Ordering and Zorn's LEMMA Folland's Real Analysis

It is easily verified that the map $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ defined by the second formula commutes with union, intersections, and complements:

$$f^{-1}\Big(\bigcup_{\alpha\in A} E_\alpha\Big) = \bigcup_{\alpha\in A} f^{-1}(E_\alpha), \qquad f^{-1}\Big(\bigcap_{\alpha\in A} E_\alpha\Big) = \bigcap_{\alpha\in A} f^{-1}(E_\alpha),$$
$$f^{-1}(E^c) = \big(f^{-1}(E)\big)^c.$$

(The direct image mapping $f : \mathcal{P}(X) \to \mathcal{P}(Y)$ commutes with unions, but in general not with intersections or complements.)

If $f : X \to Y$ is a mapping, X is called the **domain** of f and $f(X)$ is called the **range** of f. f is said to be **injective** if $f(x_1) = f(x_2)$ only when $x_1 = x_2$, **surjective** if $f(X) = Y$, and **bijective** if it is both injective and surjective. If f is bijective, it has an **inverse** $f^{-1} : Y \to X$ such that $f^{-1} \circ f$ and $f \circ f^{-1}$ are the identity mappings on X and Y, respectively. If $A \subset X$, we denote by $f|A$ the restriction of f to A:

$$(f|A) : A \to Y, \qquad (f|A)(x) = f(x) \text{ for } x \in A.$$

A **sequence** in a set X is a mapping from $\mathbb{N}$ into X. (We also use the term **finite sequence** to mean a map from $\{1, \ldots, n\}$ into X where $n \in \mathbb{N}$.) If $f : \mathbb{N} \to X$ is a sequence and $g : \mathbb{N} \to \mathbb{N}$ satisfies $g(n) < g(m)$ whenever $n < m$, the composition $f \circ g$ is called a **subsequence** of f. It is common, and often convenient, to be careless about distinguishing between sequences and their ranges, which are subsets of X indexed by $\mathbb{N}$. Thus, if $f(n) = x_n$, we speak of the sequence $\{x_n\}_1^\infty$; whether we mean a mapping from $\mathbb{N}$ to X or a subset of X will be clear from the context.

Earlier we defined the Cartesian product of two sets. Similarly one can define the Cartesian product of n sets in terms of ordered n-tuples. However, this definition becomes awkward for infinite families of sets, so the following approach is used instead. If $\{X_\alpha\}_{\alpha\in A}$ is an indexed family of sets, their **Cartesian product** $\prod_{\alpha\in A} X_\alpha$ is the set of all maps $f : A \to \bigcup_{\alpha\in A} X_\alpha$ such that $f(\alpha) \in X_\alpha$ for every $\alpha \in A$. (It should be noted, and then promptly forgotten, that when $A = \{1, 2\}$, the previous definition of $X_1 \times X_2$ is set-theoretically different from the present definition of $\prod_1^2 X_j$. Indeed, the latter concept depends on mappings, which are defined in terms of the former one.) If $X = \prod_{\alpha\in A} X_\alpha$ and $\alpha \in A$, we define the αth **projection** or **coordinate map** $\pi_\alpha : X \to X_\alpha$ by $\pi_\alpha(f) = f(\alpha)$. We also frequently write x and x_α instead of f and $f(\alpha)$ and call x_α the αth **coordinate** of x.

If the sets X_α are all equal to some fixed set Y, we denote $\prod_{\alpha\in A} X_\alpha$ by Y^A:

$$Y^A = \text{ the set of all mappings from } A \text{ to } Y.$$

If $A = \{1, \ldots, n\}$, Y^A is denoted by Y^n and may be identified with the set of ordered n-tuples of elements of Y.

0.2 ORDERINGS

A **partial ordering** on a nonempty set X is a relation R on X with the following properties:

- if xRy and yRz, then xRz;
- if xRy and yRx, then $x = y$;
- xRx for all x.

If R also satisfies

- if $x, y \in X$, then either xRy or yRx,

then R is called a **linear** (or **total**) ordering. For example, if E is any set, then $\mathcal{P}(E)$ is partially ordered by inclusion, and $\mathbb{R}$ is linearly ordered by its usual ordering. Taking this last example as a model, we shall usually denote partial orderings by $\leq$, and we write $x < y$ to mean that $x \leq y$ but $x \neq y$. We observe that a partial ordering on X naturally induces a partial ordering on every nonempty subset of X. Two partially ordered sets X and Y are said to be **order isomorphic** if there is a bijection $f : X \to Y$ such that $x_1 \leq x_2$ iff $f(x_1) \leq f(x_2)$.

If X is partially ordered by $\leq$, a **maximal** (resp. **minimal**) **element** of X is an element $x \in X$ such that the only $y \in X$ satisfying $x \leq y$ (resp. $x \geq y$) is x itself. Maximal and minimal elements may or may not exist, and they need not be unique unless the ordering is linear. If $E \subset X$, an **upper** (resp. **lower**) **bound** for E is an element $x \in X$ such that $y \leq x$ (resp. $x \leq y$) for all $y \in E$. An upper bound for E need not be an element of E, and unless E is linearly ordered, a maximal element of E need not be an upper bound for E. (The reader should think up some examples.)

If X is linearly ordered by $\leq$ and every nonempty subset of X has a (necessarily unique) minimal element, X is said to be **well ordered** by $\leq$, and (in defiance of the laws of grammar) $\leq$ is called a **well ordering** on X. For example, $\mathbb{N}$ is well ordered by its natural ordering.

We now state a fundamental principle of set theory and derive some consequences of it.

0.1 The Hausdorff Maximal Principle. *Every partially ordered set has a maximal linearly ordered subset.*

In more detail, this means that if X is partially ordered by $\leq$, there is a set $E \subset X$ that is linearly ordered by $\leq$, such that no subset of X that properly includes E is linearly ordered by $\leq$. Another version of this principle is the following:

0.2 Zorn's Lemma. *If X is a partially ordered set and every linearly ordered subset of X has an upper bound, then X has a maximal element.*

Clearly the Hausdorff maximal principle implies Zorn's lemma: An upper bound for a maximal linearly ordered subset of X is a maximal element of X. It is also not difficult to see that Zorn's lemma implies the Hausdorff maximal principle. (Apply Zorn's lemma to the collection of linearly ordered subsets of X, which is partially ordered by inclusion.)

0.3 The Well Ordering Principle. *Every nonempty set X can be well ordered.*

Proof. Let $\mathcal{W}$ be the collection of well orderings of subsets of X, and define a partial ordering on $\mathcal{W}$ as follows. If $\leq_1$ and $\leq_2$ are well orderings on the subsets E_1 and E_2, then $\leq_1$ precedes $\leq_2$ in the partial ordering if (i) $\leq_2$ extends $\leq_1$, i.e., $E_1 \subset E_2$ and $\leq_1$ and $\leq_2$ agree on E_1, and (ii) if $x \in E_2 \setminus E_1$ then $y \leq_2 x$ for all $y \in E_1$. The reader may verify that the hypotheses of Zorn's lemma are satisfied, so that $\mathcal{W}$ has a maximal element. This must be a well ordering on X itself, for if $\leq$ is a well ordering on a proper subset E of X and $x_0 \in X \setminus E$, then $\leq$ can be extended to a well ordering on $E \cup \{x_0\}$ by declaring that $x \leq x_0$ for all $x \in E$. ∎

0.4 The Axiom of Choice. *If $\{X_\alpha\}_{\alpha\in A}$ is a nonempty collection of nonempty sets, then $\prod_{\alpha\in A} X_\alpha$ is nonempty.*

Proof. Let $X = \bigcup_{\alpha\in A} X_\alpha$. Pick a well ordering on X and, for $\alpha \in A$, let $f(\alpha)$ be the minimal element of X_α. Then $f \in \prod_{\alpha\in A} X_\alpha$. ∎

0.5 Corollary. *If $\{X_\alpha\}_{\alpha\in A}$ is a disjoint collection of nonempty sets, there is a set $Y \subset \bigcup_{\alpha\in A} X_\alpha$ such that $Y \cap X_\alpha$ contains precisely one element for each $\alpha \in A$.*

Proof. Take $Y = f(A)$ where $f \in \prod_{\alpha\in A} X_\alpha$. ∎

We have deduced the axiom of choice from the Hausdorff maximal principle; in fact, it can be shown that the two are logically equivalent.

0.3 CARDINALITY

If X and Y are nonempty sets, we define the expressions

$$\operatorname{card}(X) \leq \operatorname{card}(Y), \qquad \operatorname{card}(X) = \operatorname{card}(Y), \qquad \operatorname{card}(X) \geq \operatorname{card}(Y)$$

to mean that there exists $f : X \to Y$ which is injective, bijective, or surjective, respectively. We also define

$$\operatorname{card}(X) < \operatorname{card}(Y), \qquad \operatorname{card}(X) > \operatorname{card}(Y)$$

to mean that there is an injection but no bijection, or a surjection but no bijection, from X to Y. Observe that we attach no meaning to the expression "card(X)" when it stands alone; there are various ways of doing so, but they are irrelevant for our purposes (except when X is finite — see below). These relationships can be extended to the empty set by declaring that

$$\operatorname{card}(\varnothing) < \operatorname{card}(X) \text{ and } \operatorname{card}(X) > \operatorname{card}(\varnothing) \text{ for all } X \neq \varnothing.$$

For the remainder of this section we assume implicitly that all sets in question are nonempty in order to avoid special arguments for $\varnothing$. Our first task is to prove that the relationships defined above enjoy the properties that the notation suggests.