

## Part II. Basic tools and concepts

In the next five chapters we shall assemble the concepts and tools needed for the development of a general theory of algebras and of universal constructions among them.

We begin with two chapters on ordered sets, lattices, closure operators and related concepts, since these will be used repeatedly. Because of the relation between well-ordering and the Axiom of Choice, I take the occasion to review briefly the Zermelo-Fraenkel axioms for set theory, and several statements equivalent to the Axiom of Choice.

Clearly, the general context for studying universal constructions should be some model of “a system of mathematical objects and the maps among them”. This is provided by the concept of a *category*. We develop the basic concepts of category theory in Chapter 6, and in Chapter 7 we formalize universal properties in category-theoretic terms.

Finally, in Chapter 8 we introduce the categories that will be of special interest to us: the *varieties of algebras*.

## Chapter 4. Ordered sets, induction, and the Axiom of Choice.

**4.1. Partially ordered sets.** We began Chapter 1 by making precise the concept of a group. Let us now do the same for that of a partially ordered set.

A partial ordering on a set is an instance of a “relation”. This is a different sense of the word from that of the last two chapters. These notes will deal extensively with both kinds of “relations”; which sense is meant will generally be clear from context. When there is danger of ambiguity, I will make the distinction explicit, as I do, for instance, in the index.

Intuitively, a relation on a family of sets  $X_1, \dots, X_n$  means a *condition* on  $n$ -tuples  $(x_1, \dots, x_n)$  ( $x_1 \in X_1, \dots, x_n \in X_n$ ). Since the information contained in the relation is determined by the set of  $n$ -tuples that satisfy it, this set is taken to *be* the relation in the formal definition, given below. That the relation is *viewed* as a “condition” comes out in the notation and language used.

**Definition 4.1.1.** *If  $X_1, \dots, X_n$  are sets, a relation on  $X_1, \dots, X_n$  means a subset  $R \subseteq X_1 \times \dots \times X_n$ . Relations are often written as predicates; i.e., the condition  $(x_1, \dots, x_n) \in R$  may be written  $R(x_1, \dots, x_n)$ , or  $Rx_1 \dots x_n$ , or, if  $n = 2$ , as  $x_1 R x_2$ .*

*A relation on  $X, \dots, X$ , i.e., a subset  $R \subseteq X^n$ , is called an  $n$ -ary relation on  $X$ .*

*If  $R$  is an  $n$ -ary relation on  $X$ , and  $Y$  is a subset of  $X$ , then the restriction of  $R$  to  $Y$  means  $R \cap Y^n$ , regarded as an  $n$ -ary relation on  $Y$ .*

We now recall

**Definition 4.1.2.** *A partial ordering on a set  $X$  means a binary relation “ $\leq$ ” on  $X$  satisfying the conditions*

$$(\forall x \in X) \quad x \leq x \quad \text{(reflexivity),}$$

$$(\forall x, y \in X) \quad x \leq y, \quad y \leq x \Rightarrow x = y \quad \text{(antisymmetry),}$$

$$(\forall x, y, z \in X) \quad x \leq y, \quad y \leq z \Rightarrow x \leq z \quad \text{(transitivity).}$$

*A total ordering on  $X$  means a partial ordering which also satisfies*

$$(\forall x, y \in X) \quad x \leq y \text{ or } y \leq x.$$

*A partially (respectively totally) ordered set means a set  $X$  given with a partial (total) ordering  $\leq$ .*

*If  $X$  is partially ordered by  $\leq$ , and  $Y$  is a subset of  $X$ , then  $Y$  will be understood to be partially ordered by the restriction of  $\leq$ , which will be denoted by the same symbol unless there is danger of ambiguity. This is called the induced ordering on  $Y$ .*

A more formal definition would make a partially ordered set a pair  $P = (|P|, \leq)$  where  $\leq$  is a partial ordering on  $|P|$ . But for us, partially ordered sets will in general be tools rather than the objects of our study, and it would slow us down to always maintain the distinction between  $P$  and  $|P|$ , so we shall usually take the informal approach of understanding a partially ordered set to mean a set  $P$  for which we “have in mind” a partial ordering relation  $\leq$ . Occasionally, however, we shall be more precise and refer to the pair  $(|P|, \leq)$ .

Standard examples of partially ordered sets are the set of real numbers with the usual relation

$\leq$ , the set  $\mathbf{P}(X)$  of subsets of any set  $X$  under the inclusion relation  $\subseteq$ , and the set of positive integers under the relation “ $|$ ”, where  $m|n$  means “ $m$  divides  $n$ ”.

A *total ordering* is also called a *linear ordering*. The term “ordered” without any qualifier is used by some authors as shorthand for “partially ordered”, and by others for the stronger condition “totally ordered”; we will here generally specify “partially” or “totally”. A subset  $C$  of a partially ordered set  $X$  which is totally ordered under the induced ordering is called a *chain* in  $X$ .

Note that in addition to this order-theoretic meaning of “chain”, there is a nonspecialized use of the word; for instance, one speaks of a “chain of equalities  $x_1 = x_2 = \dots = x_n$ ”. We shall at times use the term in this nontechnical way, relying on context to avoid ambiguity.

The versions of the concepts of homomorphism and isomorphism appropriate to partially ordered sets are given in

**Definition 4.1.3.** *If  $X$  and  $Y$  are partially ordered sets, an isotone map from  $X$  to  $Y$  means a function  $f: X \rightarrow Y$  such that  $x_1 \leq x_2 \Rightarrow f(x_1) \leq f(x_2)$ .*

*An invertible isotone map whose inverse is also isotone is called an order isomorphism.*

**Exercise 4.1:1.** Give an example of an isotone map of partially ordered sets which is invertible as a set map, but which is not an order isomorphism.

Some well-known notation: When  $\leq$  is a partial ordering on a set  $X$ , one commonly writes  $\geq$  for the opposite relation; i.e.,  $x \geq y$  means  $y \leq x$ . Clearly the relation  $\geq$  satisfies the same conditions of reflexivity, antisymmetry and transitivity as  $\leq$ .

This leads to a semantic problem: As long as  $\geq$  is just an auxiliary notation used in connection with the given ordering  $\leq$ , one thinks of an element  $x$  as being “smaller” (or “lower”) than an element  $y \neq x$  if  $x \leq y$ . But the preceding observation shows that one can take the opposite relation  $\geq$  as a new partial ordering on the set  $X$ , i.e., consider the partially ordered set  $(X, \geq)$ , and one should consider  $x$  as “smaller” than  $y$  in this partially ordered set if the pair  $(x, y)$  belongs to this *new* ordering. Such properties as which maps  $X \rightarrow Y$  are isotone (with respect to a fixed partial ordering on  $Y$ ) clearly change when one goes from considering  $X$  under  $\leq$  to considering it under  $\geq$ .

The set  $X$  under the opposite of the given partial ordering is called the *opposite* of the original partially ordered set. When one uses the formal notation  $P = (|P|, \leq)$  for a partially ordered set, one can write  $P^{\text{op}} = (|P|, \geq)$ . One may also replace the symbol  $\geq$  by  $\leq^{\text{op}}$ , writing  $P^{\text{op}} = (|P|, \leq^{\text{op}})$ . Thus, if  $x$  is smaller than  $y$  in  $P$ , i.e.,  $x \leq y$ , then  $y$  is smaller than  $x$  in  $P^{\text{op}}$ , i.e.,  $y \leq^{\text{op}} x$ . (“Dual ordering” is another term often used, and  $*$  is sometimes used instead of  $^{\text{op}}$ .)

In these notes we shall rarely make explicit use of the opposite partially ordered set construction. But we remark that once one gets past the notational confusion, the symmetry in the theory of partially ordered sets created by that construction is a very useful tool: after proving any result, one can say “By duality ...”, and immediately deduce the corresponding statement with all ordering relations reversed.

One also commonly uses  $x < y$  as an abbreviation for  $(x \leq y) \wedge (x \neq y)$ , and of course  $x > y$  for  $(x \geq y) \wedge (x \neq y)$ . These relations do *not* satisfy the same conditions as  $\leq$ . The conditions that they do satisfy are noted in

**Exercise 4.1:2.** Show that if  $\leq$  is a partial ordering on a set  $X$ , then the relation  $<$  is transitive and is *antireflexive*, i.e., satisfies  $(\forall x \in X) x \not< x$ . Conversely, show that any transitive antireflexive binary relation  $<$  on a set  $X$  is induced in the above way by a unique partial ordering  $\leq$ .

A relation  $<$  with these properties (transitivity and antireflexivity) might be called a “partial strict ordering”. One can thus refer to “the partial strict ordering  $<$  corresponding to the partial ordering  $\leq$ ”, and “the partial ordering  $\leq$  corresponding to the partial strict ordering  $<$ ”. Of course, for a partial ordering denoted by a symbol such as “ $|$ ” (“divides”), or  $R$  (a partial ordering written as a binary relation), there is no straightforward symbol for the corresponding partial strict ordering.

**Exercise 4.1:3.** For partially ordered sets  $X$  and  $Y$ , suppose we call a function  $f: X \rightarrow Y$  a *strict isotone map* if  $x < y \Rightarrow f(x) < f(y)$ . Show that

$$\text{one-to-one and isotone} \Rightarrow \text{strict isotone} \Rightarrow \text{isotone},$$

but that neither implication is reversible.

In contexts where “ $\leq$ ” already has a meaning, if another partial ordering has to be considered, it is often denoted by a variant symbol such as  $\preccurlyeq$ . One then uses corresponding symbols  $\succcurlyeq$ ,  $\prec$ ,  $\succ$  for the opposite order, the strict order relation, etc.. (However, order-theorists dealing with a partial ordering  $\leq$  sometimes write  $y \succ x$  to mean “ $y$  covers  $x$ ”, that is “ $y > x$  and there is no  $z$  between  $y$  and  $x$ ”. When the symbol is used this way, it cannot be used for the strict relation associated with a second ordering. We shall not use the concept of covering in these notes.)

A somewhat confused situation is that of symbols for the *subset* relation. Most often, the notation one would expect from the above discussion is followed; that is,  $\subseteq$  is used for “is a subset of”,  $\supseteq$  for the opposite relation, and  $\subset$ ,  $\supset$  for strict inclusions; we will follow these conventions here. However, many authors, especially in Eastern Europe, write  $\subset$  for “is a subset of”, a usage based on the view that since this is a more fundamental concept than that of a proper subset, it should be denoted by a primitive symbol and not by one obtained by adding an extra mark to the symbol for “proper subset”. Those authors use  $\subsetneq$  for “proper subset” (and the reversed symbols for the reversed relations). There was even at one time a movement to make “ $<$ ” mean “less than or equal to”, with  $\not\leq$  for strict inequality. Together with the above set-theoretic usage, this would have formed a consistent system, but the idea never got off the ground. Finally, many authors, for safety, use a mixed system:  $\subseteq$  for “subset” and  $\subsetneq$  for “proper subset”. (That was the notation used in the first graduate course I took, and I sometimes follow it in my papers. However, I rarely need a symbol for strict inclusion, so the question of how to write it seldom comes up.)

Although partially ordered sets are not algebras in the sense in which we shall use the term, many of the kinds of universal constructions we have considered for algebras can be carried out for them. In particular

**Definition 4.1.4.** Let  $(X_i)_{i \in I}$  be a family of partially ordered sets. Then their direct product will mean the partially ordered set having for underlying set the direct product of the underlying sets of the  $X_i$ , ordered so that  $(x_i)_{i \in I} \leq (y_i)_{i \in I}$  if and only if  $x_i \leq y_i$  for all  $i \in I$ .

**Exercise 4.1:4.** (i) Verify that the above relation is indeed a partial ordering on the product set, and that the resulting partially ordered set has the appropriate universal property to be called the direct product of the partially ordered sets  $X_i$ .

(ii) Let  $X$  be a set and  $R$  a binary relation on  $X$ . Show that there exists a universal example of a partially ordered set  $(Y, \leq)$  with a map  $u: X \rightarrow Y$  such that for all  $(x_1, x_2) \in R$  one has  $u(x_1) \leq u(x_2)$  in  $Y$ . This may be called the partially ordered set *presented* by the generators  $X$  and the relation-set  $R$ . (Cf. presentations of groups, monoids, and rings, §§3.3, 3.10, 3.12.) Will the map  $u$  in general be one-to-one? Onto?

(iii) Determine whether there exist constructions with the universal properties of the *coproduct* of two partially ordered sets, and of the *free* partially ordered set on a set  $X$ . Describe these if they exist.

(iv) Discuss the problem of *imposing* a set  $R$  of further relations on a given partially ordered set  $(X, \leq)$ ; i.e., of constructing a universal isotone map of  $X$  into a partially ordered set  $Y$  such that the images of the elements of  $X$  also satisfy the relations comprising  $R$ , and if this can be done, examine the properties of the construction.

We have noted that for any set  $X$ , the set  $\mathbf{P}(X)$  of subsets of  $X$  is partially ordered by  $\subseteq$ . Given a partially ordered set  $S$ , we may look for universal ways of representing  $S$  by subsets of a set  $X$ . Note that if  $f: X \rightarrow Y$  is a map between sets, then  $f$  induces, in natural ways, both an isotone map  $\mathbf{P}(X) \rightarrow \mathbf{P}(Y)$  and an isotone map  $\mathbf{P}(Y) \rightarrow \mathbf{P}(X)$ , the first taking subsets of  $X$  to their images under  $f$ , the second taking subsets of  $Y$  to their inverse images. Let us call these the “direction-preserving construction” and the “direction-reversing construction” respectively. Thus, given a partially ordered set  $S$ , there are four universal sets we might look for: a set  $X$  having an isotone map  $S \rightarrow \mathbf{P}(X)$  universal in terms of the direction-preserving construction of maps among power sets, a set  $X$  with such a map universal in terms of the direction-reversing construction, and sets  $X$  with isotone maps in the reverse direction,  $\mathbf{P}(X) \rightarrow S$ , universal for the same two constructions of maps among power sets.

**Exercise 4.1:5.** (i) Write out the universal properties of the four possible constructions indicated.

(ii) Investigate which of the four universal sets exist, and describe these as far as possible.

**Definition 4.1.5.** Let  $X$  be a partially ordered set,  $S$  a subset of  $X$ , and  $s$  an element of  $S$ . Then  $s$  is said to be *minimal* in  $S$  if there is no  $t \in S$  with  $t < s$ , while  $s$  is said to be the *least element* of  $S$  if for all  $t \in S$ ,  $s \leq t$ . The terms *maximal* and *greatest* are used for the dual concepts.

(There was really no need to refer to  $X$  in the above definition, since the properties in question just depend on the set  $S$  and the induced order relation on it; but these concepts are often applied to subsets of larger partially ordered sets, so I included this context in the definition.)

**Exercise 4.1:6.** Let  $X$  be a partially ordered set.

(i) Show that if  $X$  has a least element  $x$ , then  $x$  is the unique minimal element of  $X$ .

(ii) If  $X$  is finite, show conversely that a unique minimal element, if it exists, is a least element.

(iii) Give an example showing that if  $X$  is not assumed finite, this converse is false.

(I have included (iii) as a warning; I have many times found myself unwittingly writing or saying “unique minimal element” when I meant “least element”. It sounds somehow more precise, but it doesn’t mean the same thing.)

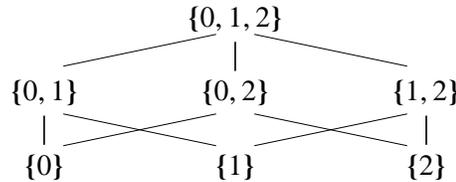
**Exercise 4.1:7.** Let  $(X, \leq)$  be a partially ordered set. Then the pair  $(X, \leq)$  constitutes a presentation of itself as a partially ordered set in the sense of Exercise 4.1:4(ii); but of course, there may be proper subsets  $R$  of the relation  $\leq$  such that  $(X, R)$  is a presentation of the

same partially ordered set. (I.e., such that  $R$  “generates”  $\leq$  in an appropriate sense.)

- (i) If  $X$  is finite, show that there exists a *least* subset of  $R$  which generates  $\leq$ .
- (ii) Show that this is not in general true for infinite  $X$ .

Point (i) of the above exercise is the basis for the familiar way of diagramming finite partially ordered sets. One draws a picture with vertices representing the elements of the set, and edges corresponding to the members of the least relation generating the partial ordering; i.e., the smallest set of order relations from which all the others can be deduced. The higher point on each edge represents the larger element under the partial ordering. This picture is called the *Hasse diagram* of the given partially ordered set.

For example, the picture below represents the partially ordered set of all nonempty subsets of  $\{0, 1, 2\}$ . The relation  $\{1\} \leq \{0, 1, 2\}$  is not shown explicitly, because it is a consequence of the relations  $\{1\} \leq \{0, 1\} \leq \{0, 1, 2\}$  (and also of  $\{1\} \leq \{1, 2\} \leq \{0, 1, 2\}$ ).



The next definition lists a few more pieces of commonly used terminology.

**Definition 4.1.6.** Let  $\leq$  be a partial ordering on a set  $X$ .

If  $x, y$  are elements of  $X$  with  $x \leq y$ , then the interval  $[x, y]$  means the subset  $\{z \in X \mid x \leq z \leq y\}$ , with the induced partial ordering  $\leq$ .

Elements  $x$  and  $y$  of  $X$  are called *incomparable* if neither  $x \leq y$  nor  $y \leq x$  holds. A subset  $Y \subseteq X$  is called an *antichain* if every pair of distinct elements of  $Y$  is incomparable.

An element  $x \in X$  is said to *majorize* a subset  $Y \subseteq X$  if for all  $y \in Y$ ,  $y \leq x$ . One similarly says  $x$  *majorizes* an element  $y$  if  $y \leq x$ .

A subset  $Y$  of  $X$  is said to be *cofinal* in  $X$  if every element of  $X$  is majorized by some element of  $Y$ .

Unfortunately, there are no standard terms for the duals of the concepts “majorize” and “cofinal”. One occasionally sees “minorize” and “coinitial”, but these seem awkward. I often write “downward cofinal” for the opposite of “cofinal”; the best circumlocutions I can suggest for the other concept are “is majorized by every element of” and “majorizes under the opposite ordering”.

The concept of cofinality defined above probably originated in topology: If  $s$  is a point of a topological space  $S$ , and  $N(s)$  the set of all neighborhoods of  $s$ , then a *neighborhood basis* of  $s$  means a subset  $B \subseteq N(s)$  cofinal in that set, under the ordering by reverse inclusion. The virtue of this concept is that one can verify that a function on  $S$  approaches some limit at  $s$  by checking its behavior on members of such a  $B$ . E.g., one generally checks continuity of a function at a point  $s$  of the real line using the cofinal system of neighborhoods  $\{(s - \epsilon, s + \epsilon) \mid \epsilon > 0\}$ .

**Exercise 4.1:8.** (i) Show that if  $X$  is a *finite* partially ordered set, then a subset  $Y$  is cofinal in  $X$  if and only if it contains all maximal elements of  $X$ .

- (ii) Show by example that this is not true for infinite partially ordered sets. Is one direction

true?

**Exercise 4.1:9.** Let  $X$  be a finite partially ordered set. One defines the *height* of  $X$  as the maximum of the cardinalities of all chains in  $X$ , and the *width* of  $X$  as the maximum of the cardinalities of all antichains in  $X$ .

(i) Show that  $\text{card}(X) \leq \text{height}(X) \cdot \text{width}(X)$ .

(The above result fails for infinite partially ordered sets, as will be shown in Exercise 4.6:8(ii).)

(ii) Must every (or some) chain in  $X$  of maximal cardinality have nonempty intersection with every (or some) antichain of maximal cardinality?

**Definition 4.1.7.** Let  $\leq$  and  $\preceq$  be partial orderings on a set  $X$ . Then one says  $\leq$  is an extension or strengthening (or sometimes, a refinement) of  $\preceq$  if it contains the latter, as subsets of  $X \times X$ ; that is, if  $x \preceq y \Rightarrow x \leq y$ .

The relation of “extension” is a partial ordering on the set of partial orderings on  $X$ . This fact can be looked at as follows. If we regard each partial ordering on  $X$  as a subset  $R \subseteq X \times X$ , and partially order the class of all subsets of  $X \times X$  by inclusion (the relation  $\subseteq$ ), then the relation of “extension” is the *restriction* of this partial ordering to the subclass of those  $R \subseteq X \times X$  which are partial orders. This observation saves us the work of verifying that “extension” satisfies the conditions for a partial order, since we know that the restriction of a partial order on a set to any subset is again a partial order. Many of the partial orderings that arise naturally in mathematics are, similarly, restrictions of the inclusion relation or of some other natural partial ordering on a larger set.

**Exercise 4.1:10.** Consider the set of all partial orderings on a set to be partially ordered as above.

(i) Show that the *maximal* elements in the set of all partial orderings on a set  $X$  are precisely the *total* orderings.

(ii) How many maximal elements does the set of partial orderings of a set of  $n$  elements have?

(iii) How many minimal elements does the set of partial orderings of a set of  $n$  elements have?

(iv) Show that every partial ordering on a finite set  $X$  is the set-theoretic intersection of a set of total orderings.

If  $\preceq$  is a partial ordering on a finite set, the smallest number of total orderings that can be intersected to get  $\preceq$  is called the “order dimension” of the partially ordered set. The next question is open-ended.

(v) What can you say about the order dimension function? (You might look for general bounds on the order dimension of a partially ordered set of  $n$  elements, try to evaluate the order dimensions of particular partially ordered sets, look at the behavior of order dimension under various constructions, etc..)

Here is an outstanding open problem.

**Exercise 4.1:11.** Let  $(X, \preceq)$  be a finite partially ordered set. Let  $N$  denote the number of total orderings “ $\leq$ ” on  $X$  extending  $\preceq$  (“linearizations of  $\preceq$ ”) and for  $x, y \in X$ , let  $N_{x,y}$  denote the number of these extensions “ $\leq$ ” which satisfy  $x \leq y$ .

(i) Prove or disprove, if you can,

*Fredman’s conjecture:* For any  $(X, \preceq)$  such that  $\preceq$  is not a total order, there exist elements  $x, y \in X$  such that

$$(4.1.8) \quad 1/3 \leq N_{x,y}/N \leq 2/3.$$

If you cannot settle this open question, here are some special cases to look at:

(ii) Let  $r$  be a positive integer, and let  $X$  be the partially ordered set consisting of a chain of  $r$  elements,  $p_1 \prec \dots \prec p_r$ , and an element  $q$  incomparable with all the  $p_i$ . What are  $N$  and the  $N_{p_i,q}$  in this case? Verify Fredman's conjecture for this partially ordered set.

(iii) Is the above example consistent with the stronger assertion that if  $X$  has no greatest element, then an  $x$  and a  $y$  satisfying (4.1.8) can be chosen from among the *maximal* elements of  $X$ ? With the assertion that for every two maximal chains in  $X$ , one can choose an  $x$  in one of these chains and a  $y$  in the other satisfying (4.1.8)? If one or the other of these possible generalizations of Fredman's Conjecture is not excluded by the above example, can you find an example that does exclude it?

(iv) Let  $r$  again be a positive integer, and let  $X$  be the set  $\{1, \dots, r\}$  partially ordered by the relation  $\prec$  under which  $i \prec j$  if and only if  $j-i \geq 2$  (where in this definition  $\geq$  has the usual meaning for integers). Verify the conjecture in this case as well. How many pairs  $(i, j)$  are incomparable under  $\prec$ , and of these pairs, how many satisfy (4.1.8)?

(v) If  $X$  is any partially ordered set such that the function  $N_{x,y}/N$  never takes on the value  $1/2$ , define a relation  $\leq_1$  on  $X$  by writing  $x \leq_1 y$  if either  $x = y$ , or  $N_{x,y}/N > 1/2$ . Determine whether this is *always*, *sometimes* or *never* a (total) ordering on  $X$ . Show that for any  $X$  which is a counterexample to Fredman's Conjecture,  $\leq_1$  must be a total ordering on  $X$ .

Fredman's conjecture arose as follows. Suppose that  $(X, \leq)$  is a finite *totally* ordered set, but that one has only partial information on its ordering; namely, one knows for certain pairs of elements  $x, y$  which element is greater, but not for all pairs. This partial information is equivalent to a partial ordering  $\prec$  on  $X$  weaker than  $\leq$ . Suppose one is capable of "testing" pairs of elements to determine their relation under  $\leq$ , and one wants to fully determine  $\leq$  using a small number of such tests. One would like to choose each test so that it approximately halves the number of candidate orderings. Examples show that one cannot do that well; but Fredman's Conjecture would imply that one can always reduce this number by at least a third at each step. For some literature on the subject, see [53] and papers referred to there.

My feeling is that it may be possible to prove Fredman's conjecture by assuming we had a counterexample, and considering the peculiar place the relation  $\leq_1$  of part (v) of the above exercise would have to have among the total orderings on  $X$  extending  $\prec$ . One can see something of the structure of the set of all total orderings on a set from the next exercise.

**Exercise 4.1:12.** Define the *distance* between two total orderings  $\leq_i, \leq_j$  on a finite set  $X$  as

$$d(\leq_i, \leq_j) = \text{number of pairs of elements } (x, y) \text{ such that } x \prec_i y, x \succ_j y.$$

Show that  $d$  is a metric on the set of all total orderings, and that for any partial ordering  $\prec$  on  $X$ , any two total orderings extending  $\prec$  can be connected by a chain (*not* meant in the order-theoretic sense!)  $\leq_1, \dots, \leq_n$  where each  $\leq_i$  is a total ordering extending  $\prec$ , and the distance between successive terms of the chain is 1.

Here is another open question.

**Exercise 4.1:13.** (*Reconstruction problem for finite partially ordered sets.*) Let  $P$  and  $Q$  be finite partially ordered sets with the same number  $n > 3$  of elements, and suppose they can be indexed  $P = \{p_1, \dots, p_n\}$ ,  $Q = \{q_1, \dots, q_n\}$  in such a way that for each  $i$ ,  $P - \{p_i\}$  and  $Q - \{q_i\}$  are isomorphic as partially ordered sets. Must  $P$  be isomorphic to  $Q$ ?

(Note that nothing is assumed about what bijections give the isomorphisms  $P - \{p_i\} \cong Q - \{q_i\}$ . We are definitely not assuming that they are the correspondences  $p_j \leftrightarrow q_j$  ( $j \neq i$ ); if we assumed this, the question would have an immediate positive answer. A way to state the

hypothesis without referring to such a correspondence is to say that the families of isomorphism classes of partially ordered  $(n-1)$ -element subsets of  $P$  and of  $Q$ , counting multiplicities, are the same.)

If the above question has an affirmative answer, then “one can reconstruct  $P$  from its  $(n-1)$ -element partially ordered subsets”, hence the name of the problem.

(The corresponding question for *graphs* with  $n > 2$  vertices is also open, and better known.)

**4.2. Digression: preorders.** One sometimes encounters binary relations which, like partial orderings, are reflexive and transitive, but which do not satisfy the antisymmetry condition. For instance, although the relation “divides” on the positive integers is a partial ordering, the relation “divides” on the set of all integers is not antisymmetric, since every  $n$  divides  $-n$  and vice versa. More generally, on the elements of any commutative integral domain, “divides” is a reflexive transitive relation, but for every element  $x$  and invertible element  $u$ ,  $x$  and  $ux$  each divide the other. Similarly, on a set of *propositions* (sentences in some formal language) about a mathematical situation, the relation  $P \Rightarrow Q$  is reflexive and transitive, but not generally antisymmetric: Distinct sentences can each imply the other, i.e., represent equivalent conditions.

To cover such situations, one makes

**Definition 4.2.1.** *A reflexive transitive (not necessarily antisymmetric) binary relation on a set  $X$  is called a preorder on  $X$ .*

The concept of a preordered set can be reduced in a natural way to a combination of two other sorts of structure that we already know:

**Proposition 4.2.2.** *Let  $X$  be a set. Then the following data are equivalent.*

- (i) *A preorder  $\preccurlyeq$  on  $X$ .*
- (ii) *An equivalence relation  $\approx$  on  $X$ , and a partial ordering  $\leq$  on the set  $X/\approx$  of equivalence classes.*

*Namely, to go from (i) to (ii), given the preorder  $\preccurlyeq$  define  $x \approx y$  to mean  $x \preccurlyeq y \wedge y \preccurlyeq x$ , and for any two elements  $[x], [y] \in X/\approx$ , write  $[x] \leq [y]$  in  $X/\approx$  if and only if  $x \preccurlyeq y$  in  $X$ .*

*Inversely, given, as in (ii), an equivalence relation  $\approx$  and a partial ordering  $\leq$  on  $X/\approx$ , one gets a preorder  $\preccurlyeq$  by defining  $x \preccurlyeq y$  to hold in  $X$  if and only if  $[x] \leq [y]$  in  $X/\approx$ .  $\square$*

**Exercise 4.2:1.** Prove the above proposition. (This requires one verification of well-definedness, and some observations showing why the two constructions, performed successively in either order, return the original data.)

This is neat: A reflexive transitive relation (a preorder) decomposes into a reflexive transitive *symmetric* relation (an equivalence relation) and a reflexive transitive *antisymmetric* relation (a partial ordering).

As an example, if we take the set of elements of a commutative ring  $R$ , preordered by divisibility, and divide out by the equivalence relation of mutual divisibility, we get a partially ordered set, which can be identified with the set of principal ideals of  $R$  partially ordered by reverse inclusion.

In view of the above proposition, there is no need for a *theory* of preorders – that is essentially subsumed in the theory of partial orderings. But it is useful to have the term “preorder” available, to refer to such relations when they arise.

The remainder of this section consists of some exercises on preorders which will not be used in subsequent sections. Exercises 4.2:2-4.2:9 concern a class of preorders having applications to ring theory, group theory, and semigroup theory. (Dependencies within that group of exercises: All later exercises depend on 4.2:2-4.2:3, and 4.2:5 is also assumed in 4.2:6-4.2:9. If you wish to hand in one of these exercises without writing out the details of others on which it depends, you should begin with a summary of the results from the latter that you will be assuming. You might check this summary with me first.)

The last exercise of the section is independent of that group.

**Exercise 4.2:2.** If  $f$  and  $g$  are nondecreasing functions from the positive integers to the nonnegative integers, let us write  $f \preccurlyeq g$  if there exists a positive integer  $N$  such that for all  $i$ ,  $f(i) \leq g(Ni)$ .

- (i) Show that  $\preccurlyeq$  is a preorder, but not a partial order, on the set of nondecreasing functions.
- (ii) On the subset of functions consisting of all polynomials with nonnegative integer coefficients, get an explicit description of  $\preccurlyeq$ , and determine its “decomposition” as in Proposition 4.2.2.
- (iii) Do the same for the set of functions consisting of the polynomials of (ii), together with the exponential functions  $i \mapsto n^i$  for all integers  $n > 1$ .
- (iv) Show that the partial ordering  $\leq$  on equivalence classes induced by the preordering  $\preccurlyeq$  on nondecreasing functions from positive integers to nonnegative integers is not a total ordering.
- (v) Regarding the nondecreasing functions from positive integers to nonnegative integers as a monoid under addition, show that the equivalence relation  $\approx$  induced by  $\preccurlyeq$  is a congruence on this monoid, so that the factor set again becomes an additive monoid.

**Exercise 4.2:3.** Let  $S$  be a monoid and  $x_1, \dots, x_n$  elements of  $S$ , and for each positive integer  $i$ , let  $g_{x_1, \dots, x_n}(i)$  denote the number of distinct elements of  $S$  which can be written as words of length  $\leq i$  in  $x_1, \dots, x_n$  (with repetitions allowed). This is a nondecreasing function from the positive integers to the nonnegative integers, the *growth function* associated with  $x_1, \dots, x_n$ .

Show that if  $S$  is generated by  $x_1, \dots, x_n$ , and if  $y_1, \dots, y_m$  is any other finite family of elements of  $S$ , then in the notation of the preceding exercise,  $g_{y_1, \dots, y_m} \preccurlyeq g_{x_1, \dots, x_n}$ . Deduce that if  $x_1, \dots, x_n$  and  $y_1, \dots, y_m$  are two generating sets for the same monoid, then  $g_{x_1, \dots, x_n} \approx g_{y_1, \dots, y_m}$ , where  $\approx$  is the equivalence relation determined as in Proposition 4.2.2 by the preorder  $\preccurlyeq$ .

Thus, if  $S$  is finitely generated, the equivalence class  $[g_{x_1, \dots, x_n}]$  is the same for all finite generating sets  $x_1, \dots, x_n$  of  $S$ . This equivalence class is therefore an invariant of the finitely generated monoid  $S$ , called its *growth rate*.

We see that if a finitely generated monoid  $S$  is embeddable in another finitely generated monoid  $T$ , then the growth rate of  $S$  must be  $\leq$  that of  $T$ .

**Exercise 4.2:4.** (i) Determine the structure of the partially ordered set consisting of the growth rates of the *free abelian* monoids on finite numbers of generators together with those of the *free* monoids on finite numbers of generators.

- (ii) With the help of the result of (i), show that the free abelian monoid on  $m$  generators is embeddable in the free abelian monoid on  $n$  generators if and only if  $m \leq n$ .
- (iii) Verify that for any positive integer  $n$ , the map from the *free monoid* on  $n$  generators  $x_1, \dots, x_n$  to the free monoid on 2 generators  $x, y$  taking  $x_i$  to  $xy^i$  ( $i = 1, \dots, n$ ) is an embedding. Is this consistent with the results of (i)?

This concept of growth rate is more often studied for groups and rings than for monoids. Note that elements  $x_1, \dots, x_n$  of a group  $G$  generate  $G$  as a group if and only if  $x_1, x_1^{-1}, \dots, x_n, x_n^{-1}$

generate  $G$  as a monoid, so the group-theoretic growth function of  $G$  with respect to  $\{x_1, \dots, x_n\}$  may be defined to be the growth function of  $G$  as a monoid with respect to the generating set  $\{x_1, x_1^{-1}, \dots, x_n, x_n^{-1}\}$ . The equivalence class of such growth functions is called the growth rate of the group  $G$ , which is thus the same as the growth rate of  $G$  as a monoid. This concept of growth rate has been used, in particular, in studying fundamental groups of manifolds [128].

If  $R$  is an algebra over a field  $k$ , then to get the ring theoretic concept of the growth rate of  $R$ , one considers, not the *number* of elements which can be written as a product of  $\leq i$  generators, but the *dimension* of the  $k$ -vector space spanned in  $R$  by such products. The remainder of the development is analogous to that of the monoid case.

Though it is a bit of a digression from the subject of preorders, I will sketch in the next few exercises an important invariant obtained from these growth rates, and some of its properties.

**Exercise 4.2:5.** If  $S$  is a monoid with finite generating set  $x_1, \dots, x_n$ , the *Gel'fand-Kirillov dimension* of  $S$  is defined as

$$(4.2.3) \quad \text{GK}(S) = \limsup_i (\ln(g_{x_1, \dots, x_n}(i)) / \ln(i)).$$

(Here “ $\ln$ ” denotes the natural logarithm, and  $\limsup_i a(i)$  means  $\lim_{j \rightarrow \infty} \sup_{i \geq j} a(i)$ . Thus, if  $a$  is a nonnegative function,  $\limsup_i a(i)$  is a nonnegative real number or  $+\infty$ .)

(i) Show that the right hand side of (4.2.3) is a function only of the *growth rate*  $[g_{x_1, \dots, x_n}]$ , hence does not depend on the choice of generators  $x_1, \dots, x_n$ , hence that the Gel'fand-Kirillov dimension is well defined.

(ii) Determine the Gel'fand-Kirillov dimensions of the free abelian monoid and the free monoid on  $n$  generators.

**Exercise 4.2:6.** (i) In the early literature on Gel'fand-Kirillov dimension, it was often stated (in effect) that for monoids  $S_1, S_2$ , one had  $\text{GK}(S_1 \times S_2) = \text{GK}(S_1) + \text{GK}(S_2)$ . Sketch an argument that seems to give this result, then point out the fallacy, and if you can, find a counterexample. (Actually, the statement was made for tensor products of algebras, rather than direct products of monoids, but one case can be reduced to the other.)

**Exercise 4.2:7.** (i) Show that if  $S$  is a finitely generated monoid and  $\text{GK}(S) < 2$ , then  $\text{GK}(S) = 0$  or  $1$ .

(ii) Show, on the other hand, that there exist finitely generated monoids having for Gel'fand-Kirillov dimensions all real numbers  $\geq 2$ , and  $+\infty$ . (Suggestion: Show that for any finite or infinite set  $S$  of elements of a free monoid  $F$ , one can construct a homomorphic image of  $F$  in which all elements not having members of  $S$  as subwords are distinct, while all elements that do have subwords in  $S$  have a common value, “0”.)

(iii) Show that there exist finitely generated monoids with distinct growth rates, but the same finite Gel'fand-Kirillov dimension.

We haven't seen any exercises on growth rates of  $k$ -algebras yet. If, as in the preceding exercise, one is only concerned with what growth rates occur, there is essentially no difference between the cases of  $k$ -algebras and of monoids, as shown in

**Exercise 4.2:8.** Let  $k$  be any field.

Show that for every monoid  $S$  with generating set  $s_1, \dots, s_n$ , there exists a  $k$ -algebra  $R$  with a generating set  $r_1, \dots, r_n$  such that  $g_{r_1, \dots, r_n} = g_{s_1, \dots, s_n}$ . Similarly, show that for every  $k$ -algebra  $R$  with generating set  $r_1, \dots, r_n$ , there exists a monoid  $S$  with a generating set  $s_1, \dots, s_{n+1}$  such that  $g_{s_1, \dots, s_{n+1}} = g_{r_1, \dots, r_n} + 1$  (where “1” denotes the constant function with value 1).

However, if one is interested in the growth of algebras with particular ring-theoretic properties, these do not in general reduce to questions about monoids. For instance, students familiar with the theory of transcendence degree of field extensions might do

**Exercise 4.2:9.** Show that if  $k$  is a field and  $R$  a finitely generated commutative  $k$ -algebra without zero-divisors, then the Gel'fand-Kirillov dimension of  $R$  as a  $k$ -algebra equals the transcendence degree over  $k$  of the field of fractions of  $R$ .

For more on Gel'fand-Kirillov dimension in ring theory, see [89].

For students familiar with the definitions of general topology, another instance of the concept of preorder is noted in:

**Exercise 4.2:10.** (i) Show that if  $X$  is a topological space, and if for  $x, y \in X$ , we define  $y \leq x$  to mean “the closure of  $\{x\}$  contains  $y$ ”, then  $\leq$  is a preorder on  $X$ .

(ii) Show that if  $X$  is *finite*, the above construction gives a bijection between topologies and preorders on  $X$ .

(iii) Under the above bijection, what classes of preorders correspond to  $T_0$ , respectively  $T_1$ , respectively  $T_2$  topologies?

(iv) If  $X$  is *infinite*, is the above map from topologies to preorders one-to-one? Onto? Can one associate to every preorder on  $X$  a strongest and/or a weakest topology yielding the given preorder under this construction?

**4.3. Induction, recursion, and chain conditions.** The familiar principle of induction on the natural numbers (nonnegative integers) that one learns as an undergraduate is based on the order properties of that set. In this and the next two sections, we shall examine more general kinds of ordered sets over which one can perform inductive proofs. We shall also see that analogous to *inductive proofs* there is a concept of *recursive constructions*, which can be performed under similar hypotheses.

Any students to whom the distinction between “minimal” and “least” elements in a partially ordered set was new should review Definition 4.1.5 before going on.

**Lemma 4.3.1.** *Let  $(X, \leq)$  be a partially ordered set. Then the following conditions are equivalent:*

(i) *Every nonempty subset of  $X$  has a minimal element.*

(ii) *For every descending chain  $x_0 \geq x_1 \geq \dots \geq x_i \geq \dots$  in  $X$  indexed by the natural numbers, there is some  $n$  such that  $x_n = x_{n+1} = \dots$ .*

(ii') *Every strictly descending chain  $x_0 > x_1 > \dots$  indexed by an initial subset of the natural numbers (that is, either by  $\{0, 1, \dots, n\}$  for some  $n$ , or by the set of all nonnegative integers) is finite (that is, is in fact indexed by  $\{0, 1, \dots, n\}$  for some  $n$ ).*

(ii'')  *$X$  has no strictly descending chains  $x_0 > x_1 > \dots$  indexed by the full set of natural numbers.*

**Proof.** (i) $\Rightarrow$ (ii'') $\Leftrightarrow$ (ii') $\Leftrightarrow$ (ii) is straightforward. Now assume (ii''), and suppose we had a nonempty subset  $Y \subseteq X$  with no minimal element. Take any  $x_0 \in Y$ . Since this is not minimal, we can find  $x_1 < x_0$ . Since this in turn is not minimal, we can find  $x_2 < x_1$ . Continuing this process, we get a contradiction to (ii'').  $\square$

**Definition 4.3.2.** A partially ordered set  $X$  is said to have descending chain condition (abbreviated “DCC”; also called “minimum condition” by some authors) if it satisfies the equivalent conditions of the above lemma.

Likewise, a partially ordered set  $X$  with the dual condition (every nonempty subset has a maximal element, equivalently,  $X$  has no infinite ascending chains) is said to have ascending chain condition (or “ACC” or “maximum condition”).

A well-ordered set means a totally ordered set with descending chain condition.

*Remark:* A chain in  $X$ , as defined following Definition 4.1.2, is a totally ordered subset, and it is meaningless to call such a subset “increasing” or “decreasing”. In the above lemma and definition, the phrases “descending chain” and “ascending chain” are used as shorthand for a totally ordered subset which can be indexed in a descending, respectively in an ascending manner by the natural numbers. (One may consider this a mixture of the order-theoretic meaning of “chain”, and the informal meaning, referring to a sequence of elements indexed by consecutive integers, with a specified relation holding between successive terms.) But note that though this shorthand is used in the fixed phrases “ascending chain condition” and “descending chain condition”, we made explicit what we meant by these phrases in the above lemma and definition.

That the natural numbers are well-ordered has been known in one form or another for millennia, but the importance of ACC and DCC for more general partially ordered sets was probably first noted in ring theory, in the early decades of the twentieth century. Rings with these conditions on their sets of *ideals*, partially ordered by inclusion, are called “Noetherian” and “Artinian” respectively, after Emmy Noether and Emil Artin who studied them.

One does not need to formally state a “principle of induction over partially ordered sets with ACC (or DCC)”. Rather, when one wishes to prove a result for all elements of a partially ordered set  $X$  with, say, DCC, one can simply begin, “Suppose there are elements of  $X$  for which the statement is false. Let  $x$  be minimal for this property”, since if the set of such elements is nonempty, it must have a minimal member. Then one knows the statement is *true* for all  $y < x$ ; and if one can show from this that it is true for  $x$  as well, one gets a contradiction, proving the desired result. Since this is a familiar form of argument, one often abbreviates it and says, “Assume inductively that the statement is true for all  $y < x$ ”, proves from this that it is true for  $x$  as well, and concludes that it is true for all elements of  $X$ .

In the most familiar sort of induction on the natural numbers, one starts by proving the desired result for 0 (or 1). Why was there no corresponding step in the schema described above? The analog of the statement that our desired result holds for 0 would be the statement that it holds for all *minimal* elements of  $X$ . But if one can prove that a statement is true for an element  $x$  whenever it is true for all smaller elements, then in particular, one can prove it in the case where the set of smaller elements is empty. Depending on the situation, the proof that a result is true for  $x$  if it is true for all smaller elements may or may not involve different arguments depending on whether  $x$  is a minimal element.

**Exercise 4.3:1.** A noninvertible element of a commutative integral domain  $C$  is called *irreducible* if it cannot be written as a product of two noninvertible elements. Give a concise proof that if  $C$  is a commutative integral domain with ascending chain condition on ideals (or even just on principal ideals), then every nonzero noninvertible element of  $C$  can be written as a product of irreducible elements.

In addition to *proofs* by induction, one often performs *constructions* in which each step requires

that a set of preceding steps already have been done. The definition of the Fibonacci numbers  $f_i$  ( $i = 0, 1, 2, \dots$ ) by the conditions

$$(4.3.3) \quad f_0 = 0, \quad f_1 = 1, \quad f_{n+2} = f_n + f_{n+1}$$

is of this sort. These are called *recursive* definitions or constructions, and we shall now see that, like inductive proofs, they can be carried out over general partially ordered sets with chain conditions.

Let us analyze what such a construction involves in general, and then show how to justify it. Suppose  $X$  is a partially ordered set with DCC, and suppose that we wish to construct recursively a certain function  $f$  from  $X$  to a set  $T$ . To say that for some  $x \in X$  the value of  $f$  has been determined for all  $y < x$  is to say that we have a function  $f_{<x}: \{y \mid y < x\} \rightarrow T$ . So “a rule defining  $f$  at each  $x$  if it is defined for all  $y < x$ ” can be formalized as a  $T$ -valued function  $r$  on the set of all pairs  $(x, f_{<x})$  consisting of an  $x \in X$  and a function  $f_{<x}: \{y \mid y < x\} \rightarrow T$ . In most applications, our rule defining  $f$  at  $x$  in terms of the values for  $y < x$  actually requires that these values satisfy some additional conditions, and we verify these conditions inductively, as the construction is described recursively. But to avoid complicating our abstract formalization, let us assume  $r$  defined for *all* pairs  $(x, f_{<x})$  where  $x \in X$  and  $f_{<x}$  is a function  $\{y \mid y < x\} \rightarrow T$ . For if we have a definition of  $r$  in “good” cases, we can extend it to other cases in an arbitrary way (e.g., assume  $0 \in T$  and send  $(x, f_{<x})$  to 0 if  $f_{<x}$  is not “good”). Then the inductive proof that  $f$  is “good” can be formally considered to come *after* the recursive construction of  $f$ .

We see that the property characterizing the function  $f$  constructed recursively as above is that for each  $x \in X$ ,  $f(x)$  is a certain function of the *restriction* of  $f$  to  $\{y \mid y < x\}$ . We recall a common notation for restrictions of functions: If  $f: X \rightarrow Y$  is a function, and  $Z$  is a subset of  $X$ , then the restriction of  $f$  to  $Z$ , a function  $Z \rightarrow Y$ , is denoted  $f \upharpoonright Z$ . (A variant symbol which we will not use is  $f \upharpoonright Z$ .)

We now justify recursive constructions by proving

**Lemma 4.3.4.** *Let  $X$  be a partially ordered set with descending chain condition,  $T$  any set, and  $r$  a function associating to every pair  $(x, f_{<x})$  such that  $x \in X$ , and  $f_{<x}$  is a function  $\{y \in X \mid y < x\} \rightarrow T$ , an element  $r(x, f_{<x}) \in T$ . Then there exists a unique function  $f: X \rightarrow T$  such that for all  $x \in X$ ,  $f(x) = r(x, f \upharpoonright \{y \mid y < x\})$ .*

**Proof.** Let  $X' \subseteq X$  denote the set of all  $x \in X$  for which there exists a unique function  $f_{\leq x}: \{y \mid y \leq x\} \rightarrow T$  with the property that

$$(4.3.5) \quad (\forall y \leq x) \quad f_{\leq x}(y) = r(y, f_{\leq x} \upharpoonright \{z \mid z < y\}).$$

We claim, first, that for any two elements  $x_0, x_1 \in X'$ , the functions  $f_{\leq x_0}, f_{\leq x_1}$  agree on  $\{y \mid y \leq x_0 \wedge y \leq x_1\}$ . For if not, choose a minimal  $y$  in this set at which they disagree. Then by (4.3.5),  $f_{\leq x_0}(y) = r(y, f_{\leq x_0} \upharpoonright \{z \mid z < y\})$ , and  $f_{\leq x_1}(y) = r(y, f_{\leq x_1} \upharpoonright \{z \mid z < y\})$ . But by choice of  $y$ , the restrictions of  $f_{\leq x_0}$  and  $f_{\leq x_1}$  to  $\{z \mid z < y\}$  are equal, hence by the above equations,  $f_{\leq x_0}(y) = f_{\leq x_1}(y)$ , contradicting our choice of  $y$ .

Next, suppose that  $X'$  were not all of  $X$ . Let  $x$  be a minimal element of  $X - X'$ . Since, as we have just seen, the functions  $f_{\leq y}$  for  $y < x$  agree on the pairwise intersections of their domains, they piece together into one function  $f_{<x}$  on the union of their domains. (Formally, this “piecing together” means taking the union of these functions, as subsets of  $X \times T$ .) If we now

define  $f_{\leq x}$  to agree with this function  $f_{< x}$  on  $\{y \mid y < x\}$ , and to have the value  $r(x, f_{< x})$  at  $x$ , we see that this function satisfies (4.3.5), and is the unique function on  $\{y \mid y \leq x\}$  which can possibly satisfy that condition. This means  $x \in X'$ , contradicting our choice of  $x$ .

Hence  $X' = X$ . Now piecing together these functions  $f_{\leq x}$  defined on the sets  $\{y \mid y \leq x\}$ , we get the desired function  $f$  defined on all of  $X$ .  $\square$

Note that the above proof was an *inductive* verification that every element of  $X$  satisfies the condition of the first sentence of the proof. So recursion is justified by induction.

Example: The Fibonacci numbers are defined recursively by using for  $X$  the ordered set of nonnegative integers, and defining  $r(n, (f_0, \dots, f_{n-1}))$  to be 0 if  $n = 0$ , to be 1 if  $n = 1$ , and to be  $f_{n-2} + f_{n-1}$  if  $n \geq 2$ .

The next exercise shows that recursive constructions are not in general possible if the given partially ordered set does not satisfy descending chain condition.

**Exercise 4.3:2.** Show that there does not exist a function  $f$  from the interval  $[0, 1]$  of the real line to the set  $\{0, 1\}$  determined by the following rules:

(a)  $f(0) = 0$ .

(b) For  $x > 0$ ,  $f(x) = 1$  if for all  $y \in [0, x)$ ,  $f(y) = 0$ ; otherwise,  $f(x) = 0$ .

If you prefer, you may replace the interval  $[0, 1]$  in this example with the countable set  $\{0\} \cup \{1/n \mid n = 1, 2, 3, \dots\}$ .

**Exercise 4.3:3.** Generalizing the above example, show that if  $(X, \leq)$  is any partially ordered set *not* satisfying descending chain condition, then

(i) There exists a function  $r$  as in the statement of Lemma 4.3.4, with  $T = \{0, 1\}$ , such that there is no function  $f$  satisfying the conditions of the conclusion of that lemma.

(ii) There exists a function  $r$  as in the statement of Lemma 4.3.4, with  $T = \{0, 1\}$ , such that there are more than one functions  $f$  satisfying the conditions of the conclusion of that lemma.

In a way, solving a differential equation with given initial conditions is like a “recursive construction over an interval of the real numbers”. But precisely because the real numbers do not have descending chain condition, the conditions for existence and uniqueness of a solution, and the arguments needed to prove these, are subtle. (A fact that often plays a role like induction in such arguments is the connectedness of the real line.)

There is a situation at the very foundation of mathematics which can be interpreted in terms of a partially ordered system with descending chain condition. The Axiom of Regularity of set theory (which will be stated formally in the next section) says that there is no “infinite regress” in the construction of sets; that is, that there are no left-infinite chains of sets under the membership relation:

$$\dots \in S_n \in \dots \in S_2 \in S_1 \in S_0.$$

This is not a difficult axiom to swallow, since if we had a set theory for which it was not true, we could pass to the “smaller” set theory consisting of those sets which admit no such chain to the left of them. The class of such sets would be closed under all the constructions required by the remaining axioms of set theory, and the “new” set theory would satisfy the Axiom of Regularity.

To interpret Regularity in the terms we have just been discussing, let us write  $A \prec B$ , for sets  $A$  and  $B$ , if there is a chain of membership-relations,  $A = S_0 \in S_1 \in \dots \in S_n = B$  ( $n > 0$ ). This relation is clearly transitive. The Regularity Axiom implies that  $\prec$  is antireflexive (if we had  $A \prec A$ , then a chain of membership relations connecting  $A$  with itself could be iterated to give an

infinite chain going to the left), hence  $\prec$  is the partial strict ordering corresponding to a partial ordering  $\preccurlyeq$ ; and Regularity applied again says that this partial ordering has descending chain condition. (Well, almost. We have only defined the concepts of partial ordering and chain condition for *sets*, and the class of all sets is not a set. To get around this problem we can translate these observations more precisely as saying that for each set  $A$ ,  $\{B \mid B \preccurlyeq A\}$  is itself a set, and has descending chain condition under  $\preccurlyeq$ .) This allows one to prove set-theoretic results inductively, and make set-theoretic definitions recursively.

We had another such situation in Chapter 1, when we talked about the set  $T = T_{X, \mu, \iota, e}$  of group-theoretic terms in a set of symbols  $X$ . These also satisfy a principle of regularity, in terms of the relation “ $s$  occurs in  $t$ ”, which we denoted  $t \succ s$  in Exercise 1.7:4. To show this, let  $T'$  denote the set of elements of  $T$  admitting no infinite descending  $\succ$ -chains to the right of them. One verifies that  $T'$  is closed under the operations of conditions (a) and (b) of the definition of  $T$  (in §1.5), and concludes that if  $T'$  were properly smaller than  $T$ , one would have a contradiction to condition (c) of that definition. We only sketched the construction of  $T$  in Chapter 1, but in §8.3 below we will introduce the concept of “term” for general classes of algebras, and the above argument will then enable us to perform recursion and induction on such terms.

One can, of course, do inductive proofs and recursive constructions over partially ordered sets with *ascending* as well as descending chain condition. These come up often in ring theory, where Noetherian rings, i.e., rings whose partially ordered set of ideals has ACC, are important. In proving that a property holds for an arbitrary ideal  $I$  of such a ring, one may assume inductively that it is true for all strictly *larger* ideals. To get the result allowing us to perform recursive constructions in such situations, i.e., the analog of Lemma 4.3.4 with  $>$  replacing  $<$ , it is not necessary to repeat the proof of that lemma; we can use duality of partially ordered sets. I will give the statement and sketch the argument this once, to show how an argument by duality works. After this, if I want to invoke the dual of an order-theoretic result previously given, I shall consider it sufficient to say “by duality”, or “by the dual of Proposition #.#.#” or the like.

**Corollary 4.3.6.** *Let  $X$  be a partially ordered set with ascending chain condition,  $T$  any set, and  $r$  a function associating to every pair  $(x, f_{>x})$  consisting of an element  $x \in X$  and a function  $f_{>x}: \{y \mid y > x\} \rightarrow T$  an element  $r(x, f_{>x}) \in T$ . Then there exists a unique function  $f: X \rightarrow T$  such that for all  $x \in X$ ,  $f(x) = r(x, f|_{\{y \mid y > x\}})$ .*

**Sketch of Proof.** The *opposite* of the partially ordered set  $X$  (the structure with the same underlying set but the opposite ordering) is a partially ordered set  $X^{\text{op}}$  with *descending chain condition*, and  $r$  can be considered to be a function  $r'$  with exactly the properties required to apply Lemma 4.3.4 to that partially ordered set. That lemma gives us a unique function  $f'$  from  $X^{\text{op}}$  to  $T$  satisfying the conclusions of that lemma relative to  $r'$ , and this is equivalent to a function  $f$  from  $X$  to  $T$  satisfying the desired condition relative to  $r$ .  $\square$

Example: The Fibonacci numbers  $f_n$  were defined above for  $n \geq 0$ . With the help of downward recursion on the set of negative integers, one can now easily verify that there is also a unique way of defining  $f_n$  for negative  $n$ , such that combining the values for positive and negative  $n$ , we get a sequence  $(f_i)_{i \in \mathbb{Z}}$  which satisfies  $f_n = f_{n-2} + f_{n-1}$  for all  $n$ .

Often the key to making an inductive argument or a recursive construction work is a careful choice of a parameter over which to carry out the induction or recursion, and an appropriate ordering on the set of values of that parameter. The next definition describes a way of constructing

partial orderings that is frequently useful for such purposes. The well-ordered index set  $I$  in that definition can be as simple as  $\{0, 1\}$ .

**Definition 4.3.7.** Let  $(X_i)_{i \in I}$  be a family of partially ordered sets, indexed by a well-ordered set  $I$ . Then lexicographic order on  $\prod_I X_i$  is defined by declaring  $(x_i) \leq (y_i)$  to hold if and only if either  $(x_i) = (y_i)$ , or for the least  $j \in I$  such that  $x_j \neq y_j$ , one has  $x_j < y_j$  in  $X_j$ .

Note that if  $I = \{1, \dots, n\}$  with its natural order, then this construction orders  $n$ -tuples  $(x_1, \dots, x_n) \in \prod_I X_i$  by the same “left-to-right” principle that is used to arrange words in the dictionary; hence the name of the construction. The usefulness of this construction in obtaining orderings with descending chain condition is indicated in part (iii) of

**Exercise 4.3:4.** Let  $(X_i)_{i \in I}$  be as in Definition 4.3.7.

- (i) Verify that the relation on  $\prod_I X_i$  given by that definition is indeed a partial order.
- (ii) Show that if each  $X_i$  is *totally* ordered, then so is their direct product under that ordering. Show, in contrast, that the corresponding statement is not in general true for the *product ordering*, described in Definition 4.1.4.
- (iii) Show that if  $I$  is finite and each of the  $X_i$  has descending (or ascending) chain condition, then so does their product under lexicographic ordering.
- (iv) Comparing lexicographic ordering with the product ordering, deduce that a direct product of finitely many partially ordered sets with descending chain condition satisfies descending chain condition under the product ordering as well.
- (v) Show that the product of a family of copies of the two-element totally ordered set  $\{0, 1\}$ , indexed by the natural numbers, does not have descending chain condition under the product ordering. Deduce that lexicographic ordering on products of infinite families of partially ordered sets with descending chain condition also fails, in general, to have descending chain condition.

In the next exercise, a lexicographic ordering is used to give a concise proof of a standard result on symmetric polynomials.

**Exercise 4.3:5.** Let  $R$  be a commutative ring, and  $R[x_1, \dots, x_n]$  the polynomial ring in  $n$  indeterminates over  $R$ . Given any nonzero polynomial  $f = \sum c_{i(1), \dots, i(n)} x_1^{i(1)} \dots x_n^{i(n)}$  (almost all  $c_{i(1), \dots, i(n)}$  zero), let us define the *leading term* of  $f$  to be the nonzero summand in this expression with the largest exponent-string  $(i(1), \dots, i(n))$  under lexicographic ordering on the set of all such strings. (Since the set of nonzero summands is finite, no chain condition is needed to make this definition.)

- (i) Let  $f$  and  $g$  be nonzero elements of  $R[x_1, \dots, x_n]$ , and suppose that the coefficient in the leading term of  $f$  is not a zero-divisor in  $R$ . (E.g., this is automatic if  $R$  is an integral domain.) Show that the leading term of  $fg$  is the product of the leading terms of  $f$  and of  $g$ .

An element of  $R[x_1, \dots, x_n]$  is called *symmetric* if it is invariant under the natural action of the group of all permutations of the index set  $\{1, \dots, n\}$  on the indeterminates  $x_1, \dots, x_n$ . For  $1 \leq d \leq n$ , the  $d$ th *elementary symmetric function*  $s_d$  is defined to be the sum of all products of exactly  $d$  distinct members of  $\{x_1, \dots, x_n\}$ .

- (ii) For nonnegative integers  $j(1) \dots j(n)$ , find the leading term of the product  $s_1^{j(1)} \dots s_n^{j(n)}$ , and the coefficient of that leading term in that product.
- (iii) Show that the following sets are the same: (a) The set of all  $n$ -tuples  $(i(1), \dots, i(n))$  of nonnegative integers such that  $i(1) \geq \dots \geq i(n)$ . (b) The set of all exponent-strings  $(i(1), \dots, i(n))$  of leading terms  $c_{i(1), \dots, i(n)} x_1^{i(1)} \dots x_n^{i(n)}$  of symmetric polynomials. (c) The set of all exponent-strings of leading terms of products  $s_1^{j(1)} \dots s_n^{j(n)}$ , as in (ii).

(iv) Deduce that any nonzero symmetric polynomial can be changed to a symmetric polynomial with lower exponent-string-of-the-leading-term, or to the zero polynomial, by subtracting a scalar multiple of a product of elementary symmetric polynomials. Conclude, by induction on this exponent-string, that the ring of symmetric polynomials in  $n$  indeterminates over  $R$  is generated over  $R$  by the elementary symmetric polynomials.

(For standard proofs of the above result, see [29, pp.252-255], or [31, Theorem IV.6.1, p.191]. For some related results on noncommutative rings, see [51].)

**Exercise 4.3:6.** For nonnegative integers  $i$  and  $j$ , let  $n_{i,j}$  be defined recursively as the least nonnegative integer not equal to  $n_{i,j'}$  for any  $j' < j$ , nor to  $n_{i',j}$  for any  $i' < i$ . (What ordering of the set of pairs  $(i,j)$  of nonnegative integers can one use to justify this recursion?)

Find and prove a concise description of  $n_{i,j}$ . (Suggestion: Calculate some values and note patterns. To find the “pattern in the patterns”, write numbers to base 2.)

We end with two miscellaneous exercises.

**Exercise 4.3:7.** For  $X$  a topological space, show that the following conditions are equivalent. (We do not understand “compact” to entail “Hausdorff” or “nonempty”.) (a) Every subset of  $X$  is compact in the induced topology. (b) Every open subset of  $X$  is compact in the induced topology. (c) The partially ordered set of open subsets of  $X$  has ascending chain condition.

**Exercise 4.3:8.** One may ask whether Exercise 4.3:1 has a converse: that if  $C$  is a commutative integral domain in which every nonzero noninvertible element can be written as a product of irreducibles, then  $C$  has ascending chain condition on principal ideals. Show by example that this is *not* true.

**4.4. The axioms of set theory.** We are soon going to look at some order-theoretic principles equivalent to the powerful Axiom of Choice. Hence it is desirable to review the statement of that axiom, and its status in relation to the other axioms of set theory. For completeness, I will record in this section the whole set of axioms most commonly used by set theorists.

Let us begin with some background discussion. In setting up a rigorous foundation for mathematics, one might expect the theory to require several sorts of “entities”: “primitive” elements such as numbers, additional *sets* formed out of these, *ordered pairs* of elements, *functions* from one set to another, etc.. But as the theory was developed, it turned out that one could get everything one wanted from a single basic concept, that of set, and a single relation among sets, that of membership. The result is a set theory in which the only members of sets are themselves sets.

As an important example of how other “primitives” are reduced to the set concept, we recall the case of the *natural numbers* (nonnegative integers). The first thing we learn in our childhood about these numbers is that they are used to count things; to say how many objects there are in a collection. The early set theorists observed that one can formalize the concept of two sets having the “same number” of elements set-theoretically, as meaning that there exists a bijection between them. This is clearly an equivalence relation on sets. Hence the natural numbers ought be some entities which one could associate to finite sets, so that two sets would get the same entity associated to them if and only if they were in the same equivalence class. The original plan was to use, as those entities, the equivalence classes themselves, i.e., to *define* the natural numbers  $0, 1, 2, \dots$ , to be the corresponding equivalence classes. Thus, the statement that a finite set had  $n$  elements would mean that it was a *member* of the number  $n$ . (Cardinalities of infinite sets were to be treated similarly.) This is good in principle – don’t create new entities to index the equivalence classes if the equivalence classes themselves will do. But in this case, the equivalence classes

turned out not to be a good choice: they are too big to be sets. So the next idea was to choose one easily described member from each such class, call these chosen elements the natural numbers  $0, 1, 2, \dots$ , and define a set to have  $n$  elements if it could be put in bijective correspondence with the ‘sample’ set  $n$ .

Where would one get these ‘sample’ finite sets from, using pure set theory? There is no problem getting a sample 0-element set – there is a unique set with 0 elements, the empty set  $\emptyset$ . Having taken this step, we have *one* set in hand –  $\emptyset$ . This means that we are in a position to create a sample one-element set, the set with that element as its one member, i.e.,  $\{\emptyset\}$ . Having found these two elements,  $\emptyset$  and  $\{\emptyset\}$ , we can define a 2-element set  $\{\emptyset, \{\emptyset\}\}$  to use as our next sample – and so on. After the first couple of steps, we are not so limited in our options. (For example, we could take for 4 the set of all subsets of 2.) However, the above approach, of always taking for the next number the set of numbers found so far, due to John von Neumann, is an elegant way of manufacturing one set of each natural-number cardinality, and it is taken as the definition of these numbers by modern set theorists:

$$(4.4.1) \quad \begin{aligned} 0 &= \emptyset, & 1 &= \{\emptyset\}, & 2 &= \{\emptyset, \{\emptyset\}\}, & 3 &= \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \\ & & & & & & & \dots \\ & & & & & & & i+1 = i \cup \{i\} = \{0, 1, 2, \dots, i\} \\ & & & & & & & \dots \end{aligned}$$

Another basic concept which was reduced to the concepts of set and membership is that of *ordered pair*. If  $X$  and  $Y$  are sets, then one can deduce from the axioms (shortly to be listed) that  $X$  and  $Y$  can each be determined uniquely from the set  $\{\{X\}, \{X, Y\}\}$ . Since all one needs about ordered pairs is that they are objects which specify their first and second components unambiguously, one *defines* the ordered pair  $(X, Y)$  to mean the set  $\{\{X\}, \{X, Y\}\}$ .

One then goes on to define the *direct product* of two sets in terms of ordered pairs, binary relations in terms of direct products, functions in terms of binary relations, etc.. From natural numbers, ordered pairs, and functions, one constructs the integers, the rational numbers, the real numbers, the complex numbers, etc., by well-known techniques, which I won’t review here.

(One also wants to define ordered  $n$ -tuples. The trick by which ordered pairs were defined turns out not to generalize in an easy fashion; the most convenient approach is to define an ordered  $n$ -tuple to mean a function whose domain is the set  $n$ . However, this conflicts with the definition of ordered pair! To handle this, a careful development of set theory must use different symbols, say  $\langle X, Y \rangle$  for the concept of ‘ordered pair’ first described, and  $(X_0, X_1, \dots, X_{n-1})$  for the ordered  $n$ -tuples subsequently defined.)

The above examples should give some motivation for the ‘sort’ of set theory described by the axioms which we shall now list. Of course, a text on the foundations of mathematics will first develop language allowing one to state these axioms precisely, and, since a statement in such language is not always easy to understand, it will precede or follow many of the precise statements by intuitive developments. I have tried below to give formulations that make it as clear as possible what the axioms assert, and have added some further remarks after the list. But for a thorough presentation, and for more discussion of the axioms, the student should see a text on the subject. Two recommended undergraduate texts are [12] and [21]. Written for a somewhat more advanced audience is [19].

Here, now, are the axioms of *Zermelo-Fraenkel Set Theory with the Axiom of Choice*, commonly abbreviated ZFC.

**Axiom of Extensionality:** *Sets are equal if and only if they have the same members, i.e.,  $X = Y$  if and only if for every set  $A$ ,  $A \in X \Leftrightarrow A \in Y$ .*

**Axiom of Regularity** (or **Well-foundedness**, or **Foundation**): *For every nonempty set  $X$ , there is a member of  $X$  which is disjoint from  $X$ .*

**Axiom of the Empty Set:** *There exists a set with no members. (Common notation:  $\emptyset$ .)*

**Axiom of Separation:** *If  $X$  is a set and  $P$  is a condition on sets, there exists a set  $Y$  whose members are precisely the members of  $X$  satisfying  $P$ . (Common notation:  $Y = \{A \in X \mid P(A)\}$ .)*

**Axiom of Doubletons** (or **Pairs**): *If  $X$  and  $Y$  are sets, there is a set  $Z$  whose only members are  $X$  and  $Y$ . (Common notation:  $Z = \{X, Y\}$ .)*

**Axiom of Unions:** *If  $X$  is a set, there is a set  $Y$  whose members are precisely all members of members of  $X$ . (Common notations:  $Y = \cup X$  or  $\cup_{A \in X} A$ .)*

**Axiom of Replacement:** *If  $f$  is an operation on sets (formally characterized by a set-theoretic proposition  $P(A, B)$  such that for every set  $A$  there is a unique set  $f(A)$  such that  $P(A, f(A))$  holds) and  $X$  is a set, then there exists a set  $Y$  whose members are precisely the sets  $f(A)$  for  $A \in X$ . (Common notation:  $Y = \{f(A) \mid A \in X\}$ . When there is no danger of confusion, this is sometimes abbreviated to  $Y = f(X)$ .)*

**Axiom of the Power Set:** *If  $X$  is a set, there exists a set  $Y$  whose members are precisely all subsets of  $X$ . (Common notations:  $Y = \mathbf{P}(X)$  or  $2^X$ .)*

**Axiom of Infinity:** *There exists a set having  $\emptyset$  as a member, closed under the construction  $i \mapsto i \cup \{i\}$  (cf. (4.4.1)), and minimal for these properties. (Common name: The set of natural numbers.)*

**Axiom of Choice:** *If  $X$  is a set, and  $f$  is a function associating to every  $x \in X$  a nonempty set  $f(x)$ , then there exists a function  $g$  associating to every  $x \in X$  an element  $g(x) \in f(x)$ .*

Explanations of some of the names: *Extensionality* means that a set is determined by its *extent*, not its *intent*. *Separation* says that one can form new sets by using any well-defined criterion to “separate out” certain elements of an existing set. The Axiom of *Infinity* is so called because if we did not assume it, the collection of all sets which can be built up from the empty set in finitely many steps would satisfy our axioms, giving an example of a set theory in which all sets are finite. So the axiom is equivalent to the statement that there exists an infinite set.

We described *Regularity* earlier as saying that there was no infinite regress under “ $\in$ ”. That formulation requires one to have the set of natural numbers to index such a regress, so we chose a formulation that can be expressed independently of the Axiom of Infinity. In the presence of the other axioms one can prove the two formulations equivalent. (Roughly, if one had an infinite chain  $\dots \in S_2 \in S_1 \in S_0$ , then  $\{S_i\}$  would be a counterexample to Regularity, while if a set  $X$  were a counterexample to Regularity, one could select such a chain from its elements.)

Actually the Axiom of Regularity makes little substantive difference for areas of mathematics other than set theory itself (e.g., see [21, p.92 et seq.]). Without it, one can have sets with exotic properties such as being members of themselves, but the properties of set-theoretic concepts used by most of mathematics – bijections, direct products, cardinality arguments, etc. – are little affected. Its absence would simply make it a bit trickier to construct, say, a family of *disjoint* copies of a given set. The Regularity Axiom seems to have crept into the Zermelo-Fraenkel axioms by the back door: It was not in the earlier formulations of those axioms, and still does not appear in some listings, such as that in [12]. But it is generally accepted, and we will count it among the axioms here, and rely on the convenience it provides. It gives one a comforting assurance that sets are

built up from earlier sets with no “vicious circles” in the process; hence the name “Well-Foundedness”. (By extension, many set-theorists call the condition of descending chain condition on any partially ordered set “well-foundedness”.)

Observe that the Axioms of Extensionality and Regularity essentially clarify what we intend to *mean* by a “set”.

The next seven axioms each say that certain sets exist. In each case these are sets which are *uniquely determined* by the conditions assumed. Those seven axioms can all be considered cases of a single axiom proposed by Frege in 1893, the *Axiom of Abstraction*, saying that “Given any property, there exists a set whose members are just those entities possessing that property”. That axiom nicely embodies the idea of a set, but it turned out to be too strong to be consistent: it allowed one to define things like “the set  $S$  of all sets which are not members of themselves,” which led to contradictions. (Russell’s Paradox: “Is that  $S$  a member of itself?” Either a positive or a negative answer implies its own negation.) The difficulty was, somehow, that the Axiom of Abstraction assumed “all sets” to be known, so that, in particular, the set  $S$  we were constructing was already “there”, to be chosen or rejected in choosing the members of that same set  $S$ , allowing one to create a self-contradictory criterion for that choice. Subsequent experience suggested that such contradictions could be avoided by requiring every set to be constructed from sets “constructed before it”; and the seven axioms in question represent sub-cases of the rejected “Axiom of Abstraction” which meet this condition.

(The Axiom of Regularity was probably another reaction against those paradoxes. Though adding an axiom can’t remove a contradiction, the encounter with Russell’s Paradox very likely led mathematicians to feel that sets that could not be “built up from scratch” were unhealthy, and should be excluded.)

The last axiom of ZFC, that of “choice”, is of a different sort from those that precede it. It asserts the existence of something not uniquely defined by the given data: a function that chooses, *in an unspecified way*, one element from each of a family of sets. This was very controversial in the early decades of the twentieth century, both because it led to consequences which seemed surprising then (such as the existence of nonmeasurable sets of real numbers), and because of a feeling by some that it represented an unjustifiable assumption that something one could do in the finite case could be done in the infinite case as well. It is a standard assumption in modern mathematics; such basic results as that every vector space has a basis, that a direct product of compact topological spaces is compact, and that a countable union of countable sets is countable cannot be proved without it. But there have been, and still are, mathematicians who reject it: the *intuitionists* of the early 1900’s, and the *constructivists* today.

Even accepting the Axiom of Choice, as we shall, it is at times instructive to note whether a result or an argument depends on it, or can be obtained from the other axioms. (This is like the viewpoint that, even if one does not accept the constructivists’ extreme claim that proofs of existence that do not give explicit constructions are *worthless*, one may consider constructive proofs to be desirable when they can be found.)

In the next two sections we shall develop several set-theoretic results whose proofs require the Axiom of Choice, and we will show that each of these statements is, in fact, equivalent to that axiom, in the presence of the other axioms. Hence, in those sections, we shall not assume the Axiom of Choice except when we state this assumption explicitly, and the arguments we give to show these equivalences will all be justifiable in terms of the theory given by the other axioms, called “Zermelo-Fraenkel Set Theory”, abbreviated ZF. (However, we shall not in general attempt

to show explicitly how the familiar mathematical techniques that we use are justified by those axioms – for that, again see a text in set theory.) In all later chapters, on the other hand, we shall freely use the Axiom of Choice, i.e., we will assume ZFC.

In the handful of results proved so far in this chapter, we have implicitly used the Axiom of Choice just once: in Lemma 4.3.1, in showing (ii'') $\Rightarrow$ (i). Hence for the remainder of this chapter, we shall forgo assuming that implication, and will understand the descending chain condition to refer to condition (i) of that lemma (which still *implies* (ii)-(ii'')).

Let us note explicitly one detail of set-theoretic language we have already used: Since all sets satisfying a given property may not together form a set, one needs a word to refer to “collections” of sets that are not necessarily themselves sets. These are called *classes*. An example is the *class of all sets*. One can think of classes which are not sets, not as actually being *mathematical objects*, but as providing a convenient *language* to use in making statements about all sets having one or another property.

Since classes are more general than sets, one may refer to any set as a “class”, and this is sometimes done for reasons not involving the logical distinction, but just to vary the wording. E.g., rather than saying “the set of those subsets of  $X$  such that ...”, one sometimes says “the class of those subsets of  $X$  such that ...”. And, for some reason, one always says “equivalence class”, “conjugacy class”, etc., though they are sets.

**Exercise 4.4:1.** Show that for every set  $X$  there exists a set  $Y \supseteq X$  such that  $a \in b \in Y \Rightarrow a \in Y$ ; in fact, that there exists a *least* such set. (Thus,  $Y$  is the “closure” of  $X$  under passing to members of sets.)

**4.5. Well-ordered sets and ordinals.** Recall (Definition 4.3.2) that a partially ordered set  $(X, \leq)$  is called *well-ordered* if it is totally ordered and has descending chain condition. In a totally ordered set, a minimal element is the same as a least element, so the condition of well-ordering says that every nonempty subset of  $X$  has a *least* member.

This condition goes a long way toward completely determining the structure of  $X$ . Applied first to  $X$  as a subset of itself, it tells us that if  $X$  is nonempty, it has a least element,  $x_0$ . If  $X$  does not consist of  $x_0$  alone, then  $X - \{x_0\}$  is nonempty, hence this set has a least element, which we may call  $x_1$ . We can go on in this fashion, and unless  $X$  is finite we will get a uniquely determined sequence of elements  $x_0 < x_1 < x_2 < x_3 < \dots$  at the “bottom” of  $X$ . This list may exhaust  $X$ , but if it does not, there will necessarily be a least element in the complement of the subset so far described, which we may call  $x_{1,0}$ , and if this still does not exhaust  $X$ , there will be a least element greater than it,  $x_{1,1}$ , etc.. We can construct in this way successive hierarchies, and hierarchies of hierarchies – I will not go into details – on the single refrain, “If this is not all, there is a least element of the complement”.

A couple of concrete examples are noted in

**Exercise 4.5:1.** If  $f$  and  $g$  are real-valued functions on the real line  $\mathbb{R}$ , let us in this exercise write  $f \leq g$  to mean that there exists some real number  $N$  such that  $f(t) \leq g(t)$  for all  $t \geq N$ .

(i) Show that this relation  $\leq$  is a preordering, that its restriction to the set of polynomial functions is a total ordering, and that on polynomials with *nonnegative integer* coefficients, it in fact gives a well-ordering. Determine, if they exist, the elements  $x_0, x_1, \dots, x_n, \dots, x_{1,0}, x_{1,1}$  of this set, in the notation of the preceding paragraphs.

(ii) Show that the set consisting of all polynomials with nonnegative integer coefficients, and also the function  $e^t$ , is still well-ordered under the above relation.

- (iii) Find a subset of the rational numbers which is order-isomorphic (under the standard ordering) to the set described in (ii).

To make precise the idea that the order structure of a well-ordered set is “unique, as far as it goes”, let us define an “initial segment” of any totally ordered set  $X$  to mean a subset  $I \subseteq X$  such that  $x \leq y \in I \Rightarrow x \in I$ . Then we have

**Lemma 4.5.1.** *Let  $X$  and  $Y$  be well-ordered sets. Then exactly one of the following conditions holds:*

- (i)  $X$  and  $Y$  are order-isomorphic.
- (ii)  $X$  is order-isomorphic to a proper initial segment of  $Y$ .
- (iii)  $Y$  is order-isomorphic to a proper initial segment of  $X$ .

*Further, in (ii) and (iii) the initial segments in question are unique, and in all three cases the isomorphism is unique.*

**Proof.** We shall construct an order isomorphism of one of these three types by a recursive construction on the well-ordered set  $X$ . Let me first describe the idea intuitively: We start by pairing the least element of  $X$  with the least element of  $Y$ ; and we go on, at every stage pairing the least not-yet-paired-off element of  $X$  with the least not-yet-paired-off element of  $Y$ , until we run out of elements of either  $X$  or  $Y$  or both.

Now in our formulation of recursive constructions in Lemma 4.3.4, we said nothing about “running out of elements”. But we can use a trick to reduce the approach just sketched to a recursion of the sort characterized by that lemma.

Form a set consisting of the elements of  $Y$  and one additional element which we shall denote  $\text{DONE}$ . Given any  $x \in X$ , and any function  $f_{<x}: \{x' \in X \mid x' < x\} \rightarrow Y \cup \{\text{DONE}\}$ , we define  $r(x, f_{<x}) \in Y \cup \{\text{DONE}\}$  as follows:

If the image of  $f_{<x}$  is a proper initial segment of  $Y$ , let  $r(x, f_{<x})$  be the least element of  $Y$  not in that segment. Otherwise, let  $r(x, f_{<x}) = \text{DONE}$ .

By Lemma 4.3.4 this determines a function  $f: X \rightarrow Y \cup \{\text{DONE}\}$ . It is straightforward to verify inductively that for those  $x$  such that  $f(x) \neq \text{DONE}$ , the restriction  $f_{\leq x}$  of  $f$  to  $\{x' \in X \mid x' \leq x\}$  will be the only order isomorphism between that initial segment of  $X$  and any initial segment of  $Y$ . From this we easily deduce that if the range of  $f$  does not contain the value  $\text{DONE}$ , exactly one of conclusions (i) or (ii) holds, but not (iii); while if the range of  $f$  contains  $\text{DONE}$ , (iii) holds but not (i) or (ii). In each case,  $f$  determines the unique order isomorphism with the indicated properties.  $\square$

**Exercise 4.5:2.** Give the details of the last paragraph of the above proof.

Since the well-ordered sets fall into such a neat array of isomorphism classes, it is natural to look for a way of choosing one “standard member” for each of these classes, just as the natural numbers are used as “standard members” for the different sizes of finite sets. Recall that in the von Neumann construction of the natural numbers, (4.4.1), each number arises as the set of all those that precede it, so that we have  $i < j$  if and only if  $i \in j$ , and  $i \leq j$  if and only if  $i \subseteq j$ . Clearly, each natural number, being finite and totally ordered, is a well-ordered set under this ordering. Let us take these von Neumann natural numbers as our standard examples of finite well-ordered sets, and see whether we can extend this family in a natural way to get models of infinite

well-ordered sets.

Following the principle that each new object should be the set of all that precede, we use *the set of natural numbers* as the standard example chosen from among the well-ordered sets which when listed in the manner discussed at the beginning of this section have the form  $X = \{x_0, x_1, \dots\}$ , with subscripts running over the natural numbers but nothing beyond those. Set theorists write this object

$$\omega = \{0, 1, 2, \dots, i, \dots\}.$$

The obvious representative for those sets having an initial segment isomorphic to  $\omega$ , and just one element beyond that segment, is written

$$\omega + 1 = \omega \cup \{\omega\} = \{0, 1, 2, \dots, i, \dots; \omega\}.$$

We likewise go on to get  $\omega + 2$ ,  $\omega + 3$  etc.. The element coming after all the  $\omega + i$ 's ( $i \in \omega$ ) is denoted  $\omega + \omega$  or  $\omega^2$ . (We will see later why it is not written "more naturally" as  $2\omega$ .) After the elements  $\omega^2 + i$  ( $i \in \omega$ ) comes  $\omega^3$ ; ... after all the elements of the form  $\omega^i$  ( $i \in \omega$ ) one has  $\omega\omega = \omega^2$ . In fact, one can form arbitrary "polynomials" in  $\omega$  with natural number coefficients, and the set of these has just the order structure that was given to the polynomials with such coefficients in Exercise 4.5:1 (though the natural number coefficients in our "polynomials" in  $\omega$  are written on the right). Then the set of *all* these polynomials in  $\omega$  is taken as the next standard sample well-ordered set...

So far, we have been sketching an idea; let us make it precise. A terminological observation first. If  $X$  is any well-ordered set, and  $\alpha$  the "standard" well-ordered set (to be constructed) that is order-isomorphic to it, then Lemma 4.5.1 shows us how to index the elements of  $X$  by the members of  $\alpha$  – i.e., by those "standard" well-ordered sets smaller than  $\alpha$ . Thus, the well-ordered sets less than  $\alpha$  serve as translations and (starting with  $\omega$ ) generalizations of the sequence of words "first, second, third, ..." which are used in ordinary language to index the elements of finite totally ordered sets. Hence, the term *ordinals*, used by grammarians for those words, is used by mathematicians for the "standard samples" of isomorphism types of well-ordered sets. Now for the formal definition.

**Definition 4.5.2.** An ordinal (or von Neumann ordinal) is a set  $\alpha$  such that  $\gamma \in \beta \in \alpha \Rightarrow \gamma \in \alpha$ , and such that if  $\beta \in \alpha$  and  $\gamma \in \alpha$ , then either  $\beta = \gamma$ , or  $\beta \in \gamma$ , or  $\gamma \in \beta$ .

It is easily deduced from the above definition that every member  $\beta$  of an ordinal  $\alpha$  is again an ordinal. (Key step: We need to know that  $\delta \in \gamma \in \beta$  implies  $\delta \in \beta$ . Here both  $\beta$  and  $\delta$  are members of  $\alpha$ , so if the desired conclusion did not hold, the last part of the statement that  $\alpha$  is an ordinal would imply either  $\beta = \delta$ , or  $\beta \in \delta$ . But combining either of these with the relations  $\delta \in \gamma \in \beta$ , we could get a contradiction to the Regularity Axiom.) Given this observation, the first part of the definition says that the relation " $\in$ " on the elements of an ordinal is transitive. By the Regularity Axiom, this relation is antisymmetric and antireflexive; hence the relation " $\in$  or =" will be a partial order. The second condition in the above definition makes that relation a *total* ordering, and again by the Regularity Axiom, this will be a well-ordering. (If one does not assume the Regularity Axiom, one adds to the definition of ordinal the various conditions deduced above from that axiom.)

The ordinals themselves *almost* form a well-ordered set under the relation " $\in$  or =". The only trouble is that they do not form a set! Here are the basic facts.

**Proposition 4.5.3.** (i) *Every member of an ordinal is an ordinal.*

(ii) *If  $\alpha$  and  $\beta$  are ordinals, then the following conditions are equivalent:*

(a)  $\alpha = \beta$  or  $\alpha \in \beta$ ,

(b)  $\alpha \subseteq \beta$ .

(iii) *If  $\alpha$  and  $\beta$  are ordinals, and  $\alpha \subseteq \beta$ , then  $\alpha$  is an initial segment of  $\beta$ . If it is a proper initial segment, it is also the least element of  $\beta$  not in that initial segment.*

(iv) *For any two ordinals  $\alpha$  and  $\beta$  one has either  $\alpha \subseteq \beta$  or  $\beta \subseteq \alpha$ , and every nonempty class of ordinals has a “ $\subseteq$ -least” member. (In other words, the class of ordinals satisfies the analog of the set-theoretic property of well-orderedness under  $\subseteq$ .) In particular, every ordinal, and more generally every set of ordinals, is well-ordered under  $\subseteq$ .*

(v) *The union of any set of ordinals is an ordinal.*

(vi) *The class of all ordinals is not a set.*

(vii) *Every well-ordered set has a unique order isomorphism with an ordinal.*

**Proof.** Both (i) and the implication (a) $\Rightarrow$ (b) of (ii) follow immediately from the definition. To get the reverse of the latter implication, suppose the ordinal  $\alpha$  is a proper subset of the ordinal  $\beta$ , and let us show that it is a member of  $\beta$ . We observed above that  $\beta$  is well-ordered under “ $\in$  or  $=$ ”, and as  $\alpha$  is closed under  $\in$ , it will form an initial segment of  $\beta$ . Let  $\gamma \in \beta$  be the least element not belonging to this initial segment. By definition of the ordering of  $\beta$ , the members of  $\gamma$  are the elements smaller than it. But these are the elements of  $\alpha$ , so  $\alpha = \gamma$ , proving (a). Note that we have also proved (iii).

To show the first assertion of (iv), we shall show that given two ordinals  $\alpha$  and  $\beta$ , their intersection  $\gamma = \alpha \cap \beta$  will coincide with one of them. If it did not, it would be a proper initial segment of each, and by (iii), it would then be a member of each, namely the first element not belonging to that initial segment. But this would make  $\gamma$  a member of each of  $\alpha$  and  $\beta$  but not of  $\alpha \cap \beta$ , an absurdity. To get the final assertion of the first sentence of (iv), let  $C$  be a nonempty class of ordinals, take any  $\beta \in C$ , note that  $C' = \{\alpha \mid \alpha \subseteq \beta, \alpha \in C\}$  is a set of ordinals, and apply Regularity to this set. The last sentence of (iv) is immediate.

From the first assertion of (iv) we easily see that the union of a set of ordinals will satisfy the definition of an ordinal, i.e., (v) holds. Note, moreover, that if we call this union  $\alpha$ , then  $\alpha$  will be  $\geq$  all members of  $S$ , hence its successor,  $\alpha \cup \{\alpha\}$ , will not be a member of  $S$ . So for any set  $S$  of ordinals, we have constructed an ordinal not in  $S$ , proving (vi).

To show (vii), let  $S$  be a well-ordered set. For convenience, let us form a new ordered set  $T$  consisting of the elements of  $S$ , ordered as in  $S$ , and one additional element  $z$ , greater than them all. It is immediate that  $T$  will again be well-ordered. I claim that for every  $t \in T$ , there is a unique order-isomorphism between  $\{s \mid s < t\}$  and some (unique) ordinal. Indeed, if not, there would be a least  $t$  for which this failed, and it is easy to check that the set of ordinals associated with all the elements  $s < t$  would then be order-isomorphic to  $\{s \mid s < t\}$  and would be the unique ordinal with this property, again by a unique isomorphism, contradicting our assumption. In particular, there is a unique order-isomorphism between  $\{s \mid s < z\} = S$  and an ordinal, as required.  $\square$

**Exercise 4.5.3.** State the version of our definition of ordinal that one would use if one did not assume the Regularity Axiom, and show how each use of that axiom in the proof of Proposition 4.5.3 could be eliminated if that definition were assumed.

**Exercise 4.5:4.** Let  $\alpha$  and  $\beta$  be ordinals. Show that if there exists a one-to-one isotone map  $f: \alpha \rightarrow \beta$ , then  $\alpha \leq \beta$ .

**Exercise 4.5:5.** If  $P$  is a partially ordered set with DCC, let the *height*  $\text{ht}(p)$  of an element  $p \in P$  be defined, recursively, as the least ordinal greater than the height of every element  $q < p$ , and define the height of  $P$  to be the least ordinal greater than the height of every element of  $P$ .

(i) Show that the height function is the *least* strict isotone ordinal-valued function on  $P$ , and that it has range precisely  $\text{ht}(P)$ .

(ii) Show that for every ordinal  $\alpha$  there exists a partially ordered set containing no infinite chains, and having height  $\alpha$ .

(iii) Suppose we define the *chain height* of  $P$ ,  $\text{chht}(P)$ , to be the least ordinal which cannot be embedded in  $P$  by an isotone map, and  $\text{chht}(p)$  for  $p \in P$  as  $\text{chht}(\{q \in P \mid q < p\})$ . What can you establish about the relation between the functions  $\text{ht}$  and  $\text{chht}$ ?

Since one considers ordinals to be ordered under the relation  $\subseteq$ , equivalently, “ $\in$  or  $=$ ”, one has the choice, in speaking about them, between writing  $\leq$  and  $\subseteq$ , and likewise between  $<$  and  $\in$ . Both the order-theoretic and the set-theoretic notation are used, sometimes mixed together.

For every ordinal  $\alpha$ , there is a least ordinal greater than  $\alpha$ , namely  $\alpha \cup \{\alpha\}$ . This is called the *successor* of  $\alpha$ , and written  $\alpha+1$ . “Most” ordinals are successor ordinals. Those, such as  $0$ ,  $\omega$ ,  $\omega^2$ , etc., which are not, are called *limit ordinals*. (Although, as I have just said,  $0$  is logically a limit ordinal, and I will consider it such here, it is sometimes treated as a special case, neither a successor nor a limit ordinal.)

**Exercise 4.5:6.** Show that an ordinal is a limit ordinal if and only if it is the least upper bound of all strictly smaller ordinals; equivalently, if and only if, as a set, it is the union of all its members.

Now that we understand why ordinals in general, and natural numbers in particular, are defined so that each ordinal is equal to the set of smaller ordinals, let us rescind the convention we set up in §1.3, where, for the sake of familiarity, we said that an  $n$ -tuple of elements of a set  $S$  would mean a function  $\{1, \dots, n\} \rightarrow S$ :

**Definition 4.5.4.** Throughout the remainder of these notes,  $n$ -tuples will be defined in the same way as  $I$ -tuples for other sets  $I$ . That is, for  $n$  a natural number, an  $n$ -tuple of elements of a set  $S$  will mean a function  $n \rightarrow S$ , i.e., a family  $(s_0, s_1, \dots, s_{n-1})$  ( $s_i \in S$ ). The set of all such functions will be denoted  $S^n$ .

We have referred to ordinals denoted by symbols such as  $\omega^2$ ,  $\omega^2+1$ , etc.. As this suggests, there is an arithmetic of ordinals. If  $\alpha$  and  $\beta$  are ordinals,  $\alpha+\beta$  represents the ordinal which has an initial segment  $\alpha$ , and the remaining elements of which form a subset order-isomorphic to  $\beta$ . This exists, since by putting an order-isomorphic copy of  $\beta$  “above” the ordinal  $\alpha$ , one gets a well-ordered set, and we know that there is a unique ordinal order-isomorphic to it. Similarly,  $\alpha\beta$  represents an ordinal which is composed of a family of disjoint well-ordered sets, each order-isomorphic to  $\alpha$ , one above the other, with the order structure of the set of copies being that of  $\beta$ . These operations are (of course) formally defined by recursion, as we will describe below.

Unfortunately, the formalization of recursion that we proved in Lemma 4.3.4 is not quite strong enough for the present purposes, because in constructing larger ordinals from smaller ones, we will not easily be able to give *in advance* a codomain set corresponding to the  $T$  of that lemma, and as a result, we will not be able to precisely specify the function  $r$  required by that lemma either.

However, there is a version of recursion based on the Replacement Axiom (Fraenkel's contribution to Zermelo-Fraenkel set theory) which gets around this problem. Like that axiom, it assumes we are given a construction which is not necessarily a function, because its range and domain are not assumed to be sets, but which nonetheless uniquely determines one element given another. I will not discuss this concept, but will state the result below. The proof is exactly like that of Lemma 4.3.4, except that the Axiom of Replacement is used to carry out the "piecing together" of partial functions.

**Lemma 4.5.5** (Cf. [21, Theorem 7.1.5, p.74]). *Let  $X$  be a partially ordered set with descending chain condition, and  $r$  a construction associating to every pair  $(x, f_{<x})$ , where  $x \in X$  and  $f_{<x}$  is a function with domain  $\{y \in X \mid y < x\}$ , a uniquely defined set  $r(x, f_{<x})$ . Then there exists a unique function  $f$  with domain  $X$  such that for all  $x \in X$ ,  $f(x) = r(x, f \upharpoonright \{y \mid y < x\})$ .  $\square$*

We can now define the operations of ordinal arithmetic. For completeness we start with the (nonrecursive) definition of the successor operation. Note that in each of the remaining (recursive) definitions, the ordinal  $\alpha$  is taken as "constant", and the ordinal over which we are doing the recursion is written  $\beta$  or  $\beta+1$ .

*Definition of the successor of an ordinal:*

$$(4.5.6) \quad \beta+1 = \beta \cup \{\beta\}.$$

*Definition of addition of ordinals:*

$$(4.5.7) \quad \alpha+0 = \alpha, \quad \alpha+(\beta+1) = (\alpha+\beta)+1, \quad \alpha+\beta = \cup_{\gamma<\beta} \alpha+\gamma \text{ for } \beta \text{ a limit ordinal } > 0.$$

*Definition of multiplication of ordinals:*

$$(4.5.8) \quad \alpha 0 = 0, \quad \alpha(\beta+1) = (\alpha\beta)+\alpha, \quad \alpha\beta = \cup_{\gamma<\beta} \alpha\gamma \text{ for } \beta \text{ a limit ordinal } > 0.$$

*Definition of exponentiation of ordinals:*

$$(4.5.9) \quad \alpha^0 = 1, \quad \alpha^{(\beta+1)} = (\alpha^\beta)\alpha, \quad \alpha^\beta = \cup_{\gamma<\beta} \alpha^\gamma \text{ for } \beta \text{ a limit ordinal } > 0.$$

**Exercise 4.5:7.** Definitions (4.5.7) and (4.5.8) do not look like the descriptions of ordinal addition and multiplication sketched informally above. Show that they do in fact have the properties indicated there.

Although the operations defined above agree with the familiar ones on the finite ordinals (natural numbers), they have unexpected properties on infinite ordinals. Neither addition nor multiplication is commutative:

$$\begin{aligned} 1+\omega &= \omega, & \text{but} & & \omega+1 &> \omega, \\ 2\omega &= \omega, & \text{but} & & \omega 2 &> \omega. \end{aligned}$$

Exponentiation is also different from exponentiation of cardinals (discussed later in this section):

$$2^\omega = \omega.$$

Students who have not seen ordinal arithmetic before might do:

**Exercise 4.5:8.** Prove the three equalities and two inequalities asserted above.

You may assume familiar facts about arithmetic of natural numbers, and that the ordinal operations agree with the familiar operations in these cases; but assume nothing about how they behave on infinite ordinals, except the definitions.

The formulas (4.5.7)-(4.5.9) define *pairwise* arithmetic operations. We can also define arithmetic operations on families of ordinals indexed by (what else?) ordinals. Let us record the case of addition, since we will need this later. Given  $(\alpha_\gamma)_{\gamma \in \beta}$ , the idea is to define  $\Sigma_{\gamma \in \beta} \alpha_\gamma$  to be the ordinal which, as a well-ordered set, is the union of a chain of disjoint subsets of respective order types  $\alpha_\gamma$  ( $\gamma \in \beta$ ), appearing in that order.

*Definition of infinite ordinal addition:*

$$(4.5.10) \quad \begin{aligned} \Sigma_{\gamma \in 0} \alpha_\gamma &= 0, & \Sigma_{\gamma \in \beta+1} \alpha_\gamma &= (\Sigma_{\gamma \in \beta} \alpha_\gamma) + \alpha_\beta, \\ \Sigma_{\gamma \in \beta} \alpha_\gamma &= \bigcup_{\gamma < \beta} \Sigma_{\delta \in \gamma} \alpha_\delta && \text{for } \beta \text{ a limit ordinal } > 0. \end{aligned}$$

Taking the  $\alpha_\gamma$ 's all equal, we see that our recursive definition of  $\Sigma_\beta \alpha_\gamma$  reduces to our definition of multiplication of ordinals; thus

$$(4.5.11) \quad \Sigma_{\gamma \in \beta} \alpha = \alpha\beta.$$

**Exercise 4.5:9.** (i) Given an ordinal-indexed family of ordinals,  $(\alpha_\gamma)_{\gamma \in \beta}$ , let  $\alpha$  denote the ordinal  $\bigcup_{\gamma \in \beta} \alpha_\gamma$  (the supremum of the  $\alpha_\gamma$ 's). Let  $P$  be the set  $\beta \times \alpha$ , lexicographically ordered. Show that the ordinal  $\Sigma_{\gamma \in \beta} \alpha_\gamma$  is isomorphic as a well-ordered set to  $\{(\gamma, \delta) \mid \gamma \in \beta, \delta \in \alpha_\gamma\} \subseteq P$ .

(ii) Deduce from this a description of a well-ordered set isomorphic to the ordinal product  $\alpha\beta$  of two arbitrary ordinals.

This description clearly extends inductively to finite products  $\prod_{\gamma \in \beta} \alpha_\gamma$  ( $\beta < \omega$ ), leading, incidentally, to an easy proof of *associativity* of multiplication of ordinals. The extension of these ideas to infinite products will be developed in a later exercise in this section.

We have seen that every well-ordered set is indexed in a canonical way by an ordinal; but we do not yet know whether we can well-order every set. It turns out that we can do so if we assume the Axiom of Choice. This is stated in the second point of the next lemma; the first point gives a key argument (not requiring the Axiom of Choice) used in the proof.

**Lemma 4.5.12.** *Let  $X$  be a set. Then*

(i) *There exists an ordinal  $\alpha$  which cannot be put in bijective correspondence with any subset of  $X$ ; equivalently, such that for any well-ordering “ $\leq$ ” of any subset of  $Y \subseteq X$ ,  $(Y, \leq)$  is isomorphic to a proper initial segment of  $\alpha$ .*

(ii) *Assuming the Axiom of Choice,  $X$  itself can be well-ordered.*

**Proof.** The class of well-orderings of subsets of  $X$  is easily shown to be a set, hence by the Replacement Axiom, the unique ordinals isomorphic to these various well-ordered sets form a set, hence the union of this set is an ordinal  $\beta$ . Take  $\alpha = \beta+1$ . By construction, any well-ordering of a subset  $Y \subseteq X$  induces a bijection of  $Y$  with an initial segment of  $\beta$ , which is a proper initial segment of  $\alpha$ , yielding the second formulation of (i). To get the first formulation, note that if  $\alpha$  could be put in bijective correspondence with a subset of  $X$ , then the ordering of  $\alpha$  would induce a well-ordering of that subset, such that  $\alpha$  was the unique ordinal isomorphic to that well-ordered set, giving a contradiction to our preceding conclusion.

Assuming the Axiom of Choice, let us now take a function  $c$  which associates to every nonempty subset  $Y \subseteq X$  an element  $c(y) \in Y$ . Let us recursively construct a one-to-one map from some initial subset of the ordinal  $\alpha$  of part (i) into  $X$  as follows: Suppose we have gotten a function  $f_\beta$  from an ordinal  $\beta < \alpha$ , regarded as a subset of  $\alpha$ , into  $X$ . If its image is  $X$ , we are done. If not, we send the element  $\beta$ , which is the first element of  $\alpha$  on which our map is not yet defined, to  $c(X - \text{image}(f_\beta))$ . It is easy to verify by induction that each map  $f_\beta$  is one-to-one. If this process went on to give a one-to-one map  $f_\alpha$  of  $\alpha$  into  $X$ , that would contradict (i). So instead, the construction must terminate at some step, which means we must get a bijection between an initial segment of  $\alpha$  and  $X$ , and hence a well-ordering of  $X$ , proving (ii). (As in the proof of Lemma 4.5.1, our use of a recursion that terminates before we get through all of  $\alpha$  can be formalized by adjoining to  $X$  an element DONE.)  $\square$

**Exercise 4.5:10.** (Boris Bukh) Let  $P$  be a partially ordered set, and  $\text{Ch}(P)$  the set of chains in  $P$ , partially ordered by writing  $A \leq B$  if  $A$  is an initial segment of  $B$ .

- (i) Show that there can be no strictly isotone map  $f: \text{Ch}(P) \rightarrow P$ . (Suggestion: If there were, show that one could recursively embed any ordinal  $\alpha$  in  $P$ , by sending each element of  $\alpha$  to the image under  $f$  of the chain of images of all preceding elements.)
- (ii) Deduce from (i) the same statement with  $\text{Ch}(P)$  ordered by inclusion. Conclude that  $\text{Ch}(P)$ , under either ordering, can never be order-isomorphic to  $P$ .
- (iii) Can you strengthen the result of (i) by replacing  $\text{Ch}(P)$  by some natural proper subset thereof?

Let us assume the Axiom of Choice for the rest of this section (though at the beginning of the next section, we will again suspend this assumption).

Recall that two sets are said to have the same *cardinality* if they can be put in bijective correspondence. We have shown that (assuming the Axiom of Choice), every set has the same cardinality as an ordinal. This means we can use appropriately chosen ordinals as “standard examples” of all cardinalities. In general, there are more than one ordinals of a given cardinality (e.g.,  $\omega$ ,  $\omega+1$ ,  $\omega^2$  and  $\omega^2$  are all countable), so the ordinal to use is not uniquely determined. The one easily specified choice is the *least* ordinal of the given cardinality; so one makes

**Definition 4.5.13.** A cardinal is an ordinal which cannot be put into bijective correspondence with a proper initial segment of itself.

For any set  $X$ , the least ordinal with which  $X$  can be put in bijective correspondence will be denoted  $\text{card}(X)$ . Thus, this is a cardinal, and is the only cardinal with which  $X$  can be put in bijective correspondence.

There is an arithmetic of cardinals: If  $\kappa$  and  $\lambda$  are cardinals,  $\kappa + \lambda$  is defined as the cardinality of the union of any two disjoint sets one of which has cardinality  $\kappa$  and the other cardinality  $\lambda$ ,  $\kappa \lambda$  as the cardinality of the direct product of a set of cardinality  $\kappa$  and a set of cardinality  $\lambda$ , and  $\kappa^\lambda$  as the cardinality of the set of all functions from a set of cardinality  $\lambda$  to a set of cardinality  $\kappa$ . Unfortunately, if we consider the class of cardinals as a subset of the ordinals, these are different operations from the *ordinal arithmetic* we have just defined! To compare these arithmetics, let us temporarily use the notations  $\alpha +_{\text{ord}} \beta$ ,  $\alpha \cdot_{\text{ord}} \beta$  and  $\alpha^{\text{ord} \beta}$  for ordinal operations, and  $\kappa +_{\text{card}} \lambda$ ,  $\kappa \cdot_{\text{card}} \lambda$  and  $\kappa^{\text{card} \lambda}$  for cardinal operations. A positive statement we can make is that for cardinals  $\kappa$  and  $\lambda$ , the computation of their cardinal sum and product can be reduced to that of their ordinal sum and product, by the formulas

$$(4.5.14) \quad \kappa +_{\text{card}} \lambda = \text{card}(\kappa +_{\text{ord}} \lambda) \quad \text{and} \quad \kappa \cdot_{\text{card}} \lambda = \text{card}(\kappa \cdot_{\text{ord}} \lambda).$$

These are cases of a formula holding for any family of ordinals  $(\alpha_\gamma)_{\gamma \in \beta}$ :

$$(4.5.15) \quad \sum_{\gamma \in \beta}^{\text{card}} \text{card}(\alpha_\gamma) = \text{card}(\sum_{\gamma \in \beta}^{\text{ord}} \alpha_\gamma).$$

On the other hand, the cardinality of an infinite ordinal product of ordinals is not in general equal to the cardinal product of the cardinalities of these ordinals; in particular, cardinal exponentiation does not in any sense agree with ordinal exponentiation:  $2^{\text{card} \omega}$  gives the cardinality of the continuum, which is uncountable, while  $2^{\text{ord} \omega} = \omega$ . There is no standard notation for distinguishing ordinal and cardinal arithmetic; authors either introduce ad hoc notations, or say in words whether cardinal or ordinal arithmetic is meant, or rely on context to show this.

**Exercise 4.5:11.** In this exercise we shall extend the results of Exercise 4.5:9, which characterized the order-types of general sums and finite products of ordinals, to general products. (I have put this off until now so that we would have notation distinguishing the ordinal product  $\prod^{\text{ord}} \alpha_\gamma$  from the set-theoretic product.) We will also note at the end a relation with cardinal arithmetic. We need to begin with a generalization of lexicographic ordering.

Suppose  $(X_i)_{i \in I}$  is a family of partially ordered sets, indexed by a totally ordered set  $I$ ; and let each  $X_i$  have a distinguished element, denoted  $0$  (or  $0_i$  if there is danger of ambiguity). Define the *support* of  $(x_i) \in \prod_I X_i$  as  $\{i \in I \mid x_i \neq 0\}$ , and let  $\prod_I^{\text{w.o.s.}} X_i$  denote the set of elements of  $\prod_I X_i$  having *well-ordered* support. Similarly, let  $\prod_I^{\text{f.s.}} X_i$  denote the set of elements of *finite* support.

(i) Show that lexicographic order, which in Definition 4.3.7 was defined on  $\prod_I X_i$  only for  $I$  well-ordered, may be defined on  $\prod_I^{\text{w.o.s.}} X_i$  for arbitrary totally ordered  $I$ , and that the resulting ordering is total if each  $X_i$  is totally ordered.

(ii) Show that if  $I$  is reverse-well-ordered (has ascending chain condition) then  $\prod_I^{\text{w.o.s.}} X_i = \prod_I^{\text{f.s.}} X_i$ .

(iii) Show that if  $I$  is reverse-well-ordered and if each  $X_i$  has descending chain condition, and has  $0$  as *least* element, then  $\prod_I^{\text{f.s.}} X_i$  has descending chain condition under lexicographic ordering.

(iv) Let us now be given an ordinal-indexed family of ordinals,  $(\alpha_\gamma)_{\gamma \in \beta}$ . Write down the definition of  $\prod_{\gamma \in \beta}^{\text{ord}} \alpha_\gamma$  analogous to (4.5.10). Verify that if any  $\alpha_\gamma$  is  $0$ , your definition gives the ordinal  $0$ . In the contrary case, show that your definition gives an ordinal order-isomorphic to  $\prod_{\gamma \in \beta^{\text{op}}}^{\text{f.s.}} \alpha_\gamma$ . (Here  $\beta^{\text{op}}$  denotes the set  $\beta$ , but with its ordering – used in defining lexicographic order on our product – reversed. Note that “ $\gamma \in \beta^{\text{op}}$ ,” means the same as “ $\gamma \in \beta$ .” For the elements  $0$  in the definition of  $\prod^{\text{f.s.}}$ , we take the ordinal  $0 \in \alpha_\gamma$ , which is why we need to assume all  $\alpha_\gamma$  nonzero.)

(v) Deduce a description of the order-type of  $\alpha^{\text{ord} \beta}$ , and conclude that  $\text{card}(\alpha^{\text{ord} \beta}) \leq \alpha^{\text{card} \beta}$ .

You might also want to do

(vi) Show by examples that (ii) above fails if any of the three hypotheses is deleted.

The concept of cardinality historically antedates the construction of the ordinals, so there is a system of names for cardinals independent of their names as ordinals. The finite cardinals are, of course, denoted by the traditional symbols  $0, 1, 2, \dots$ . The least infinite cardinal is denoted  $\aleph_0$ , the next  $\aleph_1$ , etc.. From our description of the cardinals as a subclass of the ordinals, we see that the class of cardinals is “well-ordered” (written in quotes, as we did for the class of ordinals,

because this class is not a set). Hence now that one has the concept of ordinal, one continues the above set of symbols using ordinal subscripts: The  $\alpha$ th cardinal after  $\aleph_0$  is written  $\aleph_\alpha$ .

There is a further notation for cardinals “regarded as ordinals”. Each  $\aleph_\alpha$ , regarded as an ordinal, is written  $\omega_\alpha$ . Thus one writes  $\aleph_0 = \omega_0 = \omega$ ,  $\aleph_1 = \omega_1$ , etc..

Let us recall, without repeating the proofs here, some well-known properties of cardinal arithmetic, though we will use them only occasionally.

**Theorem 4.5.16.** *Letting  $\kappa$ ,  $\lambda$ , etc., denote cardinals, and letting arithmetic notation denote cardinal arithmetic, the following statements are true.*

(i) For all  $\kappa$ ,  $\lambda$ ,  $\mu$ ,

$$\kappa + \lambda = \lambda + \kappa, \quad \kappa \lambda = \lambda \kappa, \quad (\kappa + \lambda)\mu = \kappa\mu + \lambda\mu, \quad \kappa^{\lambda+\mu} = \kappa^\lambda \kappa^\mu, \quad \kappa^{\lambda\mu} = (\kappa^\lambda)^\mu.$$

(ii) For sets  $X_i$  ( $i \in I$ ),  $\text{card}(\cup_I X_i) \leq \Sigma \text{card}(X_i)$ .

(iii) If  $\kappa_\beta \leq \lambda_\beta$  for all  $\beta \in \alpha$ , then

$$\Sigma_\alpha \kappa_\beta \leq \Sigma_\alpha \lambda_\beta, \quad \Pi_\alpha \kappa_\beta \leq \Pi_\alpha \lambda_\beta, \quad \text{and, if } \kappa_0 > 0, \quad \kappa_0^{\kappa_1} \leq \lambda_0^{\lambda_1}.$$

(iv) If  $\kappa \leq \lambda$  and  $\lambda$  is infinite, then  $\kappa + \lambda = \lambda$ . If also  $\kappa > 0$ , then  $\kappa \lambda = \lambda$ . In particular,  $\omega \omega = \omega$ , hence by (ii) and (iii), a countable union of countable sets is countable.

(v)  $2^\kappa > \kappa$ . Equivalently, the power set of any set  $X$  has strictly larger cardinality than  $X$ .

**Proof.** See [29, pp. 17-21], or [31, appendix 2, §1 and exercises at the end of that appendix].  $\square$

It is interesting that while the statement  $\omega \omega = \omega$  is easy to prove without the Axiom of Choice (by describing an explicit bijection), its consequence, “a countable union of countable sets is countable”, requires that axiom, to enable us to choose bijections between the set  $\omega$  and each of the infinitely many given countable sets.

Turning from arithmetic back to order properties, let me define a concept of interest in the general study of ordered sets, and note a specific application to cardinals.

**Definition 4.5.17.** *If  $X$  is a partially ordered set, then the cofinality of  $X$  means the least cardinality of a cofinal subset  $Y \subseteq X$  (Definition 4.1.6).*

*A cardinal  $\kappa$  is called regular if, as an ordinal, it has cofinality  $\kappa$ . A cardinal that is not regular is called singular.*

**Exercise 4.5:12.** Show that if a partially ordered set  $X$  has cofinality  $\kappa$ , then every cofinal subset  $Y \subseteq X$  also has cofinality  $\kappa$ .

**Exercise 4.5:13.** Prove:

- (i) Every cardinal of the form  $\aleph_{\alpha+1}$  (i.e., every cardinal indexed by a successor ordinal) is regular.
- (ii) The first infinite singular cardinal is  $\aleph_\omega$ .

The next exercise examines the class of regular cardinals within the class of ordinals.

**Exercise 4.5:14.** Let us call an ordinal  $\alpha$  regular if there is no set map from an ordinal  $< \alpha$  onto a cofinal subset of  $\alpha$ .

- (i) Show that regular ordinals are “sparse”, by verifying that the only regular ordinals are 0, 1, and the regular infinite cardinals.
- (ii) On the other hand, point (i) of the preceding exercise shows that within the set of infinite

cardinals, the singular cardinals are sparse: They must be limit cardinals, i.e., cardinals  $\omega_\alpha$  such that  $\alpha$  is a limit ordinal. Prove this if you did not do that exercise.

(iii) Show that among the *limit* cardinals, *regular* cardinals are again sparse, by showing that if  $\omega_\alpha$  is regular and  $\alpha$  is a limit ordinal, then  $\alpha$  must be a cardinal; in fact, 0 or a cardinal  $\kappa$  satisfying

$$\kappa = \omega_\kappa.$$

Show that the first cardinal  $\kappa$  satisfying that equation is the supremum of the chain  $\kappa(i)$  ( $i \in \omega$ ) defined by  $\kappa(0) = 0$ ,  $\kappa(i+1) = \omega_{\kappa(i)}$ , but that this cardinal is still *not* regular.

(Regular limit cardinals will come up again in §6.4.)

**Exercise 4.5:15.** Since ordinals are totally ordered sets, they are in particular partially ordered sets, and we can partially order the set-theoretic direct product of two ordinals by componentwise comparison (Definition 4.1.4). Regarding the product-sets  $\omega \times \omega$ ,  $\omega \times \omega_1$ ,  $\omega_1 \times \omega_1$ ,  $\omega \times \omega_\omega$ , and  $\omega_1 \times \omega_{\omega+1}$  as partially ordered in this manner, determine, as far as you can, whether each of these contains a *cofinal chain*. (Partial credit will be given for partial results, and additional credit for general results subsuming some of these particular cases.)

The properties of ordinals allow us to obtain a construction that we wondered about when we considered Stone-Čech compactifications in §3.17:

**Exercise 4.5:16.** Let  $S$  be a totally ordered set, and for convenience, let  $-\infty$  and  $+\infty$  be two elements outside  $S$ , the former regarded as less than all elements of  $S$  and the latter as greater than all elements of  $S$ . Then we can define the *order topology* on  $S$  to have as basis of open sets the intervals  $(r, t) = \{s \in S \mid r < s < t\}$ , where  $r, t \in S \cup \{-\infty, +\infty\}$ . (We could do without  $-\infty$  and  $+\infty$  if we knew that  $S$  had no least or greatest element. But if it does, then unless one introduces these elements, one has to define several distinct sorts of basic open sets, instead of just one.)

(i) Let an ordinal  $\alpha$  be given the order topology. Which subsets of  $\alpha$  are closed? Which are compact?

(ii) Show that under the order topology the ordinal  $\omega_1$  is not compact, but satisfies condition (b) of Exercise 3.17:7. (Thus, it satisfies condition (a) of that exercise, which was what we were interested in there.)

(iii) If you are familiar with the geometric construction of the *long line*, show that this also satisfies condition (b) of Exercise 3.17:7, and examine its relationship to the ordinal  $\omega_1$ .

**4.6. Zorn's Lemma.** Ordinals, together with the Axiom of Choice, give a powerful tool for constructing non-uniquely-determined objects in most areas of mathematics. Consider the following approach to such constructions:

As steps toward constructing a certain kind of object, one specifies what kind of objects one will consider “partial constructions”. One verifies that these form a set; hence there exists an ordinal  $\alpha$  of greater cardinality than that of this set. One then wants to recursively map an initial segment of  $\alpha$  into the set of partial constructions. Setting up this recursion involves three tasks:

(i) Getting an “initializing” partial construction to which to map 0.

(ii) Specifying what to do at a successor ordinal: If one has built up one’s partial construction through the stage indexed by  $\alpha$ , and it is still not “finished”, one shows that it can be extended further, to give an  $\alpha+1$ ’st stage. The Axiom of Choice lets one choose, for each “unfinished” construction, an extension to use.

(iii) Specifying what to do at a nonzero limit ordinal  $\alpha$ . In this case, one has a chain of preceding partial constructions each extending the one before, and it is usually easy to verify that

their “union” (in the appropriate sense) is a partial construction extending all of them.

Now note that if the resulting recursion did not lead to a “finished” construction at some step, one would get a one-to-one map from  $\alpha$  into the set of partial constructions, contradicting the choice of  $\alpha$ . Hence a finished construction must be obtained at some stage, as desired!

Example: To show that an arbitrary vector space  $V$  has a basis  $B$ , one considers as “partial constructions” all linearly independent subsets of  $V$ . One can begin with the linearly independent subset  $B_0 = \emptyset$ . If the subset  $B_\alpha$  one has obtained at a given stage does not span  $V$ , there will be some element  $v \in V$  outside the span of  $B_\alpha$ , and we can take  $B_{\alpha+1}$  to be  $B_\alpha \cup \{v\}$ , and verify that this is linearly independent. If  $\alpha$  is a limit ordinal, we let  $B_\alpha$  be the set-theoretic union of the chain of subsets  $B_\beta$  ( $\beta \in \alpha$ ), and verify linear independence for this set. The preceding argument then shows that we will eventually get a linearly independent subset which cannot be extended, i.e., which spans  $V$ , as desired.

In view of the usefulness of this technique, it is natural to seek a lemma whose proof will do the predictable part once and for all, and show us what must be proved separately for each case. In formulating this lemma, let us render the set of all “partial constructions” by a partially ordered set  $(X, \leq)$ , where  $\leq$  is thought of as the relation of one construction being a “part of” another. The condition we need to initialize our recursion ((i) above) is that  $X$  be nonempty. To say that we can extend a partial construction further if it is not yet “finished” ((ii) above) is, put in the contrapositive, to say that if  $X$  has any *maximal* element, this is an object of the sort we desire. Finally, the condition we need to be able to continue at steps indexed by limit ordinals ((iii) above), namely, that given a chain of partial constructions, we can pass to one which includes them all, is made the content of a definition:

**Definition 4.6.1.** *A partially ordered set  $X$  is called inductive if for every nonempty chain  $Y \subseteq X$ , there is an element  $z \in X$  majorizing  $Y$  (i.e.,  $\geq$  all elements of  $Y$ ).*

We can now state the desired result, *Zorn's Lemma*, and show that it and a number of other statements are equivalent to the Axiom of Choice.

**Theorem 4.6.2.** *Assuming the axioms of Zermelo-Fraenkel set theory (but not the Axiom of Choice), the following four statements are equivalent:*

- (i) **The Axiom of Choice:** *If  $X$  is a set, and  $f$  is a function associating to every  $x \in X$  a nonempty set  $f(x)$ , then there exists a function  $g$  associating to every  $x \in X$  an element  $g(x) \in f(x)$ . (Equivalently: the direct product of any family of nonempty sets is nonempty.)*
- (ii) **Zorn's Lemma:** *Every nonempty inductive partially ordered set  $(X, \geq)$  has a maximal element.*
- (iii) **The Well-ordering Principle:** *Every set can be well-ordered. (Equivalently: every set can be put in bijective correspondence with an ordinal.)*
- (iv) **Comparability of Cardinalities:** *Given any two sets  $X$  and  $Y$ , one of the sets can be put in bijective correspondence with a subset of the other. (Loosely: the class of cardinalities is totally ordered.)*

**Proof.** The scheme of proof will be  $(iv) \Leftrightarrow (iii) \Leftrightarrow (i) \Leftrightarrow (ii)$ . That the parenthetical restatement of (iii) is equivalent to the main statement follows from Proposition 4.5.3(vii).

$(iv) \Leftrightarrow (iii)$ : Assuming (iv), let  $X$  be any set and  $\alpha$  an ordinal with the property stated in Lemma 4.5.12(i). By (iv), there is either a bijection between  $X$  and a subset of  $\alpha$ , or vice versa.

By choice of  $\alpha$ , the latter case cannot occur, so there is a bijection between  $X$  and a subset  $S \subseteq \alpha$ . Since  $\alpha$  is well-ordered, so is every subset, and the well-ordering of  $S$  induces a well-ordering of  $X$ , proving (iii). Assuming (iii), statement (iv) follows from the comparability of ordinals, Proposition 4.5.3(iv), or more directly, from Lemma 4.5.1.

(iii) $\Leftrightarrow$ (i): We proved (i) $\Rightarrow$ (iii) in Lemma 4.5.12(ii). Conversely, assume (iii). Given  $X$  and  $f$  as in (i), statement (iii) tells us that we can find a well-ordering  $\leq$  on the set  $\cup_{x \in X} f(x)$ . We now define  $g$  to take each  $x$  to the  $\leq$ -least element of  $f(x)$ . (In terms of the axioms, we are using the Replacement Axiom to construct  $g$  as  $\{(x, y) \mid x \in X \text{ and } y \text{ is the least element of } f(x)\}$ .)

(i) $\Rightarrow$ (ii): Let  $(X, \leq)$  be a nonempty inductive partially ordered set, and let us choose as in Lemma 4.5.12(i) an ordinal  $\alpha$  which cannot be put in bijective correspondence with any subset of  $X$ . Note that the combination of conditions “inductive” and “nonempty” is equivalent to saying that for every chain  $C \subseteq X$ , including the empty chain, there is an element  $\geq$  all members of  $C$ . By (i), we may choose a function  $g$  associating to every nonempty subset of  $X$  a member of that subset. We will now recursively define an isotone map  $f: \alpha \rightarrow X$ . Assuming that for some  $\beta \in \alpha$  we have defined an isotone map  $f_{<\beta}: \beta \rightarrow X$ , observe that its image will be a chain  $C_\beta \subseteq X$ . If the set  $Y_\beta$  of elements of  $X$  greater than all members of  $C_\beta$  is nonempty, we define  $f(\beta) = g(Y_\beta)$ . In the contrary case, the hypothesis that  $X$  is inductive still tells us that there is an element  $\geq$  all members of  $C_\beta$ . We conclude that such an element must be equal to some member of  $C_\beta$ , which means that the chain has a largest element,  $c$ . In this case, we take  $f(\beta) = c$ . Note that in this case  $c$  must be maximal in  $X$ , for if not, any element of  $X$  greater than it would be greater than all elements of  $C_\beta$ , contradicting our assumption that  $Y_\beta$  was empty.

By choice of  $\alpha$ , the map  $f$  we have constructed cannot be one-to-one, but by the nature of our construction, the only situation in which one-one-ness can fail is if at some point we get a maximal element of  $X$ . Thus  $X$  has a maximal element, as claimed.

(ii) $\Rightarrow$ (i): This will be a typical application of Zorn’s Lemma. Let  $X$  and  $f$  be given as in (i). Let  $P$  be the set of all maps defined on subsets  $Y \subseteq X$  and carrying each  $x \in Y$  to an element of  $f(x)$ . Partially order  $P$  by setting  $g_1 \geq g_0$  if  $g_1$  is an *extension* of the map  $g_0$ .  $P$  is nonempty because it contains the empty mapping; it is easy to see that given any chain  $C$  of elements of  $P$  under the indicated partial ordering, the union of  $C$  will be an element of  $P$  that is  $\geq$  all elements of  $C$ , hence  $P$  is inductive. Thus it has a maximal element  $g$ . This maximal element must be a function defined on all of  $X$  (otherwise we could extend it further), completing the proof of (i).  $\square$

**Convention 4.6.3.** *Throughout the remainder of these notes, we shall assume the Axiom of Choice along with the other axioms of ZFC, and thus freely use any of the equivalent statements of the preceding theorem.*

Of these equivalent statements, Zorn’s Lemma is usually the most convenient.

Note that in the last paragraph of the above proof, our verification that  $P$  was nonempty was by the same method used to show that every nonempty chain had an upper bound: To show the latter, we used the union of the chain, while to get an element of  $P$  we took the empty function, which is the union of the empty chain. It is my experience that in *most* proofs using Zorn’s Lemma, the verification of nonemptiness may be achieved by the same construction that shows every *nonempty* chain has an upper bound; i.e., the assumption “nonempty” is rarely needed in the

latter verification. Hence my personal preference would be to use a definition of “inductive” that required *every* chain to have an upper bound, and eliminate “ $X$  nonempty” as a separate hypothesis of Zorn's Lemma. (Of course, in some exceptional cases, the verification that all chains have upper bounds may have to treat empty and nonempty chains separately. But curiously, even when the same verification works for both cases, many authors seem embarrassed to use the most trivial example to show the set is nonempty, and unnecessarily give a more complicated one instead.) For conformity with common usage, I have stated Zorn's Lemma in terms of the standard definition of “inductive”. But we may, at times, skip a separate verification that our inductive set is nonempty, and instead observe that some construction gives an upper bound for any chain, empty or nonempty.

The reader who has not seen proofs by Zorn's Lemma before, and does not see how to begin the next few exercises, might look at a few such proofs in a standard graduate algebra text such as [31], and/or ask his or her instructor for some elementary examples. The steps of identifying the sort of “partial constructions” one wants to use, describing the appropriate partial ordering on that set, verifying that the set is inductive, and verifying that a maximal element corresponds to an entity of the sort one was seeking, take practice to master.

**Exercise 4.6:1.** We saw in Exercise 4.1:10 that the maximal partial orderings on a set  $X$  were the total orderings. Deduce now for arbitrary  $X$  (as we were able to deduce there for finite  $X$ ) that

- (i) Every partial ordering can be extended to a total ordering.
- (ii) Every partial ordering is an intersection of total orderings.

**Exercise 4.6:2.** (i) If  $X$  is a totally ordered set, show that  $X$  has a subset  $Y$  well-ordered under the induced ordering, and cofinal in  $X$  (Definition 4.1.6).

(ii) Show that the  $Y$  of (i) can be taken order-isomorphic to a regular cardinal (Exercise 4.5:13), and that this cardinal is unique. However show that the set  $Y$  itself is not in general unique, and that if the condition of regularity is dropped, uniqueness of the cardinal is also lost.

(iii) Suppose  $(X_i)_{i \in I}$  is a family of totally ordered sets, such that for all  $i, j \in I$  the set  $X_i \times X_j$ , under the product order, contains a cofinal subchain. Show that the set  $\prod_I X_i$  under the product order likewise has a cofinal subchain.

The final part of this exercise does not depend on the preceding parts; rather, it represents a generalization of part (i).

- (iv) Prove that every *partially* ordered set has a cofinal subset with descending chain condition.

**Exercise 4.6:3.** For a partially ordered set  $X$ , show that the following conditions are equivalent:

- (i)  $X$  has no maximal element.
- (ii)  $X$  has two disjoint cofinal subsets.
- (ii')  $X$  has an infinite family of disjoint cofinal subsets.

The next exercise is an example where the “obvious” Zorn's Lemma proof does not work. The simplest valid proof in this case is by the well-ordering principle, which is not surprising since it is a result about well-orderability. However, this can also be turned into a Zorn's Lemma proof, if one is careful.

**Exercise 4.6:4.** Let  $X$  be a set, let  $P$  be the set of partial order relations on  $X$ , partially ordered by inclusion as in Exercise 4.1:10, and let  $Q \subseteq P$  consist of those partial orderings having descending chain condition.

- (i) Show that the maximal elements of  $Q$  (under the partial ordering induced from  $P$ ) are

the well-orderings of  $X$ .

(ii) Show that  $Q$  is *not* inductive.

(iii) Prove nonetheless that every element of  $Q$  is majorized by a maximal element, and deduce that every partial ordering with DCC on a set  $X$  is an intersection of well-orderings. (Hint: Take an appropriate ordinal  $\alpha$  and construct an indexing of the elements of  $X$  by an initial segment of  $\alpha$ , in a way “consistent” with the given partial order.)

The next three exercises, though not closely related to Zorn’s Lemma, explore further the relation between partially ordered sets and their well-ordered subsets.

**Exercise 4.6:5.** Let  $S$  be an infinite set, and  $\mathbf{P}(S)$  the set of all subsets of  $S$ , partially ordered by inclusion. Show by example that  $\mathbf{P}(S)$  can contain chains of cardinality  $> \text{card}(S)$ , but prove that  $\mathbf{P}(S)$  can never contain a *well-ordered* chain of cardinality  $> \text{card}(S)$ .

**Exercise 4.6:6.** (i) Show that every infinite *totally* ordered set has either a subset order-isomorphic to  $\omega$  or a subset order-isomorphic to  $\omega^{\text{op}}$ .

(ii) Show that every infinite partially ordered set  $P$  contains either a subset order-isomorphic to  $\omega$ , a subset order-isomorphic to  $\omega^{\text{op}}$ , or an infinite antichain (Definition 4.1.6). (Suggestion: If  $P$  has no infinite antichain, obtain a finite antichain  $B \subseteq P$  maximal for the property that the set  $S$  of elements incomparable with all elements of  $B$  is infinite; then study the properties this  $S$  must have. Alternatively, do the same thing with the roles of comparable and incomparable elements reversed.)

This family of three partially ordered sets is essentially unique for the above property:

(iii) Show that a set  $F$  of infinite partially ordered sets has the property that every infinite partially ordered set contains an isomorphic copy of a member of  $F$  if and only if  $F$  contains a partially ordered set order-isomorphic to  $\omega$ , a partially ordered set order-isomorphic to  $\omega^{\text{op}}$ , and a countable antichain.

An application of the preceding exercise is

**Exercise 4.6:7.** Let  $P$  be a partially ordered set.

(i) Show that the following conditions are equivalent:

(i.a)  $P$  contains no chains order-isomorphic to  $\omega^{\text{op}}$ .

(i.b) Every infinite subset of  $P$  contains either a subset order-isomorphic to  $\omega$ , or an infinite antichain.

(i.c)  $P$  satisfies the descending chain condition.

(ii) It is clear from (i) above that conditions (ii.a)-(ii.c) below are equivalent. Show that they are also equivalent to (ii.d):

(ii.a)  $P$  contains no chains order-isomorphic to  $\omega^{\text{op}}$ , and no infinite antichains.

(ii.b) Every infinite subset of  $P$  contains a subset order-isomorphic to  $\omega$ .

(ii.c)  $P$  has descending chain condition, and contains no infinite antichains.

(ii.d) Every total ordering extending the ordering of  $P$  is a well-ordering.

A partially ordered set  $P$  with the equivalent properties of (ii) is sometimes called “partially well-ordered”.

The first part of the next exercise notes that for uncountable cardinalities, things are more complicated.

**Exercise 4.6:8.** (i) Deduce from Exercise 4.6:5 that one can have a totally ordered set  $P$  of some infinite cardinality  $\kappa$  which contains no well-ordered or reverse-well-ordered subset of cardinality  $\kappa$ .

(ii) Suppose  $P$  is as in (i), and  $\varphi$  is a bijection between  $P$  and a well-ordered set  $Q$  of

cardinality  $\kappa$ . Consider  $\{(p, \varphi(p)) \mid p \in P\}$ , under the partial ordering induced by the product ordering on  $P \times Q$ . Show that this has neither chains nor antichains of cardinality  $\kappa$  (in contrast to the result of Exercise 4.1:9 for finite partially ordered sets).

But perhaps one can repair this deficiency. (I have not thought hard about the question asked below.)

(iii) Exercise 4.1:9 was based on defining the ‘‘height’’ of a partially ordered set as the supremum of the cardinalities of its chains; but a different concept of ‘‘height’’ was introduced for partially ordered sets with descending chain condition in Exercise 4.5:5. Can this definition be extended in some way to general partially ordered sets, or otherwise modified, so as to get an analog of Exercise 4.1:9 for partially ordered sets of arbitrary cardinality? (Or can the definition of ‘‘width’’ be so modified?)

For a curious application of the well-ordering principle to the study of abelian groups, see the first section of [39].

**4.7. Some thoughts on set theory.** I have mentioned that when the Axiom of Choice and various equivalent principles were first considered, they were the subject of a heated controversy.

The Axiom of Choice is now known to be *independent* of the other axioms of set theory; i.e., it has been proved that, assuming the consistency of the Zermelo-Fraenkel axioms without Choice, both the full set of axioms including Choice, and the Zermelo-Fraenkel axioms plus the *negation* of the Axiom of Choice are consistent. And there are further statements (for instance the Continuum Hypothesis, saying that  $2^{\aleph_0} = \aleph_1$ ) which have been shown independent of Zermelo-Fraenkel set theory *with* the Axiom of Choice, and which there do not seem to be any compelling reasons either for accepting or rejecting. This creates the perplexing question of what is the ‘‘true’’ set theory.

Alongside Zermelo-Fraenkel Set Theory with and without Choice, etc., there are further contenders for the ‘‘correct’’ foundations of mathematics. The Intuitionists objected not only to the Axiom of Choice, but to the ‘‘law of the excluded middle’’, the logical principle that every meaningful statement is either true or false. They claimed (if I understand correctly) that an assertion such as Fermat’s Last Theorem (the statement that there are no nontrivial integer solutions to  $x^n + y^n = z^n$ ,  $n > 2$ , which was unproven at the time) could be said to be false if a counterexample were found, or true if an argument could be found (using forms of reasoning acceptable to them) that proved it, but that it would be neither true nor false if neither a counterexample nor a proof existed. They maintained that the application of the law of the excluded middle to statements which involve infinitely many cases, and which thus cannot be checked case by case, was a fallacious extension to infinite sets of a method correct only for finite sets; in their words, that one cannot reason in this way about an infinite set such as the set of all natural numbers, because it cannot be regarded as a ‘‘completed totality’’.

Although this viewpoint is not current, note that the distinction between sets and proper classes, which got mathematics out of the paradoxes that came from considering ‘‘the set of all sets’’, leaves us wondering whether the *class* of all sets is ‘‘a real thing’’; and indeed one current textbook on set theory refers to this in terms of the question of whether mathematicians can consider such classes as ‘‘completed totalities’’.

During a painfully protracted correspondence with someone who insisted he could show that Zermelo-Fraenkel set theory was inconsistent, and that the fault lay in accepting infinite sets, which he called ‘‘mere phantasms’’, I was forced to think out my own view of the matter, and the conclusion I came to is that all sets, finite and infinite, are ‘‘phantasms’’; that none of mathematics is ‘‘real’’, so that there is no true set theory; but that this does not invalidate the practice of

mathematics, or the usefulness of choosing a “good” set theory.

To briefly explain this line of thought, let us understand the physical world to be “real”. (If your religious or philosophical beliefs say otherwise, you can nevertheless follow the regression to come.)

Our way of perceiving the world and interacting with it leads us to partition it into “objects”. This partitioning is useful, but is not a “real thing”.

To deal intelligently with objects, we think about families of objects, and, as our thinking gets more sophisticated, families of such families. Though I do not think the families, and families of families are “real things” either, they are useful – as descriptions of the way we classify the world.

Consider in particular our system of numbers, which are themselves not “real things”, but which give a model that allows us to use one coherent arithmetic system to deal with the various things in the world that one can count. Note that in spite of this motivation in terms of things one can count, in developing the numbers we use a system that is *not* bounded by the limitations of how high a person could count in a lifetime. A system with such a limitation arbitrarily imposed would be *more* difficult to define, learn, and work with than our system, in which the behavior of arithmetic is uniform for arbitrarily large values! Moreover, our unbounded system turns out to have applications to situations that a system bounded in that way would not be able to deal with: to demographic, geographical, astronomical and other data, which we compute from observations and theoretical models of our world, though no one human being could have counted the numbers involved unit by unit.

Now in thinking about our system of numbers, we are dealing with the concept of “all the numbers in the system” – even those who refuse to call that family a “completed totality” do reason about it! – so, if possible, we want our set theory to be able to cover such concepts. Just as we found it natural to extend the system of numbers beyond the sizes of sets a real person could count, so we may extend our system of “sets” beyond finite sets. This is not as simple as with the number-concept. Some plausible approaches turned out to lead to contradictions, e.g., those that allowed one to speak of “the set of all sets”. Among the approaches that do not lead to contradictions, some are more convenient than others. I think we are justified in choosing a more convenient system to work in – one in which the “unreal objects” that we are considering are easier to understand and generalize about.

It may seem pointless to work in a set theory which is to some extent “arbitrary”, and to which we do not ascribe absolute “truth”. But observe that as long as we use a system consistent with the laws of finite arithmetic, any statements we can prove in our system about arithmetic models of aspects of the real world, and which can in principle be confirmed or disproved in each case by a finite calculation, will be correct; i.e., as applicable to the real world as those models are. This is what I see as the “justification” for including the Axiom of Choice and other convenient axioms in our set theory.

For arguments in favor of adding another axiom, the *Axiom of Projective Determinacy*, to the standard axioms of set theory, see [129].

We should also note that making a choice among set theories or systems of reasoning does not consign all others to oblivion. Logicians *do* consider which statements hold if the Axiom of Choice is assumed and which hold if its negation is assumed. (E.g., [75] shows that in a model of ZF without Choice, one can have commutative rings with properties contradicting several standard theorems of ZFC ring theory.) Even intuitionistic logic is still studied – not, nowadays, as a preferred mode of reasoning, but as a *formal system*, related to objects called Brouwerian lattices (cf. [3]) in the same way standard logic is related to Boolean algebras.