Set Theory

P.D. Welch

September 22, 2020

Contents

Page

Part I

Fundamentals

Introduction

1.1 The beginnings

The *theory of sets* can be regarded as prior to any other mathematical theory: any everyday mathematical object, whether it be a group, ring or field from algebra, or the structure of the real line, the complex numbers etc., from analysis, or other mathematical construct, can be constructed from *sets*.

The apparent simplicity of sets belies a bewildering collection of *paradoxes*, and *logical antinomies* that plagued the early theory and led many to doubt that the theory could be made coherent. Set theory as we are going to study it was called into being by one man: Georg Cantor (1845-1918).

His papers on the subject appeared between 1874 to 1897. In one sense we can even date the first real result in set theory: it was his discovery of the uncountability of the real numbers, which he noted on December 7'th 1873.

His ideas met with some resistance, some of it determined, but also with much support, and his ideas won through. Chief amongst his supporters was the great German mathematician David Hilbert (1862-1943).

This course will start with the basic primitive concept of *set*, but will also make use along the way of a more general notion of *collection* or *class* of objects. We shall use the standard notation $\in$ for the *elementhood* relation: $x \in A$ will be read as " the set x is an element of the collection A". Only sets will occur to the left of the $\in$ symbol. In the above A may be a set or a class. We shall reserve lower case letters,

$a, b, \ldots\, u, v, x, y, \ldots$ for sets, and use upper case letters for collections or classes in general - but such collections will often also be sets. In the beginning of the course we shall be somewhat vague as to what objects sets are, and even more so as to what objects classes might be; we shall merely study a growing list of principles that we feel are natural properties that a notion of set should or could have. Only later shall we say precisely to what we are referring when we talk about the "domain of all sets". The notion of "class" is not a necessary one for this development, but we shall see that the concept arises naturally with certain formal questions, and it is a useful shorthand to be able to talk about classes, although our theory (and this course) is about sets, all talk about classes is fundamentally eliminable.[1]

One such basic principle is:

Principle (or Axiom) of Extensionality (for sets): *For two sets a, b, we shall say $a = b$ iff*:

$$\forall x(x \in a \longleftrightarrow x \in b).$$

Thus what is important about a set is merely its members. Whilst the Axiom of Extensionality does not tell us exactly what sets *are*, it does give us a criterion for when two sets are equal. There is a similar principle for collections or classes in general:

Principle (or Axiom) of Extensionality (for classes): *For two classes A, B, we shall say $A = B$ iff*:

$$\forall x(x \in A \longleftrightarrow x \in B).$$

Obviously there is no difference in the criterion, but we state the Principle separately for classes too, so that we know when we can write "$A = B$" for arbitrary classes. It is conventional to express a collection within curly parentheses:

- $\{2\} = \{x | x \text{ is an even prime number}\} = \{\text{Largest integer less than } \sqrt{5}\}$
- $\{\text{Morning Star}\} = \{\text{Evening Star}\} = \{x | x \text{ is the planet Venus}\}$;
- $\{\text{Lady Gaga}\} = \{\text{Stefani Joanne Angelina Germanotta}\}$.

This illustrates two points: that the description of the object(s) in the set or class is not relevant (what philosophers would call the *intension*). It is only the *extension* of the collection, that is what ends up in the collection, however it is specified, or even if unspecified, that counts. Secondly we use the *abstraction* notation when we want to specify by a description. This was seen at the first line of the above and will be familiar to you as a way of specifying collections of objects:

An *abstraction term* is written as $\{y \mid \ldots y \ldots\}$ where $\ldots y \ldots$ is some description (often in a formal language - say the first order language from a Logic course), and is used to collect together all the objects y that satisfy the description $\ldots y \ldots$ into a *class*. We use this notation flexibly and write $\{y \in A \mid \ldots y \ldots\}$ to mean the class of objects y in A that satisfy $\ldots y \ldots$

Axiom of Pair Set *For any sets x, y there is a set $z = \{x, y\}$ with elements just x and y. We call z the (unordered) pair set of x,y.*

In the above note that if $x = y$ then we have that $\{x, y\}$=$\{x, x\} = \{x\}$. (This is because $\{x, x\}$ has the same members as $\{x\}$ and so by the Axiom of Extensionality they are literally the same thing.) The Axiom asserts the existence of such a pair object as a *set*. (We could formally have written out the pair set as an exact abstraction term by writing $\{z \mid z = x \vee z = y\}$ but this would be overly pedantic at this stage.) It is our first example of a *set existence axiom*. As is usual we say that $x \subseteq y$ if any member of x is a member of y. We say "x is a subset of y", or "x is contained in y", or "y contains x". In symbols:

[1] In short we do not need a formal theory of classes for mathematics.

$$x \subseteq y \Leftrightarrow_{\mathrm{df}} \forall z(z \in x \to z \in y)\,;$$

also: $$x \subset y \Leftrightarrow_{\mathrm{df}} x \subseteq y \wedge x \neq y.$$

Definition 1.1 *We let $\mathcal{P}(x)$ denote the class $\{y | y \subseteq x\}$.*

Implicit in this is the idea that we *can* collect together all the subsets of a given set. Is this allowed? We adopt another set existence *axiom* about sets that says we can:

Axiom of Power Set *For any set x $\mathcal{P}(x)$ is a set, the power set of x.*

Notice that a set x can have only one power set (why?) which justifies our use of a special name $\mathcal{P}(x)$ for it. Another axiom asserting that a certain set exists is:

Axiom of the Empty Set *There is a set with no members.*

Definition 1.2 *The empty set, denoted by $\varnothing$, is the unique set with no members.*

• We can define $\varnothing$ as $\{x \mid x \neq x\}$ (since every object equals itself). Again note that there cannot be two empty sets (Why? Appeal to the Ax. of Extensionality).

• For any set (or class) A we have $\varnothing \subseteq A$ (just by the logic of quantifiers).

Example 1.3 *(i) $\varnothing \subseteq \varnothing$, but $\varnothing \notin \varnothing$; $\{\varnothing\} \in \{\{\varnothing\}\}$ but $\{\varnothing\} \nsubseteq \{\{\varnothing\}\}$.*
(ii) $\mathcal{P}(\varnothing) = \{\varnothing\}$; $\mathcal{P}(\{\varnothing\}) = \{\varnothing, \{\varnothing\}\}$; $\mathcal{P}(\{a,b\}) = \{\varnothing, \{a\}, \{b\}, \{a,b\}\}$.

We are going to build out of thin air, (or rather the empty set) in essence the *whole universe of mathematical discourse.* How can we do this? We shall form a *hierarchy* of sets, starting off with the empty set, $\varnothing$, and applying the axioms generate more and more sets. In fact it only requires two operations to generate all the sets we need: the power set operation, and another operation for forming unions. The picture is thus:

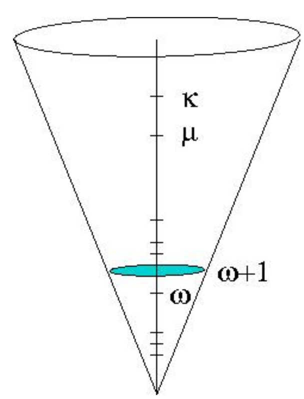

Figure 1.1: The universe V of sets

At the bottom is $V_0 =_{\mathrm{df}} \varnothing$; $V_1 =_{\mathrm{df}} \mathcal{P}(V_0) = \mathcal{P}(\varnothing)$; $V_2 =_{\mathrm{df}} \mathcal{P}(V_1)$; $V_{n+1} =_{\mathrm{df}} \mathcal{P}(V_n)\ldots$ The question arises as to what comes "next" (if there is such). Cantor developed the *theory of ordinal numbers* which extends the standard natural numbers $\mathbb{N}$. These new numbers also have an arithmetic that extends that of the usual +, $\times$ *etc.* which he developed, and which will be part of our study here. He defined a "first infinite ordinal number" which comes after all the natural numbers n and which he called ω. After ω comes $\omega + 1, \omega + 2, \ldots$. It is natural then to *accumulate* all the sets defined by the induction above, and we set $V_\omega =_{\text{df}} \{x \mid x \in V_n \text{ for some } n \in \mathbb{N}\}$. $V_{\omega+1}$ will then be defined, continuing the above, as $\mathcal{P}(V_\omega)$. However this is in the future. We first have to make sure that we have our groundwork correct, and that this is not all just fantasy.

Exercise 1.1 List all the members of V_3. Do the same for V_4. How many members will V_n have for $n \in \mathbb{N}$?

Exercise 1.2 Prove for $\alpha < 3$ that $V_{\alpha+1} = V_\alpha \cup \mathcal{P}(V_\alpha)$. (This will turn out to be true for any α.)

Exercise 1.3 We define the *rank* of a set x ('$\rho(x)$') to be the least α such that $x \subseteq V_\alpha$. Compute $\rho(\{\{\varnothing\}\})$. Do the same for $\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$.

1.2 Classes

We shall see that not all descriptions specify sets. This was a pitfall that the early workers on foundations of mathematics fell into, notably Gottlob Frege (1848-1925) The second volume of his treatise on the foundations of arithmetic (which tried to derive the laws of arithmetic from purely logical assumptions) was not far from going to press in 1903, when Bertrand Russell (1872-1970) informed him of a fundamental and, as it turned out, fatal error to his programme. Frege had, in our terms, assumed that *any* specification defined a *set* of objects. Like the Barber Paradox, Russell argued as follows.

Theorem 1.4 **(Russell)** *The collection $R = \{x | x \notin x\}$ does not define a set.*

Proof: Suppose this collection R was a set, z say. Then is $z \in R$? If so then by the description of R, $z \notin z$. However if $z \notin R$ then we should have $z \in z$! We thus have the contradiction $z \notin R \Leftrightarrow z \in R$! So there is no set z equal to $\{x | x \notin x\}$. Q.E.D.

What we have is the first example of a *class* of objects which do not form a set. When we know that a class is not, or cannot be, a set, then we call it a *proper class*. (In general we designate any collection of objects as a *class* and we reserve the term *set* for a class that we know, or posit, or define, as a set. The Russell Theorem above then proves that the Russell class R defined there is a proper class. The problem was that we were trying to define a set by looking at *every object* in the universe of sets (which we have not yet defined!). The moral of Russell's argument (which he took) is that we must restrict our ways of forming sets if we are to be free of contradictions. There followed a period of intense discussion as to how to "correctly" define sets. Once the dust eventually cleared, the following axiom scheme was seen to correctly rule out all obviously inconsistent ways of forming sets.[2] We hence adopt the following axiom scheme.

Axiom of Subsets. *Let $\Phi(x)$ be a definite, welldefined property. Let x be any set. Then*

[2] The word "obviously" is intentional: by Gödel's Second Incompleteness Theorem, we can not prove within the theory of sets that the Axiom of Subsets will always consistently yield sets. However this is a general phenomenon about formal systems, including formal number theory: such theories cannot prove their own consistency. Hence this is not a phenomenon peculiar to set theory.

$$\{y \in x \mid \Phi(y)\} \quad \textit{is a set.}$$

We call the above a *scheme* because there is one axiom for every property Φ. You might well ask what do I mean by 'a welldefined property Φ', and if we were being more formal we should specify a language in which to express such properties[3]. This axiom rules out the possibility of a "universal set" that contains all others as members.

Corollary 1.5 *Let V denote the class of all sets. Then V is a proper class.*

Proof: If V were to be a set, then we should have that $R = \{y \in V \mid y \notin y\}$ is a set by the Ax. of Subsets. However we have just shown that R is not a set. Q.E.D.

Note that the above argument makes sense, even if we have not yet been explicit as to what a set is: *whatever* we decree them to be, if we adopt the axioms already listed the above corollary holds. We want to generate more sets much as in the way mathematicians take unions and intersections. We may want to take unions of *infinite* collections of set. For example, we know how to take the union of two sets x_1 and x_2: we define $x_1 \cup x_2 =_{\mathrm{df}} \{z \mid z \in x_1 \vee z \in x_2\}$. By mathematical induction we can define $x_1 \cup x_2 \cup \cdots \cup x_k$. However we may have an infinite sequence of sets $x_1, x_2, \cdots, x_k, \ldots (k \in \mathbb{N})$ all of whose members we wish to collect together. We thus define $z = \bigcup X$, where $X = \{ x_k \mid k \in \mathbb{N}\}$, as: $\bigcup X =_{\mathrm{df}} \{t \mid \exists x \in X(t \in x)\}$.

This forms the collection we want. In fact we get a general flexible definition. Let Z be any set whatsoever. Then

Definition 1.6 $\bigcup Z =_{\mathrm{df}} \{t \mid \exists x \in Z(t \in x)\}$. *In words: for any set Z there is a class, $\bigcup Z$, which consists precisely of the members of members of Z.*

We are justified in doing this by an axiom:

Axiom of Unions: *For any set Z, $\bigcup Z$ is a set.*

This notation subsumes the more usual one as a special case: $\bigcup\{a, b\} = a \cup b$ (Check!); $\bigcup\{a, b, c, d\} = a \cup b \cup c \cup d$. Note that if $y \in x$ then $y \subseteq \bigcup x$ (but not conversely).

Example 1.7 *(i)* $\bigcup\{\{0,1,2\},\{1,2\},\{2,4,8\}\} = \{0,1,2,4,8\}$.
(ii) $\bigcup\{a\} = a$; *(iii)* $\bigcup(a \cup b) = \bigcup a \cup \bigcup b$

An extension of the above is often used:

Notation: *If I is set used to index a family of sets $\{a_j \mid j \in I\}$ we often write $\bigcup_{j\in I} A_j$ for $\bigcup\{A_j \mid j \in I\}$.*

Notice that this can be expressed as: $x \in \bigcup_{j\in I} A_j \leftrightarrow (\exists j \in I)(x \in A_j)$. We similarly define the idea of *intersection*:

Definition 1.8 *If $Z \neq \varnothing$ then $\bigcap Z =_{\mathrm{df}} \{t \mid \forall x \in Z(t \in x)\}$.*

In words: for any non-empty set Z there is another set, $\bigcap Z$, which consists precisely of the members of all members of Z. Using index sets we write

[3] It would be usual to adopt a first order language $\mathcal{L}_{\dot\in,=}$ which had $=$ plus just the single binary relation symbol $\dot\in$; then well-formed formuale of this language would be deemed to express 'well-defined properties.'

$$x \in \bigcap_{j\in I} A_j \leftrightarrow (\forall j \in I)(x \in A_j).$$

Example 1.9 $\bigcap\{\{a,b\},\{a,b,c\},\{b,c,d\}\} = \{b\}; \bigcap\{a,b,c\} = a\cap b\cap c; \bigcap\{\{a\}\} = \{a\}$.

Suppose the set Z in the above definition were empty: then we should have that for any t whatsoever that for any $x \in Z$ $t \in x$ (because there are no $x \in Z$!). However that leads us to define in this special case $\bigcap_{j\in\varnothing} A_j = V$. Note that $\bigcup_{j\in\varnothing} A_j$ makes perfect sense anyway: it is just $\varnothing$.

We have a number of basic laws that $\bigcup$ and $\bigcap$ satisfy:

(i) $I \subseteq J \to \bigcup_{i\in I} A_i \subseteq \bigcup_{j\in J} A_j$. $\quad I \subseteq J \to \bigcap_{i\in I} A_i \supseteq \bigcap_{j\in J} A_j$

(ii) $\forall i(i \in I \to A_i \subseteq C) \to \bigcup_{i\in I} A_i \subseteq C$. $\quad \forall i(i \in I \to A_i \supseteq C) \to \bigcap_{i\in I} A_i \supseteq C$.

(iii) $\bigcup_{i\in I}(A_i \cup B_i) = \bigcup_{i\in I} A_i \cup \bigcup_{i\in I} B_i$. $\quad \bigcap_{i\in I}(A_i \cap B_i) = \bigcap_{i\in I} A_i \cap \bigcap_{i\in I} B_i$.

(iv) $\bigcup_{i\in I}(A \cap B_i) = A \cap (\bigcup_{i\in I} B_i)$. $\quad \bigcap_{i\in I}(A \cup B_i) = A \cup (\bigcap_{i\in I} B_i)$.

(v) $D\backslash \bigcup_{i\in I} A_i = \bigcap_{i\in I}(D\backslash A_i)$ $\quad D\backslash \bigcap_{i\in I} A_i = \bigcup_{i\in I}(D\backslash A_i)$

(where we have written as usual for sets, $X\backslash Y = \{x \in X \mid x \notin Y\}$). You should check that you can justify these. Note that (iv) generalises a *distributive law* for unions and intersections, and (v) is a general form of *de Morgan's law*.

Exercise 1.4 Give examples of sets x, y so that $x \neq y$ but $\bigcup x = \bigcup y$.[Hint: use small sets.]

Exercise 1.5 Show that if $a \in X$ then $\mathcal{P}(a) \in \mathcal{P}(\mathcal{P}(\bigcup X))$.

Exercise 1.6 Show that for any set X: a) $\bigcup \mathcal{P}(X) = X$ b) $X \subseteq \mathcal{P}(\bigcup X)$; when do we have = here?

Exercise 1.7 Show that the distributive laws (iv) above are valid.

Exercise 1.8 Let $I = \mathbb{Q} \cap (0, 1/2)$ be the set of rationals p with $0 < p < 1/2$. Let $A_p = \mathbb{R} \cap (1/2 - p, 1/2 + p)$. Show that $\bigcup_{p\in I} A_p = (0,1)$; $\bigcap_{p\in I} A_i = \{1/2\}$.

Exercise 1.9 Let $X_0 \supseteq X_1 \supseteq \cdots$and $Y_0 \supseteq Y_1 \supseteq$ be two infinite sequences of possibly shrinking sets. Show that $\bigcap_{i\in\mathbb{N}}(X_i \cup Y_i) = \bigcap_{i\in\mathbb{N}} X_i \cup \bigcap_{i\in\mathbb{N}} Y_i$. If we take away the requirement that the sequences be shrinking, does this equality hold in general for any infinite sequences X_i and Y_i?

1.3 Relations and Functions

In this section we shall see how the fundamental mathematical notions of *relation* and *function* can be represented by sets. First relations, and we'll list various properties that relations have. In general we have sets X, Y and a relation R that holds between some of the elements of X and of Y. If X is the set of all points in the plane, and Y the set of all circles, the 'p is the centre of the circle S' determines a relation between X and Y. We shall be more interested in relations between elements of a single set, that is when $X = Y$.

We list here some properties that a relation R can have on a set X. We think of xRy as "x is related by R to y".

Type of relation	*Defining condition*
Reflexive	$x \in X \to xRx$
Irreflexive	$x \in X \to \neg xRx$ (which we may write $x \not R x$)
Symmetric	$(x, y \in X \wedge xRy) \to yRx)$
Antisymmetric	$(x, y \in X \wedge xRy \wedge yRx) \to x = y)$
Connected	$(x, y \in X) \to (x = y \vee xRy \vee yRx)$
Transitive	$(x, y, z \in X \wedge xRy \wedge yRz) \to (xRz)$.

You should recall that the definition of *equivalence relation* is that R should satisfy symmetry, reflexivity, and be transitive.

• If $X = \mathbb{R}$ and $R = \leq$ the usual ordering of the real numbers, then R is reflexive, connected, transitive, and antisymmetric. If we took $R = <$ then the relation becomes irreflexive.

• If $X = \mathcal{P}(A)$ for some set A and we took $xRy \Leftrightarrow x \subseteq y$ for $x, y \in X$ then R is reflexive, antisymmetric, and transitive. If A has at least two elements, then it is not connected since if both $x - y$ and $y - x$ are non-empty, then $\neg xRy \wedge \neg yRx$.

• If T looks like a 'tree', (think perhaps of a family tree) with an ordering aRb as 'a is a descendant of b' then we should only have irreflexivity and transitivity (and rather trivially antisymmetry because we should never have aRb and bRa simultaneously).

1.3.1 Ordering Relations

Of particular interest are *ordering relations* where R is thought of as some kind of ordering with xRy interpreted as x somehow "preceding" or "coming before" y. It is natural to adopt some kind of notation such as $\prec$ or $\preceq$ for such R. The notation of $\prec$ represents a *strict* order: given an ordering where we want reflexivity to hold, then we use $\preceq$, so that then $x \preceq x$ is allowed to hold. We may define $\preceq$ in terms of $\prec$: $x \preceq y \Leftrightarrow x \prec y \vee x = y$. Of course we can define $\prec$ in terms of $\prec$ and $=$ too, and we may want to make a choice as to which of the two relations we think of as 'prior' or more fundamental. In general (but not always) we shall tend to form our definitions and propositions in term of the "stricter" ordering $\prec$, defining $\preceq$ as and when we wish from it.

Definition 1.10 *A relation $\prec$ on a set X is a* (strict) partial ordering *if it is irreflexive and transitive. That is:*

(i) $x \in X \to \neg x \prec x$;

(ii) $(x, y, z \in X \wedge x \prec y \wedge y \prec z) \to (x \prec z)$.

Exercise 1.10 Think about how you would frame an alternative, but equivalent definition of partial order in terms of the non-strict ordering $\preceq$. Which of the defining conditions above do we need?

We saw above that for any set A that $\mathcal{P}(A)$ with $\preceq$ as $\subseteq$ was a (non-strict) partial order. If $Y \subseteq X$ then we shall call $(Y, \prec)$ a *suborder* of $(X, \prec)$. We say that an element $x_0 \in X$ is the *least element* of X (or the *minimum* of X) if $\forall x \in X(x_0 \preceq x)$ and we call it a *minimal* element if $\forall y \in X(\neg y \prec x)$. Note that a minimal element need not be a least element. (This is because a partial order need not be connected: it might have many minimal elements). *Greatest* element and *maximal* elements are defined in the corresponding way.

Notions of *least upper bound* etc. carry over to partially ordered sets:

Definition 1.11 *(i) If $\prec$ is a partial ordering of a set X, and $\varnothing \neq Y \subseteq X$, then an element $z \in X$ is a* lower bound *for Y in X if*
$$\forall y(y \in Y \rightarrow z \preceq y).$$
(ii) An element $z \in X$ is an infimum *or* greatest lower bound *(glb) for Y if (a) it is a lower bound for Y, and (b) if z' is any lower bound for Y then $z' \preceq z$.*
(iii) The concepts of upper bound *and* supremum *or* least upper bound *(lub) are defined analogously.*

• By their definitions *if* an infimum (or supremum) for Y exists, it is unique and we write $\inf(Y)$ ($\sup(Y)$) for it. Note that $\inf(Y)$, if it exists, need not be an element of Y. Similarly for $\sup(Y)$. If Y has a *least* element then in this case it is the infimum, and it obviously belongs to Y.

Definition 1.12 *(i) We say that $f : (X, \prec_1) \longrightarrow (Y, \prec_2)$ is an* order preserving *map of the partial orders $(X, \prec_1), (Y, \prec_2)$ iff*
$$\forall x, z \in X(x \prec_1 z \longrightarrow f(x) \prec_2 f(z)).$$
(ii) Orderings $(X, \prec_1)$ and *$(Y, \prec_2)$ are* (order) isomorphic, *written $(X, \prec_1) \cong (Y, \prec_2)$, if there is an order preserving map between them which is also a bijection.*
(iii) There are completely analogous definitions between nonstrict orders $\preceq_1$ and $\preceq_2$.

• Notice that ($\{$Even natural numbers$\}$,<) is order isomorphic to $(\mathbb{N}, <)$ via the function $f(2n) = n$. However $(\mathbb{Z}, <)$ is not order isomorphic to $(\mathbb{N}, <)$.

• The function $f(k) = k - 1$ is an order isomorphism of $(\mathbb{Z}, <)$ to itself. However as we shall see, there are no order isomorphisms of $(\mathbb{N}, <)$ to itself.

• For a set X with an ordering R, then we may think of the (X, R) as being officially the ordered pair $\langle X, R\rangle$ (to be defined shortly), although it is easier on the eye to simply use the curved brackets.

In one sense *any* partial order of a set X can be represented as partial order where the ordering is $\subseteq$, as the following shows.

Theorem 1.13 *(Representation Theorem for partially ordered sets) If $\prec$ partially orders X, then there is a set Y of subsets of X which is such that $(X, \preceq)$ is* order isomorphic *to $(Y, \subseteq)$.*

Proof: Given any $x \in X$ let $X^x = \{z \in X \mid z \preceq x\}$. Notice then that if $x \neq y$ then $X^x \neq X^y$. So the assignment $x \rightarrowtail X^x$ is (1-1). Let $Y = \{X^x \mid x \in X\}$. Then we have
$$x \preceq y \longleftrightarrow X^x \subseteq X^y \,;$$
consequently, setting $f(x) = X^x$ we have an order isomorphism. Q.E.D.

Often we deal with orderings where every element is comparable with every other - this is "strong connectivity" and we call the ordering "total". The picture of such an ordering has all elements strung out on a line, and so is often called (but not in this course) a 'linear order'.

Definition 1.14 *A relation $\prec$ on X is a* strict total ordering *if it is a partial ordering which is connected: $\forall x, y(x, y \in X \rightarrow (x = y \vee x \prec y \vee y \prec x))$.*

If we use $\preceq$ we call the ordering non-strict *(and the ordering is then reflexive). We can then formulate the connectedness condition as: $\forall x, y(x, y \in X \rightarrow (x \preceq y \vee y \preceq x))$.*

• In a total ordering there is no longer any difference between least and minimal elements, but that does not imply that least elements will always exist (think of the total ordering $(\mathbb{Z}, \leq)$).

• We often drop the word "strict" (or "non-strict") and leave it is as implicit when we use the symbol $\prec$ (or $\preceq$).

• Order preserving maps $f : (X, \prec_1) \longrightarrow (Y, \prec_2)$ between strict total orders must then be (1-1). (Check why?) Moreover, if f is order preserving then it also implies that $\forall x \in X \forall z \in X(x \prec_1 z \longleftarrow f(x) \prec_2 f(z))$ and so we also have equivalence here.

An extremely important notion that we shall come back to study further is that of *wellordering*:

Definition 1.15 *(i) $(A, \prec)$ is a* wellordering *if (a) it is a strict total ordering and (b) for any subset $Y \subseteq A$, if $Y \neq \varnothing$, then Y has a $\prec$-least element. We write in this case $(A, \prec) \in WO$.*

(ii) A partial ordering R on a set A, (A, R) is a wellfounded relation *if for any subset $Y \subseteq A$, if $Y \neq \varnothing$, then Y has an R-minimal element.*

Then $(\mathbb{N}, <)$ is a wellordering, but $(\mathbb{Z}, <)$ is not. Cantor's greatest mathematical contribution was perhaps recognizing the importance of this concept and generalizing it. The theory of wellorderings is fundamental to the notion of ordinal number. If (A, R) is my family tree with xRy if x is a descendant of y, then it is also wellfounded.

Exercise 1.11 If $\langle A, \prec \rangle$ is a total ordering and A is finite, show that it is a wellordering.

Notice that if $(A, \prec)$ is a wellordering, (and to avoid trivialities $A \neq \varnothing$) then we have that A must have a $\prec$-least element, a_0 say. Then $\prec$ still wellorders $A \backslash \{a_0\}$. Hence $A \backslash \{a_0\}$ must have a $\prec$-least element, a_1 say. We may continue in this way, defining $a_0 \prec a_1 \prec \cdots \prec a_n \prec \cdots$. In general we see that because $(A, \prec)$ is a wellordering, not only is there a least element, a_0, but every element $a \in A$ has an *immediate successor* $a \prec a'$, that is with no b such that $a \prec b \prec a'$. To deduce this we only used of the wellorder property that $A \backslash \{a_0, a_1, \ldots, a_n\}$ had a $\prec$-least element. We shall say that $C \subseteq A$ is an *end segment* of the strict total order $(A, \prec)$, if whenever $a \in C$ and $a \prec b$ then $b \in C$. Building on this idea we have that:

Lemma 1.16 *A strict total ordering $(A, \prec)$ is a wellordering if and only if any non-empty end segment C of A, has a $\prec$-least element.*

The proof is left as an Exercise.

1.3.2 Ordered Pairs

We have talked about relations R that may hold between objects, and even used the notation $\preceq$ if we wanted to think of the relation as an ordering. However we shall want to see how we can specify relations using sets. From that it is a short step to do the same for functions. The key building block is the notion of *ordered pair.*

Definition 1.17 *(Kuratowski) Let x, y be sets. The* ordered pair set *of x and y is the set*

$$\langle x, y \rangle =_{df} \{\{x\}, \{x, y\}\}.$$

Why do we need this? Because $\{x,y\}$ is by definition *unordered*: $\{x,y\} = \{y,x\}$. Hence $\{x,y\} = \{u,v\} \longrightarrow x = u \wedge y = v$ fails. However:

Lemma 1.18 *(Uniqueness theorem for ordered pairs)*

$$\langle x,y\rangle = \langle u,v\rangle \longleftrightarrow x = u \wedge y = v.$$

Proof: ($\longleftarrow$) is trivial. So suppose $\langle x,y\rangle = \langle u,v\rangle$. *Case 1* $x = y$. Then $\langle x,y\rangle = \langle x,x\rangle = \{\{x\},\{x,x\}\} = \{\{x\},\{x\}\} = \{\{x\}\}$. If this equals $\langle u,v\rangle$ then we must have $u = v$ (why? otherwise $\langle u,v\rangle$ would have two elements). So $\langle u,v\rangle = \{\{u\}\} = \{\{x\}\}$. Hence, by Extensionality $\{u\} = \{x\}$, and so, again using Extensionality, $u = x = y = v$.

Case 2 $x \neq y$. Then $\langle x,y\rangle$ and $\langle u,v\rangle$ have the same two elements. (Hence $u \neq v$.) Hence one of these elements has one member, and the other two. Hence we cannot have $\{x\} = \{u,v\}$. So $\{x\} = \{u\}$ and $x = u$. But that means $\{x,y\} = \{u,y\} = \{u,v\}$. So of these last two sets, if they are the same then $y = v$.
Q.E.D.

Example 1.19 *We think of points in the Cartesian plane $\mathbb{R}^2$ as ordered pairs: $\langle x,y\rangle$ with two coordinates, with x "first" on one axis, y on the other.*

Definition 1.20 *We define* ordered k-tuple *by induction:* $\langle x_1,x_2\rangle$ has been defined; if $\langle x_1,x_2,\cdots,x_k\rangle$ *has been defined, then* $\langle x_1,\cdots,x_k,x_{k+1}\rangle =_{df} \langle\langle x_1,\cdots,x_k\rangle,x_{k+1}\rangle$

- Thus $\langle x_1,x_2,x_3\rangle = \langle\langle x_1,x_2\rangle,x_3\rangle$, $\langle x_1,x_2,x_3,x_4\rangle = \langle\langle\langle x_1,x_2\rangle,x_3\rangle,x_4\rangle$ *etc.* Note that once we have the uniqueness theorem for ordered pairs, we automatically have it for ordered triples, quadruples,... that is: $\langle x_1,x_2,x_3\rangle = \langle z_1,z_2,z_3\rangle \leftrightarrow x_i = z_i (0 < i \leq 3)$ *etc.*

This leads to:

Definition 1.21 *(i) Let A, B be sets. $A \times B =_{df} \{\langle x,y\rangle \mid x \in A \wedge y \in B\}$. If $A = B$ this is often written as A^2.*

(ii) If $A_1,\ldots A_{k+1}$ are sets, we define (inductively)

$A_1 \times A_2 \times \cdots \times A_{k+1} =_{df} (A_1 \times A_2 \times \cdots \times A_k) \times A_{k+1}$

$(\textit{which equals}: \{\langle\cdots\langle\langle\langle x_1,x_2\rangle,x_3\rangle,\cdots,x_k\rangle,x_{k+1}\rangle \mid \forall i\,(1 \leq i \leq k+1 \rightarrow x_i \in A_i)\})$.

- In general $A \times B \neq B \times A$ and further, the $\times$ operation is not associative.

Exercise 1.12 Suppose for no sets x, u do we have $x \in u \in x$. Then if we define $\langle x,y\rangle_1 = \{x,\{x,y\}\}$ then show $\langle x,y\rangle_1$ also satisfies the Uniqueness statement of Lemma 1.18.

Exercise 1.13 Does $\{\{x\},\{x,y\},\{x,y,z\}\}$ give a good definition of ordered triple? Does $\{\langle x,y\rangle,\langle y,z\rangle\}$?

Exercise 1.14 Let $\mathfrak{P}$ be the class of all ordered pairs. Show that $\mathfrak{P}$ is a proper class - that is - it is not a set. [Hint: suppose for a contradiction it was a set; apply the axiom of union.]

Exercise 1.15 Show that if $x \in A, y \in A$ then $\langle x,y\rangle \in \mathcal{P}(\mathcal{P}(A))$. Deduce that if $x, y \in V_n$ then $\langle x,y\rangle \in V_{n+2}$.

Exercise 1.16 Show that $A \times (B \cup C) = (A \times B) \cup (A \times C)$. Show that if $A \times B = A \times C$ and $A \neq \varnothing$, then $B = C$.

Exercise 1.17 Show that $A \times \bigcup B = \bigcup\{A \times X \mid X \in B\}$.

Exercise 1.18 We define the 'unpairing functions' $(u)_0$ and $(u)_1$ so that if $u = \langle x,y\rangle$ then $(u)_0 = x$ and $(u)_1 = y$. Show that these can be expressed as: $(u)_0 = \bigcup\bigcap u$; $(u)_1 = \bigcup(\bigcup u - \bigcap u)$ if $\bigcup u \neq \bigcap u$; and $(u)_1 = \bigcup\bigcup u$ otherwise.

Definition 1.22 *(i) A (binary) relation R is a class of ordered pairs. R is thus any subset of some $A \times B$. (ii) We write: $R^{-1} =_{df} \{\langle y, x\rangle \mid \langle x, y\rangle \in R\}$.*

Example 1.23 *(i) $R = \{\langle 0,1\rangle, \langle 1,2\rangle, \langle 1,3\rangle, \langle 2,0\rangle\}$ is a relation. So are:*
(ii) $S_1 = \{\langle x, y\rangle \in \mathbb{R}^2 \mid x \leq y^2\}$;
(iii) $S_2 = \{\langle x, y\rangle \in \mathbb{R}^2 \mid x^2 = y\}$.
(iv) If x is any set, the identity relation *on x is $\mathrm{id}_x =_{df} \{\langle z, z\rangle \mid z \in x\}$.*
(v) A partial ordering can also be considered a relation: $R = \{\langle x, y\rangle \mid x \preceq y\}$.

Definition 1.24 *If R is a relation, then*

$$\mathrm{dom}(R) =_{df} \{x \mid \exists y \langle x, y\rangle \in R\}, \ \mathrm{ran}(R) =_{df} \{y \mid \exists x \langle x, y\rangle \in R\}.$$

The field *of a relation R,* $\mathrm{Field}(R)$, *is* $\mathrm{dom}(R) \cup \mathrm{ran}(R)$.

• With these definitions we can say that if R is a relation, then $R \subseteq \mathrm{dom}(R) \times \mathrm{ran}(R)$. Check that $\mathrm{Field}(R) = \bigcup\bigcup R$.

Notice it would be natural to want to next define a *ternary relation* as an R which is a subset of some $A \times B \times C$ say. But of course elements of this are also ordered pairs, namely something of the form $\langle\langle a, b\rangle, c\rangle$. Then $\mathrm{dom}(R) \subseteq A \times B$, $\mathrm{ran}(R) \subseteq C$. Hence ternary relations are just special cases of (binary) relations, and the same is then true for k-ary relations.

Ultimately functions are just special kinds of relations.

Definition 1.25 *(i) A relation F is a* function *("*$\mathrm{Func}(F)$*") if*
$\forall x \in \mathrm{dom}(F)$*(there is a unique y with $\langle x, y\rangle \in F$).*
(ii) If F is a function then F is (1-1) *iff* $\forall x, x'(\langle x, y\rangle \in F \wedge \langle x', y\rangle \in F \longrightarrow x = x')$.

• In the last Example (iii) and (iv) are functions; (i) and (ii) are not.

• It is much more usual to write for functions "$F(x) = y$" for "$\langle x, y\rangle \in F$". (ii) then becomes the more familiar: $\forall x \forall x'[F(x) = F(x') \to x = x']$. We also write "$F : X \to Y$" instead of "$F \subseteq X \times Y$" (with Y called the *co-domain* of f). Then "F is surjective", or "onto" becomes $\forall y \in Y(\exists x \in X(F(x) = y)$. A function $F : X \to Y$ is a *bijection* if it is both (1-1) and onto (and we write "$F : X \longleftrightarrow Y$"). If $F : A \times B \longrightarrow C$, we write $F(a, b) = c$ rather than the more formally correct $F(\langle a, b\rangle) = c$.

Notation 1.26 *Suppose $F : X \to Y$ then*
(i) $F``A =_{df} \{y \in Y \mid \exists x \in A(F(x) = y)\}$. We call $F``A$ the range of F on A.
(ii) $F \upharpoonright A =_{df} \{\langle x, y\rangle \in F \mid x \in A\}$. $F \upharpoonright A$ is the restriction of F to A.
(iii) If additionally $G : Y \longrightarrow Z$ we write $G \circ F : X \longrightarrow Z$ for the composed function defined by $G \circ F(x) = G(F(x))$.

• In this terminology $F``A = \mathrm{ran}\,(F \upharpoonright A)$.

Exercise 1.19 (i) Find a counterexample to the assertion $F \cap A^2$ equals $F \upharpoonright A$.
(ii) Show $F \upharpoonright A = F \cap (A \times \mathrm{ran}(F))$.

Exercise 1.20 As a further exercise in using this notation, suppose T is a class of functions, with the property that that for any two $f, g \in T$, $f \upharpoonright (\mathrm{dom}(f) \cap \mathrm{dom}(g)) = g \upharpoonright (\mathrm{dom}(f) \cap \mathrm{dom}(g))$ (more simply put: they both agree on the part of their domains they have in common). Then check a) $F = \bigcup T$ is a function, and b) $\mathrm{dom}(F) = \bigcup\{\mathrm{dom}(g) \mid g \in T\}$.

Again we don't need a new definition for n-ary functions: such a function $F : A_1 \times \cdots \times A_n \to B$ is again a relation $F \subseteq A_1 \times \cdots \times A_n \times B$. Then, quite naturally, $\mathrm{dom}(F) = A_1 \times \cdots \times A_n$.

As well as considering functions as special kinds of relations, which are in turn special kinds of sets, we shall want to be able to talk about sets of functions. Then:

Definition 1.27 *If X, Y are sets, then ${}^X Y =_{df} \{F \mid F : X \to Y\}$.*

Exercise 1.21 Suppose X, Y both have rank n ("$\rho(X) = n$" - see Ex.1.3). Compute a) $\rho(X \times Y); b) \rho({}^Y X)$. [Hint for b) : show first if $X, Y \in Z$ show that ${}^Y X \in \mathcal{PPPP}(Z)$.]

Definition 1.28 *(Indexed Cartesian Products). Let I be a set, and for each $i \in I$ let $A_i \neq \varnothing$ be a set; then*

$$\prod_{i \in I} A_i =_{df} \{f \mid \mathrm{Func}(f), \mathrm{dom}(f) = I \wedge \forall i \in I(f(i) \in A_i)\}$$

This allows us to take Cartesian products indexed by any set, not just some finite n.

Example 1.29 *(i) Let $I = \mathbb{N}$. Each $A_i = \mathbb{R}$. Then $\prod_{i \in I} A_i$ is the same as ${}^{\mathbb{N}}\mathbb{R}$ the set of infinite sequences of reals numbers.*

(ii) Let G_i be a group for each i in some index set I; then it is possible to put a group multiplication structure on $\prod_{i \in I} G_i$ to turn it into a group.

1.4 Transitive Sets

We think of a transitive set as one without any "$\in$-holes".

Definition 1.30 *A set x is* transitive, $\mathrm{Trans}(x)$, *iff $\forall y \in x(y \subseteq x)$. We also equivalently abbreviate* $\mathrm{Trans}(x)$ *by $\bigcup x \subseteq x$.*

- Note that easily $\mathrm{Trans}(x) \leftrightarrow \bigcup x \subseteq x$: assume $\mathrm{Trans}(x)$; if $y \in z \in x$ then, as we have $z \subseteq x$, we have $y \in x$. We conclude that $\bigcup x \subseteq x$. Conversely: if $\bigcup x \subseteq x$ then for any $y \in x$ by definition of $\bigcup$, $y \subseteq \bigcup x$, hence $y \subseteq x$ and thus $\mathrm{Trans}(x)$.

Example 1.31 *(i) $\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}$ are transitive. $\{\{\varnothing\}\}, \{\varnothing, \{\{\varnothing\}\}\}$ are not.*

Definition 1.32 *(The successor function) Let x be a set. Then $S(x) =_{df} x \cup \{x\}$.*

Exercise 1.22 Show the following: (i) Let $\mathrm{Trans}(Z) \wedge x \subseteq Z$.Then $Z \cup \{x\}$ is transitive.

(ii) If x, y are transitive, then so are: $S(x), x \cup y, x \cap y, \bigcup x$.

(iii) Let X be a class of transitive sets. then $\bigcup X$ is transitive. If $X \neq \varnothing$, then $\bigcap X$ is transitive.

(iv) Show that $\mathrm{Trans}(x) \longleftrightarrow \mathrm{Trans}(\mathcal{P}(x))$. Deduce that each V_n is transitive.

Lemma 1.33 $\mathrm{Trans}(x) \longleftrightarrow \bigcup S(x) = x$.

Proof: First note that $\bigcup S(x) = \bigcup(x \cup \{x\}) = \bigcup x \cup \bigcup\{x\} = \bigcup x \cup x$. For $(\longrightarrow)$, assume $\mathrm{Trans}(x)$; then $\bigcup x \subseteq x$. Hence $x \subseteq \bigcup S(x) \subseteq x$. Q.E.D.

Exercise 1.23 Prove the $(\longleftarrow)$ direction of the last lemma.

Exercise 1.24 (i) What sets would you have to add to $\{\{\{\varnothing\}\}\}$ to make it transitive?

(ii) In general given a set x think about how a transitive y could be found with $y \supseteq x$. (It will turn out (below) that for any set x there is a smallest $y \supseteq x$ with $\mathrm{Trans}(y)$.) [Hint: consider repeated applications of $\bigcup$: $\bigcup^0 x =_{df} x; \bigcup^1 x =_{df} \bigcup x, \bigcup^2 x =_{df} \bigcup(\bigcup^1 x), , \ldots, \bigcup^{n+1} x =_{df} \bigcup(\bigcup^n x) \ldots$ as in the next definition.]

Definition 1.34 **Transitive Closure** TC *We define by recursion on n:*

$$\bigcup^0 x = x\,;\ \bigcup^{n+1} x = \bigcup(\bigcup^n x);\ \ \mathrm{TC}(x) = \bigcup\{\bigcup^n x \mid n \in \mathbb{N}\}.$$

The idea is that by taking a $\bigcup$ we are "filling in $\in$-holes" in the sets. Informally we have thus defined $\mathrm{TC}(x) = x \cup \bigcup^1 x \cup \bigcup^2 x \cup \bigcup^3 x \cup \cdots \bigcup^n x \cup \cdots$ but the right hand side cannot be an 'official formula' as it is an infinitely long expression! But the above definition by recursion makes matters correct.

Exercise 1.25 Show that $y \in \bigcup^n x \leftrightarrow \exists x_n, x_{n-1}, \ldots, x_1(y \in x_n \in x_{n-1} \in \cdots \in x_1 \in x)$.

Note by constuction that $\mathrm{Trans}(\mathrm{TC}(x))$: $y \in \mathrm{TC}(x)$ if and only if for some n $y \in \bigcup^n x$. Then $y \subseteq \bigcup^{n+1} x \subseteq \mathrm{TC}(x)$.

Lemma 1.35 ***(Lemma on*** TC***)*** *For any set x (i)* $x \subseteq \mathrm{TC}(x)$ *and* $\mathrm{Trans}(\mathrm{TC}(x))$ *; (ii) If* $\mathrm{Trans}(t) \wedge x \subseteq t \to \mathrm{TC}(x) \subseteq t$. *Hence* $\mathrm{TC}(x)$ *is the smallest transitive set t satisfying* $x \subseteq t$. *(iii) Hence* $\mathrm{Trans}(x) \leftrightarrow \mathrm{TC}(x) = x$.

Proof (i) This clear as $x = \bigcup^0 x \subseteq \mathrm{TC}(x)$, and by the comment above.

(ii): $x \subseteq t \to \bigcup^0 x \subseteq t$. Now by induction on k, assume $\bigcup^k x \subseteq t$. Now use $A \subseteq B \wedge \mathrm{Trans}(B) \to \bigcup A \subseteq B$ to deduce $\bigcup^{k+1} x \subseteq t$ and it follows that $\mathrm{TC}(x) \subseteq t$. However t was any arbitrary transitive set containing x. (iii): $x \subseteq \mathrm{TC}(x)$ by (i). If $\mathrm{Trans}(x)$ then substitute x for t in the above: we conclude $\mathrm{TC}(x) \subseteq x$. Q.E.D

As $\mathrm{TC}(x)$ is the smallest transitive set containing x we could write this as $\mathrm{TC}(x) = \bigcap\{t \mid \mathrm{Trans}(t) \wedge x \subseteq t)\}$ (the latter is indeed transitive, see Ex. 1.22).

Exercise 1.26 (i) Show that $y \in x \to \mathrm{TC}(y) \subseteq \mathrm{TC}(x)$.

(ii) $\mathrm{TC}(x) = x \cup \bigcup\{\mathrm{TC}(y) \mid y \in x\}$ (hence $\mathrm{TC}(\{x\}) = \{x\} \cup \mathrm{TC}(x)$.)

The point to note is that taking $\mathrm{TC}(x)$ ensures that $\langle \mathrm{TC}(x), \in\rangle$ satisfies transitivity as a partial ordering.

Exercise 1.27 If f is a (1-1) function show that $f^{-1} \subseteq \mathcal{P}\mathcal{P}\left(\bigcup\{\mathrm{dom}(f), \mathrm{ran}(f)\}\right)$.

Number Systems

We see how to extend the theory of sets to build up the natural numbers $\mathbb{N}$. It was R. Dedekind (1831-1916) who was the first to realise that notions such as "infinite number system" needed proper definitions, and that the claim that a function could be defined by mathematical induction or recursion required proof. This required him to investigate the notion of such infinite systems. About the same time G. Peano (1858-1932) published a list of axioms (derived from Dedekind's work) that the structure of the natural numbers should satisfy.

Figure 2.1: Richard Dedekind

2.1 The natural numbers

Proceeding ahistorically, there were several suggestions as to how sets could represent the natural numbers $0, 1, 2, \ldots$.

E. Zermelo (1908) suggested the sequence of sets $\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\{\{\varnothing\}\}\}, \ldots$ Later von Neumann (1903-1957) suggested a sequence that has since become the usually accepted one. Recall Def.1.32.

$0 =_{df} \varnothing$,
$1 =_{df} \{0\} = \{\varnothing\} = 0 \cup \{0\} = S(0)$,
$2 =_{df} \{0, 1\} = \{\varnothing, \{\varnothing\}\} = 1 \cup \{1\} = S(1)$,
$3 =_{df} \{0, 1, 2\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} = 2 \cup \{2\} = S(2)$.

In general $n =_{df} \{0, 1, \ldots, n-1\}$. Note that with the von Neumann numbers we also have that for any n $S(n) = n+1$: $1 = S(\varnothing)$, $2 = S(1)$ etc. This latter system has the advantage that "n" has exactly n members, and is the set of all its predecessors in the usual ordering. Both Zermelo's and von Neumann's numbers have the advantage that they can be easily generated. We shall only work with the von Neumann numbers.

Definition 2.1 *A set Y is called* inductive *if (a) $\varnothing \in Y$ (b) $\forall x \in Y(S(x) \in Y)$.*

Notice that we have nowhere yet asserted that there are sets which are infinite (not that we have defined the term either). Intuitively though we can see that any inductive set which has to be closed under S cannot be finite: $\varnothing, S(\varnothing), S(S(\varnothing))$ are all distinct (although we have not proved this yet). We can remedy this through:

Axiom of Infinity: There exists an inductive set:
$$\exists Y(\varnothing \in Y \wedge \forall x \in Y(S(x) \in Y)).$$

One should note that a picture of an inductive set would show that it consists of "S-chains": $\varnothing, S(\varnothing), SS(\varnothing), \ldots$. but possibly also others of the form $u, S(u), SS(u), SSS(u) \ldots$. thus starting with other sets u. Given this axiom we can give a definition of natural number.

Definition 2.2 *(i) x is a* natural number *if $\forall Y[Y$ is an inductive set $\longrightarrow x \in Y]$.*
(ii) ω is the class of natural numbers.

We have defined: $\omega = \bigcap\{Y \mid Y \text{ an inductive set}\}$
by taking an intersection over (what one can show is a proper) class of all inductive sets. But is it a set?

Proposition 2.3 *ω is a set.*

Proof: Let z be any inductive set (by the Ax. of Inf. there is such a z). By the Axiom of Subsets: there is a set N so that:
$$N = \{x \in z \mid \forall Y[Y \text{ an inductive set} \longrightarrow x \in Y]\}.$$ Q.E.D.

Proposition 2.4 *(i) ω is an inductive set. (ii) It is thus the smallest inductive set.*

Proof: We have proven in the last lemma that ω is a set. To show it is inductive, note that by definition $\varnothing$ is in any inductive set Y so $\varnothing \in \omega$. Hence (a) of Def. 2.1 holds. Moreover, if $x \in \omega$, then for any inductive set Y, we have both x and $S(x)$ in Y. Hence $S(x) \in \omega$. So ω is closed under the S function. So *(b)* of Def. 2.1 holds. (ii) is immediate. Q.E.D.

To paraphrase the above: if we have an inductive subset of ω we know it is all of ω. It may seem odd that we define the set of natural numbers in this way, rather than as the single chain $\varnothing, S(\varnothing), \ldots$ and so on. However it is the insight of Dedekind's analysis that we obtain the powerful principle of induction, which of course is of immense utility. Note that we may *prove* this principle, which is prior to defining *order, addition,* etc. We formally state this as a principle about inductive sets given by some property Φ:

THEOREM 2.5 **(Principle of Mathematical Induction)**
Suppose Φ is a welldefined definite property of sets. Then

$$[\Phi(0) \wedge \forall x \in \omega(\Phi(x) \longrightarrow \Phi(S(x))] \longrightarrow \forall x \in \omega\Phi(x).$$

Proof: Assume the antecedent here, then it suffices to show that the set of $x \in \omega$ for which $\Phi(x)$ holds is inductive. Let $Y = \{x \in \omega \mid \Phi(x)\}$. However the antecedent then says $0 \in Y$; and moreover if $x \in Y$ then $S(x) \in Y$. That Y is inductive is then simply the antecedent assumption. Hence $\omega \subseteq Y$. And so $\omega = Y$. Q.E.D.

PROPOSITION 2.6 *Every natural number y is either 0 or is $S(x)$ for some natural number x.*

• To emphasise: this need not be true for a general inductive set: not every element can be necessarily be "reached eventually" by repeated application of S to $\varnothing$.
Proof: Let $Z = \{y \in \omega \mid y = 0 \vee \exists x \in \omega(S(x) = y)\}$. Then $0 \in Z$ and if $u \in Z$ then $u \in \omega$. Hence $S(u) \in \omega$, (as ω is inductive). Hence $S(u) \in Z$. So Z is inductive and is thus ω.
• One should note that actually the Principle of Mathematical Induction has been left somewhat vague: we did not really specify what "*a welldefined property*" was. This we can make precise just as we can for the Axiom of Subsets: it is any property that can be expressed using a formal language for sets.

EXERCISE 2.1 Every natural number is transitive. [Hint: Use Principle of Mathematical Induction - in other words, show that the set of transitive natural numbers is inductive.]

LEMMA 2.7 *ω is transitive.*

Proof: Let $X = \{n \in \omega \mid n \subseteq \omega\}$. If $X = \omega$ then by definition $\mathrm{Trans}(\omega)$. So we show that X is inductive. $\varnothing \in X$; assume $n \in X$, then $n \subseteq \omega$ and $\{n\} \subseteq \omega$, hence $n \cup \{n\} \subseteq \omega$. Hence $S(n) \in X$. So X is inductive, and $\omega = X$. Q.E.D.

2.2 Peano's Axioms

Dedekind formulated a group of axioms could that capture the important properties of the natural numbers. They are generally known as "Peano's Axioms." We shall consider general "Dedekind systems":

A *Dedekind system* is a triple $\langle N, s, e\rangle$ where
(a) N is a set with $e \in N$;
(b) $\mathrm{Func}(s) \wedge s : N \longrightarrow N$ and s is (1-1) ;
(c) $e \notin \mathrm{ran}(s)$;
(d) $\forall K \subseteq N(e \in K \wedge s``K \subseteq K \rightarrow K = N)$.

Note that $s``K \subseteq K$ is another way of saying that K is closed under the s function. We shall prove that our natural numbers form a Dedekind system; furthermore, any structure that satisfies (a) - (d) will look like ω.

Firstly then, let $\sigma=\{\langle k, S(k)\rangle \mid k \in \omega\} = S \upharpoonright \omega$ the restriction of the successor operation on sets in general, to the natural numbers.

PROPOSITION 2.8 *$\langle \omega, \sigma, 0\rangle$ forms a Dedekind system.*

Proof: We have that $0 \in \omega$, $\sigma : \omega \to \omega$, and that $0 \neq \sigma(u)(\varnothing \neq S(u))$ for any u. The axiom (d) of Dedekind system just says for $\langle \omega, \sigma, 0\rangle$ that any subset $A \subseteq \omega$, that is, of the structure's domain, that contains 0 and is closed under σ (*i.e.* that is inductive) is all of ω. But ω itself *is* the *smallest* inductive set. So certainly $A = \omega$. So (a),(c)-(e) hold and all that is left is to show that σ is (1-1). Suppose $S(m) = \sigma(m) = \sigma(n) = S(n)$. Hence $\bigcup S(m) = \bigcup S(n)$. By the last exercise $\mathrm{Trans}(m)$, $\mathrm{Trans}(n)$. By Lemma 1.33, $\bigcup S(m) = m$, and $\bigcup S(n) = n$; so $m = n$. Q.E.D.

REMARK 2.9 We shall later be showing that any two Dedekind systems are isomorphic.

2.3 THE WELLORDERING OF ω

DEFINITION 2.10 *For $m, n \in \omega$ set $m < n \iff m \in n$. Set $m \leq n \iff m = n \vee m < n$.*

Note that if $m \in \omega$ then $m < S(m)$ by definition of $<$ and S.

LEMMA 2.11 *(i) $<$ (and $\leq$) are transitive; (ii) $\forall n \in \omega \forall m(m < n \leftrightarrow S(m) < S(n))$; (iii) $\forall m \in \omega(m \nless m)$.*

Proof: (i) That $<$ is transitive follows from the fact that our natural numbers are proven (Ex.2.1) to be transitive sets: $n \in m \in k \to n \in k$.

(ii): $(\leftarrow)$If $S(m) < S(n)$ then we have $m \in S(m) \in S(n) = n \cup \{n\}$. If $S(m) = n$, then $m \in S(m) = n$, so $m < n$. If $S(m) \in n$ then as $\mathrm{Trans}(n)$ we have $m \in n$ and so $m < n$. $(\to)$ We prove the converse by the Principle of Mathematical Induction (PMI). Let $\Phi(k)$ say: "$\forall m(m < k \to S(m) < S(k))$". Then $\Phi(0)$ vacuously; and so we suppose $\Phi(k)$, and prove $\Phi(S(k))$.

Let $m < S(k)$. Then $m \in k \cup \{k\}$. If $m \in k$ then, by $\Phi(k)$ we have $S(m) < S(k) < S(S(k))$. If $m = k$ then $S(m) = S(k) < S(S(k))$. Either way we have $\Phi(S(k))$. By PMI we have $\forall n \Phi(n)$.

(iii) Note $0 \nless 0$ since $0 \notin 0$. If $k \notin k$ then $S(k) \notin S(k)$ by part (ii).

So $X = \{k \in \omega \mid k \notin k\}$ is inductive, *i.e.* all of ω. Q.E.D.

LEMMA 2.12 *$<$ is a strict total ordering.*

Proof: All we have left to prove is connectivity (often called *Trichotomy*): $\forall m, n \in \omega(m = n \vee m < n \vee n < m)$. Notice that at most one of these three alternatives can hold for m, n: if, say, the first two then we should have $n < n$, and if the second two then $m < m$ (by transitivity of $<$) and these contradict irreflexivity, *i.e.*, (iii) of the last Lemma. Let $X = \{n \in \omega \mid \forall m \in \omega(m = n \vee m < n \vee n < m)\}$. If X is inductive, the proof is complete. This is an Exercise. Q.E.D.

EXERCISE 2.2 Show this X is inductive.

EXERCISE 2.3 Show that $\forall m, n \in \omega(n < m \leftrightarrow n \subsetneq m)$.

Theorem 2.13 **(Wellordering Theorem for ω)** *Let $X \subseteq \omega$. Then either $X = \varnothing$ or there is $n_0 \in X$ so that for any $m \in X$ either $n_0 = m$ or $n_0 \in m$.*

Note: such an n_0 can clearly be called the "least element of X", since $\forall m \in X(n_0 \leq m)$. Thus the wellordering theorem, can be rephrased as:

Least Number Principle: *any non-empty set of natural numbers has a least element.*

Proof: (of 2.13) Suppose $X \subseteq \omega$ but X has no least element as above. Let

$$Z = \{k \in \omega \mid \forall n < k(n \notin X)\}.$$

We claim that Z is inductive, hence all of ω and so $X = \varnothing$. This suffices. Vacuously $0 \in Z$. Suppose now $k \in Z$. Let $n < S(k)$. Hence $n \in k \cup \{k\}$. If $n \in k$ then $n \notin X$ (as $n < k$ and $k \in Z$). But if $n \in \{k\}$ then $n = k$ and so $n \notin X$ because otherwise it would be the least element of X and X does not have such. So $S(k) \in Z$. Hence Z is inductive. Q.E.D.

Exercise 2.4 Let $X \neq \varnothing, X \subseteq \omega$. Show that there is $n \in X$, with $n \cap X = \varnothing$.

Exercise 2.5 **(Principle of Strong Induction for ω)** Suppose Φ is a definite welldefined property of natural numbers. Show that

$$\forall n[\forall k < n\Phi(k) \to \Phi(n)] \to \forall n\Phi(n).$$

[Hint: Suppose for a contradiction $X = \{n \in \omega \mid \neg\Phi(n)\} \neq \varnothing$. Apply the Least Number Principle.]

2.4 The Recursion Theorem on ω

We shall now show that it is legitimate to define functions by *recursion on ω*.

Theorem 2.14 **(Recursion theorem on ω)** *Let A be any set, $a \in A$, and $f : A \to A$, any function. Then there exists a unique function $h : \omega \to A$ so that*

(i) $h(0) = a$;

(ii) For any $n \in \omega$: $h(S(n)) = f(h(n))$.

Proof: We shall find h as a union of *k-approximations* where u is a *k-approximation* if

a) $\mathrm{Func}(u) \wedge \mathrm{dom}(u) = k$; b) If $k > 0$ then $u(0) = a$; if $k > S(n)$ then $u(S(n)) = f(u(n))$.

In other words u satisfies the defining clauses above for our intended h - without our requiring that $\mathrm{dom}(u)$ is all of ω.

Note: (i) that $\{\langle 0, a\rangle\}$ is the only 1-approximation. $\{\langle 0, a\rangle, \langle 1, f(a)\rangle\}$ is a 2-approximation. $\varnothing$ is a 0-approximation: this is because the empty set counts as a function with empty domain, hence it can be considered a 0-approximation.

(ii) If u is a k-approximation and $l \leq k$ then $u \restriction l$ is an l-approximation.

(iii) If u is a k-approximation, and $u(k-1) = c$ for some c say, then $u' = u \cup \{\langle k, f(c)\rangle\}$ is a $k+1$-approximation. Hence an approximation may always be extended.

(1) *If u is a k-approximation and v is a k'-approximation, for some $k \le k'$ then $v \upharpoonright k = u$ (and hence $u \subseteq v$).*

Proof: If not let $0 \le m < k$ be least with $u(m) \ne v(m)$. Then by b) $u(0) = a = v(0)$ so $m \ne 0$. So $m = S(m')$ and $u(m') = v(m')$. But then again by b) $u(m) = f(u(m')) = f(v(m')) = v(m)$. Contradiction! QED (1).

Exactly the same proof also shows:

(2) *(Uniqueness) If h exists, then it is unique.*

Proof: Suppose h, h' are two different functions satisfying (i) and (ii) of the theorem. Then $X = \{n \in \omega \mid h(n) \ne h'(n)\}$ is non-empty. By the least number principle, (or in other words the Wellordering Theorem for ω), there is a least number $n_0 \in X$. But then $h \upharpoonright n_0 + 1$, and $h' \upharpoonright n_0 + 1$ are two different $n_0 + 1$ approximations. This contradicts (1) which states that they must be equal. Contradiction! So $X = \varnothing$. QED (2).

(3) *(Existence). Such an h exists.*

Proof: (This is the harder part.) Let $u \in B \Longleftrightarrow \exists k \in \omega(u$ is a k-approximation). We have seen any two such approximations agree on the common part of their domains. In other words, for any $u, v \in B$ either $u \subseteq v$ or $v \subseteq u$. So we take $h = \bigcup B$.

(i) *h is a function.*

Proof: If $\langle n, c\rangle$ and $\langle n, d\rangle$ are in h, with $c \ne d$ then there must be two different approximations u with $u(n) = c$, and v with $v(n) = d$. But this is impossible by (1)!

(ii) $\mathrm{dom}(h) = \omega$.

Proof: Let $\varnothing \ne X =_{df} \{n \in \omega \mid n \notin \mathrm{dom}(h)\}$. By definition of h this means also $X = \{n \in \omega \mid$ there is no approximation u with $n \in \mathrm{dom}(u)\}$. By Note (i) above $\{\langle 0, a\rangle\}$ is the 1-approximation and is in B, so we have that the least element of X is not 0. Suppose it is $n_0 = S(m)$. As $m \notin X$, there must be an n_0-approximation u with, let us say $u(m) = c$. But then by Note (iii) above, $u \cup \{\langle n_0, f(c)\rangle\}$ is a legitimate $S(n_0)$-approximation. So $n_0 \notin X$. Contradiction! Q.E.D.

In short: $h(n)$ is that value given by $u(n)$ for any approximation with $n \in \mathrm{dom}(u)$.

EXAMPLE 2.15 *Let $n \in \omega$. We can define an "add n" function $A_n(x)$ as follows:*

$A_n(0) = n$;

$A_n(S(k)) = S(A_n(k))$.

We shall write from now "$n + 1$" for $S(n)$. Then we would more commonly write $A_n(k)$ as $n + k$. Note that the final clause of A_n then says $n + (k+1) = (n+k)+1$. Assuming we have defined the addition functions $A_n(x)$ for any n:

EXAMPLE 2.16 *(i) $M_n(x)$ function: $M_n(0) = 0; M_n(k+1) = M_n(k) + n$.*

(ii) $E_n(x)$: $E_n(0) = 1; E_n(k+1) = E_n(k) \cdot n$

Again we more commonly write these as $M_n(k)$ as $n \cdot k$, and $E_n(k)$ as n^k.

PROPOSITION 2.17 *The following laws of arithmetic hold for our definitions:*

(a) $m + (n + p) = (m + n) + p$

(b) $m + n = n + m$
(c) $m \cdot (n + p) = m \cdot n + m \cdot p$
(d) $m \cdot (n \cdot p) = (m \cdot n) \cdot p$
(e) $m{\cdot}n{=}n{\cdot}m$
(f) $m^{n+p} = m^n \cdot m^p$
(g) $(m^n)^p = m^{n \cdot p}$.

Proof: These are all proven by induction. As a sample we do (c) (assuming (a) and (b) proven). We do the induction on p. $p = 0$: then $m.(n+0) = m.n = m.n + m.0$. Suppose it holds for p. Then

$$\begin{aligned} m \cdot (n + (p+1)) &= m \cdot ((n+1) + p) \quad \text{(by (a))} \\ &= m \cdot (n+1) + m \cdot p \quad \text{(inductive hypothesis)} \\ &= (m \cdot n + m) + m \cdot p \quad \text{(by definition of } M_m\text{)} \\ &= m \cdot n + (m + m \cdot p) = m \cdot n + m \cdot (p+1) \end{aligned}$$

using (a) again and finally (b) and the definition of M_m. Q.E.D.

Exercise 2.6 Prove some of the other clauses of the last Proposition.

Now the promised isomorphism theorem on Dedekind systems.

Theorem 2.18 *Let $\langle N, s, e\rangle$ be any Dedekind system. Then $\langle \omega, \sigma, 0\rangle \cong \langle N, s, e\rangle$.*

Proof: By the Recursion Theorem on ω (2.14) there is a function $f : \langle \omega, \sigma, 0\rangle \to \langle N, s, e\rangle$ defined by: $f(0) = e$;

$$f(\sigma(k)) = f(k+1) = s(f(k)).$$

The claim is that f is a *bijection*. (This suffices since f has sent the special zero element 0 to e and preserves the successor operations σ, s.)

$\mathrm{ran}(f) = N$: because $\mathrm{ran}(f)$ satisfies (d) of Dedekind System axioms;

$\mathrm{dom}(f) = \omega$: because $\mathrm{dom}(f)$ likewise satisfies the same DS(d).

f is (1-1): let $X =_{df} \{n \in \omega \mid \forall m(m \neq n \longrightarrow f(m) \neq f(n)\}$. We shall show X is inductive and so is all of ω. By DS(c) $0 \in X$ (because $f(0) = e \neq s(u)$ for any $u \in N$, so if $m \neq 0$, $m = m^- + 1$ say, and so $f(m) = s(f(m^-)) \in \mathrm{ran}(s)$ and $s(f(m^-)) \neq e = f(0)$.) Suppose now $n \in X$. But now assume we have m with $f(m) = u =_{df} f(n+1) \in N$ (and we show that $m = n+1$), then for the same reason, namely $e \notin \mathrm{ran}(s)$ and so $u = s(f(n)) \neq e$, we have $m \neq 0$. So $m = m^- + 1$ for some m^-, and then we know $f(m) = s(f(m^-))$. But by assumption on m and definition of f: $f(m) = f(n+1) = s(f(n))$. We thus have shown $s(f(n)) = s(f(m^-))$; s is (1-1) so $f(m^-) = f(n)$. But $n \in X$ so $m^- = n$. So $m = n+1$. Hence $n + 1 \in X$. Thus X is inductive, which expresses that f is (1-1). Q.E.D.

Example 2.19 *Let $s(k) = k + 2$, let E be the set of positive even natural numbers. Then $\langle E, s, 2\rangle$ is a Dedekind system.*

Exercise 2.7 (i) Let $h : \omega \to \omega$ be given by: $h(0) = 4$ and $h(n+1) = 3 \cdot h(n)$. Compute $h(4)$.

(ii) Let $h : \omega \to \omega$ be given by $h(n) = 5 \cdot n + 2$. Express $h(n+1)$ in terms of $h(n)$ as simply as possible.

Exercise 2.8 Assume f_1 and f_2 are functions from ω to A, and that G is a function on sets, so that for every n $f_1 \restriction n$ and $f_2 \restriction n$ are in $\mathrm{dom}(G)$. Suppose also f_1 and f_2 have the property that

$f_1(n) = G(f_1 \restriction n)$ and $f_2(n) = G(f_2 \restriction n)$. Show that $f_1 = f_2$.

Exercise 2.9 Let $h : \omega \to \omega$ be given by: $h(k) = k - 10$ if k >100; and $h(k) = h(h(k+1))$ if $k \leq 100$.

Give a definition of h if possible, using the standard formulation of a definition by recursion, which involves only computing values $h(k)$from smaller values, or constants. If this is impossible show it so.

Exercise 2.10 Find (i) infinitely many functions $h : \omega \to \omega$ satisfying: $h(k) = h(k+1)$; (ii) the unique function $h : \omega \to \omega$ satisfying: (a) $h(0) = 2; h(k) = h(k+1)(h(k+1)+1)$ if $k > 0$.

Exercise 2.11 Prove that for any $n, m \in \omega$ that $n + m = 0 \leftrightarrow (n = 0 \wedge m = 0)$.

Exercise 2.12 Prove that for any $n, m, k \in \omega$ (i) $n < m \to n + k < m + k$; (ii) $k > 0 \wedge n < m \to n \cdot k < m \cdot k$.

Exercise 2.13 Prove that for any $n, m \in \omega$ that if $n \leq m$ then there is a unique $k \in \omega$ with $n + k = m$.

Exercise 2.14 $(*)$(The Ackermann function) Define using the equations the *Ackermann function:*

$A(0, x, y) = x \cdot y$

$A(k+1, x, 0) = 1$

$A(k+1, x, y+1) = A(k, A(k+1, x, y), x)$

Show that $A(k, x, y)$ is defined for all x, y, k. [Hint: Use a *double induction:* first on k assume that for all x, y $A(k, x, y)$ is defined; then assume for all $y' < yA(k+1, x, y')$ is defined.] What is $A(1, x, y)$?

Wellorderings and ordinals

In this chapter we study what was perhaps Cantor's main mathematical contribution: the theory of wellorder. He generalized the key fact about the natural numbers to allow for wellorderings on infinite sets of different type than that of $\mathbb{N}$. He noted that such wellorderings fell into equivalence classes, where all wellorderings in an equivalence class were order isomorphic. Thus each infinite wellordered set had a unique "order type". These order types could be treated like numbers and added, multiplied *etc.* A new kind of number had been invented. Later Zermelo, and then von Neumann, picked out sets to represent these new 'transfinite' numbers.

It is possible to wellorder an infinite set in many ways.

Example 3.1 *Define $\prec$ on $\mathbb{N}$ by:*

$$n \prec m \Longleftrightarrow (n \text{ is even and } m \text{ is odd}) \vee (n, m \text{ are both even or both odd, and } n < m).$$

Then $\langle \mathbb{N}, \prec \rangle$ is a wellordering.

Exercise 3.1 Let $<$ be the usual ordering on $\mathbb{N}^+ =_{df} \{n \in \omega \mid n \neq 0\}$. For $n \in \mathbb{N}^+$define $f(n)$ to be the number of distinct prime factors of n. Define a binary relation $mRn \Leftrightarrow f(m) < f(n) \vee (f(m) = f(n) \wedge m < n)$. Show that R is in fact a wellordering of $\mathbb{N}^+$. Draw a picture of it.

Example 3.2 *If $\langle A, \prec \rangle$ is a set with a wellordering and $B \subseteq A$ then $\langle B, \prec \rangle$ is also a wellordering. Note that if $y \in A$ is any element that has $\prec$-successors then it has a unique successor, namely*

$$\inf\{x \in A \mid y \prec x\}.$$

Convention*: Note that we shall use, as here, the ordering $\prec$ for B although originally it was given for A. That is, we shall not bother with writing $\langle B, \prec \cap B \times B \rangle$ but simply $\langle B, \prec \rangle$.*

Exercise 3.2 Show that $\langle A, \prec \rangle \in \mathrm{WO}$ implies there is no set $\{x_n \in A \mid n \in \omega\}$ with $\forall n(x_{n+1} \prec x_n)$. (Is there a reason one might hesitate to replace the 'implies' by '$\longleftrightarrow$' here?)

Theorem 3.3 **(Principle of Transfinite Induction)** *Let $\langle X, \prec \rangle \in \mathrm{WO}$. Then*

$$[\forall z \in X\,(\,(\forall y \prec z\Phi(y)) \to \Phi(z)\,)] \to \forall z \in X\Phi(z).$$

Proof: Suppose the antecedent holds but $\varnothing \neq Z =_{df} \{w \in X \mid \neg\Phi(w)\}$. As $\langle X, \prec \rangle \in \mathrm{WO}$ there is a $\prec$-least element $w_0 \in Z$. But then $\forall y \prec w_0\Phi(y)$. So $\Phi(w_0)$ by the antecedent. Contradiction! So $Z = \varnothing$. Q.E.D.

Definition 3.4 *If $\langle X, \prec\rangle \in$ WO then the $\prec$-initial segment X_z (or just "(initial) segment") determined by some $z \in X$ is the set of all predecessors of z: $X_z =_{df} \{u \in X \mid u \prec z\}$.*

In Example 3.1, $\mathbb{N}_1$ is the set of evens, $\mathbb{N}_4 = \{0, 2\}$. We now prove some basic facts about any wellordering.

Exercise 3.3 Show that if $\langle X, \prec\rangle$ is a total ordering, then
$\langle X, \prec\rangle \in \mathrm{WO} \quad \Leftrightarrow \quad \forall u \in X \forall Z \subseteq X_u$ (if $Z \neq \varnothing$, then Z has a $\prec$-least element).
[Thus it suffices for a total order to be a wellorder, if its restrictions to all its proper initial segments are wellorders.]

Recall the definition of *(order) isomorphism*.

Lemma 3.5 *If $f : \langle X, \prec\rangle \to \langle X, \prec\rangle$ is any order preserving map of $\langle X, \prec\rangle \in$ WO into itself, then $\forall z \in X(z \preceq f(z))$. (NB f is not necessarily an isomorphism.)*

Proof: As $\langle X, \prec\rangle$ is a wellordering, if for some z we had $f(z) \prec z$, then, there is a least element z_0 with the property. Then as f is order preserving, we should have $f(f(z_0)) \prec f(z_0) \prec z_0$ thereby contradicting the $\prec$-leastness of z_0. Q.E.D.

Note: this fails if $\langle X, \prec\rangle \notin$ WO: $f : \langle \mathbb{Z}, <\rangle \to \langle \mathbb{Z}, <\rangle$ defined by $f(k) = k-1$ is an order isomorphism.

Lemma 3.6 *If $f : \langle X, \prec\rangle \to \langle Y, \prec'\rangle$ is an order isomorphism with $\langle X, \prec\rangle, \langle Y, \prec'\rangle \in$ WO, then f is unique.*

Note: again this fails for general total orderings: $f' : \langle \mathbb{Z}, <\rangle \to \langle \mathbb{Z}, <\rangle$ is also an order isomorphism where $f'(k) = k - 2$.

Proof: Suppose $f, g : \langle X, \prec\rangle \to \langle Y, \prec'\rangle$ are two order isomorphisms. Then $h =_{df} f^{-1} \circ g : \langle X, \prec\rangle \to \langle X, \prec\rangle$ is also an order isomorphism. By Lemma 3.5 $x \preceq h(x)$ for any $x \in X$. But f is order preserving, so $f(x) \preceq' f(h(x)) = g(x)$. Applying the same argument with $h^{-1} = g^{-1} \circ f$ we get $g(x) \preceq' f(x)$. Hence $f(x) = g(x)$ for any arbitrary $x \in X$. Q.E.D.

Corollary 3.7 *If $\langle X, \prec\rangle \in$ WO and $f : \langle X, \prec\rangle \to \langle X, \prec\rangle$ is an isomorphism then $f = \mathrm{id}$.*

Proof: Since $\mathrm{id} : \langle X, \prec\rangle \to \langle X, \prec\rangle$ is trivially an isomorphism this follows from the last lemma. Q.E.D.

Exercise 3.4 Let $f : \langle X, \prec\rangle \to \langle Y, \prec'\rangle$ be an order isomorphism with $\langle X, \prec\rangle, \langle Y, \prec'\rangle \in$ WO as in the last Lemma 3.6. Show that for any $z \in X$, $f \upharpoonright X_z : \langle X_z, \prec\rangle \cong \langle Y_{f(z)}, \prec'\rangle$.

Lemma 3.8 **(Cantor 1897)** *A wellordered set is not order isomorphic to any segment of itself.*

Proof: If $f : \langle X, \prec\rangle \to \langle X_z, \prec\rangle$ is an order isomorphism then by 3.5 we have $x \preceq f(x)$ for any x, and in particular $z \preceq f(z)$. But $f(z) \in X_z$! In other words $z \preceq f(z) \prec z$! Contradiction! Q.E.D.

Lemma 3.9 *Any wellordered set $\langle X, \prec\rangle$ is order isomorphic to the set of its segments ordered by $\subset$ (recall $\subset$ means* proper subset: $\subsetneqq$).

Proof: Let $Y = \{X_a \mid a \in X\}$. Then $a \rightarrowtail X_a$ is a (1-1) mapping onto Y the set of segments, and since $a \prec b \Longleftrightarrow X_a \subset X_b$ the mapping is order preserving. Q.E.D.

Exercise 3.5 Find an example of two totally ordered sets which are not order isomorphic, although each is order isomorphic to a subset of the other. [Hint: consider subsets of $\mathbb{Q}$ with the usual order.]

Exercise 3.6 Suppose $\langle X, \prec_1 \rangle$ and $\langle Y, \prec_2 \rangle$ are wellorderings. Show that $\langle X \times Y, \prec_{\text{lex}} \rangle \in \text{WO}$ where we define $\langle u, v \rangle \prec_{\text{lex}} \langle t, w \rangle$ if $u \prec_1 t \vee (u = t \wedge v \prec_2 w)$.

3.1 Ordinal numbers

We can now introduce ordinal numbers. Recall that we generated the sequence of sets

$$\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} \ldots$$

calling these successively $0, 1, 2, 3, \ldots$where each is the set of its predecessors: each member is the set of all those sets that have gone before. We shall call such wellordered sets with this property "ordinal numbers" (or more plainly "ordinals"). We thus have seen already some examples: any natural number is an ordinal, as is ω. We first define ordinal through another property that $\langle \omega, < \rangle$ had.

Definition 3.10 *$\langle X, \in \rangle$ is an* ordinal *iff X is transitive and setting $\prec = \in$, then $\langle X, \prec \rangle$ is a wellorder of X. (In which case we also set $u \preceq v \leftrightarrow u = v \vee u \in v$, for $u, v \in X$.)*

Example 3.11 *$\langle \omega, \in \rangle$ is an ordinal, and we had* $3 = \{k \in \omega \mid k \in 3\} = \{0, 1, 2\} = (\omega)_3$.

Lemma 3.12 *$\langle X, \in \rangle$ is an ordinal implies that every element $z \in X$ is identical with the $\in$-initial segment X_z* i.e. $z = X_z = \{w \in X \mid w \in z\}$

Proof: Suppose X is transitive and $\in$ wellorders X. Let $z \in X$. Then $w \in X_z \Longleftrightarrow w \in X \wedge w \in z \Leftrightarrow w \in z$ (the last equivalence holds as $z \subseteq X$). Hence $z = X_z$. Q.E.D.

So what we are doing in defining "ordinals" is generalising what we saw obtained for the von Neumann natural numbers: that each was the set of its predecessors in the ordering $<$ that was also defined as $\in$. Since the ordering on an ordinal is always $\in$ we can drop this and simply talk about a set X being an ordinal. Note that it is somehow more natural to talk about strict total orderings when using $\in$ as the ordering relation.

We shall see that we can have many infinite ordinals. Note that if $\langle X, \in \rangle$ is an ordinal then, as $a = X_a$ for any $a \in X$ (by the last lemma), and for any other $b \in X$, we have that $a \in b \Leftrightarrow a \subsetneqq b \Leftrightarrow X_a \subsetneqq X_b$. Hence *for ordinals*, the ordering $\prec$ is also nothing other than $\subsetneqq = \subset$ *restricted to the elements* of X.

Lemma 3.13 *Any $\in$-initial segment of an ordinal $\langle X, \in \rangle$ is itself an ordinal.*

Proof: Suppose w is an element of the segment X_u. Then as $\in$ totally orders X, $t \in w \in u \to t \in u = X_u$. Hence $\text{Trans}(X_u)$. Since $\in$ wellorders X and $X_u \subseteq X$, $\in$ wellorders X_u. Hence the latter is an ordinal. Q.E.D.

Lemma 3.14 *If $Y \subset X$ is a proper subset of the ordinal X, and Y is itself an ordinal, then Y is an $\in$- initial segment of X.*

Proof: Let Y be an ordinal which is a proper subset of the ordinal X. If $a \in Y$, then as Y is an ordinal (by 3.12) $a = Y_a$, and similarly, as $a \in X$, $a = X_a$.Then $X_a = Y_a$. As Y is not all of X, then if we set $c = \inf\{z \in X \mid z \notin Y\}$, ($c$ exists as an element of X as X is wellfounded) then we have that $Y = X_c$.
Q.E.D.

Lemma 3.15 *If X, Y are ordinals, so is $X \cap Y$.*

Proof: As X, Y are transitive, so is $X \cap Y$. As $\in$ wellorders X, it wellorders $X \cap Y$, and hence the latter is an ordinal. Q.E.D.

Exercise 3.7 Show that if $\langle X, \in\rangle$ is an ordinal, then so is $\langle S(X), \in\rangle$ (where $S(X) = X \cup \{X\}$).

Theorem 3.16 **(Classification Theorem for Ordinals)** *Given two ordinals X, Y either $X = Y$ or one is an initial segment of the other (or, equivalently, one is a member of the other).*

Proof: Suppose $X \neq Y$. By the last lemma $X \cap Y$ is an ordinal. Then

Either (i) $X = X \cap Y$ or (ii) $Y = X \cap Y$ and since $X \neq Y$ (in case (i)), $X \cap Y$ is an initial segment of Y by Lemma 3.13, or (in case (ii)), using the same Lemma, an initial segment of X;

Or $X \cap Y$ is an ordinal properly contained in both X and Y. We show this is impossible. By Lemma 3.13 $X \cap Y$ is simultaneously a segment X_a say of X, and a segment Y_b say of Y for some $a \in X$ and $b \in Y$. But $a = X_a = Y_b = b$ in that case. Hence $a = b \in X \cap Y = X_a$. But then $a \in X_a$ which is absurd! Q.E.D.

Lemma 3.17 *For any two ordinals X, Y, if X and Y are order isomorphic then $X = Y$.*

Proof: Suppose $X \neq Y$. Then by the last theorem X is an initial segment of Y (or *vice versa*). However, if we had that X and Y were order isomorphic, then we should have that the wellordered set $\langle Y, \in\rangle$ isomorphic to an inital segment of itself. This is impossible by Lemma 3.8. Q.E.D.

Corollary 3.18 *If $\langle A, \prec\rangle \in$ WO then it can be isomorphic to* at most one *ordinal set.*

(Check!) We shall show that it will be so isomorphic to *at least one* ordinal. We first give an argument for what will be the inductive step in the argument to follow.

Lemma 3.19 *If every segment of a wellordered set $\langle A, \prec\rangle$ is order isomorphic to some ordinal, then $\langle A, \prec\rangle$ is itself order isomorphic to an ordinal.*

Proof: By the last Corollary we can define a function F which assigns to each element $b \in A$, a unique ordinal $F(b)$ so that $\langle A_b, \prec\rangle \cong \langle F(b), \in\rangle$. Let $Z = \text{ran}(F)$.[1] So

$$Z = \{F(b) \mid \exists b \in A \exists g_b (g_b : \langle A_b, \prec\rangle \cong \langle F(b), \in\rangle)\}.$$

(Note that for each b there can be only one such g_b by Lemma 3.6.) Now notice that if $c \prec b$, with $c, b \in A$ then $A_c = (A_b)_c$. Hence we can not have $F(c) = F(b)$, as this would imply that $g_c^{-1} \circ g_b$ would be an order

[1] Why does this set Z exist? We shall discuss later the *Axiom of Replacement* that justifies this.

isomorphism between A_b and its initial segment A_c, contradicting Lemma 3.8. Thus F is (1-1) and so a bijection between A and Z. We should have that F is an order isomorphism, *i.e.* that $F : \langle A, \prec \rangle \cong \langle Z, \in \rangle$, if it is order preserving which will be (1) below. If still $c \prec b$ then $g_b \upharpoonright A_c : \langle A_c, \prec \rangle \cong \langle (F(b))_{g_b(c)}, \in \rangle$ (by an application of Ex.3.4). So, again by uniqueness of the isomorphism of $\langle A_c, \prec \rangle$ with an ordinal, g_c is $g_b \upharpoonright A_c$ and $F(c)$ must be $(F(b))_{g_b(c)}$. Thus writing these facts out we have that

$$c \prec b \implies F(c) = (F(b))_{g_b(c)} \in F(b) \qquad (1)$$

(The latter $\in$ by Lemma 3.12.) We'd be done if we knew $\langle Z, \in \rangle$ was an ordinal. This is the case: because F is an isomorphism Z is wellordered by $\in$. All we have to check is that $\mathrm{Trans}(Z)$. But this is easy: let $u \in F(b) \in Z$ be arbitrary. As g_b is onto $F(b)$, $u = g_b(c)$ for some $c \prec b$. Then $u = F(b)_u = F(b)_{g_b(c)} = F(c)$ (the first equality holds as $F(b)$ is an ordinal, the last holds by (1) above). Hence $u \in Z$. Thus $\mathrm{Trans}(Z)$.
Q.E.D.

Theorem 3.20 **(Representation Theorem for Wellorderings, Mirimanoff 1917)** *Every wellordering $\langle X, \prec \rangle$ is order isomorphic to one and only one ordinal.*

Proof: Uniqueness follows from the Corollary 3.18. Existence will follow from the last lemma: the wellordering $\langle X, \prec \rangle$ will be order isomorphic to an ordinal, if all its initial segments are. Suppose

$$Z =_{df} \{v \in X \mid X_v \text{ is not isomorphic to an ordinal}\}.$$

If $Z = \varnothing$ then by the last Lemma we have achieved our task. Otherwise if v_0 is the $\prec$-least element of Z then $\langle X_{v_0}, \prec \rangle$ is a wellordering all of whose initial segments $(X_{v_0})_w = X_w$ for $w \prec v_0$, are isomorphic to ordinals (as such $w \notin Z$). But by the last lemma then, $\langle X_{v_0}, \prec \rangle$ is isomorphic to an ordinal. But then $v_0 \notin Z$! Contradiction! So $Z = \varnothing$. Q.E.D.

Definition 3.21 *If* $\langle X, \prec \rangle \in \mathrm{WO}$ *then the* order type *of* $\langle X, \prec \rangle$ *is the unique ordinal order isomorphic to it. We write it as* $\mathrm{ot}(\langle X, \prec \rangle)$.

Corollary 3.22 **(Classification Theorem for Wellorderings, Cantor 1897)** *Given two wellorderings $\langle A, \prec \rangle$ and $\langle B, \prec' \rangle$ exactly one of the following holds:*

(i) $\langle A, \prec \rangle \cong \langle B, \prec' \rangle$

(ii) $\exists b \in B\ \langle A, \prec \rangle \cong \langle B_b, \prec' \rangle$

(iii) $\exists a \in A \langle A_a, \prec \rangle \cong \langle B, \prec' \rangle$.

Proof: If $\langle X, \in \rangle$ and $\langle Y, \in \rangle$ are the unique ordinals isomorphic to $\langle A, \prec \rangle$, $\langle B, \prec' \rangle$ respectively, then by Theorem 3.16, either $\langle X, \in \rangle = \langle Y, \in \rangle$ (in which case (i) holds); or $\langle X, \in \rangle$ is isomorphic to an initial segment of $\langle Y, \in \rangle$ (in which case we have (ii)), or *vice versa,* and we have (iii). Q.E.D.

Definition 3.23 *Let* On *denote the class of ordinals.*

For $\alpha, \beta \in \mathrm{On}$, *we write* $\alpha < \beta =_{df} \alpha \in \beta$. $\alpha \leq \beta =_{df} \alpha < \beta \vee \alpha = \beta$.

We shall summarise below some of the basic properties of ordinals. In the sequel, as in the last definition we follow the convention of using lower case greek letters to implicitly denote ordinals.

3.2 Properties of Ordinals

We collect together:

Basic properties of ordinals: Let $\alpha, \beta, \gamma \in \mathrm{On}$.

(1) α is a transitive set, $\mathrm{Trans}(\alpha)$; $\in$ wellorders α.

(2) $\alpha \in \beta \in \gamma \to \alpha \in \gamma$.

(3) $X \in \alpha \to X \in \mathrm{On} \wedge X = \alpha_X$.

(4) $\langle \alpha, \in \rangle \cong \langle \beta, \in \rangle \to \alpha = \beta$.

(5) Exactly one of (i) $\alpha = \beta$, (ii) $\alpha \in \beta$, (iii) $\beta \in \alpha$ holds.

(1) here is Def. 3.9; (2) holds by $\mathrm{Trans}(\gamma)$; (3) is 3.12 and 3.13, and (4) is 3.17. (5) follows from 3.12 and 3.16.

Lemma 3.24 (6) **Principle of Transfinite Induction for** On *Let Φ be a well defined and definite property of ordinals.*

$$\forall \alpha \in On[(\forall \beta < \alpha \Phi(\beta)) \longrightarrow \Phi(\alpha)] \quad \longrightarrow \quad \forall \alpha \in On \Phi(\alpha)$$

Hence we have a Least Ordinal Principle *for classes:*

$$\textit{If } C \neq \varnothing, C \subseteq \mathrm{On} \textit{ then } \exists \alpha \in C \forall \beta \in C[\alpha \leq \beta].$$

Hence On is itself well-ordered.

Proof of (6): The proof of the first statement concerning Φ is exactly like, and can be considered a special case of, the Principle of Transfinite Induction Theorem 3.3. Suppose the conclusion is false and $C = \{\alpha \mid \neg\Phi(\alpha)\}$. Then reason as follows. Let $\alpha_0 \in C$ as C is assumed non-empty. If for no $\beta \in C$ do we have $\beta < \alpha_0$ then α_0 was the $\in$-minimal element of C. Otherwise we have that $C \cap \alpha_0 \neq \varnothing$. As $\alpha_0 \in \mathrm{On}$, by definition $\in$ wellorders α_0. Hence, as $C \cap \alpha_0 \subseteq \alpha_0$ is non-empty, it has an $\in$-minimal element α_1; and then α_1 is the minimal element of C. For the last sentence, we know that On is totally ordered by (5); (6) then says $<$ (or $\in$) wellorders On. Q.E.D.

Note: This last argument seems a little unnecessary, but it is not: we know any individual ordinal is wellordered: (6) implies the whole class On is wellordered. Note also that we did not require C to be a set, it could be a proper class.

The following was originally noted as a "paradox" by Burali-Forti. This was the first of the set theoretical paradoxes to appear in print. Burali-Forti noted (as in the argument below) that On itself formed a transitive class of objects well-ordered by $\in$. Hence, as On consists of *all* such transitive classes, $(\mathrm{On}, \in)$ is isomorphic to a member of itself! A plain contradiction! The reaction to this contradiction was messy: Burali-Forti thought he had shown that the class of ordinals was merely partially ordered. Russell thought that the class of ordinals was linearly ordered only (although two years later he saw the need for the distinction between sets and classes, and reasoned that On had to be a proper class, but was indeed wellordered). Again we must distinguish between sets as objects of study, and proper classes as collections of sets brought together by an arbitrary description. Burali-Forti's argument when properly dressed in its modern clothes is the following.

LEMMA 3.25 **(Burali-Forti 1897)** On *is a proper class.*

Proof. Suppose for a contradiction x is a set and $x = \text{On}$. Then as we have seen (Lemma 3.24) we can wellorder x by the ordering $\in$ on On. But then $\langle x, \in\rangle$ is itself a wellordering and furthermore $\text{Trans}(x)$. Hence $x \in \text{On}$. But then $x \in x$, and x becomes an ordinal that is a member of itself. This is nonsense as $\in$, is a *strict* ordering, and so is irreflexive, on any ordinal! QED

DEFINITION 3.26 *Let $\langle A, R\rangle, \langle B, S\rangle$ be total orderings, with $A \cap B = \varnothing$. We define the* sum *of* $\langle A, R\rangle, \langle B, S\rangle$ *to be the ordering $\langle C, T\rangle$ where $C = A \cup B$ and we set*

$$xTy \longleftrightarrow (x \in A \wedge y \in B) \vee (x, y \in A \wedge xRy) \vee (x, y \in B \wedge xSy)$$

The picture here is that we take a copy of $\langle A, R\rangle$ and place all of it *before* a copy of $\langle B, S\rangle$.

EXERCISE 3.8 Show that if $\langle A, R\rangle, \langle B, S\rangle \in \text{WO}$, then the sum $\langle C, T\rangle \in \text{WO}$.

Note that the definition required that A, B be disjoint (so that the orderings did not become "confused". We should like to use ordinals themselves for A, B but they are not disjoint. Hence it is convenient to use a simple "disjointing device" as follows. If $\alpha, \beta \in \text{On}$, then $\alpha \times \{0\}$ and $\beta \times \{1\}$ are disjoint "copies" of α and β. We could now define the "sum" of α and β as

$\alpha +' \beta =_{df} \text{ot}(\langle \alpha \times \{0\} \cup \beta \times \{1\}, T\rangle$ where $\langle \gamma, i\rangle T \langle \delta, j\rangle \leftrightarrow (i = j \wedge \gamma < \delta) \vee i < j.$

The operation $+'$ is pretty clearly associative, but it is not commutative as the following examples will show.

EXAMPLE 3.27 $2 +' 3; 2 +' \omega; \omega +' 2; \omega +' \omega; (\omega +' \omega) + 2; (\omega +' \omega) +' \omega \ldots$

$\sup\{\omega, \omega +' \omega, (\omega +' \omega) +' \omega \ldots\} = \omega.'\omega = \sup\{\omega.'n \mid n \in \omega\}.$

DEFINITION 3.28 *Let $\langle A, R\rangle, \langle B, S\rangle$ be total orderings. We define the* product *of $\langle A, R\rangle, \langle B, S\rangle$, $\langle A, R\rangle \times \langle B, S\rangle$, to be the ordering $\langle C, U\rangle$ = where $C = A \times B$ and we set U to be the* anti-lexicographic *ordering on C:*

$$\langle x, y\rangle U \langle x', y'\rangle \longleftrightarrow (ySy') \vee (y = y' \wedge xRx').$$

This is different: here we imagine taking a copy of $\langle B, S\rangle$ and *replacing* each element $y \in B$ with a copy of all of $\langle A, R\rangle$.

EXERCISE 3.9 Show that if $\langle A, R\rangle, \langle B, S\rangle \in \text{WO}$, then the product $\langle C, U\rangle = \langle A, R\rangle \times \langle B, S\rangle \in \text{WO}$.

EXERCISE 3.10 Suppressing the usual ordering $<$ on the following sets of numbers, show that in the product orderings: $\mathbb{Z} \times \mathbb{N} \ncong \mathbb{Z} \times \mathbb{Z}$. Is $\mathbb{N} \times \mathbb{Z} \cong \mathbb{Z} \times \mathbb{Z}$?. Is $\mathbb{Q} \times \mathbb{Z} \cong \mathbb{Q} \times \mathbb{N}$?

Again we could define ordinal products $\alpha \cdot' \beta$ by setting $\alpha \cdot' \beta$ to be:

$$\alpha \cdot' \beta =_{df} \text{ot}(\langle \alpha \times \beta, U\rangle) \text{ where } \langle \gamma, \delta\rangle U \langle \gamma', \delta'\rangle \longleftrightarrow (\delta < \delta') \vee (\delta = \delta' \wedge \gamma < \gamma').$$

Again $\cdot'$ will turn out to be associative (after some thought) but non-commutative.

EXAMPLE 3.29 $2 \cdot' 3; 2 \cdot' \omega; \omega \cdot' 2; \omega \cdot' \omega; (\omega \cdot' \omega) \cdot' 2; (\omega \cdot' \omega) \cdot' \omega \ldots$

Exercise 3.11 (i) Express $(\omega +' \omega) +' \omega$ using the multiplication symbol $.'$ only.
(ii) Informally express $\omega \cdot' \omega$ using the addition symbol $+'$ only.

Exercise 3.12 Show that the distributive law $(\alpha +' \beta) \cdot' \gamma = \alpha .' \gamma +' \beta .' \gamma$ is not valid. On the other hand, convince yourself that $\alpha \cdot' (\beta +' \gamma) = \alpha \cdot' \beta +' \alpha \cdot' \gamma$ will be true.

The reason we have put primes above our arithmetical operations is that we shall soon define them in another way, extending our everyday definition of $+, \cdot$ for natural numbers.

Definition 3.30 *For A a set of ordinals,* $\sup A$ *is the least ordinal $\gamma \in \mathrm{On}$ so that $\forall \delta \in A(\delta \leq \gamma)$. The* strict sup *of A,* $\sup^+ A$, $\sup A$ *is the least ordinal $\gamma \in \mathrm{On}$ so that $\forall \delta \in A(\delta < \gamma)$.*

This conforms entirely to our notion of supremum as the lub of a set. In particular:
(i) If A has a largest element μ then $\sup A = \mu$.
(ii) Suppose $A \neq \varnothing$ has no largest element; then $\sup A$ is the smallest ordinal strictly greater than all those in A.
(iii) For A any set of ordinals check that $\sup^+ A = \sup\{\delta + 1 \mid \delta \in A\}$.

Example 3.31 *$\sup 3 = 2 = \sup \{0,2\}$;* $\sup\{3\} = 3; \sup\{\mathrm{Evens}\} = \omega = \sup \omega = \sup\{\omega\}$*;*
$\sup\{0, 3, \omega + 1\} = \omega + 1$. *But* $\sup^+ 3 = 3 = \sup^+ \{0,2\}$*;* $\sup^+ \{3\} = 4; \sup^+ \{\mathrm{Evens}\} = \omega = \sup^+ \omega \neq \sup^+ \{\omega\} = \omega + 1$.

Many texts simply define $\sup(A)$ as $\bigcup A$. This makes sense:

Lemma 3.32 *Let A be a set of ordinals then* $\sup A$ *is properly defined, and equals $\bigcup A$.*

Proof: First note that $\sup(A)$ is properly defined: there is an ordinal which is an upper bound for A. Suppose not, then we have that for every $\gamma \in \mathrm{On}$ there is $\delta \in A$ with $\gamma < \delta$. By the axiom of union: as A is assumed to be a set, so is $\bigcup A$. But $\mathrm{On} = \bigcup A$! This contradicts Lemma 3.25. Hence A has an upper bound, and $\sup(A)$ exists.
Claim: $\sup A = \bigcup A$.
Proof: Let $\gamma = \sup A$. Suppose $\delta \in \bigcup A$. Then for some $\tau \in A$ we have: $\delta \in \tau \in A$. So $\delta < \gamma$ and so $\delta \in \gamma$. Hence $\bigcup A \subseteq \gamma$. Conversely suppose $\delta \in \gamma$. Then $\delta < \gamma = \sup A$ and so there is $\mu \in A$ with $\delta < \mu \leq \gamma$. Hence $\delta \in \mu \in A$ and so $\delta \in \bigcup A$. Thus $\gamma \subseteq \bigcup A$. Q.E.D.

Observe also that if $X \subseteq Y$ are sets of ordinals, then by definition, $\sup X \leq \sup Y$.

Definition 3.33 $\mathrm{Succ}(\alpha) \Leftrightarrow \exists \beta(\alpha = S(\beta))$.
We write $\beta + 1$ for $S(\beta) = \beta \cup \{\beta\}$. .

$$\mathrm{Lim}(\alpha) \Leftrightarrow \alpha \in \mathrm{On} \wedge \alpha \neq 0 \wedge \neg \mathrm{Succ}(\alpha).$$

We thus have ordinals are divided into three types: (i) 0; (ii) those of the form $\beta+1$, *i.e.* those that have an immediate predecessor, and (iii) the rest, the "limit ordinals" which have no immediate predecessors. Notice we have written $S(\beta)$ as '$\beta+1$', that is because we shall define our official '+1' operation to coincide with S (see Lemma 3.39 below) as we did for natural number addition. So we are getting slightly ahead of ourselves. Note that if $A \neq \varnothing$ has no largest element; then $\sup A$ is a limit ordinal.

Example 3.34 *Successors are: 2,* $n, \omega+1, (\omega+1)+1, \ldots$ *Limits:* ω *is the first limit ordinal; the next will be* $\omega+\omega$*, then* $(\omega+\omega)+\omega; \ldots \omega\cdot\omega, \ldots$ *when we come to define these arithmetic operations, which we shall now turn to.*

Exercise 3.13 (i) Compute $\sup(\beta+1)$ and verify that it equals $\bigcup(\beta+1)$. Suppose $0 < \lambda \in On$. Show that λ is a limit ordinal iff $\lambda = \bigcup\lambda$ iff $\sup\lambda = \lambda$. (ii) Prove that if X is a transitive set of ordinals, then X is an ordinal.

Exercise 3.14 Suppose that X, Y are two sets of ordinals, so that for every $\xi \in X$ there is $\upsilon \in Y$ with $\xi \le \upsilon$, and conversely that for every $\forall\upsilon \in Y \exists\xi \in X(\upsilon \le \xi)$. Show that $\sup X = \sup Y$. Deduce that if λ, λ' are both limit ordinals, and that $\langle \alpha_\xi \mid \xi < \lambda\rangle$ and $\langle \beta_\zeta \mid \zeta < \lambda'\rangle$ are two increasing sequences of ordinals with the property that $\forall\xi < \lambda\exists\zeta < \lambda'(\alpha_\xi < \beta_\zeta)$ and also that $\forall\zeta < \lambda'\exists\xi < \lambda(\beta_\zeta < \alpha_\xi)$, then $\sup\{\alpha_\xi \mid \xi < \lambda\} = \sup\{\beta_\zeta \mid \zeta < \lambda'\}$.

In order to give our definition of ordinal arithmetic we first prove a Recursion Theorem on ordinals, just as we did for the natural numbers ω. The structure of the proof is exactly the same. We only must take care of the fact that there now are limit ordinals as well as successors.

Theorem 3.35 **(Recursion Theorem on** On; von Neumann 1923**)** *Let* $F : V \to V$ *be any function. Then there exists a unique function* $H : \mathrm{On} \to V$ *so that:*

$$\forall\alpha(\ H(\alpha) = F(H \restriction \alpha)).$$

Proof: The reader should compare this with the proof of the Recursion Theorem on ω. As there we shall define H as a union of *approximations to* H where u is a δ*-approximation* if:

(i) $\mathrm{Func}(u), \mathrm{dom}(u) = \delta$, and (ii) $\forall\alpha < \delta(u(\alpha) = F(u \restriction \alpha))$.

Such a u satisfies the defining clauses of H throughout its domain up to δ. As before we shall combine the pieces u into the required function H. Notice how this works: (i) if $\delta > 0$ then $u(0) = F(u \restriction \varnothing)$, but $u \restriction \varnothing = \varnothing$; hence $u(0) = F(\varnothing)$ for any δ-approximation.

Note: (i) There is a single 1-approximation: it is $v = \{\langle 0, F(0)\rangle\}$. (Again $u = \varnothing$ is a 0-approximation!) (ii) if u is a δ-approximation, then, by the definition above, $u \restriction \gamma$ is a γ-approximation for any $\gamma \le \delta$.

(iii) If u is a δ-approximation, then $u \cup \{\langle\delta, F(u)\rangle\}$ is a $\delta+1$-approximation. So any approximation can be extended one step.

We let

$$B = \{u \mid \exists\delta(u \textit{ is a } \delta\textit{-approximation})\}$$

(1) *If* u *is a* δ*-approximation and* v *a* γ*-approximation, with* $\delta \le \gamma$*, then* $u = v \restriction \delta$*.*

Proof: As usual, look for a point of least difference for a contradiction: suppose τ is least with $u(\tau) \ne v(\tau)$. Then the two functions agree up to τ; *i.e.* $u \restriction \tau = v \restriction \tau$; but then $u(\tau) = F(u \restriction \tau) = F(v \restriction \tau) = v(\tau)$! Contradiction.

The import of (1) is that there can be no disagreement between approximations: they are all compatible. There are two immediate consequences of (1). Firstly, the same argument from (1) will establish:

(2) *(Uniqueness) If* H *exists then it is unique.*

(If H, H' are any two different functions that satisfy the conditions of the theorem, then let τ be the least ordinal with $H(\tau) \ne H'(\tau)$. But then $H \restriction \tau+1, H' \restriction \tau+1$ are two different $\tau+1$-approximations. This contradicts (1).)

Secondly:

(3) If $Lim(\lambda)$ and for all $\alpha < \lambda$, u_α is an α-approximation, then $\bigcup_{\alpha<\lambda} u_\alpha$ is a λ-approximation.

Proof: The union here is of an increasing sequence of sets which are approximating functions which agree on the intersecting parts of their domains. Thus $u = \bigcup_{\alpha<\lambda} u_\alpha$ is a $\bigcup_{\alpha<\lambda} \alpha = \lambda$-approximation, as it ovberys the requirements on forming approximtions.

Finally:

(4) *(Existence). Such an H exists.*

Proof: As any two approximations agree on the common part of their domains, we may sensibly define $H = \bigcup B$. Just as for the proof of recursion on ω:

(i) H is a function.

(ii) $\mathrm{dom}(H) = \mathrm{On}$.

Proof: Let C be the class of ordinals δ for which there is no δ-approximation. So if C is non-empty, by the Principle of Transfinite Induction for On, then it will have a least element ζ. By Note (i) above, $\zeta > 1$. By (3) it cannot be a limit ordinal.

If $\zeta = \mu + 1$ then there is a μ-approximation ν. But by Note (iii) we may extend ν to a $\mu + 1$-approximation u by setting $u(\mu) = F(\nu)$; *i.e.*, set $u = \nu \cup \{\langle \mu, F(\nu)\rangle\}$. Contradiction! Hence $C = \varnothing$.

Q.E.D.

Thus again, $H(\alpha)$ is defined to be that value $u(\alpha)$ given by any δ-approximation u, with $\alpha < \delta$.

Remark 3.36 As we have stated it, we have used proper classes - the function F for example is such, and On being a proper class will entail that H is too. This is not as risky as might be thought at first, since we may eliminate talk of proper classes by their defining formulae *if* we are careful. We have chosen to be a little relaxed about this, for the sake of the exposition.

Remark 3.37 Although this is the common form of the Recursion Theorem for On in text books, it is often more useful in the following form, which tends to "unpack" the function F into two different "subfunctions" and a constant depending on the type of ordinal just occurring in the definition of H. It essentially contains no more than the first theorem: one should think of it as a version of the first theorem where F is *defined by cases*.

Theorem 3.38 **(Recursion Theorem on** On**, Second Form)** *Let $a \in V$. Let $F_0, F_1 : V \to V$ be functions. Then there is a unique function $H : \mathrm{On} \to V$ so that:*

(i) $H(0) = a$;

(ii) If $\mathrm{Succ}(\alpha)$ then $H(\alpha) = F_0(H(\beta))$ where $\alpha = \beta + 1$;

(iii) If $\mathrm{Lim}(\alpha)$ then $H(\alpha) = F_1(H \restriction \alpha)$.

Proof: Define $F : V \to V$ by:

$F(x) = a$ if $x = \varnothing$,

$F(u) = F_0(u)$ if $\mathrm{Func}(u) \wedge \mathrm{dom}(u)$ is a successor ordinal,

$F(u) = F_1(u)$ if $\mathrm{Func}(u) \wedge \mathrm{dom}(u)$ is a limit ordinal,

$F(u) = \varnothing$ in all other cases.

Now apply the previous theorem to the single function F. Q.E.D.

In practise we shall be a little informal as in the following definitions of the ordinal arithmetic operations.

Definition 3.39 *We define by transfinite recursion on On:*
(Ordinal Addition) $A_\alpha(\beta) = \alpha + \beta$:
$A_\alpha(0) = \alpha$;
$A_\alpha(\beta + 1) = S(A_\alpha(\beta)) = A_\alpha(\beta) + 1$;
$A_\alpha(\lambda) = \sup\{A_\alpha(\xi) \mid \xi < \lambda\}$ *if* $\mathrm{Lim}(\lambda)$. *We write* $\alpha + \beta$ *for* $A_\alpha(\beta)$.

(Ordinal Multiplication) $M_\alpha(\beta) = \alpha \cdot \beta$:
$M_\alpha(0) = 0$;
$M_\alpha(\beta + 1) = M_\alpha(\beta) + \alpha$;
$M_\alpha(\lambda) = \sup\{M_\alpha(\xi) \mid \xi < \lambda\}$ *if* $\mathrm{Lim}(\lambda)$. *We write* $\alpha \cdot \beta$ *for* $M_\alpha(\beta)$.

(Ordinal Exponentiation) $E_\alpha(\beta) = \alpha^\beta$ *(for* $\alpha > 0$*).:*
$E_\alpha(0) = 1$;
$E_\alpha(\beta + 1) = E_\alpha(\beta) \cdot \alpha$
$E_\alpha(\lambda) = \sup\{E_\alpha(\xi) \mid \xi < \lambda\}$ *if* $\mathrm{Lim}(\lambda)$. *We write* α^β *for* $E_\alpha(\beta)$.

Compare these definitions with those for the usual arithmetic operations on the natural numbers. Note that definition of multiplication (and exponentiation) assumes that addition (respectively multiplication) has been defined for all α. They are obtained in each case by adding a third clause to cater for limit ordinals. Hence we know immediately that the ordinal arithmetic operations agree with standard ones on ω, the set of natural numbers. Note we have gone straight away to the more informal but usual notation: the second line of the above, $A_\alpha(\beta + 1) = S(A_\alpha(\beta))$, could have been stated as $\alpha + (\beta + 1) = S(\alpha + \beta) = (\alpha + \beta) + 1$ *etc.* Clearly then $\alpha + \beta < \alpha + (\beta + 1)$ for any α, β.

Lemma 3.40 *The functions* A_α *are strictly increasing and hence (1-1). That is, for any* α*:*

$$(*) \quad \beta < \gamma \to \alpha + \beta < \alpha + \gamma.$$

Proof: This is formally a proof by induction on γ, but really given the definition of the arithmetical operation A_α should be (or become) intuitively true. For suppose as an inductive hypothesis that $(*)$ holds for all $\gamma \leq \delta$. Then we show it is true for $\delta + 1$. Let $\beta < \delta + 1$. If $\beta = \delta$ then $\alpha + \delta < \alpha + (\delta + 1)$ by the comment immediately before this lemma. But if $\beta < \delta$ then by IH we know $\alpha + \beta < \alpha + \delta$ and the latter we have just argued is less than $\alpha + (\delta + 1)$.

Now suppose that $(*)$ holds for all $\gamma < \lambda$ for some limit ordinal λ. We show it holds for λ. Suppose $\beta < \lambda$. Note $\beta < \beta + 1 < \lambda$. So $\alpha + \beta < \alpha + (\beta + 1) \leq \sup\{\alpha + \gamma \mid \gamma < \lambda\} = \alpha + \lambda$ (the first $<$ holding by definition of $A_\alpha(\beta + 1)$). Q.E.D.

Lemma 3.41 *Similarly both* M_α, E_α *are also strictly increasing and hence (1-1): suppose* $\alpha, \beta, \gamma \in \mathrm{On}$ *are such that* $\beta < \gamma$. *(i) If* $\alpha > 0$ *then* $\alpha \cdot \beta < \alpha \cdot \gamma$*; (ii) if* $\alpha > 1$ *then* $\alpha^\beta < \alpha^\gamma$.

We shall not bother to do so, but we could prove that these arithmetic operations coincide with those defined earlier in terms of order types of composite orders: for any α, β, $\alpha +' \beta = \alpha + \beta$ and $\alpha \cdot' \beta = \alpha \cdot \beta$; we again emphasise that, as we remarked for the operations $+'$ and $\cdot'$, we do not have commutativity of our official operations: $2 + \omega = \sup\{2 + n \mid n \in \omega\} = \omega \neq \omega + 2$; similarly $2 \cdot \omega = \sup\{2 \cdot n \mid n \in \omega\} = \omega \neq \omega \cdot 2$;

Exercise 3.15 Prove this last lemma.

Exercise 3.16 By applying the last two lemmas, justify the following cancellation laws (and hence deduce that all these implications could be replaced by equivalences).

(a) $\alpha + \beta = \alpha + \gamma \to \beta = \gamma$.
(b) $(0 < \alpha \wedge \alpha \cdot \beta = \alpha \cdot \gamma) \to \beta = \gamma$.
(c) $\alpha^\beta = \alpha^\gamma \to \beta = \gamma$.
(d) $\alpha + \beta < \alpha + \gamma \to \beta < \gamma$.
(e) $(\alpha \cdot \beta < \alpha \cdot \gamma) \to \beta < \gamma$.
(f) $\alpha^\beta < \alpha^\gamma \to \beta < \gamma$.

Exercise 3.17 Show that for any γ and any $\alpha \leq \beta$:

(a) $\alpha + \gamma \leq \beta + \gamma$;
(b) $\alpha \cdot \gamma \leq \beta \cdot \gamma$;
(c) $\alpha^\gamma \leq \beta^\gamma$.

The following lemma gives an alternative way to view the addition and multiplication of ordinals in terms of their set elements.

Lemma 3.42 *Let $\alpha, \beta \in$ On. Then*
(i) $\alpha + \beta = \alpha \cup \{\alpha + \gamma \mid \gamma < \beta\}$;
(ii) $\alpha \cdot \beta = \{\alpha \cdot \xi + \eta \mid \xi < \beta \wedge \eta < \alpha\}$

Proof: (i) By induction on β: if $\beta = 0$ then $\alpha + 0 = \alpha \cup \varnothing = \alpha$. Suppose (i) is true for β. Then
$\alpha + (\beta + 1) = (\alpha + \beta) + 1 = S(\alpha + \beta) = \alpha + \beta \cup \{\alpha + \beta\} =$
$= \alpha \cup \{\alpha + \gamma \mid \gamma < \beta\} \cup \{\alpha + \beta\}$(by Ind Hyp.)
$= \alpha \cup \{\alpha + \gamma \mid \gamma < \beta + 1\}$.

It is thus true for $\beta + 1$.

Now suppose $\mathrm{Lim}(\lambda)$ and that (i) is true for $\beta < \lambda$. Then
$\alpha + \lambda = \sup\{\alpha + \beta \mid \beta < \lambda\}$ (by Def. of +)
$= \bigcup\{\alpha \cup \{\alpha + \gamma \mid \gamma < \beta\} \mid \beta < \lambda\}$ (by Lemma 3.32 and the Ind. Hyp.)
$= \alpha \cup \{\alpha + \gamma \mid \gamma < \lambda\}$
(as $\mathrm{Lim}(\lambda)$ implies that any $\alpha + \gamma$ for $\gamma < \lambda$ is also trivially $\alpha + \gamma$ for $\gamma < \beta$ for a $\beta < \lambda$).

It is thus true for $\mathrm{Lim}(\lambda)$ also.

(ii) Again by induction on β. For $\beta = 0$ then $\alpha \cdot 0 = 0 = \varnothing = \{\alpha \cdot \xi + \eta \mid \xi < 0 \wedge \eta < \alpha\}$. Suppose it is true for β. Then:

$\alpha \cdot (\beta + 1) = \alpha \cdot \beta + \alpha$ (by Def. of Multiplication)
$= \alpha \cdot \beta \cup \{\alpha \cdot \beta + \eta \mid \eta < \alpha\}$ (by (i) of the Lemma)
$= \{\alpha \cdot \xi + \eta \mid \xi < \beta \wedge \eta < \alpha\} \cup \{\alpha \cdot \beta + \eta \mid \eta < \alpha\}$ (by the Inductive Hypothesis)
$= \{\alpha \cdot \xi + \eta \mid \xi < \beta + 1 \wedge \eta < \alpha\}$.

Now suppose $\mathrm{Lim}(\lambda)$ and it is true for $\beta < \lambda$, we ask the reader to complete the proof as an exercise.

Exercise 3.18 Complete the proof of (ii) of the Lemma.

Corollary 3.43 *Suppose $\alpha, \beta \in \mathrm{On}$ and $0 < \alpha \leq \beta$. Then (i) there is a unique ordinal γ so that $\alpha + \gamma = \beta$; (ii) there is a unique pair of ordinals ξ, η so that $\eta < \alpha \wedge \beta = \alpha \cdot \xi + \eta$.*

Proof: (ii) By Lemma 3.41 the function $M_\alpha(\xi)$ is strictly increasing. So $\beta \leq M_\alpha(\beta) < M_\alpha(\beta+1)$ for example. So there must be a least ξ so that $\alpha \cdot \xi \leq \beta < \alpha \cdot (\xi + 1) = \alpha \cdot \xi + \alpha$. By part (i) there is a unique η so that $\beta = \alpha \cdot \xi + \eta$. So at least one pair ξ, η satisfying these requirements exists. Suppose ξ', η' is another. If $\xi = \xi'$ then $\alpha \cdot \xi = \alpha \cdot \xi'$; but then $\beta = \alpha \cdot \xi + \eta = \alpha \cdot \xi + \eta'$. By part (i) $\eta = \eta'$.

However if, say, $\xi < \xi'$ then $\xi + 1 \leq \xi'$ and so

$$\beta = \alpha \cdot \xi + \eta < \alpha \cdot \xi + \alpha = \alpha \cdot (\xi + 1) \leq \alpha \cdot \xi' \leq \alpha \cdot \xi' + \eta' = \beta$$

which is absurd. So this case cannot occur. Q.E.D.

Example: If $\alpha < \omega^2$ then $\alpha = \omega \cdot k + l$ for some $k, l \in \omega$.

Exercise 3.19 Show that if $\alpha < \omega^3$ then there exist unique $n, k, l \in \omega$ with $\alpha = \omega^2 \cdot n + \omega \cdot k + l$.

Exercise 3.20 Say that γ is an *end segment* of β if there is an α so that $\alpha + \gamma = \beta$. (Note that β is an end segment of itself.) Show that any β has at most finitely many end segments.

It is easy to see that $\sup\{\alpha + 2n \mid n \in \omega\} = \alpha + \omega$. This is an elementary example of (i) of the next exercise where we have taken X as the set of even natural numbers.

Exercise 3.21 Let X be a set of ordinals without a largest element. Show

(i) $\alpha + \sup X = \sup\{\alpha + \tau \mid \tau \in X\}$;

(ii) $\alpha \cdot \sup X = \sup\{\alpha \cdot \tau \mid \tau \in X\}$;

(iii) $\alpha^{\sup X} = \sup\{\alpha^\tau \mid \tau \in X\}$.

Lemma 3.44 *The following laws of arithmetic hold for our definitions:*

(a) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$

(b) $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$

(c) $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$

(d) $\alpha^{\beta + \gamma} = \alpha^\beta \cdot \alpha^\gamma$.

Proof: These are all proven by transfinite induction. Again we do (b) as a sample. We perform the induction on γ. For $\gamma = 0$ we have $\alpha \cdot (\beta + 0) = \alpha \cdot \beta + 0 = \alpha \cdot \beta + \alpha \cdot 0$. Suppose it is true for γ. Then $\alpha \cdot (\beta + (\gamma + 1)) = \alpha \cdot ((\beta + \gamma) + 1) = \alpha \cdot (\beta + \gamma) + \alpha = (\alpha \cdot \beta + \alpha \cdot \gamma) + \alpha$

$= \alpha \cdot \beta + (\alpha \cdot \gamma + \alpha) = \alpha \cdot \beta + \alpha \cdot (\gamma + 1)$. So it holds for $\gamma + 1$. Suppose now $\mathrm{Lim}(\gamma)$ and it holds for $\delta < \gamma$.

Then
$$\begin{aligned} \alpha \cdot (\beta + \gamma) &= \alpha \cdot \sup\{\beta + \delta \mid \delta < \gamma\} \\ &= \sup\{\alpha \cdot (\beta + \delta) \mid \delta < \gamma\} \quad \text{(by (ii) of the last Exercise} \\ &= \sup\{\alpha \cdot \beta + \alpha \cdot \delta \mid \delta < \gamma\} \quad \text{(by the Ind. Hyp.)} \\ &= \alpha \cdot \beta + \sup\{\alpha \cdot \delta \mid \delta < \gamma\} \quad \text{(by (i) of the last Exercise} \\ &= \alpha \cdot \beta + \alpha \cdot \gamma \quad \text{(by Def. of Multiplication).} \end{aligned}$$

It is sometimes useful to note that if $\beta < \omega^{\gamma}$, then there is always some $\gamma' < \gamma$ and $k < \omega$ with $\beta < \omega^{\gamma'} \cdot k$: if $\mathrm{Lim}(\gamma)$ then as $\omega^{\gamma} = \sup\{\omega^{\gamma'} \mid \gamma' < \gamma\}$ this is immediate (with $k = 1$). If $\gamma = \gamma' + 1$ then $\omega^{\gamma} = \omega^{\gamma'} \cdot \omega$, and then there is some least $k < \omega$ (possibly 1) with $\beta < \omega^{\gamma'} \cdot k$. One can use this observation without diverging into an argument by cases each time.

Exercise 3.22 Describe subsets of $\mathbb{Q}$ with order types $\omega^2, \omega^{\omega}$, and $\omega^{\omega} + \omega^3 + 17$ under the natural $<$ ordering.

Exercise 3.23 Prove that if $0 < \alpha, \beta$ then: (i) $\alpha + \beta = \beta \leftrightarrow \alpha \cdot \omega \leq \beta$.
(ii) $\alpha + \beta = \beta + \alpha \leftrightarrow \exists \gamma \exists m, n \in \omega\ (\alpha = \gamma \cdot m \wedge \beta = \gamma \cdot n)$.

Exercise 3.24 In each of (i)-(iii) find α and X a set of ordinals without a largest element with the properties
(i) $\sup X + \alpha \neq \sup\{\tau + \alpha \mid \tau \in X\}$;
(ii) $\sup X \cdot \alpha \neq \sup\{\tau \cdot \alpha \mid \tau \in X\}$;
(iii) $(\sup X)^{\alpha} \neq \sup\{\tau^{\alpha} \mid \tau \in X\}$.
[Hint: in each case a simple X can be found with $X = \{\beta_n \mid n < \omega\}$.]

Exercise 3.25 (i) Prove that if $\beta < \gamma$ then $\omega^{\beta} + \omega^{\gamma} = \omega^{\gamma}$. (ii) Prove that if $\alpha < \beta \leq \omega^{\gamma}$ then $\alpha + \beta = \omega^{\gamma}$ iff $\beta = \omega^{\gamma}$. Deduce that if for all $\alpha < \beta$ that $\alpha + \beta = \beta$ then $\beta = \omega^{\gamma}$ for some γ.

Exercise 3.26 Prove that if $\alpha \geq 2$ then $\forall \beta (\alpha \cdot \beta \leq \alpha^{\beta})$.

Exercise 3.27 If $\sigma = \omega^{\tau}$ for some $\tau > 0$, and $\alpha < \sigma$, then show that there are $\delta < \tau$, $k < \omega$, and $\gamma < \omega^{\delta}$ with $\alpha = \omega^{\delta} \cdot k + \gamma$.

Lemma 3.45 **(Cantor's Normal Form Theorem)**. *Let $1 \leq \beta$. Then there exists a unique $k \in \omega$ and unique $\gamma_0, \ldots, \gamma_{k-1}$ with $\gamma_0 > \ldots > \gamma_{k-1}$ and $d_0, \ldots, d_{k-1} \in \omega$ so that:*

$$\beta = \omega^{\gamma_0} \cdot d_0 + \omega^{\gamma_1} \cdot d_1 + \cdots + \omega^{\gamma_{k-1}} \cdot d_{k-1}.$$

The Theorem says that any ordinal $\beta \geq 1$ can be expressed "to base ω". There is nothing special about ω here: if $\alpha \leq \beta$ we could still find finitely many decreasing ordinals γ_i, and $0 < d_i < \alpha$ and have $\beta = \alpha^{\gamma_0} \cdot d_0 + \alpha^{\gamma_1} \cdot d_1 + \cdots + \alpha^{\gamma_{n-1}} \cdot d_{n-1}$. Thus β could be expressed to base α.

Proof: Let $\gamma_0 = \sup\{\gamma \mid \omega^{\gamma} \leq \beta\}$. If $\omega^{\gamma_0} < \beta$ then there is a largest $d_0 \in \omega$ so that $\omega^{\gamma_0} \cdot d_0 \leq \beta$ (thus with $\omega^{\gamma_0} \cdot (d_0 + 1) > \beta$). If $\omega^{\gamma_0} \cdot d_0 = \beta$ we are done. Otherwise there is a unique β_1 so that $\omega^{\gamma_0} \cdot d_0 + \beta_1 = \beta$. Note that in this case $\beta_1 < \beta$. Now repeat the argument: let $\gamma_1 = \sup\{\gamma \mid \omega^{\gamma} \leq \beta_1\}$; by virtue of our construction and the definition of γ_0 and d_0, we must have $\gamma_1 < \gamma_0$. If $\omega^{\gamma_1} < \beta_1$ then define $d_1 \in \omega$ as the largest natural number with $\omega^{\gamma_1} \cdot d_1 \leq \beta_1$. If we have equality here, again we are done. Otherwise there is β_2 defined to be the unique ordinal so that $\omega^{\gamma_1} + \beta_2 = \beta_1$. Since we have $\beta > \beta_1 > \beta_2 \cdots$ there must be some k with $\beta_k = 0$, that is with $\omega^{\gamma_{k-1}} \cdot d_{k-1} = \beta_{k-1}$. Thus β has the form required for the theorem, and this process uniquely determines k and the γ_i. Q.E.D.

Exercise 3.28 Convince yourself that a Cantor Normal Form theorem could be proven for other bases as indicated above: if $\alpha \leq \beta$ we may find finitely many decreasing ordinals γ_i, and $0 < d_i < \alpha$ with $\beta = \alpha^{\gamma_0} \cdot d_0 + \alpha^{\gamma_1} \cdot d_1 + \cdots + \alpha^{\gamma_{n-1}} \cdot d_{n-1}$.

Exercise 3.29 For $\alpha > 0$ show that $\omega \cdot \alpha = \alpha$ iff α is a multiple of ω^{ω}, that is for some δ, $\alpha = \omega^{\omega} \cdot \delta$.

Exercise 3.30 An ordinal σ is called *indecomposable* if $\alpha, \beta < \sigma \to \alpha + \beta < \sigma$. Show that the following are equivalent:

(i) σ is indecomposable

(ii) $\forall \alpha < \sigma (\alpha + \sigma = \sigma)$, *i.e.* σ is a fixed point of A_α for any $\alpha < \sigma$;

(iii) $\sigma = \omega^\delta$ for some ordinal δ.

Exercise 3.31 Show that least indecomposable ordinal greater than α is $\alpha \cdot \omega$.

Exercise 3.32 An ordinal σ is called *multiplicatively indecomposable* if $\alpha, \beta < \sigma \to \alpha \cdot \beta < \sigma$. Show that the following are equivalent:

(i) $1 < \sigma$ is multiplicatively indecomposable;

(ii) $\forall \alpha (0 < \alpha < \sigma \to \alpha \cdot \sigma = \sigma)$, *i.e.* σ is a fixed point of M_α for any $0 < \alpha < \sigma$;

(iii) $\sigma = \omega^{(\omega^\delta)}$ for some ordinal δ.

Exercise 3.33 Formulate a definition for an ordinal $\sigma > 2$ to be *exponentially indecomposable* and demonstrate two equivalences by analogy with the two previous exercises.

Exercise 3.34 (i) Consider the set S_0 of all finite strings of Roman letters with the dictionary or lexicographic ordering. (Thus $a <_{\text{lex}} aa <_{\text{lex}} aaa <_{\text{lex}} \cdots <_{\text{lex}} ab <_{\text{lex}} aba <_{\text{lex}} abd$ etc.) Is $\langle S_0, <_{\text{lex}} \rangle$ a wellordering?

(ii) Now consider the set S_1 of all finite strings of natural numbers (this will be denoted ${}^{<\omega}\omega$). Again consider the lexicographic ordering, where we consider also '$2 <_{\text{lex}} 3$' *i.e.*, so that $<_{\text{lex}}$ also extends the natural $<$ ordering on ω. Is $\langle S_1, <_{\text{lex}} \rangle$ a wellordering?

Exercise 3.35 Faust and Mephistopheles have coins in a currency with k denominations. Mephistopheles offers Faust the following bargain: Every day Faust must give M. a coin, and in return receives as many coins as he, Faust, demands, but only in coins of a lower denomination (except when the coin F. gave was already of the lowest denomination, in which case F. will receive nothing in return). Should Faust accept the bargain? (F. can only demand a finite number of coins each day; part of the bargain is that only M. can call a halt, F. cannot do so - thus the pact may continue indefinitely - hence we assume that F. lives for an indefinite number of days - not just three score and ten years.)

Exercise 3.36 Consider the set $\mathcal{P}$ of polynomials in the variable x with coefficients from $\mathbb{N}$. For $P, Q \in \mathcal{P}$ define $P \prec Q \leftrightarrow$ for all sufficiently large $x \in \mathbb{R}$ $P(x) < Q(x)$. Prove $\langle \mathcal{P}, \prec \rangle \in \mathrm{WO}$.

Exercise 3.37 Let ${}^{<\omega}\omega = \{ f \mid \mathrm{Fun}(f) \wedge \exists k\, (f : k \to \omega) \}$ be the set of all functions into ω with domain some $k \in \omega$. The Kleene-Brouwer ordering on ${}^{<\omega}\omega$ is defined by:

$f <_{\mathrm{KB}} g \leftrightarrow \exists n\, [\, f \upharpoonright n = g \upharpoonright n \ \wedge \ n \in \mathrm{dom}(f) \wedge (n \notin \mathrm{dom}(g) \vee f(n) < g(n))]$

Is it a total ordering? A wellordering?

Exercise 3.38 Let $\langle X, \prec \rangle \in WO$. Let $Q_X = {}^{<\omega}X$. Consider the following order $\prec_1$ on Q_X:

$$f \prec_1 g \longleftrightarrow_{\mathrm{df}} \mathrm{dom}(f) < \mathrm{dom}(g) \vee (\mathrm{dom}(f) = \mathrm{dom}(g) \wedge \exists k \le \mathrm{dom}(f)(\forall n < k f(n) = g(n) \wedge f(k) \prec g(k))).$$

Show that $\langle Q_x, \prec_1 \rangle \in WO$.

Exercise 3.39 Show that the following is a wellorder of ${}^{n}On$: for $\vec{\alpha} = \langle \alpha_0, \ldots, \alpha_{n-1} \rangle, \vec{\beta} = \langle \beta_0, \ldots, \beta_{n-1} \rangle$ set $\vec{\alpha} <^n \vec{\beta}$ iff $\max(\vec{\alpha}) < \max(\vec{\beta})$ or $(\max(\vec{\alpha}) = \max(\vec{\beta})) \wedge$ (if i is least so that $\alpha_i \neq \beta_i$ then $\alpha_i < \beta_i$).

Exercise 3.40 * Let FOn be the class of all finite sets of ordinals. Consider the following ordering $<^*$ on FOn, where as usual $p \triangle q = \{ \alpha \mid \alpha \in p \backslash q \cup q \backslash p \}$ is the *symmetric difference* of p, q:

$$p <^* q \leftrightarrow \max(p \triangle q) \in q.$$

(Or to put it another way: $\exists \beta \in q \backslash p \, (p \backslash (\beta + 1) = q \backslash (\beta + 1))$). Show that $<^*$ is a wellorder of FOn. [Hint: the given $<^*$ is just the same as the lexicographic ordering $<_{\text{lex}}$ (see above) but restricted to finite descending sequences of ordinals $p = p_0 > p_1 > \cdots > p_k$ for variable $k \in \omega$.]

Exercise 3.41 * Use the Cantor Normal Form to devise a pairing function on ordinals: that is to define a bijection $p : On \times On \leftrightarrow On$ with the additional property that $p \upharpoonright \alpha \times \alpha : \alpha \times \alpha \leftrightarrow \alpha$ is a bijection if and only if α is indecomposable (See Ex. 3.30). [Hint: Let $\beta_1 = \omega^{\gamma_0} \cdot d_0 + \omega^{\gamma_1} \cdot d_1 + \cdots + \omega^{\gamma_{k-1}} \cdot d_{k-1}$ and $\beta_2 = \omega^{\gamma_0} \cdot e_0 + \omega^{\gamma_1} \cdot e_1 + \cdots + \omega^{\gamma_{k-1}} \cdot e_{k-1}$ where, in order to match up, some of the d_i's or e_i's may have to be zero (but not both $e_i = d_i = 0$ for any i). Let $p_0 : \omega \times \omega \leftrightarrow \omega$ be any pairing function on ω - with the property that $p_0(0,0) = 0$. Then consider $\omega^{\gamma_0} \cdot p_0(d_0, e_0) + \omega^{\gamma_1} \cdot p_0(d_1, e_1) + \cdots + \omega^{\gamma_{k-1}} \cdot p_0(d_{k-1}, e_{k-1})$.]

Cardinality

"*Je le vois, mais je ne le crois pas!*"
G. Cantor 29.vii.1877. Letter to Dedekind, after discovering that $\mathbb{R} \approx \mathbb{R} \times \mathbb{R}$.

We now turn to Cantor's other major contribution to the foundations of set theory: the theory of *cardinal size* or *cardinality* of sets. Informally we seek a way of assigning a "number" to represent the size or magnitude of a set - any set whether finite or infinite. (And we have yet to define what those two words mean.) We extrapolate from our experience with finite sets when we say that two such sets have the same size when we can pair off the members one with another - just as children do arranging blocks and apples.

4.1 Equinumerosity

Definition 4.1 *Two sets A, B are* equinumerous *if there is a bijection $f : A \longleftrightarrow B$. We write then $A \approx B$ and $f : A \approx B$.*

The idea is that f is both (1-1) and onto, and thus we can "use A to count B" (more familarly from analysis we have A is a natural number or perhaps is $\mathbb{N}$ itself). An alternative word for equinumerous here (but more old-fashioned) is "equipollent". Notice that:

Lemma 4.2 *$\approx$ is an equivalence relation:*
(i) $A \approx A$; (ii) $A \approx B \to B \approx A$; (iii) $A \approx B \wedge B \approx C \to A \approx C$.

Cantor was not the first to consider using $\approx$ as a way of making a judgement about size. As Cantor acknowledged Bolzano had a few years earlier (1851) considered, but rejected it in his notes on infinite sets. Galileo had also pointed out that the squares were in (1-1) correspondence with the counting numbers, and drew the lesson that it was useless to apply concepts from the realm of the finite to talk about infinite collections. Cantor was the first to take the idea seriously.

Definition 4.3 *(i) A set B is* finite *if it is equinumerous with a natural number:*
$$\exists n \in \omega \exists f (f : n \approx B).$$
(ii) If a set is not finite then it is called infinite.

Notice that this definition makes use of the fact that our definition of natural number has built into it the fact that a natural number is the (finite) set of its predecessors, so the above definition makes sense.

Could a set be equinumerous to two different natural numbers? Well, of course not if our definitions are going to make any sense, but this is something to verify.

Lemma 4.4 **(Pidgeon-Hole Principle)** *No natural number is equinumerous to a proper subset of itself.*

Proof: Let $Z = \{n \in \omega \mid \forall f(\textit{If } f : n \to n \textit{ and } f \textit{ is } (1\text{-}1), \textit{then } \mathrm{ran}(f) = n)\}$. (Thus members of Z cannot be mapped in a (1-1) way to proper subsets of themselves.) Trivially $0 \in Z$. Suppose $n \in Z$, and prove that $n+1 \in Z$. Let f be (1-1) and $f : n+1 \to n+1$.

Case 1 $f \upharpoonright n : n \to n$.

Then by Inductive hypothesis, $\mathrm{ran}(f \upharpoonright n) = n$. Then we can only have $f(n) = n$ and thus $\mathrm{ran}(f) = n+1$.

Case 2 $f(m) = n$ for some $m \in n$.

As f is (1-1) we must have then $f(n) = k$ for some $k \in n$. We define g to be just like f but we swap around the action on n, m: define g by $g(m) = k, g(n) = n$ and $g(l) = f(l)$ for all $l \neq m, n$. Now $g : n+1 \to n+1$ and $g \upharpoonright n : n \to n$. By *Case 1* $\mathrm{ran}(g)$ equals $n+1$, but in that case so does $\mathrm{ran}(f)$. Q.E.D.

Corollary 4.5 *No finite set is equinumerous to a proper subset of itself.*

Exercise 4.1 Prove this.

Corollary 4.6 *Any finite set is equinumerous to a unique natural number.*

The next corollary is just the contrapositive of Cor. 4.5.

Corollary 4.7 *Any set equinumerous to a proper subset of itself is infinite.*

Corollary 4.8 *ω is infinite.*

Exercise 4.2 Prove the corollaries 4.6 & 4.8.

Exercise 4.3 Show that if $A \subsetneqq n \in \omega$ then $A \approx m$ for some $m < n$. Deduce that any subset of a finite set is finite.

Exercise 4.4 Suppose A is finite and $f : A \to A$. Show that f is (1-1) iff $\mathrm{ran}(f) = A$.

Exercise 4.5 Let A, B be finite. Without using any arithmetic, show that $A \cup B$ and $A \times B$ is finite.

Exercise 4.6 Show that if A is finite and $\langle A, R \rangle$ is a strict total order, then it is a wellorder (and note in this case that $\langle A, R^{-1} \rangle \in \mathrm{WO}$ too).

Theorem 4.9 **(Cantor, Dec. 7'th 1873)**

The natural numbers are not equinumerous to the real numbers: $\omega \not\approx \mathbb{R}$.

Proof: Suppose $f : \omega \to \mathbb{R}$ is (1-1). We show that $\mathrm{ran}(f) \neq \mathbb{R}$ so such an f can never be a bijection. This is the famous "diagonal argument" that constructs a number that is not on the list. We assume that the real numbers in $\mathrm{ran}(f)$ are written out in decimal notation.

$$f(0) = \quad 3.31415926\ldots$$
$$f(1) = \; -2.4245\ldots$$
$$f(2) = 176.047321\ldots \quad etc.$$

We let x be the number $0.212\ldots$ obtained by letting x have 0 integer part, and putting at the $n+1$'st decimal place a 1 if the $n+1$st decimal place of $f(n)$ is even, and a 2 if it is odd. The argument concludes by noting that x cannot be $f(n)$ for any n as it is deliberately made to differ from $f(n)$ at the $n+1$'st decimal place. Q.E.D.

Remark: in the above proof we have used the fact that if a number has a decimal representation involving only the digits 1 and 2 beyond the decimal point, then the number's representation is unique. Some authors use 0's and 9's (or binary) and then worry about the fact that $0.3999\ldots.$ is the same as 0.40000 (or, in binary, that $0.01111\ldots$ is the same as $0.1000\ldots$). The above choice of 1's and 2's avoids this. (They also, somewhat oddly, only argue with a list $f : \omega \to (0,1)$, and show first that $(0,1)$ is uncountable - which of course implies that the superset $\mathbb{R}$ is uncountable - but the restriction is unnecessary.)

Theorem 4.10 **(Cantor)** *No set is equinumerous to its power set:* $\forall X(\, X \not\approx \mathcal{P}(X))$.

Proof: Similar to the argument of the Russell Paradox: suppose for a contradiction that $f : X \approx \mathcal{P}(X)$. Let $Z = \{u \in X \mid u \notin f(u)\}$. Argue that although $Z \in \mathcal{P}(X)$ it cannot be $f(u)$ for any $u \in X$. Q.E.D.

Definition 4.11 *We define: (i)* $X \preceq Y$ *if there is a (1-1)* $f : X \to Y$ *(and write* $f : X \preceq Y$*)*
(ii) $X \prec Y$ *iff* $X \preceq Y \wedge Y \npreceq X$.

Note that then $X \approx Y \to X \preceq Y \wedge Y \preceq X$. The next theorem will show that the converse is true.

Exercise 4.7 (i) Show that $X \preceq Y$ implies that $\mathcal{P}(X) \preceq \mathcal{P}(Y)$; (ii) Show that if $X \preceq X'$ and $Y \preceq Y'$, then $X \times Y \preceq X' \times Y'$. (iii) Give an example to show that $X \prec X'$ and $Y \preceq Y'$, does not imply that $X \times Y \prec X' \times Y'$.

Theorem 4.12 **(Cantor-Schröder-Bernstein)** $X \preceq Y \wedge Y \preceq X \to X \approx Y$.

Proof: Suppose we have the (1-1) functions $f : X \to Y$ and $g : Y \to X$. We need a bijection between X and Y and we piece one together from the actions of f and g.

We define by recursion: $C_0 = X - \mathrm{ran}(g)$
$$C_{n+1} = g``f``C_n.$$

Thus $C_0$is that part of X that stops g from being a bijection. We then define

$$h(v) = \begin{cases} f(v) & \text{if } v \in C_n \text{ for some } n \quad \textit{Case 1} \\ g^{-1}(v) & \text{otherwise.} \qquad\qquad \textit{Case 2} \end{cases}$$

Note that the second case makes sense: if $v \in X$ but $v \notin C_n$ for any n, then in particular it is not in C_0, that is $v \in \mathrm{ran}(g)$.

We now define $D_n =_{df} f``C_n$. (Note that this makes $C_{n+1} = g``D_n$.) We claim that h is our required bijection.

h is (1-1): Let $u, v \in X$; as both f and g^{-1} are (1-1) the only problem is if say, $u \in \text{dom}(f)$ and $v \in \text{dom}(g^{-1})$, *i.e.*, for some m say, $u \in C_m$ and $v \notin \bigcup_{n \in \omega} C_n$ (or *vice versa)*. However then:

$h(u) = f(u) \in D_m$;

$h(v) = g^{-1}(v) \notin D_m$ (it is not in D_m because otherwise we should have $v \in C_{m+1}$ a contradiction).

Hence $h(u) \neq h(v)$.

h is onto Y: $\forall n D_n \subseteq \text{ran}(h)$. So consider $u \in Y - \bigcup_n D_n$. $g(u) \notin C_0 = X - \text{ran}(g)$ and $g(u) \notin C_{n+1}$ for any n either: this is because $C_{n+1} = g``D_n$ and $u \notin D_n$. So $g(u)$ cannot end up in C_{n+1} without it being equal to some $g(v)$ with $u \neq v \in D_n$. This would contradict the fact that g is (1-1). Therefore *Case 2* applies and $h(g(u)) = g^{-1}(g(u)) = u$. Q.E.D.

The proof of this theorem has a chequered history: Cantor proved it in 1897 but his proof used the Axiom of Choice (to be discussed later) which the above proof eschews. Schröder announced that he had a proof of the theorem in 1896 but in 1898 published an incorrect proof! He published a correction in 1911. The first fully satisfactory proof was due to Bernstein, but was published in a book by Borel, also in 1898.

Exercise 4.8 Show that (i) $(-1,1) \approx \mathbb{R}$; (ii) $(0,1) \approx [0,1]$ by finding directly suitable bijections, *without* using Cantor-Schröder-Bernstein.

Definition 4.13 *Let X be any set, we define the* characteristic function of *$Y \subseteq X$ to be the function $\chi_Y : X \to 2$ so that $\chi_Y(a) = 1$ if $a \in Y$ and $\chi_Y(a) = 0$ otherwise.*

Exercise 4.9 Show that $\mathcal{P}(\omega) \approx \mathbb{R} \approx {}^{\omega}2$. [Hint: First show that $\mathcal{P}(\omega) \approx (0,1)$. It may be easier to show that $\exists f : \mathcal{P}(\omega) \preceq (0,1)$ (by using characteristic functions of $X \subseteq \omega$ and mapping them to binary expansions). Then show that $\exists g : (0,1) \preceq \mathcal{P}(\omega)$ using a similar device. Then appeal to Cantor-Schröder-Bernstein to obtain the first $\mathcal{P}(\omega) \approx (0,1)$. Now note that $\mathcal{P}(\omega) \approx^{\omega} 2$ is easy: subsets $X \subseteq \omega$ are in (1-1) correspondence with their characteristic functions χ_X.]

Exercise 4.10 Show directly (without using that $\mathcal{P}(X) \approx {}^{X}2$ or the CSB Theorem) that $X \prec {}^{X}2$.

Definition 4.14 *A set X is* denumerably infinite *or* countably infinite *if $X \approx \omega$. It is* countable *if $X \preceq \omega$.*

Note that finite sets are countable according to this definition. Trivially from this:

Lemma 4.15 *Any subset of a countable set is countable.*

Exercise 4.11 (i) Show that $\varnothing \neq X$ is countable iff there is $f : \omega \to X$ which is onto. [Hint for ($\Leftarrow$): Construct a (1-1) map from f, demonstrating $X \preceq \omega$.]

(ii) Prove that X is countable and infinite $\Leftrightarrow$ X is countably infinite.

Lemma 4.16 *Let X and Y be countably infinite sets. Then $X \cup Y$ is countably infinite.*

By induction we could then prove for any n that if $X_{0,} \ldots, X_n$ are all countably infinite then so is their union $\bigcup_{i \leq n} X_i$.

Exercise 4.12 Show that $\omega \approx \omega \times \omega$. [Hint: consider the function $f(m,n) = 2^m(2n+1) - 1$. For future reference we let $(u)_0$ and $(v)_1$ be the (1-1) "unpairing" inverse functions from ω to ω so that $f((u)_0, (u)_1) = u$.]

Exercise 4.13 Show that $\mathbb{Z}, \mathbb{Q}$ are both countably infinite. [One way for $\mathbb{Q}$: use Ex.4.11 (i) and 4.12.]

Exercise 4.14 Prove this last lemma.

Exercise 4.15 Let X, Y, Z be sets. Either by providing suitable bijections, or by establishing injections in each direction and using Cantor-Schröder-Bernstein, in each case show that:
(i) $X \times (Y \times Z) \approx (X \times Y) \times Z$ and $X \times (Y \cup Z) \approx (X \times Y) \cup (X \times Z)$ (assume $Y \cap Z = \varnothing$) ;
(ii) ${}^{X \cup Y}Z \approx {}^{X}Z \times {}^{Y}Z$; (assume $X \cap Y = \varnothing$)
(iii) ${}^{X}(Y \times Z) \approx {}^{X}Y \times {}^{X}Z$;
(iv) ${}^{X}({}^{Y}Z) \approx {}^{(X \times Y)}Z$.

Exercise 4.16 Suppose K, L are sets bijective with (not necessarily the same) ordinals. Show that both $K \cup L$ and $K \times L$ are bijective with ordinals.

Lemma 4.17 *Let X be an infinite set, and suppose R is a wellordering of X. Then X has a countably infinite subset.*

Proof: Let x_0 be the R-least element of X. Define by recursion $x_{n+1} = R$-least element of $X - \{x_0, \ldots, x_n\}$. The latter is non-empty, because X was assumed infinite. Hence for every $n < \omega$, x_{n+1} is defined. Then $X_0 = \{x_n \mid n < \omega\}$ is a countably infinite subset of X. Q.E.D.

Without the supposition of the existence of a wellordering on X we could not run this argument. We therefore adopt the following.

Wellordering Principle (WP): *Let X be any set, then there is a wellordering R of X.*

For some sets x we know already that x can be wellordered, for example if x is finite or countably infinite (Why?). But in general this cannot be proven. It will turn out that the Wellordering Principle is equivalent to the Axiom of Choice.

Lemma 4.18 *Assume the Wellordering Principle. Then if $X_0, \ldots, X_n, \ldots (n < \omega)$ are all countably infinite then so is $\bigcup_{i<\omega} X_i$.*

Remark 4.19 Remarkably, it can be proven that without WP we are unable to prove this.

Proof: The problem is that although we are told that each X_i is bijective with ω we are not *given* the requisite functions - we are just told they exist. We must choose them, and this is where WP is involved. Let $Z = \{g \mid \exists i < \omega (g : \omega \approx X_i)\}$. Then Z is a set (it is a subset of $\bigcup\{{}^{\omega}X_i \mid i < \omega\}$). Let R be a wellordering of Z. Set our choice of g_i to be the R-least function $\bar{g} : \omega \approx X_i$. We shall amalgamate all the functions g_i for $i < \omega$, into a single function g which will be onto $\bigcup_{i<\omega} X_i$. An application of Ex.4.11 then guarantees that $\mathrm{ran}(g)$ is countable. To do the amalgamation we use the function f of Exercise 4.12, satisfying $f : \omega \times \omega \approx \omega$. Define $g : \omega \to \bigcup_{i<\omega} X_i$ by $g(f(i, n)) = g_i(n)$. Then $\mathrm{dom}(g)$ is by design $\mathrm{ran}(f) = \omega$ and now *Check* that g is onto.

4.2 Cardinal numbers

We shall assume the Wellordering Principle from now on. This means that for any set X we can find R, a wellordering of it. However if $\langle X, R\rangle \in \mathrm{WO}$ then it is isomorphic to an ordinal. If $f : \langle X, R\rangle \cong \langle \alpha, \in\rangle$ is such an isomorphism, then in particular $f : X \approx \alpha$ is a bijection. In general for a set X there will be many bijections between it and different ordinals (indeed many bijections between it and a single ordinal), but that allows for the following definition.

Definition 4.20 *Let X be any set, the* cardinality *of X, written $|X|$, is the least ordinal κ with $X \approx \kappa$.*

• This corresponds again with notion of finite cardinality. Note that if X is finite then there is just one ordinal γ with $X \approx \gamma$ (namely that $\gamma \in \omega$ with which it is bijective). This just follows from the Pidgeon-Hole Principle.

• However as already stated, a set may be bijective with different ordinals: $\omega \approx \omega + 1 \approx \omega + \omega$ for example. Still for an infinite set X, $|X|$ also makes sense.

Lemma 4.21 *For any sets $X, Y (i) X \approx Y \Leftrightarrow |X| = |Y|$; (ii) $X \preceq Y \Leftrightarrow |X| \leq |Y|$; (iii) $X \prec Y \Leftrightarrow |X| < |Y|$.*

Proof: These are really just chasing the definitions: let $\kappa = |X|$, $\lambda = |Y|$. Let $g : X \approx \kappa$, $h : Y \approx \lambda$. For (i) ($\Rightarrow$) Let $f : X \to Y$ be any bijection. Then $\lambda \nless \kappa$ since otherwise $h \circ f$ is a bijection of X with $\lambda < \kappa = |X|$ - a contradiction. Similarly $\kappa \nless \lambda$ since otherwise $g \circ f^{-1} : Y \approx \kappa < \lambda$ contradicting the definition of λ as $|Y|$. ($\Leftarrow$) Suppose $\kappa = \lambda$ and just look at $h^{-1} \circ g$. This finishes (i). Complete (ii) and (iii) is an exercise.

Q.E.D.

Exercise 4.17 Complete (ii) and (iii) of this lemma.

This last lemma (together with WP) shows that we can choose suitable ordinals as "cardinal numbers" to compare the sizes of sets. Cantor's theorems in this notation are that $|\mathbb{N}| < |\mathbb{R}|$ and in general $|X| < |\mathcal{P}(X)|$. In general when we are dealing with the abstract properties of cardinality, the lemma also shows that we might as well restrict ourselves to a discussion of the cardinalities of the ordinals themselves. All in all we end up with the following definition of *cardinal number*.

Definition 4.22 *An ordinal α is a* cardinal *or* cardinal number, *if $\alpha = |\alpha|$.*

Notice that we could have obtained an equivalent definition if we had said that an ordinal number is a cardinal if there is *some* set X with $\alpha = |X|$. (Why? Because if $\alpha = |X|$ for some set X, then we have by definition that α is least so that $\alpha \approx X$. So we cannot have the existence of a smaller $\beta \approx \alpha$ - for otherwise, by composing bijections, we should have $\beta \approx X$. Hence $\alpha = |\alpha|$. Similar arguments will be implicitly used below.)

• We tend to reserve middle of the greek alphabet letters for cardinals: $\kappa, \lambda, \mu, \ldots$

• Check that this means β is not a cardinal iff there is $\gamma < \beta$ with $\beta \preceq \gamma$.

• For any $\alpha \in \mathrm{On}$ $\alpha \geq |\alpha| = ||\alpha||$. (Check!)

Exercise 4.18 Check: each $n \in \omega$ is a cardinal, ω itself is a cardinal. [Hint: just consult the definition together with some previous lemmas and corollaries.]

Exercise 4.19 Suppose $\alpha \geq \omega$. (i) Show $\alpha \approx \alpha + 1$. (ii) Suppose that $0 < n < \omega$. Show that $\alpha + n$ is not a cardinal, nor is $\alpha + \omega$. [Hint: try it with $\alpha = \omega$ first; find a (1-1) map f from $\alpha + n$ (or $\alpha + \omega$ respectively) into α.]

Note: The last Exercise shows that infinite cardinals are limit ordinals.

Lemma 4.23 *If* $|\alpha| \leq \gamma \leq \alpha$ *then* $|\alpha| = |\gamma|$.

Proof: By definition there is $f : \alpha \approx |\alpha|$ and by Lemma 4.21(i) $||\alpha|| = |\alpha|$. Now $(\gamma \leq \alpha \longleftrightarrow \gamma \subseteq \alpha)$, hence $f \restriction \gamma : \gamma \to |\alpha|$ (1-1). Hence $\gamma \preceq |\alpha|$. But $|\alpha| \leq \gamma$ implies that $|\alpha| \subseteq \gamma$ so trivially $|\alpha| \preceq \gamma$. By CSB $\gamma \approx |\alpha|$. Hence, again by Lemma 4.21(i): $|\gamma| = ||\alpha|| = |\alpha|$. Q.E.D.

Exercise 4.20 Let S be a set of cardinals without a largest element. Show that $\sup S$ is a cardinal.

Exercise 4.21 Show that an infinite set cannot be split into finitely many sets of strictly smaller cardinality. [Hint: Suppose that Y is an infinite set. Let $X \subseteq Y$, and suppose that $|X| < |Y|$. Show that $|Y \backslash X| = |Y|$.]

4.3 Cardinal arithmetic

We now proceed to define arithmetic operations on cardinals. Note that these, other than their restrictions to finite cardinals, are very different from their ordinal counterparts.

Definition 4.24 *Let* κ, λ *be cardinals. We define*
(i) $\kappa \oplus \lambda = |K \cup L|$ *where* K, L *are any two* disjoint *sets of cardinality* κ, λ *respectively.*
(ii) $\kappa \otimes \lambda = |K \times L|$ *where* K, L *are any two sets of cardinality* κ, λ *respectively.*

Notes: (1) There is an implicit use of Exercise 4.16 to guarantee that the chosen sets indeed have cardinalities. Here it really does not matter which sets K, L one takes: if K', L' are two others satisfying the same conditions, then there are bijections $F : K \approx K'$ and $G : L \approx L'$ and thus $F \cup G : K \cup L \approx K' \cup L'$ (and similarly $K \times L \approx K' \times L'$). So simply as far as size goes it is immaterial which underlying sets we consider. (ii) can be paraphrased as $|X \times Y| = ||X| \times |Y|| = |X| \otimes |Y|$ for any sets X, Y. (See also (4) below.)

(2) Unlike ordinal operations, $\oplus$ and $\otimes$ are commutative. This is simply because in their definitions, $\cup$ is trivially commutative, and $K \times L \approx L \times K$. It is easily reasoned that they are associative too.

(3) $\kappa \oplus \kappa = |\kappa \times \{0\} \cup \kappa \times \{1\}| = |\kappa \times 2| = \kappa \otimes 2$ by definition.

(4) For any ordinals α, β: $|\alpha \times \beta| = |\alpha| \otimes |\beta|$ follows directly from the definition of $\otimes$.

Lemma 4.25 *For* $n, m \in \omega$ $m + n = m \oplus n < \omega$ *and* $m \cdot n = m \otimes n < \omega$.

Proof: We already know that $m + n, m \cdot n < \omega$. One can prove directly that $m + n = m \oplus n$ (or by induction on n), and $m \cdot n = m \otimes n$ similarly. Q.E.D.

Exercise 4.22 Complete the details of the last lemma.

Exercise 4.23 Convince yourself that for any ordinals α, β: $|\alpha +' \beta| = |\alpha| \oplus |\beta|$; $|\alpha \cdot' \beta| = |\alpha| \otimes |\beta|$ (and so the same will hold for ordinal $+$ and $\cdot$ replacing $+'$ and $\cdot'$). [Hint: This is really rather obvious given our definitions of $+'$ and $\cdot'$ using *disjoint* copies of α and β.

The next theorem shows how different cardinal multiplication is from ordinal multiplication. We shall use the following exercise in its proof.

Exercise 4.24 (i) Suppose that $\langle A,R\rangle$, $\langle A',R'\rangle$ are in WO with both A, A' uncountable, but so that every proper initial segment of $\langle A,R\rangle$ or $\langle A',R'\rangle$ is countable. Show that $\langle A,R\rangle \cong \langle A',R'\rangle \cong \langle \omega_1, <\rangle$ where ω_1 is the least uncountable ordinal (which then is the least uncountable cardinal).
(ii) Now do this for larger cardinals. Suppose $\langle A,R\rangle \in$ WO and there is a cardinal κ with $|A| \geq \kappa$, but so that for every $b \in A$ the initial segment $\langle A_b, R\rangle \cong \langle \delta, \in\rangle$ for a $\delta < \kappa$. Show that $\mathrm{ot}(\langle A,R\rangle) = \kappa$.

Theorem 4.26 **(Hessenberg)** *Let κ be an infinite cardinal. There is a bijection $\kappa\times\kappa \approx \kappa$ and thus $\kappa \otimes \kappa = \kappa$.*

Proof: By transfinite induction on κ. As $\omega\times\omega \approx \omega$ (Ex.4.12), we already know that $\omega\otimes\omega = |\omega\times\omega| = \omega$. Thus we assume the theorem holds for all smaller infinite cardinals $\lambda < \kappa$ and prove it for κ. We consider the following *Gödel ordering* on $\kappa\times\kappa$:

$$\langle\alpha,\beta\rangle \lhd \langle\gamma,\delta\rangle \Leftrightarrow_{\mathrm{df}} \max\{\alpha,\beta\} < \max\{\gamma,\delta\} \vee [\max\{\alpha,\beta\} = \max\{\gamma,\delta\} \wedge (\alpha<\gamma \vee (\alpha=\gamma\wedge\beta<\delta))]$$

(Note the last conjunct here is just the lexicographic ordering on $\kappa\times\kappa$.)

(1) *$\lhd$ is a wellorder of $\kappa\times\kappa$.*

Proof: Let $\varnothing \neq X \subseteq \kappa\times\kappa$. Let, in turn: $\gamma_0 = \min\{\max\{\alpha,\beta\} \mid \langle\alpha,\beta\rangle\in X\}$; $X_0 = \{\langle\alpha,\beta\rangle \in X \mid \max\{\alpha,\beta\} = \gamma_0\}$; $\alpha_0 = \min\{\alpha \mid \langle\alpha,\beta\rangle \in X_0\}$; and $\beta_0 = \min\{\beta \mid \langle\alpha_0,\beta\rangle\in X_0\}$. Then consider $\langle\alpha_0,\beta_0\rangle$. $\square(1)$

The ordering starts out:

$\langle 0,0\rangle \lhd \langle 0,1\rangle \lhd \langle 1,0\rangle \lhd \langle 1,1\rangle \lhd \langle 0,2\rangle \lhd \langle 1,2\rangle \lhd \langle 2,0\rangle \lhd \langle 2,1\rangle \cdots \lhd \langle 0,\omega\rangle \lhd \langle 1,\omega\rangle \lhd \langle 2,\omega\rangle \cdots \lhd \langle\omega,0\rangle \lhd \langle\omega,1\rangle \cdots \lhd \langle\omega,\omega\rangle\cdots$

(2) *Each $\langle\alpha,\beta\rangle\in\kappa\times\kappa$ has no more than $|\max(\alpha,\beta)+1\times\max(\alpha,\beta)+1| < \kappa$ many $\lhd$-predecessors.*

Proof: By looking at the square pattern that occurs, the predecessors of $\langle\alpha,\beta\rangle$ fit inside a cartesian product box of this size. To state it precisely, $A_{\langle\alpha,\beta\rangle} \subset \gamma\times\gamma$ where we set $\gamma = \max\{\alpha,\beta\}+1 < \kappa$. But by Remark (4) following on Def.4.24, $|\gamma\times\gamma| = |\gamma|\otimes|\gamma|$. As $\gamma<\kappa$,then $|\gamma|<\kappa$ and so by the inductive hypothesis we have $|\gamma|\otimes|\gamma| < \kappa$ as required. $\square(2)$

By (2) it follows that $\lhd$ has the property that every initial segment has cardinality less than κ. The whole ordering certainly has size $\geq\kappa$ since for every $\alpha<\kappa$ $\langle\alpha,0\rangle$ is in the field of the ordering! That means (by Exercise 4.24) that $\mathrm{ot}(\langle\kappa\times\kappa,\lhd\rangle) = \kappa$. But that means we have an order isomorphism between $\langle\kappa\times\kappa,\lhd\rangle$ and $\langle\kappa,\in\rangle$. But such an isomorphism is a bijection. Hence we deduce $\kappa\times\kappa\approx\kappa$, which translates to $\kappa\otimes\kappa=\kappa$. Q.E.D.

Corollary 4.27 *Let κ,λ be infinite cardinals. Then $\kappa\oplus\lambda = \kappa\otimes\lambda = \max\{\kappa,\lambda\}$.*

Proof: Assume $\lambda\leq\kappa$, so $\kappa = \max\{\kappa,\lambda\}$. Then let X,Y be disjoint with $|X| = \kappa, |Y| = \lambda$. (Then $Y \preceq X \preceq X\times\{1\}$.) Thus we have:

$$X \preceq X\cup Y \preceq X\times\{0\}\cup X\times\{1\} = X\times 2 \preceq X\times X.$$

In terms of cardinal numbers (*i.e.* Lemma 4.21) this expresses:

$$|X| \leq |X\cup Y| \leq |X\times\{0\}\cup X\times\{1\}| = |X\times 2| \leq |X\times X|, \text{ or:}$$

$$\kappa \leq \kappa\oplus\lambda \leq \kappa\oplus\kappa = \kappa\otimes 2 \leq \kappa\otimes\kappa.$$

However Hessenberg shows that $\kappa \otimes \kappa = \kappa$ so we have equality everywhere above, and in particular $\kappa = \kappa \oplus \lambda = \max\{\kappa, \lambda\}$.

Further: $X \preceq X \times Y \preceq X \times X$. Again in terms of cardinals, and quoting Hessenberg:
$\kappa \leq \kappa \otimes \lambda \leq \kappa \otimes \kappa = \kappa$ and so $\kappa \otimes \lambda = \max\{\kappa, \lambda\} = \kappa$ again. Q.E.D.

Exercise 4.25 Show that for infinite cardinals $\omega \leq \kappa \leq \lambda$ that $\kappa \oplus \lambda = \lambda$ directly, that is without use of Hessenberg's Theorem.

Exercise 4.26 Let $\lhd$ be the wellorder on $\kappa \times \kappa$ from Hessenberg's Theorem. Let $o(\alpha) =_{df} \mathrm{ot}(\alpha \times \alpha, \lhd)$. Show (i) $\{\langle \alpha, \beta \rangle \mid \langle \alpha, \beta \rangle \lhd \langle 0, \gamma \rangle\} = \gamma \times \gamma$; (ii) $o(\alpha + 1) = o(\alpha) + \alpha + \alpha + 1$; (iii) $o(\omega) = \omega$; $o(\omega \cdot 2) = \omega \cdot \omega$; (iv) $o(\alpha) = \alpha$ implies α is indecomposable; (v)* (Harder) $o(\alpha) = \alpha$ is equivalent to α being multiplicatively indecomposable (see Ex.3.32.)

Exercise 4.27 Show that if $\kappa \geq \omega$ is an infinite cardinal, then it is a fixed point of any of the ordinal arithmetic operations A_α, M_α or E_a for any $\alpha < \kappa$: $\alpha + \kappa = \kappa$; $\alpha \cdot \kappa = \kappa$ and $\alpha^\kappa = \kappa$.

Definition 4.28 *Let A be any set. Then ${}^{<\omega}A = \bigcup_n {}^nA$; this is the set of all functions $f : n \to A$ for some $n < \omega$.*

Exercise 4.28 Show that ${}^nA \approx A \times \cdots \times A$ (the n-fold cartesian product of A).

Exercise 4.29 $^{(*)}$ Assume WP. Let $|X_n| = \kappa \geq \omega$ for $n < \omega$. Show that that $|\bigcup_n X_n| = \kappa$. (This is the generalisation of Lemma 4.18 for uncountable sets X_n.) [Hint: The $^{(*)}$ means it is supposed to be slightly harder. Follow closely the format of Lemma 4.18; use the fact that we now know $\omega \times \kappa \approx \kappa$ to replace $\omega \times \omega = \omega$ in that argument.]

Corollary 4.29 *Let κ be an infinite cardinal. Then $|{}^{<\omega}\kappa| = \kappa$.*

Proof: It us enough to show that $X_n =_{df} {}^n\kappa$ has cardinality κ and then use Exercise 4.29. However ${}^n\kappa \approx \kappa \times \cdots \times \kappa \approx \kappa$ (the first $\approx$ by Exercise 4.28, the latter $\approx$ by repeated use of the Hessenberg Theorem). Q.E.D.

Definition 4.30 (WP) Let κ, λ be cardinals, then $\kappa^\lambda =_{df} |{}^LK|$, *where L, K are any sets of cardinality λ, κ respectively.*

(Recall that ${}^XY =_{df} \{f \mid f : X \to Y\}$.) We need WP here (unlike the definitions of the other cardinal arithmetic operations) since we need to know that the set of all possible functions *can* be bijective with some ordinal.

Exercise 4.30 Show that the definition of κ^λ is independent of the choices of sets L, K. Deduce that $|{}^XY| = |{}^X|Y|| = |{}^{|X|}|Y|| = |Y|^{|X|}$.

Lemma 4.31 *If κ and λ are cardinals, with $\lambda \geq \omega$, and $2 \leq \kappa \leq \lambda$, then ${}^\lambda\lambda \approx {}^\lambda\kappa \approx {}^\lambda 2 \approx \mathcal{P}(\lambda)$. Hence $2^\lambda = \kappa^\lambda = \lambda^\lambda (= |\mathcal{P}(\lambda)|)$.*

Proof: We can establish ${}^\lambda 2 \approx \mathcal{P}(\lambda)$ by identifying characteristic functions of subsets of λ with those subsets themselves. Now see that: ${}^\lambda 2 \preceq {}^\lambda\kappa \preceq {}^\lambda\lambda \preceq \mathcal{P}(\lambda \times \lambda) \approx \mathcal{P}(\lambda) \approx {}^\lambda 2$ (using Hessenberg's Theorem to see that $\lambda \times \lambda \approx \lambda$, and hence the first $\approx$ holds). Hence we have $\approx$ throughout. Q.E.D.

Lemma 4.32 (WP) *If κ, λ, μ are cardinals, then*
(i) $\kappa^{\lambda \oplus \mu} = \kappa^{\lambda} \otimes \kappa^{\mu}$; *(ii)* $(\kappa^{\lambda})^{\mu} = \kappa^{\lambda \otimes \mu}$.

Proof: (i) This is Exercise 4.15 (ii) with, for example, $X = \lambda \times \{0\}$, $Y = \mu \times \{1\}$, and $Z = \kappa$.

$\kappa^{\lambda\oplus\mu} =_{df} |{}^{\lambda\oplus\mu}\kappa| = |{}^{X\cup Y}\kappa| = |{}^{X}\kappa \times {}^{Y}\kappa|$ (the second equality by Ex 4.30, the last by Ex 4.15 (ii))
$= |{}^{X}\kappa| \otimes |{}^{Y}\kappa|$ (def. of $\otimes$)
$= \kappa^{\lambda} \otimes \kappa^{\mu}$ (using Ex. 4.30).

(ii) $(\kappa^{\lambda})^{\mu} =_{df} |{}^{\mu}(\kappa^{\lambda})| = |{}^{\mu}({}^{\lambda}\kappa)|$ (the latter equality by Ex. 4.30)
$= |{}^{\mu\times\lambda}\kappa|$ (by Exercise 4.15 (iv))
$= |{}^{\lambda\times\mu}\kappa| = \kappa^{\lambda\otimes\mu}$ (since $|{}^{A}\kappa| = \kappa^{|A|}$ - Ex.4.30 - for any set A and $|\lambda \times \mu| = \lambda \otimes \mu$). Q.E.D.

Theorem 4.33 **(Hartogs' Theorem).** *For any ordinal α there is a cardinal $\kappa > \alpha$.*

Remark: The observant may wonder why we prove this: after all Cantor's Theorem showed that for any α, $\alpha \prec \mathcal{P}(\alpha)$ and so $|\mathcal{P}(\alpha)| > \alpha$. This is true, but this required the WP (to argue that $\mathcal{P}(\alpha)$ is bijective with an ordinal, and so has a cardinality). Hartogs' theorem does not require WP - although it does require the Axiom of Replacement - which we have not yet discussed. It shows that there are arbitrarily large cardinals without appealing to Cantor's theorem.

Proof: For finite α this is trivial. Let $\alpha \geq \omega$ be arbitrary. Let $S =_{df} \{R | \langle \alpha, R\rangle \in \mathrm{WO}\}$. Then S is a set - it is a subset of $\mathcal{P}(\alpha \times \alpha)$ and so exists by Power Set and Subset Axioms. Let $\tilde{S} = \{\mathrm{ot}(\langle \alpha, R\rangle) | R \in S\}$. Then to argue that $\tilde{S}$ is a set we need to know that the range of the function that takes a wellordering to its order type, when restricted to a *set* of wellorderings yields a *set* of ordinals. To do this we appeal to the Axiom of Replacement that says that *any* definable function $F : V \to V$ when restricted to a set has a set as its range: $(\forall x \in V)(F``x \in V)$ (see next Chapter).

Then, knowing that $\tilde{S}$ is a set, we form $\sup\ \tilde{S}$ which is then an ordinal $\nu > \alpha$. As $\tilde{S}$ has no largest element (Exercise), ν is a limit ordinal (Lemma 3.31). Hence $\nu \notin \tilde{S}$. Hence there is no onto map $f : \alpha \to \nu$ (for if so we could define a wellordering R by $\gamma R \delta \leftrightarrow f(\gamma) < f(\delta)$; R is a wellordering as $\langle \nu, < \rangle$ is such, and would demonstrate that $\nu \in \tilde{S}$.) Hence $\alpha \not\approx \nu$. But then $\nu \not\approx \delta$ for any $\delta < \nu$, since for such δ there is an onto map from α onto δ (because $\delta < \delta'$ for some $\delta' \in \tilde{S}$ - in fact one may show: $\alpha \leq \delta < \nu \to \delta \in \tilde{S}$). So $|\nu| = \nu$. Q.E.D.

Corollary 4.34 Card $=_{df} \{\alpha \in \mathrm{On} \,| \alpha$ a cardinal$\}$ is also a proper class.

Proof: If there were only a set of cardinals, call it z say, then $\sup(z) \in \mathrm{On}$. By Hartogs' (or Cantor's) Theorem there is nevertheless a cardinal $> \sup(x)$! (For example $|\mathcal{P}(\sup(z))I$ if we are appealing to Cantor's Theorem.) Q.E.D.

Corollary 4.35 *For any set x there is an ordinal ν so that $\nu \not\preccurlyeq x$.*

Exercise 4.31 (Without WP, that is without assuming there is γ with $\gamma \approx x$.) Prove the last corollary. [Hint: this is really Hartogs' theorem, with the set x substituted for α throughout.]

Definition 4.36 *We define by transfinite recursion on the ordinals:*
$\omega_0 = \omega$; $\omega_{\alpha+1}$= *least cardinal number* $> \omega_\alpha$; $\mathrm{Lim}(\lambda) \to \omega_\lambda = \sup\{\omega_\alpha | \alpha < \lambda\}$.

A widely used alternative notation for ω_α uses the Hebrew letter "$\aleph_\alpha$" (read "aleph-sub-alpha"). We shall use both forms.

Definition 4.37 *An infinite cardinal ω_α with $\alpha > 0$, is called an* uncountable *cardinal; it is also called a* successor *or a* limit *cardinal, depending on whether α is a successor or limit ordinal.*

We are thus defining by transfinite recursion a function $F : \mathrm{On} \to \mathrm{On}$ which enumerates all the infinite cardinals starting with $F(0) = \omega_0 = \omega$. This function is *strictly increasing* ($\alpha < \beta \to F(\alpha) = \omega_\alpha < \omega_\beta = F(\beta)$) and it is *continuous* at limits, meaning that $F(\lambda) = \sup\{F(\alpha) \mid \alpha < \lambda\}$ for $\mathrm{Lim}(\lambda)$ - note that this supremum is certainly a cardinal (see Ex.4.20).

.

• Technically we should also call finite cardinals and zero successor cardinals as well. (Infinite) successor cardinals are however of the form $\omega_{\beta+1}$. Given any ordinal ν then, the least cardinal $> \nu$ must then be a successor cardinal, and is written ν^+.

Exercise 4.32 Are there ordinals α so that $\alpha = \omega_\alpha$? If so find one. (Such would be a *fixed point* of the cardinal enumeration function F: we should have $F(\alpha) = \alpha$.)

Cantor wrestled with the problem of whether there could be a set $X \subseteq \mathbb{R}$ that was neither countable, nor bijective with $\mathbb{R}$. Such an X would satisfy $|\mathbb{N}| < |X| < |\mathbb{R}|$. He believed this was impossible. This belief could be expressed as saying that for any infinite set $X \subseteq \mathbb{R}$, either $X \approx \mathbb{N}$ or $X \approx \mathbb{R}$.

If so, then we should have that $|\mathbb{N}| = \omega_0$ and then we must have $|\mathbb{R}|$ would be the size of the very next cardinal, so ω_1: $|\mathbb{R}| = \omega_1$. There would thus be no intermediate cardinal number for such an X to have. This is known as the *Continuum Problem*. As $\mathbb{N} \approx \omega$ and $\mathbb{R} \approx \mathcal{P}(\omega) \approx {}^\omega 2$, we can express Cantor's belief as $|\mathcal{P}(\omega)| = 2^\omega = \omega_1$, and again as $|\mathbb{R}| = \omega_1$.

Definition 4.38 ***(Cantor)* Continuum Hypothesis CH:** $2^{\omega_0} = \omega_1$;
The Generalised Continuum Hypothesis GCH: $\forall\alpha\, 2^{\omega_\alpha} = \omega_{\alpha+1}$.

• The GCH says that $\forall\alpha\, |{}^{\omega_\alpha}2|(= |\mathcal{P}(\omega_\alpha)|) = \omega_{\alpha+1}$, the exponential function $\kappa \mapsto 2^\kappa$ thus again always takes the very least possible value it could.

• As we have said, Cantor believed that CH was true but was unable to prove it. We now know why he could not: the framework within which he worked, was prior to any formalisation of axioms for sets, but even once those axioms were written down and accepted, (the "ZFC" axioms which we have introduced above) we have the following contrasting (and startling) theorems:

Theorem (Gödel 1939) In ZFC set theory we cannot prove $\neg$ CH : it is consistent that $|\mathbb{R}|$ be ω_1.

Theorem (Cohen 1963) In ZFC set theory we cannot prove CH: it is consistent that $|\mathbb{R}|$ be ω_2 (or $\omega_{17}, \omega_{\omega+5}, \ldots$).

CH on the basis of the ZFC axioms is thus an *undecidable* statement. Set theorists have searched subsequently for axioms to supplement ZFC that would decide CH but to date, in vain. We simply do not know the answer, or moreover any simple way of even trying to answer it.

Indeed the cardinal exponentiation function in general is problematic in set theory, little can definitely be said about κ^λ in general. (It is consistent with the ZFC axioms, for example, that $2^{\omega_0} = 2^{\omega_1} = \omega_{17}$, so cardinal exponentiation need not be strictly increasing: $\lambda < \kappa \not\to 2^\lambda < 2^\kappa$.) However work on this function for so-called *singular limit cardinals* κ (and $\lambda < \kappa$) has resulted in a lot of information about the universe of sets V.

Exercise 4.33 Show that CH is equivalent to the statement that every ordinal less than $2^{\aleph_0}$ is countable.

Exercise 4.34 Show that (i) the set of countable subsets of $\mathbb{R}$ has cardinality $2^{\aleph_0}$
(ii) the set of countable subsets of $\mathbb{R}$ which contain all of $\mathbb{Q}$ has cardinality $2^{\aleph_0}$;
(iii) the set of open intervals of $\mathbb{R}$ also has cardinality $2^{\aleph_0}$.

Definition 4.39 (The beth numbers) *We define by transfinite recursion on the ordinals:*
$\beth_0 = \omega$; $\beth_{\alpha+1} = 2^{\beth_\alpha}$; $\mathrm{Lim}(\lambda) \to \beth_\lambda = \sup\{\beth_\alpha | \alpha < \lambda\}$.

- Note that if the GCH holds, then $\forall\alpha(\beth_\alpha = \aleph_\alpha)$.

Exercise 4.35 Prove that there is λ with $\lambda = \beth_\lambda$.

Exercise 4.36 Show that the union of $\kappa \geq \omega$ many sets of cardinality κ is of cardinality κ. [Hint: If $\langle A_i \mid i < \kappa\rangle$ are the sets with each $|A_i| = \kappa$ then consider a (1-1) map into $\kappa \otimes \kappa$.]

Exercise 4.37 Place in correct order the following cardinals using $=, <, \leq$:
$\aleph_{13}, \aleph_{\omega^2}, \varnothing, \aleph_{\omega_1}^{\aleph_{\omega_1}}$, $\sup\{\aleph_n | n < \omega\}, \aleph_{\omega_1} \oplus \aleph_\omega$, $\aleph_\omega, \aleph_{\omega_1} \otimes \aleph_{\omega_1}, \aleph_\omega \oplus \aleph_{\omega_1}, 2^{\varnothing}, \aleph_{\omega_1}$.
You should give your reasons; apart from the 'ω^2' in the second cardinal, the arithmetic is all cardinal arithmetic.

Exercise 4.38 Simplify where possible: $2^{\aleph_0}$; $\aleph_\omega \oplus \aleph_{\omega_1}$; $(2^{\aleph_0})^{\aleph_1}$; $(\aleph_\omega)^3 \oplus (\aleph_5)^2$.
You should do this twice: the first time without assuming the Generalised Continuum Hypothesis, and the second time assuming it. (The operations are all cardinal arithmetic.)

Exercise 4.39 Show directly (without using Hessenberg's Theorem) that for $n < \omega$ $(\beth_n)^2 = \beth_n$. [Hint: use induction on n.]

Exercise 4.40 This exercise asks you to show that various classes of sequences $\{a_n\}_{n<\omega}$ with each $a_n \in \mathbb{N}$ are countable.
(i) The *eventually constant* sequences: $\exists k_0 \forall k \geq k_0\ a_k = a_{k_0}$;
(ii) The *arithmetic progressions*: $\exists p \forall n\ a_{n+1} = a_n + p$;
(iii) The *eventual geometric progressions*: $\exists k_0 \exists p \forall n \geq k_0\ a_{n+1} = a_n \cdot p$.

Exercise 4.41 A real number is said to be *algebraic* if it is a root of a polynomial $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$ where each $a_i \in \mathbb{Q}$. Show that there are only countably many algebraic numbers. A real number that is not algebraic is called *transcendental*. Deduce that almost all real numbers are transcendental, in that the set of such is equinumerous with $\mathbb{R}$.

Exercise 4.42 A *word* in an alphabet Σ is a string of symbols from Σ of finite length. Show that the number of possible words made up from the roman alphabet is countable. If we enlarge the alphabet to be now countably infinite, is the answer different?

Exercise 4.43 What is the cardinality of (i) the set of all order isomorphisms $f : \mathbb{Q} \to \mathbb{Q}$; (ii) the set of all continuous functions $f : \mathbb{R} \longrightarrow \mathbb{R}$? ; (iii) the set of all convergent sequences $\Sigma_{n=0}^{\infty} a_n$ of real numbers?

Exercise 4.44 (i) The *Cantor set* C is the set of all real numbers of the form $\Sigma_{n=0}^{\infty} a_n \cdot 3^{-(n+1)}$ with $a_n \in \{0, 2\}$. Show that $C \approx \mathbb{R}$. (ii) The *Hilbert cube* is the set $\mathcal{H} = {}^{\mathbb{N}}[0,1]$. What is $|\mathcal{H}|$?

Exercise 4.45 Let $\mathcal{V}$ be a vector space, with a basis B. We suppose B to be infinite, in which case we have that $\mathcal{V}$ is an infinite dimensional vector space. How many finite dimensional subspaces does $\mathcal{V}$ have?

Exercise 4.46 Show that the set of all permutations of $\mathbb{N}$ has cardinality $2^{\aleph_0}$.

Exercise 4.47 Show that the set of all Riemann integrable functions on $\mathbb{R}$ has cardinality $(2^{\aleph_0})^{2^{\aleph_0}}$.

Exercise 4.48 Let $(\mathbb{N}, \prec)$ be any strict total order. Show that there is a (1-1) order preserving embedding of $(\mathbb{N}, \prec)$ into $(\mathbb{Q}, <)$.

Exercise 4.49 Let $(\mathbb{N}, \prec)$ be any strict total order; show that there is a (1-1) order preserving map of $(\mathbb{N}, <)$ either into $(\mathbb{N}, \prec)$ or into $(\mathbb{N}, \succ)$.

Exercise 4.50 Let $X \subseteq \mathbb{R}$ and suppose that $(X, <) \in \mathrm{WO}$ where $<$ is the usual order on $\mathbb{R}$. Show that X is countable.

Exercise 4.51 Show that any countable ordinal $(\alpha, \in)$ can be (1-1) order-preserving embedded into $(\mathbb{R}, <)$. Show that no uncountable ordinal can be so embedded.

Definition 4.40 *(i) If $(A, \prec)$ is a strict total order, then it is called* dense*, if for any $x \prec y \in A$ there is $z \in A$ with $x \prec z \prec y$.*
(ii) If $(A, \prec)$ is a strict total order, and $B \subseteq A$ then $(B, \prec)$ is called a dense suborder *if for any $x \prec y \in A$ there is $z \in B$ with $x \prec z \prec y$.*

Exercise 4.52 * Let $(\mathbb{N}, \prec)$ be any strict total order which is dense and has no endpoints, *i.e.* no maximum nor minimum elements. Show that $(\mathbb{N}, \prec) \cong (\mathbb{Q}, <)$. Deduce that any two countable dense total orders without endpoints are isomorphic. (This is a theorem of Cantor.)

Exercise 4.53 (i) Find $(P, <)$ and $(S, <)$ two countable suborders of $(\mathbb{R}, <)$ with $(P, <) \cong (S, <)$ but $(\mathbb{R} \backslash P, <) \ncong (\mathbb{R} \backslash S, <)$.
(ii) * Show that if $(P, <)$ and $(S, <)$ are two countable dense suborders of $(\mathbb{R}, <)$ then $(\mathbb{R} \backslash P, <) \cong (\mathbb{R} \backslash S, <)$. [For (ii) use the last Exercise.]

Exercise 4.54 Suppose $(P, <)$ is a dense suborder of $(\mathbb{R}, <)$. Show that there is a countable $S \subseteq P$ with $(S, <)$ a dense suborder of $(P, <)$.

A note on* Dedekind-finite *sets:

Dedekind tried to give a direct definition of *infinite set* as any set X for which there was a (1-1) map of X to a proper subset of itself. Let us call such a set *D-infinite*. By contrast a *Dedekind finite* set, was defined as any set that was not D-infinite. However notice that this means for a particular set X, it is D-finite if there is no (1-1) map of a certain kind. The question then arises: are D-finite sets always finite (in our sense)? Or could there be a D-finite set that is infinite? It turns out that this depends on the Wellordering Principle. If WP holds then for any set X there is R so that $\langle X, R \rangle \in \mathrm{WO}$. If X is infinite then we may map X to a proper subset of itself. (How?) Thus any infinite set is also D-infinite. But what if WP fails? It turns out to be consistent with the axioms of set theory that WP fails and that there is an infinite but D-finite set. For many mathematicians this would be reason enough to add WP to our axioms of set theory - although there are many other reasons also to do so.

1599

Axioms of Replacement and Choice

We consider in this chapter the Axiom of Choice (AC) and its various equivalents, one of which we have already mentioned: the Wellordering Principle (WP). However we first look more closely at another axiom which delimits the existence of sets.

5.1 Axiom of Replacement

This axiom (which we have already used in one or two places) asserts that the action of a function on a set produces a set.

Axiom of Replacement *Let $F : V \to V$ be any function, and let x be any set. Then*

$$F``x =_{df} \{z | \exists u \in x(F(u) = z)\}$$

is a set.

The import of the axiom is one of *delimitation of size*: it says that a function applied to a set cannot produce a proper class, *i.e.* something that is too large. It thus appears *prima facie* to be different from those of the other axioms, which assert simple set existence. The 'replacement' is that of taking a set X and 'replacing' each element $u \in X$ by some other set a; and that a is specified by F: $F(u) = a$. If this is done for each $u \in X$ the resulting $X' = F``X$ should still be considered a set.

Examples: (i) Let $F(x) = \{x\}$ for any set x. Then the Axiom of Replacement ensures that $F``\,\omega = \{\{0\}, \{1\}, \{2\}, \ldots, \{n\}, \ldots\}$ is a set.

(ii) Likewise Replacement is needed to justify that $\{\aleph_0, \aleph_1, \aleph_2, \aleph_3, \ldots\}$ is a set which we can think of as $F_\aleph``\omega$ where $F_\aleph(\alpha) = \aleph_\alpha$ for $\alpha \in \mathrm{On}$. Without Replacement we cannot say the supremum of this set exists (which supremum is $\aleph_\omega$).

(iii) Similarly $V_{\omega+\omega}$, which will be defined below as $\bigcup\{V_\alpha \mid \alpha < \omega + \omega\}$, requires the use of Replacement on the function F_V where $F_V(\alpha) = V_\alpha$, in order to justify $F_V``\omega + \omega$=$\{V_\alpha \mid \alpha < \omega + \omega\}$ to be a set, before we can apply $\bigcup$ to it.

A slightly less trivial example occurs in the proof of Hartogs' Theorem (Thm. 4.33). There we had a set of wellorders S. Consider the function F that takes x to 0 unless $x = \langle A, R\rangle$ where R wellorders the set A, in which case $F(\langle A, R\rangle)$ returns the ordinal $\mathrm{ot}(\langle A, R\rangle)$. Then $F : V \to V$ is a legitimate function, and the Axiom of Replacement then asserts that $\tilde{S} = \{\mathrm{ot}(\langle \alpha, R\rangle) \mid R \in S\} = F``S$ is a *set* of ordinals.

The axiom was introduced in a paper by Zermelo who attributed it to Fraenkel (although it had been considered by several others before in various versions). In Zermelo's earlier paper there was no mention of any principle such as Replacement (in German *Ersetzung*) and thus in his axiomatic system (which was, and is, called Z for Zermelo) the set of finite numbered alephs in Example (ii) above, did not exist as a set (and nor did $V_{\omega+\omega}$). Since the set of finite numbered alephs did not exist, $\aleph_\omega$ did not exist.

Figure 5.1: Abraham Fraenkel 1891-1965

Other important examples are afforded by proofs of transfinite recursion theorems such as Theorem 3.32 (although we brushed these details under the carpet at the time). In axiomatic set theory it is usual to think of the function F as given to us defined by some formula $\varphi(u, v)$ where we have proven that $\forall u \in x \exists! v \varphi(u, v)$ (recall that $\exists! v \cdots$ is read "there exists a unique $v \cdots$"). The conclusion then can be expressed as "$\exists w \forall u \in x \exists v \in w \varphi(u, v)$" and then w in effect has been defined as a set *containing* $F``x$. (Then if we want a set that is precisely $F``x$ we may use the Axiom of Subsets to pick out from w just the set of elements in the desired range.)

5.2 Axiom of Choice

This is an axiom that is ubiquitous in mathematics. It appears in many forms: analysts use it to form sequences of real numbers, or to justify that the countable union of countable sets is countable. Algebraists use it to form maximal prime ideals in rings, and functional analysts to justify the existence of bases for infinite dimensional vector spaces. We have adopted as a basic axiom the Wellordering Principle that every set can be wellordered: for any A we may find R so that $\langle A, R \rangle \in \mathrm{WO}$. In particular this meant that $\langle A, R \rangle \cong \langle \alpha, < \rangle$ for some ordinal α, and then we could further define $|A|$ the *cardinality* of A. Without WP we could not have done this. A very common form in set theory text books of AC - the Axiom of Choice - is the following:

Axiom of Choice - AC *Let $\mathcal{G}$ be a set of non-empty sets. Then there is a choice function F so that* $\forall X \in \mathcal{G}(F(X) \in X)$.

The reason for the name "choice function" is obvious: $F(X)$ picks out for us, or chooses for us, a unique element of the set X (which is why we specify that $X \neq \varnothing$). AC turns out to be equivalent to WP. We shall prove this.

THEOREM 5.1 **(Zermelo 1908)** AC $\Longleftrightarrow$ WP.

Proof: ($\Longrightarrow$) Assume AC. Let Y be any set. We may assume that $Y \neq \varnothing$ (otherwise the result is trivial). We seek a wellordering R of Y. Let $\mathcal{G} = \{X \subseteq Y \mid X \neq \varnothing\}$. By AC let F_0 be a choice function for $\mathcal{G}$. Let u be any set not in Y. Now let $F : V \to V$ be defined by:

$$\begin{aligned} F(t) &= F_0(t) \text{ if } t \in G; \\ &= u \text{ otherwise.} \end{aligned}$$

We define by recursion $H : \alpha \to Y$ a (1-1) onto function with domain some $\alpha \in \mathrm{On}$. If we succeed here then we can define easily a wellordering R: put $xRy \longleftrightarrow H^{-1}(x) < H^{-1}(y)$ (this makes sense as H is a bijection). Define:

$$\begin{aligned} H_0(\xi) &= F(Y - \{H_0(\zeta) \mid \zeta < \xi\}) \text{ if the latter is non-empty;} \\ &= u \text{ otherwise.} \end{aligned}$$

Note that this definition implies that $H_0(0) = F(Y - \varnothing) = F(Y) \in Y$. Then by the Theorem on Transfinite Recursion Theorem on On, Theorem 3.35, there is a function $H_0 : \mathrm{On} \to Y \cup \{u\}$.

Claim There is $\beta \in \mathrm{On}$ *with* $H_0(\beta) = u$.

Proof: (The Claim says that sooner or later we exhaust Y.) Suppose not. Then H_0 is a (1-1) function sending *all* of On into the *set* Y. But then H_0^{-1} is a function. Look at H_0^{-1}"Y. By the Axiom of Replacement this is a set. But it is On itself, and by the Burali-Forti Lemma On is a proper class! This is absurd.

Q.E.D.*Claim*

Let α be least with $H_0(\alpha) = u$ and let $H = H_0 \restriction \alpha$. By the above comment this suffices.

($\Longleftarrow$) Suppose WP. Let $\mathcal{G}$ be any set of non-empty sets. Let $A =_{df} \bigcup \mathcal{G} = \{u \mid \exists X \in \mathcal{G}(u \in X)\}$. By WP suppose $\langle A, R\rangle \in \mathrm{WO}$. We need a choice function F for $\mathcal{G}$. Let $X \in \mathcal{G}$ and define $F(X)$ to be the R-least element of X. Check that this works! Q.E.D.

A collection $\mathcal{G}$ of sets is called a *chain* if $\forall X, Y \in \mathcal{G}(X \subseteq Y \vee Y \subseteq X)$.

Zorn's Lemma (ZL) *Let $\mathcal{F}$ be a set so that for every chain $\mathcal{G} \subseteq \mathcal{F}$ then $\bigcup \mathcal{G} \in \mathcal{F}$.Then $\mathcal{F}$contains a maximal element Y, that is $\forall Z \in \mathcal{F}(Y \neq Z \to Y \not\subseteq Z)$.*

THEOREM 5.2 WP $\Leftrightarrow$ AC $\Leftrightarrow$ ZL.

Proof: (ZL $\Rightarrow$ AC) Let $\mathcal{G}$ be a set of nonempty sets. We define $\mathcal{F}$ to be the set of all choice functions that exist on subsets of $\mathcal{G}$. That is we put $f \in \mathcal{F}$ if (a) $\mathrm{dom}(f) \subseteq \mathcal{G}$; (b) $\forall x \in \mathrm{dom}(f) f(x) \in x$. Such an f thus acts as a choice function on its domain, and it may only fail to be a choice function for all of $\mathcal{G}$ if $\mathrm{dom}(f) \neq \mathcal{G}$. Consider a chain $\mathcal{H} \subseteq \mathcal{F}$. $\mathcal{H}$ is thus a collection of partial choice functions of the kind $f, g \in \mathcal{F}$ with the property that either $f \subseteq g$ or $g \subseteq f$. However then if we set $h = \bigcup \mathcal{H}$ we have that h is itself a function and $\mathrm{dom}(h) = \bigcup\{\mathrm{dom}(f) \mid f \in \mathcal{H}\}$. That is h is a partial choice function, so $h \in \mathcal{F}$. Now by ZL there is a *maximal* $m \in \mathcal{F}$.

Claim m is a choice function for $\mathcal{G}$.

Proof: m is a partial choice function for $\mathcal{G}$: it satisfies (a) and (b) above. Suppose it failed to be a choice function. Then there is some $x \in \mathcal{G}$ with $x \notin \mathrm{dom}(m)$. As $\mathcal{G}$ consists of non-empty sets, pick $u \in x$. However then $m \cup \{\langle x, u\rangle\} \in \mathcal{F}$ as it is still a partial choice function, but now we see that m was not maximal. Contradiction!

(WP $\Rightarrow$ ZL) Let $\mathcal{F}$ be a set so that for every chain $\mathcal{G} \subseteq \mathcal{F}$ then $\bigcup \mathcal{G} \in \mathcal{F}$. By WP $\mathcal{F}$ can be wellordered, and *a fortiori* there is a bijection $k : \alpha \longleftrightarrow \mathcal{F}$ for some $\alpha \in \mathrm{On}$. We define by transfinite recursion on α a maximal chain $\mathcal{H}$ by inspecting the members of $\mathcal{F}$ in turn.

We start by putting $k(0)$ into $\mathcal{H}$. If $k(1) \supset k(0)$ we put $k(1)$ into $\mathcal{H}$; if not we ignore it, and consider $k(2)$. We continue in this way inspecting each $k(\alpha)$ in turn and if it extends all the previous $k(\beta)$ *which we put in* $\mathcal{H}$ then we put it into $\mathcal{H}$; and ignore it otherwise. This is an informal transfinite recursion on α. We first claim that $\mathcal{H}$ is a chain. This is obvious as we only add $X = k(\xi)$ say to $\mathcal{H}$, if it contains as subsets all the previous elements already added. We further claim that $\bigcup \mathcal{H}$ is a maximal element of $\mathcal{F}$. By our defining property of $\mathcal{F}, \bigcup \mathcal{H} \in \mathcal{F}$. If $Y \supseteq \bigcup \mathcal{H}$ then Y contains every element of $\mathcal{H}$ as a subset. However, if additionally $Y \in \mathcal{F}$ then $Y = k(\nu)$ for some ν, and so by the definition of our recursion, at stage ν we decided that $Y \in \mathcal{H}$. So $Y \subseteq \bigcup \mathcal{H}$. This suffices since we have now shown $Y = \bigcup \mathcal{H}$. QED

There are many equivalents of AC. We state without proof some more. [1]

Uniformisation Principle (UP) *If $R \subseteq X \times Y$ is any relation, then there is a function $f : X \to Y$ with (i)* $\mathrm{dom}(f) = \mathrm{dom}(r) =_{df} \{x \mid \exists y(\langle x, y\rangle \in R\}$ *and (ii)* $f \subseteq R$.

Inverse Function Principle (IFP) *For any onto function $H : X \to Y$ between sets X, Y, there is a (1-1) function $G : Y \to X$ with $\forall u \in Y(H(G(u)) = u)$.*

Cardinal Comparison *For any two sets X, Y either $X \preceq Y$* or *$Y \preceq X$.*

Hessenberg's Principle *For any infinite set $X \approx X \times X$.*

Vector Space Bases *Every vector space has a basis.*

Tychonoff Property *Let G be any set of non-empty sets. Then the direct product $\Pi_{X \in G} X \neq \varnothing$* [Here $\Pi_{i \in I} X_i =_{df} \{f \,|\mathrm{dom}(f) = I \wedge \forall i \in I(f(i) \in X_i\}$. Clearly each such f is a choice function for $\{X \mid X \in G\}$.]

Tychonoff-Kelley Property *Let X_i (for $i \in I$) be any sequence of compact topological spaces. Then the direct product space $\Pi_{i \in I} X_i$ is a compact topological space.*

It can also be shown that GCH $\Longrightarrow$ AC but this is not an equivalence.

Exercise 5.1 Show that AC $\Leftrightarrow$ UP

Exercise 5.2 Show that AC $\Rightarrow$ IFP.

In general with the above exercises the converse implications are harder.

Exercise 5.3 Show that WP $\Leftrightarrow$ Cardinal Comparison. [Hint: for ($\Leftarrow$) use the Cor. 4.35.]

Exercise 5.4 Show that AC $\Leftrightarrow$ Tychonoff Property.

Exercise 5.5 Show that WP $\Rightarrow$ Vector Space Bases. [Hint: use the argument for finite dimensional vector spaces, but transfinitely; use WP to wellorder the space, to be able to keep choosing the 'next' linearly independent element.]

[1] There is whole book devoted to listing and proving such equivalents: *Equivalents of the Axiom of Choice* by H.Rubin & J.Rubin, *Studies in Logic* Series, North-Holland Publishing, 1963.

Exercise 5.6 Show that if C is any proper class and F any (1-1) function, then $F``C$ is a proper class.

Exercise 5.7 For sets X, Y let $\mathcal{F} = \{h \mid h \subseteq X \times Y \wedge h \text{ is a (1-1) function}\}$. Assume ZL and show that there is a $g \in \mathcal{F}$ with either $\mathrm{dom}(g) = X$ or $\mathrm{ran}(g) = Y$. Deduce that using ZL we have Cardinal Comparison, that for any sets X, Y we have either $X \preceq Y$ or $Y \preceq X$.

Exercise 5.8 Use various equivalents of WP to show that if $f : X \to Y$ is an onto function, that there is $g : Y \to X$ with $\mathrm{id} = f \circ g$.

Exercise 5.9 $(*)$ ZL is often stated in an apparently stronger form, ZL^+ : *Let $\mathcal{F}$ be a set so that for every chain $\mathcal{G} \subseteq \mathcal{F}$ then $\mathcal{G}$ has an upper bound in $\mathcal{F}$. Then $\mathcal{F}$ contains a maximal element Y.* Show that this increase in strength is indeed only apparent: $\mathrm{ZL} \Leftrightarrow \mathrm{ZL}^+$.

Exercise 5.10 Use ZL to show that for any partial order $\langle A, \preceq\rangle$ there is an extension $\preceq' \supseteq \preceq$, so that $\langle A, \preceq'\rangle$ is a total order. [Hint: (i) If $\langle A, \preceq\rangle$ is not total pick $u, v \in A$ that are $\preceq$-incomparable; let $\preceq_0 = \preceq \cup \{\langle x, y\rangle \mid x \preceq u \wedge v \preceq y\}$; check that $\preceq \subset \preceq_0$ is still a partial order; (ii) apply ZL to the set of partial orders on A. This is known as the *Order Extension Principle*.] Deduce that there is a total order $\preceq$ extending the partial order $\subseteq$ on $\mathcal{P}(\mathbb{N})$.

Exercise 5.11 Show that AC is equivalent to: every family of sets contains a maximal subfamily of disjoint sets. Formally: let $\mathrm{DF}(y) \leftrightarrow_{\mathrm{df}} \forall u, v \in y\,(u \neq v \to u \cap v = \varnothing)$. Show that

$$\mathrm{AC} \leftrightarrow \forall y \exists x \subseteq y\,(\mathrm{DF}(x) \wedge \forall z \subseteq y\,(\mathrm{DF}(z) \to x \not\subset z)).$$

Exercise 5.12 Let Φ be the statement: *for any two non-empty sets X, Y, either there exists an onto map $f : X \to Y$ or there exists an onto map $g : Y \to X$.*

(i) Show that $\mathrm{WP} \Rightarrow \Phi$.

(ii) $(*)$ Show that $\Phi \Rightarrow \mathrm{WP}$. [Hint: Consider the family of maps of a set X onto an ordinal. Use a Hartogs' like argument to show that the supremum of such ordinals exists.]

Figure 5.2: Ernst Zermelo 1871-1953

5.2.1 Weaker versions of the Axiom of Choice.

Clearly AC implies the following:

Definition 5.3 (AC_ω – the countable axiom of choice) *Every countable family of non-empty sets has a choice function.*

But we cannot assume AC_ω and hope that it implies the general AC .

Theorem 5.4 *Assume* AC_ω. *Then (i) the union of countably many countable sets is countable. (ii)* (Russell & Whitehead 1912) *Every infinite set has a countably infinite subset.*

Definition 5.5 DC_ω. *Let R be a relation on a set A with the property that for any $u \in A$ there is $b \in A$ with bRa. Then there is a sequence of elements $\{u_i \mid i \in \omega\}$ of A with $u_{i+1} R u_i$ for all $i \in \omega$.*

It can be shown that $DC_\omega \Rightarrow AC_\omega$ (Bernays 1952) but not conversely (Jensen 1966). A very large part of contemporary analysis, indeed mathematics, can be done assuming only DC_ω and not the full AC or WP.

The Wellfounded Universe of Sets

At the very start of this course we introduced a picture of the universe of sets of mathematical discourse, which we dubbed V. The idea was that we could start with the empty set and build up a hierarchy of sets that would be sufficient for all of mathematics. We defined $V_0 = \varnothing$, and then $V_{n+1} = \mathcal{P}(V_n)$. The suggestion was that this idea would be continued into the transfinite. Now that we have a theory of ordinals, and theorems concerning the possibility of definitions along all the ordinals by transfinite recursion, we can make complete this picture.

Definition 6.1 **(The Wellfounded hierarchy of sets)** *We define the V_α function by transfinite recursion as:*

$$V_0 = \varnothing; \ \ V_{\alpha+1} = \mathcal{P}(V_\alpha); \mathrm{Lim}(\lambda) \to V_\lambda = \textstyle\bigcup_{\alpha<\lambda} V_\alpha; \ \ \textit{we set } V = \bigcup_{\alpha\in\mathrm{On}} V_\alpha.$$

Lemma 6.2 *For any α: (i)* $\mathrm{Trans}(V_\alpha)$
(ii) $\beta < \alpha \to V_\beta \in V_\alpha$ and hence by (i), $V_\beta \subseteq V_\alpha$.

Proof: Use transfinite induction on α: $\alpha = 0$ is trivial; if $\alpha = \beta + 1$ then $\mathrm{Trans}(X) \longrightarrow \mathrm{Trans}(\mathcal{P}(X))$ (see Exercise 1.19), thus $\mathrm{Trans}(V_\beta)$ implies $\mathrm{Trans}(V_{\beta+1})$. Then $V_\beta \in V_\alpha$ and so $V_\beta \subseteq V_\alpha$ by the latter's transitivity. If $\beta' < \beta$, then also $V_{\beta'} \in V_\beta$ by the Ind. Hyp., so $V_{\beta'} \in V_\alpha$. If $\mathrm{Lim}(\alpha)$ then as a union of transitive sets is transitive (Exercise 1.19(iii)), so $\mathrm{Trans}(V_\alpha)$ is immediate from the definition of V_α as $\bigcup_{\beta<\alpha} V_\beta$. If $\beta < \gamma < \alpha$ then by inductive hypothesis $V_\beta \in V_\gamma$. We thus have $V_\beta \in \bigcup_{\gamma<\alpha} V_\gamma = V_\alpha$. Q.E.D.

Definition 6.3 **(The rank function)** *For any $x \in V$ we let:*

$$\rho(x) = \textit{the least } \tau \textit{ so that } x \subseteq V_\tau.$$

Note that by the definition of $V_{\tau+1}$ we could just as easily have defined rank by setting $\rho(x)$ to be the least τ so that $x \in V_{\tau+1}$. (If we think of sets being formed as we ascend the V_α-hierarchy, then once all elements of a set x have appeared, say by stage τ, then x will be an element of $V_{\tau+1}$ - as the latter consists of all possible subsets of V_τ. Notice also that if $y \in x$ then $\rho(y) < \rho(x)$. As the ordinals are wellordered, this means that the $\in$-relation is a *wellfounded relation* on V (Why?).

Examples If $x, y \in V_\alpha$ then: $\{x\}, \{x, y\} \in \mathcal{P}(V_\alpha) = V_{\alpha+1}$. Hence $\langle x, y\rangle = \{\{x\}, \{x, y\}\} \in \mathcal{P}(V_{\alpha+1}) = V_{\alpha+2}$. Hence if $\rho(x) = \rho(y) = \alpha$ then $\rho(\{x, y\}) = \alpha + 1$, and $\rho(\langle x, y\rangle) = \alpha + 2$.

Hence $V_\alpha \times V_\alpha \subseteq V_{\alpha+2}$ and so $V_\alpha \times V_\alpha \in V_{\alpha+3}$. As any ordering R on V_α is a subset of $V_\alpha \times V_\alpha$ we have $R \subseteq V_{\alpha+2}$ as well, and so is also in $V_{\alpha+3}$. So $\rho(R) \leq \alpha + 2$.

Exercise 6.1 Compute (i) $\rho(S(x))$ in terms of $\rho(x)$. (ii) Show that $\rho(\bigcup x) \leq \rho(x)$, and give examples of sets x_1, x_2 with $\rho(\bigcup x_1) < \rho(x_1)$ but $\rho(\bigcup x_2) = \rho(x_2)$; can you characterise those sets z for which $\rho(\bigcup z) < \rho(z)$? (iii) Suppose $\rho(x) = \rho(y) = \alpha$ and $f : x \to y$. Compute $\rho(\langle x, y, x\rangle)$; $\rho(f)$; $\rho({}^x y)$; $\rho({}^\alpha x)$.

Exercise 6.2 What if α in the above example is a limit ordinal? Can we improve the bounds on ranks? If $\langle\omega, R\rangle$ is an ordering, what is $\rho(R)$? [Hint: compute $\rho(\omega\times\omega)$, and $\rho((\omega+1)\times(\omega+1))$.]

It is so useful to have sets organised in this hierarchical fashion that we adopt from now on one last axiom:

Axiom of Foundation: *Every set is wellfounded, that is,* $\forall x\,(x\neq\varnothing\to\exists y\in x\,(y\cap x=\varnothing))$.

Notice that such a y in the statement of the axiom, is an $\in$-minimal element of x: there cannot be any $z\in x$ which is also in y. We may thus paraphrase the Axiom of Foundation by saying that "every non-empty set x has an $\in$-minimal element". We thereby rule out by fiat the existence of sets such as x and y with the properties that $x\in x$, or $x\in y\in x$, because for such a "set", whatever "$\in$" means it is not a wellfounded relation on x. Consequently since we do adopt this axiom, we have that $\in$- is a wellfounded relation on every set, and every set appears somewhere in the V_α-hierarchy. Some texts write WF for the class of wellfounded sets in the V_α-hierarchy, prove a lemma such as 6.4 for WF, and then later introduce the Axiom of Foundation.

Lemma 6.4 *The following are equivalent: (i) The Axiom of Foundation;*

(ii) $\forall x\exists\alpha(x\in V_\alpha)$;

(iii)$\forall x\exists\alpha\,(x\subseteq V_\alpha)$.

Proof: Assume (i). We prove (ii). Let x be any set. First note that if $\mathrm{TC}(x)\subseteq V$ then for some ordinal α, $\mathrm{TC}(x)\subseteq V_\alpha$ [ρ"$\mathrm{TC}(x)$ is a set of ordinals by Ax. Replacement, and so for some α ρ"$\mathrm{TC}(x)\subseteq\alpha$]. However then we are done, since both $x\subseteq\mathrm{TC}(x)$ are elements of $V_{\alpha+1}$. Suppose $\mathrm{TC}(x)\backslash V\neq\varnothing$. Then let y be in this set, but such that $y\cap(\mathrm{TC}(x)\backslash V)=\varnothing$ by Ax. Foundation. Then any $z\in y$ is in $\mathrm{TC}(x)$ and by assumption then, $z\in V$. So $y\subseteq V$. Again ρ "y is a set of ordinals. So for some β, $y\subseteq V_\beta$. But then $y\in V_{\beta+1}$ contradicting the choice of y. (ii) $\Rightarrow$ (iii): note that the least α with $x\in V_\alpha$ is always a successor ordinal, $\alpha'+1$ say; but then $x\subseteq V_{\alpha'}$. (iii) $\Rightarrow$ (i) is also trivial: note if $x\subseteq V_\alpha$, then $\rho:\langle x,\in\rangle\to\langle\alpha,<\rangle$ is an order preserving map. Hence any element $z_0\in x$ with $\rho(z_0)$ least amongst $\{\rho(z)\mid z\in x\}$ is $\in$-minimal in x, that is $z_0\cap x=\varnothing$. Thus $\langle x,\in\rangle$ is wellfounded.

Exercise 6.3 Show that the Axiom of Foundation implies the apparently stronger statement that for any *class* $(A\neq\varnothing\to\exists y\in A\,(y\cap A=\varnothing))$.

Is the Axiom of Foundation justified? Perhaps there are mathematical objects that cannot be represented by sets or structures in V? If so this would destroy our claim that the set theory of V provides a sufficient foundation for all of mathematics. In fact this turns out not to be the case: if we assume AC we can prove that *every* structure that mathematicians invent can be seen to have an isomorphic copy in V - and since mathematicians only worry about truths in mathematical structures "up to isomorphism" this will do for us.[1]

[1]There should be a slight caveat here: some category theorists deal with proper class sized objects because they wish to work with the "category of all groups", or the "category of all sets", but there are ways of dealing also with these notions, so the spirit of the claim is true.

Exercise 6.4 Let $\mathbb{G} = \langle G, \circ, e, ^{-1} \rangle$ be a group. Assume WP, but not the Axiom of Foundation. Show that there is a group $\widetilde{\mathbb{G}} \in V$ with $\mathbb{G} \cong \widetilde{\mathbb{G}}$. [Hint: By WP find R so that $\langle G, R \rangle \in$ WO. Then "copy" $\mathbb{G}$ onto the domain $\alpha = \mathrm{ot}(\langle G, R \rangle)$.]

Exercise 6.5 Let Φ be the proposition "*There is no sequence of sets x_i for $i \in \omega$, with $x_{i+1} \in x_i$*". a) Show that the Axiom of Foundation implies Φ; b) WP together with Φ implies the Axiom of Foundation .

We now prove some properties about this hierarchy.

Lemma 6.5 *(i)* $V_\alpha = \{x \in V \mid \rho(x) < \alpha\}$;
(ii) If $x \in V$ then $\forall y \in x (y \in V \wedge \rho(y) < \rho(x))$;
(iii) If $x \in V$, then $\rho(x) = \sup\{\rho(y) + 1 \mid y \in x\} = \sup^+\{\rho(y) \mid y \in x\}$;

Proof:
For *(i)*: If $x \in V$, then $\rho(x) < \alpha \Leftrightarrow_{\mathrm{df}} \exists \beta < \alpha (x \subseteq V_\beta)) \Leftrightarrow \exists \beta < \alpha (x \in V_{\beta+1}) \Leftrightarrow x \in V_\alpha$ (by Lemma 6.4*(ii)*).
For *(ii)*: Let $\alpha = \rho(x)$. Then $x \subseteq V_\alpha$. So if $y \in x$ then $y \in V_\alpha$ and so $\rho(y) < \alpha$ by *(i)*.
For *(iii)*: Notice the second equality follows by definition of $\sup^+$. Let $\alpha = \sup^+\{\rho(y) \mid y \in x\}$. By *(ii)* if $y \in x$ then $\rho(y) < \varrho(y) + 1 \leq \rho(x)$, thus $\alpha \leq \rho(x)$. Again by *(i)* for each $y \in x$, $\rho(y) < \varrho(y) + 1 \leq \alpha$ implies $y \in V_\alpha$; so $x \subseteq V_\alpha$, *i.e.* $\rho(x) \leq \alpha$. Q.E.D.

- Note in *(iii)*, that now we may write $\rho(x) = \sup^+\{\rho(y) \mid y \in x\}$.

Lemma 6.6 *(i)* $\rho(\alpha) = \alpha$; *(ii)* $\mathrm{On} \cap V_\alpha = \alpha$.

Proof: Assume by induction for *(i)* that $\beta < \alpha \longrightarrow \rho(\beta) = \beta$. But then by Lemma 6.5 *(iii)* $\rho(\alpha) = \sup^+\{\beta \mid \beta < \alpha\} = \alpha$.

For *(ii)*: *(i)* here shows $(\supseteq)$; and $(\subseteq)$ is immediate from *(i)*, Lemma 6.5*(i)* and the inductive hypothesis. Q.E.D.

So we have a picture of sets, V, in which as an object x lives at a certain rank on the V_α-hierarchy, and its members $y \in x$ live below that at lesser levels, and in turn whose members $u \in y$ live below $\rho(y)$ and so forth.

Exercise 6.6 Show that if $\pi : \langle V, \in \rangle \to \langle V, \in \rangle$ is an isomorphism, then $\pi = \mathrm{id}$. There are thus no non-trivial isomorphisms of V with itself. [Hint: Suppose there was an x with $\pi(x) \neq x$. Choose one such x of least rank with this property. Then $y \in x \to \pi(y) = y$.] (This both generalises Cor. 3.7 and is a special case of: if $f : \langle M, R \rangle \to \langle M, R \rangle$ is an isomorphism, where $\langle M, R \rangle$ is a wellfounded relation, then $f = \mathrm{id}$.)

We can thus think of a set x as given by a graph or picture of "nodes" in a certain kind of tree where we go downwards in the $\in$-relation as we descend the tree. The tree will most likely have infinitely many nodes, and any one node may have infinitely many members immediately below it, but what it does not have is any infinitely long downwards growing branches: this is because every level of a node comes with an ordinal denoting the rank of the set attached at that point, and we can have no infinite descending chains through the ordinals. This idea provides us with a new way of defining functions or proving properties about sets: since $\in$ is wellfounded we have:

LEMMA 6.7 **Principle of $\in$-induction** *Let $\Phi(v)$ be any welldefined and definite property of sets.*

(i) (Set Form) Let $\mathrm{Trans}(X)$. *Then*

$$\forall y \in X\,((\forall x \in y\Phi(x)) \to \Phi(y)) \quad \to \quad \forall y \in X\Phi(y).$$

(ii) (V or Class form)

$$\forall y\,((\forall x \in y\Phi(x)) \to \Phi(y)) \quad \to \quad \forall y\Phi(y).$$

Proof: (i) Let $Z =_{df} \{y \in X \mid \neg\Phi(y)\}$. We prove the contrapositive and suppose $Z \neq \varnothing$ and we shall show that the antecedent of the induction scheme fails. Let y_0 be $\in$-minimal in Z (by appealing to the Axiom of Foundation). Then for any $x \in y_0$ we have $x \in X$. (Since $\mathrm{Trans}(X)$),). Hence $\Phi(x)$ holds for such x. Suppose $\forall y \in x[(\forall x \in y\,\Phi(x)) \to \Phi(y)]$ were true (for a contradiction). However we have just argued that $\forall x \in y_0\,\Phi(x)$. If this were true we'd conclude $\Phi(y_0)$- a contradiction! This finishes (i).

(ii) Notice this is exactly the same, thinking of X as the transitive class V! Instead now take $Z =_{df} \{y \mid \neg\Phi(y)\}$. The rest of the argument makes perfect sense. Q.E.D.

THEOREM 6.8 **($\in$-Recursion Theorem)** *Let $G : V \to V$ be any function. Then there is exactly one function $H : V \to V$ so that*

$$\forall x H(x) = G(H \restriction x) \quad [= G(\{\langle y, H(y)\rangle \mid y \in x\})].$$

Proof This is done in just the same format as Theorem 3.32 - the Recursion Theorem for On. As before we shall define H as a union of *approximations* where now u is an *approximation* if (a) $\mathrm{Func}(u)$, $\mathrm{Trans}(\mathrm{dom}(u))$, and (b) $\forall w \in \mathrm{dom}(u)(\, u(w) = G(u \restriction w))$. We call it an *$x$-approximation*, if additionally $x \in \mathrm{dom}(u)$. So u satisfies the defining clauses of H throughout its domain. Note for later that $\mathrm{TC}(\{x\}) \subseteq \mathrm{dom}(u)$ for any x-approximation u. Further if u is an x-approximation then the $u \restriction \mathrm{TC}(\{x\})$ is an x-approximation, and indeed is the minimal such. Lastly we may extend an approximation u in the following way: let $z \subseteq \mathrm{dom}(u)$ but $z \notin \mathrm{dom}(u)$. Then $\mathrm{Trans}(\mathrm{dom}(u) \cup \{z\})$, so we may set $v = u \cup \{\langle u, G(u)\rangle\}$.

(1) *If u and v are approximations, and we set $y = \mathrm{dom}(u) \cap \mathrm{dom}(v)$ then $u \restriction y = v \restriction y$ and is an approximation.*

Proof: Note that $\mathrm{Trans}(y)$ as the intersection of any two transitive sets is transitive. Suppose we have shown that for some $x \in y$ that $\forall z \in x(u(z) = v(z))$. Then $u \restriction x = v \restriction x$; but then $u(x) =_{df} G(u \restriction x) = G(v \restriction x) =_{df} v(x)$! We thus have shown

$$\forall x \in y(\forall z \in x(u(z) = v(z) \to u(x) = v(x)))$$

By the (set form of the) Principle of $\in$-induction applied to the transitive set $X = y$ we conclude that $\forall x \in y(u(x) = v(x))$, and we are done.

(2) *(Uniqueness) If H exists then it is unique.*

Proof: This is really the same as before but we repeat the detail: if H, H' were two such functions defined on all of V, there would be an $\in$-least set z on which they disagreed. Note that z cannot be $\varnothing$. Let $x = \mathrm{TC}(\{z\})$. So then $H \restriction x \neq H' \restriction x$ are two *different* x-approximations, which is impossible by (i).

(3) *(Existence). Such an H exists.*

Proof: Let $u \in B \Leftrightarrow \{u \mid u \text{ is an approximation}\}$. B is in general a proper class of approximations, but this does not matter as long as we are careful. As any two such approximations agree on the common transitive part of their domain, we define $H = \bigcup B$. Just as for the proof of recursion on ω:

(i) H is a function;

(ii) $\mathrm{dom}(H) = V$.

Proof: We use the principle of $\in$-induction. It suffices to show then that $\forall z(\forall y \in z(y \in \mathrm{dom}(H)) \to z \in \mathrm{dom}(H))$.

Let C be the class of sets z for which there is no z-approximation. So if we suppose for a contradiction that C is non-empty, by the Principle of $\in$-Induction, then it will have an $\in$-minimal element z such that $\forall y \in z \exists u(u$ is a y-approximation). By the remark in the first paragraph of this proof, any y-approximation restricts to a y-approximation with domain $\mathrm{TC}(\{y\})$. So now we let h be the function

$$\bigcup\{h_y \mid h_y \text{ is a } y\text{-approximation} \wedge y \in z \wedge \mathrm{dom}(h_y) = \mathrm{TC}(\{y\})\}.$$

By the above these functions h_y all agree on the parts of their domains they have in common. Note that the domain of h is a transitive set, being the union of transitive sets $\mathrm{dom}(h_y)$ for $y \in z$. Hence $z \subseteq \mathrm{dom}(h)$ and thus $\{z\} \cup \mathrm{dom}(h)$ is transitive. As noted just before (1) we can thus extend h to $h' = h \cup \{\langle z, G(h \upharpoonright z)\rangle\}$ and h' is then a z-approximation. However we assumed that $z \in C$! A contradiction. Hence $C = \varnothing$ and (ii) holds. Q.E.D.

Exercise 6.7 Show for any x that $\rho(x) = \rho(\mathrm{TC}(x))$.

Exercise 6.8 Let X be any set. Show that $\mathrm{Trans}(X) \to \rho\,``X \in \mathrm{On}$.

Exercise 6.9 Does $\mathrm{Trans}(X) \wedge X \neq \varnothing$ imply that $\varnothing \in X$?

Exercise 6.10 Show that for all α $|V_{\omega+\alpha}| = \beth_\alpha$.

Exercise 6.11 $(*)$ We say that a function $j : V \to V$ is an *elementary embedding* if it preserves the truth about objects. In other words if $\varphi(v_0, \ldots, v_n)$ is a formula expressing a property, and $x_0, \ldots, x_n$ are sets; then

$$\varphi(x_0, \ldots, x_n) \leftrightarrow \varphi(j(x_0), \ldots, j(x_n)).$$

If we assume the axioms of set theory (but not AC) our current state of knowledge allows the possibility that such a class function j could exist which is not the identity (so $j(x) \neq x$ for some $x \in V$). Show if there is such a j then for some ordinal α, $j(\alpha) \neq \alpha$. [Hint: Consider the formula "$u = \mathrm{rk}(v)$".] (It is known by a result of K. Kunen that AC rules out the existence of such a j.)

Index of Symbols

Index