# Foundations of Mathematics
# MATH 300
# Lecture Notes

F. Baudier (Texas A&M University)

November 13, 2023

# Contents

# Chapter 1

# Introduction to Mathematical Logic

## 1.1 Statements and predicates

A mathematical proof should not be subject to personal interpretation and to avoid ambiguity we need to restrict our attention to certain types of declarative sentences.

**Definition 1: Statement**

A *statement* is any declarative sentence that has a truth value (either true or false).

Characteristics of statements:

- a statement has a truth value;
- a statement is either true or false;
- a statement cannot be neither true nor false;
- a statement cannot be true and false.

We will often represent statements with capital letters, such as $P$, $Q$, ... Mathematical statements are commonly written with symbols for convenience but should be thought of as full-fledged sentences.

*Example* 1. $P : 3+5=8$, is a statement.

*Example* 2. $P : 3+5=9$, is a statement.

**Definition 2: Predicate**

A *predicate* is any declarative sentence containing one or more variables that is not a statement but becomes a statement when the variables are assigned values.

A predicate is usually written $P(n)$, $Q(x,y)$, and variants thereof, depending on the number of variables and the letters used for the variables.

*Example* 3. $P(x) : x+1 = 2$, is a predicate with one variable.

*Example* 4. $P(m,n) : n+m$ is odd, is a predicate with two variables.

*Exercise* 1. Are the following sentences statements, predicates or none of these?

1. Michael Phelps won 23 gold medals.
2. 3+5=8
3. 3+5=9
4. Today is cold.
5. x+5=8
6. table+5=8
7. This sentence is false.

## 1.2 Logical connectives

We have introduced two types of expressions that we will use in our mathematical proofs: statements and predicates. We can build more complicated expressions using the basic logical connectives: $\neg$ (negation), $\wedge$ (conjunction), $\vee$ (disjunction).

*Terminology.* Expressions of the form $P \wedge Q$, $P \vee Q$, $\neg P$, $(\neg P) \wedge (Q \vee \neg R)$, and so on, where $P$ and $Q$ are considered as variables representing statements are called statement forms. They are not actually statements themselves but become statements when the variables $P$ and $Q$ are replaced by statements.

### 1.2.1 Negation, disjunction, conjunction

**Definition 3: Negation**

If $P$ is a statement, the negation of $P$ is the statement "not $P$". We use the notation $\neg P$, which reads "not $P$" for the negation of $P$.

If $P$ is a statement only the following two cases can occur: either ($P$ is true and $\neg P$ is false) or ($P$ is false and $\neg P$ is true).

Truth tables for statement forms are tables that give the truth value of the statement form in terms of the truth values of the variables and are used to rigorously define the action of a logical connective on the statement(s) it operates.

**Definition 4: Conjunction**

Let $P$ and $Q$ be statements. The conjunction of $P$ and $Q$ is the statement "$P$ and $Q$". The notation for the conjunction of $P$ and $Q$ is $P \wedge Q$ and reads "$P$ and $Q$".

| $P$ | $\neg P$ |
|---|---|
| T | F |
| F | T |

Table 1.1: Negation truth table

| $P$ | $Q$ | $P \wedge Q$ |
|---|---|---|
| T | T | T |
| T | F | F |
| F | T | F |
| F | F | F |

Table 1.2: Conjunction truth table

**Definition 5: Disjunction**

Let $P$ and $Q$ be statements. The disjunction of $P$ and $Q$ is the statement "$P$ or $Q$". The notation for the disjunction of $P$ and $Q$ is $P \vee Q$ and reads "$P$ or $Q$".

| $P$ | $Q$ | $P \vee Q$ |
|---|---|---|
| T | T | T |
| T | F | T |
| F | T | T |
| F | F | F |

Table 1.3: Disjunction truth table

Using logical connectives one can create new statements out of given statements. One can naturally extend the definitions above to create new predicates out of given predicates.

*Example* 5. The predicate $R(x) : |x| > 3$ is the disjunction of the predicates $P(x) : x > 3$ and $Q(x) : x < -3$, *i.e.*, $R(x) = P(x) \vee Q(x)$.

*Example* 6. The system of linear equations

$$\begin{cases} 2x+1 &= 0 \\ 3y-2 &= 0 \end{cases}$$

is a predicate with two variables $R(x,y)$ which is the conjunction of the predicates $P(x) : 2x+1 = 0$ and $Q(y) : 3y-2 = 0$, *i.e.*, $R(x,y) = P(x) \wedge Q(y)$.

The disjunction is commutative, since it is plain that $P \vee Q$ and $Q \vee P$ have the same truth tables. The same remark holds for the conjunction. We can make this observation precise by defining the notion of logical equivalence between statement forms.

**Definition 6: Logically equivalent statement forms**

We say that two statement forms are *logically equivalent* if they have the same truth tables.

We sometimes use the notation $\equiv$ for logical equivalence.

*Example* 7. As we just observed $P \vee Q \equiv Q \vee P$ and $P \wedge Q \equiv Q \wedge P$.

*Example* 8. By looking at their truth tables it is easy to see that the statement forms $P$ and $\neg\neg P$ are logically equivalent.

**Theorem 1: DeMorgan's Laws**

1. $\neg(P \wedge Q)$ is logically equivalent to $(\neg P) \vee (\neg Q)$.
2. $\neg(P \vee Q)$ is logically equivalent to $(\neg P) \wedge (\neg Q)$.

*Proof.* We just need to build the truth tables of all the statement forms involved.

| $P$ | $Q$ | $P \vee Q$ | $P \wedge Q$ | $\neg[P \vee Q]$ | $\neg[P \wedge Q]$ | $\neg P$ | $\neg Q$ | $\neg P \vee \neg Q$ | $\neg P \wedge \neg Q$ |
|---|---|---|---|---|---|---|---|---|---|
| T | T | T | T | F | F | F | F | F | F |
| F | T | T | F | F | T | T | F | T | F |
| T | F | T | F | F | T | F | T | T | F |
| F | F | F | F | T | T | T | T | T | T |

□

*Exercise* 2. What is the negation of the predicate $0 < x < 1$. Find a useful denial of the predicate $0 < x < 1$?

*Solution.* By drawing a picture you can certainly guess that

$$\neg(0 < x < 1) \equiv (x \leq 0) \vee (x \geq 1)$$

(and certainly not $0 \geq x \geq 1$)!) But a formal and rigorous proof ,using the basic Boolean logic rules and calculus that we have seen so far, requires unfolding the meaning of $0 < x < 1$ and would go as follows:

$$\neg(0 < x < 1) \equiv \neg[(0 < x) \wedge (x < 1)] \equiv [\neg(0 < x)] \vee [\neg(x < 1)] \equiv (x \leq 0) \vee (x \geq 1)$$

where we have use one DeMorgan's Law. □

**Definition 7: Tautology**

A statement form that is always true no matter what are the truth values of the variables is called a *tautology*.

*Example* 9. $P \vee (\neg P)$ is a tautology.

**Definition 8: Contradiction**

A statement form that is always false no matter what are the truth values of the variables is called a *contradiction*.

*Example* 10. $P \wedge (\neg P)$ is a contradiction.

Note that if $S$ is a tautology then $\neg S$ is a contradiction and vice-versa.

### 1.2.2 Implication, contrapositive, converse, biconditional

Roughly speaking an implication is a statement with an "if-then" structure. The "if" part of the statement gives the premise or assumption that is made, and $P$ is called the hypothesis or antecedent. The "then" part is the conclusion that is asserted from the premise and $Q$ is called the conclusion or consequent.

**Definition 9: Implication**

Let $P$ and $Q$ be statements. The *implication* "$P \implies Q$" (read "$P$ implies $Q$") is the statement "If $P$, then $Q$."

There is no sense of causality in the statement "$P \implies Q$" and $P$ might be (apparently) entirely unrelated to $Q$. The *only* case when an implication is false is when $P$ is true and $Q$ is false. In particular a false proposition implies anything!

| $P$ | $Q$ | $P \implies Q$ |
|---|---|---|
| T | T | T |
| F | T | T |
| T | F | F |
| F | F | T |

Table 1.4: Implication truth table

**Theorem 2**

1. $P \implies Q$ is logically equivalent to $(\neg P) \vee Q$.
2. $\neg(P \implies Q)$ is logically equivalent to $P \wedge \neg Q$.

*Proof.* We compare the truth tables.

For 2. we could also give a proof using DeMorgan's law and (1), since $\neg[(\neg P) \vee Q] \equiv (\neg\neg P) \wedge \neg Q \equiv P \wedge \neg Q$. □

| $P$ | $Q$ | $P \implies Q$ | $\neg[P \implies Q]$ | $\neg P$ | $\neg P \vee Q$ | $\neg Q$ | $P \wedge \neg Q$ |
|---|---|---|---|---|---|---|---|
| T | T | T | F | F | T | F | F |
| F | T | T | F | T | T | F | F |
| T | F | F | T | F | F | T | T |
| F | F | T | F | T | T | T | F |

*Exercise* 3. Let

$P$ : The square function is differentiable at 0.

$Q$ : The square function is continuous at 0.

Are the implications $P \implies Q$, $Q \implies P$ true?

**Definition 10: Contrapositive**

Let $P$ and $Q$ be statements. The statement $(\neg Q) \implies \neg P$ is called the contrapositive of the statement $P \implies Q$.

**Theorem 3: Logical equivalence between an implication and its contrapositive**

$P \implies Q$ is logically equivalent to $(\neg Q) \implies (\neg P)$.

*Proof.* First observe that $(P \implies Q) \equiv (\neg P) \vee Q$. On the other hand,

$$[(\neg Q) \implies (\neg P)] \equiv [(\neg\neg Q) \vee \neg P] \equiv [Q \vee \neg P] \equiv [(\neg P) \vee Q],$$

and the conclusion follows.

We could also have compared the truth tables.

| $P$ | $Q$ | $P \implies Q$ | $\neg P$ | $\neg Q$ | $\neg Q \implies \neg P$ |
|---|---|---|---|---|---|
| T | T | T | F | F | T |
| F | T | T | T | F | T |
| T | F | F | F | T | F |
| F | F | T | T | T | T |

□

**Definition 11**

Let $P$, $Q$ be statements. The statement $Q \implies P$ is called the converse of the statement $P \implies Q$.

**Proposition 1**

$P \implies Q$ is NOT logically equivalent to $Q \implies P$.

*Proof.* We compare the truth tables.

| $P$ | $Q$ | $Q \implies P$ | $P \implies Q$ |
|---|---|---|---|
| T | T | T | T |
| F | T | F | T |
| T | F | T | F |
| F | F | T | T |

□

**Definition 12: Biconditional or equivalence**

Let $P$ and $Q$ be statements. The statement $P \iff Q$ (or $P$ iff $Q$, read $P$ if and only if $Q$) is the statement $(P \implies Q) \wedge (Q \implies P)$

The statement $(P \implies Q) \wedge (Q \implies P)$ is true when $P$ and $Q$ are simultaneously true or simultaneously false, and false otherwise. The symbol $\iff$ is called the biconditional.

| $P$ | $Q$ | $Q \implies P$ | $P \implies Q$ | $(P \implies Q) \wedge (Q \implies P)$ |
|---|---|---|---|---|
| T | T | T | T | T |
| F | T | F | T | F |
| T | F | T | F | F |
| F | F | T | T | T |

| $P$ | $Q$ | $P \iff Q$ |
|---|---|---|
| T | T | T |
| F | T | F |
| T | F | F |
| F | F | T |

Table 1.5: Biconditional truth table

Consider the implication "$P \implies Q$". We say that $P$ is a *sufficient condition* for $Q$, because in order for $Q$ to be true it is sufficient that $P$ be true. Also, we say that $Q$ is a *necessary condition* for $P$ meaning that $Q$ must be true in order for $P$ to be true, or in other words if $Q$ is false then $P$ is false.

**Theorem 4**

1. $P \iff Q$ is logically equivalent to $((\neg P) \vee Q) \wedge ((\neg Q) \vee P)$.
2. $P \iff Q$ is logically equivalent to $Q \iff P$.

*Proof.* For 1. one has

$$(P \iff Q) \equiv (P \implies Q) \wedge (Q \implies P) \equiv ((\neg P) \vee Q) \wedge ((\neg Q) \vee P)$$

For 2.

$$(P \iff Q) \equiv (P \implies Q) \wedge (Q \implies P) \equiv (Q \implies P) \wedge (P \implies Q) \equiv (Q \iff P)$$

□

**Remark 1**

The placement of the parentheses in statement forms matters. As it can be easily seen by examining their truth tables $(\neg P \vee Q) \wedge (\neg Q \vee P)$ and $\neg P \vee (Q \wedge \neg Q) \vee P$ are not logically equivalent (actually $\neg P \vee (Q \wedge \neg Q) \vee P$ is a tautology).

*Exercise* 4. Are the statement forms $P \vee Q$, $\neg P \implies Q$, and $\neg Q \implies P$ logically equivalent?

*Solution.* Yes.

| $P$ | $Q$ | $P \vee Q$ | $\neg P \implies Q$ | $\neg Q \implies P$ |
|---|---|---|---|---|
| T | T | T | T | T |
| T | F | T | T | T |
| F | T | T | T | T |
| F | F | F | F | F |

□

## 1.3 Quantifiers

We can turn a prediaste into a statement by assigning a value to the variable, e.g. if $P(x)$ is the predicate "$x^2 + 1 = 0$" then $P(1)$ is a statement (which is false) and $P(i)$ is a statement (which is true). Another way a predicate can be made into a statement is by modifying it with quantifiers that acts on the free variables which live in a certain ambient (and often implicit) universe. For example, it is clear that the declarative sentences:

"For all $x$, $x^2 + 1 = 0$."

or

"There exists $x$ such that $x^2 + 1 = 0$."

have a truth value and are thus genuine statements.

In general, if $P(x)$ is a predicate, then the mathematical expression "$(\exists x)P(x)$" (read "there exists $x$ such that $P(x)$") is also a declarative sentence with a truth value. The symbol $\exists$ is called the *existential quantifier*.

**Definition 13: Turning a predicate into a statement with an existential quantifier**

Let $P(x)$ be a predicate. The declarative sentence $(\exists x)P(x)$ is a statement that is true exactly when at least one individual element $a$ in the ambient universe has the property that $P(a)$ is true.

Similarly, if $P(x)$ is a predicate, then the mathematical expression "$(\forall x)P(x)$" (read "for all $x$, $P(x)$") is a declarative sentence with a truth value. The symbol $\forall$ is called the *universal quantifier*.

**Definition 14: Turning a predicate into a statement with a universal quantifier**

Let $P(x)$ be a predicate. The declarative sentence $(\forall x)P(x)$ is a statement that is true exactly when every element $a$ in the ambient universe has the property that $P(a)$ is true.

**Remark 2**

In practice it is usually simpler and more convenient to use the same letter for the variable and its assigned value. We will do it from now on.

*Terminology.* A variable $x$ is called a bound variable once a quantifier is applied to $x$. Otherwise we say that $x$ is a free variable.

Some statements without an explicit "for all" can also be classified as universal statements. For instance the statement

"Every number is prime."

is equivalent to

"For all number $n$, $n$ is prime."

The existential and universal quantifiers are closely related. It is clear that we want the negation of the statement

"Every number is prime."

to be

"There exists a number that is not prime."

If we let $P(n)$ be the predicate "$n$ is prime", we should all agree that the negation of $(\forall n)P(n)$ should be $(\exists n)\neg P(n)$. Similarly, the negation of an existential statement should be a universal statement and we make the following rules.

**Definition 15: Rules of negation for quantifiers**

The two basic rules to negate statements with quantifiers are:

**Rule 1** the negation of the statement "$(\exists x)P(x)$" is the statement "$(\forall x)\neg P(x)$",

**Rule 2** the negation of the statement "$(\forall x)P(x)$" is the statement "$(\exists x)\neg P(x)$".

When discussing the truth values of the following statements:

1. $(\forall x)\, x+5=8$
2. $(\exists x)\, x+5=8$
3. $(\exists x)\, x^2+1=0$
4. $(\exists n)\, n+5=\pi$,

we might disagree if, for instance, in 5. we consider $n$ to be a natural number (it would be false) or a real number (it would be true then). The truth values of statements that come from binding with a quantifier the free variable of a predicate depend on the intended universe in which the variable belong. To avoid ambiguity, we will usually make the intended universe explicit unless it is completely clear from the context.

An important notion in mathematical logic is the notion of "membership". To say that a free variable belongs to a specific universe $\mathscr{U}$, we write $x \in \mathscr{U}$. If $\mathscr{U}$ is given, $x \in \mathscr{U}$ is a predicate that can become a statement when we either assign a value to $x$ or bound the variable $x$ with a quantifier.

Given a predicate $P(x)$ and a universe $\mathscr{U}$, consider the statement

"There exists $x$ in the universe $\mathscr{U}$ such that $P(x)$."

What formal logical expression using what we have introduced so far (logical connectives, quantifiers, and membership) would convey the right meaning of the statement above? We need a finer degree of precision. More precisely, what we are trying to say is that

"There exists $x$ that is in the universe $\mathscr{U}$ and such that $P(x)$."

and the formal expression

$$(\exists x)[(x \in \mathscr{U}) \wedge P(x)]. \tag{1.1}$$

would convey the right meaning.

Similarly, we want to understand what logical expression would best describe the statement:

"For all $x$ in the universe $\mathscr{U}$, $P(x)$."

Here we are *not* trying to say that "For all $x$, $x$ is in $\mathscr{U}$ and $P(x)$.". What we are implicitly saying is that "For all $x$ that is in the universe $\mathscr{U}$, then $P(x)$.", or even more explicitly:

"For all $x$, if $x$ is in the universe $\mathscr{U}$ then $P(x)$."

The meaning of the last statement is captured by the formal expression

$$(\forall x)[(x \in \mathscr{U}) \implies P(x)]. \tag{1.2}$$

**Remark 3**

The fact that an implication is true when its assumption is false is crucial here. Indeed, if an implication could be false when its hypothesis is false then the formal statement $(\forall x)[(x \in \mathscr{U}) \implies P(x)]$ would be false each time there are elements in the ambient universe but not in the universe $\mathscr{U}$ of concern. The intended meaning of "For all $x$ in the universe $\mathscr{U}$, $P(x)$." is certainly to disregard those elements not in $\mathscr{U}$ and the logical definition of an implication is partially designed to achieve this.

**Remark 4**

The formal statement $(\exists x)[(x \in \mathscr{U}) \implies P(x)]$ would not capture the intended meaning of "There exists $x$ in the universe $\mathscr{U}$ such that $P(x)$." Indeed, each time we can find some $a$ not in the universe $\mathscr{U}$ then the implication $[(a \in \mathscr{U}) \implies P(a)]$ would be true and thus $(\exists x)[(x \in \mathscr{U}) \implies P(x)]$ would be true just because there are elements in the ambient universe but not in the universe of concern; this is not aligned with what we are trying to convey here.

Expressions of the form (1.1) and (1.2) are ubiquitous in mathematics and we need a more convenient and concise way to write them.

**Definition 16: Useful abbreviations**

Let $P(x)$ be a predicate.

1. The expression $(\exists x \in \mathscr{U})P(x)$ is an abbreviation for $(\exists x)[(x \in \mathscr{U}) \wedge P(x)]$, i.e.,
$$(\exists x \in \mathscr{U})P(x) \equiv (\exists x)[(x \in \mathscr{U}) \wedge P(x)].$$
2. The expression $(\forall x \in \mathscr{U})P(x)$ is an abbreviation for $(\forall x)[(x \in \mathscr{U}) \implies P(x)]$, i.e.,
$$(\forall x \in \mathscr{U})P(x) \equiv (\forall x)[(x \in \mathscr{U}) \implies P(x)].$$

We certainly want the negation of

"Every natural number is prime."

to be

"There exists a natural number that is not prime."

as we implicitly want to stay in the universe where the statement takes place (the set $\mathbb{N}$ of natural number here). Therefore, it would be a problem if the negation of a statement of the form $(\forall x \in \mathscr{U})P(x)$ is not $(\exists x \in \mathscr{U})\neg P(x)$.

From Rule 1 and Rule 2, we can verify that the rules of negation for the abbreviations just discussed above are actually in line with our intuition.

**Theorem 5: Negation of statements with quantifiers and membership**

1. $\neg[(\exists x \in \mathscr{U})P(x)]$ is logically equivalent to $(\forall x \in \mathscr{U})\neg P(x)$.
2. $\neg[(\forall x \in \mathscr{U})P(x)]$ is logically equivalent to $(\exists x \in \mathscr{U})\neg P(x)$.

*Proof.* 1.

$$\begin{aligned}\neg[(\exists x \in \mathscr{U})P(x)] &\equiv \neg[(\exists x)(x \in \mathscr{U}) \wedge P(x)] \\ &\equiv (\forall x)(\neg(x \in \mathscr{U})) \vee \neg P(x) \\ &\equiv (\forall x)[x \in \mathscr{U} \implies \neg P(x)] \\ &\equiv (\forall x \in \mathscr{U})\neg P(x)\end{aligned}$$

2.

$$\begin{aligned}\neg[(\forall x \in \mathscr{U})P(x)] &\equiv \neg[(\forall x)[x \in \mathscr{U} \implies P(x)]] \\ &\equiv (\exists x)\neg[x \in \mathscr{U} \implies P(x)] \\ &\equiv (\exists x)(x \in \mathscr{U}) \wedge \neg P(x) \\ &\equiv (\exists x \in \mathscr{U})\neg P(x)\end{aligned}$$

□

*Example* 11. The negation of "$(\forall x)(\exists y)P(x,y)$" is "$(\exists x)\neg[(\exists y)P(x,y)]$" which is "$(\exists x)(\forall y)\neg P(x,y)$".

Below we give example of mathematical objects whose definition involve a certain number of quantifiers. Recall that $\mathbb{Z} = \{\ldots,-2,-1,0,1,2,\ldots,\}$ denotes the set of integers, $\mathbb{N} = \{1,2,3,\ldots\}$ the set of natural numbers.

*Example* 12. Let $n \in \mathbb{Z}$. We say that $n$ is *even* if and only if $n$ is a multiple of 2, i.e., $n = 2k$ for some $k \in \mathbb{Z}$. Here it is not clearly explicit what the quantifier is, but another way to say that an integer is even is as follows: $n$ is even if and only if there exists $k \in \mathbb{Z}$ such that $n = 2k$. More formally,

$$n \in \mathbb{Z} \text{ is even} \iff (\exists k \in \mathbb{Z}) \quad n = 2k.$$

*Example* 13. Let $n \in \mathbb{Z}$. We say that $n$ is *odd* if and only if $n = 2k+1$ for some $k \in \mathbb{Z}$, or equivalently, if and only if here exists $k \in \mathbb{Z}$ such that $n = 2k+1$. More formally,

$$n \in \mathbb{Z} \text{ is odd} \iff (\exists k \in \mathbb{Z}) \quad n = 2k+1.$$

## 1.4 Statements with mixed quantifiers

It is very common for statements or definitions to involve more than one quantifier.

*Example* 14. We say that $q$ is a *rational number* (denoted $q \in \mathbb{Q}$) if and only if $n = \frac{p}{q}$ for some $p \in \mathbb{Z}$ and $q \in \mathbb{N}$. More formally,

$$q \in \mathbb{Q} \iff (\exists p \in \mathbb{Z})(\exists q \in \mathbb{N}) \quad q = \frac{p}{q}.$$

*Example* 15. The statement of the Fundamental Theorem of Arithmetic which says that every natural number is a product of prime number can be written as follows.

$$(\forall n \in \mathbb{N})\ (n = 1) \vee [(\exists k \in \mathbb{N})\ (\forall i \in \{1, \ldots, k\})\ (\exists p_i \text{ a prime number}) \quad n = p_1 \cdots p_k]$$

When a statement involves several quantifiers the order usually matters and one cannot swap quantifiers without care! Let $P(x,y)$ be a predicate with two variables. The statement

$$(\forall x)(\exists y)P(x,y)$$

is in general not logically equivalent to the statement

$$(\exists y)(\forall x)P(x,y).$$

For instance, "For all odd number $n$ there exists a number $k \in \mathbb{Z}$ such that $n = 2k+1$" is a true statement, while "there exists a number $k \in \mathbb{Z}$ such that for all odd number $n$, $n = 2k+1$" is clearly a false statement.

*Example* 16. The definition of the limit of a convergent sequence involves three quantifiers. Let $\ell$ be a fixed real number and $(x_n)_{n=1}^{\infty}$ be a sequence of real numbers. We say that $(x_n)_{n=1}^{\infty}$ converges to $\ell$, and we write $\lim_{n\to\infty} x_n = \ell$, if for all $\varepsilon > 0$ there exists a natural number $N$ such that if $n \geq N$ then $|x_n - \ell| < \varepsilon$. Symbolically,

$$\lim_{n\to\infty} x_n = \ell \iff (\forall \varepsilon > 0)(\exists N \in \mathbb{N})(\forall n \geq N)(|x_n - \ell| < \varepsilon).$$

*Example* 17. The definition of the limit of a function at a point involves three quantifiers. Let $x_0 \in (a,b)$, $\ell \in \mathbb{R}$ and $f\colon (a,x_0) \cup (x_0,b) \to \mathbb{R}$. We say that $\ell$ is the limit of $f$ at $x_0$, and we write $\lim_{x\to x_0} f(x) = \ell$, if for all $\varepsilon > 0$ there exists $\delta > 0$ such that if $x$ satisfies $0 < |x - x_0| < \delta$ then $|f(x) - \ell| < \varepsilon$. Symbolically,

$$\lim_{x\to x_0} f(x) = \ell \iff (\forall \varepsilon > 0)(\exists \delta > 0)(\forall x)[0 < |x - x_0| < \delta \implies |f(x) - \ell| < \varepsilon].$$

# Chapter 2

# Classical Proof Techniques

## 2.1 Modus Ponens and Modus Tollens

From the implication truth table we can deduce two elementary rules of inference. Recall that the truth table of the implication is:

| $P$ | $Q$ | $P \implies Q$ |
|---|---|---|
| T | T | T |
| F | T | T |
| T | F | F |
| F | F | T |

Modus Ponens is a logical argument that exploits the first row in the implication truth table and says that "$Q$ is true" is a valid conclusion based on the hypotheses that "$P$ is true" and "$P \implies Q$ is true".

*Example* 18. You know from your Calculus course that $P(f) \implies Q(f)$ is true where,

$P(f)$ :The function $f$ is differentiable at 0,

$Q(f)$ :The function $f$ is continuous at 0.

Therefore you can conclude that a function $f$ is continuous at 0 if you know that $f$ is differentiable at 0.

Modus Tollens is a logical argument that exploits the last row in the implication truth table and says that "$P$ is not true" is a valid conclusion based on the hypotheses that "$Q$ is not true" and "$P \implies Q$ is true".

*Example* 19. You know from your Calculus course that $P(f) \implies Q(f)$ is true where,

$P(f)$ :The function $f$ is differentiable at 0,

$Q(f)$ :The function $f$ is continuous at 0.

Therefore you can conclude that a function $f$ is not differentiable at 0 if you know that $f$ is not continuous at 0.

## 2.2 Proofs of existential statements

Recall that in practice it is usually simpler and more convenient to use the same letter for the variable and its assigned value and we will adopt this convention.

### 2.2.1 Existential statements of the form $(\exists x \in \mathscr{U})P(x)$

For existential statements we proceed as follows.

**Proof Technique 1: Existential statements $(\exists x \in \mathscr{U})P(x)$**

To prove directly that a statement of the form $(\exists x \in \mathscr{U})P(x)$ is true we proceed as follows:

- We must find, or simply exhibit, an element $x \in \mathscr{U}$ and demonstrate that $P(x)$ is true.

*Example* 20. Show that there exists $x \in \mathbb{R}$ such that $2x+1=0$.

*Proof of Example 20.* Let $x = -\frac{1}{2}$. Then,

$$2x+1 = 2(-\frac{1}{2})+1 = -1+1 = 0.$$

Therefore, there exists $x \in \mathbb{R}$ such that $2x+1=0$. □

*Example* 21. Show that there exists $x \in \mathbb{R}$ such that $x^2+x-1=0$.

### 2.2.2 Uniqueness in proofs of existential statements

To prove that there exists a unique $x \in \mathscr{U}$ such that $P(x)$ is true we can proceed in two different ways.

**Proof Technique 2: Uniqueness, first approach**

- We first find, or simply exhibit, an element $a \in \mathscr{U}$ and demonstrate that $P(a)$ is true.
- Then, we demonstrate that if $x$ is such that $P(x)$ is true then necessarily $x = a$.

*Example* 22. Prove that there is a unique $x \in \mathbb{R}$ such that $2x+1=0$.

*Proof of Example 22.*

**Existence:** Let $x = -\frac{1}{2}$. Then,

$$2x+1 = 2(-\frac{1}{2})+1 = -1+1 = 0.$$

Therefore, there is at least an $x \in \mathbb{R}$ such that $2x+1=0$.

**Uniqueness:** Assume for a moment that there is another $y \in \mathbb{R}$ such that $2y+1=0$. Then, $2y=-1$, and hence $y=-\frac{1}{2}=x$.

**Conclusion:** There exists a unique $x \in \mathbb{R}$ such that $2x+1=0$.

□

It is important to note that in this first approach for proving uniqueness statements the second step (uniqueness part) makes a reference to the first step (existence part), and we must first prove the existence of an element, and then prove its uniqueness.

There is another approach for proving uniqueness statements.

**Proof Technique 3: Uniqueness, second approach**

- We find, or simply exhibit, an element $a \in \mathscr{U}$ and demonstrate that $P(a)$ is true.
- We prove that if $x, y$ are such that if $P(x)$ and $P(y)$ are true then $x = y$,

We now give a different proof of Example 22

*Alternate proof of Example 22.*

**Uniqueness:** Assume for a moment that there are $y \in \mathbb{R}$ and $z \in \mathbb{R}$ such that $2y+1=0$ and $2z+1=0$. Then, $2y+1=2z+1$, and hence $2y=2z$. Simplifying by 2 on both sides we have $y=z$. Therefore, there is at most one $x \in \mathbb{R}$ such that $2x+1=0$.

**Existence:** Let $x = -\frac{1}{2}$. Then,

$$2x+1 = 2(-\frac{1}{2})+1 = -1+1 = 0.$$

Therefore, there is at least an $x \in \mathbb{R}$ such that $2x+1=0$.

**Conclusion:** There exists a unique $x \in \mathbb{R}$ such that $2x+1=0$.

□

Note that in this second approach the two steps (uniqueness and existence) can be performed independently of each other as none of them makes a reference to the other one.

*Example* 23. Prove that the equation $x^2+2x+1=0$ has a unique solution.

Let $P(x)$ be a predicate. The two approaches to prove the statement "there exists a unique $x$ such that $P(x)$" can be formally expressed using the following first order logic formulas. For the first approach

$$(\exists x)[P(x) \wedge ((\forall y)[P(y) \implies (x = y)])] \tag{2.1}$$

and for the second approach

$$[(\exists x)P(x)] \wedge [(\forall y)(\forall z)[(P(y) \wedge P(z)) \implies (y = z)]]. \tag{2.2}$$

Both logical formulas (2.1) and (2.2) can be shown to be logically equivalent and they are abbreviated as $(\exists! x)P(x)$. As you might have noticed, in the proof of the uniqueness parts we are implicitly proving universal statements.

## 2.3 Proofs of universal statements

A universal statement is a statement of the form $(\forall x)P(x)$ where $P(x)$ is some given predicate. We discuss two very common occurrences of universal statements.

### 2.3.1 Universal statements of the form $(\forall x \in \mathscr{U})P(x)$

We first describe how to prove statements with universal quantifiers.

**Proof Technique 4: Direct proof of $(\forall x \in \mathscr{U})P(x)$**

To prove directly that a statement of the form $(\forall x \in \mathscr{U})P(x)$ is true we proceed as follows:

- We begin with "Let $x$ be a fixed element of $\mathscr{U}$."
- Then, we must demonstrate that $P(x)$ is true.
- Finally, we must check that no restriction other than being in $\mathscr{U}$ has been imposed on $x$ and thus our proof is valid for an arbitrary choice of $x \in \mathscr{U}$. If this is the case, we could conclude by saying that $x$ was fixed but arbitrary.

For the following example we need to define the notion of odd number.

**Definition 17: Odd numbers**

Let $n$ be an integer. We say that $n$ is odd if there exists an integer $k$ such that $n = 2k+1$. Formally,

$$n \text{ is odd} \iff (\exists k \in \mathbb{Z})(n = 2k+1)$$

*Example* 24. Prove that for all $n \in \mathbb{Z}$, $6n+5$ is odd.

*Proof of Example 24.* Let $n \in \mathbb{Z}$ be fixed. Then,

$$6n+5 = 6n+4+1 = 2(3n+2)+1 = 2k+1,$$

where $k = 3n+2 \in \mathbb{Z}$. Therefore, $6n+5$ is odd. Since $n \in \mathbb{Z}$ was fixed but arbitrary, if follows that for all $n \in \mathbb{Z}$, $6n+5$ is odd. □

Mathematicians will usually write this proof in a more condensed way by leaving a few things implicit. They would simply write:

Let $n \in \mathbb{Z}$. Then,

$$6n+5 = 6n+4+1 = 2(3n+2)+1 = 2k+1,$$

where $k = 3n+2 \in \mathbb{Z}$. Therefore, $6n+5$ is odd and if follows that for all $n \in \mathbb{Z}$, $6n+5$ is odd.

### 2.3.2 Statements of the form $(\forall x \in \mathscr{U})[P(x) \implies Q(x)]$

Many of the statements we will have to prove are of the form $(\forall x \in \mathscr{U})[P(x) \implies Q(x)]$.

**Proof Technique 5: Direct proof of $(\forall x \in \mathscr{U})[P(x) \implies Q(x)]$**

According to the implication truth table, to prove directly that a statement of the form $(\forall x \in \mathscr{U})[P(x) \implies Q(x)]$ is true we proceed as follows:

- We begin with "Let $x \in \mathscr{U}$, such that $P(x)$ is true, be fixed".
- Then, we must demonstrate that $Q(x)$ is true.
- Finally, we must check that no restriction other than being in $\mathscr{U}$ and satisfying $P$ has been imposed on $x$ and thus our proof is valid.

For the following example we need to define the notion of even number.

**Definition 18: Even numbers**

Let $n$ be an integer. We say that $n$ is even if there exists an integer $k$ such that $n = 2k$. Formally,

$$n \text{ is even} \iff (\exists k \in \mathbb{Z})(n = 2k)$$

*Example* 25. Prove that for all integer $n$ if $n$ is even, then $n^2+5n+2$ is even.

*Proof of Example 25.* Let $n \in \mathbb{Z}$ be a fixed even integer. Then, there exists $k \in \mathbb{Z}$ such that $n = 2k$, and hence

$$n^2+5n+2 = (2k)^2+5(2k)+2 = 4k^2+10k+2 = 2(2k^2+5k+1) = 2r,$$

where $r = 2k^2+5k+1 \in \mathbb{Z}$. Therefore, $n^2+5n+2$ is even. Since $n \in \mathbb{Z}$ was fixed but arbitrary even integer, if follows that for all $n \in \mathbb{Z}$, $n^2+5n+2$ is even. □

Again a seasoned mathematicians will simply write the same proof as follows: Let $n \in \mathbb{Z}$ such that $n = 2k$ for some $k \in \mathbb{Z}$. Then,

$$n^2 + 5n + 2 = (2k)^2 + 5(2k) + 2 = 4k^2 + 10k + 2 = 2(2k^2 + 5k + 1) = 2r,$$

where $r = 2k^2 + 5k + 1 \in \mathbb{Z}$. Therefore $n^2 + 5n + 2$ is odd and it follows that for all $n \in \mathbb{Z}$, $n^2 + 5n + 2$ is odd.

The mechanism of the proof technique above can be adjusted to handle statements involving several universal quantifiers and implications where the assumption in the implication does not necessarily involve the variables. For the following example we need to define the notion of divisibility.

**Definition 19: Divisibility**

Let $n$ be an integer. We say that $n$ is divisible by the integer $k$ (or that $k$ divides $n$), and we write $k \mid n$, if there exists an integer $r$ such that $n = rk$. Formally,

$$k \mid n \iff (\exists r \in \mathbb{Z})(n = rk)$$

*Example* 26. Let $a$ and $b$ be integers. Prove that for all integers $m$ and $n$, if $7 \mid a$ and $7 \mid b$, then $7 \mid (am + bn)$.

### 2.3.3 Disproving universal statements: counterexamples

The negation of the statement $(\forall x)P(x)$ is the statement $(\exists x)\neg P(x)$. Therefore to show that a statement of the form $(\forall x)P(x)$ is false we need to find an assignment of $x$ (still denoted by $x$) such that $P(x)$ is false.

*Terminology.* An assignment of the variable $x$ such that $\neg P(x)$ is true, is called a counterexample for the statement $(\forall x)P(x)$.

We now discuss how to disprove some of the most common universal statements.

**Proof Technique 6: Disproving $(\forall x \in \mathscr{U})P(x)$**

To prove that a statement of the form $(\forall x \in \mathscr{U})P(x)$ is false we proceed as follows:

- We find an assignment of the variable $x \in \mathscr{U}$ (still denoted by $x$) such that $P(x)$ is false.

For the following example we need to define the notion of prime number.

**Definition 20: Prime numbers**

Let $p$ be a natural number. We say that $p$ is a prime number if it is only divisible by 1 and $p$ itself. Formally,

$p$ is a prime number $\iff$
$[(p \in \mathbb{N}) \wedge (p > 1) \wedge [(\forall m \in \mathbb{N})(\forall n \in \mathbb{N})[p = mn \implies ((m = 1) \vee (n = 1))]]$.

*Example* 27. Is the following statement true or false?
For all positive integers $n$, $n^2 + n + 41$ is prime.

The negation of the statement $(\forall x \in \mathscr{U})[P(x) \implies Q(x)]$ is the statement $(\exists x \in \mathscr{U})\neg[P(x) \implies Q(x)]$, and thus the negation of $(\forall x \in \mathscr{U})[P(x) \implies Q(x)]$ is logically equivalent to $(\exists x \in \mathscr{U})P(x) \wedge \neg Q(x)$. Note that the negation of an implication is *not* an implication!

**Proof Technique 7: Disproving $(\forall x \in \mathscr{U})[P(x) \implies Q(x)]$**

To prove that a statement of the form $(\forall x \in \mathscr{U})[P(x) \implies Q(x)]$ is false we proceed as follows:

- We find an assignment of the variable $x \in \mathscr{U}$ (still denoted by $x$) such that $P(x)$ is true and $Q(x)$ is false.

*Example* 28. Prove or disprove that for all integer $n$, if $n$ is even then $n^2 + 1$ is even.

## 2.4 Proofs by contrapositive

The statement forms $P \implies Q$ and $\neg Q \implies \neg P$ are logically equivalent.

**Proof Technique 8: Proving the contrapositive**

To prove $P \implies Q$ one may choose instead to prove $\neg Q \implies \neg P$.

*Example* 29. Let $n$ be an integer. If $n^3$ is odd, then $n$ is odd.

*Example* 30. For this example you can use the following fact that will be proven later: 5 does not divides $n$ if and only if there exists an integer $k$ and an integer $i \in \{1,2,3,4\}$ such that $n = 5k + i$.
Prove that for every integer $n$, if 5 divides $n^2$ then 5 divides $n$.

## 2.5 Proof by contradiction

A proof by contradiction is based on the observation that the statement form $(\neg P) \implies [Q \wedge \neg Q]$ is logically equivalent to $P$.

| $P$ | $Q$ | $Q \wedge \neg Q$ | $\neg P$ | $(\neg P) \implies [Q \wedge \neg Q]$ |
|---|---|---|---|---|
| T | T | F | F | T |
| T | F | F | F | T |
| F | T | F | T | F |
| F | F | F | T | F |

Therefore, in order to prove a statement $P$, for example, we could assume that $P$ is false and deduce a statement that we know is false (like $0 = 1$ or $\frac{1}{2}$ is an integer...).

**Proof Technique 9: Proof by contradiction**

To prove a statement $P$ is true by contradiction we proceed as follows:

- We begin first with "Assume $\neg P$ is true for the sake of contradiction."
- Then, we deduce a contradiction.
- Finally, we conclude that $P$ must be true.

*Example* 31. Prove that there does not exist integers $m$ and $n$ such that $15m + 5n = 81$.

*Example* 32. Let $x \in \mathbb{R}$. If for all $\varepsilon > 0$, $|x| < \varepsilon$, then $x = 0$.

A classical use of a proof by contradiction allows us to show that some real numbers are irrational.

**Theorem 6: Irrationality of $\sqrt{2}$**

The real number $\sqrt{2}$ is irrational.

Recall that a number $x$ is irrational if it is not rational, *i.e.,*

$$\neg \left[ (\exists p \in \mathbb{N})(\exists q \in \mathbb{Z}^+) \left[ \frac{p}{q} = x \right] \right]$$

Another celebrated proof by contradiction is a proof of Euclid's Theorem. Euclid's Theorem says that there are infinitely many prime numbers. We recall the formal definition of a prime number.

**Definition 21: Prime numbers**

A natural number $p$ is prime if

$$p > 1 \text{ and } (\forall m, n \in \mathbb{N})[p = mn \implies (m = 1 \vee n = 1)].$$

For now we will assume the Fundamental Theorem of Arithmetic, but we will prove it later once we have learned what a proof using strong induction is.

**Theorem 7: Fundamental Theorem of Arithmetic**

Every positive integer greater than 1 can be written as a product of primes. Furthermore, this product of primes is unique, except for the order in which the factors appear.

**Theorem 8: Euclid's Theorem**

There are infinitely many prime numbers.

## 2.6 Other useful proof techniques

### 2.6.1 Proving biconditional statements

Since $P \iff Q$ is logically equivalent to $(P \implies Q) \wedge (Q \implies P)$, in order to prove that a statement of the form $P \iff Q$ is true, we need to prove that $P \implies Q$ AND that $Q \implies P$.

**Proof Technique 10: Biconditional statements $P \iff Q$**

1. Prove $P \implies Q$

   and

2. prove $Q \implies P$.

*Example* 33. Prove that for all integer $n$,

$$n \text{ is even} \iff n+2 \text{ is even.}$$

*Example* 34. Prove that for all numbers $x, y \in \mathbb{R}$ with $y \geq 0$, $|x| \leq y$ if and only if $-y \leq x \leq y$.

### 2.6.2 Proving disjunction statements

Let $P$ and $Q$ be statements. To prove disjunction statements we can use the observation that $P \vee Q$, $\neg P \implies Q$, and $\neg Q \implies P$ are logically equivalent.

| $P$ | $Q$ | $P \vee Q$ | $\neg P \implies Q$ | $\neg Q \implies P$ |
|---|---|---|---|---|
| T | T | T | T | T |
| T | F | T | T | T |
| F | T | T | T | T |
| F | F | F | F | F |

**Proof Technique 11: Proving disjunction statements**

To prove that a statement of the form $P \vee Q$ is true, we may choose either one of the following to options:

1. Assume $\neg P$ and prove $Q$,

   or

2. assume $\neg Q$ and prove $P$.

*Example* 35. Prove that for all real numbers $x$ and $y$ with $y \geq 0$, if $x^2 \geq y$, then $x \geq \sqrt{y}$ or $x \leq -\sqrt{y}$

### 2.6.3 Proof by cases

*Example* 36. Prove that for all integer $k$, $k(k+1)$ is even.

*Example* 37. Prove that for all real numbers $x$ and $y$, $|x+y| \leq |x| + |y|$.

*Hint.* □

### 2.6.4 Working backwards

This type of technique is usually applied when proving inequalities. Here is an example.

*Example* 38. Prove that for every positive real number $x$, one has $\frac{x}{x+1} < \frac{x+1}{x+2}$.

We need to understand why this inequality holds but when we write our proof the inequality should only be written as the conclusion of our proof. If we want to prove that $\frac{x}{x+1} < \frac{x+1}{x+2}$ for all $x > 0$, then it would equivalent to proving the inequality $(x+1)(x+2) < x(x+1)$

# Chapter 3

# Induction

## 3.1 Principle of Mathematical Induction

The principle of mathematical induction is a very powerful tool to deal with infinite objects and to prove rigorously infinitely many (in the sense that they can be enumerated) statements.

**Theorem 9: Principle of Mathematical Induction**

Let $P(n)$ be a predicate where the variable takes integer values. Suppose that there exists $k_0 \in \mathbb{Z}$ such that

$P(k_0)$ is true (the base case)

and

for all $k \geq k_0$, $P(k+1)$ is true under the assumption that $P(k)$ is true (the induction step),

then for all $k \geq k_0$ $P(k)$ is true (the conclusion).

*Proof.* Follows from the Induction Axiom applied to the set

$$Y := \{n \in \mathbb{N} | P(k_0 + n) \text{ is true}\}.$$

□

The principle of mathematical induction is most commonly used with $k_0 = 0$ or $k_0 = 1$.

*Example* 39. Show that for all integers $n \geq 1$, $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.

*Solution:* For all integers $n \geq 1$, let $P(n) : \sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.

**Base case:** Since $\sum_{i=1}^{1} i = 1$ and $\frac{1(1+1)}{2} = 1$, one has that $\sum_{i=1}^{1} i = \frac{1(1+1)}{2}$ and $P(1)$ is true.

**Induction step:** Let $k \geq 1$ and assume that $P(k)$ is true, i.e. we assume that $\sum_{i=1}^{k} i = \frac{k(k+1)}{2}$. Then,

$$\begin{aligned}\sum_{i=1}^{k+1} i &= \sum_{i=1}^{k} i + (k+1) \\ &= \frac{k(k+1)}{2} + (k+1) \text{ (by the induction hypothesis)} \\ &= \frac{(k+1)(k+2)}{2},\end{aligned}$$

and hence $P(k+1)$ is true.

**Conclusion:** By the Principle of Mathematical Induction, one can conclude that $\forall n \geq 1$, $P(n)$ is true, which means that for all $n \geq 1$, $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.

□

*Example* 40. Show that the following equalities hold.

1. for all $n \geq 1$, $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$.
2. for all $n \geq 1$, $\sum_{i=1}^{n} i^3 = \frac{n^2(n+1)^2}{4}$.
3. for all $n \geq 1$, $\sum_{i=1}^{n} (2i)^2 = \frac{2n(2n+1)(2n+2)}{6}$.

*Solutions.* 1. Let $P(n)$ be the statement "$\sum_{i=1}^{n} (2i-1) = n^2$".

**Base case** Since $1 = 1^2$, $P(1)$ is true.

**Induction Step** Assume that $P(n)$ is true, i.e. we assume that $\sum_{i=1}^{n} (2i-1) = n^2$. Then,

$$\begin{aligned}\sum_{i=1}^{n+1} (2i-1) &= \sum_{i=1}^{n} (2i-1) + (2(n+1)-1) \\ &= n^2 + (2n+1) \text{ (by the induction hypothesis)} \\ &= (n+1)^2,\end{aligned}$$

and hence $P(n+1)$ is true. By the Principle of Mathematical Induction, one can conclude that $\forall n \geq 1$, $P(n)$ is true, which means that for all $n \geq 1$, $\sum_{i=1}^{n} (2i-1) = n^2$.

2. Let $P(n)$ be the statement "$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$".

**Base case** Since $1^2 = \frac{1(1+1)(2+1)}{6}$, $P(1)$ is true.

**Induction Step** Assume that $P(n)$ is true, i.e. we assume that $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$. Then,

$$\begin{aligned}
\sum_{i=1}^{n+1} i^2 &= \sum_{i=1}^{n} i^2 + (n+1)^2 \\
&= \frac{n(n+1)(2n+1)}{6} + (n+1)^2 \text{ (by the induction hypothesis)} \\
&= \frac{n(n+1)(2n+1)+6(n+1)^2}{6} \\
&= \frac{(n+1)(n(2n+1)+6(n+1))}{6} \\
&= \frac{(n+1)(2n^2+7n+6)}{6} \\
&= \frac{(n+1)(n+2)(2n+3)}{6} \\
&= \frac{(n+1)((n+1)+1)(2(n+1)+1)}{6}
\end{aligned}$$

and hence $P(n+1)$ is true. By the Principle of Mathematical Induction, one can conclude that $\forall n \geq 1$, $P(n)$ is true, which means that for all $n \geq 1$, $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$.

3. exercise
4. exercise
5. exercise

□

## 3.2 Principle of Strong Mathematical Induction

**Theorem 10: Principle of Strong Mathematical Induction**

Let $P(n)$ be a predicate where the variable takes integer values. Suppose that there exists an integer $k_0$ such that

$P(k_0)$ is true (the base case),

and

for all $k \geq k_0$, $P(k+1)$ is true under the assumption that for all $r \in \{k_0, k_0+1, \ldots, k\}$ $P(r)$ is true (the induction step),

then for all $n \geq k_0$ $P(n)$ is true (the conclusion).

**Theorem 11: Fundamental Theorem of Arithmetic**

Every positive integer greater than 2 can be written as a product of primes. Furthermore, this product of primes is unique, except for the order in which the factors appear.

*Proof.* Formally the statement says that for all integer $n \geq 2$ there exists $p_1, p_2, \ldots, p_k$ prime numbers for some $k \in \mathbb{N}$ such that $n = p_1 p_2 \cdots p_k$. We will show that it is indeed true using the principle of strong mathematical induction. For $n = 2$ the statement is clearly true since 2 is a prime number. Let $n \geq 2$ and assume that for all integer $r$ such that $2 \leq r \leq n$, $r$ is a product of prime numbers. If $n+1$ is prime then the conclusion holds. If $n+1$ is not prime then there are integers $1 < a < n+1$ and $1 < b < n+1$ such that $n+1 = ab$ . Since $2 \leq a \leq n$ and $2 \leq b \leq n$, $a$ and $b$ are products of prime numbers, say $a = p_1 p_2 \cdots p_k$ and $b = q_1 q_2 \cdots q_s$ for some prime numbers $p_1, p_2, \ldots, p_k, q_1, q_2, \ldots, q_s$. Thus, $n+1 = ab = (p_1 p_2 \cdots p_k)(q_1 q_2 \cdots q_s) = p_1 p_2 \cdots p_k q_1 q_2 \cdots q_s$ which is a product of prime numbers. We conclude by invoking the principle of strong mathematical induction. □

*Exercise* 5. Consider the sequence $(a_n)_{n=1}^{\infty}$ recursively defined as $a_1 = 1$, $a_2 = 5$ and for all $n \geq 2$, $a_{n+1} = a_n + 2a_{n-1}$. Show that for all $n \geq 1$, $a_n = 2^n + (-1)^n$.

*Solution:* For all $n \in \mathbb{N}$, let $P(n)$ be the predicate $a_n = 2^n + (-1)^n$.

**Base case:** Since $a_1 = 1$ and $2^1 + (-1)^1 = 2 - 1 = 1$, one has that $a_1 = 2^1 + (-1)^1$ and $P(1)$ is true.

**Induction step:** Let $k \geq 1$ and assume that for all $r \in \{1, 2, \ldots, k\}$ $P(r)$ is true, i.e. we assume that for all $r \in \{1, 2, \ldots, k\}$ $a_r = 2^r + (-1)^r$. We want to show that $P(k+1)$ is true. In this problem, the case $k = 1$ has to be treated separately. If $k = 1$, observe that $P(2)$ is true (regardless of the truth value of $P(1)$) since $2^2 + (-1)^2 = 5 = a_2$ and thus in particular if $P(1)$ is true then $P(2)$ is true. Otherwise, if $k \geq 2$, assuming $P(1), P(2), \ldots, P(k)$ are true, then

$$\begin{aligned} a_{k+1} &= a_k + 2a_{k-1} \text{ (here we need } k \geq 2 \text{ since } a_0 \text{ is not defined)} \\ &= 2^k + (-1)^k + 2(2^{k-1} + (-1)^{k-1}) \text{ (by the induction hypothesis)} \\ &= 2 \cdot 2^k + (-1)^{k-1}(-1+2) \\ &= 2^{k+1} + (-1)^{k+1} \text{ (since } (-1)^{k+1} = (-1)^{k-1}), \end{aligned}$$

and hence $P(k+1)$ is true.

**Conclusion:** By the Principle of Strong Mathematical Induction, one can conclude that for all $n \geq 1$, $P(n)$ is true, which means that for all $n \geq 1$, $a_n = 2^n + (-1)^n$.

□

The more traditional way to write your solution is as follows.

*Alternate Solution:* For all $n \in \mathbb{N}$, let $P(n)$ be the predicate $a_n = 2^n + (-1)^n$.

Since $a_1 = 1$ and $2^1 + (-1)^1 = 2 - 1 = 1$, one has that $a_1 = 2^1 + (-1)^1$ and $P(1)$ is true. Since $2^2 + (-1)^2 = 5 = a_2$, $P(2)$ is also true. Let $k \geq 2$, and assume $P(1), P(2), \ldots, P(k)$ are true, then

$$\begin{aligned} a_{k+1} &= a_k + 2a_{k-1} \text{ (here we need } k \geq 2 \text{ since } a_0 \text{ is not defined)} \\ &= 2^k + (-1)^k + 2(2^{k-1} + (-1)^{k-1}) \text{ (by the induction hypothesis)} \\ &= 2 \cdot 2^k + (-1)^{k-1}(-1+2) \\ &= 2^{k+1} + (-1)^{k+1} \text{ (since } (-1)^{k+1} = (-1)^{k-1}), \end{aligned}$$

and hence $P(k+1)$ is true. By the Principle of Strong Mathematical Induction, one can conclude that for all $n \geq 1$, $P(n)$ is true, which means that for all $n \geq 1$, $a_n = 2^n + (-1)^n$. □

# Chapter 4

# Introduction to Elementary Set Theory

## 4.1 Sets and subsets

We won't give a formal definition of the notion of a set but we will understand the word set as an undefined term which refers to a collection of objects. The objects in a set are called elements and we use the notation $x \in X$ to express that the element $x$ is in the set $X$. The notion of membership is also not formally defined and is part of the concept of a set. We use the abbreviation $x \notin X$ for $\neg(x \in X)$.

**Axiom** There is a set with no elements which is called the empty set and is denoted by $\emptyset$.

Observe that $x \in \emptyset$ is always false regardless of the element $x$ that is under consideration, and thus $x \notin \emptyset$ is always true.

*Example* 41. Classical sets.

1. $\mathbb{N} := \{1, 2, 3, \dots\}$, the natural numbers.
2. $\mathbb{Z} := \{\dots, -2, -1, 0, 1, 2, \dots\}$, the integers
3. $\mathbb{Q} := \{\frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N}\}$, the rational numbers.
4. $\mathbb{R}$, the real numbers

**Definition 22: Truth set of a predicate**

Let $P(x)$ be a predicate and $\mathscr{U}$ be the ambient set. The set

$$A := \{x \in \mathscr{U} \mid P(x) \text{ is true}\}$$

is called the truth set of the predicate $P(x)$.

*Example* 42.

1. $\{x \in \mathbb{Z} \mid (\exists k \in \mathbb{Z})[x = 5k]\}$
2. $\{x \in \mathbb{R} \mid (x > 0) \wedge (x^2 \in \mathbb{Z}^+)\}$

**Definition 23: Sets of the form $n\mathbb{Z}$**

Let $n \in \mathbb{Z}$. We define a set denoted $n\mathbb{Z}$ as follows:

$$n\mathbb{Z} := \{x \in \mathbb{Z} \mid (\exists k \in \mathbb{Z})[x = nk]\}$$

For instance, $5\mathbb{Z}$ is the set $\{x \in \mathbb{Z} \mid (\exists k \in \mathbb{Z})[x = 5k]\}$ which is also sometimes simply described as $\{5k \mid k \in \mathbb{Z}\}$.

*Example* 43. Let $a < b$ be real numbers. There are 9 types of elementary intervals of real numbers.

1. $[a,b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ the closed interval.
2. $(a,b) = \{x \in \mathbb{R} : a < x < b\}$ the open interval.
3. $(a,b] = \{x \in \mathbb{R} : a < x \leq b\}$ half-open, half-closed
4. $[a,b) = \{x \in \mathbb{R} : a \leq x < b\}$ half-open, half-closed
5. $[a,+\infty) = \{x \in \mathbb{R} : x \geq a\}$ unbounded
6. $(a,+\infty) = \{x \in \mathbb{R} : x > a\}$ unbounded
7. $(-\infty,a] = \{x \in \mathbb{R} : x \leq a\}$ unbounded
8. $(-\infty,a) = \{x \in \mathbb{R} : x < a\}$ unbounded
9. $(-\infty.+\infty) = \mathbb{R}$ unbounded

**Definition 24: Subset**

Let $X$ and $Y$ be sets. We say that $X$ is a subset of $Y$, and write $X \subseteq Y$, if every element of $X$ is also an element of $Y$. Formally,

$$X \subseteq Y \iff (\forall x)[x \in X \implies x \in Y].$$

*Remark* 1. The expression $X \subseteq Y$ is a very convenient abbreviation for the statement $(\forall x)[x \in X \implies x \in Y]$. To prove that $X \subseteq Y$ you need to prove an *implication* with a *universal* quantifier.

*Example* 44. $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$

*Example* 45. We write $X \nsubseteq Y$ for $\neg(X \subseteq Y)$. Give a formal statement expressing $X \nsubseteq Y$.

*Example* 46. Let $X = \{n \in \mathbb{Z} \mid n \text{ is a multiple of } 4\}$ and $Y = \{n \in \mathbb{Z} \mid n \text{ is even}\}$. Prove that $X \subseteq Y$.

**Proposition 2**

Let $X$ be a set. Then,

1. $\emptyset \subseteq X$.
2. $X \subseteq X$

*Proof.*

1. $\emptyset \subset X$ follows from the fact that the implication $x \in \emptyset \implies x \in X$ is always true (i.e. a tautology) since $x \in \emptyset$ is always false (i.e. a contradiction).

2. $X \subseteq X$ follows from the fact that $(x \in X) \implies (x \in X)$ is always true (indeed $P \implies P$ is a tautology).

□

**Proposition 3: Transitivity of the subset relation**

Let $X$, $Y$, and $Z$ be sets. If $X \subseteq Y$ and $Y \subseteq Z$, then $X \subseteq Z$.

*Proof.* (Hint: Direct proof.) Assume that $X \subseteq Y$ and $Y \subseteq Z$. If $X = \emptyset$ then $X \subseteq Z$ holds by Proposition 2. Otherwise, let $x \in X$ then it follows from $X \subseteq Y$ that $x \in Y$. Moreover, it follows from $Y \subseteq Z$ that $x \in Z$. Therefore $X \subseteq Z$. □

**Definition 25: Equality between sets**

We say that two sets $X$ and $Y$ are equal, written $X = Y$, if they have the same elements. Formally,

$$X = Y \iff (\forall x)(x \in X \iff x \in Y).$$

**Proposition 4: Double inclusion**

Let $X$ and $Y$ be sets. Then,

$$X = Y \iff (X \subseteq Y) \wedge (Y \subseteq X).$$

*Proof.* The statement follows from the fact that $x \in X \iff x \in Y$ is logically equivalent to $(x \in X \implies x \in Y) \wedge (x \in Y \implies x \in X)$. □

*Example* 47. Prove that $X = \{n \in \mathbb{Z} \mid n+5 \text{ is odd}\}$ is the set of all even integers.

**Definition 26: Proper subsets**

Let $X$ be a subset of $Y$. We say that $X$ is a proper subset of $Y$, and we write $X \subset Y$, if $X \neq Y$. Formally,

$$X \subset Y \iff (X \subseteq Y) \wedge (X \neq Y).$$

*Example* 48. Show that the set $X = 33\mathbb{Z}$ is a proper subset of $\mathbb{Z}$.

## 4.2 Operation on sets

In this section we describe several natural operations on sets that can be used to create new sets out of given sets.

### 4.2.1 Union and intersection of two sets

We start with the most natural operations on a pair of sets; union and intersection.

**Definition 27: Union of two sets**

Let $X$ and $Y$ be sets. The union of $X$ and $Y$, denoted $X \cup Y$, is the set of all elements that belong to $X$ or to $Y$. Formally,

$$X \cup Y = \{z \mid (z \in X) \vee (z \in Y)\}.$$

Taking the union of two sets provides a set that is "bigger" in the sense that it contains both sets. The following two properties can be deduced from logical principles.

**Proposition 5**

Let $X, Y, Z$ be sets. Then,

1. $X \cup \emptyset = X$
2. $X \cup Y = Y \cup X$ (commutativity of the union operation)
3. $(X \cup Y) \cup Z = X \cup (Y \cup Z)$ (associativity of the union operation)

*Proof.*

1. $X \cup \emptyset = X$ follows from the fact that $(x \in X) \vee (x \in \emptyset)$ is logically equivalent to $(x \in X)$ since $(x \in \emptyset)$ is always false.

2. $X \cup Y = Y \cup X$ follows from the fact that $(z \in X) \vee (z \in Y)$ is logically equivalent to $(z \in Y) \vee (z \in X)$ (indeed $P \vee Q \equiv Q \vee P$).

3. $(X \cup Y) \cup Z = X \cup (Y \cup Z)$ follows from the fact that $((a \in X) \vee (a \in Y)) \vee (a \in Z)$ is logically equivalent to $(a \in X) \vee ((a \in Y) \vee (a \in Z))$ (indeed $(P \vee Q) \vee R \equiv P \vee (Q \vee R)$).

□

**Proposition 6**

Let $X$ and $Y$ be sets. Then,

1. $X \subseteq X \cup Y$
2. $Y \subseteq X \cup Y$
3. $X \subseteq Y \iff X \cup Y = Y$

*Proof.*

1. If $X = \emptyset$ the inclusion holds, otherwise let $x \in X$. Then $x \in X \cup Y$ by definition of the union, and thus $X \subseteq X \cup Y$.

2. If $Y = \emptyset$ the inclusion holds, otherwise let $y \in Y$. Let $y \in Y$. Then $y \in X \cup Y$ by definition of the union, and thus $Y \subseteq X \cup Y$.

3. We first prove $\implies$:

   Assume that $X \subseteq Y$. Observe first that $X \subseteq X \cup Y$ always holds. If $X \cup Y = \emptyset$ the reverse inclusion holds, otherwise let $z \in X \cup Y$. Then either $z \in Y$ or $z \in X$. But in the latter case it follows from $X \subseteq Y$ that $z \in Y$. In all cases $z \in Y$ and thus $X \cup Y \subseteq Y$. Combining the two inclusions we have $X \cup Y = Y$.

   We now prove $\impliedby$:

   Assume that $X \cup Y = Y$. If $X = \emptyset$ the inclusion holds, otherwise let $x \in X$. Then $x \in X \cup Y$ by definition of the union and thus $x \in Y$ follows from the assumption $X \cup Y = Y$. Therefore, $X \subseteq Y$.

□

**Definition 28: Intersection of two sets**

Let $X$ and $Y$ be sets. The intersection of $X$ and $Y$, denoted $X \cap Y$, is the set is the set of all elements that belong to $X$ and to $Y$. Formally,

$$X \cap Y = \{z \mid (z \in X) \wedge (z \in Y)\}.$$

Taking the intersection of two sets provides a set that is "smaller" in the sense that it is contained in both sets. The following two properties can be deduced from logical principles.

**Proposition 7**

Let $X, Y, Z$ be sets. Then,

1. $X \cap \emptyset = \emptyset$
2. $X \cap Y = Y \cap X$ (commutativity of the intersection operation)
3. $(X \cap Y) \cap Z = X \cap (Y \cap Z)$ (associativity of the union operation)

*Proof.*

1. $X \cap \emptyset = \emptyset$ follows from the fact that $(x \in X) \wedge (x \in \emptyset)$ is always false since $(x \in \emptyset)$ is always false.
2. $X \cap Y = Y \cap X$ follows from the fact that $(z \in X) \wedge (z \in Y)$ is logically equivalent to $(z \in Y) \wedge (z \in X)$ (indeed $P \wedge Q \equiv Q \wedge P$).
3. $(X \cap Y) \cap Z = X \cap (Y \cap Z)$ follows from the fact that $((a \in X) \wedge (a \in Y)) \wedge (a \in Z)$ is logically equivalent to $(a \in X) \wedge ((a \in Y) \wedge (a \in Z))$ (indeed $(P \wedge Q) \wedge R \equiv P \wedge (Q \wedge R)$).

□

**Proposition 8**

Let $X$ and $Y$ be sets. Then,

1. $X \cap Y \subseteq X$,
2. $X \cap Y \subseteq Y$,
3. $X \subseteq Y \iff X \cap Y = X$.

*Proof.*

1. If $X \cap Y = \emptyset$ then $X \cap Y \subseteq X$. Otherwise let $z \in X \cap Y$, and then $z \in X$ by definition of the intersection, and thus $X \cap Y \subseteq X$.
2. If $X \cap Y = \emptyset$ then $X \cap Y \subseteq Y$. Otherwise let $z \in X \cap Y$, and then $z \in Y$ by definition of the intersection, and thus $X \cap Y \subseteq Y$.
3. We first prove $\Longrightarrow$:

   Assume that $X \subseteq Y$. Observe first that $X \cap Y \subseteq X$ always holds. If $X = \emptyset$ the reverse inclusion holds, otherwise let $z \in X$. Then $z \in Y$ follows from the assumption $X \subseteq Y$, and hence $z \in X \cap Y$. Therefore $X \subseteq X \cap Y$ and combining the two inclusions we have $X \cup Y = Y$.

   We now prove $\Longleftarrow$:

Assume that $X \cap Y = X$. If $X = \emptyset$ the inclusion holds, otherwise let $x \in X$. Then it follows from the assumption $X \cap Y = X$ that $x \in X \cap Y$, and hence $x \in Y$ by definition of the intersection. Therefore, $X \subseteq Y$.

□

**Definition 29: Disjoint sets**

We say that two sets $X$ and $Y$ are *disjoint* if they have no element in common, or equivalently if their intersection is the empty set. Formally,

$$X \text{ and } Y \text{ are disjoint} \iff X \cap Y = \emptyset.$$

The distributivity properties of the union operation over the intersection operation, and vice versa, follow from the two logical equivalences $P \wedge (Q \vee R) \equiv (P \wedge Q) \vee (P \wedge R)$ and $P \vee (Q \wedge R) \equiv (P \vee Q) \wedge (P \vee R)$, but you could try to write "double-inclusion mathematician's proofs".

**Proposition 9: Distributivity Properties**

Let $X, Y, Z$ be sets. Then,

1. $X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$,
2. $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$.

### 4.2.2 Complement

The notion of complement is described in this section.

**Definition 30: Complement**

Let $X$ and $Y$ be subsets of some ambient set $U$. The complement of $X$ in $Y$, denoted $Y \setminus X$, is the set of elements that are in $Y$ but not in $X$. Formally,

$$Y \setminus X = \{z \in U \mid (z \in Y) \wedge (z \notin X)\}.$$

For convenience, the set $U \setminus X$ will be simply denoted by $\overline{X}$, and called the complement of $X$. Formally,

$$\overline{X} = \{z \in U \mid z \notin X\}.$$

*Remark* 2.

1. The set $Y \setminus X$ is always a subset of $Y$.
2. We always have that $x \in \overline{X} \iff x \notin X$.

3. The definition of the complement of $X$ in $Y$ does NOT assume that either set be a subset of the other.

4. The alternative notation $Y - X$ is also used for $Y \setminus X$.

In the following proposition we record some elementary properties that can be obtained from logical principles.

**Proposition 10**

Let $X$ and $Y$ be subsets of a universal set $U$. Then

1. $\overline{U} = \emptyset$,
2. $\overline{\emptyset} = U$.
3. $X \setminus Y = X \cap \overline{Y}$,
4. $\emptyset \setminus X = \emptyset$,
5. $X \setminus \emptyset = X$.
6. $\overline{\overline{X}} = X$.

*Proof.*

1. $\overline{U} = \{z \in U : z \notin U\} = \emptyset$, since $(z \in U) \wedge (z \notin U)$ is always false.
2. $\overline{\emptyset} = \{z \in U : z \notin \emptyset\} = U$, since $(z \in U) \wedge (z \notin \emptyset)$ is always true.
3. $X \setminus Y = \{z \in U : (z \in X) \wedge (z \notin Y)\} = \{z \in U : (z \in X) \wedge (z \in \overline{Y})\} = X \cap \overline{Y}$, (by definition of the complement and the intersection).
4. $\emptyset \setminus X = \{z \in \emptyset : z \notin X\} = \emptyset$, since $(z \in \emptyset) \wedge (z \notin X)$ is always false.
5. $X \setminus \emptyset = \{z \in U : (z \in X) \wedge (z \notin \emptyset)\} = \{z \in U : z \in X\} = X$, since $z \notin \emptyset$ is always true.
6. $\overline{\overline{X}} = \{z \in U : z \notin \overline{X}\} = \{z \in U : z \in X\} = X$ (by definition of the complement).

□

Taking complements reverse the inclusion relationship.

**Proposition 11: Complements of subsets**

Let $X$ and $Y$ be subsets of some universal set $U$. Then $X \subseteq Y$ if and only if $\overline{Y} \subseteq \overline{X}$.

*Proof.* We first prove the "if" part. Assume that $\overline{Y} \subseteq \overline{X}$. If $X = \emptyset$ then $X \subseteq Y$. Otherwise, let $x \in X$ then $x \notin \overline{X}$ by definition of the complement, and it follows from our assumption that $x \notin \overline{Y}$. Therefore, $x \in Y$ and thus $X \subseteq Y$.

The proof of the "only if" part goes as follows. Assume that $X \subseteq Y$. If $\overline{Y} = \emptyset$ then $\overline{Y} \subseteq \overline{X}$. Otherwise, let $z \in \overline{Y}$ then $z \notin Y$ by definition of the complement, and hence $z \notin X$ by our assumption. Therefore, $z \in \overline{X}$ and thus $\overline{Y} \subseteq \overline{X}$.

□

We now prove De Morgan's laws, which state that the complement of the union is the intersection of the complements, and that the complement of the intersection is the union of the complements.

**Theorem 12: DeMorgan's Laws**

Let $X$ and $Y$ be subsets of a universal set $U$. Then

1. $\overline{X \cup Y} = \overline{X} \cap \overline{Y}$,
2. $\overline{X \cap Y} = \overline{X} \cup \overline{Y}$.

*Proof.*

1. We first prove the inclusion $\overline{X \cup Y} \subseteq \overline{X} \cap \overline{Y}$. If $\overline{X \cup Y} = \emptyset$ then the inclusion holds, otherwise let $z \in \overline{X \cup Y}$. Then $z \notin X \cup Y$ (by definition of the complement), and it follows that $z \notin X$ and $z \notin Y$ (by definition of the union). Thus, $z \in \overline{X}$ and $x \in \overline{Y}$ (by definition of the complement), which means that $z \in \overline{X} \cap \overline{Y}$ (by definition of the intersection).

   For the reverse inclusion, if $\overline{X} \cap \overline{Y} = \emptyset$ then the inclusion holds, otherwise let $z \in \overline{X} \cap \overline{Y}$. Then $z \in \overline{X}$ and $z \in \overline{Y}$ (by definition of the intersection), and thus $z \notin X$ and $z \notin Y$ (by definition of the complement). It follows that $z \notin X \cup Y$ (by definition of the union), and hence $z \in \overline{X \cup Y}$ (by definition of the complement).

   Therefore, it follows from the definition of equality between sets that $\overline{X \cup Y} = \overline{X} \cap \overline{Y}$.

2. We first prove the inclusion $\overline{X \cap Y} \subseteq \overline{X} \cup \overline{Y}$. If $\overline{X \cap Y} = \emptyset$ then the inclusion holds, otherwise let $z \in \overline{X \cap Y}$. Then $z \notin X \cap Y$ (by definition of the complement), and it follows that $z \notin X$ or $z \notin Y$ (by definition of the intersection). Thus, $z \in \overline{X}$ or $z \in \overline{Y}$ (by definition of the complement), which means that $z \in \overline{X} \cup \overline{Y}$ (by definition of the union). We just proved that $\overline{X \cap Y} \subseteq \overline{X} \cup \overline{Y}$. For the reverse inclusion, if $\overline{X} \cup \overline{Y} = \emptyset$ then the inclusion holds, otherwise let $z \in \overline{X} \cup \overline{Y}$. Then $z \in \overline{X}$ or $z \in \overline{Y}$ (by definition of the union), and thus $z \notin X$ or $x \notin Y$ (by definition of the complement). It follows that $z \notin X \cap Y$ (by definition of the intersection), and hence $z \in \overline{X \cap Y}$ (by definition of the complement). This shows the reverse inclusion.

□

*Remark* 3. The proof above of the De Morgan's laws, which is a basic-double inclusion proof and uses the logic of connectives implicitly, is the typical proof a mathematician would write. However, a logician will argue that the equality holds since he, or she, will recognize that the truth of the statement follows from the logical equivalence of two statement forms. Indeed, consider the predicates $P(z)$ : "$z \in X$" and $Q(z)$ : "$z \in Y$". From the logic standpoint $\overline{X \cup Y} = \overline{X} \cap \overline{Y}$ is actually a convenient abbreviation for the proposition

$$(\forall z \in U)[\neg(P(z) \vee Q(z)) \iff (\neg P(z) \wedge \neg Q(z))].$$

But, we have proven that $\neg(P \vee Q)$ is logically equivalent to $(\neg P \wedge \neg Q)$ no matter what statements are substituted for $P$ and we can conclude that the equality actually holds! Similarly, the second equality holds since from the logic standpoint $\overline{X \cap Y} = \overline{X} \cup \overline{Y}$ is actually a convenient abbreviation for the statement

$$(\forall z \in U)[\neg(P(z) \wedge Q(z)) \iff (\neg P(z) \vee \neg Q(z))].$$

But, as we have proven that $\neg(P \wedge Q)$ is logically equivalent to $(\neg P \vee \neg Q)$ we conclude as above.

### 4.2.3 Arbitrary unions and intersections

For all $i \in I$, where $I$ is called the indexing set, let $X_i$ be a subset of some universal set. We use the notation $\{X_i \mid i \in I\}$ or $(X_i)_{i \in I}$ to denote the collection of such sets. In the previous section we defined the union of two sets. Based on the definition of the union of two sets we can naturally recursively define the union of finitely many sets $X_1, X_2, \ldots, X_n$, for $n \geq 2$, this new set will be denoted by $\bigcup_{k=1}^{n} X_k$, as follows:

$$\bigcup_{k=1}^{2} X_k = X_1 \cup X_2,$$

and for $n \geq 3$

$$\bigcup_{k=1}^{n} X_k = (\bigcup_{k=1}^{n-1} X_k) \cup X_n.$$

Since the operation of taking union is associative these new sets are unambiguously defined. Using a similar approach we can define the intersection of finitely many sets. Unfortunately, we cannot use a recursive definition to define arbitrary infinite unions or intersections (e.g. if the index $I = \mathbb{R}$) and we need to proceed differently and define arbitrary unions as the truth set of a certain predicate.

**Definition 31: Arbitrary unions**

Let $I$ be a set and $(X_i)_{i \in I}$ be a collection of sets. The union of the collection $(X_i)_{i \in I}$, denoted $\bigcup_{i \in I} X_i$ is the set of all elements that belong to at least one set of the collection. Formally,

$$\bigcup_{i \in I} X_i = \{x \mid (\exists i \in I)[x \in X_i]\}.$$

*Remark* 4. We can easily show using the principle of mathematical induction that the set $\bigcup_{k=1}^{n} X_k$ that was recursively defined and the set $\bigcup_{i\in\{1,2,\ldots,n\}} X_i$ where $I=\{1,2,\ldots,n\}$ defined using the truth set coincide and the two definitions are compatible. Since $\bigcup_{k=1}^{n} X_k = \bigcup_{i\in\{1,2,\ldots,n\}} X_i$ we will use both notations interchangeably.

*Remark* 5. If $I=\mathbb{N}$ we write $\bigcup_{n=1}^{\infty} X_n$ for $\bigcup_{n\in\mathbb{N}} X_i$.

**Proposition 12**

Let $(X_i)_{i\in I}$ be a collection of sets. Then, for all $j\in I$ one has $X_j \subseteq \bigcup_{i\in I} X_i$.

*Exercise* 6. Let $X_n=[1,1+\frac{1}{n}]$ for $n\in\mathbb{N}$. Compute $\bigcup_{i=n}^{\infty} X_n$.

*Exercise* 7. Let $X_n=(\frac{3}{n},5n]$ for $n\geq 1$. Compute $\bigcup_{n=1}^{\infty} X_n$.

*Solution:* We will show that $\bigcup_{n=1}^{\infty} X_n=(0,\infty)$.

- First, we show that $\bigcup_{n=1}^{\infty} X_n \subseteq (0,\infty)$.

  Let $x\in\bigcup_{n=1}^{\infty} X_n$, then there exists $k\geq 1$ such that $x\in X_k=(\frac{3}{k},5k]$ and hence $\frac{3}{k}<x\leq 5k$. Since it follows from $k\geq 1$ that $\frac{3}{k}\geq 3>0$ and $5k<\infty$ one has $0<x<\infty$ and thus $x\in(0,\infty)$. Therefore $\bigcup_{n=1}^{\infty} X_n\subseteq(0,\infty)$

- We now show that $(0,\infty)\subseteq\bigcup_{n=1}^{\infty} X_n$.

  Assume now that $x\in(0,\infty)$, then $x>0$ and also $\frac{x}{5}>0$. On the one hand, if follows from the Archimedean principle that there is some $n_1\in\mathbb{N}$ such that $n_1>\frac{x}{5}$, so $5n_1\geq x$. On the other hand, $\frac{3}{x}>0$ and it follows from the Archimedean principle that there exists $n_2\in\mathbb{N}$ such that $\frac{3}{x}<n_2$ and hence $x>\frac{3}{n_2}$. Let $k=\max\{n_1,n_2\}\geq 1$ then $\frac{3}{k}\leq\frac{3}{n_2}<x\leq 5n_1\leq k$ and hence $x\in X_k$. Therefore, $(0,\infty)\subseteq\bigcup_{n=1}^{\infty} X_n$.

By combining the two inclusions we get $\bigcup_{n=1}^{\infty} X_n\subseteq(0,\infty)$. □

Using a similar approach we can define arbitrary intersections.

**Definition 32: Arbitrary intersections**

Let $I$ be a set and $\{X_i \mid i\in I\}$ be a collection of sets. The intersection of the collection, denoted $\bigcap_{i\in I} X_i$ is the set of all elements that belong to all sets of the collection. Formally,

$$\bigcap_{i\in I} X_i=\{x\mid(\forall i\in I)[x\in X_i]\}.$$

*Remark* 6. If $I=\mathbb{N}$ we write $\bigcap_{n=1}^{\infty} X_i$ for $\bigcap_{n\in\mathbb{N}} X_n$.

**Proposition 13**

Let $(X_i)_{i\in I}$ be a collection of sets. Then, for all $j\in I$ one has $\bigcap_{i\in I} X_i\subseteq X_j$.

*Exercise* 8. Let $X_n = [1, 1+\frac{1}{n}]$ for $n \in \mathbb{N}$. Compute $\bigcap_{n=1}^{\infty} X_n$.

*Exercise* 9. Let $X_n = (\frac{3}{n}, 4n]$ for $n \geq 1$. Compute $\bigcap_{n=1}^{\infty} X_n$.

**Theorem 13: DeMorgan's Laws for arbitrary unions and intersections**

Let $(X_i)_{i\in I}$ be a collection of set. Then

1. $\overline{\bigcup_{i\in I} X_i} = \bigcap_{i\in I} \overline{X_i}$,
2. $\overline{\bigcap_{i\in I} X_i} = \bigcup_{i\in I} \overline{X_i}$.

*Proof.*

1. We first prove the inclusion $\overline{\bigcup_{i\in I} X_i} \subseteq \bigcap_{i\in I} \overline{X}_i$.

   If $\overline{\bigcup_{i\in I} X_i} = \emptyset$ then the inclusion holds, otherwise let $z \in \overline{\bigcup_{i\in I} X_i}$. Then $z \notin \bigcup_{i\in I} X_i$ (by definition of the complement), and it follows that $z \notin X_i$ for all $i \in I$ (by definition of the union). Thus, $z \in \overline{X_i}$ for all $i \in I$ (by definition of the complement), which means that $z \in \bigcap_{i\in I} \overline{X}_i$ (by definition of the intersection).

   For the reverse inclusion, if $\bigcap_{i\in I} \overline{X}_i = \emptyset$ then the inclusion holds, otherwise let $z \in \bigcap_{i\in I} \overline{X}_i$. Then $z \in \overline{X_i}$ for all $i \in I$ (by definition of the intersection), and thus $z \notin X_i$ for all $i \in I$ (by definition of the complement). It follows that $z \notin \bigcup_{i\in I} X_i$ (by definition of the union), and hence $z \in \overline{\bigcup_{i\in I} X_i}$ (by definition of the complement).

   Therefore, it follows that $\overline{\bigcup_{i\in I} X_i} = \bigcap_{i\in I} \overline{X}_i$.

2. We first prove the inclusion $\overline{\bigcap_{i\in I} X_i} \subseteq \bigcup_{i\in I} \overline{X}_i$.

   If $\overline{\bigcap_{i\in I} X_i} = \emptyset$ then the inclusion holds, otherwise let $z \in \overline{\bigcap_{i\in I} X_i}$. Then $z \notin \bigcap_{i\in I} X_i$ (by definition of the complement), and it follows that $z \notin X_i$ for some $i \in I$ (by definition of the intersection). Thus, $z \in \overline{X_i}$ for some $i \in I$ (by definition of the complement), which means that $z \in \bigcup_{i\in I} \overline{X}_i$ (by definition of the union).

   For the reverse inclusion, if $\bigcup_{i\in I} \overline{X}_i = \emptyset$ then the inclusion holds, otherwise let $z \in \bigcup_{i\in I} \overline{X}_i$. Then $z \in \overline{X_i}$ for some $i \in I$ (by definition of the union), and thus $z \notin X_i$ for some $i \in I$ (by definition of the complement). It follows that $z \notin \bigcap_{i\in I} X_i$ (by definition of the intersection), and hence $z \in \overline{\bigcap_{i\in I} X_i}$ (by definition of the complement).

   Therefore, it follows that $\overline{\bigcap_{i\in I} X_i} = \bigcup_{i\in I} \overline{X}_i$.

□

### 4.2.4 Power set

We will now consider sets whose elements are sets themselves.

**Definition 33: Power set of a set**

Let $X$ be a set. The power set of $X$, denoted $\mathscr{P}(X)$ or $2^X$, is the set of all subsets of $X$. Formally,

$$\mathscr{P}(X) = \{Y \mid Y \subseteq X\}.$$

**Remark 5**

- Do not forget the empty set and the set itself in the power set! In particular, the power set of a set is *never* empty.
- If follows from the definition that

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

*Example* 49. The power set of $X = \{1,2,3\}$ is

$$\mathscr{P}(X) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}.$$

*Example* 50. The power set of $X = \emptyset$ is

$$\mathscr{P}(\emptyset) = \{\emptyset\},$$

and

$$\mathscr{P}(\mathscr{P}(\emptyset)) = \mathscr{P}(\{\emptyset\}) = \big\{\emptyset, \{\emptyset\}\big\},$$

and

$$\mathscr{P}(\mathscr{P}(\mathscr{P}(\emptyset))) = \mathscr{P}(\mathscr{P}(\{\emptyset\})) = \mathscr{P}(\{\emptyset, \{\emptyset\}\}) = \big\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\big\},$$

etc...

**Theorem 14**

Let $X$ and $Y$ be sets. Then,

$$X \subseteq Y \iff \mathscr{P}(X) \subseteq \mathscr{P}(Y).$$

*Proof.* We will prove the two implications separately.

- $\Longrightarrow$ : Assume that $X \subseteq Y$. Let $A \in \mathscr{P}(X)$ then $A \subseteq X$ and by transitivity of the subset relation since $X \subseteq Y$ one has $A \subseteq Y$. Therefore $A \in \mathscr{P}(Y)$, and $\mathscr{P}(X) \subseteq \mathscr{P}(Y)$.
- $\Longleftarrow$: Assume that $\mathscr{P}(X) \subseteq \mathscr{P}(Y)$. Since $X \subseteq X$ then $X \in \mathscr{P}(X)$ and thus $X \in \mathscr{P}(Y)$ by our assumption. Therefore, $X \subseteq Y$.

□

*Exercise* 10. Show that for all $n \geq 0$, if $X$ is a set with exactly $n$ elements then the number of sets in the power set of $X$ is equal to $2^n$.

*Hint.* You could give a proof using induction. □

### 4.2.5 Cartesian products

To define the concept of Cartesian product we need to understand what is an ordered pair. Consider a set with two elements $\{x,y\}$. This set does not convey a notion of order since $\{x,y\} = \{y,x\}$. If one wants to introduce a notion of order we can formally define the ordered pair $(x,y)$ as the set $\{\{x\},\{x,y\}\}$. With this definition the characteristic property of ordered pairs holds. Indeed,

$$(x_1,y_1) = (x_2,y_2) \iff (x_1 = x_2) \wedge (y_1 = y_2).$$

Also, with this definition it is clear that $(x,y) \neq (y,x)$ since $\{\{x\},\{x,y\}\}$ is obviously not the same set as $\{\{y\},\{y,x\}\} = \{\{y\},\{x,y\}\}$. We will not use the formal definition of an order pair but we will use the concept of ordered pairs as well as the characteristic property.

**Definition 34: Cartesian product of two sets**

Let $X$ and $Y$ be sets. The Cartesian product of $X$ and $Y$, written $X \times Y$, is the set of all *ordered pairs* $(x,y)$ with $x \in X$ and $y \in Y$. Formally,

$$X \times Y = \{(x,y) \mid (x \in X) \wedge (y \in Y)\}.$$

The Cartesian product is named after René Descartes. It is a generalization of the Cartesian coordinate system in the context of arbitrary sets (not just the real numbers).

**Remark 6**

It follows from the definition that

$$w \in X \times Y \iff (\exists x \in X)(\exists y \in Y)[w = (x,y)].$$

*Example* 51. The Cartesian product $\mathbb{R} \times \mathbb{R}$ is nothing else but the 2-dimensional plane usually simply denoted by $\mathbb{R}^2$.

The following property can be easily deduced form logical principles.

**Proposition 14**

Let $X$ be a set. Then $X \times \emptyset = \emptyset$

**Remark 7**

In general, the Cartesian product is not a commutative operation. This is clear by considering the following elementary example. Let $X = \{0,1\}$ and $Y = \{2,3\}$ then

$$X \times Y = \{(0,2),(0,3),(1,2),(1,3)\}$$

but

$$Y \times X = \{(2,0),(2,1),(3,0),(3,1)\},$$

and clearly $X \times Y \neq Y \times X$.
In general, the Cartesian product is also not an associative operation but this is slightly more subtle. Let $X = \{0\}$, $Y = \{1\}$, and $Z = \{2\}$ then

$$(X \times Y) \times Z = \{((0,1),2)\}$$

but

$$X \times (Y \times Z) = \{(0,(1,2))\},$$

and clearly $(X \times Y) \times Z \neq X \times (Y \times Z)$.
However these two sets seem so similar that we want to identify them. This will be done precisely using bijective functions in the next chapter.

# Chapter 5

# Relations

## 5.1 Definitions and basic properties

**Definition 35: Relations**

Let $X$ and $Y$ be sets. A relation $\mathscr{R}$ from $X$ to $Y$ is a subset of $X \times Y$, *i.e.,* $\mathscr{R} \subseteq X \times Y$. If $(x,y) \in \mathscr{R}$ we simply write $x\mathscr{R}y$. We simply say that $\mathscr{R}$ is a relation on $X$ if it is a relation from $X$ to $X$. In other words, a relation $\mathscr{R}$ on a set $X$ is a subset of $X \times X$

*Example* 52. Below are classical examples of relations.

1. The relation of divisibility on $\mathbb{Z}$ defined as

$$\mathscr{R} = \{(x,y) \in \mathbb{Z} \times \mathbb{Z} \mid y = kx, \text{ for some } k \in \mathbb{Z}\}.$$

   This relation is usually denoted by $|$. Then $2 \mid 4$, $19 \mid 0$, but $5 \nmid 2$, etc.

2. The relation of congruence modulo 5 on $\mathbb{Z}$ defined as $x\mathscr{R}y \iff x - y = 5k$ for some $k \in \mathbb{Z}$. This relation is usually denoted by $\equiv$, and $12 \equiv 7 \bmod 5$, but $1 \not\equiv 3 \bmod 5$, etc.

3. The relation of equality at 0 on $F(\mathbb{R})$, the set of functions $f\colon \mathbb{R} \to \mathbb{R}$, defined as $f\mathscr{R}g \iff f(0) = g(0)$.

4. The subset relation on $P(X)$ defined as $A\mathscr{R}B \iff A \subseteq B$.

5. The strict subset relation on $P(X)$ defined as $A\mathscr{R}B \iff A \subset B$.

6. The "less or equal than" relation of $\mathbb{R}$, defined as $x\mathscr{R}y \iff x \leq y$.

7. The "strictly less than" relation of $\mathbb{R}$, defined as $x\mathscr{R}y \iff x < y$.

We now define a couple of important properties that a relation on a set can have.

**Definition 36: Reflexivity**

A relation $\mathscr{R}$ on a set $X$ is *reflexive* if every element in $X$ is in relation with itself. Formally,

$$\mathscr{R} \text{ is a reflexive relation on } X \iff (\forall x \in X)\, x\mathscr{R}x.$$

Divisibility, congruence modulo $k$, subset relation, "less or equal than" relation are all reflexive relations. Strict subset relation and "strictly less than" relation are not reflexive in general.

**Definition 37: Symmetry**

A relation $\mathscr{R}$ on a set $X$ is *symmetric* if for all elements $x, y \in X$ such that $x$ is in relation with $y$ then $y$ is in relation with $x$. Formally,

$$\mathscr{R} \text{ is a symmetric relation on } X \iff (\forall x \in X)(\forall y \in X)[x\mathscr{R}y \implies y\mathscr{R}x].$$

Congruence modulo $k$ is a symmetric relation. The divisibility , subset, strict subset, "less or equal than", and "strictly less than" relations are not symmetric in general.

**Definition 38: Antisymmetry**

A relation $\mathscr{R}$ on a set $X$ is *antisymmetric* if for all elements $x, y \in X$ such that $x$ is in relation with $y$ and $y$ is in relation with $x$ then $x = y$. Formally,

$$\mathscr{R} \text{ is an antisymmetric relation on } X$$
$$\iff$$
$$(\forall x \in X)(\forall y \in X)[((x\mathscr{R}y) \wedge (y\mathscr{R}x)) \implies (x = y)].$$

The divisibility, subset, "less or equal than" relations are symmetric relations. The strict subset relation and "strictly less than" relation are not symmetric in general.

**Definition 39: Transitivity**

A relation $\mathscr{R}$ on a set $X$ is *transitive* if for all elements $x, y, z \in X$, whenever $x$ is in relation with $y$ and $y$ is in relation with $z$, then $x$ is in relation with $z$. Formally,

$$\mathscr{R} \text{ is a transitive relation on } X$$
$$\iff$$
$$(\forall x \in X)(\forall y \in X)(\forall z \in X)[((x\mathscr{R}y) \wedge (y\mathscr{R}z)) \implies (x\mathscr{R}z)].$$

All the relations in Example 63 are transitive.

## 5.2 Equivalence relations and partitions

**Definition 40: Equivalence relation**

A relation $\mathscr{R}$ on a set $X$ is an *equivalence relation* if it is reflexive, symmetric and transitive.

For an equivalent relation $\mathscr{R}$, $x\mathscr{R}y$ is often denoted by $x \sim y$ and reads $x$ is equivalent to $y$.

**Definition 41: Equivalence classes**

If $\sim$ is an equivalence relation on a nonempty set $X$, and $x \in X$, the set $[x] = \{y \in X \mid y \sim x\}$ is called the *equivalence class* of $x$. Elements of the same class are said to be equivalent.

The set of all the equivalence classes is called the quotient set of $X$ and denoted $X_{/\sim}$, *i.e.,* $X_{/\sim} = \{[x] \colon x \in X\}$. Observe that if two elements are equivalent then their equivalence classes coincide.

**Lemma 1**

Let $\sim$ be an equivalence relation on a nonempty set $X$. Let $x, y \in X$, then $x \sim y$ if and only if $[x] = [y]$.

*Proof.* If $x \sim y$ then $[x] \subseteq [y]$, indeed if $z \in [x]$ then $z \sim x$ and by transitivity $z \sim y$, and thus $z \in [y]$. By symmetry it follows that $[x] = [y]$. As for the converse, assume that $[x] = [y]$, then since $x \in [x]$ we have $x \in [y]$ and thus $x \sim y$. □

The main purpose of defining an equivalence relation is to classify elements of a set according to a certain property, and in some sense to generalize the concept of equality. As we will see having an equivalence relation provides a procedure to partition a set. We now introduce the concept of a partition. Let $Y$ be a set and $\mathfrak{P}$ a subset of $\mathscr{P}(Y)$. We use the notation $\bigcup_{A \in \mathfrak{P}} A$ for $\bigcup_{A \in \mathfrak{P}} X_A$ where $X_A = A$. In other words, the set $\bigcup_{A \in \mathfrak{P}} A$ is the set of all elements that belong to at least one set of $\mathfrak{P}$, *i.e.,* $\bigcup_{A \in \mathfrak{P}} A := \{z \colon (\exists A \in \mathfrak{P})\ (z \in A)\}$.

**Definition 42: Partitions**

Let $X$ be a nonempty set. A partition of $X$ is a subset $\mathfrak{P}$ of $\mathscr{P}(X)$ such that

1. For all $C \in \mathfrak{P}$, $C \neq \emptyset$ (non-empty clusters).
2. $\bigcup_{C \in \mathfrak{P}} C = X$ (covering property).
3. For all $A, B \in \mathfrak{P}$, if $A \neq B$, then $A \cap B = \emptyset$ (disjointness property).

**Theorem 15: Canonical partition generated by an equivalence relation**

If $\sim$ is an equivalence relation on a nonempty set $X$, then the collection of equivalence classes of $\sim$, namely $\{[x]\}_{x\in X}$, forms a partition of $X$.

*Proof.* Thanks to reflexivity, every equivalence class $[x]$ is nonempty since $x \in [x]$. Moreover, it also follows from the previous observation that $X \subseteq \bigcup_{x\in X}[x]$ and thus the covering property holds (the other inclusion always holds). It remains to show the disjointness property. Let $[x]$ and $[y]$ be equivalence classes such that $[x] \neq [y]$. Assume for the sake of a contradiction that $[x]\cap[y] \neq \emptyset$. Let $z \in [x]\cap[y]$ then $z \in [x]$ and $z \in [y]$, and hence $z \sim x$ and $z \sim y$. By transitivity $x \sim y$ and it follows from the lemma above that $[x] = [y]$, a contradiction. □

**Theorem 16: Canonical equivalence relation generated by a partition**

Let $\mathfrak{P}$ be a partition of a nonempty set $X$. Define a relation $\sim_{\mathfrak{P}}$ on $X$ by $x \sim_{\mathfrak{P}} y$ if and only if there exists $C \in \mathfrak{P}$ such that $x \in C$ and $y \in C$ (in other words $x$ and $y$ are in the same cluster). Then $\sim_{\mathfrak{P}}$ is an equivalence relation on $X$.

*Proof.* Observe that $x \sim_{\mathfrak{P}} x$ simply means that $x$ is in the same cluster as itself, which is plainly true, and hence $\sim_{\mathfrak{P}}$ is clearly reflexive. Also, $\sim_{\mathfrak{P}}$ is patently symmetric since if $x \sim_{\mathfrak{P}} y$ then $y \sim_{\mathfrak{P}} x$, indeed if $x$ is in the same cluster as $y$ then $y$ is evidently in the same cluster as $x$. It remains to observe that transitivity follows from the disjointness property. Assume that $x, y, z$ are pairwise distinct elements (*i.e.,* $x \neq y$, $y \neq z$, and $x \neq z$) such that $x \sim_{\mathfrak{P}} y$ and $y \sim_{\mathfrak{P}} z$. Then by definition of $\sim_{\mathfrak{P}}$ there exist $C, D \in \mathfrak{P}$ such that $x, y \in C$ and $y, z \in D$. Since $y \in C$ and $y \in D$ we have $y \in C \cap D$. Necessarily $C = D$ otherwise it would contradict the disjointness property, and thus $x \sim_{\mathfrak{P}} z$. □

# Chapter 6

# Functions

## 6.1 Definition and Basic Properties

A function between two sets is a correspondence between elements of these two sets that enjoy some special properties.

**Definition 43: Functions**

Let $X$ and $Y$ be nonempty sets. A *function* $f$ from $X$ to $Y$, and we write $f\colon X \to Y$, is a relation that assigns to *every* element in $X$ *one and only one* element in $Y$. Formally, a function $f$ from $X$ to $Y$ is a subset $f \subseteq X \times Y$ such that

$$[(\forall x \in X)(\exists! y \in Y)\ (x,y) \in f].$$

Let $X$ and $Y$ be nonempty sets. We denote $F(X,Y) = \{f \mid f\colon X \to Y\}$, the set of all functions from $X$ to $Y$. If $X = Y$, we simply write $F(X)$.

**Remark 8: Convention**

In the sequel when we say that $f\colon X \to Y$ is a function we always assume, unless stated otherwise, that $X$ and $Y$ are nonempty.

Note that the logical formula $[(\forall x \in X)(\exists! y \in Y)\ (x,y) \in f]$ is equivalent to the logical formula

$$[(\forall x \in X)(\exists y \in Y)\ (x,y) \in f]$$

$$\wedge$$

$$[(\forall x \in X)[((x,y_1) \in f) \wedge ((x,y_2) \in F)] \implies (y_1 = y_2)]].$$

To verify that a relation $f\colon X \to Y$ is a function it is usually immediate that check that every $x \in X$ has at least an image in $Y$, but it might require some extra effort to show that there is at most one image.

**Remark 9: Functional notation**

Functions play a central role in set theory and in mathematics in general, and we use specific terminology and notation. Since for every $x \in X$ there is a *unique* element $y \in Y$ such that $(x,y) \in f$, we prefer a much more convenient functional notation. Therefore, we will denote by $f(x)$ the unique element that is in relation with $x$. If $f(x) = y$ we say that $y$ is the image of $x$ or that $x$ is the preimage of $y$. We call $X$ the domain of $f$ and $Y$ the codomain. Therefore, a map $f\colon X \to Y$ is a function if and only if

$$[(\forall x \in X)(\exists y \in Y)\ y = f(x)]$$

$$\wedge$$

$$(\forall x_1 \in X)(\forall x_2 \in X)[(x_1 = x_2) \implies (f(x_1) = f(x_2))].$$

*Example* 53. Let $X = \{1,2,3\}$ and $Y = \{5,8,10\}$. The relation $f$ defined by $f(1) = f(2) = 10$, $f(3) = 8$ is a function from $X$ to $Y$.

*Example* 54. The relation $f\colon \mathbb{Z} \to \mathbb{Z}$ that is defined by

$$f(k) = \begin{cases} 0 \text{ if } k \text{ is even,} \\ 1 \text{ if } k \text{ is odd,} \\ 2 \text{ if } k \text{ is a multiple of 4.} \end{cases}$$

is not a function from $\mathbb{Z}$ to $\mathbb{Z}$ since $f(8) = 0$ but also $f(8) = 2$.

*Example* 55. The identity function on $X$ is the function $i_X\colon X \to X$ such that for all $x \in X$, $i_X(x) = x$.

*Example* 56. For all $a,b \in \mathbb{R}$ the functions $f_{a,b}\colon \mathbb{R} \to \mathbb{R}$, defined by $f_{a,b}(x) = ax + b$ are called linear functions.

*Example* 57. Let $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$ and consider the relation $\sim$ on $\mathbb{Z} \times \mathbb{Z}^*$ defined as $(a,b) \sim (c,d)$ iff $ad = bc$. The relation $\sim$ is an equivalence relation on $\mathbb{Z} \times \mathbb{Z}^*$. Define $f\colon (\mathbb{Z} \times \mathbb{Z}^*)_{/\sim} \to (\mathbb{Z} \times \mathbb{Z}^*)_{/\sim}$ as $f([(a,b)]) = [(2a,b)]$. Is $f$ a function?

We now define what it means for two functions to be equal.

**Definition 44: Equality for functions**

Two functions $f_1\colon X_1 \to Y_1$ and $f_2\colon X_2 \to Y_2$ are *equal*, denoted $f_1 = f_2$, if they have the *same domain*, the *same codomain*, and the *same action* on elements in $X$. Formally,

$$f_1 = f_2 \iff (X_1 = X_2) \wedge (Y_1 = Y_2) \wedge ((\forall z \in X_1)[f_1(z) = f_2(z)]).$$

The next definition introduces the concept of image, or range, of a function.

**Definition 45: Image (or range) of a function**

Let $f\colon X \to Y$ be a function. The *image* (or the *range*) of the function $f$ is the set, denoted $\mathrm{Im}(f)$, of all elements in the codomain that are the image of an element in the domain. Formally,

$$\mathrm{Im}(f) = \{y \in Y \mid (\exists x \in X)[y = f(x)]\}.$$

**Remark 10**

The image of a function is a subset of the *codomain* of the function. It follows from the definition that

$$y \in \mathrm{Im}(f) \iff (\exists x \in X)[y = f(x)].$$

Note also that an alternative description of the image of $f$ is

$$\mathrm{Im}(f) = \{f(x) \mid x \in X\}.$$

*Exercise* 11. Let $f\colon \mathbb{Z} \to \mathbb{Z}$ defined by $f(k) = \begin{cases} k-1 \text{ if } k \text{ is even,} \\ k+3 \text{ if } k \text{ is odd.} \end{cases}$

Then $\mathrm{Im}(f) = \mathbb{Z}$.

The next definition introduces the concept of the graph of a function.

**Definition 46: Graph of a function**

Let $f\colon X \to Y$ be a function. The *graph* of the function $f$ is the set, denoted $G_f$, of all ordered pairs $(x,y)$ of elements $x \in X$ and $y \in Y$ that are in relation. Formally,

$$G_f = \{(x,y) \in X \times Y \mid y = f(x)\}.$$

**Remark 11**

The graph of a function is a subset of the Cartesian product of its domain with its codomain. It follows from the definition that

$$z \in G_f \iff (\exists x \in X)[z = (x, f(x))].$$

*Exercise* 12. Let $f(x) = \dfrac{3x+5}{x-2}$. Determine the domain, codomain, and graph of $f$.

## 6.2 Composition of Functions

Assume we are given two functions $f$ and $g$. If the codomain of $f$ coincides with the domain of $g$ then it is make sense to look at what element is obtained if we first apply

$f$ and then $g$ to an element in the domain of $f$. This procedure gives a function from the domain of $f$ in the codomain of $g$.

**Definition 47: Composition of functions**

Let $f\colon X \to Y$, $g\colon Y \to Z$. We define a function $g \circ f\colon X \to Z$, called the composition of $f$ and $g$, by $g \circ f(x) = g(f(x))$, $\forall x \in X$.

Note that for the composition to be defined we just need the image of $f$ to be a subset of the domain of $g$.

**Remark 12**

In general, $g \circ f \neq f \circ g$ and the composition is not a commutative operation! Indeed, consider the function $f\colon \mathbb{R} \to \mathbb{R}$ defined for all $x \in \mathbb{R}$ by $f(x) = 3x$ and the function $g\colon \mathbb{R} \to \mathbb{R}$ defined for all $x \in \mathbb{R}$ by $g(x) = x^2$. It is easy to see that $g \circ f$ and $f \circ g$ have the same domain and codomain, but for instance $g \circ f(1) = 9 \neq 3 = f \circ g(1)$.

**Proposition 15**

Let $f\colon X \to Y$ be a function. Then $f \circ i_X = f$ and $i_Y \circ f = f$.

*Proof.* First we prove that $(f \circ i_X) = f$. Observe that $X$ is the domain of both $(f \circ i_X)$ and $f$, and that $Y$ is the codomain of both $f \circ i_X$ and $f$. It remains to show that for all $x \in X$, $(f \circ i_X)(x) = f(x)$. By definition of the composition operation and of $i_X$, it follows that if $x \in X$ then $(f \circ i_X)(x) = f(i_X(x)) = f(x)$.

The proof is similar for the second statement. Observe that $X$ is the domain of both $i_Y \circ f$ and $f$, and that $X$ is the codomain of both $i_Y \circ f$ and $f$. It remains to show that for all $x \in X$, $(i_Y \circ f)(x) = f(x)$. By definition of the composition operation and of $i_Y$, it follows that if $x \in X$ then $(i_Y \circ f)(x) = i_Y(f(x)) = f(x)$, since $f(x) \in Y$.

□

The composition operation is associative.

**Proposition 16: Associativity of the composition**

Let $f\colon W \to X$, $g\colon X \to Y$, and $h\colon Y \to Z$ be functions. Then,

$$(h \circ g) \circ f = h \circ (g \circ f).$$

*Proof.* Observe that $W$ is the domain of both $(h \circ g) \circ f$ and $h \circ (g \circ f)$, and that $Z$ is the codomain of both $(h \circ g) \circ f$ and $h \circ (g \circ f)$. It remains to show that for all $w \in W$, $((h \circ g) \circ f)(w) = (h \circ (g \circ f))(w)$. By definition of the composition operation it follows that if $x \in X$ then

$$((h \circ g) \circ f)(w) = (h \circ g)(f(w)) = h(g(f(w)))$$

and

$$(h\circ(g\circ f))(w) = h((g\circ f)(w)) = h(g(f(w))).$$

Therefore, $((h\circ g)\circ f)(w) = h(g(f(w))) = (h\circ(g\circ f))(w)$ and the two functions are equal. □

## 6.3 Surjectivity, injectivity, and bijectivity of functions

### 6.3.1 Definitions and examples

A surjective function (or onto function) is a function whose image fills in completely the codomain.

**Definition 48: Surjective function**

A function $f\colon X\to Y$ is *surjective* (or onto, or a surjection) if every element in the codomain of $f$ admits a preimage in the domain of $f$. Formally,

$$f\colon X\to Y \text{ is surjective} \iff (\forall y\in Y)(\exists x\in X)[y=f(x)].$$

The following proposition is a characterization of surjectivity in terms of the image of the function.

**Proposition 17: Characterization of surjectivity in terms of the image**

Let $f\colon X\to Y$ be a function. Then, $f$ is surjective if and only if $\mathrm{Im}(f)=Y$.

*Proof.* We know that $\mathrm{Im}(f)\subseteq Y$ always holds, but the definition of surjectivity says that $Y\subseteq \mathrm{Im}(f)$. Therefore $Y=\mathrm{Im}(f)$. □

*Example* 58. The identity function on $X$ is surjective.

*Example* 59. Let $f\colon (-\infty,2)\cup(2,\infty)\to\mathbb{R}$, defined by $f(x)=\dfrac{3x+5}{x-2}$. The function $f$ is not surjective since $\mathrm{Im}(f)=(-\infty,3)\cup(3,\infty)$.

However, the function $g\colon (-\infty,2)\cup(2,\infty)\to(-\infty,3)\cup(3,\infty)$, defined by $g(x)=\dfrac{3x+5}{x-2}$ is surjective.

*Exercise* 13. Let $f\colon\mathbb{Z}\to\mathbb{Z}$ defined by $f(k)=\begin{cases}k-1 \text{ if } k \text{ is even},\\ k+3 \text{ if } k \text{ is odd}.\end{cases}$

Show that the function $f$ is surjective.

*Exercise* 14. Let $f\colon\mathbb{R}\to\mathbb{R}$, defined by $f(x)=x+2|x|$. Is $f$ surjective?

A function is injective (or one-to-one often abbreviated as $1-1$) if no two distinct elements in the domain are assigned the same element in the codomain.

**Definition 49: Injective function**

A function $f: X \to Y$ is *injective* (or one-to-one, or an injection) if every two distinct elements in the domain have distinct images in the codomain. Formally,

$$f: X \to Y \text{ is injective}$$
$$\iff$$
$$(\forall x_1 \in X)(\forall x_2 \in X)[(x_1 \neq x_2) \implies (f(x_1) \neq f(x_2))].$$

**Remark 13**

In practice, to show that a function is injective we need to prove *either* one of the following two logically equivalent statements (the second statement is the contrapositive of the first statement.):

- for all $x_1, x_2 \in X$ if $x_1 \neq x_2$ then $f(x_1) \neq f(x_2)$.
- for all $x_1, x_2 \in X$ if $f(x_1) = f(x_2)$ then $x_1 = x_2$.

*Example* 60. The identity function on $X$ is injective.

*Example* 61. The projections $\pi_X : X \times Y \to X, (x,y) \mapsto x$ and $\pi_Y : X \times Y \to Y, (x,y) \mapsto y$ are surjective.

*Example* 62. Let $f: \mathbb{Z} \to \mathbb{Z}$ defined by $f(k) = \begin{cases} k-1 \text{ if } k \text{ is even,} \\ k+3 \text{ if } k \text{ is odd.} \end{cases}$

The function $f$ is injective.

*Example* 63. Let $f: (-\infty, 2) \cup (2, \infty) \to \mathbb{R}$, defined by $f(x) = \dfrac{3x+5}{x-2}$. Is the function $f$ injective?

*Example* 64. Let $f: \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x + 2|x|$. Is $f$ injective?

**Definition 50: Bijective function**

Let $f: X \to Y$ be a function. Then $f$ is *bijective* (or a bijection) if $f$ is simultaneously injective and surjective. In the case where $X = Y$ a bijection is simply called a permutation.

*Example* 65. The identity function $i_X : X \to X$ is a permutation.

*Example* 66. Let $f: (-\infty, 2) \cup (2, \infty) \to \mathbb{R}$, defined by $f(x) = \dfrac{3x+5}{x-2}$. Is $f$ bijective?

### 6.3.2 Injectivity, surjectivity and composition

In this section we show that injectivity, surjectivity, and bijectivity are stable under composition.

**Proposition 18: Stability of injectivity under composition**

Let $f: W \to X$ and $g: X \to Y$ be functions. If $f$ and $g$ are injective, then $g \circ f$ is also injective.

*Proof.* Assume that $f$ and $g$ are injective. Let $w_1, w_2 \in W$ such that $g \circ f(w_1) = g \circ f(w_2)$, then $g(f(w_1)) = g(f(w_2))$ (by definition of the composition) and $f(w_1) = f(w_2)$ (by injectivity of $g$). Now it follows from the injectivity of $f$ that $w_1 = w_2$, and $g \circ f$ is injective.

□

**Proposition 19: Stability of surjectivity under composition**

Let $f: W \to X$ and $g: X \to Y$ be functions. If $f$ and $g$ are surjective, then $g \circ f$ is also surjective.

*Proof.* Assume that $f$ and $g$ are surjective. Let $y \in Y$, then there exists $x \in X$ such that $g(x) = y$ (by surjectivity of $g$). Since $x \in X$, there exists $w \in W$ such that $x = f(w)$ (by surjectivity of $f$). And hence, $y = g(x) = g(f(w)) = g \circ f(w)$ (by definition of the composition). We have just shown that for every $y \in Y$ there exists $w \in W$ such that $y = g \circ f(w)$, which means that $g \circ f$ is surjective. □

**Proposition 20: Stability of bijectivity under composition**

Let $f: W \to X$ and $g: X \to Y$ be functions. If $f$ and $g$ are bijective, then $g \circ f$ is also bijective.

*Proof.* Assume that $f$ and $g$ are bijective, then in particular they are both injective . By Theorem 15, $g \circ f$ is then injective. A similar reasoning using Theorem 16 will show that $g \circ f$ is surjective, and hence $g \circ f$ is bijective. □

## 6.4 Invertible functions

In this section we take a look at those functions whose actions can be "undone".

**Definition 51: Invertibility**

Let $f: X \to Y$ be a function. We say that $f$ is *invertible* (or admits an inverse) if there exists a function $g: Y \to X$ such that $f \circ g = i_Y$ and $g \circ f = i_X$.

Invertibility and bijectivity are intimely connected. Indeed, as we will see shortly being invertible and being bijective are actually equivalent notions for functions! Invertibility and bijectivity is thus the same concept but in two different disguises. Bijectivity is more intrinsic and analytic in the sense that it can be checked directly on

the function, whereas invertibility has a more extrinsic and algebraic flavor since it involves another function and the composition operation. The goal of this section is to prove this equivalence. We first show that invertibility implies injectivity.

**Theorem 17**

Let $f\colon X \to Y$ be a function. If $f$ is invertible then $f$ is injective.

*Proof.* Assume that $f$ is invertible. Then there is a function $g : Y \to X$ such that $g \circ f = i_X$ and $f \circ g = i_Y$. If $x_1, x_2 \in X$ and $f(x_1) = f(x_2)$, then $x_1 = g(f(x_1)) = g(f(x_2)) = x_2$. Thus $f$ is injective. □

A nice consequence of the injectivity of invertible functions is that the inverse of an invertible function is uniquely determined.

**Proposition 21: Uniqueness of the inverse**

Let $f\colon X \to Y$ be a function. If $f$ is invertible then $f$ has a unique inverse.

*Proof.* Let $f\colon X \to Y$ be a function. Our goal is to show that if there are two functions $g_1, g_2\colon Y \to X$ such that $f \circ g_1 = i_Y$, $g_1 \circ f = i_X$, $f \circ g_2 = i_Y$, and $g_2 \circ f = i_X$, then $g_1 = g_2$. Let $y \in Y$ then $(f \circ g_1)(y) = i_Y(y) = y$ and $(f \circ g_2)(y) = i_Y(y) = y$, thus $(f \circ g_1)(y) = (f \circ g_2)(y)$. It follows from the definition of the composition that $f(g_1(y)) = f(g_2(y))$ and since $f$ is invertible, $f$ is injective and hence $g_1(y) = g_2(y)$. Therefore, $g_1 = g_2$. □

**Remark 14**

If $f$ is invertible, by Proposition 21 the unique function satisfying the conditions of the definition is called *the* inverse of $f$ and is denoted $f^{-1}$.

We now show that invertibility implies surjectivity.

**Theorem 18**

Let $f\colon X \to Y$ be a function. If $f$ is invertible then $f$ is surjective.

*Proof.* Assume that $f$ is invertible. Then there is a function $g : Y \to X$ such that $g \circ f = i_X$ and $f \circ g = i_Y$. Let $y \in Y$, and put $x = g(y)$. Then by definition of $g$, one has $x \in X$, and thus

$$\begin{aligned} f(x) &= f(g(y)) \text{ (because } f \text{ is a function)} \\ &= (f \circ g)(y) \text{ (by definition of the composition)} \\ &= i_Y(y) \text{ (since } f \circ g(y) = i_Y(y) \text{ by our assumption)} \\ &= y \text{ (by definition of the identity function on } Y.) \end{aligned}$$

Therefore $f$ is surjective. □

Finally, we prove that bijectivity implies invertibility. This result is slightly more difficult to obtain since it involves constructing an inverse for the function.

**Theorem 19**

Let $f\colon X \to Y$ be a function. If $f$ is bijective then $f$ is invertible.

*Proof.* Assume $f$ is bijective. Given $y \in Y$, since $f$ is surjective there is some $x \in X$ such that $y = f(x)$, and since $f$ is injective this $x$ is unique. Indeed if there are $x_1, x_2 \in X$ such that $f(x_1) = y = f(x_2)$, then $x_1 = x_2$ by injectivity of $f$. So for every $y \in Y$ there is a unique $x_y \in X$ such that $y = f(x_y)$. We will define a function $g : Y \to X$ by assigning to every element $y \in Y$ to the unique element $x_y \in X$ such that $f(x_y) = y$, i.e. $g(y) = x_y$. By uniqueness of $x_y$, $g$ is a function.

Given $y \in Y$, then $g(y) = x_y$ where $f(x_y) = y$, and thus $f(g(y)) = f(x_y) = y$ (since $f$ is a function). It follows from the definition of the composition that $(f \circ g)(y) = y$, and by definition of the identity function that $(f \circ g)(y) = i_Y(y)$. Since $y \in Y$ was arbitrary, one has $f \circ g = i_Y$.

It remains to show that $(g \circ f) = i_X$. Now given $x \in X$, $g(f(x))$ is the element $x_0 \in X$ such that $f(x_0) = f(x)$. That is, $g(f(x)) = x_0 = x$, since $f$ is injective. Thus $g \circ f = i_X$, and therefore $f$ is invertible. □

Combining the last three theorems we obtain the following corollary which was the main goal of this section.

**Corollary 1**

Let $f\colon X \to Y$ be a function. Then,

$f$ is invertible if and only if $f$ is bijective.

*Proof.* Assume that $f$ is invertible, then it follows by Theorem 17 that $f$ is injective and by Theorem 18 that $f$ is surjective. Therefore, $f$ is bijective. The converse is Theorem 19. □

Since bijectivity is stable under composition it follows from Corollary 1 that invertibility is also stable under composition. We give a direct and elementary proof of this statement without invoking Corollary 1.

**Proposition 22: Stability of invertibility under composition**

Let $f\colon X \to Y$ and $g\colon Y \to Z$ be functions. If $f$ and $g$ are invertible, then the function $g \circ f\colon X \to Z$ is invertible and its inverse is given is $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

*Proof.* Let $g^{-1}$ and $f^{-1}$ be the inverses of $g$ and $f$ respectively. It follows from the associativity of the composition operation that,

$$\begin{aligned}(g\circ f)\circ(f^{-1}\circ g^{-1}) &= g\circ(f\circ f^{-1})\circ g^{-1}\\ &= g\circ i_Y\circ g^{-1}\\ &= g\circ g^{-1}\\ &= i_Z\end{aligned}$$

and similarly,

$$\begin{aligned}(f^{-1}\circ g^{-1})\circ(g\circ f) &= f^{-1}\circ(g^{-1}\circ g)\circ f)\\ &= f^{-1}\circ i_Y\circ f)\\ &= f^{-1}\circ f\\ &= i_X.\end{aligned}$$

Therefore, $g\circ f$ is invertible and $(g\circ f)^{-1} = f^{-1}\circ g^{-1}$.

□

## 6.5 Functions and sets

Recall that the image of a function $f\colon X\to Y$ is the set $\mathrm{Im}(f) = \{y\in Y \mid (\exists x\in X)[y = f(x)]\}$. The image of a function is a particular case of the direct image of a subset under the function.

### 6.5.1 Direct image

**Definition 52: Direct image of a set**

Let $f\colon X\to Y$ be a function. If $Z\subseteq X$, the *image of $Z$ under $f$* is the set, denoted $f(Z)$, of all elements in the codomain that are the image of at least one element in $Z$. Formally,

$$f(Z) = \{y\in Y \mid (\exists z\in Z)[y = f(z)]\}.$$

**Remark 15**

- Note that $f(X)$ is simply the image of $f$, *i.e.*, $\mathrm{Im}(f) = f(X)$.
- It follows from the definition that

$$v\in f(Z) \iff (\exists z\in Z)\ [v = f(z)].$$

The following proposition states that inclusion is preserved under taking direct images.

**Proposition 23: Direct image and inclusion**

Let $f\colon X \to Y$ be a function. Let $W$ and $Z$ be subsets of $X$. If $W \subseteq Z$, then $f(W) \subseteq f(Z)$.

*Proof.* If $f(W)$ is empty then the conclusion holds. Otherwise, let $v \in f(W)$ then there exists $w \in W$ such that $v = f(w)$ (by definition of the direct image). But since $W \subseteq Z$ it follows that $w \in Z$ and thus $v \in f(Z)$ (by definition of the direct image). Therefore, $f(W) \subseteq f(Z)$. □

The following proposition states that the direct image of an union is the union of the direct images.

**Proposition 24: Direct image and union**

Let $f\colon X \to Y$ be a function and $W$ and $Z$ be subsets of $X$. Then,

$$f(W \cup Z) = f(W) \cup f(Z).$$

*Proof.* The proof is a classical double-inclusion argument.

- We first show that $f(W \cup Z) \subseteq f(W) \cup f^{-1}(Z)$. If $f(W \cup Z)$ is empty then the conclusion holds. Otherwise, let $y \in f(W \cup Z)$, then there exists $x \in W \cup Z$ such that $y = f(x)$ (by definition of the direct image) thus $y = f(x)$ for some $x \in W$ or $y = f(x)$ for some $x \in Z$ (by definition of the union) and hence $y \in f(W)$ or $y \in f(Z)$ (by definition of the image) and $y \in f(W) \cup f(Z)$ (by definition of the union). Therefore $f(W \cup Z) \subseteq f(W) \cup f(Z)$.
- Now we show that $f(W) \cup f(Z) \subseteq f(W \cup Z)$. If $f(W) \cup f(Z) = \emptyset$ the inclusion holds. Otherwise, let $y \in f(W) \cup f(Z)$, then $y \in f(W)$ or $y \in f(Z)$ (by definition of the union) thus $y = f(x)$ for some $x \in W$ or $y = f(x)$ for some $x \in Z$ (by definition of the image) and $y = f(x)$ for some $x \in W \cup Z$ (by definition of the union) thus $y \in f(W \cup Z)$ (by definition of the direct image). Therefore $f(W) \cup f(Z) \subseteq f(W \cup Z)$.

□

The situation is slightly different as far as intersection is concerned.

**Proposition 25: Direct image and intersection**

Let $f\colon X \to Y$ be a function and $W$ and $Z$ be subsets of $X$. Then,

$$f(W \cap Z) \subseteq f(W) \cap f(Z).$$

*Proof.* If $f(W \cap Z) = \emptyset$ then the inclusion holds. Otherwise, let $y \in f(W \cap Z)$, then there exists $x \in W \cap Z$ such that $y = f(x)$ (by definition of the direct image), thus $y = f(x)$ for some $x \in W$ and $y = f(x)$ for some $x \in Z$ (by definition of the intersection), and

hence $y \in f(W)$ and $y \in f(Z)$ (by definition of the direct image), and $y \in f(W) \cap f(Z)$ (by definition of the intersection). Therefore $f(W \cap Z) \subseteq f(W) \cap f(Z)$. □

### 6.5.2 Inverse image

**Definition 53: Inverse image of a set**

Let $f\colon X \to Y$ be a function. Let $Z$ be a subset of $Y$. Then the *inverse image of $Z$ with respect to $f$*, denoted $f^{-1}(Z)$, is the set of all elements in $X$ that have their image in $Z$. Formally,

$$f^{-1}(Z) := \{x \in X \mid f(x) \in Z\}.$$

**Remark 16**

- In this context the symbol $f^{-1}$ does not refer to the inverse of the function $f$ (which might not exist in the first place).
- If follows from the definition that $v \in f^{-1}(Z) \iff f(v) \in Z$.

The following proposition states that inclusion is preserved under taking inverse images.

**Proposition 26: Inverse image and inclusion**

Let $f\colon X \to Y$ be a function. Let $W$ and $Z$ be subsets of $Y$. If $W \subseteq Z$, then $f^{-1}(W) \subseteq f^{-1}(Z)$

*Proof.* If $f^{-1}(W)$ is empty then the conclusion holds. Otherwise, let $v \in f^{-1}(W)$ then $f(v) \in W$ (by definition of the inverse image) and it follows from $W \subseteq Z$ that $f(v) \in Z$, and hence $v \in f^{-1}(Z)$ (by definition of the inverse image). Therefore, $f^{-1}(W) \subseteq f^{-1}(Z)$. □

The following proposition states that the inverse image of a union is the union of the inverse images.

**Proposition 27: Inverse image and union**

Let $f\colon X \to Y$ be a function. Let $W$ and $Z$ be subsets of $Y$. Then,

$$f^{-1}(W \cup Z) = f^{-1}(W) \cup f^{-1}(Z).$$

*Proof.* The proof is another classical double-inclusion argument.

- We first show the inclusion $f^{-1}(W \cup Z) \subseteq f^{-1}(W) \cup f^{-1}(Z)$. If $f^{-1}(W \cup Z) = \emptyset$ the inclusion holds. Otherwise, let $x \in f^{-1}(W \cup Z)$, then $f(x) \in W \cup Z$ (by definition of the inverse image) thus $f(x) \in W$ or $f(x) \in Z$ (by definition of the union) and hence $x \in f^{-1}(W)$ or $x \in f^{-1}(Z)$ (by definition of the inverse image) and $x \in f^{-1}(W) \cup f^{-1}(Z)$ (by definition of the union). Therefore $f^{-1}(W \cup Z) \subseteq f^{-1}(W) \cup f^{-1}(Z)$.

- Then we show that $f^{-1}(W) \cup f^{-1}(Z) \subseteq f^{-1}(W \cup Z)$. If $f^{-1}(W) = \emptyset$ the inclusion hols. Otherwise, let $x \in f^{-1}(W) \cup f^{-1}(Z)$, then $x \in f^{-1}(W)$ or $x \in f^{-1}(Z)$ (by definition of the union) and $f(x) \in W$ or $f(x) \in Z$ (by definition of the inverse image) and hence $f(x) \in W \cup Z$ (by definition of the union) thus $x \in f^{-1}(W \cup Z)$ (by definition of the inverse image). Therefore $f^{-1}(W) \cup f^{-1}(Z) \subseteq f^{-1}(W \cup Z)$.

□

The following proposition states that the inverse image of an intersection is the intersection of the inverses images.

**Proposition 28: Inverse image and intersection**

Let $f : X \to Y$ be a function. Let $W$ and $Z$ be subsets of $Y$. Then,

$$f^{-1}(W \cap Z) = f^{-1}(W) \cap f^{-1}(Z).$$

*Proof.* As you would expect the proof goes through a double inclusion argument.

- First of all the inclusion $f^{-1}(W \cap Z) \subseteq f^{-1}(W) \cap f^{-1}(Z)$. If $f^{-1}(W \cap Z) = \emptyset$ the inclusion holds. Otherwise, let $x \in f^{-1}(W \cap Z)$, then $f(x) \in W \cap Z$ (by definition of the inverse image) thus $f(x) \in W$ and $f(x) \in Z$ (by definition of the intersection) and hence $x \in f^{-1}(W)$ and $x \in f^{-1}(Z)$ (by definition of the inverse image) and $x \in f^{-1}(W) \cap f^{-1}(Z)$ (by definition of the intersection). Therefore $f^{-1}(W \cap Z) \subseteq f^{-1}(W) \cap f^{-1}(Z)$.

- Second of all, the inclusion $f^{-1}(W) \cap f^{-1}(Z) \subseteq f^{-1}(W \cap Z)]$. If $f^{-1}(W) \cap f^{-1}(Z) = \emptyset$ the inclusion holds. Otherwise, let $x \in f^{-1}(W) \cap f^{-1}(Z)$, then $x \in f^{-1}(W)$ and $x \in f^{-1}(Z)$ (by definition of the intersection) and $f(x) \in W$ and $f(x) \in Z$ (by definition of the inverse image) and hence $f(x) \in W \cap Z$ (by definition of the intersection) thus $x \in f^{-1}(W \cap Z)$ (by definition of the inverse image). Therefore $f^{-1}(W) \cap f^{-1}(Z) \subseteq f^{-1}(W \cap Z)$.

□

### 6.5.3 Remarks about the notation

We use the notation $f(A)$ for the direct image of a subset of the domain, or $f^{-1}(B)$ for the inverse image of a subset of the codomain. These are very convenient, intuitive, and classical notations, however you should never forget what is their exact meaning. It would actually be more appropriate and rigorous to use a different notation that would make the distinction between the function $f\colon X \to Y$ and its direct image ,or inverse image, that are functions defined on sets of sets. For instance we could use the following notation :

- direct image:

$$\begin{array}{rccl} f_{direct}\colon & \mathscr{P}(X) & \to & \mathscr{P}(Y) \\ & A & \mapsto & f_{direct}(A) := \{y \in Y \mid (\exists z \in A)[y = f(z)]\} \end{array}$$

- inverse image:

$$\begin{array}{rccl} f_{inverse}\colon & \mathscr{P}(Y) & \to & \mathscr{P}(X) \\ & B & \mapsto & f_{inverse}(B) := \{x \in X \mid f(x) \in B]\} \end{array}$$

  In the previous section we settled on the simpler notation $f(A)$ instead of $f_{direct}(A)$ and $f^{-1}(B)$ in lieu of $f_{inverse}(B)$. It is crucial that you understand that talking about $f^{-1}(B)$ does *not* imply that $f$ is invertible, and $f^{-1}(B)$ does not refer to the inverse of $f$ as a function but is only meant to refer to the inverse image, which according to the definition makes sense for *any function regardless of its invertibility*. In the case where $f$ is actually invertible then the inverse image of a subset becomes $f^{-1}(B) = \{f^{-1}(b) \mid b \in B]\}$ where in the expression $f^{-1}(b)$, $f^{-1}$ *is* the inverse of the function $f$ which goes from $X$ onto $Y$.

# Chapter 7

# Introduction to the Cardinality of Sets

## 7.1 Finite and Infinite sets

It is usual to denote by $|X|$ the cardinality of $X$, i.e., the number of elements of $X$. If we are dealing with finite sets we intuitively understand what $|X| = |Y|$ means, but what if the sets are infinite. Understanding a formal definition of "cardinality" and the concept of "infinity" is the goal of this chapter.

**Definition 54**

A set $X$ is said to be *finite* if there exist a natural number $n \geq 1$ and a bijection between $X$ and $\{1, 2, \ldots, n\}$. The number $n$ is called the cardinality of $X$, and is denoted by $|X|$.

**Definition 55**

A set $X$ is said to be infinite if it is not finite, and we use the notation $|X| = \infty$ to express that $X$ is infinite.

**Proposition 29**

If $X$ and $Y$ are finite then $|X \times Y| = |X||Y|$.

**Proposition 30**

Let $X$ and $Y$ be finite *disjoint* sets. Then $|X \cup Y| = |X| + |Y|$.

Using the Principle of Mathematical Induction one can prove the following corol-

lary.

**Corollary 1.** *Let $X_1, X_2, \ldots, X_n$ be a collection of finite mutually disjoint sets, i.e. $X_i \cap X_j = \emptyset$ if $i \neq j$. Then*

$$\left|\bigcup_{i=1}^{n} X_i\right| = \sum_{i=1}^{n} |X_i|.$$

**Corollary 2.** *Let $X$ and $Y$ be finite sets. Then $|X \cup Y| = |X| + |Y| - |X \cap Y|$.*

### 7.1.1 The Pigeonhole Principle

The pigeonhole principle, in its simplest form, says that if $k$ objects are places in $n$ containers and $k > n$, then at least one container will have more than one object in it. The mathematical formulation is as follows.

**Theorem 20: Pigeonhole Principle**

Let $X_1, X_2, \ldots, X_n$ be a collection of finite mutually disjoint sets. Let $X = \bigcup_{i=1}^{n} X_i$. If $|X| = k$ and $k > n$, then, for some $i$, $|X_i| \geq 2$.

**Theorem 21: Generalized Pigeonhole principle**

Let $X_1, X_2, \ldots, X_n$ be a collection of finite mutually disjoint sets. Let $X = \bigcup_{i=1}^{n} X_i$. If $|X| > nr$ for some positive integer $r$. Then, for some $i$, $|X_i| \geq r+1$.

## 7.2 Comparing the size of sets

**Definition 56: Equinumerability**

Let $X$ and $Y$ be sets. We say that $X$ and $Y$ are *equinumerous* (or equipotent) if there exists a bijection from $X$ onto $Y$. If $X$ is equinumerous to $Y$ we write $X \approx Y$.

**Proposition 31**

Let $X, Y, Z$ be sets. Then,

1. $X \approx X$.
2. If $X \approx Y$ then $Y \approx X$.
3. If $X \approx Y$ and $Y \approx Z$ then $X \approx Z$.

*Proof.* The three assertions are merely a reformulation, using the language of equinumerous relation, of results about bijective maps that have already been proven.

1. The identity function $i_X$ on any set $X$ is a bijection.
2. If there is a bijection $f$ from $X$ to $Y$ then the inverse of $f$, namely $f^{-1}$, is a bijection from $Y$ to $X$.
3. The composition of two bijections is a bijection.

□

The proposition above suggests that $\approx$ seems to have all the attributes of an equivalence relation that would be defined on the collection of all sets ... but there is an issue with this last statement since the collection of all sets CANNOT be a set (this is Russell's Paradox), and we only introduced relations on sets! This is a delicate point that can only be overcome using a rigorous axiomatic approach for set theory instead of the elementary naive approach that we have undertaken.

In the next proposition we compare the size of a set to the size of its power set.

**Proposition 32**

Let $X$ be a set.

1. There exists an injection from $X$ into $\mathscr{P}(X)$.
2. There does not exist a surjection from $X$ onto $\mathscr{P}(X)$.

*Sketch of proof.*

1. Consider the function $f\colon X \to \mathscr{P}(X)$ defined by $f(x) = \{x\}$ for all $x \in X$. This function which assigns to an *element* $x$ in $X$, the *subset* of $X$ that consists of the set $\{x\}$ (the set with exactly one element that is equal to $x$) is easily seen to be injective.
2. Let $f$ be a function from $X$ to $\mathscr{P}(X)$. Consider the set

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

By definition $Z$ is a subset of $X$ and hence $Z \in \mathscr{P}(X)$. Assume that $Z$ has a preimage, *i.e.,* $Z = f(a)$ for some $a \in X$. By examining the two cases, $a \in Z$ or $a \notin Z$, you will realize that it leads to a contradiction in both cases and thus $Z$ cannot have a preimage. Therefore, no function $f$ from $X$ to $\mathscr{P}(X)$ can ever be surjective since the set $Z$ will never have a preimage.

□

We write $X \preceq Y$ if there is an injection from $X$ into $Y$, and $X \prec Y$ if $X \preceq Y$ and $X \not\approx Y$. Considering $X = \mathbb{N}$, Proposition 32 tells you that $\mathbb{N} \prec \mathscr{P}(\mathbb{N})$. It can be shown that $\mathscr{P}(\mathbb{N}) \approx \mathbb{R}$ and hence $\mathbb{N} \prec \mathbb{R}$. The Continuum Hypothesis states that there does not exist a set $X$ such that $\mathbb{Q} \prec X \prec \mathbb{R}$. More informally, there is no set whose cardinality lies strictly between the cardinality of the "discrete" set $\mathbb{N}$ and the "continuous" set $\mathbb{R}$.