(0)}$ and $x \in B_{j(0)}$, and hence $x \in A_{i(0)} \cap B_{j(0)}$, so that $x \in \cap_{i,j} (A_i \cup B_j)$. — Conversely, if $x$ lies in the latter set, then one can find indices $i(0)$ and $j(0)$ such that $x \in A_{i(0)} \cap B_{j(0)}$. Since $A_{i(0)}$ is a subset of $\cup_i A_i$ and $B_{j(0)}$ is a subset of $\cup_j B_j$, it follows that the intersection $A_{i(0)} \cap B_{j(0)}$ is a subset of $(\cup_i A_i) \cap (\cup_j B_j)$. This proves the first identity in $(a)$.■

We now turn to the second identity. Suppose that $x \in (\cap_i A_i) \cup (\cap_j B_j)$. Then either $x \in \cap_i A_i$ or $x \in \cap_j B_j$. In the first case we have $x \in A_i$ for all $i$ and in the second we have $x \in B_j$ for all $j$. Therefore in both cases we have $x \in A_i \cup B_j$ for all $i$ and $j$., so that $x \in \cap_{i,j} (A_i \cup B_j)$. — Conversely, if $x$ lies in the latter set, then for each ordered pair $(i, j)$ we either have $x \in A_i$ or

$x \in B_j$. It suffices to show that if $x \notin \cap_i\ A_i$, then we must have $x \in \cap_j\ B_j$. However, if $x$ does not belong to the first intersection, then for some $i(0)$ we have $x \notin A_{i(0)}$, and thus for all ordered pairs $(\,i(0), j\,)$ we must have $x \in B_j$. The last statement implies that $x \in \cap_j\ B_j$, which is what we needed to verify.■

($b$) Suppose that $x \in A_k$ for some $k$, and choose $j$ such that $k \in I_j$. We then have $x \in \cup\ \{A_i \mid i \in I_j\}$, which in turn implies

$$x \quad \in \quad \bigcup_{j \in J} \left( \bigcup_{i \in I_j} A_i \right) .$$

Conversely, if $x$ belongs to the latter set, then for some $j$ we have $x \in \cup\ \{A_i \mid i \in I_j\}$, which in turn means that $x \in A_i$ for some $i$, so that $x \in \cup_k\ A_k$. This proves the first identity in ($b$).■

We now turn to the second identity. Suppose that $x \in A_k$ for all $k$. Then for each $j$ we have $x \in \cap\ \{A_i \mid i \in I_j\}$, and hence we also have

$$x \quad \in \quad \bigcap_{j \in J} \left( \bigcap_{i \in I_j} A_i \right) .$$

Conversely, if $x$ belongs to the latter set, then for all $j$ we have $x \in \cap\ \{A_i \mid i \in I_j\}$, which in turn means that $x \in A_i$ for all $i$, so that $x \in \cap_k\ A_k$. This proves the second identity in ($b$).■

**4.** ($a$) Suppose first that $(x, y) \in (\cup_i\ A_i)\ \times\ (\cup_j\ B_j)$. Then one can find some indices $i(0)$ and $j(0)$ such that $x \in A_{i(0)}$ and $y \in B_{j(0)}$. Therefore we have $(x, y) \in A_{i(0)} \times B_{j(0)}$. Since the latter is contained in $\cup_{i,j}\ (A_i \times B_j)$ it follows that $(x, y) \in \cup_{i,j}\ (A_i \times B_j)$. — Conversely, it $(x, y)$ belongs to the latter set, then one can find some indices $i(0)$ and $j(0)$ such that $(x, y) \in A_{i(0)} \times B_{j(0)}$, and therefore it follows that $(x, y)$ belongs to $(\cap_i\ A_i)\ \times\ (\cap_j\ B_j)$. This proves the first identity in ($a$).■

We now turn to the second identity. Suppose that $(x, y) \in (\cap_i\ A_i)\ \times\ (\cap_j\ B_j)$. Then for all $i$ and $j$ we have $x \in A_i$ and $y \in B_j$, so that $x \in A_i \times B_j$ for all $i$ and $j$, and hence we have $(x, y)\ \in\ \cap_{i,j}\ (A_i \times B_j)$. — Conversely, if $(x, y)$ belongs to the latter set, then for each $i$ and $j$ we know that $x \in A_i$ and $y \in B_j$, so that $(x, y) \in (\cap_i\ A_i)\ \times\ (\cap_j\ B_j)$. This proves the second identity in ($a$).■

($b$) First of all, we need to show that for each $j$ we have $\cap_i\ X_i \subset X_j \subset \cup_i\ X_i$. If $y$ lies in the intersection on the left hand side, then it lies in each $X_i$ and in particular it lies in $X_j$, so the first inclusion is true. Likewise, if $y \in X_j$ then trivially we have $f \in X_i$ for some $i$ and hence $y \in \cup_i\ X_i$.

To complete the second part of the problem, we need to show that if the sets $U$ and $V$ satisfy

$$U \quad \subset \quad X_j \quad \subset \quad V$$

for every $j$, then we have $U \subset \cap_i\ X_i$ and $\cup_i\ X_i \subset V$. — If $y \in U$, then by hypothesis we have $y \in X_i$ for all $i$ and since the intersection is defined by this condition we have $U \subset \cap_i\ X_i$. Also, if $y \in \cup_i\ X_i$, then for some $j$ we have $y \in X_j$. By assumption $X_j \subset V$, and therefore we also have $y \in V$. But this means that every element of $\cup_i\ X_i$ is also in $V$, so that $X_j \subset V$ as required.■

## VI.2 : Infinite Cartesian products

*Exercises to work*

**1.** The main idea is to apply the Universal Mapping Property:

*Let $\{A_\alpha \mid \alpha \in \Lambda\}$ be a family of nonempty sets, and suppose that we are given data consisting of a set $P$ and functions $h_\alpha : P \to A_\alpha$ such that for* EVERY *collection of data $(S, \{f_\alpha : S \to A_\alpha\})$ there is a unique function $f : S \to P$ such that $h_\alpha \circ f = f_\alpha$ for all $\alpha$. Then there is a unique $1-1$ correspondence $\Phi : \prod_\alpha A_\alpha \to P$ such that $h_\alpha \circ \Phi$ is projection from $\prod_\alpha X_\alpha$ onto $A_\alpha$ for all $\alpha$.*

**Application to the exercise.** For each $k$ let $P_k$ denote the product of objects whose index belongs to $J_k$ and denote its coordinate projections by $p_i$. The conclusions amount to saying that there is a canonical morphism from $\prod_k P_k$ to $\prod_i X_i$ that has an inverse morphism. Suppose that we are given morphisms $f_i$ from the same set $S$ to the various sets $X_i$. If we gather together all the morphisms for indices $i$ lying in a fixed subset $J_k$, then we obtain a unique map $g_k : S \to P_k$ such that $p_i \circ g_k = f_i$ for all $i \in J_k$.

Let $q_k : \prod_\ell P_\ell \to P_k$ be the coordinate projection. Taking the maps $g_k$ that have been constructed, one obtains a unique map $F : S \to \prod_k P_k$ such that $q_k \circ F = g_k$ for all $k$. By construction we have that $p_i \circ q_k \circ F = f_i$ for all $i$. If there is a unique map with this property, then $\prod_k P_k$ will be isomorphic to $\prod_i X_i$ by the Universal Mapping Property. But suppose that $\theta$ is any map with this property. Once again fix $k$. Then $p_i \circ q_k \circ F = p_i \circ q_k \circ \theta = f_i$ for all $i \in J_k$ implies that $q_k \circ F = q_k \circ \theta$, and since the latter holds for all $k$ it follows that $F = \theta$ as required.■

**2.** Once again we use the Universal Mapping Property. If we are given a sequence of set-theoretic functions $f_i : X_i \to Y_i$, then we obtain a corresponding set of functions $f_i^* : \prod_j X_j \to Y_i$ defined by the identities

$$f_i^* \quad = \quad f_i \circ p_i^X$$

where $p_i^x : \prod_j X_j \to X_i$ is projection. Thus the Universal Mapping Property yields a unique mapping

$$F = \prod_i f_i \ : \ \prod_i X_i \quad \to \quad \prod_i Y_i$$

such that $p_i^Y \circ F = f_i^* = f_i \circ p_i^X$ for each $i$, where $\pi_i^X$ and $\pi_i^Y$ denote the $i^{\text{th}}$ coordinate projections for $\prod_i X_i$ and $\prod_i Y_i$ respectively.

The assertion that $F$ is an identity mapping if each $f_i$ is an identity mapping follows because from the uniqueness part of the Universal Mapping Property, for the identity mapping on the product satisfies the displayed equation if each of the mappings $f_i$ is an identity mapping.

Finally, the assertion about composites can be verified as follows: Let $H = \prod_j g_j \circ f_j$. Then for each $i$ we have

$$p_i^Z \circ H \quad = \quad g_i \circ f_i \circ \pi_i^X \quad = \quad g_i \circ p_i^Y \circ F \quad = \quad p_i^Z \circ G \circ F$$

and therefore $H = G \circ F$ by the Universal Mapping Property.■

**3.** Follow the hint. Since each $f_j$ is a bijection we have inverse mappings $g_j = f_j^{-1}$. By the preceding exercise we then have

$$\prod_j f_j \ \circ \ \prod_j g_j \quad = \quad \prod_j (f_j \circ g_j) \quad = \quad \prod_j \mathrm{id}(Y_j) \quad = \quad \mathrm{id}\left(\prod_j Y_j\right)$$

and we also have

$$\prod_j g_j \circ \prod_j f_j \quad = \quad \prod_j (g_j \circ f_j) \quad = \quad \prod_j \mathrm{id}(X_j) \quad = \quad \mathrm{id}\left(\prod_j X_j\right)$$

so that the product of the inverses $\prod_j g_j$ is an inverse to $\prod_j f_j$.■

**4.** Both statements are **TRUE.** — To prove the first one, let $\mathbf{u}, \mathbf{v} \in \prod_j X_j$ with coordinates $u_j$ and $v_j$ respectively, and suppose that $\prod_j f_j(\mathbf{u}) = \prod_j f_j(\mathbf{v})$. This means that the $j^{\mathrm{th}}$ coordinates are equal for every $j$. But the $j^{\mathrm{th}}$ coordinates of the given elements are $f_j(u_j)$ and $f_j(v_j)$ respectively. Since each $f_j$ is 1–1 it follows that $u_j = v_j$ for all $j$, which in turn means that $\mathbf{u} = \mathbf{v}$. Therefore $\prod_j f_j$ is 1–1 if each $f_j$ is 1–1.

Suppose now that each map $f_j$ is onto, and let $\mathbf{y} \in \prod_j Y_j$ with coordinates $y_j$. Since each $f_j$ is onto for each $j$ there is an element $x_j \in X_j$ such that $f_j(x_j) = y_j$. If we take $\mathbf{x} \in \prod_j X_j$ such that the $j^{\mathrm{th}}$ coordinate is $x_j$ for each $j$, then it follows that $\prod_j f_j(\mathbf{x}) = \mathbf{y}$ and hence $\prod_j f_j$ is onto.■

**5.** We shall use the following result from Unit IV: *Let $X$ and $Y$ be sets, let $\varphi : X \to Y$ be a function, let $R$ be a binary relation on $X$, and let $\mathbf{E}$ be the equivalence relation generated by $R$. Suppose that for all $u, v \in X$ we know that $u\,R\,v$ implies $\varphi(u) = \varphi(v)$. Then for all $x, y \in X$ such that $x\,\mathbf{E}\,y$ we have $\varphi(x) = \varphi(y)$.*

To solve the problem, let $R$ be the binary relation on $B$ such that $uBv$ if and only if there is some $x \in A$ such that $u = f(x)$ and $v = g(x)$, let $\mathbf{E}$ be the equivalence relation generated by $R$, let $C$ be the corresponding set of equivalence classes, and let $p : B \to C$ be the equivalence class projection. By construction we have $p \circ f(x) = p \circ g(x)$ for all $x \in A$.

Suppose now that we have a function $q : B \to D$ such that $q \circ f = q \circ g$. We need to define a function $h : C \to D$ such that $h$ sends the equivalence class $[b]$ of $b$ to $q(b)$. The main problem is to verify that $h$ is well defined; *i.e.*, it does not depend upon the choice of an element of $b$ representing a given equivalence class. If we can show that $h$ is well defined, the it will follow that $h \circ p = q$; furthermore if we also have $k \circ p = q$, then for each $c \in C$ we may write $c = p(b)$ for some $b$ and hence

$$k(c) \quad = \quad k \circ p(b) \quad = \quad q(b) \quad = \quad h \circ p(b) \quad = \quad h(x)$$

so that $h = k$. — By the proposition quoted in the first paragraph, it suffices to show that if $u\,R\,v$ then $q(u) = q(v)$, and by definition of $R$ this reduces to showing that for each $x \in A$ we have $q \circ f(x) = q \circ g(x)$; this equation holds by our hypothesis on $q$, and therefore by the proposition we know that $h$ is well defined. As noted before, this completes the proof.■

**6.** First of all, we observe a consequence of the uniqueness statement. Namely, the only maps $\varphi : C \to C$ and $\psi : E \to E$ such that $p \circ \varphi = p$ and $q \circ \psi = q$ are the identities on $C$ and $E$ respectively.

By the universal mapping properties for coequalizers, there are unique maps $H : C \to E$ such that $r = H \circ p$ and $K : E \to C$ such that $p = K \circ r$. It follows that $r = K \circ H \circ r$ and $p = H \circ K \circ p$, and therefore by the first sentence we conclude that $H \circ K$ and $K \circ H$ are the identity mappings on $C$ and $E$ respectively. As usual, this implies that both $H$ and $K$ are bijections.■

## VI.3 : Transfinite cardinal numbers

*Exercises to work*

**1.** Let $S$ be the set in question, and let $S[n]$ denote the set of subsets with $n$ elements. It will suffice to show that each $A[n]$ is countable, because a countable union of countable sets is countable. In fact, since there is only one subset with no elements, we might as well assume that $n \geq 1$.

Since $S$ is countable it has a well-ordering. Define a map $F$ from $S[n]$ to $\Pi^n S$ — the product of $n$ copies of $S$ with itself — such that the coordinates of $F(B)$ are the elements of $B$ in order; *i.e.*, the first coordinate is the least element $b_0$, the second is the least element of those remaining after $b_0$ is removed, and so on. This defines a 1–1 mapping into $\Pi^n S$, which is a countable set. Hence each set $S[n]$ is countable as required.

To see the final assertion, note that if $S$ is finite then the set $P(S)$ of all subsets is finite, while if $S$ is infinite, then the set $\mathbf{F}_1(S)$ of subsets with exactly one element is in 1–1 cooresponcence with $S$, and hence both $\mathbf{F}_1(S)$ and $P(S) \supset \mathbf{F}_1(S)$ are infinite.∎

**2.** The equivalence class projection from $S$ to $S/E$ is an onto mapping, and since $S$ is countable the results of Section 3 imply that $S/E$ must also be countable.∎

## VI.4 : Countable and uncountable sets

*Exercises to work*

**1.** Let $A$, $B$, $C$, $D$ be sets such that $|A| = \alpha$, $|B| = \beta$, $|C| = \gamma$, and $|D| = \delta$. The cardinal number inequalities imply the existence of 1–1 mappings $f : A \to B$ and $g : C \to D$. These maps in turn define mappings $f \amalg g : A \amalg C \to B \amalg D$ and $f \times g : A \times C \to B \times D$ as follows:

$$[f \amalg g]\,(a, 1) \quad = \quad (\,f(a), 1\,)$$

$$[f \amalg g]\,(c, 2) \quad = \quad (\,g(c), 2\,)$$

$$[f \times g]\,(c, b) \quad = \quad (\,f(a),\, g(c)\,)$$

Since $\alpha + \gamma = |A \amalg C|$ and $\alpha \cdot \gamma = |A \times C|$ and similarly if $\beta, \delta, B, D$ replace $\alpha, \gamma, A, C$ the conclusion of the problem will follow if we can verify that $f \amalg g$ and $f \times g$ are both 1–1.

To see that $f \amalg g$ is 1–1, suppose that we have two classes $(u, i)$ and $(v, j)$ which have the same image under this map. By definition of $f \amalg g$ we know that the second coordinates satisfy $i = j$ so that either this second coordinate is 1 and $u, v \in A$ or else this second coordinate is 2 and $u, v \in B$. In each case the injectivity of $f$ and $g$ imply that the images of $(u, i)$ and $(v, i)$ are the same if and only if $u = v$. Therefore $f \amalg g$ is injective if $f$ and $g$ are.

To see that $f \times g$ is 1–1, suppose that we have $f \times b(a, c) = f \times g(a', c')$. By definition of $f \times g$ we conclude that

$$\big(\,f(a),\, g(c)\,\big) \quad = \quad \big(\,f(a'),\, g(c')\,\big)$$

and since ordered pairs are determined by their coordinates the latter implies that $f(a) = f(a')$ and $g(c) = g(c')$. Since $f$ and $g$ are injective this implies that $a = a'$ and $c = c'$ so that $(a, c) = (a', c')$ and hence $f \times g$ is injective if $f$ and $g$ are.■

**2.** Choose $A$ so that $|A| = \alpha$. Then $A \times \emptyset = \emptyset$ implies that $\alpha \cdot 0 = 0$. Also $\alpha^1$ is the cardinality of the set of all functions from $\{1\}$ to $A$, which is in 1–1 correspondence with $A$ under the mapping which sends $f : \{1\} \to A$ to the value $f(1) \in A$. Therefore we have $\alpha^1 = |A| = \alpha$. Finally, $1^\alpha$ is the cardinality of the set of all functions from $A$ to $\{1\}$, and since there is a unique function of this type (the function whose value at every element of $A$ is equal to 1), it follows that $1^\alpha = 1$.■

**3.** We shall split the proof into several steps.

$(i)$ Suppose that $\alpha = \aleph_0$ or $2^{\aleph_0}$. Prove that $\alpha^\alpha = 2^\alpha$. Suppose that $\alpha = \aleph_0$ and $\beta = 2^\alpha$. Then $\beta^\alpha = 2^\alpha$.

$(ii)$ Let $X$ be a set, and let $\Sigma(X)$ denote the set of bijections from $X$ to itself. Suppose that $\varphi : X \to Y$ is a bijection of sets. Prove that there is a bijection $\varphi_* : \Sigma(X) \to \Sigma(Y)$ such that $\varphi_*(h) = \varphi \circ h \circ \varphi^{-1}$ for all $h \in \Sigma(X)$.

$(iii)$ Suppose that $|X| = \alpha$, where $\alpha = \aleph_0$ or $2^{\aleph_0}$. Prove that $|\Sigma(X)| = \alpha^\alpha = 2^\alpha$.

*Proof of* $(i)$. We shall prove the statements in the individual sentences separately. For the first sentence, we have $2^\alpha \leq \alpha^\alpha$; since $\alpha \cdot \alpha = \alpha$ for these cardinal numbers we also have

$$\alpha^\alpha \quad \leq \quad (2^\alpha)^\alpha \quad = \quad 2^{\alpha \cdot \alpha} \quad = \quad 2^\alpha$$

and hence the Schröder-Bernstein Theorem implies that the left and right hand sides are equal.■

IMPORTANT GENERALIZATION. This argument works for an infinite cardinal number $\alpha$ if we know that $\alpha \cdot \alpha = \alpha$. By the results of Section VII.4 this equation holds for all infinite cardinal numbers, and therefore it follows that **the conclusion of Part** $(i)$ **is true for all infinite cardinal numbers.**■

*Proof of* $(ii)$. It follows immediately that the construction $\Phi_*(f) = \varphi \circ f \circ \varphi^{-1}$ defines a mapping of function sets from $\mathbf{F}(X, X)$ to $\mathbf{F}(Y, Y)$. We need to show that it sends the subset $\Sigma(X)$ to $\Sigma(Y)$. In other words, if $f$ is a bijection we need to check that $\varphi \circ f \circ \varphi^{-1}$ is also a bijection. But this follows immediately because the latter is a composite of bijections and a composite of bijections is also a bijection.

Let $\psi : Y \to S$ be the inverse of $\varphi$. Then by the same reasoning as above we have a map $\psi_* : \Sigma(Y) \to \Sigma(X)$, and it will suffice to show that the composites $\psi_* \circ \varphi_*$ and $\varphi_* \circ \psi_*$ are both identity mappings. Consider the following chains of equations:

$$\psi_* \circ \varphi_*(f) \quad = \quad \psi \circ \varphi \circ f \circ \varphi^{-1} \circ \psi^{-1} \quad = \quad \psi \circ \varphi \circ f \circ \psi \circ \varphi \quad = \quad \mathrm{id}_X \circ f \circ \mathrm{id}_X \quad = \quad f \quad = \quad \mathrm{Identity}(f)$$

$$\varphi_* \circ \psi_*(g) \quad = \quad \varphi \circ \psi \circ g \circ \psi^{-1} \circ \varphi^{-1} \quad = \quad \varphi \circ \psi \circ g \circ \varphi \circ \psi \quad = \quad \mathrm{id}_Y \circ g \circ \mathrm{id}_Y \quad = \quad g \quad = \quad \mathrm{Identity}(g)$$

It follows that $\varphi_*$ is bijective and $\psi_*$ is its inverse.■

*Proof of* $(iii)$. Assume that $X$ is either $\mathbf{N}$ or $\mathbf{R}$. Since $\Sigma(X) \subset \mathbf{F}(X, X)$ by definition, it follows that if $|X| = \alpha$ then $|\Sigma(X)| \leq \alpha^\alpha$. By the Schröder-Bernstein Theorem it will suffice to prove the reverse inequality.

Define a map $\sigma : \mathbf{F}(X, X) \to \Sigma(X \times X)$ so that for each $u : X \to X$ we have the following description of $\sigma_u \in \Sigma(X \times X)$:

[1] $\sigma_u(y, 0) = (\,y,\, u(y)\,)$

[2] $\sigma_u(\,y,\, u(y)\,) = (y, 0)$

[3] $\sigma_u(y, z) = (y, z)$ otherwise.

In words, $\sigma_u$ interchanges elements of the form $(y, 0)$ with elements of the form $(\,y.\, u(y)\,)$ and leaves everything else fixed.— Note that if $X = \mathbf{R}$ the mapping $\sigma_u$ is almost never going to be continuous. By construction each $\sigma_u$ defines a mapping from $X$ to itself, and in fact the composites $\sigma_u \circ \sigma_u$ are all equal to the identity map on $X \times X$. This shows that each $\sigma_u$ is a bijection, and in fact each such map is equal to its own inverse.

We now need to show that $\sigma$ is an injection. However, if $\sigma_u = \sigma_v$ then for every $y \in X$ we have $\sigma_u(y, 0) = \sigma_v(y, 0)$, which implies that $u(y) = v(y)$; it follows that if $\sigma_u = \sigma_v$, then $u = v$ as required.

The preceding argument shows that $\alpha^\alpha \leq |\Sigma(X \times X)|$. Since $\alpha \cdot \alpha = \alpha$ for the sets $X$ we are considering, we may now apply $(i)$ to conclude that $|\Sigma(X \times X)| = |\Sigma(X)|$, and therefore we also have $\alpha^\alpha \leq |\Sigma(X)|$. As noted previously in this exercise, we may now conclude that equality actually holds in the latter expression. Finally, we may now apply Part $(i)$ to see that $2^\alpha = |\Sigma(X)|$ also holds.■

**Determination of $|\Sigma(X)|$ for an arbitrary set $X$.** More generally, it is possible to describe $|\Sigma(X)|$ as a function of $|X|$ in a very straightforward manner. Not surprisingly, the finite and transfinite cases must be handled separately.

*The finite case.* If $X$ is finite and $|X| = n > 0$, then there is a 1–1 correspondence between $\Sigma(X)$ and the symmetric group $\Sigma_n$ of permutations of $\{1, 2, \cdot, n\}$. It is well known that $\Sigma_n$ contains $n!$ elements. Further information on this may be found in Section 4.3 of Rosen, and particularly on page 321.

*The transfinite — or infinite — case.* The proof of the preceding exercise is valid for all infinite sets $X$ whose cardinal numbers satisfy $|X \times X| = |X|$. At one point in the argument we defined a map using $0 \in \mathbf{R}$, but in general one can carry out the construction replacing 0 by some arbitrary fixed element $x_0 \in X$. As noted above, by the results of Section VII.4 the identity $|X \times X| = |X|$ holds for all infinite sets, and consequently the proof implies that for every infinite set $X$ we have $|\Sigma(X)| = 2^{|X|} = |X|^{|X|}$.

*Relations between the finite and transfinite cases.* There is a loose connection between the computations for $|\Sigma(X)|$ in the finite and transfinite cases ($n!$ versus $\alpha^\alpha$) in terms of a classical asymptotic formula for estimating $n!$ discovered by A. de Moivre (1667–1754) and J. Stirling (1692–1770), which is usually known as **Stirling's Formula**:

$$\lim_{n \to \infty} \frac{n!}{n^n \sqrt{2\pi n}\, e^{-n}} \quad = \quad 1$$

A discussion of this formula and a relatively elementary derivation of it may be found at the following online site:

`http://en.wikipedia.org/wiki/Stirling's_formula`

The formula implies that for all large values of $n$ the percentage error for estimating $n!$ by the denominator goes to 0 as $n \to \infty$. However, for several reasons it would be stretching things too far to assert that the results $|\Sigma_n| = n!$ and $|\Sigma(\mathbf{N})| = |\mathbf{N}|^{|\mathbf{N}|}$ somehow "fit together continuously" in some precise matehamtical sense.■

**4.** If $\mathbf{c} = |\mathbf{R}|$ then we have

$$\mathbf{c} \quad \leq \aleph_0 \cdot \mathbf{c} \quad \leq \quad \mathbf{c} \cdot \mathbf{c} \quad = \quad \mathbf{c}$$

and therefore it is enough to show that for each nonnegative integer $n$ the set of all subsets with $n$ elements has cardinality $\mathbf{c}$, and likewise for the set of all countably infinite subsets.

Let $n$ be a positive integer. As in Exercise VI.3.1, define a map from subsets of $\mathbf{R}$ with $n$ elements into $\mathbf{R}^n$ such that the first coordinate is the least element of the set, the second coordinate is the smallest of the remaining elements, and so on. We can always find least elements for finite subsets because the real numbers are linearly ordered. This shows that the cardinality of the set of subsets with $n$ elements is less than or equal to the cardinality of $\mathbf{R}^n$, which is $\mathbf{c}$. On the other hand, given a real number $r_0$, it is easy to find a subset with $n$ elements whose least element is $r_0$, so this gives a 1–1 mapping from $\mathbf{R}$ into the set of all such subsets (specifically, let the second element be $r_0 + 1$, *etc.*). Therefore the cardinality of the set of all subsets with exactly $n$ elements is $\mathbf{c}$ provide $n$ is a positive integer. Of course, if $n = 0$ this cardinality is 1. It then follows that the set of all finite subsets of $\mathbf{R}$ has cardinality equal to $\aleph_0 \cdot \mathbf{c} + 1 = \mathbf{c}$.

To complete the proof it will suffice to show that the set $\mathbf{D}$ of all countably infinite subsets of $\mathbf{R}$ also has cardinality $\mathbf{c}$. We can define a 1–1 map from $\mathbf{R}$ into this set as before, sending $r_0$ to the set of all numbers of the form $r_0 + k$ where $k$ is a nonnegative integer. This mapping is injective because each set in the range has a least element and for different real numbers one obtains different least elements.

Thus it only remains to show the cardinality of this set of subsets is less than or equal to $\mathbf{c}$.

Suppose that $B$ is a countably infinite subset of $\mathbf{R}$. Then there is a 1–1 correspondence from $\mathbf{N}$ to $B$, so we pick such a mapping $h_B : \mathbf{N} \to B$ (*we are using the Axiom of Choice to do this*). We may now compose this chosen bijection with inclusion to obtain a mapping $f_B$ from $\mathbf{N}$ into mapping from $\mathbf{D}$ to $\mathbf{F}(\mathbf{N}, \mathbf{R})$.

If we take different subsets we obtain different mappings because their ranges are unequal, and this means that there is a 1–1 map from $\mathbf{D}$ to $\mathbf{F}(\mathbf{N}, \mathbf{R})$, so that $|\mathbf{D}| \leq |\mathbf{R}|^{|\mathbf{N}|}$. Since we already know that $\mathbf{c} \leq |\mathbf{D}|$, everything reduces to proving that $|\mathbf{R}|^{|\mathbf{N}|} = \mathbf{c}$. This is a consequence of the following chain of equations:

$$\left(2^{\aleph_0}\right)^{\aleph_0} \quad = \quad 2^{\aleph_0 \cdot \aleph_0} \quad = \quad 2^{\aleph_0}$$

It follows that the set of countably infinite subsets of $\mathbf{R}$ has the same cardinality as $\mathbf{R}$ itself.■

**5.** This is very similar to the preceding example. Let $X$ denote the set of continuous real valued functions on the unit interval, and let $Y$ denote the set of functions defined at rational points of that interval. There is a natural map from $X$ to $Y$ defined by restricting to the rational points of the interval, and the statement in the exercise means that this mapping is injective. The reasoning of the previous problem shows that $|Y| = \mathbf{c}$ and hence that $|X| \leq \mathbf{c}$. On the other hand it is easy to show that $\mathbf{c} \leq |X|$; for example, we may define a 1–1 mapping from $\mathbf{R}$ into $X$ sending $r$ to the constant function whose value at every point is equal to $r$. Therefore it follows that $|X| = \mathbf{c}$.■

## VI.6 : Transfinite induction and recursion

*Exercises to work*

**1.** It will be necessary to assume Axiom of Choice and Well-Ordering Principle from Section VII.1 of the lecture notes.

Let $A$ be the original partially ordered set, and let $P$ be a well-ordered set which is in 1–1 correspondence with $\mathbf{P}(A)$. Let $A^+ = A \cup \{A\}$ and extend the partial ordering on $A$ to $A^+$ by making $A \in A^+$ the maximal element. We shall define a nondecreasing map $f : P \to A^+$ by transfinite recursion such that $f$ is strictly increasing on $f^{-1}[A]$.

Denote the minimal element of $P$ by 0, and define $f(0)$ by picking a point in $A$ using a choice function. Suppose now that we have defined $f(\beta)$ for all $\beta < \alpha$; we need to define $f(\alpha)$. There are two cases. If there is some $z \in A$ such that $z > f(\beta)$ for all $\beta < \alpha$, define $f(\alpha)$ by choosing such a value of $z$ (again, this requires a choice function). If no such value of $z$ exists, let $f(\alpha) = A$.

Let $B = f[\,f^{-1}[A]\,]$; since $P$ is well-ordered and $f$ is strictly increasing on on $f^{-1}[A]$, it follows that $B$ is a well-ordered subset of $A$. Thus it will suffice to show that $B$ is cofinal in $A$. Suppose that $x \in A$; we need to show that there is some $b \in B$ such that $b > x$. Assume this does not hold for some particular choice of $x$. If this happens then the recursive definition yields a strictly increasing map from $P$ into $A$, and in fact the image is contained in the set of all elements less than $x$. Since $f$ is strictly increasing it follows that $|P| \leq |A|$. However, by construction we have $|P| = |\mathbf{P}(A)| > |A|$, which yields a contradiction. This means that for each $x \in A$ there must be some $b$ such that $b > x$, so that $B$ is a cofinal well-ordered subset.■

**2.** Let $A$ be the linearly ordered subset. If $A$ is well-ordered, then the conclusion of the exercise is true because every nonempty subset of a well-ordered set is well-ordered. — Conversely, suppose that for each $x \in A$ the set of all strict predecessors of $A$ is well-ordered, and let $B$ be a nonempty subset of $A$. We need to show that $B$ has a minimal element. Let $b_0 \in B$; if $b_0$ is a minimal element of $B$, we are done. On the other hand, if $b_0$ is not a minimal element and $L(b_0)$ is the set of strict predecessors of $b_0$, then $B \cap L(b_0)$ is nonempty, and since $L(b_0)$ is well-ordered it follows that $B \cap L(b_0)$ has a minimal element $b_1$. We claim that $b_1$ is a minimal element of $B$. For each $y \in B$ we have $y = b_0$, $y > b_0$ or $y < b_0$. In the first two cases we have $y \geq b_0 > b_1$, and in the last case we have $b_1 \leq y$ because then $y \in B \cap L(b_0)$ and $b_1$ is a minimal element of the intersection.■

**3.** Let $A$ be a well-ordered set, and let $A^{\mathbf{op}}$ denote $A$ with the reverse ordering. If $A$ is infinite, then $A$ contains a subset that has the same order type as

$$\omega \ := \ \{\, 0 \ < \ 1 \ < \ 2 \ < \ 3 \ < \ 4 \ < \ 5 \ < \ 6 \ \cdots \}$$

and since $\omega^{\mathbf{op}}$ does not have a minimal element it follows that $A^{\mathbf{op}}$ is not well-ordered. — Suppose now that $A$ is finite. In order to prove that $A^{\mathbf{op}}$ is well-ordered, we need to show that every nonempty subset of $A$ has a maximal element.

We shall prove this by induction on $|A|$. If $|A| = 0$ the statement is vacuously true. Similarly, if $|A| = 1$, then $A$ has a unique nonempty subset, and its unique element is a maximal element. Suppose now that we know the result for $|A| = n \geq 1$, and suppose that $B$ is a well-ordered set with $(n+1)$ elements. Let 0 be the minimal element of $B$, and let $B_1 = B - \{0\}$. Given a nonempty subset $C \subset B$, let $C_1 = C \cap B_1$. If $C_1$ is nonempty, then by the induction hypothesis it follows that $C_1$ has a maximal element $m$. Since $C \subset C_1 \cup \{0\}$ and $0 < m$, it follows that $m$ is also a maximal element of $C$. On the other hand, if $C_1 = \emptyset$ then $C$ must be equal to $\{0\}$ and hence 0 is the maximal element of $C$.■

# VII. The Axiom of Choice and related topics

## VII.1 : Nonconstructive existence statements

*Exercises to work*

**1.** Define $g : B \to A$ on $f[A]$ by taking $b \in f[A]$ and picking $g(b) \in A$ such that $f(g(b)) = b$, and define $g$ on $B - f[A]$ by setting $g(b) = z$ for some chosen element $z \in A$. We need to show that $f = f \circ g \circ f$. By construction, if $a \in A$, then $g(f(a))$ satisfies $f(g(f(a))) = f(a)$, so the condition $f = fgf$ is satisfied.■

**2.** If $|A| \leq |B|$ then there is a 1–1 mapping $f : A \to B$, and by Exercise III.4.13 there is a mapping $g : B \to A$ such that $g \circ f = \mathrm{id}_A$. The mapping $g$ is onto because $a \in A$ can be written as $g(b)$ where $b = f(a)$. — Conversely, if there is a surjection $f : B \to A$, then by the Axiom of Choice there is a function $s : A \to B$ such that $s(a) \in f^{-1}[\{a\}]$ for all $a \in A$. It follows that $f \circ s(a) = a$ for all $a$. We claim that $s$ is 1–1; if $s(u) = s(v)$, then $u = f(s(u)) = f(s(v)) = v$. Therefore we have $|A| \leq |B|$.■

**3.** To see that $W_r \cap W_q = \emptyset$ if $q \neq r$, observe that the second coordinates of elements in $W_r$ are all equal to $r$ while the second coordinates of elements in $W_r$ are all equal to $q$. Therefore the second coordinates of elements of $W_r$ and $W_q$ are distinct, so that $W_r \cap W_q = \emptyset$. — The union of the sets $W_q = \cup_k Y_k \times \{k\}$ is equal to $\amalg_k Y_k$ by the definition of the latter. — For each $q$ there is a 1–1 correspondence between $Y_q$ and $W_q = Y_q \times \{q\}$ sending $y$ to $(y, q)$.■

**4.** For each $k$ we are given a bijection $f_k : Y_k \to V_k$; denote the respective inverses by $g_k$. If we define $f : \amalg_k Y_k \to \amalg_k V_k$ by $f(y, k) = (f_k(y), k)$, then the map $g : \amalg_k V_k \to \amalg_k Y_k$ defined by $g(v, k) = (g_k(v), k)$ satisfies $f \circ g = \mathrm{id}$ and $g \circ f = \mathrm{id}$, so that $g$ is an inverse to $f$ and both maps are bijections.■

## VII.2 : Extending partial orderings

*Exercises to work*

**1.** The standard alphabetical ordering is a linear ordering that contains the given partial ordering. One way to visualize this is to move the pieces of the Hasse diagram slightly so that $a$ falls below $b$, etc. — One can check this more methodically by constructing a matrix whose rows and columns correspond to the points of the original set in alphabetical order. Saying that the usual alphabetical ordering contains the given one amounts to saying that all ordered pairs for which the original relation holds must lie on or above the main diagonal. The following chart indicates this. In the latter some elements on or above the main diagonal are marked with numbers. If a number appears in the position $(x, y)$, it means that $xRy$ and there is a chain of length $k$ such that $xRx_1R\ \cdots\ Rx_k = y$; if nothing appears, then no such chain exists.

| | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ | $g$ | $h$ | $i$ | $j$ | $k$ | $l$ | $m$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $a$ | 0 | | | 1 | | | | 2 | 2 | 3 | 3 | 4 | 4 |
| $b$ | | 0 | | 1 | 1 | | | 2 | 2 | 3 | 3 | 4 | 4 |
| $c$ | | | 0 | | | 1 | 2 | | | | 3 | 4 | 4 |
| $d$ | | | | 0 | | | | 1 | 1 | 2 | 2 | 3 | 3 |
| $e$ | | | | | 0 | | | 1 | | 2 | 2 | 3 | 3 |
| $f$ | | | | | | 0 | 1 | | | | 2 | 3 | 3 |
| $g$ | | | | | | | 0 | | | | 1 | 2 | 2 |
| $h$ | | | | | | | | 0 | | 1 | 1 | 2 | 2 |
| $i$ | | | | | | | | | 0 | 1 | | 2 | |
| $j$ | | | | | | | | | | 0 | | 1 | |
| $k$ | | | | | | | | | | | 0 | 1 | 1 |
| $l$ | | | | | | | | | | | | 0 | |
| $m$ | | | | | | | | | | | | | 0 |

**2.** The usual ordering works because if $a$ and $b$ are positive integers such that $a$ divides $b$, then $a \leq b$.■

**3.** One way of finding a suitable linear ordering is to draw a Hasse diagram as in Exercise 1. The file `hasse-VII-2-3.JPG` in the online directory depicts one such possibility; namely, the linear ordering given by the following chain:

$$A < G < B < L < C < H < D < M < K < E < F$$

As before, one way to visualize this is to move the pieces of the Hasse diagram slightly. More methodically, this can be checked by constructing a matrix whose rows and columns correspond to the points of the original set in the order listed above. Saying that the new linear ordering contains the given partial ordering amounts to saying that all ordered pairs for which the original relation holds must lie on or above the main diagonal.

Here is the chart which corresponds to the one in Exercise 1; the notation for the entries of the chart is the same as for the earlier exercise.

| | $A$ | $G$ | $B$ | $L$ | $C$ | $H$ | $D$ | $M$ | $K$ | $E$ | $F$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $A$ | 0 | 1 | 1 | 1 | 2 | 2 | 3 | 2 | 3 | 4 | 3 |
| $G$ | | 0 | | | 1 | 1 | 2 | | 3 | 3 | 4 |
| $B$ | | | 0 | | 1 | | 2 | | | 3 | 4 |
| $L$ | | | | 0 | | | | 1 | | | 2 |
| $C$ | | | | | 0 | | 1 | | | 2 | 3 |
| $H$ | | | | | | 0 | 1 | | 1 | 2 | 2 |
| $D$ | | | | | | | 0 | | | 1 | 2 |
| $M$ | | | | | | | | 0 | | | 1 |
| $K$ | | | | | | | | | 0 | | 1 |
| $E$ | | | | | | | | | | 0 | 1 |
| $F$ | | | | | | | | | | | 0 |

**4.** The main thing to do is to order the subsets so that the subsets with $p$ elements come before the subsets with $q$ elements if $p < q$. For example, this can be done as follows:

$$\emptyset \;<\; \{1\} \;<\; \{2\} \;<\; \{3\} \;<\; \{1,2\} \;<\; \{1,3\} \;<\; \{2,3\} \;<\; \{1,2,3\}$$

One interesting exercise might be to determine exactly how many linear orderings contain the given partial ordering.■

**5.** This is similar to Exercise 1. Once again, the standard alphabetical ordering is a linear ordering that contains the given partial ordering, and one way to visualize this is to move the pieces of the Hasse diagram slightly so that $a$ falls below $b$, etc. — Once again, it is possible to check this more methodically by constructing a matrix whose rows and columns as in Exercise 1. Here is what one obtains for the exercise we are now considering:

| | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ | $g$ | $h$ | $i$ | $j$ | $k$ | $l$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $a$ | 0 | | | | 1 | | | 2 | | 3 | | 4 |
| $b$ | | 0 | | | 1 | 1 | | 2 | 2 | 3 | 3 | 4 |
| $c$ | | | 0 | | | 1 | 1 | | 2 | 3 | 3 | 4 |
| $d$ | | | | 0 | | | 1 | | 2 | 3 | 3 | 4 |
| $e$ | | | | | 0 | | | 1 | | 2 | | 3 |
| $f$ | | | | | | 0 | | | 1 | 2 | 2 | 3 |
| $g$ | | | | | | | 0 | | 1 | 2 | 2 | 3 |
| $h$ | | | | | | | | 0 | | 1 | | 2 |
| $i$ | | | | | | | | | 0 | 1 | 1 | 2 |
| $j$ | | | | | | | | | | 0 | | 1 |
| $k$ | | | | | | | | | | | 0 | 1 |
| $l$ | | | | | | | | | | | | 0 |

## VII.3 : Equivalence proofs

*Exercises to work*

**1.** Let **F** be a (nonempty) family of subsets of $A$ of finite character. We want to apply Zorn's Lemma to this family.

In order to do so, we need to show that linearly ordered subsets of **F** have upper bounds in **F**. Suppose that $\mathbf{L} \subset F$ is a linearly ordered subfamily consisting of the sets $L_\alpha$. We claim that $M = \cup_\alpha L_\alpha$ also belongs to **F**, and we shall prove this using the finite character assumption.

Suppose that $C \subset M$ is finite with elements $c_1 , \cdots , c_k$. Then we can find $L_j \in \mathbf{L}$ such that $c_j \in L_j$ for all $j$. Given any finite subset of a linearly ordered set, it is always possible to find a maximal element; applying this to the present situation, we can find some $m$ such that $L_m$ contains $L_j$ for all $j$. It follows that $C \subset L_j$, so the finite character assumption implies that $C \in \mathbf{F}$. Thus we have shown that every finite subset of $M$ belongs to **F**, and since the latter has finite character it follows that $M$ itself belongs to **F**. As noted before, one can now apply Zorn's Lemma to find a maximal subset in **F**.■

**2.** The hypotheses should have included an assumption that all the subsets in $\mathbf{F}$ are nonempty (the conclusion makes no sense if $A$ is the empty set). — Let $P_+(S)$ be the set of all nonempty subsets of $S$, and let $c : P_+(S) \to S$ be a choice function on $P_+(S)$. Let $C \subset S$ be the image of $c|\mathbf{F}$. Suppose now that $A \subset \mathbf{F}$. Then we know that $c(a) \in C \cap A$. On the other hand, if $x \in C \cap A$, then $x = c(b)$ for some subset $B$, and it follows that $x \in B$ as well. Since $A \cap B = \emptyset$ if $A \neq B$ (this is the condition on $\mathbf{F}$), it follows that $x$ must be equal to $c(a)$, and therefore we know that $C \cap A = \{ c(a) \}$.∎

## VII.4 : Additional consequences

*Exercises to work*

**1.** Once again we shall follow the hint, starting by choosing $X_i$ and $Y_j$ such that $|X_i| = \alpha_i$ and $|Y_j| = \beta_j$. We need to prove that there is no surjection from $\coprod_i X_i$ to $\prod_i Y_i$. In other words, given an injective mapping $f : \coprod_i X_i \to \prod_i Y_i$, we need to find a point in the codomain which does not lie in the image of $f$.

For each $i$ let $h_i$ be the composite

$$X_i \longrightarrow \coprod_j X_i \longrightarrow \prod_j Y_j \longrightarrow Y_i$$

where the first map is the standard injection of $X_i$ into the disjoint union and the last map is the projection onto $Y_i$. The cardinality inequality implies that $h_i$ cannot be surjective, so there is some $y_i \in Y_i$ which does not lie in the image of $h_i$. Let $\mathbf{y} \in \prod_i Y_i$ be the element whose coordinates are given by the corresponding elements $y_i$.

However, if $\mathbf{y}$ did lie in this image, then for some $k$ in the indexing set $J$ the element $\mathbf{y}$ would be the image of an element coming from $X_k$. This would mean that the coordinate $y_k$ would be equal to $h_k(x)$, where $x \in X_k$ and the image of $x$ in $\amalg_j X_j$ maps to $\mathbf{y}$. Since $y_k$ does not lie in the image of $h_k$ by construction, it follows that $\mathbf{y}$ cannot lie in the image of $f$. Therefore we know that

$$\sum_i \alpha_i \not\geq \prod_i \beta_i$$

and by the linear ordering property for cardinal numbers it follows that the number on the left hand side is strictly less than the number on the right hand side.∎

**2.** Under the conditions of this exercise one can prove the following weaker inequality:

$$\sum_i \alpha_i \leq \prod_i \beta_i$$

PROOF. As suggested by the hint, the first step is to note that if $\alpha$ is an infinite cardinal number, then we have

$$\alpha \leq \alpha + 1 \leq \alpha + \aleph_0 = \alpha$$

because $\alpha + \aleph_0 = \alpha$ for every infinite cardinal number, so that $\alpha + 1 = \alpha$ also holds.

As in the preceding exercise, choose sets $X_i$ and $Y_j$ such that $|X_i| = \alpha_i$ and $|Y_j| = \beta_j$. As elsewhere, we let $\sigma(C) = C \cup \{C\}$ and use the fact that $C \notin C$ to conclude that $\sigma(C)$ is given by $C$

plus some point which does not lie in $C$. Since all cardinal numbers in sight are infinite, it follows that for all $i$ the sets $\sigma(X_i)$ amd $\sigma(Y_i)$ have the same cardinalities as $X_i$ and $Y_i$ respectively. The assumption that $\alpha_i \leq \beta_i$ for all $i$ implies that we have injections $g_i : X_i \to Y_i$.

We shall now define a mapping

$$F : \prod_j \sigma(X_j) \quad \longrightarrow \quad \prod_j \sigma(Y_j)$$

such that on the image of each $X_i$ the map takes the point corresponding to $x \in X_i$ to the point in $\prod_j Y_j$ whose $j^{\text{th}}$ coordinate equals $g_i(x)$ if $j = i$ and $X_j$ if $j \neq i$. Since the images of $X_u$ and $X_v$ are disjoint if $u \neq v$, it follows that we obtain a well defined function in this manner.

It will suffice to prove that $F$ is injective. Suppose that $F(p) = F(q)$. By definition, for all but one choice of indexing variable, the $k^{\text{th}}$ coordinate of $F(p)$ is equal to the extra point $X_k \in \sigma(X_k)$, and a similar statement holds for $F(q)$. Therefore the exceptional coordinate is the same for both $p$ and $q$. However, if $\ell$ is this exceptional coordinate, then by construction $p$ and $q$ both lie in the image of $X_\ell$. The latter implies that $F(p) = F(q)$ if and only if $g_\ell(p) = g_\ell(q)$. Since $g_\ell$ is injective, it follows that $p = q$. Therefore $F$ is 1–1 as required.■

**3.** 9*i*) Let $A, B, C$ be sets such that $|A| = \alpha$, $|B| = \beta$, and $|C| = \gamma$. The condition $\alpha \leq \beta$ means there is an injection $j : A \to B$. Define an associated map of function sets

$$j_\# : \mathbf{F}(C, A) \quad \longrightarrow \quad \mathbf{F}(C, B)$$

by the formula $j_\#(f) = j \circ f$. The assertion about cardinal numbers will follow if we can prove that $j_\#$ is injective.

Suppose that $f_1, f_2 \in \mathbf{F}(C, A)$ satisfy $j_\#(f_1) = j_\#(f_2)$. Then $j \circ f_1 = j \circ f_2$, so that $j(f_1(x)) = j(f_2(x))$ for all $x \in C$. Since $j$ is injective, this means that $f_1(x) = f_2(x)$ for all $x \in C$, which in turn implies that $f_1 = f_2$. Therefore the map $j_\#$ is injective as required.■

(*ii*) One way to do this is to start with the special cases where $\alpha$ and $\beta$ are (finite) powers of 2. More precisely, if $\beta = 2^k$ for some $k \geq 1$ we shall prove that $\beta^\gamma = 2^\gamma$.

By the transfinite laws of exponents the left hand side is equal to $2^{k\gamma}$, and since $k \cdot \gamma = \gamma$ for every positive integer $k$ and transfinite cardinal $\gamma$, the desired conclusion follows immediately when $\beta$ is a positive integral power of 2.

For general choices of $\beta \geq 2$ we can find powers of 2 such that

$$2^p \quad = \quad \beta_0 \quad \leq \quad \beta \quad \leq \quad \beta_1 \quad = \quad 2^q$$

and if we combine this with the first part of the exercise, we see that

$$2^\gamma \quad \leq \quad \beta_0^\gamma \quad \leq \quad \beta^\gamma \quad \leq \quad \beta_1^\gamma \quad = \quad 2^\gamma$$

where the first and last equations follow from the discussion in the preceding paragraph. We can now use the Schröder-Bernstein Theorem to conclude that $\beta^\gamma = 2^\gamma$.■

**4.** Suppose that we have an infinite non-limit ordinal $\mu + 1$ and a 1–1 correspondence between the corresponding set $S[\mu + 1]$ and some other set $X$. Then we have the cardinal number identities

$$|X| \quad = \quad \big|S[\mu + 1]\big| \quad = \quad \big|S[\mu]\big| \; + \; 1 \quad = \quad \big|S[\mu]\big|$$

so that there is a 1–1 correspondence between $X$ and the elements of the ordinal number $\mu$. Therefore $\mu + 1$ cannot be the least ordinal for which there is a 1–1 correspondence, and it follows that if $\lambda_0$ is the minimum such ordinal, then $\lambda_0$ must be a limit ordinal.■

**5.** All finite ordinals are cardinal numbers, and the first infinite ordinal $\omega$ is equal to $\aleph_0$. Thus the next infinite ordinal, namely $\omega + 1$, is the first ordinal number that is not also a cardinal number.■

**6.** We shall solve this in a sequence of steps:

*First, we shall use a previous exercise to prove the result when* $|A| \leq 2^{\aleph_0}$.

Let $D(A)$ denote the countably infinite subsets. Since $A$ contains a countably infinite subset $B$, it follows that $2^{\aleph_0} = |D(B)| \leq |D(A)|$, and since there is an injective mapping from $A$ to $\mathbf{R}$ it also follows that $|D(A)| \leq D(\mathbf{R})| = 2^{\aleph_0}$, Therefore the number of countably infinite subsets is $2^{\aleph_0}$ by the Schröder-Bernstein Theorem.

*From this point on assume that* $|A| \geq 2^{\aleph_0}$. *Next, we shall prove that the set of countably infinite subsets is in* $1-1$ *correspondence with the set* $A^{\mathbf{N}}$ *of functions from* $\mathbf{N}$ *to* $A$ *as follows: Given a countably infinite subset* $E$ *and a specific* $1-1$ *correspondence from* $\mathbf{N}$ *onto* $E$ *we shall obtain a map from* $\mathbf{N}$ *to* $A$ *that turns out to be* $1-1$.

Since the image of the map associated to a subset $B$ is equal to $B$ by construction, it follows that different subsets determine functions with different images. Thus the functions must also be different.

*By the Schröder-Bernstein Theorem it will be enough to define a map from* $A^{\mathbf{N}}$ *to countably infinite subsets of A. There is a* $1-1$ *map from such functions to countable subsets of* $\mathbf{N} \times A$ *given by taking the graphs of functions. We shall use this to define the desired map to countable subsets of A? A comparison of* $|A|$ *and* $|\mathbf{N} \times A|$ *will be helpful here.*

Since $A$ is infinite, the rules for transfinite cardinal arithmetic imply that $|A| = |\mathbf{N} \times A|$. Thus it is also enough to prove that the number of countably infinite subsets of $\mathbf{N} \times A$ is at least as large as the cardinality of $A^{\mathbf{N}}$. But given two functions from $\mathbf{N}$ to $A$, their images are distinct countably infinite subsets of the product $\mathbf{N} \times A$. Note that the graphs are always infinite because for each $n \in \mathbf{N}$ we have a point in $\mathbf{N} \times A$ whose first coordinate is equal to $n$.

*Finally, we shall use a modified version of the Zorn's Lemma argument proving* $\alpha \cdot \alpha = \alpha$ *for infinite cardinals* $\alpha$ *to prove that* $\alpha^\omega = \alpha$. *Specifically, we shall consider the collection of all pairs* $(B, \varphi)$ *consisting of* $B \subset A$ *satisfying* $|B| \geq 2^{\aleph_0}$ *and a bijection* $\varphi : B^{\mathbf{N}} \to B$, *with a partial ordering such that* $(B, \varphi) \leq (D, \psi)$ *if and only if* $B \subset D$ *and* $\psi = \varphi$ *on* $B^{\mathbf{N}}$. *We may use the previously established fact that* $2^{\mathbf{c}} = \mathbf{c}^{\mathbf{c}}$ *(where* $\mathbf{c} = 2^{\aleph_0}$*) from the proof of Exercise* **rm VI.4.3** *to show this set is nonempty.*

By assumption $A$ has a subset $B$ such that $|B| = |\mathbf{R}|$, and we know there is a 1–1 correspondence between $\mathbf{R}$ and $\mathbf{R}^{\mathbf{N}}$ by the earlier exercise.

*We shall verify that Zorn's Lemma applies and hence there is a maximal pair, say* $(B, \varphi)$.

Given a linearly ordered collection $(B_\alpha, \varphi_\alpha)$, we need to show that their union belongs to the given collection. Let $B^* = \cup B_\alpha$; then it follows immediately that one obtains a well defined mapping $\varphi^* : (B^*)^{\mathbf{N}} \to B^*$ from the mappings $\varphi_\alpha$, so one needs to check that this map is bijective. To see it is injective, suppose $x, y \in (B^*)^{\mathbf{N}}$. Then there is some $\alpha$ such that $x, y \in (B_\alpha)^{\mathbf{N}}$, and $\varphi^*(x) = \varphi^*(y)$ implies $\varphi_\alpha(x) = \varphi_\alpha(y)$. Since $\varphi_\alpha$ is injective, this means $x - y$. To see that $\varphi^*$ is surjective, let $z \in B^*$, so that $z \in B_\alpha$ for some $\alpha$ and hence lies in the image of $\varphi_\alpha$, which means it also lies in the image of $\varphi^*$. Thus we have shown that in our partially ordered set, every linearly ordered subset has an upper bound, and this means that Zorn's Lemma applies.

*If $|B| = |A|$ we are done, so suppose instead that $|B| < |A|$. In this case we shall show there is some $C \subset A$ such that $C \subset A - B$ and $|C| = |B|$.*

If this were false then the cardinality of every subset of $A - B$ would be strictly less than $|B|$, and in particular $|A - B| < |A|$, so that $|A|$ would be less than $|B|$, which is less than $|A|$. This contradiction shows that there must be some subset of $A - B$ whose cardinality is equal to $|B|$. In fact, one has $|A - B| = |A| > |B|$ in our situation, but we shall not need the full strength of this conclusion.

*We now explain why $(B \cup C)^{\mathbf{N}}$ can be written as a union of pairwise disjoint subsets $S_Y$, where $Y$ runs over all subsets of $\mathbf{N}$ such that $\mathbf{a} \in S_Y$ if and only if $a_k \in B$ for $k \in Y$ and $a_k \in C$ otherwise. We shall show that each such set in 1–1 correspondence with $B^{\mathbf{N}}$ and $C^{\mathbf{N}}$.*

When we write the product as a union of the pairwise disjoint subsets $S_Y$, we are merely sorting the elements of the product into subsets depending upon which coordinates lie in $B$ and which lie in $C$. Since $B$ and $C$ are disjoint, these two properties are mutually exclusive. Each of the sets in question is a product for which every factor is either $B$ or $C$. Therefore all the factors are in 1–1 correspondence with both $B$ and $C$, and it follows (from an exercise in Section V.1) that every set $S_Y$ is in 1–1 correspondence with $B^{\mathbf{N}}$ and $C^{\mathbf{N}}$

*If $\mathbf{P}_1(\mathbf{N})$ denotes the proper subsets of $\mathbf{N}$, we shall construct a bijection from $\mathbf{P}_1(\mathbf{N}) \times C$ to $C$.*

The set $\mathbf{P}_1(\mathbf{N})$ is obtained from the entire power set $\mathbf{P}(\mathbf{N})$ by deleting one subset, and since we are working with infinite sets the cardinalities of $\mathbf{P}_1(\mathbf{N})$ and $\mathbf{P}(\mathbf{N})$ are equal. Since $|B| = |C| \geq |\mathbf{P}(\mathbf{N})|$, it follows that $|\mathbf{P}_1(\mathbf{N}) \times C| = |\mathbf{P}(\mathbf{N}) \times C| = |C|$.

*We shall show there is a bijection from $(B \cup C)^{\mathbf{N}}$ to $B \cup C$ sending $S_{\mathbf{N}} = B^{\mathbf{N}}$ to $B$ by the maximal map and sending the other sets $S_Y$ to $C$ by the composites of $S_Y \to \{Y\} \times C$ and $\mathbf{P}_1(\mathbf{N}) \times C \to C$.*

It is only necessary to define the bijection on the pieces. The assertion gives the definition on $S_{\mathbf{N}} = B^{\mathbf{N}}$, and it describes the map on the remaining pieces as well. We need to check this map is bijective. It will suffice to show that the partial composite $\cup_{Y \neq \mathbf{N}} S_Y \to \mathbf{P}_1(\mathbf{N}) \times C$ is bijective because the total composite sends the codomain to $C$, which is disjoint from $B$.

By construction the map sends the pairwise disjoint subsets $S_Y$ into the pairwise disjoint subsets $\{Y\} \times C$, so the proof of the bijectivity assertion reduces to verifying the latter for each of the pieces we have described. But on these pieces the map is a bijection by construction.

*We claim this a contradiction, and we shall determine the source of the contradiction.*

We have constructed a bijection from $(B \cup C)^{\mathbf{N}}$ to $B \cup C$ which properly contains a maximal bijection. The problem arose in our assumption that $|B|$ was strictly less than $|A|$, so the latter must be false and we must have $|B| = |A|$. As noted before, this proves the identity $|A^{\mathbf{N}}| = |A|$ when $|A| \geq |\mathbf{R}|$ and thus also completes the proof of the exercise.■

**7.** Since every set can be well-ordered, this is true for $\mathbf{R}$ in particular. Therefore the set of all uncountable ordinals is nonempty, so it must contain a least element, which we are calling $\Lambda_1$ here.

To prove the assertion about least upper bounds, we use the standard von Neumann model for the ordinals (ordered by $\alpha < \beta \Longleftrightarrow \alpha \in \beta$). Suppose that we have a countable family of ordinals $\alpha_k \in \Lambda_1$ and we consider the union $C$ of all ordinals $\beta$ such that $\beta \leq \alpha_k$ for some $k$. Now each $\alpha_k$ is countable, si this means that $C$ is countable. By construction, if $\gamma \leq \beta$ for some $\beta \in C$, then $\gamma \in C$. Now $\Lambda_1 - C$ cannot be countable since $C$ is countable but $\Lambda_1$ is not. Therefore $\Lambda_1 - C$ is nonempty and as such has a least element. By a previous sentence we know that this least element $\delta_0$ must be strictly greater than every element of $C$ and hence $\delta_0$ is an upper bound for the set of all ordinals $\alpha_k$.

To conclude we must find a least upper bound for this set of ordinals. Suppose that $\delta_0$ is not a least upper bound. Then there is some $\delta_1 < \delta_0$ that is also an upper bound. We claim that $\delta_1$ must be a least upper bound. By construction of $\delta_0$ we know that anything strictly less than $\delta_0$ is not greater than every element of $C$. Hence there is some $\gamma \in C$ such that $\delta_1 \leq \gamma < \delta_0$. Now $\gamma < \alpha_m$ for some $m$ and by the defining properties of $\delta_0$ we also have $\delta_1 \leq \alpha_m < \delta_0$. On the other hand, since $\delta_1$ is an upper bound we also have the reverse inequality $\alpha_m \leq \delta_1$ so that equality must hold. Thus we have shown that $\alpha_m \geq \alpha_k$ for all $k$, which means that $\alpha_m$ must be the least upper bound for the original set of ordinals.$\blacksquare$

*Postscript.* In fact, if $\delta_0$ is not the least upper bound, then we have $\delta_0 = \alpha_m + 1$ because $\delta_0$ is the least element that is greater than each of the elements in $C$.$\blacksquare$

**8.** We shall need the following elementary fact about linearly ordered sets:

**LEMMA.** *If $Y$ is a linearly ordered set and $y_1, \cdots, y_n \in Y$, then there is some $k$ such that $y_k \geq y_j$ for all $j$.*

**Proof.** This is trivial if $n = 1$; assume it is true for $n = m$, suppose we are given $y_1, \cdots, y_{m+1} \in Y$, and let $Y_0 = Y - \{y_{m+1}\}$. Then by the induction hypothesis there is some $q$ such that $y_q \geq y_j$ for $j \leq m$. Since $Y$ is linearly ordered we know that either $y_q \leq y_{m+1}$ or $y_{m+1} = y_q$. In the first case it follows that $y_{m+1} \geq y_j$ for all $j \leq m+1$, and in the second it follows that $y_q \geq y_j$ for all $j \leq m+1$.$\blacksquare$

*Solution to the exercise.* Let $F$ be a family of subsets of some set $S$ with the finite intersection property, and let $\mathcal{G}$ be the set of all families $G$ of subsets of $S$ such that $F \subset G$ and $G$ has the finite intersection property; then $\mathcal{G}$ is partially ordered with respect to inclusion. The proof of the statement in the exercise reduces to showing that the hypothesis in Zorn's Lemma is true for $\mathcal{G}$.

Let $\mathcal{L}$ be a nonempty linearly ordered subset of $\mathcal{G}$, and let $L^* = \cup\{L \in \mathcal{L}\}$. Clearly $L \subset L^*$ for all $L \in \mathcal{L}$; we claim that $L^* \in \mathcal{G}$; *i.e.*, $F \subset L^*$ and $L^*$ has the finite intersection property. The first statement is clear since $F$ is contained in every $L \in \mathcal{L}$. To prove the second, suppose we are given $A_1, \cdots, A_n \in \mathcal{L}^*$. For each $j$ there is some $L_j \in \mathcal{L}$ such that $A_j \in L_j$. Since $\mathcal{L}$ is linearly ordered, the lemma shows there is some $q$ such that $L_j \subset L_q$ for all $j$. Therefore we have $A_1, \cdots, A_n \in L_q$, and since $L_q$ has the finite intersection property we conclude that $\cap_j A_j \neq \emptyset$. — We have now shown that $\mathcal{G}$ is a partially ordered set in which linearly ordered subsets have upper bounds, and therefore $\mathcal{G}$ has a maximal element $G^*$ by Zorn's Lemma. By construction $G^*$ is a maximal family of subsets with the finite intersection property such that $F \subset G^*$.$\blacksquare$

**Further comment.** In many uses of Zorn's Lemma, it is important to understand what maximality implies for a set $H$ which properly contains the maximal set $G^*$. In this problem, it means that one can find a finite collection of subsets $B_t$ in $H$ such that $\cap_t B_t = \emptyset$. Here are two other facts about the maximal family $G^*$ in the exercise that are true: ($i$) The family $G^*$ is closed under finite intersections. ($ii$) If we are given $A \in G^*$ and $C \subset S$ such that $A \subset C$, then $C \in G^*$. — Both of these follow because $G^* \cup \{C\}$ has the finite intersection property; writing up this argument in detail is left to the reader.$\blacksquare$