

(New) CORRECTIONS:

p 98
p 77

MATH 318/Fall, 2003

Mathematical Logic

PROF. M. MAKKA

p. 311

Table of contents

2006 Jan: 48: (4p diag)
64: l. -6

Chapter 1 Sets and functions

Section 1.1	Sets	p. 1
Section 1.2	Subsets and the Boolean operations on sets	p. 7
Section 1.3	Ordered pairs and functions	p. 15

Chapter 2 Binary relations

Section 2.1	Kinds of binary relation	p. 28
Section 2.2	Equivalence relations	p. 40
Section 2.3	Operations on binary relations	p. 47

Chapter 3 Orders and lattices

Section 3.1	Orders	p. 60
Section 3.2	Lattices	p. 74

Chapter 4 Boolean algebras and propositional logic

Section 4.1	Boolean algebras	p. 92
Section 4.2	Generating Boolean subalgebras	p. 105
Section 4.3	Boolean functions	p. 122

Chapter 5 Predicate logic

Section 5.1	Quantifiers	p. 136
Section 5.2	Formal specification of computer programs	p. 151
Section 5.3	Entailment in predicate logic	p. 163

Chapter 6 The mathematics of the natural numbers

Section 6.1	The system of the natural numbers	p. 174
Section 6.2	Divisibility among the integers	p. 190
Section 6.3	Counting	p. 201

Chapter 1 Sets and functions

Section 1.1 Sets

The concept of set is a very basic one. It is simple; yet, it suffices as the basis on which all abstract notions in mathematics can be built.

A set is *determined by its elements*.

If A is a set, we write $x \in A$ to say that x is an *element of* A . Other readings for " $x \in A$ " are: " x belongs to the set A ", " x is in A ".

Anything may be an element of a set; any two, possibly unrelated, things may be elements of the same set. In fact, any way of collecting things into a whole results in a set; the things collected are the *elements* of the set.

To say that a set is *determined* by its elements is to say that any set is completely given by specifying what its elements are. We may express this in the following mathematical style:

Principle of extensionality. Two sets A and B are equal if for all things x , x is an element of A if and only if x is an element of B .

We will frequently use the following logical abbreviations:

$\forall x$: "for all x ...".

$\exists x$: "there is x such that ..."

\longleftrightarrow : "if and only if"

\implies : "if ..., then ..."

$\&$: "... and..."

Principle of Extensionality, in abbreviated form: For sets A and B ,

$$\forall x (x \in A \leftrightarrow x \in B) \implies A=B.$$

A set may be given by *listing* its elements. A list enclosed in curly brackets denotes the set whose elements are the things in the list. E.g.,

$$\{0, 2, 142, 96, 3\} \tag{1}$$

denotes the set whose elements are the five integers listed. The order in which the elements are listed is immaterial. Thus,

$$\{2, 142, 0, 3, 96\}$$

is a notation for the same set as (1). Also, if a list contains repetitions, when enclosed in curly brackets, it will denote the same set as the list with the repetitions removed. E.g.,

$$\{2, 142, 0, 0, 3, 96, 3\}$$

still denotes the same set as the previous two notations.

There are sets that cannot be listed; they are *infinite*. E.g., the set of all non-negative integers, or natural numbers, denoted by \mathbb{N} , is such an infinite set. We may write

$$\mathbb{N} = \{0, 1, 2, 3, \dots\},$$

but this is a very incomplete notation! Other important sets of numbers are as follows:

$$\mathbb{Z} = \{0, 1, -1, 2, -2, \dots\},$$

the set of all (positive, negative and zero) integers;

$$\mathbb{Q} = \text{the set of all rational numbers}$$

(a number is *rational* if it is of the form $\frac{p}{q}$ for some integers p and q ($q \neq 0$));

\mathbb{R} = the set of all real numbers

and

\mathbb{C} = the set of all complex numbers

($\sqrt{2}$, π and e are real numbers, but they are not rational; $i = \sqrt{-1}$ is a complex number which is not real).

A set may be specified by giving a property or condition that its elements, and only its elements, have or satisfy; the elements of the set are exactly the things that have the property (satisfy the condition) in question. In fact, the five number-sets just introduced are given in this way. E.g., any thing x is a member of \mathbb{Q} if and only if it has the property that there are integers p and q , $q \neq 0$, such that $x = \frac{p}{q}$. By using arbitrary conditions, we may define an endless variety of sets.

The curly brackets are also used for specifying a set by a condition. E.g.,

$$\mathbb{Q} = \left\{ \frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0 \right\},$$

$$\mathbb{C} = \{x + y\sqrt{-1} : x, y \in \mathbb{R}\}.$$

In each of these formulas, in front of the colon (:), one finds an expression denoting a quantity depending on certain variables; in the first case, these variables are p and q , in the second x and y . The complete symbol denotes the set of all values of the expression when the variables range over all values satisfying the condition stated after the vertical line. Thus, one should read the first formula as "the set of all $\frac{p}{q}$ such that p and q are integers and $q \neq 0$ ".

A set given in the form of a list may be specifiable more conveniently by a condition. E.g., the set

$\{0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32\}$

is the same as the set

$$\{n \in \mathbb{N} \mid n \text{ is even and } n < 34\}.$$

This latter symbolic expression should be read as:

"the set of all n in \mathbb{N} such that n is even and $n < 34$ ".

Note the use of the symbol " \in " in this set-notation.

Differently worded conditions may give rise to the same set. E.g., the following two brackets define the same set:

$$\{n \in \mathbb{N} \mid n \text{ is a prime number and } 11 \leq n \leq 120\} \quad (2)$$

$$\{n \in \mathbb{N} \mid n \text{ is not divisible by } 2, 3, 5 \text{ and } 7, \text{ and } 1 < n \leq 120\} \quad (3)$$

(can you prove this?).

It is a fundamental point that sets themselves may be elements of other sets. E.g., the set

$$A = \{\{1, 2\}, \{4, 7\}\} \quad (4)$$

has two elements, $\{1, 2\}$ and $\{4, 7\}$, both of which are sets themselves. In axiomatic set theory, mathematical objects are constructed as sets of sets of sets ... with "arbitrary" complexity.

Note that the set

$$B = \{1, 2, 4, 7\}$$

is something very different from A in (4); whereas the elements of A are sets, the elements of B are numbers, not sets. (Unlike in the usual versions of axiomatic set theory, we do not identify numbers with sets. Foundationally, our set theory has "urelements", non-sets with no

internal structure; natural numbers are such "ur-elements"). Also note that the set $\{\{7, 4\}, \{2, 1\}\}$ is equal to A , but the set $\{\{1, 4\}, \{2, 7\}\}$ is not (why?).

One specific set, the *empty set*, has to be pointed out. The empty set is the set which has no elements; it is given by any condition that is contradictory, that is, a condition that has nothing to satisfy it. E.g., the set

$$\{n \in \mathbb{N} \mid n^2 + 1 \text{ is divisible by } 4\}$$

is empty, since there is no natural number n for which $n^2 + 1$ is divisible by 4. (If n is even, $n = 2k$, then $n^2 + 1 = 4k^2 + 1$, which is not divisible by 4 (gives the remainder 1 when divided by 4); if n is odd, $n = 2k + 1$, then $n^2 + 1 = 4(k^2 + k) + 2$, which is not divisible by 4 either.)

There is just one empty set; if both A and B are empty, then for any x , $x \in A$ just in case $x \in B$, namely never; hence, $A = B$ by the principle of extensionality. The symbol for the empty set is \emptyset .

Let us state the general principle behind the curly-bracket notation.

Principle of Comprehension

The set

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

that is, the set whose elements are those, and only those, x that have property $P(x)$, *exists*.

Sadly, there are some exceptions to the validity of the principle of comprehension. Notably, the *vacuous* condition $P(x)$ that is identically true (which can be represented by the expression $x=x$, since everything is equal to itself) cannot be used in the principle. It would give rise to the set $\{x: x=x\}$, that is, the set of all things; and the set of all things does not exist; it gives rise to the famous paradoxes (contradictions) of set theory.

Unfortunately, it is quite difficult to describe precisely the extent of the validity of the principle of comprehension; this is done in the discipline called *axiomatic set theory*. For us, it

should suffice to say that *normally*, the principle of comprehension is valid; the exceptions are rare.

Note the following consequences of the meaning of the term $\{x: P(x)\}$:

If $P(x)$, then $x \in \{x: P(x)\}$;
 If not $P(x)$, then $x \notin \{x: P(x)\}$;
 $x \in \{x: P(x)\}$ if and only if $P(x)$.

We use the logical abbreviation \neg for "not"; we read $\neg P(x)$ as: "it is not the case that $P(x)$ "; briefly: $\text{not}(P(x))$.

Using logical abbreviations, we have:

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

$$\neg P(x) \implies x \notin \{x: P(x)\}$$

$$x \in \{x: P(x)\} \iff P(x) .$$

Variants of the comprehension notation can be explained thus:

$$\{x \in X: P(x)\} \stackrel{\text{def}}{=} \{x: x \in X \ \& \ P(x)\}$$

$$\{f(x): P(x)\} \stackrel{\text{def}}{=} \{y: \exists x. y=f(x) \ \& \ P(x) .\} .$$

Section 1.2 Subsets and the Boolean operations on sets

If every element of the set A is an element of the set B , we say that A is a *subset of* B , or that A is *contained in* B , or that B *contains* A , and we write $A \subseteq B$, or $B \supseteq A$.

Any set is a subset of itself: $A \subseteq A$. When we want to say that A is a subset of B and A is different from B , we say that A is a *proper subset of* B , and we write $A \subset B$.

For instance, $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$; the five basic number-sets are increasingly more comprehensive sets. $\mathbb{Z} \subset \mathbb{Q}$ because every integer is a rational number ($n = \frac{n}{1}$), but *not* ($\mathbb{Q} \subseteq \mathbb{Z}$), since $\frac{1}{2}$ is a rational number but not an integer.

Note that for any set A , $\emptyset \subseteq A$: any thing that belongs to \emptyset (there is nothing like that) belongs to A as well.

The following general law, the

Antisymmetry law for \subseteq :

$$A \subseteq B \ \& \ B \subseteq A \implies A = B \tag{1}$$

(here A and B are sets) is equivalent to the principle of extensionality (see the first section), namely the principle that says that sets are determined by their elements. In fact, the left-hand side of the implication in (1) says that all elements of A are elements of B and vice versa, which is to say that A and B have the same elements.

The antisymmetry law serves as the formal setup for showing that two sets given by conditions are the same (if that is the case). One shows two things: **one**, that the first set is contained in the other, **two**, that the other is contained in the first. The reader should try this on the example of the sets given in (2) and (3) in Section 1.1. Later in this section, we will see another example.

An obvious law for \subseteq is the law of

Transitivity for \subseteq :

$$A \subseteq B \text{ and } B \subseteq C \implies A \subseteq C .$$

Although this is completely obvious, it serves as the basis for an important generalization, in the notion of ordering, considered in the next chapter.

Note that the relations "being an element of", denoted by \in , and "being a subset of", denoted by \subseteq , are very different. E.g., $2 \subseteq \{2, 7\}$ does not hold; it does not even make sense, since 2 is not a set. Of course, $2 \in \{2, 7\}$ does hold. Also, $\{2\} \subseteq \{2, 7\}$ holds, and this is the same as $2 \in \{2, 7\}$. On the other hand, $\{2\} \in \{2, 7\}$ *does not* hold; in fact, the set $\{2, 7\}$ has no element that is a set.

It is possible that both $A \in B$ and $A \subseteq B$ hold at the same time; e.g.,

$$\{1, 2\} \in \{\{1, 2\}, 1, 2\} , \quad \{1, 2\} \subseteq \{\{1, 2\}, 1, 2\} ;$$

however, such situations are rather rare in practice.


The set of all subsets of a given set A is called the *power set* of A , and it is denoted by $\mathcal{P}(A)$. That is to say,

$$X \in \mathcal{P}(A) \iff X \subseteq A .$$

E.g.,

$$\mathcal{P}(\{1, 2\}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\} .$$

$\mathcal{P}(A)$ is never empty; the empty set is always in it. Also, $A \in \mathcal{P}(A)$, and thus, if A is non-empty, $\mathcal{P}(A)$ has at least two elements.



With A an alphabet (see the first section), A^* the set of strings over A , a subset of A^* is called a *language* over A . The idea is that the strings in the language are the *well-formed sentences* of the language (a sentence is considered a single string, by treating the blanks as occurrences of a special character called **blank**). Initially, we do not put any condition on how the sentences should be formed; hence the complete generality of the definition of language. The theory of formal languages, an important part of theoretical computer science, deals mainly with how one can generate a language by rules. The legal programs of PASCAL form a (formal) language; the relevance of formal language theory should be indicated by this remark alone. Later we will also see particular formal languages related to logic.

The bracket notation for sets is used in a modified form to denote a subset of a given set.

$$\{n \in \mathbb{N} \mid n \text{ is prime}\}$$

is the same as

$$\{n \mid n \in \mathbb{N} \text{ and } n \text{ is prime}\},$$

the set of positive primes, and of course, it is a subset of \mathbb{N} .

In what follows, capital letters always denote sets.

The *intersection* of X and Y , in symbols $X \cap Y$, is the set whose elements are the things that are elements of both X and Y at the same time:

$$a \in X \cap Y \iff a \in X \text{ and } a \in Y.$$

E.g.,

$$\{1, 2, 4, 7, 10\} \cap \{3, 4, 10, 13\} = \{4, 10\},$$

$$\{1, 2, 4\} \cap \{3, 7\} = \emptyset,$$

$$\{n \in \mathbb{N} \mid n \text{ is even}\} \cap \{n \in \mathbb{N} \mid n \text{ is prime}\} = \{2\}.$$

If A_i is a set for all values $i = 1, 2, \dots, n$ of the subscript i , then $\bigcap_{i=1}^n A_i$, the intersection of the A_i , is the set of all those x which are elements of all A_i :

$$x \in \bigcap_{i=1}^n A_i \iff \text{for all } i \text{ such that } 1 \leq i \leq n, \text{ we have } x \in A_i.$$

E.g., if A_i is the set of natural numbers divisible by i , then

$$\bigcap_{i=1}^{10} A_i = A_{8 \cdot 9 \cdot 5 \cdot 7} = A_{2520}$$

(why?).

One can take the intersection of any family of sets. If I is any set, and A_i is a particular set for each $i \in I$ (in which case we talk about the *family* $\langle A_i \rangle_{i \in I}$ of sets), then $\bigcap_{i \in I} A_i$, the intersection of the A_i , is the set of all things that belong to every A_i , $i \in I$. The

notation $\bigcap_{i=1}^n A_i$ means the same as $\bigcap_{i \in \{1, \dots, n\}} A_i$.

E.g., $\bigcap_{n \in \mathbb{N} - \{0\}} \{k \cdot n \mid k \in \mathbb{N}\} = \{0\}$ (why?)

(here, of course, $\mathbb{N} - \{0\} = \{n \in \mathbb{N} \mid n > 0\}$).

The *union* of two sets X and Y , denoted $X \cup Y$, is the set of all things that are elements of either X , or Y , or both:

$$a \in X \cup Y \iff \text{either } a \in X, \text{ or } a \in Y \text{ (or both).}$$

E.g.,

$$\{1, 2, 4, 7, 10\} \cup \{3, 4, 10, 13\} = \{1, 2, 3, 4, 7, 10, 13\},$$

$$\{n \in \mathbb{N} \mid n \text{ is even}\} \cup \{n \in \mathbb{N} \mid n \text{ is odd}\} = \mathbb{N}.$$

We may take the union of more than two sets. $\bigcup_{i \in I} A_i$ denotes the set, called the *union* of the sets A_i , $i \in I$, whose elements are those things that belong to at least one of the sets A_i . E.g., if $B_i = \{5k+i \mid k \in \mathbb{N}\}$, then

$$\bigcup_{i \in \{0, 1, 2, 3, 4\}} B_i = \mathbb{N}$$

(why is that?). We write $\bigcup_{i=0}^4$ instead of $\bigcup_{i \in \{0, 1, 2, 3, 4\}}$, thus $\bigcup_{i=0}^4 B_i = \mathbb{N}$. Or, to give another example: if A_i is the set

$$A_i = \{n \in \mathbb{N} \mid n \text{ is divisible by } i, \text{ and } n \leq 120\},$$

then

$$\bigcup_{i=2}^7 A_i = \{n \in \mathbb{N} \mid n \leq 120 \text{ and } n \text{ is not prime}\} \cup \{2, 3, 5, 7\}$$

(this is related to the equality of the sets (2) and (3) in Section 1.1).

Here is another way we can use the union (\bigcup) and intersection (\bigcap) symbols.

Assume that X is a set of sets: that is, all elements of X are themselves sets. Then $\bigcup X$ denotes the union of all the sets in X .

For instance, if $X = \{A_i \mid i \in \{2, 3, 4, 5, 6, 7, \dots\}\}$, then $\bigcup X$ is the same as $\bigcup_{i=2}^7 A_i$ considered earlier.

We may give the definition of the notation $\bigcup X$ thus: for any x ,

$$x \in \bigcup X \iff \text{there is } A \in X \text{ such that } x \in A.$$

Note that $\bigcup \emptyset = \emptyset$.

The notation $\bigcap X$ is similar: it denotes the intersection of all the sets in the set X . In symbols:

$$x \in \bigcap X \iff \text{for all } A \in X, \text{ we have } x \in A.$$

Here, there is an exclusion: $\bigcap \emptyset$ does not make sense. The reason is that the last display would give that *all* things x belong to $\bigcap \emptyset$; however, "the set of *all* things" is not a legitimate concept.

The third operation we consider here is the *difference* of two sets. $X - Y$ denotes the set of those things that are in X , but are not in Y :

$$a \in X - Y \iff a \in X \text{ and } a \notin Y.$$

E.g.,

$$\{1, 2, 4, 7, 10\} - \{3, 4, 10, 13\} = \{1, 2, 7\}.$$

Let A be a fixed set, and consider the operations of intersection, union and difference performed on subsets of A ; the result is always a subset of A again. This is clear since those operations never involve elements that are not in either of the sets in question. We say that the collection of all subsets of A is *closed* under the operation of intersection, union and difference; the latter operations are collectively called the *Boolean operations*.

If all sets under consideration are understood to be subsets of the fixed set A , then the difference $A - X$ is abbreviated as $-X$, and it is called the *complement* of X , or the *complement of X with respect to A* in more detail. E.g., if we are talking about subsets of \mathbb{N} , that is, $A = \mathbb{N}$, then

$$- \{n \in \mathbb{N} \mid n \text{ is even}\} = \{n \in \mathbb{N} \mid n \text{ is odd}\}.$$

The Boolean operations obey certain laws. Here is a list of them. In what follows, A is a fixed set, X , Y and Z denote arbitrary subsets of A ; $-X$ means $A - X$.

Commutative laws:

$$X \cap Y = Y \cap X ,$$

$$X \cup Y = Y \cup X .$$

Associative laws:

$$X \cap (Y \cap Z) = (X \cap Y) \cap Z ,$$

$$X \cup (Y \cup Z) = (X \cup Y) \cup Z .$$

Just like in the case of addition and multiplication of numbers, these laws have the consequence that in expressions using several intersection operations, or alternatively, several union operations, parentheses may be omitted or restored in any meaningful way, and the order of terms may be changed, without altering the value of the expression. E.g.,

$$(X \cap (U \cap V)) \cap Z = (U \cap Z) \cap (V \cap X) = U \cap V \cap X \cap Z .$$

Absorption laws:

$$X \cap (X \cup Y) = X ,$$

$$X \cup (X \cap Y) = X .$$

Idempotent laws:

$$X \cap X = X$$

$$X \cup X = X$$

Distributive laws:

$$(X \cup Y) \cap Z = (X \cap Z) \cup (Y \cap Z) ,$$

$$(X \cap Y) \cup Z = (X \cup Z) \cap (Y \cup Z) .$$

Laws for complements:

$$X \cap -X = 0 ,$$

$$X \cup -X = A ;$$

De Morgan's laws:

$$-(X \cap Y) = (-X) \cup (-Y) ,$$

$$-(X \cup Y) = (-X) \cap (-Y) .$$

The proofs of these identities can be done via antisymmetry for \subseteq . We consider the first distributive law.

To show $(X \cup Y) \cap Z \subseteq (X \cap Z) \cup (Y \cap Z)$, let x belong to the left-hand-side, to show that it belongs to the right-hand side. Then x belongs both to $X \cup Y$ and Z . Since it belongs to $X \cup Y$, it either belongs to X (**Case 1**), or to Y (**Case 2**), or possibly both. In **Case 1**, x belongs both to X and Z , hence, to $X \cap Z$, and thus to $(X \cap Z) \cup (Y \cap Z)$. In **Case 2**, we obtain the same conclusion similarly. We have shown that x belongs to the right-hand side in any case.

To show $(X \cup Y) \cap Z \supseteq (X \cap Z) \cup (Y \cap Z)$, let x belong to the right-hand side, to show that it belongs to the left-hand side. Since $x \in (X \cap Z) \cup (Y \cap Z)$, either $x \in X \cap Z$ (**Case 1**), or $x \in Y \cap Z$ (**Case 2**). In the first case, $x \in X$ and $x \in Z$; from $x \in X$ it follows that $x \in X \cup Y$; hence, x belongs to both $X \cup Y$ and Z , and thus to the left-hand side. In Case 2, the argument is similar.

Section 1.3 Ordered pairs and functions

The ordered pair (a, b) of two things a and b is another thing that contains the information of both a and b , together the information that " a comes first, b second". Mathematically expressed, the essential property of the ordered-pair construction is

$$(a, b) = (c, d) \iff a = c \text{ and } b = d. \quad (1)$$

It is possible to construct the ordered pair set-theoretically; however, we will not do so here; all we ever use about ordered pairs is the fact expressed in (1). Let us note though that the pair-set $\{a, b\}$ would *not* work as the ordered pair: we have

$$\{0, 1\} = \{1, 0\},$$

but we want

$$(0, 1) \neq (1, 0).$$

The use of the ordered pair is familiar in coordinate geometry; the points in the plane equipped with a Cartesian coordinate system are represented by ordered pairs of real numbers. Various geometric figures become sets of ordered pairs. Denoting the set of ordered pairs of real numbers by \mathbb{R}^2 , the set

$$\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$$

is the circle with center the origin, and radius the unit length; that is, the set of points on that circle. The set

$$\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$$

is the open disc of radius 1 around the origin;

$$\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$$

is the closed disc; the open disc does not, the closed one does, contain the circumference.

For sets A and B , $A \times B$, the *Cartesian product of A and B* , is the set of all ordered pairs (a, b) with first element a from A , second element b from B :

$$A \times B \stackrel{\text{def}}{=} \{(a, b) : a \in A \text{ and } b \in B\}.$$

Thus, what we wrote as \mathbb{R}^2 above is the same as $\mathbb{R} \times \mathbb{R}$; in general, we may write A^2 for $A \times A$.

A *function f* from a set A to another set B is a rule that assigns, to every element a of A , a definite element of B ; this element is denoted by $f(a)$; it is called the *value* of the function f at the *argument* a . We write

$$f: A \longrightarrow B$$

to indicate that f is a function from A to B ; A is the *domain* of f , B is the *codomain* of f . The codomain of f has to be distinguished from the *range* of f ; the latter is the set $\{f(a) : a \in A\}$ of all values of f :

$$\text{range}(f) \stackrel{\text{def}}{=} \{f(a) : a \in A\}.$$

The range of f is a subset of the codomain of f ; the range and the codomain may or may not be the same.

It is possible to construe functions as sets, in particular, as sets of ordered pairs: with $f: A \longrightarrow B$, we may consider the set of all pairs $(a, f(a))$ with $a \in A$; this set is called the *graph* of the function f :

$$\text{graph}(f) \stackrel{\text{def}}{=} \{(a, f(a)) : a \in A\}.$$

This is exactly the representation of functions that we use in coordinate geometry and calculus.

For instance, with the exponential function $\exp: \mathbb{R} \rightarrow \mathbb{R}$ assigning e^x to x for all $x \in \mathbb{R}$, we associate its graph which is the exponential curve in the Cartesian plane.

Note that the range of $\exp: \mathbb{R} \rightarrow \mathbb{R}$ is the set of all positive real numbers,

$\mathbb{R}^+ = \{y \in \mathbb{R} : y > 0\}$. This is true since the values of the exponential function are all positive, and every positive real number is the value of \exp at a suitable argument $x \in \mathbb{R}$: if $y > 0$, then there is $x \in \mathbb{R}$ namely, $x = \ln(y)$, for which $y = f(x)$. The range of $\exp: \mathbb{R} \rightarrow \mathbb{R}$ does not coincide with its codomain: $\mathbb{R}^+ \subsetneq \mathbb{R}$.

Usually, we do not distinguish between the function and its graph; the exponential function and the exponential curve are considered to be the same thing. There is one qualification to this rule though: two functions $f: A \rightarrow B$ and $g: A \rightarrow C$, with the same domain but with different codomains, may have the same graph. E.g., the \sin function may be construed as $\sin: \mathbb{R} \rightarrow \mathbb{R}$, from \mathbb{R} to \mathbb{R} , or as $\sin: \mathbb{R} \rightarrow [-1, 1]$, from \mathbb{R} to the closed interval $[-1, 1]$ (since all values of \sin are in the latter interval); these two functions have the same graph. For us, these two functions are technically different; the specification of a function includes the specification of its domain as well as its codomain.

When is a set, say A , is the graph of a function? There are two conditions that are necessary and sufficient for this to hold:

- (i) Every element a of A must be an ordered pair: a must equal to (x, y) for suitable (uniquely determined) x and y ;
- (ii) For all x, y, z , $(x, y) \in A$ and $(x, z) \in A$ imply that $y = z$ [note the same x as first component in the two ordered pairs].

The second condition expresses the fact that for a function f , the value $y = f(x)$ is *uniquely determined* by the argument x . If (i) and (ii) hold true, then there is a function $f: X \rightarrow Y$ for which $\text{graph}(f) = A$. Here, X , the domain of f , is the set of all x for which there is y such that $(x, y) \in A$; Y , the codomain, is any set that *contains* as a subset the set R of all y for which there is x such that $(x, y) \in A$ (R is the *range* of f); and we have

$$y = f(x) \iff (x, y) \in A.$$

The usual notation for a function is to give its value at an indeterminate argument; thus, e^x

denotes the exponential function. This notation is ambiguous, however; it may also mean the value of the function at a certain argument-value of x . A more explicit notation e.g. for the exponential function is

$$x \longmapsto e^x \quad (x \in \mathbb{R})$$

Note here the vertical line at the beginning of the arrow; this kind of arrow is to be distinguished from the arrow that connects the domain and codomain of the function. If we write \exp for the exponential function, a full notation and description of the function \exp is this:

$$\begin{array}{ccc} \exp : \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & e^x . \end{array}$$

If we have two functions $f: A \longrightarrow B$ and $g: A \longrightarrow B$ between the same two sets, f and g are the *same function*, $f = g$, just in case for all $a \in A$, $f(a) = g(a)$:

$$f = g \iff \text{for all } a \in A, f(a) = g(a) .$$

This is in agreement with the construal of functions as sets of ordered pairs: $f = g$ just in case $\text{graph}(f) = \text{graph}(g)$; note that this is valid only if the two functions f and g are given already with the same domain and the same codomain.

Here is a notation for specifying a function when the domain of the function is a reasonably small finite set. I'll explain this on an example. For instance, the symbolic expression

$$\begin{array}{cccccc} (1 & 3 & 5 & 7 & 9 & 11) \\ 0 & 5 & 4 & 20 & 3 & 3 \end{array} \quad (2)$$

denotes the function whose domain is the set $\{1, 3, 5, 7, 9, 11\}$, the set which is listed in the upper row, and whose value for each argument in the domain is given in the second row underneath the particular argument; in the case of (2), if the function is called f , then $f(1)=0$, $f(3)=5$, $f(5)=4$, etc.

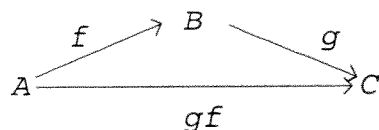
To be precise, we should note that this notation exhibits only the *graph* of the function. In the

example (2), the function f may have any codomain (which then has to be specified separately) that contains the set $\{0, 5, 4, 20, 3\}$, the range of the function f .

If we have two functions, $f: A \rightarrow B$ and $g: B \rightarrow C$, such that the codomain of the first is the same as the domain of the second, we can form their *composite* $g \circ f: A \rightarrow C$; the definition of $g \circ f$ is:

$$(g \circ f)(a) \stackrel{\text{def}}{=} g(f(a)) \quad (a \in A).$$

We may omit the circle in the notation of composition, and write simply gf . To see the domain/codomain relationships of the functions involved, we may draw the three functions f , g , and gf in the diagram



The composite of two functions is defined only if the codomain of one coincides with the domain of the other.

E.g., consider the functions

$$\begin{array}{ccc} f: \mathbb{N} & \longrightarrow & \mathcal{P}(\mathbb{N}) \\ n & \longmapsto & \{n\} \end{array} \quad \text{and} \quad \begin{array}{ccc} g: \mathcal{P}(\mathbb{N}) & \longrightarrow & \mathcal{P}(\mathbb{N}) \\ X & \longmapsto & \mathbb{N} - X \end{array}$$

Then, gf is the following function:

$$\begin{array}{ccc} gf: \mathbb{N} & \longrightarrow & \mathcal{P}(\mathbb{N}) \\ n & \longmapsto & \{x \in \mathbb{N} \mid x \neq n\} \end{array}$$

When, in the calculus, we talk about a function like $\sin(e^x)$, we have in mind a composite; in the case at hand, the composite $\sin \circ \exp$:

$$\begin{array}{ccccc}
 \mathbb{R} & \xrightarrow{\exp} & \mathbb{R} & \xrightarrow{\sin} & \mathbb{R} \\
 x & \longmapsto & e^x & & \\
 & & y & \longmapsto & \sin(y) \\
 x & \longmapsto & e^x & \longmapsto & \sin(e^x)
 \end{array}$$

The operation of composition of functions satisfies the *associative law*: in the situation

$$A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D,$$

we have that

$$h(gf) = (hg)f.$$

Indeed, $h(gf)$ applied to any $a \in A$ gives

$$[h(gf)](a) = h([gf](a)) = h(g(f(a))) ;$$

and, $(hg)f$ applied to a gives

$$[(hg)f](a) = (hg)(f(a)) = h(g(f(a))) ,$$

which is the same value. Since the two functions, both from A to D , give the same value at each argument $a \in A$, they are equal.

With any set A , there is a particular function associated, namely the *identity function on A* :

$$\begin{array}{ccc}
 1_A : A & \longrightarrow & A \\
 a & \longmapsto & a
 \end{array}$$

($1_A(a) = a$ for $a \in A$). This has the property that its composite with any function, provided it is well-defined, is the function itself:

$$B \xrightarrow{f} A \xrightarrow{1_A} A \quad :: \quad 1_A \circ f = f ,$$

$$A \xrightarrow{1_A} A \xrightarrow{g} C \quad :: \quad g \circ 1_A = g .$$

Another operation on functions is *restriction*. Suppose $f: A \rightarrow B$ and $A' \subseteq A$. Then the *restriction of f to A'* is the function denoted as $f \upharpoonright A' : A' \rightarrow B$ for which $(f \upharpoonright A')(a) = f(a)$ for all $a \in A'$. E.g., for the absolute-value function $|-| : \mathbb{Z} \rightarrow \mathbb{N}$, its restriction to the subset \mathbb{N} of its domain is $|-| \upharpoonright \mathbb{N} = 1_{\mathbb{N}}$, the identity function on \mathbb{N} ; the reason is that $|n| = n$ for all $n \in \mathbb{N}$.

With any subset A' of any set A , one can associate the *inclusion function* $\varphi: A' \rightarrow A$, which acts like the identity: $\varphi(a) = a$ ($a \in A'$); what makes it different from the identity function is that its domain and codomain are not (necessarily) equal. Note that, with the notation of this and the previous paragraph, $f \upharpoonright A' = f \circ \varphi$.

A function $f: A \rightarrow B$ is *injective*, or *one-to-one*, or f is an *injection*, if it maps distinct arguments to distinct values:

$$a \neq a' \implies f(a) \neq f(a') \quad \text{for any } a, a' \in A .$$

A more positive, but equivalent, way of putting the definition of injectivity is that

$$f(a) = f(a') \implies a = a' \quad \text{for any } a, a' \in A .$$

E.g., the exponential function $\exp: \mathbb{R} \rightarrow \mathbb{R}$ is injective: if x and y are two distinct real numbers, then either $x < y$, or $y < x$; in the first case $e^x < e^y$ (the exponential function is strictly increasing), in the second case the other way around; thus, at any rate, $e^x \neq e^y$. But, the \sin function is not injective: $\sin(0) = \sin(\pi) = 0$.

$f: A \rightarrow B$ is *surjective*, or *onto*, or f is a *surjection*, if for any $b \in B$, there is at least one $a \in A$ such that $f(a) = b$. f is surjective just in case its range equals its codomain.

E.g., the range of $\sin: \mathbb{R} \rightarrow \mathbb{R}$ is

$$[-1, 1] \stackrel{\text{def}}{=} \{y \in \mathbb{R} \mid -1 \leq y \leq 1\} ;$$

thus, $\sin: \mathbb{R} \rightarrow \mathbb{R}$ is not surjective (e.g., for $y = 2$, there is no x such that $\sin(x) = y = 2$). But, if we consider \sin to be a function from \mathbb{R} to the interval $[-1, 1]$, $\sin: \mathbb{R} \rightarrow [-1, 1]$, then \sin , in this sense, is surjective. (Note that for us, the information of the codomain is part of the data defining the function. Thus, $\sin: \mathbb{R} \rightarrow \mathbb{R}$ and $\sin: \mathbb{R} \rightarrow [-1, 1]$ are, strictly speaking, not the same function.)

If $A \xrightleftharpoons[f]{g} B$, and $gf = 1_A$, we say that g is a *left inverse* of f , or that f is a *right inverse* of g . If $gf = 1_A$ and $fg = 1_B$ both hold, g is a *two-sided inverse*, or simply, an *inverse*, of f (and then, of course, f is an inverse of g).

Consider

$$\begin{array}{ccc} \left\lfloor \frac{k}{2} \right\rfloor & \xleftarrow{\quad g \quad} & k \\ \mathbb{N} & \xrightleftharpoons[f]{\quad} & \mathbb{N} \\ n & \xrightarrow{\quad f \quad} & 2n \end{array}$$

(here, $\left\lfloor \frac{k}{2} \right\rfloor$ denotes the largest integer not greater than $\frac{k}{2}$). Then $gf = 1_{\mathbb{N}}$, since

$$(gf)(n) = g(2n) = \left\lfloor \frac{2n}{2} \right\rfloor = [n] = n = 1_{\mathbb{N}}(n).$$

However, $fg \neq 1_{\mathbb{N}}$; e.g., $(fg)(1) = 2 \left\lfloor \frac{1}{2} \right\rfloor = 2 \cdot 0 = 0 \neq 1$. Thus, in this case, g is a left inverse of f , but it is not a right inverse of it.

We claim that in the situation:

$$A \xrightleftharpoons[f]{g} B, \text{ and } gf = 1_A,$$

f is injective and g is surjective. Indeed, if $a, a' \in A$, and $f(a) = f(a')$, then

$$\begin{array}{c} a = g(f(a)) = g(f(a')) = a', \\ \uparrow \\ gf = 1_A \end{array}$$

which shows the injectivity of f . On the other hand, if $a \in A$ is an arbitrary element of A , then for $b = f(a)$, we have $g(b) = g(f(a)) = a$ (again since $gf = 1_A$); this shows that g is surjective.

We have shown that

if a function (f in the previous situation) has a left inverse, then it is injective, and if a function (g above) has a right inverse, it is surjective.

The converses of the last two assertions are *almost* true. First,

if $f: A \rightarrow B$ is injective, and if A is not empty, then f has a left inverse:

given any $b \in B$, define $g(b)$ to be $a \in A$ for which $f(a) = b$ if there is (necessarily at most) one such a ; if however there is no such a , let $g(b)$ be any element in A (since A is not empty, there is at least one such). Then $(gf)(a) = g(f(a)) = a$ by the definition of g on $b = f(a)$; thus $gf = 1_A$.

Secondly,

if $g: B \rightarrow A$ is surjective, then it has a right inverse.

Namely, we define $f: A \rightarrow B$ in the following way. Given any $a \in A$, we pick an arbitrary $b \in B$ such that $g(b) = a$; by the assumption of g being surjective, there is certainly at least one such b ; we make $f(a)$ equal this b . Then, with $f: A \rightarrow B$ so defined, $(gf)(a) = g(b)$ for the b described above; but the choice of that b was such that $g(b) = a$; this shows that $(gf)(a) = a$ for any $a \in A$, which is to say that f is a right inverse of g . [In a foundational setting, this argument requires the so-called *Axiom of*

Choice.]

Returning to the previous assertion, the additional assumption of A being non-empty is necessary: any $f: \emptyset \rightarrow B$ is injective, but there is a function $g: B \rightarrow \emptyset$ at all only if B is also empty.

If a function is both injective and surjective, it is called *bijective*, or a *bijection*. Here are two examples for bijection:

$$\begin{array}{ccc} f: \mathbb{Z} & \longrightarrow & \mathbb{Z} \\ x & \longmapsto & x - 1 \end{array}$$

$$\begin{array}{ccc} g: \mathbb{Q} & \longrightarrow & \mathbb{Q} \\ x & \longmapsto & 2x \end{array}$$

The symbol \cong is used to indicate a bijection: $f: A \xrightarrow{\cong} B$.

To say that a function is a bijection is the same as to say that it has an inverse.

Indeed, if it has a (two-sided) inverse, then, by what we said above, it is both injective and surjective. On the other hand, if $f: A \rightarrow B$ is bijective, and for a moment, we assume that A is non-empty, then f has a left inverse $g: B \rightarrow A$ and a right inverse $h: B \rightarrow A$: $gf = 1_A$, $fh = 1_B$. But then

$$h = 1_A \circ h = (gf)h = g(fh) = g \circ 1_B = g,$$

which shows that $h = g$ is a two-sided inverse of f . If A happens to be empty, then, with $f: A \rightarrow B$ bijective, in particular, surjective, B must also be empty; in this case, $f = 1_{\emptyset}$, the "empty function", is a two-sided inverse of itself.

The last argument also shows that

the (two-sided) inverse, if exists, is uniquely determined:

if g and h are both inverses of f , then g is a left inverse, h is a right inverse, of f , and the calculation above shows that $g = h$. Moreover, we have also shown that

if f has a left inverse g , and also a right inverse h , then $g=h$, and thus f has a two-sided inverse, namely $g = h$.

The inverse of $f: A \rightarrow B$, if exists, is denoted by f^{-1} . Thus, the defining properties of $f^{-1}: B \rightarrow A$ are:

$$f \circ f^{-1} = 1_B \text{ and } f^{-1} \circ f = 1_A.$$

The composite of two injections (if well-defined) is an injection; the composite of two surjections is a surjection; the composite of two bijections is a bijection.

We leave the easy proof to the reader.

A *permutation* of a set A is any bijection from A to A itself. The following denotes a permutation of the set $\{1, 2, 3, 4, 5\}$:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 3 & 1 \end{pmatrix};$$

(Recall this notation for a function from above, at (2): if σ is the name of the permutation at hand, then $\sigma(1)=4$, $\sigma(2)=2$, $\sigma(3)=5$, $\sigma(4)=3$, $\sigma(5)=1$.)

The set of all functions $A \longrightarrow B$ is denoted by the exponential notation B^A .

Sequences are particular functions. E.g., the 5-term sequence $\langle a_1, a_2, a_3, a_4, a_5 \rangle$ may be identified with the function whose domain is the set $\{1, 2, 3, 4, 5\}$, and whose value at i is a_i . The notation $\langle a_i \rangle_{1 \leq i \leq n}$ means the n -term sequence whose i^{th} term is a_i . A^n denotes the set of all n -term sequences of members of A . Thus,

$$A^n = \{ \langle a_i \rangle_{1 \leq i \leq n} \mid a_i \in A \text{ for all } i < n \}.$$

Note that for $n=0$, for any set A , there is exactly one 0-term sequence of elements of A , the *empty sequence* \perp ; $A^0 = \{\perp\}$.

If A is an *alphabet*, that is, a set of characters, then *strings* over A are essentially the same as finite sequences of elements of A ; strings of length n are the same as n -term sequences of elements of A . $A^* \stackrel{\text{def}}{=} \bigcup_{n \in \mathbb{N}} A^n$ is the set of all strings over A . Note that A^* always contains \perp , the empty string, as an element.

If, for instance, $A = \{a, b, c\}$, then the strings $aabccba$ and $bccb$ are members of A^* ; the first belongs to A^7 , the second to A^4 .

Infinite sequences are the same as functions with domain \mathbb{N} ; $\mathbb{R}^{\mathbb{N}}$ is the set of all infinite sequences $\langle r_i \rangle_{i \in \mathbb{N}} = \langle r_0, r_1, \dots, r_n, \dots \rangle$ of reals. Infinite sequences of reals are important in the calculus.

Some more notation related to functions. Let $f: A \longrightarrow B$. If $X \subseteq A$, the *image of X under f* , denoted $f[X]$, is the set of all values of f while the argument of f ranges over X :

$$f[X] \stackrel{\text{def}}{=} \{ f(a) \mid a \in X \}.$$

E.g., when

$$\begin{aligned} f: \mathbb{N} &\longrightarrow \mathbb{N} \\ n &\longmapsto 2n \end{aligned}$$

and $X_1 = \{n \in \mathbb{N} \mid n \text{ is even}\}$, $X_2 = \{n \in \mathbb{N} \mid n \text{ is odd}\}$, then

$$f[X_1] = \{n \in \mathbb{N} \mid n \text{ is divisible by } 4\},$$

$$f[X_2] = \{n \in \mathbb{N} \mid n \text{ is even, but not divisible by } 4\}.$$

If $Y \subseteq B$, the *inverse image of Y under f* , $f^{-1}[Y]$ (warning: this notation does not imply that the inverse of f , f^{-1} , exists!) is the set of all $a \in A$ that are mapped into Y by f :

$$f^{-1}[Y] \stackrel{\text{def}}{=} \{a \in A : f(a) \in Y\}.$$

E.g., with continuing the previous example, $f^{-1}[X_1] = \mathbb{N}$ and $f^{-1}[X_2] = \emptyset$.

In the general case $f: A \longrightarrow B$, let $b \in B$. Then $f^{-1}[\{b\}]$ is the set of those $a \in A$ whose f -image is b , $f(a) = b$. Thus, f is injective iff for all $b \in B$, $f^{-1}[\{b\}]$ has at most one element; f is surjective iff for all $b \in B$, $f^{-1}[\{b\}]$ has at least one element; and f is bijective iff for all $b \in B$, $f^{-1}[\{b\}]$ has exactly one element.