b1 SET THEORY

Under normal circumstances, these notes would be my script for in-person lectures, which would be fleshed out by ad libs, questions and comments from the audience, extra diagrams, digressions and so forth. This year none of that is possible. I'm trying to supply some of that extra material in the videos. The biggest thing missing is you, the audience. As a very distant second-best, you can email me comments, questions, requests for clarification etc., and I will try to respond as quickly as I can (but I'm afraid that sometimes, that may not be very).

The videos will not be a read-through of the notes. I intend in the videos to concentrate on the parts that I think are most difficult, passing over some of the easier stuff. The tone of the videos will also be less formal than that of the notes, and they will concentrate on trying to get over the intuitions and some of the motivation.

It looks as though they are going to end up taking less time than the lectures would have done, possibly partly because I would have said absolutely everything at least once and probably as much as three times (before writing it out, while writing it, and afterwards), and would in addition have talked while waiting for people to finish writing up whatever they were writing.

I hope you find these notes, and the videos, useful. Please, as I say, get in touch if you have any questions.

I may continue to edit this document throughout the term.

Latest edit: 12th February 2021.

The material contained in the questions on the problem sheets, and their solutions, is on the syllabus.

Prerequisites: None.

0. What is Set Theory, and why study it?

Set theory is the study of sets, and of the relation $\in$ which connects the elements of a set to the set itself. It has two main purposes.

1. The purpose for which set theory was originally developed by Cantor was to provide **tools** for addressing mathematical problems. The most important of these are:

Cardinal numbers. These answer the question: how many? For example, *How many wives did Henry VIII have?* Answer: *six.* Or, *how many natural numbers are there?* Answer: $\aleph_0$. Infinite cardinal numbers, and their arithmetic, were revolutionary.

Ordinal numbers. These answer questions about position in an arrangement. For example, Anne Boleyn was the *second* wife of Henry *VIII.* Infinite ordinal numbers allow one to generalise mathematical induction and recursion.

2. Set theory was subsequently developed as a **foundation for mathematics.** There are reasons why the optimism of the early days of set theory has abated—we will come to some of these—but I propose two ways in which set theory is a foundation for mathematics.

Ontological: We can use the theory of sets to construct, for example, a complete ordered field. Complete ordered fields are very complicated while sets are relatively simple, so this is a good thing. Indeed, in the theory of sets we can construct mathematical structures of all kinds.

In this lecture course, we will construct the natural numbers; this then permits the construction of $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$.

Epistemological: We can reduce questions in mathematics to questions in set theory. For example, *is it the case that a product of compact spaces is compact?* Answer: *Yes, if the Axiom of Choice is true.* Or, *can Lebesgue measure be extended to a countably additive measure on all subsets of the real line?* Answer: *Yes, if there is a real-valued measurable cardinal.*†

In this lecture course, we will state axioms of set theory which are powerful enough to do much of modern mathematics, and which are at the same time relatively simple. **But** we will not really address the question of what, for example, the natural numbers really are. *As mathematicians,* we don't really care; we only care about their structure, and about what theorems we can prove about them.

We will be studying pure sets, that, is things like this:

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

Intuitively, pairs of curly brackets with nothing at the "bottom level". The reason is that these are enough to provide the technology and the foundation that we want, and they are relatively easy to deal with. If we allow our sets to contain, for example, ducks, Mount Everest or the Moon Landings, our formalism becomes much more complicated with no real mathematical gain.

Some authors have been impressed by the pure set program, believing that it licenses a philosophy in which everything (including people, galaxies etc) is built up out of nothing. This is misguided, because (a) people are clearly not pure sets, and (b) a pair of curly brackets is not nothing.

1. Russell's Paradox

The first attempts at establishing set theory as a foundation for mathematics involved axioms like

Naive collection axiom *If $\phi(x)$ is any property of sets, then*

$$\{x : \phi(x)\}$$

exists.

One might object to this axiom on the grounds that "property of sets" is vague, but it can be made precise in a rigorous way.

Now one can certainly prove a lot of theorems from an axiom like this. Unfortunately it is too good to be true:

THEOREM 1.1. *(Russell's Paradox) The naive collection axiom is inconsistent.*

PROOF: Assume the naive collection axiom. Let

$$D = \{x : x \notin x\}.$$

† Lebesgue measure, and real-valued measurable cardinals, are not on the syllabus.

This exists, by the naive collection axiom.

But $D \in D$ iff $D \notin D$, $\dot{\times}$.

So the naive collection axiom is inconsistent. □

This is a devastating result, and was unexpected (though Cantor possibly knew about it).

When we set out to axiomatise set theory, we will have to be much more cautious. We will only say that sets exist if we have to.

2. The basic axioms, and some basic sets

The axioms of set theory are all expressible in the *language of set theory*, which is a language of first-order predicate calculus with equality, with a binary predicate symbol $\in$. This language is adequate for the purpose, but it is difficult to use. For the purposes of this course, we will assume that any statement in mathematical English can be expressed in the language of set theory. (See the first handout.)

2.1. The first few axioms

We begin with what we hope will prove to be harmless axioms about sets.

Axiom of extensionality *Two sets are equal if and only if they have the same elements.*

This axiom is really the definition of equality of sets. It is not obvious; it would lead in ordinary life to us saying that the set of mermaids is equal to the set of unicorns (both sets being empty), which is weird. In mathematics, however, it works sufficiently well.

DEFINITION 2.1.1. *The* empty set*, written $\varnothing$, is the set with no elements.*

PROPOSITION 2.1.2. *There is at most one empty set.*

PROOF: Suppose a and b are both empty. Then a and b have the same elements, for every element of a is (vacuously) an element of b, and every element of b is an element of a.

Hence by Extensionality, $a = b$. □

Empty set axiom *The empty set $\varnothing$ exists.*

It is in fact the case that this axiom can be proved from the other axioms; but we will not trouble to make our axioms independent of each other.

We need some more axioms to construct some more interesting sets.

Axiom of Pairs *If a and b are sets, then so is $\{a, b\}$.*

PROPOSITION 2.1.3. *If a is a set, then so is $\{a\}$.*

PROOF: Apply the Axiom of Pairs to a and a; then $\{a, a\}$ is a set, and this is equal to $\{a\}$ by Extensionality. □

It is vitally important to keep track of "the number of curly brackets". For instance, $\varnothing$, $\{\varnothing\}$, $\{\{\varnothing\}\}$,... are all different from each other.

PROPOSITION 2.1.4. *$\varnothing$ and $\{\varnothing\}$ are not equal.*

PROOF: $\varnothing \in \{\varnothing\}$, so $\{\varnothing\}$ is not empty. □

We can now construct infinitely many sets.

DEFINITION 2.1.5. *Suppose A is a set. Then $\bigcup A$ is defined to be*

$$\{x : \exists a \in A \, x \in a\}.$$

Suppose a and b are sets. Then we define $a \cup b$ to be $\bigcup\{a, b\}$. Suppose $A = \{a_i : i \in I\}$ is a set of sets. Then we define $\bigcup_{i \in I} a_i$ to be $\bigcup A$.

Intuitively, an application of the pairing operation adds curly brackets and the union operation strips one away.

Axiom of Unions *Suppose A is a set. Then so is the union $\bigcup A$ of its elements.*

PROPOSITION 2.1.6. *If a and b are sets, then $a \cup b$ is a set.*

PROOF: $\{a, b\}$ exists by the Axiom of Pairs. $a \cup b = \bigcup\{a, b\}$ then exists by the Axiom of Unions. □

Now for a cautious version of the naive collection axiom.

Subset axiom scheme *Suppose A is a set and $\phi(x)$ is a statement in the language of set theory. Then*

$$\{x \in A : \phi(x)\}$$

is a set.

Note that we are allowing the statement $\phi(x)$ to mention sets other than x.

The Subset axiom scheme is actually an infinite set of axioms, one for each $\phi(x)$. It is also known as the Separation Scheme and the Comprehension Scheme.

The subset axiom scheme allows us to do a number of useful operations.

PROPOSITION 2.1.7. *If a and b are sets, then so is their difference $a \setminus b$.*

PROOF: $a \setminus b = \{x \in a : x \notin b\}$, which exists by the Subset Axiom Scheme. □

PROPOSITION 2.1.8. *Let a be a non-empty set. Then $\bigcap a$, the intersection of all the elements of a, is a set.*

PROOF: $\bigcap a = \{x \in \bigcup a : \forall y \in a \, x \in y\}$, which exists by the Union Axiom and the Subset Axiom Scheme. □

2.2. Classes

A very strange and striking consequence of the Subset Axiom Scheme:

THEOREM 2.2.1. *There is no set of all sets.*

PROOF: Let V be a set containing all sets.

Let $D = \{x \in V : x \notin x\}$. This is a set, by the Subset Axiom Scheme.

But since V contains all sets as elements, $D \in V$.

Hence $D \in D$ iff $D \notin D$, $\dot{\times}$. □

This is a nuisance. Indeed, there are many useful sets which provably can't exist. We get around this by introducing the notion of a *class*. A class is anything of the form

$$\{x : \phi(x)\}$$

where $\phi(x)$ is a property of sets. So, for instance, $\{x : x = x\}$ is the class of all sets, $\{x : x$ has only one element$\}$ is the class of one-element sets, and so on. A set is a class which is an element of some class (see the Axiom of Pairs). A class which is not a member of another class is a *proper class*. So we may paraphrase the above theorem as: the class of all sets is a proper class.

We have not given a formal definition of a proper class, since (at least in the way we are presenting set theory) they are an abuse of terminology, rather than a precisely definable concept.

2.3. The Axiom of Foundation

Finally in this chapter, a rather odd axiom which forbids certain sets from existing.

Foundation axiom *Suppose A is a non-empty set. Then A has an $\in$-minimal element; that is, there exists $m \in A$ such that $m \cap A = \varnothing$.*

Set theorists use the Foundation axiom because it makes the structure of the universe of sets tidier, and is really part of the definition of what we mean by a "set". The best way for the moment to indicate how, is to prove some theorems.

PROPOSITION 2.3.1. *Let a be any set. Then $a \notin a$.*

PROOF: Suppose $a \in a$. Let $A = \{a\}$. If m is any element of A, then $m = a$, but then $a \in m \cap A$.

This contradicts the Axiom of Foundation. □

Sets like this arguably "ought not to exist".

COROLLARY 2.3.2. *For all sets a, $a \neq \{a\}$.*

So, NEVER DROP CURLY BRACKETS. Do not write, for example, 1 when you mean $\{1\}$; also, do not write $\subset$ when you mean $\in$ or vice versa.

PROPOSITION 2.3.3. *Let a be any set. Then there does not exist a set $\{a_0, a_1, a_2, \ldots\}$ such that $a = a_0 \ni a_1 \ni a_2 \ni \cdots$*

In the presence of all the other axioms of Set Theory, including the Axiom of Choice, this Proposition is equivalent to the Axiom of Foundation. One of the main purposes of the Axiom of Foundation is to prevent infinite descending chains from existing.

2.4. Ordered pairs

We define our first non-trivial piece of mathematical structure.

DEFINITION 2.4.1. *Suppose a and b are sets. We define the* ordered pair $\langle a, b\rangle$ *to be*

$$\{\{a\}, \{a, b\}\}.$$

Note carefully what we are doing here. We are *not* saying that the ordered pair $\langle a, b\rangle$ is "really" the rather strange set $\{\{a\}, \{a, b\}\}$. We are only claiming that it is *convenient* to define it in this way, because, firstly, $\{\{a\}, \{a, b\}\}$ exists as a set, and secondly, because it works. We explain what we mean by this.

Firstly, existence:

PROPOSITION 2.4.2. *If a and b are sets, then so is $\langle a, b\rangle$.*

PROOF: By the Axiom of Pairs, $\{a, b\}$ is a set. Applying the Axiom of Pairs to a and a, $\{a, a\} = \{a\}$ is a set. Now applying the Axiom of Pairs again,

$$\langle a, b\rangle = \{\{a\}, \{a, b\}\}$$

is a set. □

Second, proper functioning:

PROPOSITION 2.4.3. *$\langle a, b\rangle = \langle c, d\rangle$ if and only if $a = c$ and $b = d$.*

PROOF: $\Leftarrow$) Trivial.

$\Rightarrow$) Suppose $\langle a, b\rangle = \langle c, d\rangle$. Then

$$\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}.$$

Applying the Axiom of Unions, take the union of both sides:

$$\bigcup\{\{a\}, \{a, b\}\} = \bigcup\{\{c\}, \{c, d\}\},$$

that is,

$$\{a, b\} = \{c, d\}.$$

Now if $a = b$, then since $\{c, d\} = \{a, b\} = \{a\}$, $c = a$ and $d = a$, so $a = c$ and $b = d$ as required. Similarly if $c = d$, $a = c$ and $b = d$ follow.

If $a \neq b$ and $c \neq d$, then $\{a, b\} \neq \{a\}$ and $\{c, d\} \neq \{c\}$, so

$$\{\{a\}, \{a, b\}\} \setminus \{\{a, b\}\} = \{\{a\}\},$$

and

$$\{\{c\}, \{c, d\}\} \setminus \{\{c, d\}\} = \{\{c\}\}.$$

($\{\{a, b\}\}$ and $\{\{c, d\}\}$ exist by Proposition 2.1.3, and $\{\{a\}, \{a, b\}\}\setminus\{\{a, b\}\}$ and $\{\{c\}, \{c, d\}\}\setminus\{\{c, d\}\}$ exist by Proposition 2.1.8.)

Since

$$\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$$

and

$$\{a, b\} = \{c, d\},$$

we have that

$$\{\{a\}\} = \{\{c\}\}$$

and hence $\{a\} = \{c\}$ and then $a = c$; and

$$\{b\} = \{a, b\} \setminus \{a\} = \{c, d\} \setminus \{c\} = \{d\},$$

so

$$b = d.$$

□

3. Powersets and products

We now introduce the first of a series of very powerful axioms.

3.1. The powerset axiom

DEFINITION 3.1.1. *Let X be a set. The* powerset *of X, written $\wp X$, is the class of all subsets of X.*

Powerset axiom *Let X be a set. Then $\wp X$ is a set.*

One reason that the powerset axiom is so powerful is that it gets us sets of lots of different sizes. Thus:

THEOREM 3.1.2. *Let X be any set. Then there is an injection from X to $\wp X$.*

Strictly speaking, we shouldn't prove this theorem until we have defined what functions are, and said what we are and are not allowed to do with them. But let us suppose for the moment that we have.

PROOF: Define $f : X \to \wp X$ as follows:

$$f : a \mapsto \{a\}.$$

This is clearly well-defined and one-to-one. □

However,

THEOREM 3.1.3. *(Cantor's Theorem) Let X be any set. Then there is no surjection from X to $\wp X$.*

PROOF: Suppose g is a surjection from X to $\wp X$. Let

$$D = \{a \in x : a \notin g(a)\}.$$

Then, by the Subset Axiom Scheme, D is a set.

Since g is onto, $D = g(a)$ for some a. But then $a \in D$ iff $a \notin D$, $\dot{\times}$. □

Notice the similarity between this and Russell's Paradox.

COROLLARY 3.1.4. *There is no injection from $\wp X$ to X.*

PROOF: Suppose $f : \wp X \to X$ is one-to-one. Then

$$g : a \mapsto \begin{cases} b & \text{if } a = f(b), \\ \varnothing & \text{if } a \notin \operatorname{ran} f \end{cases}$$

is a surjection from X to $\wp X$. □

When we have defined cardinal numbers, we will express this by saying "$\wp X$ has more elements than X".

3.2. Cartesian products

Now another piece of formal construction.

DEFINITION 3.2.1. *Suppose A and B are sets. Their* Cartesian product *is*

$$A \times B = \{\langle a, b\rangle : a \in A,\, b \in B.\}$$

☡ Notation: $A \times B$ is a set of ordered pairs $\langle a, b\rangle$. Do NOT write $a \times b$ for a typical element $\langle a, b\rangle$ of $A \times B$; this is WRONG.

PROPOSITION 3.2.2. *If A and B are sets, then $A \times B$ is a set.*

PROOF: Applying the Axioms of Pairs, Unions and the Powerset,

$$\wp\wp\bigcup\{A, B\} = \wp\wp(A \cup B)$$

is a set.

Applying the Subset Axiom Scheme,

$$\{x \in \wp\wp(A \cup B) : \exists a \in A, b \in B\, x = \langle a, b\rangle\}$$

is a set.

But this set is equal to $A \times B$. The only non-trivial thing we need to prove is that any element of $A \times B$ is in $\wp\wp(A \times B)$.

But if $x \in A \times B$, then for some $a \in A$ and $b \in B$, $x = \langle a, b\rangle = \{\{a\}, \{a, b\}\}$. Both $\{a\}$ and $\{a, b\}$ are subsets of $A \cup B$ and so are elements of $\wp(A \cup B)$; so $\{\{a\}, \{a, b\}\}$ is a subset of $\wp(A \cup B)$ and so an element of $\wp\wp(A \cup B)$. □

3.3. Relations and functions

Recall from the first year:

DEFINITION 3.3.1. *A* relation *between sets A and B is a subset of $A \times B$.*

A relation f between A and B is a function *iff for all $a \in A$, there exists unique b in B such that $\langle a, b\rangle \in f$. We write $b = f(a)$.*

DEFINITION 3.3.2. *Suppose A and B are sets. Then we write ${}^AB$ for the set of all functions from A to B.*

PROPOSITION 3.3.3. *Suppose A and B are sets. Then ${}^AB$ is a set.*

PROOF: On the problem sheets. □

Finally, some miscellaneous, but useful, function-related definitions.

DEFINITION 3.3.4. *Suppose that f is a function from A to B. Then A is the domain of f.*

DEFINITION 3.3.5. *Suppose A and B are sets. Then a* partial function *from A to B is a function f having the property that for some subset C of A, f is a function from C to B.*

We are accustomed to defining inverses of (bijective) functions. This notion can be generalised.

Definition 3.3.6. *Suppose R is a relation between A and B. Then R^{-1} is the relation between B and A defined so that $\langle b, a\rangle \in R^{-1}$ iff $\langle a, b\rangle \in R$.*

We are also accustomed to defining the image of a function under a set. In the context of set theory, we have to be careful about this. Again we do it in a bit more generality.

Definition 3.3.7. *Suppose R is a relation between A and B. We define a function R" from $\wp A$ to $\wp B$ such that for all $X \in \wp A$,*

$$R"(X) = \{b \in B : \exists a \in X\ \langle a, b\rangle \in R\}.$$

We write $R[X]$ (with square brackets) for $R"(X)$.

If f is a function from A to B, if $X \in \wp A$, then we refer to $f[X]$ as the image *of X under f, and if $Y \in \wp B$, then we refer to $f^{-1}[Y]$ as the* preimage *of Y under f. If $b \in B$, then $f^{-1}[\{b\}]$ is often known as the* fibre *of b under f.*

3.4. Order relations

One particular kind of relation which we will have cause to use often is the *order relation*. The basic definition is the following:

Definition 3.4.1. *A relation R is a partial order on a set A iff R is a relation on A, and it is*

1. *Reflexive: for all $a \in A$, $a\ R\ a$.*
2. *Antisymmetric: for all $a, b \in A$, if $a\ R\ b$ and $b\ R\ a$, then $a = b$.*
3. *Transitive: for all $a, b, c \in A$, if $a\ R\ b$ and $b\ R\ c$, then $a\ R\ c$.*

Examples 3.4.2. *$\subseteq$ on $\wp X$.*

$\leq$ on $\mathbb{R}$.

$|$ on $\mathbb{N}$.

Definition 3.4.3. *A partial order R on a set A is a total order iff for all $a, b \in A$, either $a\ R\ b$ or $b\ R\ a$.*

4. The natural numbers

Frege defined the natural numbers as follows (in modern notation): $0 = |\varnothing|$, $1 = |\{0\}|$, $2 = |\{0, 1\}|$, $3 = |\{0, 1, 2\}|$.

In modern set theory, we do something simpler but less natural, namely define $0 = \varnothing$, $1 = \{0\}$, $2 = \{0, 1\}$, and in general, $n = \{0, 1, 2, \ldots, n-1\}$.

If we do this, then

$$n + 1 = \{0, 1, 2, \ldots, n\} = \{0, 1, 2, \ldots, n-1\} \cup \{n\} = n \cup \{n\}.$$

So we begin our study of the natural numbers with the operation $x \mapsto x \cup \{x\}$.

4.1. The axiom of infinity

DEFINITION 4.1.1. *If x is a set, then define x^+ to be $x \cup \{x\}$.*

DEFINITION 4.1.2. *Define a* successor set *to be a set Ω such that $\varnothing \in \Omega$, and whenever $x \in \Omega$, $x^+ \in \Omega$ as well.*

The axioms we have so far do not suffice to prove that there is a successor set. So we have:

Axiom of Infinity *There is a successor set.*

DEFINITION 4.1.3. *Let Ω be a successor set. We define the set ω of natural numbers to be*

$$\{x \in \Omega : x \text{ is in every successor set.}\}$$

This definition is essentially due to Dedekind.

THEOREM 4.1.4. *ω exists, and is included in every successor set.*

PROOF: ω exists by the Subset Axiom Scheme. That it is included in every successor set is trivial, by the definition. □

THEOREM 4.1.5. *ω is a successor set.*

PROOF: $\varnothing \in \omega$, because $\varnothing$ belongs to every successor set, and belongs to Ω in particular.

If $x \in \omega$, then x belongs to every successor set Ω'. Since Ω' is a successor set, x^+ belongs to Ω' also. That is, x^+ belongs to every successor set.

Thus ω is a successor set, as required. □

THEOREM 4.1.6. *ω is the only successor set that is included in every successor set.*

PROOF: Obvious. For if ω and ω' both satisfy this, then $\omega \subseteq \omega'$, since ω' is a successor set and ω is included in every successor set, and likewise $\omega' \subseteq \omega$. So $\omega = \omega'$. □

The point is that our definition of ω appeared to depend on the choice of a particular successor set Ω. In fact it does not.

INFORMAL DEFINITION 4.1.7. *We define 0 to be $\varnothing$, 1 to be 0^+, 2 to be 0^{++}, etc.*

THEOREM 4.1.8. *(Proof by induction) Suppose $P(x)$ is a property of natural numbers such that $P(0)$ holds, and for all n, if $P(n)$ holds, then $P(n+1)$ holds.*

Then $P(n)$ holds for every n.

PROOF: Let $A = \{n \in \omega : P(n)\}$. Then A is a successor set. Since ω is included in every successor set, $\omega \subseteq A$. Obviously $A \subseteq \omega$. So $\omega = A$. □

DEFINITION 4.1.9. *If $m, n \in \omega$, then we say that $m \leq n$ iff $m \in n$ or $m = n$.*

PROPOSITION 4.1.10. *If $m, n, k \in \omega$ and $m \in n \in k$, then $m \in k$.*

PROOF: By induction on k.

If $k = 0$, the statement is vacuous.

Suppose it is true for k. Suppose $m \in n \in k^+$.

Since $k^+ = k \cup \{k\}$, there are two cases. The first is if $n \in k$. Then by the inductive hypothesis, $m \in k \subseteq k \cup \{k\} = k^+$.

The second is if $n \in \{k\}$. Then $n = k$. Hence $m \in k \subseteq k \cup \{k\} = k^+$.

Either way, $m \in k^+$ as required. □

PROPOSITION 4.1.11. *$\leq$ is a partial order on ω.*

PROOF: $\leq$ is reflexive by definition.

Antisymmetry: suppose $m \leq n$ and $n \leq m$, but $m \neq n$. Then $m < n$ and $n < m$; that is, $m \in n$ and $n \in m$. Then $m \in m$, $\dot{\times}$ to Proposition 2.3.1.

Transitivity: suppose $m \leq n$ and $n \leq k$. If $m = n$ or $n = k$, then trivially $m \leq k$. Otherwise, $m < n < k$, that is, $m \in n \in k$, so $m \in k$ by Proposition 4.1.10. Hence $m \leq k$ as required. □

PROPOSITION 4.1.12. *For all $m, n \in \omega$, if $m < n^+$ then $m \leq n$.*

PROOF: If $m \in n^+$ then $m \in n \cup \{n\}$. Thus either $m \in n$ or $m = n$, that is, $m \leq n$. □

PROPOSITION 4.1.13. *For all m and n in ω, $m \leq n$ iff $m \subseteq n$.*

PROOF: $\Rightarrow$) We use induction on n.

If $n = 0$, then $m \leq 0$ iff $m \in 0$ or $m = 0$. But $m \in 0$ is impossible, since 0 is empty. So $m \leq 0$ iff $m = 0$. But also, since 0 is empty, $m \subseteq 0$ iff $m = 0$. So $m \leq 0$ iff $m \subseteq 0$.

Now suppose that for all m, $m \leq n$ implies $m \subseteq n$.

Suppose $m \leq n^+$. Then either $m = n^+$, when certainly $m \subseteq n^+$, or $m < n^+$, so $m \leq n$, so by the inductive hypothesis, $m \subseteq n \subseteq n \cup \{n\} = n^+$.

$\Leftarrow$) Again we use induction on n.

Suppose $m \subseteq 0$. Then since 0 is empty, so is m, so $m = 0$, so $m \leq 0$.

Now suppose $m \subseteq n^+ = n \cup \{n\}$. If $n \notin m$, then $m \subseteq n$, so by the inductive hypothesis, $m \leq n$, so $m \leq n^+$ (since $n \leq n^+$, using Proposition 4.1.11). If $n \in m$, then $n \leq m$. Hence by the above argument $n \subseteq m$. Thus in fact $n^+ \subseteq m$, so $m = n^+$, so $m \leq n^+$. □

PROPOSITION 4.1.14. *$\leq$ is a total order on ω.*

PROOF: We prove by induction on n: For all m, if $m \not\leq n$, then $n \leq m$.

If $n = 0$, then n is empty, so trivially $n \subseteq m$ for all m, so $n \leq m$.

Now suppose the result true for n. Suppose $m \not\leq n^+$.

Then $m \not\leq n$ by Proposition 4.1.10, so by the inductive hypothesis, $n \leq m$. Hence $n \subseteq m$. Since $m \not\leq n$, $n \neq m$, so $n < m$, so $n \in m$. Hence $n^+ = n \cup \{n\} \subseteq m$. So $n^+ \leq m$, as required. □

PROPOSITION 4.1.15. *Every element of ω is either 0 or the successor of a unique natural number.*

PROOF: It is trivial to prove by induction on n the statement "either $n = 0$ or for some m, $n = m^+$".

As for uniqueness, if $n = m^+$, then $m = \bigcup n$ (question 2. on sheet 1, and transitivity of m), so the operator $n \mapsto n^+$ on ω is one-to-one. □

THEOREM 4.1.16. *(Proof by strong induction) Suppose $P(x)$ is a property of natural numbers such that for each $n \in \omega$, if $P(m)$ holds for every $m \in n$, then $P(n)$ holds.*

Then $P(n)$ holds for every $n \in \omega$.

PROOF: Let $Q(n)$ be the property "For every $m \in n$, $P(m)$ holds".

Then $Q(0)$ holds vacuously, since there are no $m \in 0$.

Now, if $Q(n)$ holds, then $P(m)$ holds for every $m \in n$, so by assumption, $P(n)$ holds, so $P(m)$ holds for every $m \in n \cup \{n\} = n^+$. So $Q(n^+)$ holds.

Hence by induction, $Q(n)$ holds for every n. In particular, for each n, $Q(n^+)$ holds. But $n \in n^+$, so $P(n)$ holds for every n as required. □

Whether it is better to use induction or strong induction to prove a result depends on circumstances.

THEOREM 4.1.17. *Every non-empty subset of ω has a least element.*

PROOF: Let S be a subset of ω with no least element. Define $P(n)$ to be the statement "$n \notin S$".

Suppose that $P(m)$ holds for all $m < n$, but $P(n)$ does not hold. Then n is the least element of S, $\dot{\times}$. Thus $P(n)$ must in fact hold.

Now, by strong induction, $P(n)$ holds for all n; that is, S is empty. □

THEOREM 4.1.18. *(Definition by recursion) Suppose A is a set, $a \in A$, and $g : A \to A$. Then there exists a unique function $f : \omega \to A$ such that $f(0) = a$ and $f(n^+) = g(f(n))$ for all n.*

PROOF: Let us say a function ϕ is *nice* if, for some n, $\phi : n \to A$, and, if $n \neq \varnothing$, then $\phi(0) = a$, and if $k^+ \in n$, then $\phi(k^+) = g(\phi(k))$.

Then any two nice functions agree on their common domain; that is, if ϕ and ψ are nice, and if $k \in \operatorname{dom}\phi \cap \operatorname{dom}\psi$, then $\phi(k) = \psi(k)$. We prove this by induction on k; the base case, $k = 0$, is trivial since $\phi(0) = \psi(0) = a$, and as for the inductive step, if $\phi(k) = \psi(k)$, then $\phi(k^+) = g(\phi(k)) = g(\psi(k)) = \psi(k^+)$.

Now we show that for every $n \in \omega$, there exists a nice function with domain n. This is trivial if $n = 0$, since $\varnothing$ is a function from $\varnothing$ to A, and it vacuously is nice.

Now if there is a nice function ϕ with domain n, then the following is a nice function with domain $n + 1$:

$$\phi \cup \{\langle 0, a\rangle\},$$

if $n = 0$, and if $n = k^+$, then

$$\phi \cup \{\langle n, g(\phi(k))\rangle\}.$$

Now let B be the set of all nice functions; it is a set since we can express

$$B = \{\phi \in \wp(\omega \times A) : \phi \text{ is a nice function}\},$$

and we can apply the Subset Axiom Scheme.

Now let $f = \bigcup B$.

Certainly $f \subseteq \omega \times A$, since all elements of A are subsets of $\omega \times A$, being partial functions from ω to A.

If $\langle k, b\rangle$ and $\langle k, c\rangle$ are elements of f, then there exist nice functions ϕ and ψ such that $\langle k, b\rangle \in \phi$ and $\langle k, c\rangle \in \psi$. But we have just proved that in that case, $b = c$. So f is a function.

Now f has domain ω, because for every n, there is a nice function ϕ whose domain is $n^+ = n \cup \{n\}$. Then n is in the domain of ϕ; and then $\langle n, \phi(n)\rangle \in f$, so n is in the domain of f.

Next we show that f has the required properties. $f(0) = a$, since there is a nice function ϕ with 0 in its domain; then $\phi(0) = a$. Hence $\langle 0, a\rangle \in \phi$, so $\langle 0, a\rangle \in f$, so $f(0) = a$.

Now for each n, $\langle n, f(n)\rangle \in f$, so $\langle n, f(n)\rangle \in \phi$ for some nice function ϕ. Hence $\phi(n) = f(n)$. Now suppose ψ is a nice function with domain n^{++}. Then $\psi(n) = \phi(n) = f(n)$ since nice functions agree, and then $\psi(n^+) = g(\psi(n)) = g(f(n))$ by niceness. So $\langle n^+, g(f(n))\rangle \in \psi$, so $\langle n^+, g(f(n))\rangle \in f$, so $f(n^+) = g(f(n))$.

So we have established that a function f with the required properties exists. Now we show that it is unique, by showing by induction on n that if f' is another such function, then $f(n) = f'(n)$.

Clearly $f(0) = f'(0)$ because both are equal to a.

Suppose $f(n) = f'(n)$. Then $f(n^+) = g(f(n)) = g(f'(n)) = f'(n^+)$.

Hence $f = f'$, and we have established uniqueness. □

COROLLARY 4.1.19. *Suppose A and B are sets, $h : B \to A$ is a function, and $g : A \to A$ is a function. Then there is a unique function $f : B \times \omega \to A$ such that for all $b \in B$, $f(b, 0) = h(b)$, and for all $b \in B$ and $n \in \omega$, $f(b, n^+) = g(f(b, n))$.*

PROOF: For each $b \in B$, let f_b be the unique function from ω to A such that $f_b(0) = h(b)$, and $f_b(n^+) = g(f_b(n))$ for all $n \in \omega$.

Now define $f : B \times \omega \to A$ by $f(b, n) = f_b(n)$. f exists since

$$f = \{x \in (B \times \omega) \times A : \exists b \in B\, \exists n \in \omega\, x = \langle\langle b, n\rangle, f_b(n)\rangle\}.$$

Clearly f has the right properties. □

4.2. Arithmetic on the natural numbers

DEFINITION 4.2.1. *We use recursion to define addition on the natural numbers as follows:*

1. *For all $n \in \omega$, $n + 0 = n$;*
2. *For all $m, n \in \omega$, $m + n^+ = (m + n)^+$.*

THEOREM 4.2.2. *Addition is associative.*

PROOF: We prove by induction on k that for all m and n, $m + (n + k) = (m + n) + k$.

$m + (n + 0) = m + n = (m + n) + 0$.

$m + (n + k^+) = m + (n + k)^+ = (m + (n + k))^+ = ((m + n) + k)^+$ by the inductive hypothesis, and this is equal to $(m + n) + k^+$. □

LEMMA 4.2.3. *For all m, $n \in \omega$, $m + n^+ = m^+ + n$.*

PROOF: We do induction on n.

$m + 0^+ = (m + 0)^+ = m^+ = m^+ + 0$.

$m + n^{++} = (m + n^+)^+ = (m^+ + n)^+$ by the inductive hypothesis, and this is equal to $m^+ + n^+$.

LEMMA 4.2.4. *For all m, $m + 0 = 0 + m$.*

PROOF: Induction on m. The case $m = 0$ is trivial.

$0 + m^+ = (0 + m)^+ = (m + 0)^+$ by the inductive hypothesis, and this is $m^+ = m^+ + 0$. □

THEOREM 4.2.5. *Addition is commutative.*

PROOF: We prove by induction on n that for all m, $m + n = n + m$.

We have already done the case $n = 0$.

Now $m + n^+ = (m + n)^+ = (n + m)^+$ by the inductive hypothesis, and $(n + m)^+ = n + m^+ = n^+ + m$. □

DEFINITION 4.2.6. *We define multiplication by recursion on ω as follows.*

1. $m.0 = 0$ *for all* m,
2. $m.n^+ = m.n + m$ *for all* m *and* n.

THEOREM 4.2.7. *For all $m, n, k \in \omega$, $m.(n + k) = m.n + m.k$.*

PROOF: By induction on k.

$m.(n + 0) = m.n = m.n + 0 = m.n + m.0$.

$m.(n + k^+) = m.(n + k)^+ = m.(n + k) + m = (m.n + m.k) + m = m.n + (m.k + m) = m.n + m.k^+$. □

THEOREM 4.2.8. *Multiplication is associative.*

PROOF: We prove by induction on k that for all m and n, $m.(n.k) = (m.n).k$.

$m.(n.0) = m.0 = 0 = (m.n).0$.

$m.(n.k^+) = m.(n.k + n) = m.(n.k) + m.n = (m.n).k + m.n = (m.n).k^+$. □

THEOREM 4.2.9. *Multiplication is commutative.*

PROOF: Problem sheets. □

On this basis, it is now possible to construct the ring $\mathbb{Z}$ and the fields $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$, and the various other structures of classical mathematics.

DEFINITION 4.2.10. *We define exponentiation by recursion on ω as follows.*

1. $m^0 = 1$ *for all* $m \neq 0$,
2. $m.n^+ = m^n.m$ *for all* $m \neq 0$ *and* n.

(We steer clear of 0^0 for the usual reasons.)

THEOREM 4.2.11. *For all $m \neq 0$ and for all n and p, $m^{n+p} = m^n.m^p$.*

PROOF: By induction on p.

Base case: $m^{n+0} = m^n = m^n.1 = m^n.m^0$.

Inductive step: $m^{n+p^+} = m^{n+p^+} = m^{n+p}.m = (m^n.m^p).m = m^n.(m^p.m) = m^n.m^{p^+}$. □

THEOREM 4.2.12. *For all $m \neq 0$ and for all n and p, $m^{n.p} = (m^n)^p$.*

PROOF: By induction on p.

Base case: $m^{n.0} = m^0 = 1 = (m^n)^0$.

Inductive step: $m^{n.p^+} = m^{(n.p)+n} = m^{n.p}.m^n = (m^n)^p.m^n = (m^n)^{p^+}$. □

4.3. Peano Arithmetic.

We have now proved that something is consistent. But what exactly? We have been attempting to justify the assertion that the natural numbers, with their arithmetic, can sensibly be said to exist. We are in fact interested mainly in the properties of the natural numbers, rather than in what they really are. The properties might include some such list as the following:

DEFINITION 4.3.1. *The axioms of* Peano Arithmetic *are the following list of axioms about a structure containing an element* 0*, a function* $x \mapsto x^+$*, and binary functions* $+$ *and* $.$*:*

1. *The function* $x \mapsto x^+$ *is one-to-one and* y *belongs to its range iff* $y \neq 0$.
2. *For all* x, $x + 0 = x$.
3. *For all* x, y, $x + y^+ = (x + y)^+$.
4. *For all* x, $x.0 = 0$.
5. *For all* x, y, $x.y^+ = x.y + x$.
6. *Let* A *be any set such that* $0 \in A$ *and* $\forall x\,(x \in A \to x^+ \in A)$. *Then every element belongs to* A.

THEOREM 4.3.2. *ω, with the operations of $+$ and $.$ already defined, satisfies the axioms of Peano Arithmetic.*

PROOF: Now obvious. □

THEOREM 4.3.3. *Suppose $\mathfrak{M}$ and $\mathfrak{N}$ satisfy the axioms of Peano Arithmetic.. Then $\mathfrak{M}$ and $\mathfrak{N}$ are isomorphic.*

PROOF: We suppose $\mathfrak{M}$ is a model of the Peano Axioms, and show that $\mathfrak{M}$ is isomorphic to ω.

Define a function $f : \omega \to \mathfrak{M}$ by recursion as follows:

$f(0) = 0_{\mathfrak{M}}$, $f(n^+) = f(n)^+$. Then by a result on the problem sheets, f is a bijection.

We now prove by induction that for all m and n, $f(m + n) = f(m) + f(n)$ and $f(m.n) = f(m).f(n)$. □

5. Cardinal arithmetic

In mathematics—and in language—a cardinal number is a quantity used to measure the number of elements in a set. Thus the cardinal number of the set $\{-1, 1\}$ is two, and so on.

One of Cantor's strokes of genius was to define infinite cardinal numbers. We begin by defining the notion that two sets have the same number of elements, and we do this in the intuitive way: by saying two sets have the same number of elements iff there is a bijection between them

5.1. Equicardinality

DEFINITION 5.1.1. *Suppose X and Y are sets. We say X and Y* have the same cardinality*, and write $X \sim Y$ or $|X| = |Y|$ iff there is a bijection between them.*

We say X has cardinality less than or equal to that of Y*, and write $X \preceq Y$ or $|X| \leq |Y|$, iff there is an injection from X to Y.*

If $X \preceq Y$ and $X \not\sim Y$, we write $X \prec Y$ and $|X| < |Y|$.

THEOREM 5.1.2. *$\preceq$ is transitive.*

PROOF: Trivial. For if $X \preceq Y \preceq Z$, suppose $f : X \to Y$ and $g : Y \to Z$ are one-to-one. Then so is $g \circ f : X \to Z$. □

THEOREM 5.1.3. *$\preceq$ is reflexive.*

PROOF: The identity is an injection from X to itself. □

We would like some result approximating a proof that X is antisymmetric; more realistically, that if $X \preceq Y$ and $Y \preceq X$, then $X \sim Y$, or in the other notation, if $|X| \leq |Y|$ and $|Y| \leq |X|$, then $|X| = |Y|$. This is true, but is surprisingly difficult to prove.

We begin with a lemma.

LEMMA 5.1.4. *(Tarski's Fixed Point Theorem) Suppose X is a set, and $F : \wp X \to \wp X$ is a function having the property that if $A \subseteq B$, then $F(A) \subseteq F(B)$.*

Then there exists $Y \in \wp X$ such that $F(Y) = Y$.

PROOF: Define Y as follows:

$$Y = \bigcup \{A \in \wp X : A \subseteq F(A)\}.$$

We show first that $Y \subseteq F(Y)$. For, if $y \in Y$, then there exists A such that $y \in A$ and $A \subseteq F(A)$. Since $A \subseteq Y$ by definition of Y, $F(A) \subseteq F(Y)$. So $y \in A \subseteq F(A) \subseteq F(Y)$. Thus $Y \subseteq F(Y)$ as required.

Now we show that $F(Y) \subseteq Y$. For, since $Y \subseteq F(Y)$, also $F(Y) \subseteq F(F(Y))$; thus $F(Y) \subseteq Y$ by definition of Y.

Hence $Y = F(Y)$, as required. □

THEOREM 5.1.5. *(The Schröder-Bernstein Theorem) If X and Y are sets, $X \preceq Y$ and $Y \preceq X$, then $X \sim Y$.*

PROOF: Let $f : X \to Y$ and $g : Y \to X$ be one-to-one. Define $f" : \wp X \to \wp Y$ and $g" : \wp Y \to \wp X$ in the obvious way.

Define $F : \wp X \to \wp X$ by

$$F(A) = g"(Y \setminus f"(X \setminus A)).$$

We show that if $A \subseteq B$, then $F(A) \subseteq F(B)$.

For, if $A \subseteq B$, then

$$X \setminus B \subseteq X \setminus A,$$

so

$$f"(X \setminus B) \subseteq f"(X \setminus A),$$

so

$$Y \setminus f"(X \setminus A) \subseteq Y \setminus f"(X \setminus B),$$

so

$$g"(Y \setminus f"(X \setminus A)) \subseteq g"(Y \setminus f"(X \setminus B)),$$

as required.

Now by the Tarski Fixed Point Theorem, there exists $Z \in \wp X$ such that $F(Z) = Z$.

Now define

$$h : x \mapsto \begin{cases} f(x) & \text{if } x \notin Z, \\ y & \text{if } x \in Z \text{ and } g(y) = x. \end{cases}$$

We show that h is a bijection from X to Y.

Firstly, h is well-defined, because if $x \in Z$, then $x \in \operatorname{ran} g$, and because g is one-to-one.

h is one-to-one: suppose $h(x) = h(x')$.

If $x, x' \notin Z$, then $h(x) = f(x)$ and $h(x') = f(x')$. Thus $f(x) = f(x')$. But f is one-to-one, so $x = x'$.

If $x, x' \in Z$, suppose $x = g(y)$ and $x' = g(y')$. Then $y = y'$, so $x = x'$.

Now suppose $x \in Z$ and $x' \notin Z$. Then $x = g(h(x))$, and since $x \in Z$, g is one-to-one, and $Z = F(Z)$, $h(x) \in Y \setminus f"(X \setminus Z)$. Now $h(x) = h(x')$, and since $x' \notin Z$, $h(x') = f(x')$. But then $h(x) = h(x') = f(x') \in f"(X \setminus Z)$, $\dot{\times}$.

h is onto: if $y \in Y$, then there are two cases.

The first is when $y \in Y \setminus f"(X \setminus Z)$. Then $g(y) \in g"(Y \setminus f"(X \setminus Z))$, so $h(g(y)) = y$.

The second is when $y \in f"(X \setminus Z)$. Then there exists $x \in X \setminus Z$ such that $y = f(x)$. But then $y = h(x)$ also. □

But is it the case that if X and Y are sets, then either $|X| \leq |Y|$ or $|Y| \leq |X|$? Common sense might say yes... but the axioms we have so far are not sufficient to prove it. We will return to this later.

5.2. Finiteness

DEFINITION 5.2.1. *Suppose $n \in \omega$, and X is a set. We say X* has cardinality n, *and write $|X| = n$, iff $X \sim n$.*

We say X is finite *iff X has cardinality n, for some $n \in \omega$. Otherwise we say X is* infinite.

We need to know that the above notions are well-defined.

First a lemma:

THEOREM 5.2.2. *Suppose X is a finite set. Then if $f : X \to X$ is one-to-one, then it is onto.*

PROOF: It is sufficient to prove this for $X = n$ for some $n \in \omega$, for if $f : X \to X$ is one-to-one and not onto, and $g : n \to X$ is a bijection, then $g^{-1} \circ f \circ g : n \to n$ is a one-to-one function that is not onto.

We prove this by induction on n.

For $n = 0$, every function from $\varnothing$ to $\varnothing$ is both one-to-one and onto, so there is nothing to prove.

Now assume the result for n, and suppose $f : n^+ \to n^+$ is one-to-one but not onto. Suppose $k \notin \operatorname{ran} f$.

If $n = k$, let $g = f$. Otherwise,

$$g : i \mapsto \begin{cases} f(i) & \text{if } f(i) \neq n, \\ k & \text{if } f(i) = n. \end{cases}$$

Either way, g is a one-to-one function from n^+ to n.

Then $g{\restriction}n$ is a one-to-one function from n to $n \setminus \{g(n)\}$; that is, it is a one-to-one function from n to n that is not onto, $\dot{\times}$. □

THEOREM 5.2.3. *(The Pigeonhole Principle) If $m, n \in \omega$, and $m \neq n$, then $m \not\sim n$.*

PROOF: Suppose $m < n$, and $f : n \to m$ is a bijection.

Then f is a one-to-one function from n to n that is not onto, $\dot{\times}$. □

THEOREM 5.2.4. *If A and B are sets and $A \subseteq B$, then $A \preceq B$.*

PROOF: Trivial. □

COROLLARY 5.2.5. *If $m, n \in \omega$, then $m \leq n$ iff $m \preceq n$.*

5.3. Countability

DEFINITION 5.3.1. *We say that a set X is* countably infinite, *and write $|X| = \aleph_0$, iff $X \sim \omega$.*

We say X is countable *iff X is finite or countably infinite.*

$\aleph_0$ is our first example of an infinite cardinal number. Obviously, $|\omega| = \aleph_0$.

THEOREM 5.3.2. *Every subset of ω is countable.*

PROOF: Suppose $A \subseteq \omega$, and that A is not finite.

Define $h : \omega \to A$ by recursion as follows: $h(0) = \min A$, and $h(n^+) = \min\{x \in A : x > h(n)\}$ for each n, if this set is non-empty.

Now h is onto. For, let $B = A \setminus \operatorname{ran} f$. If $a \in B$, then for all natural numbers n, $h(n) < a$. We prove this by induction on n. For $h(0) = \min A < a$; and if $h(n) < a$, then $h(n^+) < a$ also.

Also, by induction, for all n, $n \leq h(n)$. Clearly $0 \leq h(0)$. If $n \leq h(n)$, then, because $h(n) < h(n^+)$, $n^+ \leq h(n^+)$ also.

So for all n, $n < a$; and this is a contradiction.

Hence h is onto. Clearly h is one-to-one, so h is a bijection. □

COROLLARY 5.3.3. *A set X is countable iff $X \preceq \omega$.*

PROOF: $\Rightarrow$): Obvious.

$\Leftarrow$): suppose $f : X \to \omega$ is one-to-one. Then f is a bijection between X and $\operatorname{ran} f \subseteq \omega$. By the previous theorem, $\operatorname{ran} f$ is countable, so we are done. □

THEOREM 5.3.4. *A non-empty set X is countable iff there is an onto function from ω to X.*

PROOF: $\Rightarrow$): Suppose X is countable. If X is countably infinite, then there is a bijection from X to ω, which is onto. If $|X| = n$, where $n > 0$, and $f : n \leftrightarrow X$, define $g(i) = f(i)$ if $i < n$ and $g(i) = f(0)$ otherwise. Then g is onto.

$\Leftarrow$): Suppose $f : \omega \to X$ is onto. Define $g : X \to \omega$ as follows:

$$g(x) = \min\{n : f(n) = x\}.$$

Then g is one-to-one, so $X \preceq \omega$, so X is countable. □

5.4. Basic cardinal arithmetic

DEFINITION 5.4.1. *If A and B are disjoint sets, we define $|A| + |B|$ to be $|A \cup B|$.*

If A and B are any sets, we define $|A|.|B|$ to be $|A \times B|$, and we define $|A|^{|B|}$ to be $|{}^B A|$.

PROPOSITION 5.4.2. *These operations are well-defined.*

PROOF: Obvious. □

THEOREM 5.4.3. *Suppose κ, λ and μ are cardinal numbers. Then:*

1. $\kappa + \lambda = \lambda + \kappa$,
2. $\kappa + (\lambda + \mu) = (\kappa + \lambda) + \mu$,
3. $\kappa.\lambda = \lambda.\kappa$,
4. $\kappa.(\lambda.\mu) = (\kappa.\lambda).\mu$,
5. $\kappa.(\lambda + \mu) = \kappa.\lambda + \kappa.\mu$,
6. $\kappa^{\lambda+\mu} = \kappa^\lambda.\kappa^\mu$,
7. $(\kappa^\lambda)^\mu = \kappa^{\lambda.\mu}$,
8. $(\kappa.\lambda)^\mu = \kappa^\mu.\lambda^\mu$.

PROOF: 1. If A and B are disjoint, then $A \cup B = B \cup A$, so obviously $|A \cup B| = |B \cup A|$.

3. For any sets A and B, we exhibit a bijection between $A \times B$ and $B \times A$ thus:

$$\Phi : \langle a, b\rangle \mapsto \langle b, a\rangle,$$

for any $a \in A$ and $b \in B$.

It is clear that Φ is a bijection. □

5.5. More results in cardinal arithmetic

THEOREM 5.5.1. *For any cardinals $\kappa \le \kappa'$ and $\lambda \le \lambda'$, $\kappa + \lambda \le \kappa' + \lambda'$ and $\kappa.\lambda \le \kappa'.\lambda'$.*

THEOREM 5.5.2. *For any cardinal κ, $\kappa < 2^\kappa$.*

PROOF: Theorem 3.1.2 and Corollary 3.1.4. □

COROLLARY 5.5.3. $\aleph_0 < 2^{\aleph_0} < 2^{2^{\aleph_0}} < \cdots$. *So there are infinitely many infinite cardinal numbers.*

THEOREM 5.5.4. $\aleph_0.\aleph_0 = \aleph_0$.

PROOF: Recall that $\omega \times \omega \sim \omega$; that is, $|\omega \times \omega| = |\omega|$; so $|\omega|.|\omega| = |\omega|$, or $\aleph_0.\aleph_0 = \aleph_0$. □

COROLLARY 5.5.5. $\aleph_0 + \aleph_0 = \aleph_0$; *and for all* $n \in \omega$, $n > 0$, $\aleph_0 + n = \aleph_0.n = \aleph_0$.

PROOF: By a result on the problem sheets, $\aleph_0 + \aleph_0 = \aleph_0.2$. Thus $\aleph_0 \leq \aleph_0 + n \leq \aleph_0 + \aleph_0 = \aleph_0.2 \leq \aleph_0.m \leq \aleph_0.\aleph_0 = \aleph_0$ (if $m > 1$).

By the Schröder-Bernstein Theorem, all these are equal. □

THEOREM 5.5.6. $\mathbb{Z}$ *and* $\mathbb{Q}$ *are countable.*

PROOF: See the solutions to Problem Sheet 0. □

THEOREM 5.5.7. *For any set* X, $|\wp X| = 2^{|X|}$.

PROOF: We set up a bijection between $\wp X$ and ${}^X 2$.

If $S \subseteq X$, define χ_S so that

$$\chi_S : x \mapsto \begin{cases} 1 & \text{if } x \in S, \\ 0 & \text{if } x \notin S. \end{cases}$$

Then we claim that the map $S \mapsto \chi_S$ is the bijection we are after.

One-to-oneness: If $S \neq S'$, say wlog that $x \in S$ but $x \notin S'$. Then $\chi_S(x) = 1$, but $\chi_{S'}(x) = 0$. So $\chi_S \neq \chi_{S'}$.

Ontoness: let $f : X \to 2$. Let $S = f^{-1}{}"\{1\}$. Then $\chi_S(x) = 1$ iff $x \in S$ iff $f(x) = 1$—so $\chi_S = f$. □

THEOREM 5.5.8. $|\mathbb{R}| = 2^{\aleph_0}$.

PROOF: We first show that $|\mathbb{R}| \geq 2^{\aleph_0}$.

We define a function $\Phi : {}^\omega 2 \to \mathbb{R}$ as follows:

$$\Phi(f) = \sum_{i=0}^{\infty} \frac{2.f(i)}{3^{i+1}};$$

then Φ is a bijection between ${}^\omega 2$ and the Cantor set, and so is an injection from ${}^\omega 2$ to $\mathbb{R}$.

Next we show that $|\mathbb{R}| \leq 2^{\aleph_0}$.

We define a function $\Psi : \mathbb{R} \to \wp\mathbb{Q}$ as follows.

$$\Psi(x) = \{q \in \mathbb{Q} : q < x\}.$$

Ψ is one-to-one, because if $x \neq y$, say $x < y$, then there exists $q \in \mathbb{Q}$ such that $x < q < y$; so $q \in \Psi(y)$ but $q \notin \Psi(x)$, so $\Psi(y) \neq \Psi(x)$.

Finally, by the Schröder-Bernstein Theorem, $|\mathbb{R}| = 2^{\aleph_0}$. □

6. Ordinals

There is something that cardinal arithmetic does not capture about the way we count in real life, which is, effectively, to assign *ordinal* labels to the things we are counting: when we say "one, two, three..." we are effectively designating one object to be the *first*, another to be the *second*, and so on. At the end of the process, the number of ordinals we have used tells us how many things we have counted. This process is known as *ordination*.

To generalise, we need new notation. It proceeds:

$0, 1, 2, 3, 4, 5, \ldots, \omega, \omega + 1, \omega + 2, \ldots, \omega + \omega = \omega 2, \omega 2 + 1, \omega 2 + 2, \ldots$

Every so often we need to introduce a new notation to label the next sequence of ordinals.

The infinite ordinals generated by this process have two uses. The first is to do an infinite version of ordination, so that we can, as it were, count infinite sets. The second is to do an infinite version of induction and recursion.

But in order to define the ordinals formally, we need a rather abstruse piece of theory.

6.1. Well-ordered sets

DEFINITION 6.1.1. *A relation $\leq$ on a set X is a* well-ordering *of X iff it is a total ordering, in which every non-empty subset S of X has a least element.*

The concept of a well-ordering is an abstraction from the ordinals. To illustrate this we have:

PROPOSITION 6.1.2. *ω is well-ordered by $\overset{\in}{=}$.*

PROOF: Theorem 4.1.17. □

PROPOSITION 6.1.3. *Every finite totally ordered set is well-ordered.*

PROOF: Suppose X is a totally ordered, but not well-ordered, set. We prove that for every $n \in \omega$, $|X| \geq n$; it follows that X cannot be finite.

Let S be a non-empty subset of X with no least element. We proceed by induction on n, showing that for all n, there is a function $f : n \to S$ such that $f(i+1) < f(i)$ for all $i \in n - 1$.

If $n = 0$, this is trivial, since the empty function vacuously has the given property.

Now suppose we know that $|S| \geq n$; suppose $f : n \to S$ is a one-to-one function with the given property. Then $f(n-1)$ is not the least element of S; let $a \in S$ be less. Let

$$g = f \cup \{\langle n, a \rangle\};$$

then $g : n + 1 \to S$, and $g(n) < g(n-1)$, so for all $i \in n$, $g(i+1) < g(i)$. □

COROLLARY 6.1.4. *Each natural number is well-ordered by $\overset{\in}{=}$.*

It is NOT SUFFICIENT for a relation $\leq$ to be a well-ordering of a set X for $\leq$ to be a total order on X in which X itself has a least element. $[0, 1]$, with the usual order, is a counterexample. $[0, 1]$ has a least element, but $(0, 1]$ is a subset without one. In fact:

THEOREM 6.1.5. *Every well-ordered subset of $\mathbb{R}$ is countable.*

PROOF: Problem sheets. □

The point about well-orderings is the following theorem:

THEOREM 6.1.6. *(Induction on a well-ordering) Suppose X is a set, $\leq$ is a well-ordering of X, and $\phi(x)$ is a property of elements of x. Suppose that for all x, if $\phi(y)$ holds for all $y < x$, then $\phi(x)$ holds.*

Then $\phi(x)$ holds for all x.

This is a generalisation of strong induction.

PROOF: Let $S = \{x \in X : \phi(x) \text{ does not hold.}\}$.

Suppose $\phi(x)$ does not hold for all x; then S is non-empty, so it has a least element a. Since $a \in S$, $\phi(a)$ does not hold. Since a is the least element of S, for all $y < a$, $\phi(y)$ holds; but then by assumption $\phi(a)$ holds, $\dot{\times}$. $\square$

DEFINITION 6.1.7. *Suppose X is a set, and $\leq$ is a well-ordering of X. If $x \in X$, define*

$$\operatorname{seg} x = \{y \in X : y < x\}.$$

We also have

THEOREM 6.1.8. *(Recursion on a well-ordered set) Suppose X is a set, $\leq$ is a well-ordering of X, A is a set, and Φ is a function which, for any $x \in X$, inputs a function from $\{y \in X : y < x\}$ to A, and outputs an element of A; formally,*

$$\Phi : \bigcup_{x \in X} {}^{\{y \in X : y < x\}}A \to A.$$

Then there exists a unique function $f : X \to A$ such that for all $x \in X$, $f(x) = \Phi(f \restriction \{y \in X : y < x\})$.

PROOF: (not examinable) Imitate the proof of the corresponding theorem for ω. $\square$

We will introduce more palatable versions of both these theorems in due course.

Well-ordered sets are much more similar to each other than general totally ordered sets are. First we define what we mean for two of them to be the same:

DEFINITION 6.1.9. *Suppose $\langle X, \leq_X \rangle$ and $\langle Y, \leq_Y \rangle$ are totally ordered sets. Then they are* order-isomorphic *iff there is a bijection $f : X \to Y$ such that for all $x, x' \in X$, $x \leq_X x'$ iff $f(x) \leq_Y f(x')$.*

THEOREM 6.1.10. *Suppose X and Y are sets, and $\leq_X$ and $\leq_Y$ are well-orderings of X and Y respectively. Then one of the following holds:*

1. *X is order-isomorphic to Y.*
2. *There exists $y \in Y$ such that X is order-isomorphic to $\operatorname{seg} y$.*
3. *There exists $x \in X$ such that Y is order-isomorphic to $\operatorname{seg} x$.*

We are using here a standard abuse of notation, that of writing X to refer to $\langle X, \leq_X \rangle$.

PROOF: We define a function $f : X \to Y$ by recursion as follows:

$$f(x) = \min\{y \in Y : \nexists z < x \; y = f(z)\}$$

if this set is non-empty; we leave $f(x)$ undefined otherwise.

Now if $y < x$, and $f(y)$ and $f(x)$ are both defined, then

$$\{w \in Y : \nexists z < y\, w = f(z)\} \supseteq \{w \in Y : \nexists z < x\, w = f(z)\},$$

so

$$\min\{w \in Y : \nexists z < y\, w = f(z)\} \leq \min\{w \in Y : \nexists z < x\, w = f(z)\},$$

so $f(y) \leq f(x)$. Indeed,

$$f(y) \notin \{w \in Y : \nexists z < x\, w = f(z)\},$$

so $f(y) \neq f(x) = \min\{w \in Y : \nexists z < x\, w = f(z)\}$, so $f(y) < f(x)$.

We now ask whether f is defined everywhere; that is, whether the set $\{y \in Y : \nexists z < x\, y = f(z)\}$ is non-empty for all x.

If not, then let x be least such that $\{y \in Y : \nexists z < x\, y = f(z)\}$ is empty. Then

$$g = f \restriction \operatorname{seg} x$$

is a one-to-one, onto, order-preserving function from $\operatorname{seg} x$ to Y, so $\operatorname{seg} x$ is order-isomorphic to Y.

So now suppose $f(x)$ is defined for all x. Then f is an order-isomorphism between X and $\operatorname{ran} f$.

If $\operatorname{ran} f = Y$, then X is order-isomorphic to Y.

Otherwise, let $y = \min(Y \setminus \operatorname{ran} f)$; then for all $z \geq y$, $z \notin \operatorname{ran} f$, for otherwise, suppose $z = f(x)$. Then $z = \min\{w \in Y : \nexists u < x\, f(u) = w\}$. This is clearly untrue, as $y \in \{w \in Y : \nexists u < x\, f(u) = w\}$, and $y < z$.

So $\operatorname{ran} f = \operatorname{seg} y$. So X is order-isomorphic to $\operatorname{seg} y$. □

6.2. The ordinals

Definition 6.2.1. *A set X is* transitive *iff whenever $a \in b \in X$, $a \in X$.*

Definition 6.2.2. *A set α is an* ordinal *iff it is transitive and well-ordered by $\overset{\in}{=}$.*

Theorem 6.2.3. *ω is an ordinal. So is n, for every $n \in \omega$.*

Proof: ω is transitive because if $m \in n \in \omega$, then $m \in \omega$. The statement that natural numbers are transitive is proved on the problem sheets.

We have also proved that ω and its elements are well-ordered by $\overset{\in}{=}$. □

Theorem 6.2.4. *If α is an ordinal, then so is α^+.*

Proof: First we show that α^+ is transitive.

Suppose $\gamma \in \beta \in \alpha^+$. $\alpha^+ = \alpha \cup \{\alpha\}$, so there are two cases.

If $\beta \in \{\alpha\}$, then $\beta = \alpha$, so $\gamma \in \alpha$, so $\gamma \in \alpha^+$.

If $\beta \in \alpha$, then $\gamma \in \alpha$ because α is transitive, so $\gamma \in \alpha^+$.

Now we show well-ordering. Suppose S is a non-empty subset of α^+.

There are two cases. If $S \cap \alpha \neq \varnothing$, then let β be its least element. Since $\beta \in \alpha$, β is the least element of S.

If $S \cap \alpha = \varnothing$, then $S = \{\alpha\}$, and obviously α itself is the least element of S. $\square$

THEOREM 6.2.5. *If α and β are ordinals, then $\alpha \subseteq \beta$ or $\beta \subseteq \alpha$.*

PROOF: Suppose not. Suppose $\alpha \not\subseteq \beta$ and $\beta \not\subseteq \alpha$.

Then $\alpha \cap \beta$ is transitive, for if $\xi \in \eta \in \alpha \cap \beta$, then $\xi \in \eta \in \alpha$, so $\xi \in \alpha$ since α is transitive, and $\xi \in \eta \in \beta$, so $\xi \in \beta$ since β is transitive. So $\xi \in \alpha \cap \beta$.

Let $\gamma = \min(\alpha \setminus \beta)$, and $\delta = \min(\beta \setminus \alpha)$.

We argue that $\gamma = \alpha \cap \beta$. If $\xi \in \gamma$, then $\xi \notin \alpha \setminus \beta$. On the other hand $\xi \in \gamma \in \alpha$, so $\xi \in \alpha$. So $\xi \in \alpha \cap \beta$. Now if $\xi \notin \gamma$, then $\xi \notin \alpha \cap \beta$, for otherwise, $\xi \in \alpha$, so since α is totally ordered, $\xi = \gamma$ or $\xi \ni \gamma$; and since $\alpha \cap \beta$ is transitive, $\gamma \in \alpha \cap \beta$, $\dot{\times}$.

Similarly $\delta = \alpha \cap \beta$. Hence $\gamma = \delta \in \alpha \cap \beta$, $\dot{\times}$. $\square$

THEOREM 6.2.6. *If α and β are ordinals, then $\alpha \subseteq \beta$ iff $\alpha \overset{\in}{=} \beta$.*

PROOF: $\Leftarrow$) If $\alpha = \beta$, then obviously $\alpha \subseteq \beta$. If $\alpha \in \beta$, then for all $\gamma \in \alpha$, $\gamma \in \beta$ by transitivity; so $\alpha \subseteq \beta$.

$\Rightarrow$) If $\alpha \subseteq \beta$, but $\alpha \neq \beta$, let $\gamma = \min(\beta \setminus \alpha)$.

As above, $\gamma = \alpha$, so $\alpha \in \beta$. $\square$

DEFINITION 6.2.7. *If α and β are ordinals, then $\alpha \leq \beta$ iff $\alpha \overset{\in}{=} \beta$, iff $\alpha \subseteq \beta$.*

THEOREM 6.2.8. *Suppose A is a set of ordinals. Then $\bigcup A$ is an ordinal.*

PROOF: We show first that $\bigcup A$ is transitive. Suppose $\gamma \in \beta \in \bigcup A$. Since $\beta \in \bigcup A$, there exists $\alpha \in A$ such that $\beta \in \alpha$. Since α is an ordinal, it is transitive, so $\gamma \in \alpha$. Then $\gamma \in \bigcup A$.

Now we show that $\bigcup A$ is well-ordered. Suppose S is a non-empty subset of $\bigcup A$. Let $\alpha \in S$. If α is the least element of S, then we are done, so suppose not. Then there exists $\beta \in \alpha$ such that $\beta \in S$; that is, $S \cap \alpha$ is a non-empty subset of α. Since α is an ordinal, it is well-ordered. Let β be the least element of $S \cap \alpha$. If γ now is any element of S, then either $\gamma \in \alpha$, when $\gamma \geq \beta$ as shown, or $\gamma \notin \alpha$. Then $\gamma \overset{\ni}{=} \alpha$, so $\gamma \overset{\ni}{=} \beta$, as required. $\square$

DEFINITION 6.2.9. *An ordinal α is said to be a* successor ordinal *iff there exists an ordinal β such that $\alpha = \beta^+$.*

An ordinal α is said to be a limit ordinal *iff it is not* 0 *and is not a successor ordinal.*

EXAMPLE 6.2.10. *ω is a limit ordinal. All non-zero natural numbers are successor ordinals.*

THEOREM 6.2.11. *λ is a limit ordinal iff for all $\alpha \in \lambda$, $\alpha^+ \in \lambda$.*

PROOF: If $\alpha \in \lambda$, then $\lambda \neq \alpha^+$ or λ would be a limit. Also $\alpha \subseteq \lambda$ but $\lambda \not\subseteq \alpha$. Let $\beta \in \lambda \setminus \alpha$. Then $\beta \notin \alpha$. So $\alpha \overset{\in}{=} \beta$. So $\alpha \in \lambda$, so $\alpha^+ \subseteq \lambda$. Hence $\alpha^+ \leq \lambda$, so $\alpha^+ \in \lambda$. $\square$

THEOREM 6.2.12. *Suppose A is a non-empty set of ordinals.*

If A has a greatest element α, then $\bigcup A = \alpha$.

Otherwise, $\bigcup A$ is a limit ordinal, and is the least ordinal greater than all elements of α.

PROOF: If α is the greatest element of A, then for all $\beta \in A$, $\beta \subseteq \alpha$. Thus $\bigcup A \subseteq \alpha$. The reverse is clear.

Now suppose A has no greatest member.

If $\beta \in A$, then there exists $\gamma > \beta$ such that $\gamma \in A$. So $\gamma \subseteq \bigcup A$, so $\gamma \leq \bigcup A$, so $\beta < \bigcup A$.

Now suppose $\gamma > \beta$ for all $\beta \in A$. Then $\gamma \supseteq \beta$ for all $\beta \in A$, so $\gamma \supseteq \bigcup A$. So $\gamma > \bigcup A$. □

PROPOSITION 6.2.13. *Let α be an ordinal. Then α^+ is the least ordinal greater than α.*

PROOF: Certainly $\alpha^+ > \alpha$. Suppose $\gamma > \alpha$. Then $\gamma \ni \alpha$, so $\{\alpha\} \subseteq \gamma$. Also $\gamma \supseteq \alpha$. So $\gamma \supseteq \alpha \cup \{\alpha\} = \alpha^+$. So $\gamma \geq \alpha^+$ as required. □

PROPOSITION 6.2.14. *Let α be an ordinal, and let $f : \alpha \to \alpha$ be an order-preserving function (ie. $\gamma \in \delta$ implies $f(\gamma) \in f(\delta)$).*

Then for all $\beta \in \alpha$, $f(\beta) \geq \beta$.

PROOF: Suppose not. Let $S = \{\beta \in \alpha : f(\beta) < \beta\}$.

By assumption, $S \neq \varnothing$, so S has a least element γ. Since $\gamma \in S$, $f(\gamma) < \gamma$. Since $f(\gamma) < \gamma$ and γ is the least element of S, $f(f(\gamma)) \geq f(\gamma)$.

Now, $f(\gamma) < \gamma$ while $f(f(\gamma)) \geq f(\gamma)$, $\dot{\times}$ to the assumption that f is order-preserving. □

So $\bigcup$ is the supremum operator on sets of ordinals.

COROLLARY 6.2.15. *λ is a limit ordinal iff λ is non-empty and*

$$\lambda = \bigcup \{\alpha : \alpha < \lambda\}.$$

(More snappily, $\lambda = \bigcup \lambda$.)

$\bigcup \alpha^+ = \alpha$.

6.3. Transfinite induction

THEOREM 6.3.1. *Suppose $\phi(\alpha)$ is a property of ordinals, such that, for some α, $\phi(\alpha)$ is true. Then there is a least α such that $\phi(\alpha)$ holds.*

We say "the class of ordinals is well-ordered".

PROOF: Suppose $\phi(\alpha)$. If, for all $\beta < \alpha$, $\phi(\beta)$ does not hold, then, whenever $\phi(\beta)$ holds, $\beta \geq \alpha$, so α is the least ordinal satisfying $\phi(.)$.

Now suppose there exists $\beta < \alpha$ such that $\phi(\beta)$ holds; then there is $\beta \in \alpha$ such that $\phi(\beta)$; thus

$$S = \{\beta \in \alpha : \phi(\beta)\}$$

is non-empty. Now α is well-ordered, so S has a least element. Let $\gamma = \min S$. Then if $\phi(\beta)$ holds, then either $\beta \geq \alpha$, when $\beta \geq \gamma$ also, or $\beta < \alpha$, ie $\beta \in \alpha$, so $\beta \in S$, so $\beta \geq \gamma$.

So γ is least such that $\phi(\gamma)$. □

COROLLARY 6.3.2. *(Strong transfinite induction) Suppose $\phi(\alpha)$ is a property of ordinals such that for every α, if, for all $\beta < \alpha$, $\phi(\beta)$ holds, then $\phi(\alpha)$ holds.*

Then $\phi(\alpha)$ holds for all ordinals α.

PROOF: Suppose $\phi(\alpha)$ does not hold for all ordinals α. Then $\neg\phi(\alpha)$ holds for some α. So there is a least such. But then $\phi(\beta)$ holds for all $\beta < \alpha$, $\dot{\times}$. □

A more palatable version:

THEOREM 6.3.3. *(Transfinite induction) Suppose that $\phi(\alpha)$ is a property of ordinals satisfying the following conditions:*

1. $\phi(0)$ holds.

2. For all α, if $\phi(\alpha)$ holds, then $\phi(\alpha^+)$ holds.

3. If λ is a limit ordinal, and for all $\alpha < \lambda$, $\phi(\alpha)$ holds, then $\phi(\lambda)$ holds.

Then $\phi(\alpha)$ holds for all ordinals α.

PROOF: Suppose that α is an ordinal, and for all $\beta < \alpha$, $\phi(\beta)$ holds.

Then there are three cases.

If $\alpha = 0$, then $\phi(\alpha)$ by assumption.

If α is a successor ordinal β^+, then since $\beta < \alpha$, $\phi(\beta)$ holds; so $\phi(\beta^+)$ holds by assumption; that is, $\phi(\alpha)$ holds.

If λ is a limit, then $\phi(\lambda)$ holds by assumption.

Now by strong transfinite induction, $\phi(\alpha)$ holds for all ordinals α. □

6.4. Replacement, and Transfinite Recursion

We have developed a theory of the ordinals which could, however, have only a very meagre subject matter. With the axioms of set theory we have so far, we can prove the existence of the following ordinals:

$$0,\ 1,\ 2,\ 3,\ \ldots,\ \omega,\ \omega^+,\ \omega^{++},\ \omega^{+++},\ \ldots$$

but we *cannot prove the existence of any others.* In particular, we cannot prove that there is any limit ordinal other than ω.

Replacement Axiom Scheme *Given a set X, and a rule which associates, with each element x of X, a set $\Phi(x)$,*

$$\{y : \exists x \in X\, y = \Phi(x)\}$$

is a set.

Speaking very informally, we say "the range of a function is a set". This is informal, because Φ is not a set of ordered pairs, so it is not a function; but this locution captures the basic idea of Replacement.

Using the Replacement Scheme, we can deduce the existence of lots of ordinals, as follows:

THEOREM 6.4.1. *If $\alpha \neq \beta$, then there is no order-isomorphism between α and β.*

PROOF: Suppose $f : \alpha \to \beta$ is an order-isomorphism. Then also so is $f^{-1} : \beta \to \alpha$.

Then by Proposition 6.2.14, $f(\gamma) \geq \gamma$ for all γ. The same is also true for f^{-1}; so $f(\gamma) \leq \gamma$. So $f(\gamma) = \gamma$. □

THEOREM 6.4.2. *(Hartogs' Theorem) Let X be any set. Then there exists an ordinal α such that $|\alpha| \not\leq |X|$.*

PROOF: Let $\mathscr{Y}$ be the set of all well-orderings of subsets of X such that if $\leq \in \mathscr{Y}$ is a well-ordering of Y, then there exists an ordinal $\beta_\leq$ (necessarily unique) such that $\beta_\leq$ is order-isomorphic to $\langle Y, \leq \rangle$.

Let

$$\alpha = \{\beta_{\leq} : \leq \in \mathscr{Y}\};$$

this is a set by Replacement.

We show that $|\alpha| \not\leq |X|$.

For, if $f : \alpha \to X$ is one-to-one, let $Y = \operatorname{ran} f$; define $\leq$ on Y such that $f(\gamma) \leq f(\delta)$ iff $\gamma \leq \delta$. Then $\leq$ is a well-ordering on Y, and $\langle Y, \leq\rangle$ is order-isomorphic to α. Clearly $\alpha = \beta_{\leq}$. So $\alpha \in \alpha$, $\dot{\times}$. □

COROLLARY 6.4.3. *There exists an uncountable ordinal.*

PROOF: Apply the above with $X = \omega$. □

COROLLARY 6.4.4. *Let X be a set, and let $\leq$ be a well-ordering of X. Then there exists an ordinal α such that α is order-isomorphic to X.*

PROOF: By Hartogs' Lemma, there is an ordinal α such that there is no one-to-one function $f : \alpha \to X$.

By Theorem 6.1.10, there must be an order-isomorphism between $\langle X, \leq\rangle$ and some initial segment $\operatorname{seg}\beta$ of α; $\operatorname{seg}\beta = \beta$ and β is now order-isomorphic to $\langle X, \leq\rangle$. □

We can now state a full-blooded notion of Transfinite Recursion:

THEOREM 6.4.5. *Suppose we have a rule which associates, with any function f whose domain is an ordinal, a set $\Phi(f)$.*

Then there is a unique rule which associates, with each ordinal α, a set $\Psi(\alpha)$ such that for all α,

$$\Psi(\alpha) = \Phi(\Psi\restriction\alpha).$$

[More formally, let g be the function on α such that for all $\beta \in \alpha$, $g(\beta) = \Psi(\beta)$. Then $\Psi(\alpha) = \Phi(g)$.]

In order to prove this, we would need slightly more sophistication in logic than we have at our disposal at the moment; but the proof does depend on the Axiom of Replacement (otherwise, the function g alluded to in the bit with square brackets might not be a set).

We deduce the following more appealing version:

THEOREM 6.4.6. *(Transfinite recursion) Given a set a, a rule associating a set $\Theta(x)$ with each set x, and a rule associating a set $\Phi(f)$ with each function f whose domain is a limit ordinal, there exists a unique rule associating with each ordinal α a set $\Psi(\alpha)$ such that*

1. $\Psi(0) = a$,
2. $\Psi(\alpha^+) = \Theta(\Phi(\alpha))$,
3. $\Psi(\lambda) = \Phi(\Psi\restriction\lambda)$.

6.5. Ordinal arithmetic

We now give lots of applications of transfinite recursion and induction.

We will define operations of addition and multiplication on the ordinals, generalising the operations on the natural numbers.

N But the operations of ordinal and cardinal arithmetic ARE ALMOST TOTALLY UNLIKE EACH OTHER. They are different in almost every respect except that (unfortunately) they use the same notation.

DEFINITION 6.5.1. *(Ordinal addition)*

1. $\alpha + 0 = \alpha$ *for all* α.
2. $\alpha + \beta^+ = (\alpha + \beta)^+$ *for all* α *and* β.
3. $\alpha + \lambda = \bigcup_{\beta<\lambda} \alpha + \beta$, *if* λ *is a limit ordinal.*

EXAMPLE 6.5.2. $\omega + 1 = \omega + 0^+ = (\omega + 0)^+ = \omega^+ \neq \omega$.

But $1 + \omega = \bigcup_{n\in\omega} 1 + n = \bigcup_{n\in\omega} n + 1 = \varnothing \cup \bigcup_{n\in\omega} n + 1 = \bigcup_{n\in\omega} n = \omega$.

So $\omega + 1 \neq 1 + \omega$.

So, ordinal addition is not commutative.

DEFINITION 6.5.3. *(Ordinal multiplication)*

1. $\alpha.0 = 0$ *for all* α.
2. $\alpha.\beta^+ = \alpha.\beta + \alpha$ *for all* α *and* β.
3. $\alpha.\lambda = \bigcup_{\beta<\lambda} \alpha.\beta$, *if* λ *is a limit ordinal.*

EXAMPLE 6.5.4. $2.\omega = \bigcup_{n\in\omega} 2.n = \omega$.

$\omega.2 = \omega.1 + \omega = \omega + \omega \neq \omega$.

So ordinal multiplication is not commutative. Moreover $2.\omega = (1 + 1).\omega \neq \omega + \omega$, *so the left distributive law does not hold.*

LEMMA 6.5.5. *Ordinal addition is strictly monotonic in the second argument.*

COROLLARY 6.5.6. *For all* α *and* β, $\alpha + \beta \geq \beta$.

THEOREM 6.5.7. *Suppose* $\alpha \leq \beta$. *Then there exists* γ *such that* $\beta = \alpha + \gamma$.

PROOF: Suppose this is not true, and that β is the least ordinal such that $\alpha \leq \beta$ and there does not exist γ such that $\beta = \alpha + \gamma$.

Clearly $\beta \neq \alpha$, since $\alpha = \alpha + 0$.

Also β is not a successor δ^+ with $\delta \geq \alpha$, for then if $\delta = \alpha + \gamma$, then $\delta^+ = \alpha + \gamma^+$.

Also β cannot be a limit greater than α either, for if it were, well consider $A = \{\gamma \in \beta^+ : \alpha + \gamma \leq \beta\}$. A is a set of ordinals; let $\lambda = \bigcup A = \sup A$. If A has a greatest element, then λ is that greatest element; $\alpha + \lambda \leq \beta$; but $\alpha + \lambda^+ \not\leq \beta$. Since $\alpha + \lambda^+ = (\alpha + \lambda)^+$, we must have $\beta = \alpha + \lambda$. On the other hand if A has no greatest element, then λ is a limit ordinal greater than all elements of A; and $\alpha + \lambda = \bigcup_{\gamma\in A} \alpha + \gamma$. The $\alpha + \gamma$ are all less than β; $\alpha + \lambda$ is the least ordinal greater than all of them; so we must have that $\alpha + \lambda = \beta$. □

THEOREM 6.5.8. *Ordinal addition is associative.*

PROOF: We prove by transfinite induction on γ that, for all α and β, $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$.

$\alpha + (\beta + 0) = \alpha + \beta = (\alpha + \beta) + 0$.

$\alpha + (\beta + \gamma^+) = \alpha + (\beta + \gamma)^+ = (\alpha + (\beta + \gamma))^+ = ((\alpha + \beta) + \gamma)^+ = (\alpha + \beta) + \gamma^+$.

Suppose λ is a limit. Then $(\alpha + \beta) + \lambda = \bigcup_{\gamma \in \lambda} (\alpha + \beta) + \gamma = \bigcup_{\gamma \in \lambda} \alpha + (\beta + \gamma) = \alpha + \bigcup_{\gamma \in \lambda} (\beta + \gamma) = \alpha + (\beta + \lambda)$. □

THEOREM 6.5.9. *For all ordinals α, β and γ, $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.*

PROOF: $\alpha(\beta.0) = \alpha.0 = 0 = (\alpha\beta).0$.

$\alpha(\beta.\gamma^+) = \alpha(\beta\gamma + \beta) = \alpha(\beta\gamma) + \alpha\beta = (\alpha\beta)\gamma + \alpha\beta = (\alpha\beta)\gamma^+$.

If λ is a limit, then $\alpha(\beta + \lambda) = \alpha.\bigcup_{\gamma \in \lambda} \beta + \gamma = \bigcup_{\gamma \in \lambda} \alpha(\beta + \gamma) = \bigcup_{\gamma \in \lambda} \alpha\beta + \alpha\gamma = \alpha\beta + \bigcup_{\gamma \in \lambda} \alpha\gamma = \alpha\beta + \alpha\lambda$. □

THEOREM 6.5.10. *Ordinal multiplication is associative.*

PROOF: We prove by transfinite induction on γ that, for all α and β, $\alpha.(\beta.\gamma) = (\alpha.\beta).\gamma$.

$\alpha.(\beta.0) = \alpha.0 = 0 = (\alpha.\beta).0$.

$\alpha.(\beta.\gamma^+) = \alpha.(\beta.\gamma + \beta) = (\alpha.(\beta.\gamma)) + \alpha.beta$ by the previous lemma; and is equal to $((\alpha.\beta).\gamma) + \alpha.\beta = (\alpha.\beta).\gamma^+$.

Suppose λ is a limit. Then $(\alpha.\beta).\lambda = \bigcup_{\gamma \in \lambda} (\alpha.\beta).\gamma = \bigcup_{\gamma \in \lambda} \alpha.(\beta.\gamma) = \alpha.\bigcup_{\gamma \in \lambda} (\beta.\gamma) = \alpha.(\beta.\lambda)$. □

LEMMA 6.5.11. *Suppose $\alpha \neq 0$ and $\beta < \gamma$. Then $\alpha\beta < \alpha\gamma$.*

PROOF: If $\beta < \gamma$, then there exists $\delta > 0$ such that $\gamma = \beta + \delta$. Hence if $\alpha \neq 0$, then $\alpha\gamma = \alpha\beta + \alpha\gamma > \alpha\beta$. □

7. The Axiom of Choice

The axioms we have defined so far—Extensionality, Empty set, Pairs, Unions, Subset, Foundation, Power Set, Infinity and Replacement—comprise the *Zermelo-Fraenkel* or ZF axioms. But they leave some questions unanswered, and some theorems unproved. The gap is filled by one more axiom, which was controversial in the earlier part of the last century. Some mathematicians are still uncomfortable with this extra axiom—the Axiom of Choice—today. To many set theorists, this discomfort seems strange; much modern mathematics depends on the Axiom of Choice; while the axioms of Replacement, Infinity and the Power Set pose much more serious difficulties. Nevertheless the feeling of discomfort does exist.

7.1. Cardinals and ordinals

Let us begin our tour of unproved, but reasonable, statements. We have not yet proved the following statement:

Cardinal comparability, or CC *If X and Y are sets, then $|X| \leq |Y|$ or $|Y| \leq |X|$.*

In fact, the ZF axioms are not powerful enough to prove this. We cannot even prove that every infinite set has cardinality at least $\aleph_0$.

One of our motivations for introducing the ordinals was that we would use them to "count" infinite sets. This only works if the following statement is true:

Well-ordering principle, or WO *If X is any set, then there is a well-ordering of X.*

THEOREM 7.1.1. *Suppose that WO is true. Then every set is equinumerous with some ordinal.*

PROOF: Let X be a set. Then by Hartogs' Theorem, there is an ordinal α such that $\alpha \not\preceq X$. Let $\leq$ be a well-ordering of X. Then since α is not order-isomorphic to an initial segment of X, and α is not equinumerous with X, X is order-isomorphic to an initial segment $\operatorname{seg}\beta$ of α. But $\operatorname{seg}\beta = \beta$, so $X \sim \beta$. □

COROLLARY 7.1.2. *WO implies CC.*

PROOF: Suppose X and Y are sets. Let α and β be ordinals such that $X \sim \alpha$ and $Y \sim \beta$. Then either $\alpha \subseteq \beta$, or $\beta \subseteq \alpha$. Hence $|X| \leq |Y|$ or $|Y| \leq |X|$. □

THEOREM 7.1.3. *CC implies WO.*

PROOF: Let X be any set. By Hartogs' Theorem, there exists an ordinal α such that $\alpha \not\preceq X$. By CC, $X \preceq \alpha$. Let $f : X \to \alpha$ be one-to-one. Define $\leq$ on X such that $x \leq y$ iff $f(x) \leq f(y)$. Then $\langle X, \leq\rangle$ is order-isomorphic to $\operatorname{ran} f$, so $\leq$ is a well-ordering of X. □

But... now try imagining a well-ordering of $\mathbb{R}$. The usual order is certainly not one. Neither is any other order that is easy to describe.

7.2. The Axiom of Choice

Reasonable as CC and WO are, we do not (for historical reasons) add them to our list of axioms. We add instead the following:

Axiom of Choice (AC) *Let $\mathscr{A}$ be a non-empty set of disjoint non-empty sets. Then there exists a set B such that for all $A \in \mathscr{A}$, $|A \cap B| = 1$.*

THEOREM 7.2.1. *The Axiom of Choice is equivalent to the following statement: for every set X, there is a function (a* choice function*) f from $\wp X \setminus \{\varnothing\}$ to X such that for all $A \in \operatorname{dom} f$, $f(A) \in A$.*

A choice function chooses a particular element of each set.

PROOF: $\Rightarrow$) For each $A \in \wp X \setminus \{\varnothing\}$, let $A^* = \{A\} \times A$. Then if $A_1 \neq A_2$, $A_1^* \cap A_2^* = \varnothing$.

Now let $\mathscr{A} = \{A^* : A \in \wp X \setminus \{\varnothing\}\}$, and let B be a set such that for each $A^* \in \mathscr{A}$, $|B \cap A^*| = 1$.

Then B is a choice function, for B is a set of ordered pairs of the form $\langle A, a\rangle$ where $a \in A$, and for each A, there exists a unique $a \in A$ such that $\langle A, a\rangle \in B$.

$\Leftarrow$) Let f be a choice function on $\bigcup \mathscr{A}$.

Let $B = \{f(a) : A \in \mathscr{A}\}$. Then $|B \cap A| = 1$ for each $A \in \mathscr{A}$, because the elements of $\mathscr{A}$ are disjoint. □

THEOREM 7.2.2. *AC implies WO.*

PROOF: Let X be a set. Let $f : \wp X \setminus \{\varnothing\} \to X$ be a choice function. Let $*$ be some set not belonging to X (eg, X itself).

By transfinite recursion, define x_α for each ordinal α as follows.

If $X \setminus \{x_\beta : \beta < \alpha\}$ is non-empty, let $x_\alpha = f(X \setminus \{x_\beta : \beta < \alpha\})$. Otherwise let $x_\alpha = *$.

Now let β be least such that $x_\beta = *$. Then $X \subseteq \{x_\alpha : \alpha < \beta\}$, but if $\beta < \alpha$, then by minimality of β, $x_\alpha \in X$; and if $\alpha < \gamma < \beta$, then $x_\alpha \neq x_\beta$.

Define $f : \beta \to X$ by $f(\alpha) = x_\alpha$. Then f is a bijection between β and X. Now define $\leq$ on X by $f(\beta) \leq f(\gamma)$ iff $\beta \leq \gamma$; then $\leq$ is a well-ordering of X. □

THEOREM 7.2.3. *WO implies AC.*

PROOF: Let X be a set. We show that there is a choice function for X.

Let $\leq$ be a well-ordering of X. If A is a non-empty subset of X, let $f(A) = \min A$. $\square$

COROLLARY 7.2.4. *WO, CC and AC are all equivalent.*

7.3. Zorn's Lemma

Our last principle is extraordinarily useful but rather hard to understand.

DEFINITION 7.3.1. *Let $\mathscr{A}$ be a set of sets. A* chain *in $\mathscr{A}$ is a non-empty subset $\mathscr{C}$ of $\mathscr{A}$ such that whenever $C_1, C_2 \in \mathscr{C}$, either $C_1 \subseteq C_2$ or $C_2 \subseteq C_1$.*

A maximal element *of $\mathscr{A}$ is an element A of $\mathscr{A}$ such that for all $B \in \mathscr{A}$, if $A \subseteq B$, then $A = B$.*

Zorn's Lemma (ZL) *Suppose $\mathscr{A}$ is a non-empty set of sets such that whenever $\mathscr{C}$ is a chain in $\mathscr{A}$, $\bigcup \mathscr{C} \in \mathscr{A}$. Then $\mathscr{A}$ has a maximal element.*

An example of an application of ZL:

THEOREM 7.3.2. *Suppose ZL. Then every vector space has a basis.*

PROOF: Let V be a vector space over a field $\mathfrak{F}$. Say a subset B of V is *linearly independent* iff every finite subset of it is linearly independent in the usual sense, and that it is a *spanning set* iff for all $v \in V$, there exist $v_1, \ldots, v_n \in B$, $\alpha_1, \ldots, \alpha_n \in \mathfrak{F}$ such that $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Say B is a *basis* iff it is a linearly independent spanning set.

Now let $\mathscr{B}$ be the set of all linearly independent subsets of V.

Suppose $\mathscr{C}$ is a chain in $\mathscr{B}$. We argue that $\bigcup \mathscr{C} \in \mathscr{B}$. For, suppose $v_1, \ldots, v_n \in \bigcup \mathscr{C}$. Then there exist $C_1, \ldots, C_n \in \mathscr{C}$ such that $v_i \in C_i$ for all i. Since $\mathscr{C}$ is a chain, one of the C_i is the biggest, say C_j. Then $v_1, \ldots, v_n \in C_j$. Since $C_j \in \mathscr{B}$, C_j is linearly independent; hence every finite subset is linearly independent; hence $\{v_1, \ldots, v_n\}$ is linearly independent in particular. So indeed $\bigcup \mathscr{C}$ is linearly independent, so it is an element of $\mathscr{B}$.

So ZL can be applied to $\mathscr{B}$. Let B be a maximal element.

Suppose B is not a spanning set. Let $v \in V$ have the property that there do not exist $v_1, \ldots, v_n \in B$ and $\alpha_1, \ldots, \alpha_n \in \mathfrak{F}$ such that $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Then for all $v_1, \ldots, v_n \in B$, $\{v_1, \ldots, v_n, v\}$ is linearly independent. Hence $B \cup \{v\}$ is linearly independent. So B is not maximal after all, $\dot{\times}$.

Hence B is a basis, as required. $\square$

This proof exemplifies applications of ZF:

1. Identify a suitable $\mathscr{A}$.
2. Show that ZF can be applied to $\mathscr{A}$.
3. Find a maximal element using ZL.
4. Show the maximal element has the property you want, by showing that if it did not, then it could not be maximal.

THEOREM 7.3.3. *ZL implies AC.*

PROOF: Problem sheets. $\square$

THEOREM 7.3.4. *WO implies ZL.*

PROOF: Suppose $\mathscr{A}$ is a family of sets having the property than any union of a chain in $\mathscr{A}$ is in $\mathscr{A}$.

Let $<$ be a well-ordering of $\mathscr{A}$. By Theorem 6.4.1, let α be an ordinal order-isomorphic to $\langle \mathscr{A}, < \rangle$, let $\pi : \alpha \to \mathscr{A}$ be the isomorphism, and write A_α for $\pi(\alpha)$.

We define $f : \mathscr{A} \to \{\mathsf{yes}, \mathsf{no}\}$ by recursion, with the inductive hypothesis being that for all β, $\{\gamma < \beta : f(\gamma) = \mathsf{yes}\}$ is a chain.

Suppose we have defined $f(A_\gamma)$ for all $\gamma < \beta$. If $A_\beta \supseteq A_\gamma$ whenever $\gamma < \beta$ and $f(A_\gamma) = \mathsf{yes}$, then define $f(A_\beta) = \mathsf{yes}$, otherwise define $f(A_\beta) = \mathsf{no}$.

It is thus clear that the inductive hypothesis goes through, and $\mathscr{C} = f^{-1}\{\mathsf{yes}\}$ is a chain. Let $A = \bigcup \mathscr{C}$. Then $A \in \mathscr{A}$, so $A = A_\beta$ for some β. Now for all $\gamma < \beta$, if $f(A_\gamma) = \mathsf{yes}$, then $A_\gamma \in \mathscr{C}$ so $A_\gamma \subseteq A = A_\beta$. Hence $f(A_\beta) = \mathsf{yes}$ by definition of f. Thus $A = A_\beta \in \mathscr{C}$.

We now show that A is a maximal element of $\mathscr{A}$. For if not, suppose $A \subset A'$, and that $A' = A_\delta$. Then for all $\gamma < \delta$ for which $f(A_\gamma) = \mathsf{yes}$, $A_\gamma \in \mathscr{C}$, so $A_\gamma \subseteq A$, so $A_\gamma \subseteq A_\delta$. So by definition of f, $f(A_\delta) = \mathsf{yes}$. Hence $A_\delta \in \mathscr{C}$, so $A_\delta \subseteq A$, $\dot{\times}$.

So $\mathscr{A}$ has a maximal element, as required. □

COROLLARY 7.3.5. *CC, WO, AC and ZL are all equivalent.*

8. The theory of the cardinal numbers

In this section, we assume ZFC.

8.1. Initial ordinals

DEFINITION 8.1.1. *An infinite ordinal α is said to be an initial ordinal if and only if there does not exist $\beta < \alpha$ such that β is equinumerous with α.*

THEOREM 8.1.2. *Every infinite set is equinumerous with a unique initial ordinal.*

PROOF: By Theorem 7.1.1, any infinite set X is equinumerous with some ordinal γ. Let α be the least element of the set

$$\{\delta \in \gamma + 1 : \delta \sim \gamma\}.$$

It is obvious that α is equinumerous with X. We now observe that α is an initial ordinal. For, suppose $\epsilon \sim \alpha$. Then either $\epsilon \in \alpha$ or $\alpha \in \epsilon$, or $\alpha = \epsilon$, by Theorem 6.2.5. If $\epsilon \in \alpha$, then $\epsilon \in \gamma + 1$, so by definition of α, $\epsilon \geq \alpha$. In the other two cases, of course, $\epsilon \geq \alpha$ already. □

By transfinite recursion, we can give names to the initial ordinals as follows.

DEFINITION 8.1.3. *Define ω_α, by recursion on ordinals α, as follows.*

1. ω_0 is the smallest initial ordinal (namely ω itself).

2. $\omega_{\alpha+1}$ is the smallest initial greater than ω_α.

3. $\omega_\lambda = \bigcup_{\alpha<\lambda} \omega_\alpha$ (if λ is a limit); this is the smallest initial ordinal greater than ω_α for all $\alpha < \lambda$.

THEOREM 8.1.4. *ω_α exists for all ordinals α.*

PROOF: We require to prove that if ω_β exists for all $\beta < \alpha$, then there is an initial ordinal greater than all of them.

But, the set $A = \{\omega_\beta : \beta < \alpha\}$ exists by the Replacement Schema. It is a set of ordinals, so its union $\lambda = \bigcup A$ is an ordinal, and is in fact the supremum of A. Now $|\wp\lambda| > |\lambda| \geq |\omega_\beta|$ for all $\beta < \alpha$. There is a unique initial ordinal equinumerous with $\wp\lambda$; and this cannot be any of the ω_β for $\beta < \alpha$. □

DEFINITION 8.1.5. *We define $\aleph_\alpha$ to be the cardinal number of ω_α, for all α.*

We can now give a (conventional) definition of a cardinal number.

DEFINITION 8.1.6. *A sets X is said to be a cardinal number iff either $X \in \omega$, or X is an initial ordinal. We identify $\aleph_\alpha$ with ω_α.*

THEOREM 8.1.7. *(Burali-Forti Paradox) There is no set of all ordinals.*

PROOF: Suppose O is the set of all ordinals. Then $\sup O + 1$ is another ordinal greater than all the elements of O, which is impossible. □

COROLLARY 8.1.8. *There is no set of all cardinal numbers.*

PROOF: Suppose C is the set of all cardinal numbers. Then $\{\alpha : \aleph_\alpha \in C\}$ is a set by Replacement, and is the set of all ordinals, $\dot{\times}$. □

8.2. More cardinal arithmetic

The Axiom of Choice has the effect of reducing cardinal addition and multiplication to triviality.

THEOREM (ZFC) 8.2.1. *Suppose X is an infinite set. Then X is equinumerous with $X \times X$.*

PROOF: Let $\mathscr{A}$ be the set of all functions f such that for some subset Y of X, f is a bijection from Y to $Y \times Y$. Then if $\mathscr{C}$ is a chain in $\mathscr{A}$, then $\bigcup \mathscr{C} \in \mathscr{A}$ also.

Also $\mathscr{A}$ is non-empty, because by (CC), $|X| \geq \aleph_0$, so there exists a countably infinite subset Y of X; then we can certainly find a bijection f between Y and $Y \times Y$, and then $f \in \mathscr{A}$.

So by Zorn's Lemma, $\mathscr{A}$ has a maximal element F. Suppose F is a bijection between Y and $Y \times Y$.

Suppose that $|X \setminus Y| \geq |Y|$. Let Z be a subset of $X \setminus Y$ equinumerous with Y. Now Y is infinite, and $|Y| \leq 3|Y| \leq |Y|.|Y| = |Y|$; so $|Y| = 3|Y| = 3\big(|Y|\big)^2$. So there is a bijection g between Z and $Y \times Z \cup Z \times Y \cup Z \times Z$. Then $F \cup g$ is a bijection between $Y \cup Z$ and $(Y \cup Z) \times (Y \cup Z)$. Now $F \cup g \in \mathscr{A}$, thus F is not maximal after all, $\dot{\times}$.

Thus by (CC), $|X \setminus Y| < |Y|$.

Now $|X| = |Y| + |X \setminus Y| \leq 2|Y| \leq |Y|.|Y| = |Y|$.

Since $|Y| = |Y|.|Y|$, $|X| = |X|.|X|$, so X is equinumerous with $X \times X$. □

COROLLARY 8.2.2. *Suppose κ and λ are cardinal numbers, κ is infinite, and $1 \leq \lambda \leq \kappa$. Then $\kappa + \lambda = \kappa.\lambda = \kappa$.*

PROOF: $\kappa \leq \kappa + \lambda \leq \kappa + \kappa = \kappa.2 \leq \kappa.\lambda \leq \kappa.\kappa = \kappa$ by the above theorem. □

8.3. Cardinal exponentiation (not on the syllabus)

So, under ZFC, cardinal addition and multiplication are very unexciting. What about exponentiation? Let's take the simplest example possible. What exactly is $2^{\aleph_0}$? Which of the $\aleph_\alpha$ is it?

CONJECTURE 8.3.1. *(Continuum Hypothesis)* $2^{\aleph_0} = \aleph_1$.

THEOREM 8.3.2. *(Gödel) From ZFC, it is impossible to disprove the Continuum Hypothesis.*

THEOREM 8.3.3. *(Cohen) From ZFC, it is impossible to prove the Continuum Hypothesis.*

To get a feel for the range of uncertainty that exists, we need another definition.

DEFINITION 8.3.4. *Suppose α is a limit ordinal. The cofinality of α is the least ordinal κ such that there exists a function $f : \kappa \to \alpha$ such that for all $\beta < \alpha$, there exists $\mu < \kappa$ such that $\beta < f(\mu)$.*

The cofinality of an ordinal is always an initial ordinal, that is, a cardinal.

EXAMPLES 8.3.5. *Every countable limit ordinal has cofinality ω. The first uncountable ordinal ω_1 has cofinality ω_1 (otherwise it would be a countable union of countable sets). However ω_ω has cofinality ω.*

DEFINITION 8.3.6. *A cardinal κ is a successor cardinal iff it is $\aleph_{\alpha+1}$ for some α. It is a limit cardinal iff it is $\aleph_\lambda$ for some limit ordinal λ.*

THEOREM 8.3.7. *Let κ be a cardinal. Then the cofinality of 2^κ is greater than κ.*

THEOREM 8.3.8. *Let κ be a cardinal of uncountable cofinality. Then it is impossible to prove from ZFC that $2^{\aleph_0} \neq \kappa$.*

For example, $2^{\aleph_0}$ could be $\aleph_3$, $\aleph_{57}$, $\aleph_{\omega+1}$,... But it couldn't be $\aleph_\omega$.

Actually it gets worse.

THEOREM 8.3.9. *For each non-limit ordinal α, let β_α be an ordinal such that*

1. Either β_α is a successor ordinal, or the cofinality of β_α is greater than that of ω_α,

2. If $\alpha \leq \gamma$, then $\beta_\alpha \leq \beta_\gamma$,

3. $\beta_\alpha > \alpha$.

Then the following statement cannot be disproved from ZFC: for all α such that α is not a limit ordinal, $2^{\aleph_\alpha} = \aleph_{\beta_\alpha}$.

For example, we could put $\beta_\alpha = \alpha+1$. Then we have the following statement consistent with ZFC: for all α, $2^{\aleph_\alpha} = \aleph_{\alpha+1}$.

The situation is less totally chaotic at limit ordinals. At limit ordinals of uncountable cofinality, we have:

THEOREM 8.3.10. *Let λ be a limit ordinal of uncountable cofinality. Suppose that for all $\alpha < \lambda$, $2^{\aleph_\alpha} = \aleph_{\alpha+1}$. Then $2^{\aleph_\lambda} = \aleph_{\lambda+1}$.*

This method of proof absolutely does not work at cardinals like $\aleph_\omega$. The best we have to date is the rather striking

THEOREM 8.3.11. *Suppose that for all $n \in \omega$, $2^{\aleph_n} = \aleph_{n+1}$. Then $2^{\aleph_\omega} \leq \aleph_{\omega_4}$.*