PMA603: Set Theory — 1998-99
by P. G. Dixon

Contents

1 Introduction

Question: are Fourier series unique?

To be precise: let $(a_n), (b_n)$ be sequences of complex numbers such that

$$\sum_{n=-\infty}^{\infty} a_n e^{int} = \sum_{n=-\infty}^{\infty} b_n e^{int} \tag{1}$$

for all $t \in [0, 2\pi)$, with both series converging; does it follows that $a_n = b_n$ for all n? Cantor (1870) showed that the answer is affirmative and he went on to ask: what happens if you only know (1) for all t lying outside some small exceptional set E? He showed that the answer is again affirmative for countable closed sets E. (This was later extended by W. H. Young (1909) to arbitrary countable sets.) It was through this work that Cantor came to create what we now know as infinite set theory.

We shall leave the Fourier series question here . A reference which continues that story is: Alexander S. Kechris and Alain Louveau, *Descriptive set theory and the structure of sets of uniqueness,* (Cambridge University Press, LMS Lecture Note Series **128**, 1987) ISBN 0-521-35811-6 367pp.

We shall be sketching a development of set theory from the axioms of Zermelo and Fraenkel. We assume that everyone understands what is meant by the word 'set' (actually it holds the record for the word with the most meanings listed in the Oxford English Dictionary — 464 entries!). Cantor's definition (1895) of the word set (= Menge) was as follows.

> 'Unter einer "Menge" verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objecten m unsrer Anschauung oder unseres Denkens (welche die "Elemente" von M gennant werden) zu einen Ganzen.'

Roughly,

> 'By a "class" we understand a combination (summary) M of definite, well-distinguished objects m of our intuition or of our thought (which are called "elements" of M) into one whole.'

But E. T. Bell comments helpfully that

> 'No two of a dozen mathematicians and scientists bilingual in English and German agreed on the meaning of this definition; two said it was meaningless. This definition is one source of trouble in the foundations of mathematics.'

We shall modestly take the view that set theory encompasses all of mathematics. But, set of what? Sets of sets, of course! The only things which exist in our universe are sets. We shall see that this is not as stupid as it sounds.

Our theory then is about certain undefined objects called 'sets' between any two of which a certain relation, denoted $\in$ and called 'membership' may or may not hold. We now present the axioms.

2 The axioms of Zermelo–Fraenkel set theory

To formally develop ZF set theory, one must start with formal logic: the Propositional Calculus (the study of the connectives $\&$, $\vee$ (or), $\neg$ (not), $\Rightarrow$, $\iff$), the Predicate Calculus ($\forall x P(x)$, $\exists x P(x)$), the Predicate Calculus with Equality (distinguished predicate $E(x, y)$ written $x = y$). The theory ZF is developed from the Predicate Calculus with Equality by adding the distinguished binary predicate $\in$. The variables are thought of as sets and $x \in y$ is to mean 'x is a member of y'. In this informal introduction to the formal theory, we shall assume that the reader has a vague notion of what the axioms and theorems of the Predicate Calculus with Equality should be, based on his/her long experience of reasoning with statements involving the logical notions mentioned above. We shall state the axioms precisely, but we shall not attempt full formal proofs.

We write $x \notin y$ as an abbreviation for $\neg(x \in y)$.

We write $\exists!$ to mean 'there exists a unique…', i.e. $\exists! x Q(x)$ is an abbreviation for

$$(\exists x Q(x)) \,\&\, \forall x \forall y (Q(x) \,\&\, Q(y) \;\Rightarrow\; x = y).$$

The Axiom of Extensionality

This says that a set is determined by its *extension* — that is, by the elements it contains.

$$(AE) \qquad \forall z (z \in x \iff z \in y) \;\Rightarrow\; x = y.$$

This axiom contains free variables x, y. There is an implicit $\forall x \forall y$ applying to the whole axiom. Note that the converse implication

$$x = y \;\Rightarrow\; \forall z (z \in x \iff z \in y)$$

is a consequence of one of the axioms of our underlying logic (the Predicate Calculus with Equality); to be precise, it is an instance of the axiom scheme

$$x = y \;\Rightarrow\; (P(x) \iff P(y))$$

where P is any predicate.

The major consequence of this is that there is at most one set with no elements (the fact that such a set exists is axiom (AN) below). An alternative would be to allow lots of such 'atomic' sets (or *urelemente*). The earliest proofs of the independence of the Axiom of Choice (the Fraenkel–Mostowski models) worked with a set theory allowing infinitely many urelemente.

The Subset Axiom

This is really an *axiom scheme*: one axiom for each predicate $P(.)$.

$$(AS) \qquad \exists y \forall z (z \in y \iff z \in x \,\&\, P(z));$$

i.e. the set

$$\{z \in x : P(z)\}$$

exists when x exists.

This is a cut-down version of the axiom scheme we should like to have, namely that, for each predicate $P(.)$, the set $\{z : P(z)\}$ exists; formally,

$$(AS') \qquad \exists y \forall z (z \in y \iff P(z)).$$

The reason we cannot have this scheme is that it leads to a contradiction.

Russell's Paradox

Suppose (AS′) and let $P(z) \equiv (z \notin z)$, so

$$\exists y \forall z (z \in y \iff z \notin z).$$

In particular, for $z = y$ we get

$$z \in z \iff z \notin z.$$

which is impossible.

The restrictive nature of (AS) is apparent when one considers how often in mathematics one wishes to define sets just by a property: 'the set of all groups', 'the set of all topological spaces', *et cetera.* There can be no 'set of all sets', for if there were such a set V, then $\{x \in V : x \notin x\}$ would be a set by (AS) and would give us Russell's Paradox. We can get round this by saying that the object $Y = \{z : P(z)\}$ is a *class*, the class of all sets z such that $P(z)$. A class differs from a set in that although you can talk about sets being members of classes ($z \in Y$ simply means $P(z)$), you cannot necessarily talk of a class being a member of a set. Of course, some classes will correspond to sets, but others (*proper classes*) will not. We shall, in the future use proper classes with the assumption that if a proper class has been defined by, say, $A = \{x : P(x)\}$, then $a \in A$ is simply an *abbreviation* for $P(a)$. These abbreviations must be expanded in order to reconstruct the true formulae of ZF. (The same will be true of all definitions of new symbols.)

Here, we are expounding Zermelo–Fraenkel (ZF) set theory; an alternative axiomatization of set theory due to von Neumann, Bernays and Gödel (NBG set theory) has classes as the objects of the theory and a distinguished unary predicate 'x is a set', with axioms ensuring that only sets can be members of classes. The other noteworthy feature of NBG set theory is that it has a finite set of axioms and no axiom schemes. Details may be found in Gödel's book on the consistency of the Continuum Hypothesis [1] and in Mendelson's textbook [2].

The Axiom of Replacement

This is again an axiom scheme. To state it succinctly, we first introduce an definition. (Notice that definitions in an axiomatic theory are simply abbreviations.)

Definition 2.1 For a predicate $P(.)$ we write $\mathcal{M}_x (P(x))$ for $\exists y \forall x (P(x) \Rightarrow x \in y)$.

Thus $\mathcal{M}_x (P(x))$ means 'there is a set which contains every x such that $P(x)$. By (AS) this implies that the class $\{x : P(x)\}$ is a set.

$$(AR) \qquad \forall x \exists y \forall z (P(x,z) \iff z \in y) \Rightarrow \forall u \mathcal{M}_y (\exists x (x \in u \,\&\, P(x,y)))\,.$$

The hypothesis is that $P(x, y)$ defines a set-valued 'function' $F : x \mapsto y = \{z : P(x, z)\}$. (The inverted commas around the word 'function' are essential: we shall later define the notion of function and it will be a *set* of ordered pairs. The thing we have here is a proper class.) The conclusion is that, given a set u, the class $\bigcup_{x \in u} F(x)$ is a set. (Actually, it only says that there is a set containing this class but, as noted above, (AS) then implies that the class itself is a set.)

The Power Set Axiom

[1] K. Gödel, 'The consistency of the continuum hypothesis' (Princeton University Press, Annals of Mathematics Studies, **3**, 1940) ML 3 PER 510.5

[2] E. Mendelson, 'Introduction to mathematical logic' (van Nostrand, The University Series in Undergraduate Mathematics, 1963).

Definition 2.2 We write $x \subseteq y$ for $\forall z(z \in x \ \Rightarrow \ z \in y)$ and $x \subset y$ for $(x \subseteq y \ \& \ x \neq y)$.

The Power Set Axiom is then just

$$(AP) \qquad \forall x \mathcal{M}_y\,(y \subseteq x)\,.$$

Thus, for every set x the class $\{y : y \subseteq x\}$ is a set. This set is unique (by (AE)); we call it the *power set* $\mathcal{P}(x)$ of x

The Power Set Axiom is very useful in combination with (AS): for example, if x is a set, so is $\{\{y\} : y \in x\}$ because

$$\mathcal{M}_z\,(z \subseteq x \ \& \ \exists u(u \in z) \ \& \ \forall u \forall v(u \in z \ \& \ v \in z \ \Rightarrow \ u = v))\,.$$

The Axiom of the Null Set

The axioms so far can be satisfied if no sets exist! We rectify this.

$$(AN) \qquad \exists x \forall y(y \notin x).$$

This set x is unique, by (AE); we call it the *null set* or *empty set* and denote it $\emptyset$.

Having brought this set into being, we have quite a number of sets derived from it. By (AP), we have, successively,

$$\{\emptyset\},$$
$$\{\emptyset, \{\emptyset\}\},$$
$$\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\},$$

et cetera , and by (AS) we have all the subsets of these. However, all of these are finite.

The Axiom of Infinity

Next, we need an axiom to say that infinite sets exist. The following axiom, which subsumes the Axiom of the Null Set, does that.

$$(AI) \qquad \exists x(\exists y(y \in x \ \& \ \forall z(z \notin y)) \ \& \ \forall y(y \in x \ \Rightarrow \ \exists z(z \in x \ \& \ \forall w(w \in z \iff w \in y \lor w = y)))).$$

A definition will help us make sense of this.

Definition 2.3 For a set x we define

$$S(x) := \{y : y \in x \lor y = x\},$$

i.e. ,

$$S(x) = x \cup \{x\}.$$

To see that $S(x)$ is a set looks a little complicated. As it is defined as a union, we need (AR). Define $P(a, b)$ to be true if $a = \emptyset$ and $b = x$ or if $a = \{\emptyset\}$ and $b = \{x\}$ and false otherwise. Then apply (AR) with $u = \{\emptyset, \{\emptyset\}\}$ to get that $x \cup \{x\}$ is a set. Formally,

$$P(a, b) \equiv ((\forall z(z \notin a)) \ \& \ b = x) \lor ((\forall z(z \in a \iff \forall w(w \notin z))) \ \& \ (\forall z(z \in b \iff z = x))).$$

We can now state the Axiom of Infinity briefly:

$$(AI) \qquad \exists x(\emptyset \in x \ \& \ \forall y(y \in x \ \Rightarrow \ S(y) \in x)).$$

Thus (AI) asserts the existence of a set x which contains the following as elements.

Definition 2.4 The natural numbers:

$$\begin{array}{rcl} 0 & := & \emptyset, \\ 1 & := & S(0) = \{0\} = \{\emptyset\}, \\ 2 & := & S(1) = \{0, 1\} = \{\emptyset, \{\emptyset\}\}, \\ 3 & := & S(2) = \{0, 1, 2\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \\ 4 & := & S(3) = \{0, 1, 2, 3\}, \\ \ldots & & \end{array}$$

The Axiom of Foundation or Axiom of Regularity

$$(AF) \qquad \exists y(y \in x) \;\Rightarrow\; \exists y(y \in x \,\&\, \neg\exists z\,(z \in x \,\&\, z \in y))\,.$$

In conventional language: every non-empty set x has an element y which is disjoint from x. Equivalently, it is not possible to have a non-empty set x such that each element y of x has an element z which is also an element of x.

If, on the contrary, this were possible, we could define x_1 to be any element $y \in x$ and $x_2 = z \in x \cap x_1$, as above. This would start a sequence $x_1, x_2, x_3, \ldots$ of elements of x with

$$x \ni x_1 \ni x_2 \ni x_3 \ni \ldots.$$

Having found such $x_1, \ldots, x_n \in x$, we could choose $x_{n+1} \in x \cap x_n$ and proceed by induction.

Conversely, suppose we had such a descending sequence. Then $\{x, x_1, x_2, \ldots\}$ is a set, no element of which is disjoint from the whole set. Thus (AF) is a way of excluding such infinite descending sequences without having to talk about sequences and hence natural numbers, functions, *et cetera* .

A particular consequence of this axiom is the following proposition.

Proposition 2.5 $\forall x(x \notin x)$.

Proof. Suppose otherwise, that we have an x with $x \in x$. Now apply (AF) to the non-empty set $\{x\}$. It says that there is a set $y \in \{x\}$ such that no member of y is a member of $\{x\}$. But $y \in \{x\}$ means $y = x$ and so $x \in y$ and $x \in \{x\}$; contradiction. $\Diamond$

This completes our list of the axioms of ZF. There are other axioms we shall need to add to it: the Axiom of Choice, the Axiom of Determinateness, the Continuum Hypothesis, large cardinal axioms, the Axiom of Constructibility, but these have always been regarded as less secure than the basic axioms. Indeed the additional axioms we have just listed are mutually incompatible. We shall now develop as much set theory as we can using ZF alone.

3 Development

Definitions are essentially just abbreviations. In some cases, this is straightforward. We have already met

$$x \subseteq y \quad \equiv \quad \forall z(z \in x \;\Rightarrow\; z \in y).$$

It is a simple matter to expand out this abbreviation.

Others are more complicated: thus if A is a proper class we have to find a definition of A as $\{x : P(x)\}$ in order to expand $a \in A$ as $P(a)$. Note that if $\forall x(P(x) \iff Q(x))$ then $A = \{x : P(x)\} = \{x : Q(x)\}$ and $a \in A$ may be rendered by the equivalent formula $Q(a)$.

Definitions of sets of the form $a := \{x : P(x)\}$ rely on an existence statement $\mathcal{M}_x\,(P(x))$ having been proved. To expand out such an abbreviation, a formula $Q(a)$ becomes

$$\forall a(\forall x(x \in a \iff P(x)) \;\Rightarrow\; Q(a)).$$

More generally, if we define $a := \iota x P(x)$ (the x such that $P(x)$), having first proved that there is one and only one x such that $P(x)$, then $Q(a)$ becomes

$$\forall x(P(x) \;\Rightarrow\; Q(x)).$$

These considerations apply to our definition $\emptyset := \iota x(\forall y(y \notin x))$, whose existence is guaranteed by (AN) and uniqueness by (AE).

We use the notation $\{x \in s : P(x)\}$ as an abbreviation for $\{x : x \in s \,\&\, P(x)\}$. This is the safest form of definition since it automatically produces a set by (AS).

Definitions 3.1

$$\begin{aligned}
\{x\} &:= \{y : y = x\};\\
\{x, y\} &:= \{z : z = x \vee z = y\};\\
\bigcup x &:= \{z : \exists y(z \in y \ \&\ y \in x)\};\\
\bigcap x &:= \{z : \forall y(y \in x \ \Rightarrow \ z \in y)\};\\
x \cup y &:= \bigcup\{x, y\} = \{z : z \in x \vee z \in y\};\\
x \cap y &:= \bigcap\{x, y\} = \{z : z \in x \ \&\ z \in y\};\\
x \setminus y &:= \{z : z \in x \ \&\ z \notin y\}.
\end{aligned}$$

Note that $\{x, x\} = \{x\}$.

We shall define functions by identifying the function $f : x \to y$ with its graph, which is a subset of $x \times y$. To define $x \times y$ we need the concept of an *ordered pair*.

Definition 3.2 (Kuratowski)

$$(x, y) := \{\{x\}, \{x, y\}\}.$$

Note that $(x, x) = \{\{x\}\}$.

We check the basic property of ordered pairs.

Proposition 3.3

$$(x, y) = (u, v) \iff x = u \ \&\ y = v.$$

Proof. The proof that $x = u \ \&\ y = v \ \Rightarrow \ (x, y) = (u, v)$ is easy, using (AE) repeatedly.

Conversely, if $(x, y) = (u, v)$ then we have

$$\{x\} \in (x, y) = (u, v) = \{\{u\}, \{u, v\}\},$$

so $\{x\} = \{u\}$ or $\{x\} = \{u, v\}$. Hence $u \in \{u\} = \{x\}$ or $u \in \{u, v\} = \{x\}$. Therefore $u = x$.

Also

$$\{x, y\} \in (x, y) = (u, v) = \{\{u\}, \{u, v\}\} = \{\{x\}, \{x, v\}\},$$

so $\{x, y\} = \{x\}$ or $\{x, y\} = \{x, v\}$. Hence $y \in \{x, y\} = \{x\}$ or $y \in \{x, y\} = \{x, v\}$. Therefore $y = x$ or $y = v$.

Interchanging the rôles of (x, y) and (u, v), we have that $v = u$ or $v = y$. Thus, either $y = v$, or we have $y = x$ and $u = v$, i.e. $y = x = u = v$.

$\Diamond$

Definition 3.4 For any two sets x, y we want to define their *Cartesian product*, which we want to be a *set*. We must use (AS) to define it as a subset. Now if $u \in x$ and $v \in y$, then $\{u\}$ and $\{u, v\}$ are elements of $\mathcal{P}(x \cup y)$, so $(u, v) \in \mathcal{P}(\mathcal{P}(x \cup y))$. We define

$$x \times y := \{(u, v) \in \mathcal{P}(\mathcal{P}(x \cup y)) : u \in x \ \&\ v \in y\}.$$

Definition 3.5 The ternary predicate $\mathrm{Funct}(f, x, y)$, henceforth written $f : x \to y$ meaning 'f is a function from the set x into the set y' is defined as

$$f \subseteq x \times y \ \&\ \forall u(u \in x \ \Rightarrow \ \exists v(v \in y \ \&\ (u, v) \in f)) \ \&\ \forall u \forall v \forall w((u, v) \in f \ \&\ (u, w) \in f \ \Rightarrow \ v = w)$$

We call x the *domain* and y the *codomain* of f. If $z \subseteq x$, we write $f \restriction z$ for $\{(u, v) \in f : u \in z\}$, the *restriction* of f to z.

Definitions 3.6 Given a function $f : x \to y$ and $u \in x$ we define $f`u$ by

$$f`u := \iota v((u,v) \in f).$$

Given $f : x \to y$ and $u \subseteq x$ we define $f``u$ by

$$\begin{aligned} f``u &:= \{v \in y : \exists w(w \in u \,\&\, (w,v) \in f)\} \\ &= \{f`w : w \in u\}. \end{aligned} \tag{2}$$

In conventional mathematics, both of these are denoted $f(u)$, but in set theory, where all objects are explicitly sets this is ambiguous. So as not to depart too far from conventional mathematical notation, we shall generally use $f(u)$ for $f`u$, but not for $f``u$.

Expressions such as (3) can be abbreviated by writing

$$\begin{array}{lll} \exists w \in u(P(w)) & \text{for} & \exists w(w \in u \,\&\, P(w)), \\ \forall w \in u(P(w)) & \text{for} & \forall w(w \in u \;\Rightarrow\; P(w)). \end{array}$$

We shall do this in future.

Definition 3.7 A function $f : x \to y$ is *injective* if

$$\forall u_1 \in x \forall u_2 \in x(f`u_1 = f`u_2 \;\Rightarrow\; u_1 = u_2). \tag{3}$$

Formally, this definition is a definition of '$f : x \to y$ is injective' as an abbreviation for (3), but we are gradually dropping the rigidly formal approach.

4 Ordinals and Transfinite Induction

We have already indicated how to define the natural numbers 0,1,2,3, ... (Definition 2.4). Starting with $0 = \emptyset$, we defined each successive number to be the set of all the preceding ones.

Continuing in the same way, we can define

$$\begin{aligned} \omega &:= \{0,1,2,\ldots\}, \text{ the set of all natural numbers,} \\ \omega+1 &:= \omega \cup \{\omega\} = \{0,1,2,\ldots,\omega\}, \\ \omega+2 &:= (\omega+1) \cup \{\omega+1\} = \{0,1,2,\ldots,\omega,\omega+1\}, \\ &\ldots \end{aligned}$$

These are the *infinite ordinals*. We now define them more systematically.

Definition 4.1 A set x is $\in$-*transitive* if every member of x is a subset of x. Written symbolically:

$$\forall y \forall z(z \in y \,\&\, y \in x \;\Rightarrow\; z \in x).$$

A set x is an *ordinal* if x and every member of x is $\in$-transitive.

Remark 4.2 Every member of an ordinal is an ordinal. (Proof: easy exercise)

Definition 4.3 For ordinals α, β we define

$$\alpha < \beta \iff \alpha \in \beta$$

and

$$\alpha \leq \beta \iff \alpha \in \beta \text{ or } \alpha = \beta.$$

We show that $\leq$ is a partial ordering of the ordinals. The $\in$-transitivity of γ gives us

$$\alpha < \beta \mathbin{\&} \beta < \gamma \;\Rightarrow\; \alpha < \gamma,$$

so $<$ is transitive. Anti-reflexivity, the fact that $\neg(\alpha < \alpha)$, follows from $\alpha \notin \alpha$ (Proposition 2.5). Transitivity plus anti-reflexivity imply antisymmetry:

$$\alpha < \beta \;\Rightarrow\; \neg(\beta < \alpha).$$

These three properties easily translate into the three axioms for the partial order $\leq$:

1. $\alpha \leq \alpha$;
2. $\alpha \leq \beta \mathbin{\&} \beta \leq \gamma \;\Rightarrow\; \alpha \leq \gamma$;
3. $\alpha \leq \beta \mathbin{\&} \beta \leq \alpha \;\Rightarrow\; \alpha = \beta$.

We want to show that the ordinals are well-ordered: i.e. that they are totally ordered and that any non-empty class of ordinals has a least. We begin by proving that any non-empty class of ordinals has a minimal element (which is not, at this stage, known to be unique).

Proposition 4.4 *For any predicate P,*

$$\exists\alpha P(\alpha) \;\Rightarrow\; \exists\mu(P(\mu) \mathbin{\&} \forall\beta(\beta < \mu \;\Rightarrow\; \neg P(\beta))).$$

Proof. Let α be any ordinal such that $P(\alpha)$. If $\forall\beta(\beta < \alpha \;\Rightarrow\; \neg P(\beta))$, then $\mu = \alpha$ is the desired ordinal. If not, then

$$x = \{\beta : \beta < \alpha \mathbin{\&} P(\beta)\}$$

is nonempty. We now use (AF):

$$(AF) \qquad \exists y(y \in x) \;\Rightarrow\; \exists y\,(y \in x \mathbin{\&} \neg\exists z(z \in x \mathbin{\&} z \in y))\,.$$

Let μ be the set y given by (AF). Since $y \in x$, we know that y is an ordinal less that α. The statement $\neg\exists z(z \in x \mathbin{\&} z \in \mu)$ then says that no member z of μ is in x. That is, no ordinal less than μ has property P. $\Diamond$

Now we show that the ordering is total.

Proposition 4.5 *For all ordinals α, β just one of the following holds: $\alpha < \beta$, $\alpha = \beta$, $\beta < \alpha$.*

Proof. Clearly at most one of $\alpha < \beta$, $\alpha = \beta$, $\beta < \alpha$ can hold. Let us say that α and β are *comparable* if one does hold. Suppose that it is not true that every pair of ordinals are comparable. Then there is an ordinal which does not have the property of being comparable with every other ordinal. Therefore, by the previous proposition, there is a minimal such ordinal; call it μ. There is an ordinal not comparable with μ, and so there is a minimal such ordinal; call it ν.

We show that $\nu \subseteq \mu$. Let $\alpha \in \nu$. Then α is an ordinal less than ν so, by the minimality of ν, we have that α is comparable with μ. Now $\mu = \alpha$ and $\mu < \alpha$ both imply $\mu < \nu$, contrary to hypothesis, so we must have $\alpha < \mu$, i.e. $\alpha \in \mu$.

In fact, $\nu \subset \mu$, since μ and ν are incomparable. Let $\rho \in \mu \setminus \nu$. Since μ is an ordinal and $\rho \in \mu$, we know that ρ is an ordinal less than μ. By the minimality of μ, every ordinal is comparable with ρ. In particular, ρ and ν are comparable. Now $\rho \notin \nu$, so either $\nu = \rho$ or $\nu < \rho$. However, since $\rho < \mu$, either of these implies $\nu < \mu$, contradicting the choice of ν. This final contradiction proves the result. $\Diamond$

Corollary 4.6 *Every non-empty class of ordinals has a unique minimal element.*

Corollary 4.7 *Every ordinal is well-ordered by $\in$.*

Our development of the theory of ordinals has followed that of Schoenfield[3]. An alternative (used in Levy's book[4]) is to define ordinals as sets which are $\in$-transitive and well-ordered by $\in$. The above corollary, together with Exercise 6.2 below, show that the two definitions are equivalent.

[3] J. R. Schoenfield, 'Mathematical Logic', (Addison-Wesley, 1967) §9.3.
[4] Azriel Levy 'Basic Set Theory'(Springer 1979) Definition II.3.8.

Proposition 4.8 *If α, β are ordinals, then $\alpha \leq \beta$ iff $\alpha \subseteq \beta$.*

Proof. The fact that $\alpha \leq \beta \Rightarrow \alpha \subseteq \beta$ follows from the $\in$-transitivity of β. Conversely, if $\alpha \subseteq \beta$ then $\alpha \leq \beta$, since otherwise Proposition 4.5 would imply $\beta < \alpha$, so $\beta \in \alpha \subseteq \beta$, which implies $\beta \in \beta$, contrary to Proposition 2.5. $\Diamond$

The Principle of Transfinite Induction

Theorem 4.9 *For any predicate P, if*

$$\forall\alpha(\forall\beta(\beta < \alpha \Rightarrow P(\beta)) \Rightarrow P(\alpha))$$

then $\forall\alpha P(\alpha)$.

Proof. Suppose the conclusion is false. Let μ be the least ordinal such that $\neg P(\mu)$. Then $\forall\beta(\beta < \mu \Rightarrow P(\beta))$ holds, but $P(\mu)$ fails, contradicting the hypothesis. $\Diamond$

We recall Definition 2.3 $S(\alpha) = \alpha \cup \{\alpha\}$.

Exercise 4.10 Show that if α is an ordinal, then $S(\alpha)$ an ordinal and it is the least ordinal greater than α.

Exercise 4.11 Show that if α, β are ordinals with $S(\alpha) = S(\beta)$, then $\alpha = \beta$.

Definition 4.12 We say that α is a *successor ordinal* if there is some ordinal β with $\alpha = S(\beta)$. If α is neither a successor ordinal nor 0, we say that α is a *limit ordinal.*

We need to show that limit ordinals exist. (We have already introduced $\omega = \{0, 1, 2, \ldots\}$, the set of all natural numbers, informally. Now we construct it more formally.)

Proposition 4.13 *If x is a set of ordinals, then $\bigcup x$ is an ordinal.*

We recall that

$$\bigcup x := \{z : \exists y(z \in y \ \&\ y \in x)\}.$$

Proof. (Exercise) $\Diamond$

By Proposition 4.8, $\bigcup x$ is the least ordinal μ such that $\alpha \leq \mu$ for all $\alpha \in x$. We shall write $\mu = \sup\{\alpha : \alpha \in x\}$.

Corollary 4.14 *The class* On *of all ordinals is a proper class.*

Proof. If On were a set, $\alpha = \bigcup \mathrm{On}$ would be an ordinal greater than or equal to every ordinal. This is impossible since $\alpha < S(\alpha)$. $\Diamond$

The is the Burali–Forti paradox (1897); (it was a paradox before the distinction between sets and proper classes was introduced).

Proposition 4.15 *There is at least one limit ordinal.*

Proof. We use the Axiom of Infinity:

$$(AI) \qquad \exists x(\emptyset \in x \ \&\ \forall y(y \in x \Rightarrow S(y) \in x)).$$

Let

$$\mu = \bigcup\{\alpha : \alpha \in \mathrm{On} \ \&\ \alpha \in x\}.$$

Then μ is an ordinal and for all ordinals $\alpha \in x$, we have $\alpha \subseteq \mu$ and so $\alpha \leq \mu$. Since $0 = \emptyset \in x$ we have $1 = S(0) \in x$ and so $1 \leq \mu$. If $\mu = S(\beta)$, then $\beta \in \mu$, so $\beta \in \alpha \in x$ for some ordinal α; but then $S(\beta) = \mu \geq \alpha > \beta$, so $\mu = \alpha$. By the definition of x, we have $S(\alpha) \in x$, so $\mu \geq S(\alpha)$; contradiction. Thus μ is neither 0 nor a successor ordinal; i.e. it is a limit ordinal. $\Diamond$

The successor form of Transfinite Induction

Theorem 4.16 *For any predicate P, if*

1. $P(0)$,
2. $\forall\alpha(P(\alpha) \;\Rightarrow\; P(S(\alpha)))$ *and*
3. *for every limit ordinal λ,*
$$(\forall\alpha < \lambda)P(\alpha) \;\Rightarrow\; P(\lambda),$$

then $\forall\alpha P(\alpha)$.

Definition 4.17 The least limit ordinal is denoted ω. Its members are the *natural numbers*. An ordinal is said to be *finite* if it is a natural number and *infinite* otherwise.

Proposition 4.18 The Peano axioms for arithmetic.

1. $0 \in \omega$
2. $n \in \omega \;\Rightarrow\; S(n) \in \omega$
3. $A \subseteq \omega \;\&\; 0 \in A \;\&\; (n \in A \;\Rightarrow\; S(n) \in A) \;\Rightarrow\; A = \omega$
4. $n \in \omega \;\Rightarrow\; S(n) \neq 0$
5. $(n, m \in \omega \;\&\; S(n) = S(m)) \;\Rightarrow\; n = m$

Definition by transfinite induction

We wish to define a 'function' F inductively, using a formula G which, for each ordinal α, gives $F(\alpha)$ in terms of α and the values $F(\beta)$ for $\beta < \alpha$. To be more precise, we show that for each ordinal α there is a unique function f with domain α such that for every $\beta \in \alpha$ we have

$$f(\beta) = G(\beta, f \restriction \beta). \tag{4}$$

Let us write $P(\alpha)$ for this statement. Thus

$$P(\alpha) \equiv (\exists! f)(\mathcal{D}(f) = \alpha \;\&\; (\forall\beta \in \alpha)(f(\beta) = G(\beta, f \restriction \beta))).$$

We write f_α for the unique f associated with the ordinal α.

We prove $P(\alpha)$ by transfinite induction. Suppose α is an ordinal such that $P(\beta)$ holds for all $\beta < \alpha$. Then if $\gamma < \beta < \alpha$, the uniqueness part of the hypothesis $P(\gamma)$ implies that $f_\beta \restriction \gamma = f_\gamma$. We define f_α by:

(a) if $S(\beta) < \alpha$ then $f_\alpha(\beta) = f_{S(\beta)}(\beta)$;

(b) if $S(\beta) = \alpha$ then $f_\alpha(\beta) = G(\alpha, f_\beta)$.

Then (a), together with our previous remark, ensures that f_α is an extension of the f_β for $\beta < \alpha$. If α is a limit ordinal, this is enough to prove the existence part of $P(\alpha)$. The uniqueness assertion in $P(\alpha)$ follows from the uniqueness assertion in $P(\beta)$ applied to $f_\alpha \restriction \beta$, for each $\beta < \alpha$. If $\alpha = S(\beta)$, then (a) and (b) together prove the existence part of $P(\alpha)$. The uniqueness of $f_\alpha(\gamma)$ for $\gamma < \beta$ follows as before, and the uniqueness of $f_\alpha(\beta)$ is immediate because the requirement on f_α defines $f_\alpha(\beta)$ in terms of the $f_\alpha(\gamma)$ for $\gamma < \beta$.

The following example illustrates definition by transfinite induction in a ring-theoretic context.

Example 4.19 Definition of the Baer lower radical.

Let A be a (non-commutative) ring. We seek to define an ideal $R \lhd A$ such that the quotient ring A/R has no non-zero nilpotent ideals and such that R is minimal with this property.

We start by defining R_1 to be the sum of all the nilpotent ideals in A. Unfortunately, A/R_1 could contain non-zero nilpotent ideals. We therefore look at the sum of all the nilpotent ideals of A/R_1. This is R_2/R_1 for an ideal $R_2 \supseteq R_1$. Continuing in this way, we define $R_3, R_4, \ldots$, an ascending sequence of ideals of A. If, at any stage A/R_n has no nilpotent ideals, then we get $R_{n+1} = R_n$ and we stop and put $R = R_n$. Otherwise, having defined R_n for all $n \in \omega$, we define $R_\omega = \bigcup_{n\in\omega} R_n$ and then carry on, defining $R_{S(\omega)}$ such that $R_{S(\omega)}/R_\omega$ is the sum of the nilpotent ideals of A/R_ω.

More formally, the definition of the R_α and R is as follows. The R_α $\alpha \in \mathrm{On}$ are defined by:

- $R_0 = \{0\}$;
- $R_{S(\alpha)}$ is the ideal of A containing R_α such that $R_{S(\alpha)}/R_\alpha$ is the sum of the nilpotent ideals of A/R_α.
- for limit ordinals λ,
$$R_\lambda = \bigcup_{\alpha\in\lambda} R_\alpha.$$

Lemma 4.20 *Let R_α ($\alpha \in \mathrm{On}$) be an increasing family of subsets of a set A. Then $R_{S(\alpha)} = R_\alpha$ for some $\alpha \in \mathrm{On}$.*

Proof of Lemma. Suppose $R_{S(\alpha)} \supset R_\alpha$ for all ordinals α. Let

$$R = \{x \in A : x \in R_\alpha \text{ for some ordinal } \alpha\}.$$

Then R is a set — because it is a subclass of the set A and there is a 'function' $F : R \to \mathrm{On}$ given by

$$F(x) = \min\{\alpha : x \in R_{S(\alpha)}\}.$$

Then $F``R = \mathrm{On}$, for, given $\alpha \in \mathrm{On}$, there is an $x \in R_{S(\alpha)} \setminus R_\alpha$ and $x \notin R_\alpha$ implies $x \notin R_{S(\beta)}$ for all $\beta < \alpha$, so $\alpha = F(x)$. It follows that $\bigcup F``R = \mathrm{On}$. However, the Axiom of Replacement tells us that $\bigcup F``R$ must be a set. (The above has been written in terms of a 'function' F rather than the predicate P as in the formal statement of (AR), but it is a simple matter to translate into those terms.) We therefore conclude that $R_{S(\alpha)} = R_\alpha$ for some ordinal α. $\Diamond$

Define $R = R_\alpha$, (it follows that $R = R_\beta$ for all $\beta \geq \alpha$, but this is not needed). Then R is an ideal such that A/R has no non-zero nilpotent ideals.

If S is another ideal of A for which A/S has no non-zero nilpotent ideals, then we show by transfinite induction that $R_\beta \subseteq S$ for all $\beta \leq \alpha$. Clearly $R_0 \subseteq S$. Now suppose $R_\beta \subseteq S$, and consider a nilpotent ideal N/R_β of A/R_β. Then there exists r such that $N^r \subseteq R_\beta \subseteq N \cap S$. Therefore $N/N \cap S$ is nilpotent and so $(N + S)/S$, which is isomorphic to $N/N \cap S$, is a nilpotent ideal of A/S, which must therefore be trivial; i.e. $N \subseteq S$. Thus S/R_β contains every nilpotent ideal of A/R_β. Therefore $S/R_\beta \supseteq R_{S(\beta)}/R_\beta$ and so $S \supseteq R_{S(\beta)}$. Finally, if λ is a limit ordinal and $R_\beta \subseteq S$ for all $\beta < \lambda$, then clearly $R_\lambda \subseteq S$. Hence $R = R_\alpha \subseteq S$, as desired.

5 Ordinal arithmetic

We define the basic arithmetic operations for ordinals by transfinite induction in the same way as one would define these operations for natural numbers when developing number theory from Peano's axioms.

Definition 5.1 For ordinals α, β we define $\alpha + \beta$ by induction on β:

- $\alpha + 0 := \alpha$;
- $\alpha + S(\beta) := S(\alpha + \beta)$;

- if β is a limit ordinal, $\alpha+\beta := \sup\{\alpha+\gamma : \gamma<\beta\}$.

Thus $\alpha+1 = S(\alpha)$, and we can drop the notation $S(\alpha)$.

For ordinals α, β we define $\alpha.\beta$ by induction on β:

- $\alpha.0 := 0$;
- $\alpha.S(\beta) := \alpha.\beta+\alpha$;
- if β is a limit ordinal, $\alpha.\beta := \sup\{\alpha.\gamma : \gamma<\beta\}$.

In particular, $\alpha.1 = \alpha$, $\alpha.2 = \alpha+\alpha$.

For ordinals α, β we define α^β by induction on β:

- $\alpha^0 := 1$;
- $\alpha^{S(\beta)} := \alpha^\beta.\alpha$;
- if β is a limit ordinal, $\alpha^\beta := \sup\{\alpha^\gamma : \gamma<\beta\}$.

In particular, $\alpha^1 = \alpha$, $\alpha^2 = \alpha.\alpha$.

Remark 5.2 We can now picture the ordinals thus:
$0, 1, 2, \ldots \omega, \omega{+}1, \omega{+}2, \ldots, \omega{+}\omega{=}\omega.2, \omega.2{+}1, \ldots, \omega.3, \ldots\ldots, \omega.\omega{=}\omega^2, \ldots, \omega^3, \ldots, \omega^\omega, \ldots, \omega^{\omega^\omega}, \ldots.$

Proposition 5.3 *The following laws hold, for ordinals $\alpha, \beta\gamma$:*

$$(\alpha+\beta)+\gamma = \alpha+(\beta+\gamma); \qquad (\alpha.\beta).\gamma = \alpha.(\beta.\gamma);$$

$$\alpha.(\beta+\gamma) = \alpha.\beta+\alpha.\gamma; \qquad \alpha^{\beta+\gamma} = \alpha^\beta.\alpha^\gamma.$$

In general, however we have

$$\alpha+\beta \neq \beta+\alpha; \qquad \alpha.\beta \neq \beta.\alpha;$$

$$(\alpha+\beta).\gamma \neq \alpha.\gamma+\beta.\gamma; \qquad (\alpha.\beta)^\gamma \neq \alpha^\gamma.\beta^\gamma.$$

The proofs are by transfinite induction. (Exercise.) The following examples show the failures.

Examples 5.4 (a) $1+\omega = \omega \neq \omega+1$;

(b) $2.\omega = \omega \neq \omega.2$;

(c) $(\omega+1).2 = (\omega+1)+(\omega+1) = (\omega+(1+\omega))+1 = (\omega+\omega)+1 = \omega.2+1 \neq \omega.2+1.2$;

(d) $(\omega.2)^2 = (\omega.2).(\omega.2) = (\omega.(2.\omega)).2 = (\omega.\omega).2 = \omega^2.2 \neq \omega^2.2^2$.

Proposition 5.5 *If $\alpha \in \mathrm{On}$ and A is a set of ordinals with supremum σ, then*

(a) $\alpha+\sigma = \sup\{\alpha+\beta : \beta\in A\}$*;*

(b) $\alpha.\sigma = \sup\{\alpha.\beta : \beta\in A\}$*;*

(c) $\alpha^\sigma = \sup\{\alpha^\beta : \beta\in A\}$*;*

Proof. These results are trivial if $\sigma\in A$. Otherwise, σ must be a limit ordinal, in which case they follow from the definitions of the arithmetic operations at limit ordinals. $\Diamond$

Theorem 5.6 (Subtraction, division and taking logs.) *Given ordinals α, β:*

(a) $\beta\le\alpha \;\Rightarrow\; (\exists!\gamma)(\beta+\gamma=\alpha)$*;*

(b) $\beta\neq 0 \;\Rightarrow\; (\exists!\gamma)(\exists!\delta)(\beta.\gamma+\delta=\alpha \;\&\; \delta<\beta)$;

(c) $\beta\neq 0,1 \;\&\; \alpha\neq 0 \;\Rightarrow\; (\exists!\gamma)(\exists!\delta)(\exists!\varepsilon)(\beta^\gamma.\delta+\varepsilon=\alpha \;\&\; 0<\delta<\beta \;\&\; \varepsilon<\beta^\gamma)$.

In all three cases, $\gamma \le \alpha$.

Proof.

(a) We first prove existence (with $\gamma \le \alpha$) by transfinite induction on α, the ordinal β being thought of as fixed. The case $\alpha = 0$ can only occur when $\beta = 0$ and clearly the result holds with $\gamma = 0$. Suppose the result holds for α; we show that it holds for $S(\alpha)$. If $S(\alpha) < \beta$ there is nothing to prove. If $S(\alpha) = \beta$, then $\beta + 0 = S(\alpha)$. If $S(\alpha) > \beta$, then $\alpha \ge \beta$, so by the induction hypothesis there exists γ such that $\beta + \gamma = \alpha$ and $\gamma \le \alpha$. Then $\beta + S(\gamma) = S(\alpha)$ and $S(\gamma) \le S(\alpha)$.

Assuming that $\lambda \ge \beta$ is a limit ordinal and the result holds for all $\alpha < \lambda$, we show that it holds for λ. The case $\lambda = \beta$ is trivial, as before, so suppose $\lambda > \beta$. Let $\mu = \sup A$ where $A = \{\gamma : \beta + \gamma < \lambda \ \& \ \gamma \le \lambda\}$. Then $\mu \le \lambda$ and

$$\begin{aligned} \beta + \mu &= \sup\{\beta + \gamma : \gamma \in A\}, \text{ by Proposition 5.5,} \\ &= \sup\{\alpha : \alpha < \lambda\} = \lambda, \end{aligned}$$

since every $\alpha < \lambda$ is expressible in the form $\alpha = \beta + \gamma$ for some $\gamma \in A$.

For uniqueness, we first observe that $\beta + \gamma > \beta$ for all $\gamma \ne 0$. (This is easily proved by TI on γ.) If $\beta + \gamma = \beta + \delta$ with, say, $\gamma < \delta$, then, by the existence proof above, there is an ordinal σ such that $\delta = \gamma + \sigma$. Then

$$\beta + \delta = \beta + (\gamma + \sigma) = (\beta + \gamma) + \sigma > \beta + \gamma.$$

Thus $\beta + \gamma = \beta + \delta$ implies $\gamma = \delta$ and the uniqueness assertion is proved.

(b) Let γ' be the least ordinal such that $\beta.\gamma' > \alpha$. (There certainly are ordinals with this property since $\beta.S(\alpha) \ge S(\alpha) > \alpha$, so γ' exists and $\gamma' \le S(\alpha)$.) Clearly $\gamma' \ne 0$. Further, γ' cannot be a limit ordinal since if it were, then we should have

$$\begin{aligned} \beta.\gamma' &= \sup\{\beta.\varepsilon : \varepsilon < \gamma'\}, \text{ by the definition of multiplication,} \\ &\le \alpha, \end{aligned}$$

since each $\beta.\varepsilon \le \alpha$.

Therefore γ' is a successor, say $\gamma' = S(\gamma)$. Then $\gamma' \le S(\alpha)$ implies $\gamma \le \alpha$. By the minimality of γ',

$$\beta.\gamma \le \alpha < \beta.\gamma' = \beta.S(\gamma) = \beta.\gamma + \beta. \tag{5}$$

By the first part of this theorem, $\alpha = \beta.\gamma + \delta$ for some ordinal δ. If $\delta \ge \beta$ then $\delta = \beta + \varepsilon$, for some ε, so

$$\alpha = \beta.\gamma + (\beta + \varepsilon) = (\beta.\gamma + \beta) + \varepsilon \ge \beta.\gamma + \beta,$$

contradicting (5). Therefore $\delta < \beta$. The uniqueness proof is left as an exercise.

(c) Exercise. (Begin by letting γ' be the least ordinal such that $\beta^{\gamma'} > \alpha$. The fact that $\alpha \ne 0$ ensures that $\gamma' \ne 0$. Note that $\delta > 0$ since $\delta = 0$ would imply $\alpha = \varepsilon < \beta^{\gamma}$, contradicting the minimality of $\gamma' = S(\gamma)$.)

$\Diamond$

Theorem 5.7 *For $\beta \ge 2$ every ordinal α has a unique expansion in the form*

$$\alpha = \beta^{\gamma_0}.\delta_0 + \beta^{\gamma_1}.\delta_1 + \ldots + \beta^{\gamma_{n-1}}.\delta_{n-1},$$

where $n \in \omega$, $0 < \delta_i < \beta$ for all i and $\alpha \ge \gamma_0 > \gamma_1 > \ldots > \gamma_{n-1}$. (The case $\alpha = 0$ corresponds to $n = 0$.)

Proof. Suppose $\alpha \neq 0$. The proof consists of iterating the third part of Theorem 5.6. Thus, there exist unique $\gamma_0, \delta_0, \varepsilon_0$ such that

$$\alpha = \beta^{\gamma_0}.\delta_0 + \varepsilon_0$$

with $\gamma_0 \leq \alpha$, $0 < \delta_0 < \beta$ and $\varepsilon_0 < \beta^{\gamma_0}$. Then either $\varepsilon_0 = 0$, in which case we stop having proved the result with $n = 1$, or

$$\varepsilon_0 = \beta^{\gamma_1}.\delta_1 + \varepsilon_1$$

with $0 < \delta_1 < \beta$ and $\varepsilon_1 < \beta^{\gamma_1}$. Moreover, $\beta^{\gamma_1} \leq \beta^{\gamma_1}.\delta_1 \leq \varepsilon_0 < \beta^{\gamma_0}$ and so $\gamma_1 < \gamma_0$.

Proceeding thus, we obtain a strictly decreasing sequence of ordinals

$$\alpha \geq \gamma_0 > \gamma_1 > \ldots > \gamma_n$$

which stops only when $\varepsilon_n = 0$. Since we cannot have an infinite strictly decreasing sequence of ordinals, this must happen for some $n \in \omega$. The existence result follows. Uniqueness is easily deduced from the uniqueness assertion in Theorem 5.6. $\Diamond$

Corollary 5.8 *Every ordinal α has a unique expansion in the form*

$$\alpha = \omega^{\gamma_0}.a_0 + \omega^{\gamma_1}.a_1 + \ldots + \omega^{\gamma_{n-1}}.a_{n-1},$$

where $n \in \omega$, the a_i are positive integers and $\alpha \geq \gamma_0 > \gamma_1 > \ldots > \gamma_{n-1} \geq 0$.

6 Ordinals and well-ordered sets

Theorem 6.1 *Every well-ordered set is order-isomorphic to a unique ordinal by a unique isomorphism.*

Proof. Let $(X, \leq)$ be a well-ordered set. We define $F(\alpha) \in X$ for $\alpha \in \mathrm{On}$ by $F(\alpha) = \min(X \setminus F``\alpha)$, i.e. $F(\alpha)$ is the least member of $X$ not yet in the image of $F$. This defines $F$ by transfinite induction, up to an ordinal $\alpha$ such that $X \setminus F``\alpha = \emptyset$, unless this never happens, in which case we have $F(\alpha)$ defined for all α.

Suppose $F(\alpha)$ is defined for all $\alpha \in \mathrm{On}$. Note that each $F(\alpha)$ is defined to be different from all the $F(\beta)$ with $\beta < \alpha$. Then the sets $F``\alpha$ ($\alpha \in \mathrm{On}$) form a strictly increasing family of subsets of X, indexed by the ordinals. By Lemma 4.20, this is impossible.

Therefore, for some α, the map $F : \alpha \to X$ is surjective. By definition, it is injective and order-preserving. Further, if $G : \beta \to X$ is another order-isomorphism, then it is easy to show, by transfinite induction on γ, that $G(\gamma) = F(\gamma)$ for all $\gamma < \alpha$ and it follows that $\alpha = \beta$. $\Diamond$

One result of this theorem is that we can complete the proof of the equivalence of the two common definitions of the ordinals.

Exercise 6.2 Show that every set which is $\in$-transitive and well-ordered by $\in$ is an ordinal.

Another consequence of Theorem 6.1 is a more conceptual interpretation of the operations of ordinal arithmetic.

Addition Given two disjoint totally ordered sets A, B, we define an ordering on $A \cup B$. We define, for $a_1, a_2 \in A$, $b_1, b_2 \in B$,

- $a_1 \leq a_2$ in $A \cup B$ iff $a_1 \leq a_2$ in A;
- $b_1 \leq b_2$ in $A \cup B$ iff $b_1 \leq b_2$ in B;
- $a_1 \leq b_2$ in $A \cup B$ for all $a_1 \in A$, $b_1 \in B$.

If A and B are well-ordered, then so is $A \cup B$. In particular, for ordinals α, β, let A, B be disjoint well-ordered sets isomorphic to α, β. Then $A \cup B$ is isomorphic to some ordinal γ and it is easy to see (proof by transfinite induction on β) that $\gamma = \alpha + \beta$.

Multiplication Let A, B be two totally ordered sets. On $A \times B$, the Cartesian product, we define the reverse-lexicographic ordering: $(a_1, b_1) \le (a_2, b_2)$ iff $b_1 < b_2$ or $b_1 = b_2$ and $a_1 \le a_2$. Again, it is easy to see that if α, β are ordinals, then $\alpha \times \beta$ corresponds to the ordinal $\alpha.\beta$.

Exponentiation Let A, B be two well-ordered sets, with $A \neq \emptyset$, and let a_0 be the smallest element of A. We shall use the notation $A \uparrow B$ to mean the set of all those functions $f : B \to A$ for which $\{b \in B : f(b) \neq a_0\}$ is finite. (We reserve the simpler notation A^B for the set of *all* functions $f : B \to A$.) We define an ordering on $A \uparrow B$ by: $f_1 \le f_2$ iff $f_1 = f_2$ or $f_1(b) < f_2(b)$ where b is the largest element of B for which $f_1(b) \neq f_2(b)$. Notice that $f_1(b) \neq f_2(b)$ for only finitely many b, so we can find a greatest. If α, β are ordinals, then the ordered set $\alpha \uparrow \beta$ is order-isomorphic to the ordinal α^β; the proof is by transfinite induction on β.

7 Cardinals

Definition 7.1 For set x, y we define

$$\begin{aligned} x \asymp y &\iff \exists \text{ bijective } f : x \to y \\ x \preceq y &\iff \exists \text{ injective } f : x \to y \end{aligned}$$

Proposition 7.2 *The relation $\asymp$ is an 'equivalence relation'.*

Proof. Obvious. $\Diamond$

Essentially, cardinal numbers are the equivalence classes, or representatives of the equivalence classes, of $\asymp$.

Let $|x|$ denote the *cardinality* of x — i.e. the cardinal number associated with the equivalence class containing x. We note that

$$x_1 \asymp x_2 \;\&\; y_1 \asymp y_2 \;\&\; x_1 \preceq y_1 \quad\Rightarrow\quad x_2 \preceq y_2,$$

so we can define the relation $\le$ between cardinals by

$$|x| \le |y| \iff x \preceq y.$$

It is easy to see that $\le$ is reflexive and transitive; the fact that it is antisymmetric is the following famous theorem.

Theorem 7.3 (The Schröder–Bernstein Theorem.) *For all sets X, Y*

$$X \preceq Y \;\&\; Y \preceq X \quad\Rightarrow\quad X \asymp Y.$$

Proof. Let $f : X \to Y$, $g : Y \to X$ be injective. For $x \in X$ consider the following sequence of elements

$$\begin{aligned} g^{-1}(x), &\qquad (\text{if } x \in g``Y), \\ f^{-1}(g^{-1}(x)), &\qquad (\text{if } g^{-1}(x) \in f``X), \\ g^{-1}(f^{-1}(g^{-1}(x))), &\qquad (\text{if } f^{-1}(g^{-1}(x)) \in g``Y), \\ &\ldots \end{aligned}$$

Let the *order* $o(x)$ of x be defined as the number of such preimages which exist. Likewise we define the *order* $o(y)$ of an element $y \in Y$ by considering the sequence of preimages $f^{-1}(y)$, $g^{-1}(f^{-1}(y))$, ...

Now define $h : X \to Y$ by

(1) If $o(x) = \infty$, then $h(x) := f(x)$. In this case $o(f(x)) = \infty$. Moreover, if $y \in Y$ with $o(y) = \infty$, then $x := f^{-1}(y)$ exists and has infinite order, so $y = h(x)$. Thus h maps the infinite order elements of X bijectively onto the infinite order elements of Y.

(2) If $o(x) = 2n$ $(n = 0, 1, 2, \ldots)$, then $h(x) := f(x)$. In this case, $o(f(x)) = 2n+1$ and if $y \in Y$ with $o(y) = 2n+1$, then $x := f^{-1}(y)$ exists with $o(x) = 2n$ and $y = h(x)$. Again, h maps the even order elements of X bijectively onto the odd order elements of Y.

(3) If $o(x) = 2n+1$ $(n = 0, 1, 2, \ldots)$, then $h(x) := g^{-1}(x)$. In this case, $o(g^{-1}(x)) = 2n$ and if $y \in Y$ with $o(y) = 2n$, then $x := g(y)$ has $o(x) = 2n+1$ and $y = h(x)$. Thus, h maps the odd order elements of X bijectively onto the even order elements of Y.

This completes the proof. $\Diamond$

Definition 7.4

$$2^{|X|} := |\mathcal{P}(X)|.$$

Theorem 7.5 (Cantor's Theorem) *For every set X,*

$$2^{|X|} > |X|.$$

Of course, it is obvious that $\mathcal{P}(X) \succeq X$, by the imbedding

$$x \mapsto \{x\} : X \to \mathcal{P}(X).$$

The real content of Cantor's Theorem is that

$$\mathcal{P}(X) \not\asymp X.$$

Proof. Suppose $f : X \to \mathcal{P}(X)$ is bijective. Let

$$N := \{x \in X : x \notin f(x)\}.$$

Then $N \in \mathcal{P}(X)$, so $N = f(n)$ for some $n \in X$. But then

$$n \notin f(n) \iff n \in N \iff n \in f(n).$$

This contradiction completes the proof. $\Diamond$

Without a distinction between sets and proper classes, we would have
Cantor's Paradox: let κ be the cardinality of the set of all sets; then κ is the largest cardinal; but $2^\kappa > \kappa$.

By following through the proof of Cantor's Theorem with X = the set of all sets, we see that the proof is just Russell's Paradox in disguise.

We must now be explicit about what our cardinal numbers actually *are.* We do not want them to be the equivalence classes of $\asymp$, since these are proper classes. Instead, we select representatives from these proper classes. These representatives will be ordinals, but to make this plan work (and for most of cardinal arithmetic) we need the following statement, the truth of which we shall discuss later.

The Well-Ordering Principle.

(WO) Every set can be well-ordered.

In view of Theorem 6.1, this is equivalent to

$$(\forall x)(\exists \alpha \in \mathrm{On})(x \asymp \alpha).$$

There may be several ordinals equivalent to a given set: for example, $\omega \asymp \omega + 1$ by the map

$$\begin{array}{rcll} 0 & \mapsto & \omega & \\ n & \mapsto & n-1 & \quad (n = 1, 2, 3, \ldots). \end{array}$$

Definition 7.6

$$|x| := \min\{\alpha \in \mathrm{On} : x \asymp \alpha\}.$$

This definition is valid since any class of ordinals has a least.

Proposition 7.7 *For all sets x, y*

$$|x| = |y| \iff x \asymp y.$$

Proof. Obvious. $\Diamond$

We could formally define inequality between cardinals by

$$|x| \leq |y| \iff x \preceq y. \tag{6}$$

However, with cardinals defined to be ordinals, we can use $\leq$ with its usual ordinal meaning and prove (6) as a proposition.

Proposition 7.8 (WO) *For any sets x, y we have*

$$|x| \leq |y| \iff x \preceq y.$$

Proof. If $|x| \leq |y|$, then $|x| \subseteq |y|$; let $i : |x| \to |y|$ be the inclusion map. We have $f : x \asymp |x|$ and $g : y \asymp |y|$, so $g^{-1}if : x \to y$ is injective, and $x \preceq y$.

Suppose $h : x \preceq y$. Again, we have $g : y \asymp |y|$. Let $u = g``h``x \subseteq |y|$. We need the following lemma.

Lemma 7.9 *If u is a subset of an ordinal β, then u is a well-ordered set and is therefore order-isomorphic to some ordinal α. Then $\alpha \leq \beta$.*

Proof of Lemma. Exercise $\Diamond$

Continuing the proof of the proposition: we have $x \asymp u \asymp \alpha$ for some ordinal $\alpha \leq |y|$. From the definition of $|x|$, we must have $|x| \leq \alpha$. Therefore $|x| \leq |y|$. $\Diamond$

Proposition 7.10 *If X is a non-empty set, then $|X| \leq |Y|$ iff there is a surjective map $f : Y \twoheadrightarrow X$.*

Proof. Since $X \neq \emptyset$, there is at least one element $x_0 \in X$. If $|X| \leq |Y|$, then there exists $g : X \to Y$ injective. We define $f : Y \twoheadrightarrow X$ by

$$f(y) := \begin{cases} g^{-1}(y) & (y \in g(X)) \\ x_0 & (y \in Y \setminus g(X)) \end{cases}$$

Conversely, if $f : Y \twoheadrightarrow X$ is surjective, we use (WO) to well-order Y and define $g : X \to Y$ injective by

$$g(x) := \min\{y \in Y : f(y) = x\}.$$

$\Diamond$

Definition 7.11 An ordinal number α is said to be *a cardinal* if there is a set x such that $|x| = \alpha$; equivalently, if $|\alpha| = \alpha$.

Thus, since $|\omega + 1| = \omega$ as shown above, $\omega + 1$ is not a cardinal.

Theorem 7.12 *Every natural number (i.e. every $n \in \omega$) is a cardinal.*

Lemma 7.13 *For all ordinals α, β*

$$\alpha + 1 \asymp \beta + 1 \Rightarrow \alpha \asymp \beta.$$

Proof of Lemma. Suppose $f : \alpha + 1 \to \beta + 1$ is bijective, i.e. $f : \alpha \cup \{\alpha\} \to \beta \cup \{\beta\}$. If $f(\alpha) = \mu \neq \beta$ and $f(\lambda) = \beta$, say, then $\lambda \in \alpha$ and $\mu \in \beta$, so we construct a bijective map $g : \alpha \to \beta$ by: $g(\lambda) = \mu$ and $g(\gamma) = f(\gamma)$ $(\gamma \neq \lambda)$. $\Diamond$

Lemma 7.14

$$n \in \omega \Rightarrow n \not\asymp n + 1.$$

Proof of Lemma. We use ordinary induction: $0 \not\asymp 1$ since $f : \emptyset \to \{\emptyset\}$ cannot be surjective. The induction step

$$n \not\asymp n+1 \quad \Rightarrow \quad n+1 \not\asymp n+1+1$$

follows from the previous lemma. $\Diamond$

Proof of Theorem. Certainly, 0 is a cardinal, as it is the least ordinal α with $\alpha \asymp \emptyset$. Suppose $n \in \omega$ is such that $n+1$ is not a cardinal. Then $n+1 \asymp m$ for some $m \leq n$. But then $n+1 \preceq n$ and $n \preceq n+1$, so $n \asymp n+1$, contradicting Lemma 7.14. $\Diamond$

Proposition 7.15 *The first limit ordinal ω is a cardinal.*

Proof. If $\omega \asymp n < \omega$ then $n+1 \preceq \omega \preceq n$ and $n \preceq n+1$, so $n \asymp n+1$, contradicting Lemma 7.14 again. $\Diamond$

Definition 7.16 We write $\aleph_0$ for ω considered as a cardinal.

Generally, we observe that the infinite cardinals form a subclass of On which is well-ordered with initial segments which are sets. Therefore (exercise) there is an order isomorphism

$$\text{On} \to \text{Infinite Cardinals}$$

which we write

$$\alpha \mapsto \aleph_\alpha.$$

We write ω_α for $\aleph_\alpha$ considered as an ordinal. The distinction between ω_α and $\aleph_\alpha$ is non-existent in our theory, but would appear in any variant in which cardinals were not identified with specific ordinals. We shall retain this historic distinction, partly because itis common in the literature, but also because it enables us to distinguish cardinal arithmetic (defined below) from ordinal arithmetic without inventing new operation symbols. Thus '$\aleph_\alpha + \aleph_\alpha = \aleph_\alpha$' will signify addition of cardinals; '$\omega_\alpha + \omega_\alpha = \omega_\alpha.2$' will signify ordinal addition (of the corresponding ordinals). In the terminology of computer languages, 'ordinal' and 'cardinal' are two different 'data types' and we are 'overloading' the arithmetic operation symbols, as can be done in Ada.

The Operations of Cardinal Arithmetic

Definition 7.17 If $\boldsymbol{m}, \boldsymbol{n}$ are two cardinals, $\boldsymbol{m}+\boldsymbol{n}$ is defined to be the cardinality of the disjoint union of $\boldsymbol{m}$ and $\boldsymbol{n}$:

$$\boldsymbol{m}+\boldsymbol{n} = |(\boldsymbol{m} \times \{0\}) \cup (\boldsymbol{n} \times \{1\})|,$$

where, on the right, we are thinking of $\boldsymbol{m}$ and $\boldsymbol{n}$ as ordinals and so as sets of ordinals. Thus

$$\aleph_0 + \aleph_0 = |\{(n,0) : n \in \omega\} \cup \{(n,1) : n \in \omega\}| = \aleph_0.$$

The product $\boldsymbol{m}.\boldsymbol{n}$ is the cardinality of the Cartesian product

$$\boldsymbol{m}.\boldsymbol{n} = |\boldsymbol{m} \times \boldsymbol{n}|.$$

Powers are defined by

$$\boldsymbol{m}^{\boldsymbol{n}} = |\boldsymbol{m}^{\boldsymbol{n}}|,$$

where, on the right hand side, we are using the notation A^B, for sets A, B, to denote the set of all functions $f : A \to B$.

It is easy to see that for $\boldsymbol{m}, \boldsymbol{n} \in \omega$ the operations of ordinal and cardinal arithmetic coincide and coincide with our usual notions. The distinction between infinite ordinals and cardinals becomes dramatically apparent when we ask the following question.

How big is ω_1?

The ordinal ω_1 is the first ordinal with $\omega_1 \not\asymp \omega$. We have already seen that $\omega+1 \asymp \omega$, so $\omega_1 > \omega+1$.

Definition 7.18 We say that a set x is *countable* if $|x| \leq \aleph_0$ and *countably infinite* if $|x| = \aleph_0$.

The following theorem is easily proved.

Theorem 7.19 *(AC) Every countable union of countable sets is countable.*

It follows that $\omega^2 = \omega.\omega$ (ordinal exponentiation and multiplication) is countable, as is ω^3 and, indeed, every ω^n ($n \in \omega$). But then

$$\omega^\omega = \bigcup_{n\in\omega} \omega^n$$

is countable. From this we obtain that ω^{ω^ω} is countable (this means $\omega^{(\omega^\omega)}$, as usual). Hence $\omega^{\omega^{\cdots^\omega}}$, with n nested exponents is countable for each $n \in \omega$, and hence $\omega^{\omega^{\cdot^{\cdot^{\cdot}}}}$, with ω nested exponents. Call this last ordinal ω^* (temporarily) and we can go on to prove the countability of $(\omega^*)^*$, which is ω^* to the power ω^* to the power ω^* ..., with ω^* nested exponents; *et cetera, et cetera, et cetera.* In *ordinal* terms, ω_1 is BIG!

In fact, ω_1 cannot be the limit of any sequence of ordinals less than ω_1.

Application.

In functional analysis, a useful technique for producing counterexamples involves defining Banach spaces E_α ($\alpha \leq \omega_1$) by transfinite induction: $E_0 := \mathbb{C}$; $E_{\alpha+1} :=$ some construction from E_α such that $E_\alpha \subseteq E_{\alpha+1}$ isometrically; and, for limit ordinals λ,

$$E_\lambda := \widetilde{\left(\bigcup_{\alpha<\lambda} E_\alpha\right)}, \qquad \text{(completion).}$$

Proposition 7.20

$$E_{\omega_1} = \bigcup_{\alpha<\omega_1} E_\alpha,$$

(no completion).

Proof. Let (x_n) be a Cauchy sequence in $E = \bigcup_{\alpha<\omega_1} E_\alpha$. Then each $x_n \in E_{\alpha_n}$ for some $\alpha_n < \omega_1$. Now $\alpha_n < \omega_1$ implies that α_n is countable. Therefore

$$\beta := \sup_{n\in\omega} \alpha_n = \bigcup_{n\in\omega} \alpha_n$$

is countable, so $\beta < \omega_1$. Thus $x_n \in E_\beta$ for all n. Now E_β is a Banach space, so $x_n \to x \in E_\beta \subseteq E$. Thus E is already complete. $\Diamond$

8 The Axiom of Choice

Before proceeding further with the development of cardinal arithmetic, we return to the proof of (WE): Every set can be well-ordered.

It is now known that this cannot be proved within (ZF). We need a new axiom: Zermelo's Axiom of Choice (AC).

Definition 8.1 We say that a function $f : x \to \bigcup x$ is a **choice function** on x if $f(y) \in y$ for all $y \in x$.

(AC) For every set x with $\emptyset \notin x$ there is a choice function on x.

Colloquially: given a collection x of non-empty sets it is always possible to make a simultaneous choice of one element from each set $y \in x$.

Now in many situations, it is not necessary to invoke (AC) to get a choice function. This is particularly the case if the sets y have additional structure. For example: if the sets y are groups, we can simply choose $f(y)$ as the identity element of each y; if the y are sets of ordinals, we can let $f(y)$ be the least element of y. The Axiom of Choice only comes into play when we need to make *arbitrary* choices.

Furthermore, (AC) is only needed to make *infinitely many* arbitrary choices *simultaneously.* If $x = \{y\}$, so that we only want to make a choice $f(y)$ from a single non-empty set y, then the fact that $y \neq \emptyset$ means that $(\exists z)(z \in y)$ so $(\exists(y,z))(z \in y)$, so $(\exists f : x \to \bigcup x)(f(y) \in y)$. A (finite) induction argument then shows that if x is any finite set of non-empty sets, then there is a choice function on x. However, we cannot make the step to infinite x by TI; for example, if $x = \{y_0, y_1, y_2, \ldots\}$ is a sequence of non-empty sets then there exist choice functions f_n on $\{y_0, y_1, \ldots, y_n\}$ for each n, i.e. the set F_n of all such choice functions f_n is non-empty, but we cannot simultaneously *choose* one such $f_n \in F_n$ for each n without invoking (AC).

Theorem 8.2 *(AC)* $\Rightarrow$ *(WE)*

Proof. It suffices to show that, given an arbitrary non-empty set a, there is a bijection from some ordinal α onto a. Let x be the set of all non-empty subsets of a and let $f : x \to \bigcup x = a$ be a choice function on x. We define $F : \mathrm{On} \to a$ inductively by:

$$F(\beta) = f(a \setminus (F``\beta)),$$

while $F``\beta \neq a$ and, say, $F(\beta) = f(a)$ otherwise. That is, each $F(\beta)$ is chosen from the remaining elements of $a$, whilst this is still possible. If $F``\beta \neq a$ for all β, then $F(\beta)$ is defined as $f(a \setminus (F``\beta))$ for all ordinals $\beta$. This provides a strictly increasing family of subsets $F``\beta$ ($\beta \in \mathrm{On}$) of a, which is impossible by Lemma 4.20. Therefore we must have $F``\beta = a$ for some ordinal β. Choosing α to be the least such β, we have $F : \alpha \twoheadrightarrow a$ surjective and since

$$F(\beta) = f(a \setminus (F``\beta)) \qquad (\beta < \alpha),$$

the function F is also injective. $\Diamond$

Given that it can be shown that (AC) is not provable in (ZF), we can show that (WE) is not provable without (AC) by the following easy proposition.

Proposition 8.3 *Using only (ZF),*

$$(WE) \quad \Rightarrow \quad (AC).$$

Proof. Given a set x with $\emptyset \notin x$, we use (WE) to well-order $\bigcup x$. This provides simultaneously a well-order on each set $y \in x$. We then define the choice function on x by

$$f(y) := \min y \qquad (y \in x).$$

$\Diamond$

There are many other statements which are similarly equivalent in (ZF) to (AC). In fact there are 3 books full of them! (See the Bibliography.)

Probably the most important equivalent statement is *Zorn's Lemma.* Before we state it, we need some terminology concerning partially ordered sets (posets).

Definition 8.4 A *chain* in a poset is a totally ordered subset. An element x in a poset $(S, \leq)$ is an *upper bound* for a subset $T \subseteq S$ if $x \geq t$ for all $t \in T$. (It is not required that x be in T.) A *maximal* element of a poset $(S, \leq)$ is an element $s \in S$ such that for all $t \in S$, $t \geq s \;\Rightarrow\; t = s$.

Theorem 8.5 (ZFC) (Zorn's Lemma). *Every non-empty poset in which every chain has an upper bound has a maximal element.*

Proof. Let $(S, \leq)$ be the given poset. Let f be a choice function on $\mathcal{P}(S)$. For every $T \subseteq S$, let $\mathcal{U}(T)$ denote the set of strict upper bounds for T: i.e.

$$\mathcal{U}(T) := \{s \in S : s > t \quad (t \in T)\}.$$

We define $F : \mathrm{On} \to S$ inductively by

$$\begin{array}{rcll} F(0) & := & f(S), & \\ F(\alpha) & := & f(\mathcal{U}(F``\alpha)) & \quad (\alpha > 0), \end{array}$$

whilst $\mathcal{U}(F``\alpha) \neq \emptyset$ Then $F$ is a strictly increasing function, for if $\alpha < \beta$, then $F(\beta)$ is a strict upper bound for $F``\beta$ and $F(\alpha) \in F``\beta$, so $F(\alpha) < F(\beta)$. In particular, $F$ is an injection. Therefore the definition of $F$ must stop at some point (otherwise $F``\alpha$ would be a strictly increasing family of subsets of a set, indexed by the ordinals, contradicting Lemma 4.20. Therefore we must have $\mathcal{U}(F``\alpha) = \emptyset$ for some ordinal $\alpha$. Now $F``\alpha$ is a chain in S, so it has an upper bound $s \in S$, but there is no element of S strictly greater than s. Thus s is the desired maximal element. $\Diamond$

Theorem 8.6 *Within (ZF), Zorn's Lemma implies (AC).*

Proof. Let X be a set all of whose members are non-empty. Let S be the set of all pairs (f, Y) such that f is a choice function on a non-empty subset Y of X. If $y \in X$, then $y \neq \emptyset$, so there is a choice function $f : \{y\} \to y$. Thus $(f, \{y\}) \in S$, so $S \neq \emptyset$. We partially order S by defining

$$(f_1, Y_1) \leq (f_2, Y_2) \iff Y_1 \subseteq Y_2 \ \&\ f_2 \restriction Y_1 = f_1.$$

Let $\{(f_i, Y_i)\}_{i \in I}$ be a chain in S. Then we can define an upper bound (f, Y) for the chain by making

$$Y = \bigcup_{i \in I} Y_i$$

and defining f on Y by $f(y) = f_i(y)$ if $y \in Y_i$. This is consistent because $\{(f_i, Y_i)\}_{i \in I}$ is a chain. Thus if $y \in Y_i \cap Y_j$, then either $(f_i, Y_i) \leq (f_j, Y_j)$ or $(f_i, Y_i) \geq (f_j, Y_j)$; say the former. Then $f_i = f_j \restriction Y_i$, so $f_i(y) = f_j(y)$.

Zorn's Lemma implies that S has a maximal element (f, Y). We show that $Y = X$. Suppose not; let $x \in X \setminus Y$. Then $x \neq \emptyset$; so there is an element $\xi \in x$. We can then define a choice function g on $Y \cup \{x\}$ by $g(y) = f(y)$ $(y \in Y)$; $g(x) = \xi$. Then $(g, Y \cup \{x\}) > (f, Y)$, contradicting the maximality of (f, Y). Thus we have a choice function on X, as desired. $\Diamond$

Application: Existence of a Hamel basis.

Often it is possible to tackle a problem quite naturally using either Zorn's Lemma or well-ordering and transfinite recursion. Consider the following two proofs of a basic theorem in algebra.

Theorem 8.7 *Every vector space has a Hamel basis.*

First proof.

Let E be the given vector space. Well-order $E \setminus \{0\}$ so that

$$E = \{0\} \cup \{x_\alpha : \alpha < \mu\}$$

for some ordinal μ. We construct a basis $(e_\beta)_{\beta<\lambda}$ by transfinite induction: we let $e_0 = x_0$ and having constructed e_β $(\beta < \alpha)$, we find the smallest γ with $x_\gamma \notin \mathrm{span}\{e_\beta : \beta < \alpha\}$ and then let e_α be this x_γ. The process continues until $\mathrm{span}\{e_\beta : \beta < \alpha\} = E$, at which point we have a basis.

Second proof.

Let $(S, \subseteq)$ be the poset consisting of all linearly independent subsets of the given vector space E. Then $\{x_0\} \in S$. so $S \neq \emptyset$. If $(S_i)_{i \in I}$ is a chain in S, then $\bigcup_{i \in I} S_i$ is an upper bound in S. By Zorn's Lemma, S has a maximal element B say. If B is not a basis for E, then there exists $x \in E \setminus \mathrm{span}(B)$. However, this would imply that $\{x\} \cup B$ is a linearly independent set, so $\{x\} \cup B$ is an element of S strictly above B, contradicting the maximality of B. Therefore, B must be a basis.

Weaker versions of Choice.

There are various restricted form of (AC).

Countable Choice $(\mathbf{AC})_\omega$: (AC) for x countable.

Countable Choice from Sets of Reals $(\mathrm{AC})_\omega(\mathcal{P}(\mathbb{R}))$: (AC) for x countable and $x \subseteq \mathcal{P}(\mathbb{R})$.

Dependent Choice $(\mathrm{DC})_\omega$: given $x = \{y_0, y_1, y_2, \ldots\}$ and a function $d : \omega \times \bigcup x \to \bigcup x$ such that for all $z \in y_n$, $d(n, z)$ is a non-empty subset of y_{n+1}, there is a function $f : \omega \to \bigcup x$ such that $f(n) \in y_n$ and $f(n+1) \in d(n, f(n))$ for each $n \in \omega$. (The $(n+1)$st is dependent on the nth.)

Choice from Finite Sets (AC)(F): (AC) with all the $y \in x$ finite.

The Ultrafilter Theorem (UFT): every filter in a non-trivial Boolean algebra is contained in an ultrafilter.

The Boolean Prime Ideal Theorem (BPIT): every ideal in a non-trivial Boolean algebra is contained in a prime ideal.

The first three of these are self-explanatory; the terminology of the last two will be explained shortly. None of these restricted forms is provable within (ZF), but they are weaker than the full form of (AC). The conventional practice in mathematics is to assume Dependent Choice (which includes Countable Choice) without comment, but to point out uses of (AC) where uncountably many simultaneous choices are made.

Choice from finite sets is not usually isolated as a special case, but there is some beautiful work of J. H. Conway *et al.* on the relationship between the axioms '(AC) for sets of n elements' for different values of n. Notice that choice from finite sets would be provable in (ZF) if the finite sets were to come equipped with some structure which enabled a choice to be made—e.g. if they were finite sets of reals, where we could choose the smallest number in each set—but this is not generally the case.

The Tale of the Millionaire's Socks illustrates this. There once was a millionaire (actually, an infinitely rich millionaire) who owned infinitely many pairs of shoes and socks. One day, for want of anything better to do, he decided to try to select one of each pair of shoes and one of each pair of socks. He selected the shoes easily enough, by just taking the left shoe in each case, but (working in (ZF)) he was unable to simultaneously select one from each pair of socks! (Exercise for the reader: work this into a convincing fairy tale!)

Having noted that Choice from finite sets of reals is provable within (ZF), we see that the point of $(\mathrm{AC})_\omega(\mathcal{P}(\mathbb{R}))$ is that the sets of reals involved are not, generally, finite. Notice that even if they were all countable, you would need to have a choice of a counting of each set before you could get a choice function by selecting the first member of each set.

Example 8.8 The theorem which says that a function $f : X \to Y$ between two metric spaces is continuous iff $f(x_n) \to f(x)$ whenever $x_n \to x$ requires $(\mathrm{AC})_\omega$ in its proof. Suppose f is not continuous at x; then there is some $\varepsilon > 0$ such that for each n the set $S_n = \{s \in X : d(s, x) < 1/n \;\&\; d(f(s), f(x)) > \varepsilon\}$ is non-empty. We choose $x_n \in S_n$ $(n = 1, 2, 3, \ldots)$, simultaneously, to get $x_n \to X$ with $f(x_n) \not\to f(x)$.

Example 8.9 The theorem which says that a function $f : \mathbb{R} \to \mathbb{R}$ is continuous iff $f(x_n) \to f(x)$ whenever $x_n \to x$ requires $(\mathrm{AC})_\omega(\mathcal{P}(\mathbb{R}))$.

Example 8.10 Baire's Category Theorem requires Dependent Choice $(\mathrm{DC})_\omega$.

Baire's Category Theorem states that in a complete metric space X the intersection of any countable family $(G_n)_{n\in\omega}$ of dense open sets is non-empty. In the proof, one constructs sets $B(x_n; \delta_n)$ such that

$$\overline{B}(x_n; \delta_n) \subseteq B(x_{n-1}; \delta_{n-1}) \cap G_n \qquad (n = 1, 2, 3, \ldots);$$

it follows that the sequence (x_n) is Cauchy; its limit is in all the sets $\overline{B(x_n; \delta_n)}$ and hence in all the sets G_n, as desired. Note how each (x_n, δ_n) is chosen from a set of possible pairs which depends on the pair (x_{n-1}, δ_{n-1}).

9 Cardinal Arithmetic

Proposition 9.1 *For all cardinals $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$,*

$$\begin{array}{rcl} (\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c} & = & \boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c}), \\ \boldsymbol{a}+\boldsymbol{b} & = & \boldsymbol{b}+\boldsymbol{a}, \\ \boldsymbol{a}+0 & = & \boldsymbol{a}, \\ \boldsymbol{a} \leq \boldsymbol{b} & \iff & (\exists \boldsymbol{c})(\boldsymbol{a}=\boldsymbol{b}+\boldsymbol{c}), \\ \boldsymbol{a} \leq \boldsymbol{b} \,\&\, \boldsymbol{c} \leq \boldsymbol{d} & \Rightarrow & \boldsymbol{a}+\boldsymbol{c} \leq \boldsymbol{b}+\boldsymbol{d}, \\ \boldsymbol{a}.(\boldsymbol{b}+\boldsymbol{c}) & = & \boldsymbol{a}.\boldsymbol{b}+\boldsymbol{a}.\boldsymbol{c}. \end{array} \qquad \begin{array}{rcl} (\boldsymbol{a}.\boldsymbol{b}).\boldsymbol{c} & = & \boldsymbol{a}.(\boldsymbol{b}.\boldsymbol{c}), \\ \boldsymbol{a}.\boldsymbol{b} & = & \boldsymbol{b}.\boldsymbol{a}, \\ \boldsymbol{a}.1 & = & \boldsymbol{a}, \\ \boldsymbol{a}.0 & = & 0, \\ \boldsymbol{a} \leq \boldsymbol{b} \,\&\, \boldsymbol{c} \leq \boldsymbol{d} & \Rightarrow & \boldsymbol{a}.\boldsymbol{c} \leq \boldsymbol{b}.\boldsymbol{d}, \end{array}$$

Proposition 9.2

$$\boldsymbol{a}+1=\boldsymbol{a} \iff \boldsymbol{a} \geq \aleph_0.$$

Proof. If $\boldsymbol{a} \geq \aleph_0$ then $\boldsymbol{a} = |A|$ for some set $A \supseteq \{a_0, a_1, a_2, \ldots\}$. If $b \notin A$, then $A \cup \{b\} \asymp A$ by the bijection

$$\begin{array}{rcll} b & \mapsto & a_0 & \\ a_i & \mapsto & a_{i+1} & (i = 0,1,2,\ldots) \\ a & \mapsto & a & (a \in A \setminus \{a_0, a_1, a_2, \ldots\}). \end{array}$$

$\Diamond$

Corollary 9.3 *For every ordinal α, the ordinal ω_α is a limit ordinal.*

Corollary 9.4 *For every ordinal α and every $n \in \omega$,*

$$\omega_\alpha + n = \omega_\alpha.$$

Proof. by (finite) induction on n. $\Diamond$

Proposition 9.5 *For all ordinals α,*

(i) $\aleph_\alpha + \aleph_\alpha = \aleph_\alpha$,

(ii) $\aleph_\alpha . \aleph_\alpha = \aleph_\alpha$.

Proof.

(i) We construct a bijection

$$f : (\omega_\alpha \times \{0\}) \cup (\omega_\alpha \times \{1\}) \to \omega_\alpha$$

as follows. If β is an ordinal whose base ω expansion is

$$\beta = \omega^{\gamma_1} b_1 + \ldots + \omega^{\gamma_n} b_n + b_{n+1},$$

where $\gamma_1 > \gamma_2 > \ldots > \gamma_n \geq 1$ and $b_1, \ldots, b_n \in \omega \setminus \{0\}$, $b_{n+1} \in \omega$, then we define

$$\begin{array}{rcl} f(\beta, 0) & = & \omega^{\gamma_1} b_1 + \ldots + \omega^{\gamma_n} b_n + 2b_{n+1}, \\ f(\beta, 1) & = & \omega^{\gamma_1} b_1 + \ldots + \omega^{\gamma_n} b_n + 2b_{n+1} + 1. \end{array}$$

Note that because ω_α is a limit ordinal,

$$\beta < \omega_\alpha \;\Rightarrow\; f(\beta, 0), f(\beta, 1) < \omega_\alpha.$$

(ii) We prove that $\omega_\alpha \times \omega_\alpha \asymp \omega_\alpha$ by induction on α. For $\alpha = 0$, the fact that $\omega \times \omega \asymp \omega$ is easily proved. Suppose the result is known for all ordinals less than α. To show that $\omega_\alpha \times \omega_\alpha \asymp \omega_\alpha$, we define an ordering $\lhd$ on $\omega_\alpha \times \omega_\alpha$ by
$(\alpha,\beta) \lhd (\gamma,\delta) \iff$

$$(\max\{\alpha,\beta\} < \max\{\gamma,\delta\}) \vee ((\max\{\alpha,\beta\} = \max\{\gamma,\delta\}) \;\&\; (\alpha < \gamma \vee (\alpha = \gamma \;\&\; \beta < \delta)))\,.$$

It is easily checked that $\lhd$ is a well-ordering of $\omega_\alpha \times \omega_\alpha$, so there is an order isomorphism

$$F : (\mu, <) \to (\omega_\alpha \times \omega_\alpha, \lhd),$$

for some ordinal μ. Then

$$\mu \xrightarrow{F} \omega_\alpha \times \omega_\alpha \xrightarrow{\pi_1} \omega_\alpha$$

is surjective, so $|\mu| \geq \aleph_\alpha$, so $\mu \geq \omega_\alpha$. On the other hand, if $\nu < \mu$ then

$$\begin{aligned} F``\nu &= \{(\beta,\gamma) \in \omega_\alpha \times \omega_\alpha : (\beta,\gamma) \lhd (\delta,\varepsilon) = F(\nu)\} \\ &\subseteq \max\{\delta+1,\varepsilon+1\} \times \max\{\delta+1,\varepsilon+1\} \\ &\asymp \max\{\delta+1,\varepsilon+1\} < \omega_\alpha, \end{aligned}$$

by the induction hypothesis, since $\max\{\delta+1,\varepsilon+1\} \prec \omega_\alpha$. Therefore $F``\nu \prec \omega_\alpha$, so $\nu < \omega_\alpha$. Therefore $\mu = \omega_\alpha$ and the result is proven.

$\Diamond$

Corollary 9.6

$$\aleph_\alpha + \aleph_\beta = \max\{\aleph_\alpha, \aleph_\beta\} = \aleph_\alpha.\aleph_\beta.$$

Proof. We have

$$\max\{\aleph_\alpha,\aleph_\beta\} \leq \aleph_\alpha + \aleph_\beta \leq \max\{\aleph_\alpha,\aleph_\beta\} + \max\{\aleph_\alpha,\aleph_\beta\} = \max\{\aleph_\alpha,\aleph_\beta\}$$

and the equality follows from the Schröder-Bernstein Theorem. The proof that $\aleph_\alpha.\aleph_\beta = \max\{\aleph_\alpha,\aleph_\beta\}$ is similar. $\Diamond$

Corollary 9.7 *For every ordinal α and every $n \in \omega$,*

$$n.\aleph_\alpha = \aleph_\alpha.$$

Proof. by induction on n. $\Diamond$

Next, we consider cardinal exponentiation.

Proposition 9.8 *For cardinals $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$,*

$$\begin{array}{rclcrcl} \boldsymbol{a}^{\boldsymbol{b}+\boldsymbol{c}} &=& \boldsymbol{a}^{\boldsymbol{b}}.\boldsymbol{a}^{\boldsymbol{c}}, & & \left(\boldsymbol{a}^{\boldsymbol{b}}\right)^{\boldsymbol{c}} &=& \boldsymbol{a}^{\boldsymbol{b}.\boldsymbol{c}}, \\ (\boldsymbol{a}.\boldsymbol{b})^{\boldsymbol{c}} &=& \boldsymbol{a}^{\boldsymbol{c}}.\boldsymbol{b}^{\boldsymbol{c}}, & & 1^{\boldsymbol{a}} &=& 1, \\ 0^{\boldsymbol{a}} &=& 0 \quad (\boldsymbol{a} \neq 0), & & 0^0 &=& 1, \\ \boldsymbol{a}^1 &=& a, & & \boldsymbol{a}^2 &=& \boldsymbol{a}.\boldsymbol{a}, \\ & & \boldsymbol{a} \leq \boldsymbol{b} \;\&\; \boldsymbol{c} \leq \boldsymbol{d} & \Rightarrow & \boldsymbol{a}^{\boldsymbol{c}} \leq \boldsymbol{b}^{\boldsymbol{d}}. & & \end{array}$$

Proposition 9.9 *For every ordinal α and every $n \in \omega$,*

$$\aleph_\alpha{}^n = \aleph_\alpha.$$

Proof. by induction on n using Corollary 9.6. $\Diamond$

For any set x, subsets y of x can be identified with their characteristic functions

$$\chi_y(t) = \begin{cases} 0 & (t \in x \setminus y) \\ 1 & (t \in y), \end{cases}$$

hence

$$|\mathcal{P}(x)| = 2^{|x|}.$$

Thus Cantor's Theorem tells us that

$$2^{\boldsymbol{n}} > \boldsymbol{n}$$

for every cardinal $\boldsymbol{n}$.

This brings us to a famous conjecture.

The Generalized Continuum Hypothesis (GCH)

For every infinite cardinal $\boldsymbol{n}$, there is no cardinal $\boldsymbol{m}$ with $\boldsymbol{n} < \boldsymbol{m} < 2^{\boldsymbol{n}}$. With (AC), this becomes

$$2^{\aleph_\alpha} = \aleph_{\alpha+1}.$$

A special case of this is:

The Continuum Hypothesis (CH)

There is no cardinal $\boldsymbol{m}$ with $\aleph_0 < \boldsymbol{m} < 2^{\aleph_0}$. Again, with (AC) this becomes

$$2^{\aleph_0} = \aleph_1.$$

Put in less set-theoretic terms: for every subset $S \subseteq \mathbb{R}$, either S is countable or $S \asymp \mathbb{R}$. The problem of proving (CH) was stated as the first of his famous list of 23 problems by Hilbert in 1900. He also mentioned under the same heading the problem of whether the reals can be well-ordered. He even seems to entertain the possibility of a constructive proof of the latter. The idea that these problems might be independent of (ZF) was inconceivable.

In fact, both (GCH) and (CH) are consistent with (ZF) and independent of (ZFC). Here, (ZFC) denotes (ZF) + (AC). The consistency result: if (ZF) is consistent, then so is (ZFC) + (GCH) was proved by Gödel (1940). The independence result: if (ZF) is consistent, then so is (ZFC) + $\neg$(GCH) was proved by Cohen (1963). Note that by Gödel's Incompleteness Theorem, if (ZF) is consistent then we cannot prove that it is consistent within (ZF). If (ZF) is not consistent, we can prove anything in it.

10 Applications and non-applications of the Continuum Hypothesis

Whilst it is often conventient to visualise a world in which (CH) holds, it is generally desirable to prove results without using it, if possible. We regard it in the same way as (AC), but with even more suspicion. Very often it is sufficient to work with the cardinal $\mathfrak{c} = |\mathbb{R}|$. The following exercise shows just a few of the many sets encountered in routine analysis whose cardinality is $\mathfrak{c}$.

Exercise 10.1 We write $\mathfrak{c}$ for the cardinal $|\mathbb{R}|$. Without using either (AC) or (CH), show that the following sets have cardinality $\mathfrak{c}$:

1. any non-trivial open, closed or half-open interval of $\mathbb{R}$;
2. the irrationals;
3. the set $2^{\mathbb{N}}$ of all infinite sequences of zeros and ones;
4. the set $\mathbb{N}^{\mathbb{N}}$ of all infinite sequences of positive integers;
5. the set of all open subsets of $\mathbb{R}$; (hint: every open subset of $\mathbb{R}$ is a countable union of open intervals with rational end-points).

Corollary 10.2

$$\mathfrak{c} = 2^{\aleph_0}.$$

Proof. The cardinal $2^{\aleph_0}$ is the cardinality of the set $2^{\mathbb{N}}$ of the exercise. $\Diamond$

The following is an example of the use of transfinite induction up to cardinality $\mathfrak{c}$. In what follows, we use the word 'circle' to mean a circle of radius 1 in some plane in $\mathbb{R}^3$.

Theorem 10.3 *There is a family $\{C_\alpha\}_{\alpha<\mathfrak{c}}$ of disjoint circles (of radius 1) in $\mathbb{R}^3$ whose union is the whole of $\mathbb{R}^3$.*

Lemma 10.4 *Let $\omega \leq \alpha < \mathfrak{c}$.*

1. *Given a family $\{C_\xi\}_{\xi<\alpha}$ of circles, there is a point $p \in \mathbb{R}^3$ disjoint from all of them.*
2. *For any such point p, there is a circle C passing through p that is disjoint from each of the circles C_ξ.*

Proof of Lemma. Let P_ξ denote the plane containing C_ξ ($\xi < \alpha$). The set of all planes in $\mathbb{R}^3$ has cardinality $\mathfrak{c} > \alpha$, so there must be at least one plane P not containing any of the circles C_ξ. Each of the C_ξ meets this plane in at most two points, so the set $P \cap \bigcup_{\xi<\alpha} C_\xi$ has cardinality at most $2|\alpha| = |\alpha| < \mathfrak{c}$. Therefore there are points in P not belonging to any of the circles C_ξ, $\xi < \alpha$.

Given such a point p, the set of all planes through p has cardinality $\mathfrak{c} > \alpha$. Therefore there is a plane P through the point p that does not contain any of the circles C_ξ.

For each n, since C_ξ is not contained in P it must meet P in at most two points a_ξ, b_ξ. Now for each point $x \in P \setminus \{p\}$ there are at most two circles (of radius 1) in the plane P passing through both p and x. Therefore, there are at most $4|\alpha| = |\alpha| < \mathfrak{c}$ circles in P passing through p that contain some of the points a_ξ, b_ξ ($\xi < \alpha$).

There are $\mathfrak{c}$ circles (of radius 1) in P passing through p. Therefore, there must be circles in P passing through p and missing all the points a_ξ, b_ξ ($\xi < \alpha$), and consequently disjoint from all the circles C_ξ ($\xi < \alpha$). $\Diamond$

Proof of Theorem. We well-order $\mathbb{R}^3$: let $\{p_\alpha\}_{\alpha<\mathfrak{c}}$ be an enumeration of $\mathbb{R}^3$. We construct, by transfinite induction, a family $\{C_\alpha\}_{\alpha<\mathfrak{c}}$ of disjoint circles such that $p_\alpha \in \bigcup_{\xi\leq\alpha} C_\xi$ for all $\alpha < \mathfrak{c}$. This will be the desired family.

Suppose C_ξ ($\xi < \alpha$) have been found such that $p_\beta \in \bigcup_{\xi\leq\beta} C_\xi$ for all $\beta < \alpha$. Let β be the least ordinal such that $p_\beta \notin \bigcup_{\xi<\alpha} C_\xi$. There must be such an ordinal by the first part of the lemma. By the induction hypothesis, we must have $\beta \geq \alpha$, for if $\beta < \alpha$ then

$$p_\beta \in \bigcup_{\xi\leq\beta} C_\xi \subseteq \bigcup_{\xi<\alpha} C_\xi.$$

We define C_α to be the circle C given by the lemma when $p = p_\beta$. Then either $p_\alpha \in \bigcup_{\xi<\alpha} C_\xi$ or $p_\alpha = p_\beta \in C_\alpha$, so the induction step is proved.

(Note that the choice of C_α is to some extent arbitrary and so this proof involves making infinitely many arbitrary choices, which requires another application of (AC) in addition to that used to well-order $\mathbb{R}^3$.) $\Diamond$

Sometimes, one comes across theorems in everyday mathematics which do depend on (CH). Here is one such.

Theorem 10.5 *(AC) For a set $A \subseteq \mathbb{R}^2$ and $x, y \in \mathbb{R}$, we define*

$$A^y = \{x \in \mathbb{R} : (x, y) \in A\} \qquad A_x = \{y \in \mathbb{R} : (x, y) \in A\}.$$

Then (CH) is equivalent to the existence of a set $A \subseteq \mathbb{R}^2$ such that A^y and $(\mathbb{R}^2 \setminus A)_x$ are both countable, for every $x, y \in \mathbb{R}$.

Proof.

1. Suppose (CH); so $\mathfrak{c} = \aleph_1$. Thus we can write $\mathbb{R} = \{x_\alpha : \alpha < \omega_1\}$. Let $A = \{(x_\alpha, x_\beta) : \alpha \leq \beta\}$.

   Given $y \in \mathbb{R}$, we have $y = x_\alpha$ for some $\alpha < \omega_1$. Then $A^y = \{x_\beta : \beta \leq \alpha\}$. Since $\alpha < \omega_1$, it follows that A^y is countable.

   Given $x \in \mathbb{R}$, we have $x = x_\alpha$ for some $\alpha < \omega_1$; and $(\mathbb{R}^2 \setminus A)_x = \{x_\beta : \beta < \alpha\}$. Again, since $\alpha < \omega_1$, this set is countable.

2. Let $A \subseteq \mathbb{R}^2$ be as described and suppose (CH) fails, so $\aleph_1 < \mathfrak{c}$. We well-order $\mathbb{R}$ as $\{x_\alpha : \alpha < \mathfrak{c}\}$ and let $X = \bigcup_{\alpha<\omega_1} A^{x_\alpha}$. By assumption, each A^{x_α} is countable, so $|X| \leq \aleph_1 < \mathfrak{c}$. It follows that we can find some $x \in \mathbb{R} \setminus X$. Then, for every $\alpha < \omega_1$, we have $x \notin A^{x_\alpha}$, so $(x, x_\alpha) \notin A$, so $x_\alpha \in (\mathbb{R} \setminus A)_x$. Thus $|(\mathbb{R} \setminus A)_x| \geq \aleph_1$, contrary to hypothesis.

$\diamondsuit$

Another more significant example of an application of (CH) is the solution by Garth Dales and Jean Esterle of Kaplansky's Problem.

Consider the algebra $A = C[0,1]$ of all continuous complex-valued functions on $[0,1]$. This is an associative algebra under pointwise multiplication, scalar multiplication and addition. It is a normed algebra in the norm

$$\|f\| = \sup\{|f(t)| : 0 \leq t \leq 1\};$$

that is, the function $f \mapsto \|f\| : A \to \mathbb{R}^+$ satisfies

1. $\|f\| = 0 \;\Rightarrow\; f = 0$ $(f \in A)$,
2. $\|f + g\| \leq \|f\| + \|g\|$ $(f, g \in A)$,
3. $\|\lambda f\| = |\lambda|\,\|f\|$ $(\lambda \in \mathbb{C}, f \in A)$,
4. $\|fg\| \leq \|f\|\,\|g\|$ $(f, g \in A)$.

This norm is complete (every Cauchy sequence in A has a limit in A) and it may be shown that every complete algebra norm $|.|$ on A is equivalent to $\|.\|$ in the sense that there exist constants $c, C > 0$ such that

$$c\|f\| \leq |f| \leq C\|f\| \qquad (f \in A).$$

Kaplansky's Problem was whether *every* algebra norm on A is equivalent to $\|.\|$; equivalently, does there exist an incomplete algebra norm on A? Dales and Esterle showed that, assuming (CH), there does exist an incomplete algebra norm on A. Solovay then showed that this result is not provable in (ZFC) alone: there is a model of (ZFC) in which there is no incomplete algebra norm on A. Details of the independence result, using Woodin's simplification of Solovay's proof, may be found in the book:

> H. G. Dales and W. H. Woodin, 'An Introduction to Independence for Analysts', (Cambridge University Press, LMS Lecture Note Series no.115, 1987).

11 Cardinal exponentiation: Cofinality, Regular and Singular Cardinals

One might wonder whether all expressions of the form ${\aleph_\alpha}^{\aleph_\beta}$ are determined once the **continuum function** $\aleph_\alpha \mapsto 2^{\aleph_\alpha}$ is known. The answer is positive if (GCH) holds, but negative in general. Under a technical assumption (that the existence of a supercompact cardinal is consistent with (ZFC)), Magidor[5] proved that (ZFC) is consistent with either of the scenarios:

(i) $2^{\aleph_0} = \aleph_1$, $2^{\aleph_n} = \aleph_{\omega+2}$ $(n \leq \omega)$, $2^{\aleph_\alpha} = \aleph_{\alpha+2}$ $(\alpha > \omega)$, ${\aleph_\omega}^{\aleph_0} = \aleph_{\omega+1}$;

(ii) as (i) but ${\aleph_\omega}^{\aleph_0} = \aleph_{\omega+2}$.

[5] M. Magidor 'On the singular cardinals problem I' *Israel J. Math.,* **28** (1977), 1–31.

We now investigate the function $(\alpha, \beta) \mapsto {\aleph_\alpha}^{\aleph_\beta}$. In this discussion, we shall assume (ZFC), but not (GCH) unless explicitly stated.

Proposition 11.1 *If $\boldsymbol{a}, \boldsymbol{b}$ are cardinals such that $\boldsymbol{b}$ is infinite and $2 \leq \boldsymbol{a} \leq 2^{\boldsymbol{b}}$, then $\boldsymbol{a}^{\boldsymbol{b}} = 2^{\boldsymbol{b}}$.*

Proof.

$$2^{\boldsymbol{b}} \leq \boldsymbol{a}^{\boldsymbol{b}} \leq \left(2^{\boldsymbol{b}}\right)^{\boldsymbol{b}} = 2^{\boldsymbol{b}.\boldsymbol{b}} = 2^{\boldsymbol{b}}.$$

$\Diamond$

Corollary 11.2 *(i)*

$${\aleph_\alpha}^{\aleph_\beta} = 2^{\aleph_\beta} \qquad (\aleph_\alpha \leq \aleph_\beta);$$

(ii)

$${\aleph_\alpha}^{\aleph_\beta} \leq 2^{\aleph_\alpha} \qquad (\aleph_\alpha \geq \aleph_\beta).$$

Proof. (i) is from the previous proposition and (ii) follows from (i):

$${\aleph_\alpha}^{\aleph_\beta} \leq {\aleph_\alpha}^{\aleph_\alpha} = 2^{\aleph_\alpha}.$$

$\Diamond$

To proceed any further, we need an important new concept.

Definition 11.3 Let α, β be limit ordinals; we say that α is **of cofinality** β and write $\mathrm{cf}(\alpha) = \beta$ if there is a family of ordinals $\{\mu_\gamma\}_{\gamma<\beta}$ with $\mu_\gamma < \alpha$ $(\gamma < \beta)$ and $\sup_\gamma \mu_\gamma = \alpha$ and β is the least ordinal for which such a family exists.

It is easy to see that in this situation the family $\{\mu_\gamma\}_{\gamma<\beta}$ can be chosen to be increasing.

Properties of cofinality:

(i) $\mathrm{cf}(\alpha) \leq \alpha$, since $\alpha = \sup_{\gamma<\alpha} \gamma$ is one expression of α as a sup of strictly smaller ordinals;

(ii) $\mathrm{cf}(\mathrm{cf}(\alpha)) = \mathrm{cf}(\alpha)$;

(iii) $\mathrm{cf}(\alpha) \leq |\alpha|$, since if $f : |\alpha| \asymp \alpha$ then $\alpha = \sup_{\gamma<|\alpha|} f(\gamma)$ is one expression of α as a sup of strictly smaller ordinals;

(iv) if $\mathrm{cf}(\alpha) = \alpha$ then α is a cardinal, (by (iii));

(v) $\mathrm{cf}(\alpha)$ is a cardinal, (by (ii) and (iv)).

Examples 11.4 (i) $\mathrm{cf}(\omega^\omega) = \omega$, since $\omega^\omega = \sup_{n\in\omega} \omega^n$, (or by property (iii) above).

(ii) $\mathrm{cf}(\omega_\omega) = \omega$, since $\omega_\omega = \sup_{n\in\omega} \omega_n$, which shows that the converse of property (iv) is false.

The following proposition generalizes Example (ii).

Proposition 11.5 *If α is a limit ordinal, then* $\mathrm{cf}(\omega_\alpha) = \mathrm{cf}(\alpha)$.

Proof. If $\beta = \mathrm{cf}(\alpha)$, we have $\alpha = \sup_{\gamma<\beta} \mu_\gamma$ for some $\mu_\gamma < \alpha$ $(\gamma < \beta)$. Then

$$\omega_\alpha = \sup_{\gamma<\beta} \omega_{\mu_\gamma},$$

so $\mathrm{cf}(\omega_\alpha) \leq \mathrm{cf}(\alpha)$.

Conversely, if $\beta = \mathrm{cf}(\omega_\alpha)$ and $\omega_\alpha = \sup_{\gamma<\beta} \mu_\gamma$ for some $\mu_\gamma < \alpha$ $(\gamma < \beta)$, then we can write $\alpha = \sup_{\gamma<\beta} \nu_\gamma$, where ν_γ is such that $|\mu_\gamma| = \omega_{\nu_\gamma}$. $\Diamond$

Definition 11.6 We say that an infinite cardinal $\boldsymbol{a}$ is **regular** if $\mathrm{cf}(\boldsymbol{a}) = \boldsymbol{a}$ and **singular** if $\mathrm{cf}(\boldsymbol{a}) < \boldsymbol{a}$.

Examples 11.7 $\aleph_0$ is regular, $\aleph_\omega$ is singular.

Proposition 11.8 *$\aleph_{\alpha+1}$ is regular, for all α.*

Lemma 11.9 *If $\aleph_\alpha, \aleph_\beta$ are cardinals and $\{A_\delta\}_{\delta<\omega_\beta}$ is a family of sets with $|A_\delta| \leq \aleph_\alpha$ $(\delta < \omega_\beta)$, then*

$$\left|\bigcup_{\delta<\omega_\beta} A_\delta\right| \leq \aleph_\alpha.\aleph_\beta$$

Proof of Lemma. The fact that $|A_\delta| \leq \aleph_\alpha$ implies that there is a surjection $f_\delta : \omega_\alpha \to A_\delta$ for each δ. Putting these together, we construct a surjection

$$(\gamma, \delta) \mapsto f_\delta(\gamma) : \omega_\alpha \times \omega_\beta \to \bigcup_{\delta<\omega_\beta} A_\delta,$$

whence the result. $\Diamond$

Proof. of Proposition 11.8.

Suppose the proposition is false, that cf $(\aleph_{\alpha+1}) = \aleph_\beta$ with $\beta \leq \alpha$. Then

$$\omega_{\alpha+1} = \bigcup_{\delta<\omega_\beta} \mu_\delta$$

with $\mu_\delta \leq \omega_\alpha$ $(\delta < \omega_\beta)$. The lemma then gives

$$|\omega_{\alpha+1}| \leq \aleph_\alpha.\aleph_\beta = \aleph_\alpha,$$

which is a contradiction. $\Diamond$

Lemma 11.10 (Zermelo) *If A_γ, B_γ $(\gamma < \omega_\beta)$ are sets with $|A_\gamma| < |B_\gamma|$ $(\gamma < \omega_\beta)$, then*

$$\left|\bigcup_\gamma A_\gamma\right| < \left|\prod_\gamma B_\gamma\right|.$$

Proof. The conclusion certainly holds with a $\leq$ in place of $<$; the point of this lemma is that the inequality is strict. Suppose otherwise; then there is a surjection

$$f : \bigcup_\gamma A_\gamma \to \prod_\gamma B_\gamma.$$

Let

$$C_\gamma := \pi_\gamma\left(f``A_\gamma\right) \qquad (\gamma < \omega_\beta),$$

where π_γ is the projection onto the γth factor of the infinite product. Then $|C_\gamma| \leq |A_\gamma| < |B_\gamma|$, so $B_\gamma \setminus C_\gamma \neq \emptyset$. Choose $y_\gamma \in B_\gamma \setminus C_\gamma$ for each $\gamma < \omega_\beta$, so that we get an element

$$y = (y_\gamma)_{\gamma<\omega_\beta} \in \prod_{\gamma<\omega_\beta} B_\gamma.$$

Then, since f is surjective, $y = f(x)$ for some $x \in \bigcup_\gamma A_\gamma$. Then x is in some A_δ, so

$$y = f(x) \in f``A_\delta,$$

so

$$y_\delta = \pi_\delta(f(x)) \in C_\delta,$$

which contradicts the definition of y_δ. $\Diamond$

Theorem 11.11

$$\aleph_\alpha{}^{\mathrm{cf}\,(\aleph_\alpha)} > \aleph_\alpha.$$

Proof. If α is either 0 or a successor ordinal, then, using Proposition 11.8, we have $\mathrm{cf}\,(\aleph_\alpha) = \aleph_\alpha$, so

$$\aleph_\alpha{}^{\mathrm{cf}\,(\aleph_\alpha)} = \aleph_\alpha{}^{\aleph_\alpha} = 2^{\aleph_\alpha} > \aleph_\alpha.$$

If α is a limit ordinal, with $\mathrm{cf}\,(\aleph_\alpha) = \aleph_\beta$, then

$$\omega_\alpha = \bigcup_{\gamma<\omega_\beta} \mu_\gamma,$$

with $\mu_\gamma < \omega_\alpha$, so $|\mu_\gamma| < \aleph_\alpha$, for all $\gamma < \omega_\beta$. Zermelo's Lemma then implies

$$\aleph_\alpha = \left|\bigcup_{\gamma<\omega_\beta} \mu_\gamma\right| < \left|\prod_{\gamma<\omega_\beta} \omega_\alpha\right| \le \aleph_\alpha{}^{\aleph_\beta}.$$

$\Diamond$

Corollary 11.12

$$\mathrm{cf}\left(\aleph_\beta{}^{\aleph_\alpha}\right) > \aleph_\alpha$$

Proof. Suppose otherwise, that

$$\mathrm{cf}\left(\aleph_\beta{}^{\aleph_\alpha}\right) = \boldsymbol{n} \le \aleph_\alpha.$$

Then $\aleph_\alpha.\boldsymbol{n} = \aleph_\alpha$, so

$$\aleph_\beta{}^{\aleph_\alpha} < \left(\aleph_\beta{}^{\aleph_\alpha}\right)^{\boldsymbol{n}} = \aleph_\beta{}^{(\aleph_\alpha.\boldsymbol{n})} = \aleph_\beta{}^{\aleph_\alpha}.$$

$\Diamond$

We are now in a position to give a complete description of cardinal exponentiation assuming (GCH).

Theorem 11.13 *If we assume (GCH), then*

$$\aleph_\alpha{}^{\aleph_\beta} = \begin{cases} \aleph_\alpha & \text{if } \aleph_\beta < \mathrm{cf}\,(\aleph_\alpha) \\ \aleph_{\alpha+1} & \text{if } \mathrm{cf}\,(\aleph_\alpha) \le \aleph_\beta \le \aleph_\alpha \\ \aleph_{\beta+1} & \text{if } \aleph_\beta \ge \aleph_\alpha \end{cases}$$

Proof.

(i) Suppose $\aleph_\beta < \mathrm{cf}\,(\aleph_\alpha)$. We consider the three types of ordinal α.

(a) $\alpha = 0$ is impossible since $\mathrm{cf}\,(\aleph_0) = \aleph_0$.

(b) If α is a successor ordinal, say $\alpha = \gamma + 1$, then $\aleph_\alpha$ is regular, so

$$\aleph_\beta < \mathrm{cf}\,(\aleph_\alpha) = \aleph_\alpha = \aleph_{\gamma+1},$$

so $\aleph_\beta \le \aleph_\gamma$. By (GCH),

$$2^{\aleph_\gamma} = \aleph_{\gamma+1} = \aleph_\alpha,$$

so

$$\aleph_\alpha{}^{\aleph_\beta} = \left(2^{\aleph_\gamma}\right)^{\aleph_\beta} = 2^{\aleph_\gamma.\aleph_\beta} = 2^{\aleph_\gamma} = \aleph_\alpha.$$

(c) Now suppose α is a limit ordinal. Clearly $\aleph_\alpha{}^{\aleph_\beta} \geq \aleph_\alpha$; we prove the reverse inequality. Since $\aleph_\beta < \mathrm{cf}\,(\aleph_\alpha)$, any function $f : \omega_\beta \to \omega_\alpha$ must have $\sup f``\omega_\beta < \omega_\alpha$; otherwise we should be contradicting the minimality of $\mathrm{cf}\,(\aleph_\alpha)$. Therefore the set A of all such functions is

$$\bigcup_{\gamma<\alpha} A_\gamma,$$

where A_γ is the set of all functions $f : \omega_\beta \to \omega_\gamma$. Thus, if $\gamma < \beta$, then

$$\begin{aligned} |A_\gamma| &= \aleph_\gamma{}^{\aleph_\beta} \\ &= 2^{\aleph_\beta}, \text{ by Corollary 11.2(i)}, \\ &= \aleph_{\beta+1}, \text{ by (GCH)}, \\ &< \aleph_\alpha, \end{aligned}$$

since $\beta < \alpha$ and α is a limit ordinal.
If $\gamma \geq \beta$, then

$$\begin{aligned} |A_\gamma| &= \aleph_\gamma{}^{\aleph_\beta} \\ &\leq 2^{\aleph_\gamma}, \text{ by Corollary 11.2(ii)}, \\ &= \aleph_{\gamma+1}, \text{ by (GCH)}, \\ &< \aleph_\alpha, \end{aligned}$$

since $\gamma < \alpha$ and α is a limit ordinal.
Lemma 11.9 then gives

$$\aleph_\alpha{}^{\aleph_\beta} = |A| \leq \aleph_\alpha . \aleph_\alpha = \aleph_\alpha .$$

(ii) Suppose $\mathrm{cf}\,(\aleph_\alpha) \leq \aleph_\beta \leq \aleph_\alpha$. Then

$$\begin{aligned} \aleph_\alpha &< \aleph_\alpha{}^{\mathrm{cf}\,(\aleph_\alpha)}, \text{ by Theorem 11.11} \\ &\leq \aleph_\alpha{}^{\aleph_\beta} \\ &\leq 2^{\aleph_\alpha}, \text{ by Corollary 11.2(ii)}, \\ &= \aleph_{\alpha+1}, \text{ by (GCH)}. \end{aligned}$$

Since $\aleph_{\alpha+1}$ is the next cardinal above $\aleph_\alpha$, it follows that $\aleph_\alpha{}^{\aleph_\beta} = \aleph_{\alpha+1}$.

(iii) If $\aleph_\beta \geq \aleph_\alpha$, then Corollary (11.2)(i) gives

$$\aleph_\alpha{}^{\aleph_\beta} = 2^{\aleph_\beta} = \aleph_{\beta+1}.$$

$\Diamond$

Let us now forget (GCH) and investigate the continuum function.

Theorem 11.14 (König's Lemma)

$$\mathrm{cf}\left(2^{\aleph_\alpha}\right) > \aleph_\alpha. \tag{7}$$

Proof. The result follows from Corollary 11.12 with $\beta = 0$, since

$$\aleph_0{}^{\aleph_\alpha} = 2^{\aleph_\alpha},$$

by Corollary 11.2(i). $\Diamond$

This is a very important constraint on the behaviour of the continuum function $\aleph_\alpha \mapsto 2^{\aleph_\alpha}$; for example, it tells us that $2^{\aleph_0}$ cannot be $\aleph_\omega$. The other, rather obvious, constraint is that

$$\alpha \leq \beta \;\;\Rightarrow\;\; 2^{\aleph_\alpha} \leq 2^{\aleph_\beta}. \tag{8}$$

Of course we also know that $2^{\aleph_\alpha} > \aleph_\alpha$, but this can be viewed as consequence of (7).

Easton (1964) showed that (7) and (8) are the only statements you can make about $2^{\aleph_\alpha}$ for *regular* cardinals $\aleph_\alpha$. That is, if P is a function on the class of all regular cardinals $\aleph_\alpha$ into the class of all cardinals such that

(i) $\mathrm{cf}\,(P\,(\aleph_\alpha)) > \aleph_\alpha$ and

(ii) $\alpha \leq \beta \;\;\Rightarrow\;\; P\,(\aleph_\alpha) \leq P\,(\aleph_\beta)$,

then it is consistent with (ZFC) that

$$2^{\aleph_\alpha} = P\,(\aleph_\alpha)$$

for all regular cardinals $\aleph_\alpha$.

The **Singular Cardinals Problem** asks what constraints apply to $2^{\aleph_\alpha}$ for singular cardinals $\aleph_\alpha$.

Before we proceed, we need the notion of infinite sums and products of cardinals.

Definition 11.15 Let $\{\boldsymbol{n}_i\}_{i\in I}$ be a set of cardinals and let A_i be a set of cardinality $\boldsymbol{n}_i$, for each i. Then the infinite sum $\sum_{i\in I}\boldsymbol{n}_i$ is defined to be the cardinality of the disjoint union

$$\bigcup_{i\in I}(A_i \times \{i\})$$

and the infinite product $\prod_{i\in I}\boldsymbol{n}_i$ is defined to be the cardinality of the cartesian product set $\prod_{i\in I}A_i$, that is, the set of all choice functions

$$f : I \to \bigcup_{i\in I}A_i$$

with $f(i) \in A_i$ $(i \in I)$.

Infinite sums and products obey the obvious generalizations of the laws in (9.1).

Exercise 11.16 1. Show that if I is infinite and $\boldsymbol{n}_i > 0$ for all $i \in I$, then

$$\sum_{i\in I}\boldsymbol{n}_i = |I|.\sup_{i\in I}\boldsymbol{n}_i.$$

2. Show that

$$2^{\sum_{i\in I}\boldsymbol{n}_i} = \prod_{i\in I}2^{\boldsymbol{n}_i}.$$

Definition 11.17 For α a limit ordinal, we define

$$2^{<\,\aleph_\alpha} = \sup\left\{2^{\aleph_\beta} : \beta < \alpha\right\}.$$

Note that (GCH) would imply that

$$2^{<\,\aleph_\alpha} = \sup\{\aleph_{\beta+1} : \beta < \alpha\} = \aleph_\alpha.$$

Lemma 11.18 *If α is a limit ordinal, then*

$$2^{\aleph_\alpha} = \left(2^{<\,\aleph_\alpha}\right)^{\mathrm{cf}\,(\aleph_\alpha)}.$$

Proof. Let $\kappa = \mathrm{cf}\,(\aleph_\alpha) = \mathrm{cf}\,(\alpha)$. Then $\aleph_\alpha = \sup_{i<\kappa} \aleph_{\alpha_i}$, for some $\alpha_i < \alpha$ $(i < \kappa)$. Therefore

$$\aleph_\alpha = \max\{\kappa, \aleph_\alpha\} = \kappa.\aleph_\alpha = \sum_{i<\kappa} \aleph_{\alpha_i}.$$

Then we have

$$\begin{aligned}
2^{\aleph_\alpha} &= 2^{\sum_{i<\kappa} \aleph_{\alpha_i}} \\
&= \prod_{i<\kappa} 2^{\aleph_{\alpha_i}} \\
&\le \prod_{i<\kappa} 2^{<\aleph_\alpha} \\
&= \left(2^{<\aleph_\alpha}\right)^\kappa \\
&\le \left(2^{\aleph_\alpha}\right)^\kappa \\
&= 2^{\aleph_\alpha . \kappa} \\
&= 2^{\aleph_\alpha}.
\end{aligned}$$

$\Diamond$

The following theorem shows that there are constraints on the continuum function at singular cardinals beyond those we have already considered. In fact, the full list of such constraints is still unknown.

Theorem 11.19 (Bukovský[6]—Hechler[7]) *Let $\aleph_\alpha$ be a singular cardinal such that the continuum function on $[0, \aleph_\alpha)$ is eventually constant; that is, there exists $\beta < \alpha$ such that*

$$2^{\aleph_\gamma} = 2^{\aleph_\beta} \qquad (\beta \le \gamma < \alpha);$$

then

$$2^{\aleph_\alpha} = 2^{\aleph_\beta}.$$

Proof. Let $\kappa = \mathrm{cf}\,(\alpha)$ and let $\gamma = \max\{\beta, \kappa\}$. Then

$$\begin{aligned}
2^{\aleph_\alpha} &= \left(2^{<\aleph_\alpha}\right)^\kappa, \text{ by the lemma,} \\
&= \left(2^{\aleph_\gamma}\right)^\kappa, \text{ since } \gamma > \beta, \\
&= 2^{\aleph_\gamma}, \text{ since } \gamma > \kappa, \\
&= 2^{\aleph_\beta}, \text{ since } \gamma > \beta.
\end{aligned}$$

$\Diamond$

Another constraint on the continuum function is the following theorem of Silver[8].

Theorem 11.20 *Let $\aleph_\alpha$ be a singular cardinal with* $\mathrm{cf}\,(\alpha) > \omega$. *Then $\aleph_\alpha$ cannot be the first point at which (GCH) fails; i.e.*

$$2^{\aleph_\beta} = \aleph_{\beta+1}\ (\beta < \alpha) \ \Rightarrow\ 2^{\aleph_\alpha} = \aleph_{\alpha+1}.$$

[6] L. Bukovsky, 'The continuum problem and the powers of alephs', *Comment. Math. Univ. Carolinae*, **6** (1965), 181–197.

[7] S.H. Hechler, 'Powers of singular cardinals and a strong form of the negation of the generalized continuum hypothesis', *Z. Math. Logik Grundlagen Math.*, **19** (1973), 83–84.

[8] J. Silver, 'On the singular cardinals problem' (Proc. Int. Cong. Math. Vancouver (1974), 265–268

The hypothesis $\mathrm{cf}(\alpha) > \omega$ is needed, since Magidor[9] proved (assuming that there is a huge cardinal larger than a supercompact cardinal) that (ZFC) is consistent with the scenario: $2^{\aleph_n} = \aleph_{n+1}$ $(n < \omega)$, but $2^{\aleph_\omega} = \aleph_{\omega+2}$.

Finally, in this chapter, we remark that it is possible for (GCH) to fail everywhere: more precisely, Woodin[10] has shown that (if there is a supercompact cardinal) there is a model of (ZFC) in which

$$2^{\aleph_\alpha} = \aleph_{\alpha+2} \qquad (\alpha \in \mathrm{On}).$$

12 Large Cardinals

One way to construct a fairly large cardinal is to ask: can you have $\aleph_\alpha = \alpha$? The answer is 'yes'. To construct such an α, define a sequence of cardinals α_n by:

$$\alpha_0 = \aleph_0, \qquad \alpha_{n+1} = \aleph_{\alpha_n} \qquad (n \in \omega).$$

Then $\alpha = \sup_{n\in\omega} \alpha_n$ satisfies $\aleph_\alpha = \alpha$. However, in one sense at least, this α is small; it is a singular cardinal with $\mathrm{cf}(\alpha) = \omega$.

12.1 Inaccessible cardinals

Let us recall the definitions of regular and singular cardinals and add some closely related definitions.

Definitions 12.1 A cardinal $\aleph_\alpha$ is said to be:

1. *regular* if $\mathrm{cf}(\aleph_\alpha) = \aleph_\alpha$;
2. *singular* if $\mathrm{cf}(\aleph_\alpha) < \aleph_\alpha$;
3. a *(weak) limit cardinal* if α is a limit ordinal; equivalently if
$$\aleph_{\beta+1} < \aleph_\alpha \qquad (\beta < \alpha).$$
4. a *strong limit cardinal* if
$$2^{\aleph_\beta} < \aleph_\alpha \qquad (\beta < \alpha).$$

Clearly, every strong limit cardinal is a weak limit cardinal and, if (GCH) holds, then 'weak limit' and 'strong limit' are synonymous.

The standard example (the smallest example) of a singular cardinal is $\aleph_\omega$; the standard (and smallest) example of a weak limit cardinal is also $\aleph_\omega$. In fact, we have shown (Proposition 11.8) that all singular cardinals are weak limit cardinals. The converse implication is questionable.

Definitions 12.2 An uncountable cardinal is said to be:

1. *weakly inaccessible* if it is a regular weak limit cardinal;
2. *(strongly) inaccessible* if it is a regular strong limit cardinal.

Again, 'strong' implies 'weak' and they are equivalent if (GCH) holds.

The idea here is that there are two ways of getting a large cardinal $\boldsymbol{a}$ from cardinals $\boldsymbol{b} < \boldsymbol{a}$. One is by taking limits of transfinite sequences of length $< \boldsymbol{a}$ of cardinals $< \boldsymbol{a}$; this gives access to singular cardinals. The other is by going up to the next largest cardinal (in the weak case) or from $\boldsymbol{a}$ to $2^{\boldsymbol{a}}$ (in the strong case); this gives access to the non-limit cardinals. Cardinals which are not 'accessible' from below by either of these methods are called 'inaccessible'.

The example with which we started this section, the smallest α with $\aleph_\alpha = \alpha$ is not inaccessible (it is singular); but it is true that every inaccessible has that property.

[9] M. Magidor, 'On the singular cardinals problem II' *Ann. of Math.,* **106** (1977), 517–547.

[10] see M. Foreman and H. Woodin, 'The generalized continuum hypothesis can fail everywhere', *Ann.Math.*, **133** (1991), 1–35; **MR** 91k:03130

Proposition 12.3 *If $\aleph_\kappa$ is weakly inaccessible, then $\aleph_\kappa = \kappa$.*

Proof.

$$\begin{aligned} \aleph_\kappa &= \mathrm{cf}\,(\aleph_\kappa)\,, \text{ since } \aleph_\kappa \text{ is regular,} \\ &= \mathrm{cf}\,(\kappa)\,, \text{ since } \kappa \text{ is a limit ordinal,} \\ &\leq \kappa \\ &\leq \aleph_\kappa. \end{aligned}$$

Therefore $\kappa = \aleph_\kappa$. $\Diamond$

Theorem 12.4 (GCH) *A limit cardinal κ is inaccessible if and only if $\sum_{\lambda<\kappa} \kappa^\lambda = \kappa$.*

Proof. Suppose κ is inaccessible. If $\lambda < \kappa$, then $\lambda < \mathrm{cf}\,(\kappa)$, since κ is regular, so $\kappa^\lambda = \kappa$ by Theorem 11.13(i). Therefore

$$\sum_{\lambda<\kappa} \kappa^\lambda = \sum_{\lambda<\kappa} \kappa = \kappa.\kappa = \kappa.$$

(Actually, with more care, the above can be proved without (GCH).)

Conversely, suppose κ is a limit cardinal which is not inaccessible and therefore, by (GCH), not weakly inaccessible. Then κ must be singular, so $\mathrm{cf}\,(\kappa) < \kappa$, so

$$\sum_{\lambda<\kappa} \kappa^\lambda \geq \kappa^{\mathrm{cf}\,(\kappa)} > \kappa.$$

$\Diamond$

Thus inaccessibles are characterized, among limit cardinals, as the fixed points of a certain 'function' f from cardinals to cardinals. Other types of large cardinal are similarly defined as fixed points of suitable functions.

12.2 The Cumulative Hierarchy

Definition 12.5 We define the **cumulative hierarchy** of sets V_α ($\alpha \in \mathrm{On}$) by transfinite induction:

$$\begin{aligned} V_0 &= \emptyset, \\ V_{\alpha+1} &= \mathcal{P}(V_\alpha), \\ V_\mu &= \bigcup_{\alpha<\mu} V_\alpha \text{ for } \mu \text{ a limit ordinal.} \end{aligned}$$

In fact, all sets occur somewhere in this hierarchy. In our presentation, this is a consequence of the Axiom of Foundation (AF), but one could present it as an alternative axiom, in place of (AF).

Proposition 12.6

$$V = \bigcup_{\alpha\in\mathrm{On}} V_\alpha.$$

Proof. Let $x \in V \setminus \bigcup_{\alpha\in\mathrm{On}} V_\alpha$. We shall define another such set y with $y \subseteq \bigcup_{\alpha\in\mathrm{On}} V_\alpha$. If $x \subseteq \bigcup_{\alpha\in\mathrm{On}} V_\alpha$, then we set $y := x$. Otherwise we proceed as follows.

We form the **$\in$-transitive closure** t of x by setting $t_0 := x$, then $t_{n+1} := \bigcup t_n$ (meaning $\bigcup_{u\in t_n} u$) ($n \in \omega$), and finally

$$t := \bigcup_{n\in\omega} t_n.$$

Then t is a set and it is the smallest $\in$-transitive set containing x.

Let $s = t \setminus \bigcup_{\alpha\in\mathrm{On}} V_\alpha$. This is a set, by the Axiom of Subsets, since t is a set. and since $x \not\subseteq \bigcup_{\alpha\in\mathrm{On}} V_\alpha$, it is non-empty. By (AF), there is a $y \in s$ such that $y \cap s = \emptyset$. By the construction of t, we have $y \subseteq t$; therefore

$$y \setminus \bigcup_{\alpha\in\mathrm{On}} V_\alpha = y \cap s = \emptyset,$$

i.e.

$$y \subseteq \bigcup_{\alpha\in\mathrm{On}} V_\alpha.$$

By the definition of s, we have $y \notin \bigcup_{\alpha\in\mathrm{On}} V_\alpha$.

(If we had stated (AF) in the form 'every non-empty *class* has an element which is disjoint from it' then we should have achieved this stage more quickly by applying (AF) to $V \setminus \bigcup_{\alpha\in\mathrm{On}} V_\alpha$ directly; what we have done above is essentially to prove the class form. We used the set form in our axiomatization because (ZF) does not, formally, allow class variables.)

Setting $y_\alpha = y \cap V_\alpha$ ($\alpha \in \mathrm{On}$) we have

$$y = \bigcup_{\alpha\in\mathrm{On}} y_\alpha.$$

This is the situation described in Lemma 4.20. The conclusion is that the transfinite sequence (y_α) is eventually constant; that is, $y \subseteq V_\beta$ for some ordinal β. Then $y \in V_{\beta+1}$, contradicting $y \notin \bigcup_{\alpha\in\mathrm{On}} V_\alpha$. This completes the proof. $\Diamond$

Lemma 12.7 *Let κ be the first inaccessible cardinal. Then $|V_\alpha| < \kappa$ ($\alpha < \kappa$), and $|V_\kappa| = \kappa$.*

Proof. We prove $|V_\alpha| < \kappa$ by induction on α. It is certainly true for $V_0 = \emptyset$. If $|V_\alpha| = \boldsymbol{n} < \kappa$ then $V_{\alpha+1}| = 2^{\boldsymbol{n}} < \kappa$ since κ is a strong limit cardinal. If α is a limit ordinal and $|V_\beta| < \kappa$ for all $\beta < \alpha$, then

$$|V_\alpha| = \sup_{\beta<\alpha} |V_\beta| \le \kappa,$$

but we cannot have $\kappa = \sup_{\beta<\alpha} |V_\beta|$, since this would imply $\mathrm{cf}\,(\kappa) \le \alpha < \kappa$, contradicting the regularity of κ. This completes the induction proof of our first assertion.

For the second assertion, since

$$V_\kappa = \bigcup_{\alpha<\kappa} V_\alpha$$

we have

$$|V_\kappa| \le \kappa.\kappa = \kappa,$$

by Lemma 11.9. Conversely, since $|V_{\alpha+1} \setminus V_{\alpha+1}| \ge 1$ ($\alpha < \kappa$), we have $|V_\kappa| \ge \kappa$. $\Diamond$

12.3 Models of set theory

Let $\mathcal{L}$ be a **language**; that is, a finite set of distinguished predicates. For example, the language of set theory consists just of the distinguished binary predicate $\in$, the language of ring theory has two binary predicates $+$ and '.'. A **theory** is a set of sentences in the predicate calculus with equality and the distinguished predicates of the language. By a **sentence** we mean a **well-formed formula**; roughly, this means a string of symbols which makes sense — formally, there is an obvious inductive definition which may be found in any text on mathematical logic.

Definition 12.8 A **model** for a language $\mathcal{L}$ is a set $\mathfrak{M}$ equipped with a relation $P^{\mathfrak{M}}$ corresponding to each of the predicates P of the language. (Usually we omit the superscript.)

A model $\mathfrak{M}$ is a model for the theory T if whenever each of the sentences is **interpreted** in the model $\mathfrak{M}$, using the relations $P^{\mathfrak{M}}$ for the predicates P, the sentence is found to be true. We write $\mathfrak{M} \vDash T$ and say '$\mathfrak{M}$ **satisfies** T'.

The notion of **interpretation** of a sentence in a model is again something which we take as obvious, but which can be defined formally by induction on the length of the sentence. The formal definition can be found in any text on mathematical logic covers some model theory. It is easiest to think in terms of examples. If T is consists of the axioms for associative rings, in the language of ring theory, then a 'model of T' is just a ring.

A **model of set theory** is, therefore, a pair $\mathfrak{M} = (M, E)$ consisting of a class M (we vary the definition used for models of algebraic theories to allow M to be a proper class) and a binary relation E on M corresponding to the binary predicate $\in$. We can consider some special kinds of models of set theory.

Definition 12.9 A model $\mathfrak{M} = (M, E)$ of set theory is an **$\in$-model** if $E = \in \cap M^2$, that is, the elements of M are sets (of course — the model is built inside the usual universe of sets) and the relation E that represents $\in$ when the language is interpreted in $\mathfrak{M}$ is just the membership relation $\in$ restricted to M. For $\in$-models, we can just speak of 'the model M'.

An $\in$-model M is said to be **transitive** if M is an $\in$-transitive class; that is, $x \in y \in M \quad \Rightarrow \quad x \in M$.

Let us call expressions of the form $(\exists x \in X)$ and $(\forall x \in X)$ **restricted quantifiers**. We shall say that a formula ϕ is a **restricted formula** if it contains no quantifiers other than restricted quantifiers.

Lemma 12.10 *If M is a transitive model and ϕ is a restricted formula, then for all $x_1, \ldots, x_n \in M$,*

$$M \vDash \phi(x_1, \ldots, x_n) \text{ if and only if } \phi(x_1, \ldots, x_n). \tag{9}$$

Proof. The proof is by induction on the length of the formula ϕ. We suppose that (9) holds for all formulae of length strictly smaller than ϕ and prove it for ϕ.

If ϕ is an 'atomic formula', i.e. a formula $x = y$ or $x \in y$, then (9) holds. If ϕ is of the form $\neg\psi$, $\psi \& \chi$, $\psi \vee \chi$ or $\psi \Rightarrow \chi$, where ψ and χ are shorter formulae, then (9) holds for ψ and χ by the induction hypothesis, and therefore (9) holds for ϕ.

The crucial part of the proof concerns the case when $\phi(X, y, \ldots)$ is of the form $(\exists x \in X)\psi(x, X, y, \ldots)$ or $(\forall x \in X)\psi(x, X, y, \ldots)$. Again, ψ is a shorter formula, so (9) holds for ψ. We prove the first of these cases and leave the second as an exercise. (Alternatively, one can use the rule '$\forall \equiv \neg\exists\neg$' to dispense with universal restricted quantifiers.)

The statement $M \vDash \phi(X, y, \ldots)$, is then

$$M \vDash \exists x (x \in X \ \& \ \psi(x, X, y, \ldots)).$$

By the induction hypothesis on ψ, this is equivalent to

$$(\exists x \in M)(x \in X \ \& \ M \vDash \psi(x, X, y, \ldots)). \tag{10}$$

Clearly (10) implies

$$(\exists x \in M)(x \in X \ \& \ \psi(x, X, y, \ldots)),$$

so $(\exists x \in X)\psi(x, X, y, \ldots)$, i.e. $\phi(X, y, \ldots)$. Conversely, suppose $\phi(X, y, \ldots)$, i.e. $(\exists x \in X)\psi(x, X, y, \ldots)$. Here, $X, y, \ldots \in M$, but we do not necessarily have $x \in M$. However, since M is transitive and $x \in X \in M$, we do have $x \in M$, so (10) follows. $\Diamond$

Corollary 12.11 *Every transitive model satisfies the Axiom of Extensionality and the Axiom of Foundation.*

We rewrite (AE) and (AF) as restricted formulae:

(AE)
$$\forall z (z \in x \iff z \in y) \quad \Rightarrow \quad x = y$$

becomes

$$((\forall z \in x)(z \in y) \ \& \ (\forall z \in y)(z \in x)) \quad \Rightarrow \quad x = y;$$

and

(AF)
$$\exists y (y \in x) \quad \Rightarrow \quad \exists y (y \in x \ \& \ \neg \exists z \, (z \in x \ \& \ z \in y))$$

becomes

$$(\exists y \in x)(y = y) \quad \Rightarrow \quad (\exists y \in x)(\forall z \in x)(z \notin y).$$

12.4 Independence of the existence of inaccessibles

Theorem 12.12 *(ZFC) If κ is the first inaccessible cardinal, then V_κ is a transisitve $\in$-model for ZFC+(there is no inaccessible cardinals).*

Proof. We write $M = V_\kappa$. First we observe that V_κ is transitive: it is easy to prove by transfinite induction that V_α is transitive for all ordinals α. It then follows from Corollary 12.11 that $M \models (AE) + (AF)$.

We must show that M satisfies the remaining axioms of (ZFC).

$$(AS) \qquad \exists y \forall z (z \in y \iff z \in x \,\&\, P(z));$$

Given $x \in M$, the axiom (AS) provides a set y; we have to show that $y \in M$. Now $x \in M$ implies $x \in V_\alpha$ for some $\alpha < \kappa$. Since V_α is transitive, every $z \in x$ is a member of V_α. Therefore $y \subseteq V_\alpha$, so $y \in V_{\alpha+1} \subseteq M$. Thus $M \models (AS)$.

$$(AR) \qquad \forall x \exists y \forall z (P(x,z) \iff z \in y) \;\Rightarrow\; \forall u \mathcal{M}_y \left(\exists x (x \in u \,\&\, P(x,y))\right).$$

Suppose that

$$M \models \forall x \exists y \forall z (P(x,z) \iff z \in y).$$

Let

$$F = \{(x,y) \in M \times M : y = \{z \in M : P(x,z)\}\}.$$

Then $F : M \to M$ is a function. To show

$$M \models \forall u \mathcal{M}_y \left((\exists x \in u) P(x,y))\right)$$

it suffices to show that if $u \in M$ then $F``u \in M$; (using the transitivity of $M$ again). If $u \in M$, then $u \in V_\alpha$ for some $\alpha < \kappa$, so $u \subseteq V_\alpha$ (since $V_\alpha$ is transitive), so $|u| \le |V_\alpha| < \kappa$, by Lemma 12.7. Then $|F``u| \le |u| < \kappa$. Each $s \in F``u$ is a member of some $V_{\alpha_s}$ with $\alpha_s < \kappa$. We cannot have $\sup_s \alpha_s = \kappa$, since that would be to express $\kappa$ as the sup of $|F``u| < \kappa$ smaller ordinals. Therefore $\sup_s \alpha_s = \alpha < \kappa$. Thus $s \in V_\alpha$ $(s \in F``u)$. This means that $F``u \subseteq V_\alpha$, so $F``u \in V_{\alpha+1}$, so $F``u \in M$, as desired.

$$(AP) \qquad \forall x \mathcal{M}_y \left(y \subseteq x\right).$$

The Power Set Axiom is satisfied in M; essentially, because $x \in M$ implies $x \in V_\alpha$ for some $\alpha < \kappa$ and so $\mathcal{P}(x) \in V_{\alpha+1} \subseteq M$. However, this argument disguises a subtlety: we need to check that, for $x \in M$, $\mathcal{P}(x)$ within the model means the same as $\mathcal{P}(x)$ in V.

For this, we need to know is that, for $x, y \in M$,

$$(M \models y \subseteq x) \iff y \subseteq x.$$

The point here is that $M \models y \subseteq x$ means that for every $u \in M$, if $u \in y$ then $u \in x$, whereas $y \subseteq x$ in V means that the same is true for all $u \in V$. The fact that these two are equivalent comes from the transitivity of M, since

$$u \in y \subseteq x \in M \;\Rightarrow\; u \in x \in M \;\Rightarrow\; u \in M.$$

We could skip the the Null Set Axiom

$$(AN) \qquad \exists x \forall y (y \notin x),$$

as it was superseded by the Axiom of Infinity. However, we may as well observe that M satisfies (AN), since all this involves is noting that the null set $\emptyset$ of V is a member of M and consequently forms the null set for M.

(AI) $$\exists x(\emptyset \in x \ \& \ \forall y(y \in x \quad \Rightarrow \quad S(y) \in x)).$$

We have just noted that $\emptyset$ is the same in M and V. It is easy to see that

$$S(x) = \iota y(z \in y \iff (z \in x \vee z = x))$$

satisfies $S(x) \in M$ whenever $x \in M$ and that

$$M \vDash [S(x) = \iota y(z \in y \iff (z \in x \vee z = x))];$$

in other words, $S(x)$ in V is the same as $S(x)$ interpreted in the model M. An example of an x proving (AI) is the ordinal ω. Since $\omega \in M$, it follows that $M \vDash$ (AI).

We can proceed in this fashion, showing the ideas of 'f being a function' and f'x are interpreted in the same way in M as in V. We must show that $M \vDash (AC)$, assuming the Axiom of Choice (AC) holds in V. Given a set $x \in M$, we need only show that the choice function f on x produced by (AC) in V belongs to M. Now if $x, y \in V\alpha$ then all $\xi \in x$ and $\eta \in y$ are in V_α, so unordered pairs $\{\xi, \eta\}$ are in $V_{\alpha+1}$, ordered pairs (ξ, η) are in $V_{\alpha+2}$, and functions $f : x \to y$ are in $V_{\alpha+3}$. In particular, if $x \in V_\alpha$, then $\bigcup x \in V_\alpha$ and so any choice function on x is in $V_{\alpha+3}$. Thus $M \vDash (AC)$.

Working further through the definitions of (ZFC), it can easily be shown that:

1. $M \vDash \alpha \in \mathrm{On}$ iff $\alpha \in \mathrm{On}$;
2. $M \vDash \alpha$ is a cardinal iff α is a cardinal;
3. $M \vDash \alpha$ is a regular cardinal iff α is a regular cardinal;
4. $M \vDash \alpha$ is an inaccessible cardinal iff α is an inaccessible cardinal.

Therefore, since κ is the least inaccessible cardinal and $M = V_\kappa$, we have

$$M \vDash \text{ there is no inaccessible cardinal.}$$

$\Diamond$

Remark 12.13 In view of what we have observed about the power set operation in M, it is clear that if we assume (GCH) in V, then $M \vDash$ (GCH).

13 Constructibility

In this section we shall describe briefly Gödel's model which showed the relative consistency of (AC) and (GCH). Another view is that we describe a version of set theory in which the sets which exist are those which we can describe. In the construction of the Cumulative Hierarchy, we put into $V_{\alpha+1}$ *all* the subsets of V_α, whether they were in any sense describable or not. Now we shall be more careful!

Definition 13.1 Given a set X we denote by $\mathcal{L}_X$ the language of set theory with a constant $\overline{x}$ for each $x \in X$.

We shall say that a set Y is an X**-definable** subset of X if, for some formula ϕ of $\mathcal{L}_X$ with one free variable,

$$Y = \{x \in X : X \vDash \phi(\overline{x})\},$$

where $\vDash$ has the obvious meaning.

We write $\mathrm{Def}(X)$ for the set of all X-definable subsets of X.

Definition 13.2 We define the **constructible hierarchy** L_α ($\alpha \in \mathrm{On}$) by induction on α:

$$\begin{aligned} L_0 &= \emptyset; \\ L_{\alpha+1} &= \mathrm{Def}(L_\alpha); \\ L_\lambda &= \bigcup_{\alpha<\lambda} L_\alpha \text{ for limit ordinals } \lambda. \end{aligned}$$

We write

$$L = \bigcup_{\alpha\in\mathrm{On}} L_\alpha,$$

which is a proper class.

Lemma 13.3 *(i) For all $\alpha < \beta$ we have $L_\alpha \subseteq L_\beta$ and $L_\alpha \in L_\beta$.*

(ii) For all α the set L_α is $\in$-transitive. Hence the class L is $\in$-transitive, and so $L \vDash$ (AE)+(AF).

(iii) For all α we have $\alpha \in L_{\alpha+1}$. Hence $\mathrm{On} \subseteq L$.

Proof. The proofs are similar to those for the cumulative hierarchy (V_α). $\Diamond$

Theorem 13.4 Gödel[11] *The class L is a transitive model for (ZFC) +(CH). Consequently, if (ZF) is consistent then so is (ZFC).*

Proof. We do not have time for a full proof, for which we refer to Keith Devlin's book *Constructibility*[12] or *Aspects of Constructibility*[13]. A key point of difference from the proof with V_κ is that $L \vDash$ (AC). This is shown by showing $L \vDash$ (WO), where (WO) is the statement 'every set can be well-ordered', which is equivalent to (AC). The idea, essentially, is that L has a well-ordered grading $\bigcup L_\alpha$ and on each of the L_α one defines a well-order using transfinite induction, the ordering on $L_{\alpha+1}$ being formed from that on L_α using the fact that each element of $L_{\alpha+1}$ is a definable subset of L_α and so corresponds to a formlua $\phi(x, \overline{x_1}, \dots, \overline{x_n})$. The proof that $L \vDash$ (GCH) is more complicated. $\Diamond$

The constructible hierarchy, introduced for these independence proofs, has considerable interest in its own right. It gives us a new axiom, the **Axiom of Constructibility**, which needs no abbreviation as it is just

$$V = L,$$

that is: every set is constructible. Thus what Gödel showed was that L is a model for (ZF)+($V = L$) and that

$$(V = L) \quad \Rightarrow \quad \text{(AC) \& (GCH)}.$$

The Axiom of Constructibility has other consequences.

Definition 13.5 We say that a abelian group G is a **W-group** if $\mathrm{Ext}(G, \mathbb{Z}) = 0$; that is, if whenever H is an abelian group with a subgroup $K \equiv \mathbb{Z}$ such that $H/K \equiv G$, we must have $H = K \oplus M$ with $M \equiv G$.

Every free abelian group is a W-group; the **Whitehead Problem (1951)** asks whether every W-group is free. Stein showed, in 1951, that every countable W-group is free. In 1975, Shelah solved the Whitehead problem positively — assuming $V = L$. For further details we refer to Keith Devlin's book *The Axiom of Constructibility: a guide for the mathematician*[14].

[11] Kurt Gödel, *The Consistency of the Axiom of Choice and of the Generalized Continuum-Hypothesis with the Axioms of Set Theory,* (Princeton University Press, Annals of Math. Studies **3**, 1940).

[12] Springer, Perspectives in Mathematical Logic, 1984

[13] Springer, Lecture Notes in Math. **354**, 1973

[14] Springer, Lecture Notes in Math. **617**, 1977

14 Appendix: weak forms of (AC)

To describe the Ultrafilter Theorem and the Boolean Prime Ideal Theorem we need some definitions.

Definition 14.1 A **Boolean algebra** is a set B with binary operations $\vee$ and $\wedge$, a unary operation $x \mapsto \overline{x}$ and distinguished elements (0-ary operations) 0 and 1, satisfying the following axioms:

$$\begin{array}{rclcrcl}
a \vee a &=& a & \qquad & a \wedge a &=& a \\
a \vee b &=& b \vee a & & a \wedge b &=& b \wedge a \\
(a \vee b) \vee c &=& a \vee (b \vee c) & & (a \wedge b) \wedge c &=& a \wedge (b \wedge c) \\
a \wedge (b \vee c) &=& (a \wedge b) \vee (a \wedge c) & & a \vee (b \wedge c) &=& (a \vee b) \wedge (a \vee c) \\
a \vee \overline{a} &=& 1 & & a \wedge \overline{a} &=& 0 \\
a \vee 0 &=& a & & a \wedge 1 &=& a.
\end{array}$$

We say B is **non-trivial** if $0 \neq 1$.

Definition 14.2 In a Boolean algebra B we introduce a relation $\leq$ by:

$$a \leq b \iff a \vee b = b.$$

Proposition 14.3 *In a Boolean algebra B:*

1. *$\leq$ is a partial order;*
2. *$a \leq b \iff a \wedge b = a$;*
3. *$a \wedge b = \inf\{a, b\}$ and $a \vee b = \sup\{a, b\}$ in the sense of the partial order $\leq$;*
4. *$0 = \inf B$; $1 = \sup B$;*
5. *$a \leq b \;\Rightarrow\; a \vee c \leq b \vee c$ and $a \leq b \;\Rightarrow\; a \wedge c \leq b \wedge c$;*
6. *$a \vee b = 1 \iff a \geq \overline{b}$ and $a \wedge b = 0 \iff a \leq \overline{b}$; hence*
$$a = \overline{b} \iff a \wedge b = 0 \;\&\; a \vee b = 1;$$
7. *$\overline{0} = 1$ and $\overline{1} = 0$;*
8. *$\overline{\overline{a}} = a$;*
9. *$\overline{a \wedge b} = \overline{a} \vee \overline{b}$ and $\overline{a \vee b} = \overline{a} \wedge \overline{b}$;*
10. *$a \leq b \iff \overline{b} \leq \overline{a}$.*

Definition 14.4 We say that a non-empty subset I of a Boolean algebra B is an **ideal** if:

(a) $a, b \in I \;\Rightarrow\; a \vee b \in I$;

(b) $a \in I \;\&\; b \leq a \;\Rightarrow\; b \in I$;

(c) $1 \notin I$.

We say that a non-empty subset F of a Boolean algebra B is an **filter** if:

(a) $a, b \in F \;\Rightarrow\; a \wedge b \in F$;

(b) $a \in F \;\&\; b \geq a \;\Rightarrow\; b \in F$;

(c) $0 \notin F$.

Examples 14.5 Given $a \in B$ with $a \neq 1$, the set $I_a = \{x \in B : x \leq a\}$ is an ideal, called the **principal ideal** generated by a. Given $a \in B$ with $a \neq 0$, the set $F_a = \{x \in B : x \geq a\}$ is a filter, called the **principal filter** generated by a.

Proposition 14.6 *For an ideal I of a Boolean algebra B the following are equivalent:*

(a) for all $a \in B$, either $a \in I$ or $\overline{a} \in I$;

(b) I is maximal in the set of all ideals of B, ordered by inclusion.

(c) if $a \wedge b \in I$ then at least one of a, b is in I;

Some authors write $+$ for $\vee$ and . for $\wedge$ which makes the term 'ideal' natural in analogy with ring theory (though it makes the existence of two distributive laws look odd!). Continuing this analogy, we say that I is a **prime ideal** if it satisfies the conditions of this theorem.

Proof.

$(a) \Rightarrow (b)$ Notice that, for any ideal I we can never have both $a \in I$ and $\overline{a} \in I$. The fact that $(a) \Rightarrow (b)$ follows immediately.

$(b) \Rightarrow (c)$ Suppose I is maximal and $a \wedge b \in I$ with neither a nor b belonging to I. Let

$$J = \{x \in B : x \leq a \vee y \text{ for some } y \in I\}.$$

Then it is easy to check that J satisfies the first two conditions for being an ideal: first, if $x_1, x_2 \in J$, then $x_1 \leq a \vee y_1$, $x_2 \leq a \vee y_2$, with $y_1, y_2 \in I$, so

$$x_1 \vee x_2 \leq (a \vee y_1) \vee (a \vee y_2) = a \vee (y_1 \vee y_2)$$

and $y_1 \vee y_2 \in I$, so $x_1 \vee x_2 \in J$; secondly, if $z \leq x$ and $x \in J$ with, say, $x \leq a \vee y$ and $y \in I$, then $z \leq a \vee y$ and so $z \in J$. Now J properly contains I and I is a maximal ideal, so we must have $1 \in J$. Therefore $1 \leq a \vee x$ for some $x \in I$, so $1 = a \vee x$. Likewise $1 = b \vee y$ for some $y \in I$. It follows that

$$1 = (a \vee x) \wedge (b \vee y) = (a \wedge b) \vee (a \wedge y) \vee (x \wedge (b \vee y)) \in I,$$

since $a \wedge b \in I$, by hypothesis,

$$a \wedge y \leq y \in I \quad \Rightarrow \quad a \wedge y \in I$$

and

$$x \wedge (b \vee y) \leq x \in I \quad \Rightarrow \quad x \wedge (b \vee y) \in I.$$

This contradiction proves (c).

$(c) \Rightarrow (a)$ This implication is immediate because $a \wedge \overline{a} = 0 \in I$.

$\Diamond$

Clearly the definition of 'filter' is dual to that of 'ideal', so the set $\overline{I} := \{x \in B : \overline{x} \in I\}$ is a filter if and only if I is an ideal. Thus we have the following dual proposition.

Proposition 14.7 *For an filter F of a Boolean algebra B the following are equivalent:*

(a) for all $a \in B$, either $a \in F$ or $\overline{a} \in F$;

(b) F is maximal in the set of all filters of B, ordered by inclusion.

(c) if $a \vee b \in F$ then at least one of a, b is in F;

Definition 14.8 We say that F is an **ultrafilter** if the equivalent conditions of this proposition hold.

It is now easy to see that (BPIT) and (UFT) are just dual statements. The proof of either, by applying Zorn's Lemma to the set of all ideals/filters ordered by inclusion, is straightforward.

To illustrate the use of (UFT), it is convenient to refer to the filter approach to general topology. The term 'filter' here will refer to a filter in the Boolean algebra $\mathcal{P}(X)$ of subsets of the given topological space X.

Definition 14.9 Given a point $x \in X$, we write $\mathcal{N}_x$ for the **neighbourhood filter** of x, that is, the set of all neighbourhoods of x. We say that a filter $\mathcal{F}$ **converges** to a point $x \in X$ if $\mathcal{F} \supseteq \mathcal{N}_x$.

Theorem 14.10 *(a) (ZF) A topological space is Hausdorff if and only if no filter converges to more than one point.*

(b) (ZF)+(UFT) A topological space is compact if and only if every ultrafilter converges.

(c) (ZF) A function $f : X \to Y$ is continuous if and only if $f(\mathcal{F}) \to f(x)$ whenever $\mathcal{F} \to x$.

Proof. The proofs of (a) and (c) are straightforward; we prove (b).

Suppose that X is compact and that $\mathcal{U}$ is an ultrafilter on X. Let $\overline{\mathcal{U}}$ be the set of all closures of members of $\mathcal{U}$. Then, since $\mathcal{U}$ has the finite intersection property, so does $\overline{\mathcal{U}}$. Since X is compact, there exists a point $x \in \bigcap \overline{\mathcal{U}}$. Therefore, for every neighbourhood N of x, $X \setminus N \notin \mathcal{U}$. Since $\mathcal{U}$ is an ultrafilter $X \setminus N \notin \mathcal{U}$ implies $N \in \mathcal{U}$. Thus $\mathcal{U} \to x$. We have only used (ZF).

For the converse, suppose that every ultrafilter on X converges. Let $\mathcal{C}$ be a family of closed subsets of X with the finite intersection property. Let

$$\mathcal{F} = \{Y \subseteq X : Y \supseteq F_1 \cap \ldots \cap F_n \text{ for some } F_1, \ldots \mathbb{F}_n \in \mathcal{C}\}.$$

Then $\mathcal{F}$ is a filter on X. Using (UFT), let $\mathcal{U}$ be an ultrafilter containing $\mathcal{F}$. Then $\mathcal{U}$ converges; say $\mathcal{U} \to x$. For every neighbourhood N of x, $N \in \mathcal{U}$, so $N \cap F \neq \emptyset$ for every $F \in \mathcal{U}$ and, in particular, for every $F \in \mathcal{C}$. Since the sets in $\mathcal{C}$ are closed, this implies $x \in \bigcap \mathcal{C}$. $\diamondsuit$

Application: Tychonoff's Product Theorem

The Axiom of Choice inevitably plays a rôle in the discussion of infinite product spaces. After all, it is equivalent to the statement that every Cartesian product of a family of non-empty sets is non-empty. We shall now prove that it is equivalent to the famous theorem of Tychonoff that every product of compact spaces is compact.

We recall the definition of the product topology. Given a family $\{X_i\}_{i \in I}$ of topological spaces, let $X = \prod_{i \in I} X_i$ denote the Cartesian product and $\pi_i : X \to X_i$ $(i \in I)$ the coordinate projections. The **product topology** on X is the weakest topology making all the maps π_i continuous. In terms of filters it is the topology such that $\mathcal{F} \to x$ in X iff $\pi_i(\mathcal{F}) \to \pi_i(x)$ for each $i \in I$.

Theorem 14.11 Tychonoff's Product Theorem *In (ZF)+(AC), every product of compact spaces is compact.*
In (ZF)+(UFT), every product of compact Hausdorff spaces is compact.

Proof. With the notation above, we have to show that if all the X_i are compact then every ultrafilter $\mathcal{U}$ in X converges. If $\mathcal{U}$ is an ultrafilter in X, then each filter $\pi_i(\mathcal{U})$ is an ultrafilter and therefore converges, say $\pi_i(\mathcal{U}) \to x_i$. If all the X_i are Hausdorff, then the limits x_i are unique. Otherwise, we use (AC) to choose, for each $i \in I$, a limit point x_i of $\pi_i(\mathcal{U})$. Let $x = (x_i)_{i \in I}$. Then $\mathcal{U} \to x$, by the definition of the product topology. $\diamondsuit$

Theorem 14.12 (Kelley) *Tychonoff's Product Theorem implies (AC), within (ZF).*

Proof. Assume Tychonoff's Product Theorem. Let x be a set on which we wish to have a choice function. For each $y \in x$, define the topological space $S_y = y \cup \{p_y\}$, where $p_y \notin y$, with the topology in which a set $G \subseteq S_y$ is open iff either

(i) $G \subseteq \{p_y\}$ or

(ii) $S_y \setminus G$ is finite.

It is easy to see that this is a compact (and T_1) topology on S_y. Tychonoff's Theorem then implies that the product space $X = \prod_{y \in x} S_y$ is compact.

In X we consider the family of closed sets $F_y = \pi_y^{-1}(y)$. This family has the finite intersection property: to show $F_{y_1} \cap \ldots \cap F_{y_n} \neq \emptyset$ we choose $s_i \in y$ $(1 \leq i \leq n)$, and define $x \in F_{y_1} \cap \ldots \cap F_{y_n}$ by

$$\pi_y(x) = \begin{cases} s_i & (y = y_i,\ 1 \leq i \leq n) \\ p_y & \text{otherwise.} \end{cases}$$

X is compact, there exists $z \in \bigcap_{y\in x} F_y$. Then $\pi_y(z) \in y$ $(y \in x)$; that is, $y \mapsto \pi_y(z)$ is a choice function on x. $\diamondsuit$

Actually, we have proved slightly more than stated: Tychonoff's Theorem for T_1-spaces implies (AC). It can be shown that Tychonoff's Theorem for Hausdorff spaces implies (UFT).

Theorem 14.12 is typical of many proofs that statements in ordinary mathematics are equivalent to (AC) or to one of its weaker forms.

15 Bibliography.

1. Krzysztof Ciesielski, 'Set theory for the working mathematician', (CUP, LMS Student Texts, **39**, 1997) 240pp., ISBN 0-521-59465-0, £13.95 (less 25% discount for LMS members); Library ref. ML 512.811(C). [A good small book on the main part of the course; after the basic theory it explores Martin's Axiom and associated models.]

2. R. L. Vaught, 'Set theory: an introduction' (Birkhäuser, 1985) ISBN 3-7643-3238-7, ML 512.811(V) [A very gentle introduction.]

3. Jean-Louis Krivine, 'Introduction to axiomatic set theory', (D. Reidel, Synthese Library, , 1971) 98pp., ISBN 90-227-0169-5. [A basic introduction.]

4. Devlin,K.J., 'Fundamentals of contemporary set theory', (Springer, Universitext, , 1979) 182pp., ISBN 3540904417. [Another 'set theory for working mathematicians': basic theory plus some remarks on constructibility and independence proofs.]

5. Thomas Jech, 'Set Theory 2/e', (Springer, Perspectives in Mathematical Logic, , 1997) 634pp., ISBN 3-540-63048-1, £64.50. [Probably the best modern text for further reading, but rather condensed.]

6. Azriel Levy, 'Basic Set Theory', (Springer, Perspectives in Mathematical Logic, , 1979) 391pp., ISBN 3-540-08417-7. [A good general text on the course, covering all aspects at a gentler pace than Jech.]

7. Gaisi Takeuti and Wilson M. Zaring, 'Introduction to Axiomatic Set Theory 2/e', (Springer, Graduate Texts in Mathematics, **1**, 1982) 246pp., ISBN 3-540-90683-5.

8. Gaisi Takeuti and Wilson M. Zaring, 'Axiomatic Set Theory', (Springer, Graduate Texts in Mathematics, **8**, 1973) 238pp., ISBN 3-540-90050-0. [A sequel to their 'Introduction'. The two books together give an introduction to basic set theory plus independence proofs, up to 1973.]

9. E. Mendelson, 'Introduction to mathematical logic' (van Nostrand, The University Series in Undergraduate Mathematics, 1963) [Logic and set theory. Reprinted by another publisher?]

10. J. R. Schoenfield, 'Mathematical logic', (Addison-Wesley, 1967). [Another good reference for logic and set theory.]

11. H.G.Dales and W.H.Woodin, 'An Introduction to Independence Theory for Analysts', (CUP, London Mathematical Society Lecture Note Series, **115**, 1987) 241pp., ISBN 0-521-33996-0. [What its title says, but centered around Dales' solution of Kaplansky's problem and the authors' subsequent study of its logical status.]

12. Moore,G.H., 'Zermelo's Axiom of Choice: its origins, development and influence', (Springer, Studies in the History of Mathematics and Physical Sciences, **8**, 1982) 410pp., ISBN 3-540-90670-3, ML 512.811(Z).

13. Thomas J. Jech, 'The Axiom of Choice', (North-Holland, Studies in Logic and the Foundations of Mathematics, **75**, 1973) 202pp., ISBN 0-444-10484-4, ML B 510.1(J). [A good general text on (AC); its various forms, basic independence proofs, and the possibility of life without (AC). In fact, everything you would want to know about (AC) (up to 1973).]

14. Paul Howard, (Eastern Michigan U.), and Jean E. Rubin, (Purdue U.), 'Consequences of the Axiom of Choice', (AMS, SURV, **59**, 1998) 432pp., ISBN 0-8218-0977-6, $89. [Everything you would not want to know about (AC)! A database (floppy disk attached) of equivalents and consequences of (AC).]

15. H. Rubin & J. Rubin, 'Equivalents of the Axiom of Choice' (North-Holland, Studies in Logic and the Foundations of Mathematics, 1963).

16. H. Rubin & J. Rubin, 'Equivalents of the Axiom of Choice II' (North-Holland). [The Rubin & Rubin books are the original studies of this subject—more readable than Howard & Rubin.]

17. Kanamori,A., 'The higher infinite: large cardinals in set theory from their beginnings', (Springer, Perspectives in Mathematical Logic, , 1994) 536pp., ISBN 3-540-57071-3, £77.50, ML 512.811 (K). [*The* book on large cardinals; but definitely a 'further reading' book, not a course text.]

18. Frank R. Drake, 'Set Theory: an introduction to large cardinals', (North-Holland, Studies in Logic and the Foundations of Mathematics, **76**, 1974) 351pp., ISBN 0-444-10535-2. ML B 512.811(D). [An older and much gentler introduction to large cardinals.]

19. K. Gödel, 'The consistency of the continuum hypothesis' (Princeton University Press, Annals of Mathematics Studies, **3**, 1940) ML 3 PER 510.5 [Original proof of the relative consistency of (AC)+(GCH).]

20. Devlin,K.J., 'The Axiom of Constructibility: a Guide for the Mathematician', (Springer, LNM, **617**, 1977) 96pp., ISBN 3-540-08520-3. [A short guide to $V = L$ with emphasis on applications.]

21. Devlin,K.J., 'Constructibility', (Springer, Perspectives in Mathematical Logic, 1984) 425pp., ISBN 3-540-13258-9. ['This book is intended to give a fairly comprehensive account of the theory of constructible sets at an advanced level'.]

22. Devlin,K.J., 'Aspects of Constructibility', (Springer, LNM, **354**, 1973) 240pp., ISBN 3-540-06522-9. [Essentially a first draft for 'Constructibility' above.]

23. Andras Hajnal and Peter Hamburger, 'Set Theory', (CUP, LMS Student Texts, **48**, 1999), 316pp, ISBN 0-521-59667-X.