

INTRODUCTION
TO
PURE
MATHEMATICS

MICHAELMAS TERM
2007

BRIAN STEWART
EXETER COLLEGE

Introduction

The purpose of the five lectures in the first week of the Mathematics course is to introduce some standard mathematical ideas and notations.

It is important that you understand the status of this document.

It is not intended to define the syllabus—see your **Course Handbook** for that. It is not intended as a textbook; follow the links from <http://www.maths.ox.ac.uk/current-students/> for that.

I have prepared these notes mainly for myself. When I give the lectures I certainly don't intend to just copy them on to the board, but what I say won't stray far beyond what I have typed up, and won't omit very much of it. The examples will be whatever seems right at the time.

The material does not come in five blocks of equal weight, so the sections probably don't correspond to lectures.

It would be a mistake for you not to take your own notes in lectures—you will find it much easier to keep engaged with the subject and learn from lectures if you do take notes. The lectures in the first week are an opportunity to develop note-taking skills.

Students who need to have the source (\LaTeX) of the document, or who need to generate very large print versions, should get in touch with this year's lecturer — Michael Vaughan-Lee.

Please let me know of any necessary corrections: preferably by email.

Michael Vaughan-Lee
vlee@maths.ox.ac.uk
18th July 2007

Contents

1	Sets	1
1.1	Membership is all	1
1.2	Logic	1
1.3	Naive Notation for Sets	2
1.4	Some Famous Sets	2
1.5	Set Inclusion	3
1.6	(Simple) Unions and Intersections	4
1.7	Set Minus	4
1.8	An exercise	4
2	More Recipes for Sets	5
2.1	The Power Set	5
2.2	Families of Sets	6
2.3	Unions	6
2.4	Intersections	6
2.5	Ordered Pairs	7
2.6	Cartesian Product	7
2.7	An Exercise	8
3	Functions	8
3.1	Preliminary Ideas	8
3.2	Functions: the definition	9
3.3	Other Names	10
3.4	Important Notation	10
3.5	Some Famous Functions	10
3.6	Composition of Functions	11
3.7	Restrictions of Functions	11
3.8	Injective Functions	12
3.9	Surjective Functions	12
3.10	Bijjective Functions	12
3.11	Identity Maps	12
3.12	Invertible Functions	13
3.13	Images and Pre-images	14
4	Induction	15
4.1	Well-ordering of \mathbb{N}	15
4.2	Least Criminals	15
4.3	Mathematical Induction	15
4.4	An Application	16
5	How big are sets?	18
5.1	Equinumerous sets	18
5.2	Finite Sets	18
5.3	Infinite sets	18
5.4	Countably infinite sets	18
5.5	Two equivalent definitions	20
5.6	Uncountable sets	21

1 Sets

1.1 Membership is all

Sets are the building-blocks of mathematics.

Intuitively they are ‘collections’ of ‘elements’. But we don’t care what things are, we only care how they are related. What we need to know is which elements in the mathematical universe belong to which sets.

We capture this key idea by writing $a \in A$, or by saying:
‘(the element) a is a member of (the set) A ’.

Moreover, in mathematics we can only distinguish sets if they have at least some different members. More formally,

$$A = B \quad \text{if and only if} \quad \forall x (x \in A \iff x \in B).$$

Here (and elsewhere) the ‘ \forall ’ symbol means ‘for all’. So this mathematical formula translates to ‘for all x , x is a member of A if and only if x is a member of B ’.

1.1.1 Notation

In these five lectures quite a lot of notation will be introduced: part of our work this week is to develop a *common* unambiguous language which we can use to talk about mathematics. It would be overwhelming to introduce it all with no motivation at the very outset, so we’ll meet it bit by bit as we go along. After each lecture you should be checking up what’s new and trying it out. Equally important, if you think you already know something, check that you use it in the same way as we will do this term — we need a *common* language.

So far we have met

$x \in X$: pronounced ‘ x is a member of X ’.

$\forall x$: pronounced ‘for all x ’.

Here is one more:

$\exists x$: pronounced ‘there exists x ’.

1.2 Logic

We need to take care that we use some of the shortest words in exactly the same way. Here is how mathematicians use certain of these.

1.2.1 And

‘ \mathcal{P} and \mathcal{Q} is true’ means that both \mathcal{P} is true and \mathcal{Q} is true.

1.2.2 Or

‘ \mathcal{P} or \mathcal{Q} is true’ means that at least one of \mathcal{P} and \mathcal{Q} is true. (The so-called ‘*weak or*’.)

1.2.3 Implies

‘ $\mathcal{P} \implies \mathcal{Q}$ is true’ means that either \mathcal{P} is false or \mathcal{Q} is true.

There are many ‘elegant variations’ of this: all the following mean the same:

‘ \mathcal{P} implies \mathcal{Q} is true’

‘If \mathcal{P} then \mathcal{Q} ’

‘ \mathcal{Q} provided \mathcal{P} ’

‘ \mathcal{P} only if \mathcal{Q} ’

It’s important to note that these all do mean the same, and that such sentences as ‘If $A \neq A$ then $A = A$ ’ are taken in mathematics to be *true*.

1.2.4 Iff

‘ $\mathcal{P} \iff \mathcal{Q}$ is true’ means that either both \mathcal{P} and \mathcal{Q} are true, or both \mathcal{P} and \mathcal{Q} are false.

Elegant variations of this include:

‘ \mathcal{P} if and only if \mathcal{Q} is true’

‘ \mathcal{P} iff \mathcal{Q} is true’

‘ \mathcal{P} and \mathcal{Q} are equivalent’.

1.3 Naive Notation for Sets

We may indicate which set we want to talk about by using notation like this:

$\{a, b, c, \dots\}$ — note that the braces must be braces, not brackets $[]$ or parentheses $()$ or whatever. We pronounce this ‘the set whose members are a, b, c and so on’.

$\{x : x \text{ has the property } \mathcal{P}\}$, pronounced ‘the set of all x such that x satisfies the property \mathcal{P} ’. This is sometimes written $\{x | x \text{ has the property } \mathcal{P}\}$, or even $\{x | \mathcal{P}(x)\}$. It is still pronounced the same way.

Actually, this notation is highly dangerous and may lead us into contradiction: much safer is

$\{x \in X : \mathcal{P}(x)\}$, pronounced ‘the set of all elements which are members of the set X and which satisfy the property \mathcal{P} ’.

1.3.1 Examples

$$\{n | \exists x, \exists y, \exists z (x^n + y^n = z^n)\}$$

$$\{z | z^2 + z + 1 = 0\}$$

1.4 Some Famous Sets

Here are the basic sets of mathematics with which we will be concerned this term.

\mathbb{N} : the *natural numbers* $0, 1, 2, \dots$

1.4.1 Warning

Despite my hope for no ambiguity, we have to accept that not all mathematicians speak exactly the same dialect. There is a difference of opinion about which set we should denote by \mathbb{N} . Some people think 0 should not be a member. It makes absolutely no difference to the mathematics we can do, and allowing 0 full membership rights will save us a lot of special cases later in the term. Of course I also think I am *right*¹ but that's irrelevant.

The *warning* is this: check up which notation any textbook you read is using, and adapt accordingly.

Now back to our list of sets.

\mathbb{N} : the *natural numbers* $0, 1, 2, \dots$

\mathbb{Z} : the *integers* $0, \pm 1, \pm 2, \dots$

\mathbb{Q} : the *rational numbers*, $\{x \mid \exists m \exists n (x = \frac{m}{n}, m \in \mathbb{Z}, n \in \mathbb{Z}, n \neq 0)\}$.

\mathbb{R} : the *real numbers*, of which much, much more — later!

\mathbb{C} : the *complex numbers*, $\{z : \exists a \in \mathbb{R}, \exists b \in \mathbb{R} (z = a + bi)\}$.

Last, least, but highly important:

\emptyset : the (only) set with no members.

Note that this is *not* ϕ . It is pronounced ‘(the) empty set’ and not ‘phi’.

1.4.2 Intervals

If $a, b \in \mathbb{R}$ then we define intervals as follows

$(a, b) := \{x \in \mathbb{R} \mid a < x < b\}$, pronounced ‘the *open interval* a, b ’ is defined to be the set of all real numbers x such that x is greater than a and less than b .

$[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$, pronounced ‘the *closed interval* a, b ’ is defined to be the set of all real numbers x such that x is greater than or equal to a and less than or equal to b .

Another *warning* on ambiguity: the notation (a, b) has another meaning which we will come to soon. The French have attempted to avoid the ambiguity by writing this one $]a, b[$ but this hasn't caught on.

1.5 Set Inclusion

Given sets A and B we write $A \subseteq B$ (pronounced ‘ A is a subset of B ’) if every member of A is also a member of B .

From our very first paragraph, then,

$$A = B \iff A \subseteq B \text{ and } B \subseteq A.$$

We can illustrate this in a picture: [omitted]

Example. $\{n \in \mathbb{N} \mid n \text{ is even}\} \subseteq \mathbb{R}$

¹I side with those who think that the answer to the questions like ‘How many two-headed Dr Stewarts do you know?’ and ‘How many square triangles are there?’ is ‘None’ and not ‘I can't begin to count them.’

1.5.1 Notation

Suppose that A is a subset of B . We can also use the notation $B \supseteq A$.

Although we always ought to take care to define all new notation we do sometimes slip up. If we've defined $X \preceq Y$ we will sometimes write $A \succ B$ instead of $B \preceq A$ and leave the reader to make the obvious guesses.

1.6 (Simple) Unions and Intersections

Given any sets A and B we define

$$A \cup B := \{x \mid x \in A \text{ or } x \in B\}$$

$$A \cap B := \{x \mid x \in A \text{ and } x \in B\}$$

These are called the '*union*' and the '*intersection*' and we can draw pictures so: [omitted].

Example. $\{0, 1, 2\} \cup \{0, 2, 4, 8\} = \{0, 1, 2, 4, 8\}$ and $\{0, 1, 2\} \cap \{0, 2, 4, 8\} = \{0, 2\}$.

1.7 Set Minus

Given two sets A and B we write $A \setminus B$ (pronounced '*A less B*', '*A remove B*', '*A setminus B*') for the set whose members are precisely those elements of A which are not members of B .

More succinctly, $A \setminus B := \{x \in A \mid x \notin B\}$.

Here again is a picture: [omitted].

Example. $\{0, 1, 2\} \setminus \{0, 2, 4, 6\} = \{1\}$.

1.7.1 Notation

Note the new abbreviation: $x \notin X$ pronounced '*x is not a member of X*'.

This is standard: if we have defined some notion $a \simeq b$ then we write $a \not\simeq b$ to mean that it is not the case that $a \simeq b$.

1.8 An exercise

The purpose of this is to give you a model for writing out proofs of elementary things about sets. The actual result is very dull.

Example. For any sets A, B, C we have

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C).$$

There are two ways we can set the proof out. Both depend on the property we stated in the very first paragraph: two sets are equal if and only if they have exactly the same members. It is worthwhile learning to write out both sorts of proof

Proof. (a) We will prove first that every member of $(A \cup B) \cap C$ is a member of $(A \cap C) \cup (B \cap C)$.

If $x \in (A \cup B) \cap C$ then by definition of \cap we have that $x \in (A \cup B)$ and $x \in C$. Using the definition of \cup , we have that $x \in A$ or $x \in B$, and under all circumstances $x \in C$. Hence $x \in A$ and $x \in C$, or $x \in B$ and $x \in C$. Using the definition of \cap twice we see that $x \in A \cap C$ or $x \in B \cap C$. Using the definition of \cup we see that $x \in (A \cap C) \cup (B \cap C)$ as required.

(b) Now we will prove that every member of $(A \cap C) \cup (B \cap C)$ is a member of $(A \cup B) \cap C$.

... [omitted]

(c) Hence the sets $(A \cap C) \cup (B \cap C)$ and $(A \cup B) \cap C$ are equal. □

Proof. We (systematically) list the possibilities for the membership of a general element x , using the definitions of \cup and \cap . This is called a ‘truth-table’ proof.

A	B	C	$A \cup B$	$(A \cup B) \cap C$	$A \cap C$	$B \cap C$	$(A \cap C) \cup (B \cap C)$
Y	Y	Y	Y	Y	Y	Y	Y
Y	Y	N	Y	N	N	N	N
Y	N	Y	Y	Y	Y	N	Y
Y	N	N	Y	N	N	N	N
N	Y	Y	Y	Y	N	Y	Y
N	Y	N	Y	N	N	N	N
N	N	Y	N	N	N	N	N
N	N	N	N	N	N	N	N

The fact that $x \in (A \cup B) \cap C$ if and only if $x \in (A \cap C) \cup (B \cap C)$ proves the result. □

2 More Recipes for Sets

So far we’ve introduced some of the basic sets of Mathematics — $\emptyset, \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and \mathbb{C} — and defined some easy ways to make new sets from old ones. We now want to give some more complicated ways of manufacturing new (and interesting) sets.

2.1 The Power Set

Let X be any set. We write $\mathcal{P}(X)$, which we pronounce as ‘the power set of X ’, for the set whose members are precisely the subsets of X :

$$\mathcal{P}(X) := \{A \mid A \subseteq X\}.$$

Example. Suppose that $X = \{0, 1\}$ is the set which has just two members, namely 0 and 1. Then

$$\mathcal{P}(X) = \{\emptyset, \{0\}, \{1\}, X\}$$

a set with four elements.

One example to be absolutely sure about is this:

Example. $\mathcal{P}(\emptyset) = \{\emptyset\} \neq \emptyset$.

2.2 Families of Sets

Suppose we have a set \mathcal{X} whose members are themselves sets. We can often avoid confusing ourselves if we say that \mathcal{X} is a ‘family’ of sets, rather than a set of sets. Mathematically it makes no difference, but psychologically it may help.

Example. *We might speak of ‘the family of closed and bounded intervals of \mathbb{R} ’ instead of writing*

$$\{I : \exists a \in \mathbb{R} \quad \exists b \in \mathbb{R} \quad \text{with } a < b \text{ and } I = [a, b] \quad \}$$

2.3 Unions

Suppose that \mathcal{X} is a set, and that each member of \mathcal{X} is itself a set. We then define the set

$$\bigcup \mathcal{X} := \{x \mid x \in X \text{ for some } X \in \mathcal{X}\}$$

which we call ‘the union of \mathcal{X} ’.

It is often easier to use slightly different notation. Suppose that A is a set, and that for each $\alpha \in A$ there is a set X_α ‘indexed by’ α . Then

$$\bigcup_{\alpha \in A} X_\alpha := \{x \mid x \in X_\beta \text{ for some } \beta \in A\}.$$

Example. *For each $m \in \mathbb{R}$ let L_m be the line in the plane through the origin whose slope (gradient) is m . Then*

$$\bigcup_{m \in \mathbb{R}} L_m = \{P : P \text{ does not lie on the } y\text{-axis}\} \cup \{\text{the origin}\}.$$

2.4 Intersections

Now suppose that A is a *non-empty* set, and that for each $\alpha \in A$ there is a set X_α indexed by α . Then

$$\bigcap_{\alpha \in A} X_\alpha := \{x \mid x \in X_\beta \text{ for all } \beta \in A\}.$$

Example. *For each $m \in \mathbb{R}$ let L_m be the line in the plane through the origin whose slope (gradient) is m . Then*

$$\bigcap_{m \in \mathbb{R}} L_m = \{\text{the origin}\}.$$

2.4.1 Notation

In the notations $\bigcup_{\alpha \in A} X_\alpha$ and $\bigcap_{\alpha \in A} X_\alpha$ the α is a ‘dummy’:

$$\bigcup_{\alpha \in A} X_\alpha = \bigcup_{\beta \in A} X_\beta = \bigcup_{\mathbb{J} \in A} X_{\mathbb{J}}$$

and so on.

2.5 Ordered Pairs

If a and b are mathematical objects we write (a, b) for the pair taken in that order, ‘the ordered pair’. The characteristic property is that $(x, y) = (a, b)$ if and only if both $x = a$ and $y = b$.

Example.

$$(\emptyset, \{\emptyset\}) \neq (\{\emptyset\}, \emptyset)$$

Proof. It is not possible that $\emptyset = \{\emptyset\}$, since sets are equal if and only if they have the same members. Now $\emptyset \in \{\emptyset\}$, but $\emptyset \notin \emptyset$ since \emptyset has no members. \square

2.5.1 Ordered k -tuples

Suppose that a_1, \dots, a_k are mathematical objects. We use the notation (a_1, \dots, a_k) for the k -tuple taken in that order. Again, what matters is that

$$(a_1, \dots, a_k) = (b_1, \dots, b_k) \iff a_j = b_j \text{ for all } j = 1, 2, \dots, k.$$

We could, if we were very keen, build everything up from ordered pairs by putting $(x, y, z) := ((x, y), z)$ and so on. But we will not do this.

2.5.2 An Ambiguity

This is the promised ambiguity: if a and b are real numbers then (a, b) may denote either the ordered pair as here, or the interval as earlier. We rely on the context to make it clear what’s intended. If you think there might be a doubt in your reader’s mind, then say ‘the ordered pair (a, b) ’ or ‘the interval (a, b) ’.

2.6 Cartesian Product

Let A and B be sets. The Cartesian product of the sets is defined to be

$$A \times B := \{(a, b) \mid a \in A, b \in B\}.$$

Example. $\mathbb{R} \times \mathbb{R}$ is ‘just’ the set of points in the usual real plane.

Example. $\{0, 1, 2\} \times \{0, 1\}$ has precisely six members, namely $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$, $(2, 0)$, and $(2, 1)$.

Here is a picture: [omitted]

Note that we get moderately good pictures of $A \times B$ by assuming that $A \subset \mathbb{R}$, and $B \subset \mathbb{R}$ and drawing the usual picture in the plane, with A a subset of the x -axis, and B a subset of the y -axis.

2.6.1 A generalisation

Suppose that A_1, \dots, A_k are sets; then

$$A_1 \times \dots \times A_k := \{(a_1, \dots, a_k) \mid a_j \in A_j \text{ for } j = 1, 2, \dots, k\}.$$

2.6.2 Cartesian Powers

Suppose that A is a set, and $k \geq 2$ a natural number. We then define A^k to be the set $A_1 \times \cdots \times A_k$ where every $A_j = A$.

Example. \mathbb{R}^2 is the usual plane, \mathbb{R}^3 the usual 3-space.

2.6.3 Special Cases

If you are a careful reader you will be protesting. What about $k = 1$? And why exclude the case $k = 0$?

Well for $k = 1$ just define $A^1 := A$, suppress any distinction between a and (a) and no problems will arise.

For $k = 0$, take a deep breath and then define $A^0 := \{\emptyset\}$.

2.7 An Exercise

Example. $(A \times B) \cap (B \times A) = (A \cap B)^2$.

Proof. To prove two sets are equal we need to show that they have exactly the same members. So we must show two things.

First show $(A \times B) \cap (B \times A) \subseteq (A \cap B)^2$.

Suppose x is a member of the left-hand side.

Then x must (by definition of \cap) be a member of $A \times B$ and also of $B \times A$. That is, $x = (\alpha, \beta)$, where $\alpha \in A$ (as $x \in A \times B$) and $\alpha \in B$ (as $x \in B \times A$). That is, $\alpha \in A \cap B$.

Similarly $\beta \in A \cap B$.

Hence $x \in (A \cap B) \times (A \cap B) = (A \cap B)^2$ and we are done.

And next show $(A \times B) \cap (B \times A) \supseteq (A \cap B)^2$.

Take x a member of the right-hand side. Then $x = (\gamma, \delta)$ for some $\gamma, \delta \in A \cap B$. Then, as $\gamma \in A$ and $\delta \in B$, we have that $x \in A \times B$.

A similar argument shows that $x \in B \times A$, so that $x \in (A \times B) \cap (B \times A)$ as required. □

3 Functions

At this stage we have set up all our notation about sets, and are ready to use it in the first non-trivial way — giving a definition of ‘function’. Mathematics is the study of ‘functions’ so it is worth getting this idea straight.

3.1 Preliminary Ideas

Without thinking too much about it, most people would start off explaining that a function is a ‘rule’ that gives ‘values’ to ‘variables’. Things like x^2 , or like $\sqrt{x^2 + x + 1}$ or $\frac{1}{x^2 - 1}$ or $\log(x^{23} - 42)$ or $A \cos \omega t$ or $\frac{az + b}{cz + d}$ or whatever.

Soon we run into problems. What sort of values do we allow? What sort of values can we give the variables? What counts as an admissible rule?²

Note. *Occasionally, in order to make progress, we need to clear our minds and make a fresh start. This is one of these places. You will make better progress if you put out of your mind for a week or two all pre-conceptions about what a function ‘is’ and work with the definition that you are about to learn. Once you have learned to work with this new definition then you can go back to thinking about functions in the intuitive way you have got used to. But from time to time, when you need to prove something carefully, then you will need to use this new definition.*

3.2 Functions: the definition

A function consists of three things:

- (i) a set A , called the *domain* of f ;
- (ii) a set B , called the *codomain* of f ;
- (iii) a set $f \subseteq A \times B$ (sometimes called *the rule*) satisfying
 - (a) for each $a \in A$ there is at least one $b \in B$ such that $(a, b) \in f$;
 - (b) if $(a, b) \in f$ and $(a, b') \in f$ then $b = b'$.

Intuitively, then, the set f associates with each element $a \in A$ precisely one $b \in B$, namely the unique one with $(a, b) \in f$.

Note. *Two functions are equal if and only if they have the same domain, the same codomain, and the same rule.*

Example. *The following is a function: the domain is \mathbb{R} , the codomain is \mathbb{R} , the rule is $\{(x, y) | y = x^2 + x + 1\}$.*

The following is also a function: the domain is \mathbb{R} , the codomain is \mathbb{C} , the rule is $\{(x, y) | y = x^2 + x + 1\}$.

These are not the same function.

Example. *The following is a function: the domain is \mathbb{Z} , the codomain is $\{0, 1\}$, the rule is $\{(n, 0) | n \text{ is even}\} \cup \{(n, 1) | n \text{ is odd}\}$.*

Example. *There are precisely 8 functions with domain $\{0, 1, 2\}$ and codomain $\{0, 1\}$.*

(There are two choices for an element to pair with each of the three elements of the domain).

Example. *Let $A = B = \mathbb{R}$ and let the rule be $\{(x, x^2) | x \in \mathbb{R}\}$. If we plot the set of points $\{(x, x^2) | x \in \mathbb{R}\}$ in the (x, y) -plane in the usual way then we get what at school you might have called ‘the graph of the function $y = x^2$ ’. So in a very real sense, the definition of a function which we are introducing here is a definition of what we normally call the graph of a function. We normally think that if you have a function then you can plot the graph. But equally, if you have the graph then you know what the function is. Which comes first? The chicken or the egg?*

²A late 19th century analysis book (Forsyth) gives a sound definition of $z^{\sqrt{2}}$, but feels obliged to comment ‘which is scarcely a function in the usual sense’. Mathematicians have become more open-minded and are now prepared to speak openly about things their predecessors thought in bad taste.

3.3 Other Names

Instead of ‘*function*’ we may use the following:

mapping

map.

These words mean no more and no less. Mathematically a map is a function (as defined above), and so too is a mapping.

3.4 Important Notation

From now on we write $f : A \rightarrow B$ (which we pronounce as ‘*f is a function from A to B*’) to mean that we have a function whose domain is A , codomain B and rule named f .

As there is a unique element paired with each $a \in A$ by the rule f it is convenient to give it a name that reminds us of its role: so we write it $f(a)$, and call it ‘*the value of f at a*’.

Often the rule we use is determined by a formula. In such cases we will write things like ‘let $f : \mathbb{R} \rightarrow \mathbb{R}$ be given by $x \mapsto (x^2 + 23)$ ’. We pronounce the last phrase ‘*x maps to (x² + 23)*’.

Sometimes we also write things like ‘ f is a function with domain A and codomain B such that $f : x \mapsto f(x)$ ’.

Example. *What is the largest subset of \mathbb{R} on which the rule $x \mapsto \frac{x+1}{x-1} \in \mathbb{R}$ defines a function?*

The answer is $\mathbb{R} \setminus \{1\}$. As we will see next week, $\frac{a+1}{a-1}$ is a (unique) real number whenever $a \in \mathbb{R} \setminus \{1\}$. For codomain we can take \mathbb{R} .

3.5 Some Famous Functions

3.5.1 Factorials

We define the factorial function to be the function whose domain is \mathbb{N} , codomain \mathbb{N} , such that $0 \mapsto 1$ and $n \mapsto n(n-1)(n-2) \dots 2 \cdot 1$ otherwise.

The notation for this function is $!$, but we don’t write $!(n)$, we write $n!$. (There are other cases where we can’t rewrite history, and just have to accept the notation in common use.)

Example. $0! = 1, 1! = 1, \dots, 5! = 120, 6! = 720$.

(We all know the first six by heart.)

3.5.2 Binomials

We define the binomial function to be the function whose domain is $\mathbb{N} \times \mathbb{N}$ and codomain \mathbb{N} , and such that

$$(n, k) \mapsto \begin{cases} \frac{n!}{k!(n-k)!} & \text{if } k \leq n; \\ 0 & \text{otherwise.} \end{cases}$$

The notation for this function is even more idiosyncratic: we write $\binom{n}{k}$ for its value at the point (n, k) .

Example. $\binom{4}{2} = 6$, $\binom{5}{3} = 10$, \dots

Example. $\binom{n}{2} = \frac{n(n-1)}{2}$ provided $n \geq 2$.

3.5.3 Powers

By way of example let us define natural powers of real numbers.

For domain let us take $\mathbb{R} \times \mathbb{N}$, and for codomain \mathbb{R} . Then we define the function by

$$(x, n) \mapsto \begin{cases} x \cdot x \cdots x & (n \text{ times}) & \text{if } n \neq 0; \\ 1 & & \text{if } n = 0. \end{cases}$$

Our notation for this function is: x^n .

We could define similar functions replacing \mathbb{R} by any of \mathbb{N} , \mathbb{Z} , \mathbb{Q} , \mathbb{C} ; indeed by any set where we have a multiplication defined. We call them all x^n ; this is ambiguous but leads to no confusion.

Note. Yes, I did intend to define $0^0 = 1$.

3.6 Composition of Functions

Now let $f : A \rightarrow B$ and $g : B \rightarrow C$ be functions. We define a function $g \circ f : A \rightarrow C$ by specifying the rule

$$g \circ f := \{(a, c) \mid a \in A, c \in C \text{ and } \exists b \in B \text{ such that } (a, b) \in f \text{ and } (b, c) \in g\}.$$

We call this ‘the composition of f and g ’.

Note that $(g \circ f)(a) = g(f(a))$.

Example. Let $f : \mathbb{N} \rightarrow \mathbb{N}$ be given by $n \mapsto n + 1$; and let $g : \mathbb{N} \rightarrow \mathbb{N}$ be given by $n \mapsto (n - 1)$ when $n > 0$ and $0 \mapsto 0$. Then $g \circ f : n \mapsto n$.

(What about $f \circ g$?)

3.7 Restrictions of Functions

Suppose that $f : A \rightarrow B$ is a function, and that $X \subseteq A$. Then it is easy to see that the rule $f \cap X \times B$ defines a function with domain X and codomain B . We call this ‘the restriction of f to X ’, and write it $f|_X : X \rightarrow B$.

There is no corresponding notation for what happens if we only look at subsets of the codomain.

Example. Let $\phi : \mathbb{C} \setminus \{0\} \rightarrow \mathbb{C}$ be given by $z \mapsto \frac{1}{z}$. Let $T := \{z \in \mathbb{C} : |z| = 1\}$. Then $\phi|_T : z \mapsto \frac{1}{z}$ for all $z \in T$.

3.8 Injective Functions

Let $f : A \rightarrow B$ be a function. We say that ' f is injective' under these circumstances:

$$\forall a_1 \in A, \forall a_2 \in A \quad f(a_1) = f(a_2) \implies a_1 = a_2$$

We may also say that f is 1—1; it means the same thing.

Example. The mapping $s : \mathbb{R} \rightarrow \mathbb{R}$ given by $s : t \mapsto t^2$ is not 1—1.

Proof. $s(1) = 1^2 = 1$ and $s(-1) = (-1)^2 = 1$, yet $1 \neq -1$. □

Note. If you have any doubts about the truth of any of these statements, don't worry: we will prove them from some simple assumptions next week. But do notice this: when we prove something in this way, by giving a counterexample, we must give an explicit counterexample. It's no good saying things like 'Well $x^2 = (-x)^2$ ' and leaving your audience to find such an x for themselves. You must do all the work.

3.9 Surjective Functions

Let $f : A \rightarrow B$ be a function. We say that ' f is surjective' under these circumstances:

$$\forall b \in B \quad \exists a \in A \text{ such that } b = f(a)$$

We may also say that ' f is onto'; it means the same thing.

Example. The mapping $s : \mathbb{R} \rightarrow \{y \in \mathbb{R} \mid y \geq 0\}$ is onto.

Intuitively this is clear if we draw a picture of the set $\mathbb{R} \times \{y \in \mathbb{R} \mid y \geq 0\}$ and mark on the picture the set s (that is, the set of pairs in s). But to turn that picture into a proof needs some very careful analysis later in the term.

3.10 Bijective Functions

A function $f : A \rightarrow B$ is called *bijective* if it is both injective and surjective.

Example. The mapping $r : \mathbb{Z} \rightarrow \mathbb{Z}$ given by $x \mapsto (-x)$ is bijective.

3.10.1 More jargon

A bijective map is sometimes called a *bijection*; likewise an injective map is called an *injection*, a surjective map a *surjection*. Don't let the multiplicity of names confuse you.

3.11 Identity Maps

Let A be a set; then the rule $x \mapsto x$ defines a function with domain A and codomain A . We denote this map by $I_A : A \rightarrow A$ and call it the identity map on A .

Each I_A is clearly bijective.

3.12 Invertible Functions

Let $f : A \rightarrow B$ be a function. We say that ‘ f is invertible’ if there exists a function $g : B \rightarrow A$ such that $f \circ g = I_B$ and $g \circ f = I_A$. We then call g an *inverse* of f .

It is trivial to see that if $f : A \rightarrow B$ has an inverse $g : B \rightarrow A$, then $g : B \rightarrow A$ has an inverse $f : A \rightarrow B$.

Theorem. *Suppose that $f : A \rightarrow B$ is invertible. Then the inverse of f is unique.*

Proof. Suppose that both g_1 and g_2 are inverses, and that $b \in B$. Then

$$g_1(b) = g_1(I_B(b)) = g_1(f \circ g_2)(b) = (g_1 \circ f)(g_2(b)) = I_A(g_2(b)) = g_2(b),$$

so that $g_1 = g_2$. □

Theorem. *Let $f : A \rightarrow B$ be a function. Then f is invertible if and only if f is a bijection.*

Proof. Suppose first that f is invertible with inverse g . To see that f is injective note that

$$f(a_1) = f(a_2) \implies g(f(a_1)) = g(f(a_2)) \implies I_A(a_1) = I_A(a_2) \implies a_1 = a_2.$$

To see that f is surjective, let $b \in B$; then $b = I_B(b) = f(g(b))$.

Now suppose that f is a bijection. To construct a function $g : B \rightarrow A$ we work with our ‘ordered pair’ definition. That is, we define

$$g := \{(x, y) \in B \times A \mid (y, x) \in f\}.$$

We then need positive answers to the questions: is g a function? Does $f \circ g = I_B$ and $g \circ f = I_A$?

To answer the first, we have certainly got the right domain and codomain, it is the two parts (a) and (b) of the part (iii) of the definition that need checked. For (a) we need to make sure that given any $b \in B$ there is some $a \in A$ with $(b, a) \in g$, that is with $(a, b) \in f$. This is precisely what ‘ f is surjective’ guarantees. For (b) we need to make sure that $a = a'$ whenever we have $b \in B$ and $a, a' \in A$ with $(b, a), (b, a') \in g$. Well, this happens exactly when $(a, b), (a', b) \in f$. The statement ‘ f is injective’ guarantees that $a = a'$.

It is now an easy exercise³ to prove that g is the inverse of f . □

3.12.1 Notation for Inverses

If $f : A \rightarrow B$ has an inverse $g : B \rightarrow A$ it is very tempting to write f^{-1} for g . This can mislead. As we’ll see f^{-1} has a meaning in general, and even when f is invertible, the general meaning is not ‘the inverse of f ’. This is why using ‘ $\arctan y$ ’ and not ‘ $\tan^{-1} y$ ’ is a good idea.

³This is code for: write out the proof yourself.

3.13 Images and Pre-images

Let $f : A \rightarrow B$ be a function, and let $X \subseteq A$, $Y \subseteq B$.

We define the set

$$f(X) := \{b \in B : \exists a \in X \text{ such that } f(a) = b\}$$

which we call ‘*the image of X under f* ’.

We also define the set

$$f^{-1}(Y) := \{a \in A : f(a) \in Y\}$$

which we call ‘*the pre-image of Y under f* ’.

Example. Suppose that $s : \mathbb{R} \rightarrow \mathbb{R}$ is given by $x \mapsto x^2$. Let $X := (-1, 2)$ and $Y_1 := [-3, 4]$, $Y_2 := [-100, -1]$. (Here we’re using the notation for intervals.)

Then $s(X) = \{x \in \mathbb{R} : 0 \leq x < 4\}$ and $s^{-1}(Y_1) = [-2, 2]$, $s^{-1}(Y_2) = \emptyset$.

3.13.1 The Range of a Function

Suppose that $f : A \rightarrow B$ is a function. The ‘*range of f* ’ is the set $f(A)$.

Confusingly, in some parts of the subject this is called the *image of f* .

3.13.2 Notational Nightmares

All this is fine if we use notation that distinguishes clearly between elements and subsets of a set. But real life is more complex; sometimes we do have some X which is both an element and a subset of A . How do we then know what $f(X)$ means? If there is such ambiguity we need to spell out which we mean.

Example. Let $A := \{\emptyset\}$ and $B := \{1\}$; define the function $i : A \rightarrow B$ by $\emptyset \mapsto 1$. What is $i(\emptyset)$? If we mean the element \emptyset then it is $\{1\}$; if we mean the subset \emptyset then it is \emptyset .

4 Induction

In this section we discuss mathematical induction, the key property of natural numbers. We base this on an assumption about subsets of natural numbers.

4.1 Well-ordering of \mathbb{N}

We take the natural numbers as a given, but we single out one of its key properties, the fact that it is ‘well-ordered’. By this we mean that the following holds:

WOP If $X \subseteq \mathbb{N}$, and $X \neq \emptyset$, then there exists some $x_0 \in X$ such that:

(E) for all $x \in X$, $x_0 \leq x$;

(U) if $b \in X$ is such that for all $x \in X$ we have $b \leq x$ then $b = x_0$

Less formally, we are asserting that every non-empty subset of \mathbb{N} has a unique least element.

Note. From WOP we can prove that for each pair of natural numbers $m, n \in \mathbb{N}$ we have either $m \leq n$ or $n \leq m$. Think about the subset $\{m, n\}$.

4.2 Least Criminals

Suppose that we want to prove that some property $\mathcal{P}(n)$ holds for all natural numbers $n \in \mathbb{N}$.⁴

One method of proof is to suppose that the result is false, and to look at the non-empty set $X := \{n \in \mathbb{N} \mid \mathcal{P}(n) \text{ is false}\}$. This is the set of ‘criminals’. Use the WOP to pick x_0 , the ‘least criminal’.

Now, in the specific case we have to argue to a contradiction.

Exercise. Every natural number greater than 1 is a product of primes.

Proof. Suppose false and let n be the least criminal. Then certainly n is not prime, so that $n = k\ell$, with $k < n$ and $\ell < n$. As n is the least criminal both k and ℓ are products of primes; and the same is then true about $n = k\ell$, a contradiction. \square

4.3 Mathematical Induction

We can use the WOP in a very specific way.

Theorem (The Principle of Mathematical Induction). *Suppose we have a property $\mathcal{P}(n)$ of natural numbers, satisfying*

(The root, or base step) $\mathcal{P}(0)$ is true;

(The inductive step) for every $k \in \mathbb{N}$, $\mathcal{P}(k) \implies \mathcal{P}(k + 1)$ is true.

⁴In passing, note that we can express some very deep theorems in this form:

$$n \leq 2 \quad \text{or} \quad \forall x \in \mathbb{N} \quad \forall y \in \mathbb{N} \quad \forall z \in \mathbb{N} \quad (x^n + y^n = z^n \implies x, y, z \text{ are not coprime})$$

for instance.

Then for every $n \in \mathbb{N}$, the property $\mathcal{P}(n)$ is true

Proof. If this theorem is not true then $\{n \in \mathbb{N} | \mathcal{P}(n) \text{ is false}\}$ is non-empty and so has a least element m . Clearly, by the assumption ‘root’, $m \neq 0$; so $m = k + 1$ for some $k \in \mathbb{N}$. As $k < m$ we have that $\mathcal{P}(k)$ is true. By assumption ‘inductive step’ we have that $\mathcal{P}(k) \implies \mathcal{P}(k + 1)$ is true. Hence $\mathcal{P}(m)$, that is $\mathcal{P}(k + 1)$, is true: a contradiction. \square

Example. For all $n \in \mathbb{N}$, $\sum_{k=0}^{k=n} k = \binom{n+1}{2}$.

Proof. The **root** is easy: $\sum_{k=0}^{k=0} k = 0$, and also $\binom{0+1}{2} = 0$ as $1 < 2$.

For the **inductive step**, we assume that $\sum_{k=0}^{k=n} k = \binom{n+1}{2}$. Then

$$\sum_{k=0}^{k=n+1} k = \sum_{k=0}^{k=n} k + (n+1) = \binom{n+1}{2} + (n+1) = \frac{(n+1)n}{2} + (n+1) = \dots = \frac{(n+2)(n+1)}{2}$$

and so we have proved the result holds for $(n + 1)$. So we are done by the Principle of Mathematical Induction. \square

4.3.1 Strong Induction

Theorem (Strong Induction). Suppose we have a property $\mathcal{P}(n)$ of natural numbers, satisfying

(The inductive jump) For all natural numbers k , $\mathcal{P}(k)$ is true provided $\mathcal{P}(\ell)$ is true for every $\ell < k$.

Then for every $n \in \mathbb{N}$, the property $\mathcal{P}(n)$ is true.

Proof. The proof is left as an easy exercise. \square

Note. Note that the omission of the requirement to prove $\mathcal{P}(0)$ is only apparent — the inductive jump, with $k = 0$, requires us to prove $\mathcal{P}(0)$ with no assumptions at all to help us.

4.4 An Application

Theorem (The Binomial Theorem).

$$\text{For every natural number } n, \quad (1 + x)^n = \sum_{k=0}^{k=n} \binom{n}{k} x^k.$$

Proof. We will argue by Mathematical Induction on n .

For the **root** things are easy: $(1 + x)^0 = 1$ by definition, and $\sum_{k=0}^{k=0} \binom{0}{k} x^k = \binom{0}{0} x^0 = \frac{1}{1-1} \cdot 1 = 1$ also.

For the **inductive step** we assume that the result is true for n , and we must prove it is true for $(n + 1)$.

$$\begin{aligned}
(1 + x)^{n+1} &= (1 + x)(1 + x)^n \quad (\text{definition of powers}) \\
&= (1 + x) \sum_{k=0}^{k=n} \binom{n}{k} x^k \quad (\text{inductive hypothesis}) \\
&= \sum_{k=0}^{k=n} \binom{n}{k} x^k + \sum_{k=0}^{k=n} \binom{n}{k} x^{k+1} \quad (\text{rules of algebra}) \\
&= \binom{n}{0} x^0 + \sum_{k=1}^{k=n} \binom{n}{k} x^k + \sum_{k=0}^{k=n-1} \binom{n}{k} x^{k+1} + \binom{n}{n} x^{n+1} \\
&= x^0 + \sum_{k=1}^{k=n} \left(\binom{n}{k} + \binom{n}{k-1} \right) x^k + x^{n+1} \\
&= \binom{n+1}{0} x^0 + \sum_{k=1}^{k=n} \binom{n+1}{k} x^k + \binom{n+1}{n+1} x^{n+1} \quad (\text{see below!}) \\
&= \sum_{k=0}^{k=n+1} \binom{n+1}{k} x^k \quad \text{as required.}
\end{aligned}$$

We see that to complete the result we need a lemma⁵ which we will prove in a moment. Before we do, note that our proof is complete by the Principle of Mathematical Induction. □

Lemma (Pascal's Triangle).

$$\text{For every pair } n, k \text{ of natural numbers, } \binom{n+1}{k+1} = \binom{n}{k+1} + \binom{n}{k}.$$

Proof. If $k > n$ then all the terms are 0 and we are done. If $k = n$ then $\binom{n}{k+1} = 0$ and the other two terms are equal to 1 and we are done. Otherwise this is pure calculation:

$$\begin{aligned}
\binom{n}{k+1} + \binom{n}{k} &= \frac{n!}{(k+1)!(n-k-1)!} + \frac{n!}{k!(n-k)!} \\
&= \frac{n!}{k!(n-k-1)!} \left(\frac{1}{k+1} + \frac{1}{n-k} \right) \\
&= \frac{n!}{k!(n-k-1)!} \cdot \frac{n+1}{(k+1)(n-k)} \\
&= \frac{(n+1)!}{(k+1)!(n-k)!} \\
&= \binom{n+1}{k+1}.
\end{aligned}$$

□

⁵A mini-theorem used as a stepping stone in the proof of a theorem. But some lemmas don't deserve their poor-relation status; this is one such.

5 How big are sets?

In this section we introduce some of the simplest ideas about the ‘size’ of a set.

5.1 Equinumerous sets

Let A and B be sets. We will say ‘ A and B are equinumerous’, and will write $A \approx B$, if there is a bijection $\theta : A \rightarrow B$.

There are — but we can only hint at the full story in this section — some typical sets, with which we compare other sets in this way. These are called the cardinals.

Note that

$$A \approx B \iff B \approx A;$$

this is just an easy application of our results on inverses and bijections; check it out.

5.2 Finite Sets

A set A is called finite if either $A = \emptyset$, or for some non-zero natural number k we have that $A \approx \{0, 1, \dots, k-1\}$. In the former case we say that A has size 0, in the latter size k . We may denote the size of A by $|A|$, or $\sharp A$.

Note. *It would be the same thing (wouldn't it?) to use the set $\{1, 2, \dots, k\}$ as the typical set of size k . Use whichever is convenient.*

5.2.1 A small doubt

Could a set A be of size k and of size ℓ , with $k \neq \ell$? The answer is ‘certainly not’.

We sketch a ‘least criminal’ proof. Consider the smallest k for which this happens. We can then construct a bijection φ from $\{0, 1, \dots, k\}$ to $\{0, 1, \dots, \ell\}$. Remove k and $\varphi(k)$, and (with some adjustment) we’ll get a set of size $k-1$ and also size $\ell-1$. This contradicts our definition of k .

5.3 Infinite sets

A set which is not finite is called ‘infinite’.

Example. *The sets \mathbb{N} , \mathbb{Z} , \mathbb{Q} , \mathbb{R} , and \mathbb{C} are all infinite.*

We are going to see that they are not all equinumerous.

5.4 Countably infinite sets

A set A is called ‘countably infinite’ if $\mathbb{N} \approx A$.

Example. *The set of even natural numbers is countably infinite.*

Proof. The map $n \mapsto 2n$ is the required bijection. □

5.4.1 Removing an Element

Theorem. *Suppose A is countably infinite, and $a \in A$. Then $A \setminus \{a\}$ is countably infinite.*

Proof. Suppose that we have a bijection $\theta : \mathbb{N} \rightarrow A$ with $\theta(k) = a$. Then

$$\psi : \mathbb{N} \rightarrow A \setminus \{a\}, \text{ where } \psi : n \mapsto \begin{cases} \theta(n) & n < k \\ \theta(n+1) & n \geq k \end{cases}$$

is clearly the bijection we need. □

Corollary. *The set $\mathbb{N} \setminus \{0\}$ is countably infinite.*

(This should reassure those who prefer to exclude 0 from its place among the natural numbers.)

5.4.2 Interlacing

Theorem. *Suppose that A_1 and A_2 are countably infinite, and that $A_1 \cap A_2 = \emptyset$. Then $A_1 \cup A_2$ is countably infinite.*

Proof. Suppose we have bijections $\theta_i : \mathbb{N} \rightarrow A_i$ for $i = 1, 2$. Then

$$\theta : \mathbb{N} \rightarrow A_1 \cup A_2, \text{ where } \theta : n \mapsto \begin{cases} \theta_1(\frac{n}{2}) & n \text{ even} \\ \theta_2(\frac{n-1}{2}) & n \text{ odd} \end{cases}$$

is clearly the bijection we need. □

Corollary. *The set \mathbb{Z} of integers is countably infinite.*

Proof. We write $\mathbb{Z} = \mathbb{N} \cup \{-(n+1) : n \in \mathbb{N}\}$, both of which are countably infinite. □

5.4.3 Products

Theorem. *Suppose that A_1 and A_2 are countably infinite. Then $A_1 \times A_2$ is countably infinite.*

Proof. Suppose we have bijections $\varphi_i : A_i \rightarrow \mathbb{N}$ for $i = 1, 2$. Suppose that we can deal with the case when $A_1 = A_2 = \mathbb{N}$, that is suppose we have a bijection $q : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$. Then the map $\theta : A_1 \times A_2 \rightarrow N$ by $(x_1, x_2) \mapsto q(\varphi_1(x_1), \varphi_2(x_2))$ is a bijection.

So let us construct the map q . To do this think of the pairs of natural numbers plotted in the usual way in the plane.

[picture]

On the line $x + y = k$ — let's call it L_k — there are $k + 1$ such 'lattice points'. Our plan is to deal with points on the lines L_0 , then L_1 , then L_2 , and so on; and on each line to count off the points from $(0, k)$ down to $(k, 0)$.

So on the lines L_0, \dots, L_{k-1} we have already counted $\frac{k(k+1)}{2}$ points; and we see that on the line L_k the point (a, b) (where of course $a + b = k$) is the $a + 1$ -th.

So to get a bijection from $\mathbb{N} \times \mathbb{N}$ to \mathbb{N} we can use the map

$$(m, n) \mapsto \frac{(m+n)(m+n+1)}{2} + m.$$

□

5.4.4 More jargon

We say that a set is ‘countable’ when it is either finite or countably infinite. Often, when the set is clearly not finite we omit any proof of that fact.

We sometimes say that a countably infinite set ‘has cardinality \aleph_0 ’.

5.5 Two equivalent definitions

Theorem. *An infinite set A is countably infinite if and only if there is an injection $j : A \rightarrow \mathbb{N}$.*

Proof. The ‘only if’ direction is trivial. So suppose there is an injection $j : A \rightarrow \mathbb{N}$. We can think of this as giving an order on our set A . We manufacture a bijection $\theta : A \rightarrow \mathbb{N}$ by associating with each $a \in A$ the number of $x \in A$ which come before a in the order. That is, we define

$$\theta(a) := |\{x \in A \mid j(x) < j(a)\}|;$$

we don’t carry out the tedious checks that this is a function, is 1–1, and is onto. □

Example. *The set $\mathbb{N} \times \mathbb{N}$ is (as we’ve seen) countable, as we have an injection $(m, n) \mapsto 2^m 3^n$.*

Theorem. *An infinite set A is countably infinite if and only if there is a surjection $\pi : \mathbb{N} \rightarrow A$.*

Proof. The ‘only if’ direction is trivial. So suppose there is a surjection $\pi : \mathbb{N} \rightarrow A$. Although every element $a \in A$ is associated with a natural number (‘onto’), it may be associated with many. But we can pick the least of these. That is, we can define a map $j : A \rightarrow \mathbb{N}$ by

$$j(a) := \min \pi^{-1}(a);$$

again we omit the tedious details of the proof that j is a function, and that it is 1–1. The result now follows by the previous theorem. □

Corollary. *The set \mathbb{Q} of rational numbers is countable.*

Proof. As $\mathbb{Q} = \{x \in \mathbb{Q} \mid x > 0\} \cup \{0\} \cup \{x \in \mathbb{Q} \mid x < 0\}$ it is not hard to see that it is enough to prove that $\mathbb{Q}^{>0} := \{x \in \mathbb{Q} \mid x > 0\}$ is countable.

We have a bijection $p : \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$ (the inverse of the map $q : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ we defined before), so if we give a surjection $r : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{Q}^{>0}$ then $r \circ p$ will surject \mathbb{N} on $\mathbb{Q}^{>0}$ and we will have the result.

Define $r(m, n) := \frac{m+1}{n+1}$; this is the onto map we need. □

5.6 Uncountable sets

A set which is not countable is called ‘*uncountable*’.

Once we know what we mean by the real numbers we will prove:

Theorem. *The set of real numbers, \mathbb{R} , is uncountable.*

If a set A is equinumerous with \mathbb{R} we may say that ‘ A has the same cardinality as \mathbb{R} ’. (What that cardinality *is* depends what your basic mathematical assumptions are.)

5.6.1 Bigger and Bigger Sets

The following result due to Cantor proves that we don’t ever run out of sets which are not equinumerous to ones we’ve met already.

Theorem. *Let A be a set, and let \mathcal{A} be the set of subsets of A . Then these sets are not equinumerous.*

Proof. Suppose we have a map $\pi : A \rightarrow \mathcal{A}$. We show that π cannot be a surjection. This shows that there are no bijections from A to \mathcal{A} , so that the sets are not equinumerous.

Form the set

$$X := \{a \in A \mid a \notin \pi(a)\}.$$

We show that $X \notin \pi(A)$.

Suppose to the contrary that $X = \pi(x)$ for some $x \in A$. Then ask yourself the question ‘is $x \in X$ or not?’. From the definition of X we know that $x \in X$ if and only if $x \notin \pi(x)$. This is contradictory if $X = \pi(x)$, and the only way to escape the contradiction is to abandon our assumption that $X = \pi(x)$. So $X \notin \pi(A)$, as claimed. \square

The Greek Alphabet

Name	Upper Case	Lower Case
Alpha	A	α
Beta	B	β
Gamma	Γ	γ
Delta	Δ	δ
Epsilon	E	ε
Zeta	Z	ζ
Eta	H	η
Theta	Θ	θ
Iota	I	ι
Kappa	K	κ
Lambda	Λ	λ
Mu	M	μ
Nu	N	ν
Xi	Ξ	ξ
Omicron	O	\omicron
Pi	Π	π
Rho	P	ρ
Sigma	Σ	σ
Tau	T	τ
Upsilon	Υ	υ
Phi	Φ	φ
Chi	X	χ
Psi	Ψ	ψ
Omega	Ω	ω

And here, out of place, is one Hebrew letter: aleph, \aleph .
 (An infinite countable set is said 'to have cardinality \aleph_0 '.)