Chapter 2
The Art of Proof

In Example 1.2 we gave a proof of $1+1=2$ in 17 (!) proof steps. At that point you may have probably asked yourself that if it takes that much effort to prove such a simple statement, how can one ever prove any non-trivial mathematical result using formal proofs. This objection is of course justified; however we will show in this chapter how one can simplify formal proofs using some methods of proof such as proofs by cases or by contradiction. It is crucial to note that the next results are not theorems of a formal theory, but theorems about formal proofs. In particular, they show how — under certain conditions — a formal proof can be transformed into another.

The Deduction Theorem

In common mathematics, one usually proves implications of the form

$$\text{IF } \Phi \text{ THEN } \Psi$$

by simply assuming the truth of Φ and deriving from this the truth of Ψ. When writing formal proofs, the so-called DEDUCTION THEOREM enables us to use a similar trick: Rather than proving $\mathbf{\Phi} \vdash \varphi \to \psi$ we simply add φ to our set of formulae $\mathbf{\Phi}$ and prove $\mathbf{\Phi} \cup \{\varphi\} \vdash \psi$.

If $\mathbf{\Phi}$ is a set of formulae and $\mathbf{\Phi}'$ is another set of formulae in the same language as $\mathbf{\Phi}$, then we write $\mathbf{\Phi} + \mathbf{\Phi}'$ for $\mathbf{\Phi} \cup \mathbf{\Phi}'$. In the case when $\mathbf{\Phi}' = \{\varphi\}$ consists of a single formula, we write $\mathbf{\Phi} + \varphi$ for $\mathbf{\Phi} \cup \mathbf{\Phi}'$.

THEOREM 2.1 (DEDUCTION THEOREM). *If $\mathbf{\Phi}$ is a set of formulae and $\mathbf{\Phi}+\psi \vdash \varphi$, then $\mathbf{\Phi} \vdash \psi \to \varphi$; and vice versa, if $\mathbf{\Phi} \vdash \psi \to \varphi$, then $\mathbf{\Phi} + \psi \vdash \varphi$:*

$$\mathbf{\Phi} + \psi \vdash \varphi \quad \Longleftrightarrow \quad \mathbf{\Phi} \vdash \psi \to \varphi \tag{DT}$$

Proof. It is clear that $\mathbf{\Phi} \vdash \psi \to \varphi$ implies $\mathbf{\Phi} + \psi \vdash \varphi$. Conversely, suppose that $\mathbf{\Phi} + \psi \vdash \varphi$ holds and let the sequence $\varphi_0, \ldots, \varphi_n$ with $\varphi_n \equiv \varphi$ be a formal proof for φ from $\mathbf{\Phi} + \psi$. For each $i \leq n$ we will replace the formula φ_i by a sequence of formulae which ends with $\psi \to \varphi_i$. Let $i \leq n$ and assume $\mathbf{\Phi} \vdash \psi \to \varphi_j$ for every $j < i$.

- If φ_i is a logical axiom or $\varphi_i \in \mathbf{\Phi}$, we have

$\varphi_{i,0}$:	φ_i	$\varphi_i \in \mathbf{\Phi}$ or φ_i is a logical axiom
$\varphi_{i,1}$:	$\varphi_i \to (\psi \to \varphi_i)$	instance of L$_1$
$\varphi_{i,2}$:	$\psi \to \varphi_i$	from $\varphi_{i,1}$ and $\varphi_{i,0}$ by (MP)

- The case $\varphi_i \equiv \psi$ follows directly from Example 1.1.
- If φ_i is obtained by Modus Ponens from φ_j and $\varphi_k \equiv (\varphi_j \to \varphi_i)$ for some $j < k < i$, we have

$\varphi_{i,0}$:	$\psi \to \varphi_j$	since $j < i$
$\varphi_{i,1}$:	$\psi \to (\varphi_j \to \varphi_i)$	since $k < i$
$\varphi_{i,2}$:	$\varphi_{i,1} \to ((\psi \to \varphi_j) \to (\psi \to \varphi_i))$	instance of L$_2$
$\varphi_{i,3}$:	$(\psi \to \varphi_j) \to (\psi \to \varphi_i)$	from $\varphi_{i,2}$ and $\varphi_{i,1}$ by (MP)
$\varphi_{i,4}$:	$\psi \to \varphi_i$	from $\varphi_{i,3}$ and $\varphi_{i,0}$ by (MP)

- If $\varphi_i \equiv \forall x \varphi_j$ with $j < i$ and $x \notin \text{free}(\psi)$, the claim follows from

$\varphi_{i,0}$:	$\psi \to \varphi_j$	since $j < i$
$\varphi_{i,1}$:	$\forall x(\psi \to \varphi_j)$	from $\varphi_{i,0}$ by $(\forall)$
$\varphi_{i,2}$:	$\forall x(\psi \to \varphi_j) \to (\psi \to \varphi_i)$	instance of L$_{12}$
$\varphi_{i,3}$:	$\psi \to \varphi_i$	from $\varphi_{i,2}$ and $\varphi_{i,1}$ by (MP)

⊣

As a first application, note that $\vdash \varphi \to \varphi$ is a trivial consequence of the Deduction Theorem, whereas its formal proof in Example 1.1 has five steps.

As a further application of the Deduction Theorem, we show that the equality relation is symmetric. We first show that $\{x = y\} \vdash y = x$:

φ_0:	$(x = y \wedge x = x) \to (x = x \to y = x)$	instance of L$_{15}$
φ_1:	$x = x$	instance of L$_{14}$
φ_2:	$x = y$	$x = y$ belongs to $\{x = y\}$
φ_3:	$x = x \to (x = y \to (x = y \wedge x = x))$	instance of L$_5$
φ_4:	$x = y \to (x = y \wedge x = x)$	from φ_3 and φ_1 by (MP)
φ_5:	$x = y \wedge x = x$	from φ_4 and φ_2 by (MP)
φ_6:	$x = x \to y = x$	from φ_0 and φ_5 by (MP)
φ_7:	$y = x$	from φ_6 and φ_1 by (MP)

Thus, we have $\{x = y\} \vdash y = x$, and by the Deduction Theorem 2.1 we see that $\vdash x = y \to y = x$, and finally, by Generalisation we get

$$\vdash \forall x \forall y (x = y \to y = x).$$

We leave it as an exercise to the reader to show that the equality relation is also transitive (see EXERCISE 1.1).

Example 2.1. We prove the tautology $\neg\neg\varphi \to \varphi$. By the DEDUCTION THEOREM it suffices to prove $\{\neg\neg\varphi\} \vdash \varphi$.

φ_0:	$\neg\neg\varphi \to (\neg\varphi \to \varphi)$	instance of L_9
φ_1:	$\neg\neg\varphi$	$\neg\neg\varphi \in \{\neg\neg\varphi\}$
φ_2:	$\neg\varphi \to \varphi$	from φ_0 and φ_1 by (MP)
φ_3:	$(\varphi \to \varphi) \to ((\neg\varphi \to \varphi) \to ((\varphi \vee \neg\varphi) \to \varphi))$	instance of L_8
φ_4:	$\varphi \to \varphi$	by Example 1.1
φ_5:	$(\neg\varphi \to \varphi) \to ((\varphi \vee \neg\varphi) \to \varphi)$	from φ_3 and φ_4 by (MP)
φ_6:	$(\varphi \vee \neg\varphi) \to \varphi$	from φ_5 and φ_2 by (MP)
φ_7:	$\varphi \vee \neg\varphi$	instance of L_0
φ_8:	φ	from φ_6 and φ_7 by (MP)

Natural Deduction

We have introduced predicate logic so that there are many logical axioms and only two inference rules. However, it is also possible to introduce calculi with an opposite approach: few axioms and many inference rules. In the calculus of **natural deduction** there are, in fact, no axioms at all. Its inference rules essentially state how to transform a given formal proof to another one. We write $\boldsymbol{\Phi} \vdash \varphi$ to state that there is a formal proof of φ in the calculus of natural deduction with the non-logical axioms given by $\boldsymbol{\Phi}$.

Let $\boldsymbol{\Phi}$ be a set of formulae and let φ, ψ, χ be any formulae. The first rule states how formal proofs can be initialized.

$$\text{INITIAL RULE (IR):} \quad \frac{}{\boldsymbol{\Phi} \vdash \varphi} \quad \text{for } \varphi \in \boldsymbol{\Phi}.$$

In the calculus of natural deduction, for each logical symbol, there are **introduction rules** and **elimination rules**.

$$(\mathrm{I}\wedge): \quad \frac{\boldsymbol{\Phi} \vdash \varphi \text{ and } \boldsymbol{\Phi} \vdash \psi}{\boldsymbol{\Phi} \vdash \varphi \wedge \psi} \qquad (\mathrm{E}\wedge): \quad \frac{\boldsymbol{\Phi} \vdash \varphi \wedge \psi}{\boldsymbol{\Phi} \vdash \varphi \text{ and } \boldsymbol{\Phi} \vdash \psi}$$

$$(\mathrm{I}\vee): \quad \frac{\boldsymbol{\Phi} \vdash \varphi \text{ or } \boldsymbol{\Phi} \vdash \psi}{\boldsymbol{\Phi} \vdash \varphi \vee \psi} \qquad (\mathrm{E}\vee): \quad \frac{\boldsymbol{\Phi} \vdash \varphi \vee \psi, \boldsymbol{\Phi} + \varphi \vdash \chi \text{ and } \boldsymbol{\Phi} + \psi \vdash \chi}{\boldsymbol{\Phi} \vdash \chi}$$

$$(\mathrm{I}\!\to): \quad \frac{\boldsymbol{\Phi} + \{\varphi\} \vdash \psi}{\boldsymbol{\Phi} \vdash \varphi \to \psi} \qquad (\mathrm{E}\!\to): \quad \frac{\boldsymbol{\Phi} \vdash \varphi \to \psi \text{ and } \boldsymbol{\Phi} \vdash \varphi}{\boldsymbol{\Phi} \vdash \psi}$$

$$\text{(I}\neg\text{):}\quad \frac{\boldsymbol{\Phi} + \varphi \boldsymbol{\vdash} \psi \wedge \neg\psi}{\boldsymbol{\Phi} \boldsymbol{\vdash} \neg\varphi} \qquad\qquad \text{(E}\neg\text{):}\quad \frac{\boldsymbol{\Phi} \boldsymbol{\vdash} \neg\neg\varphi}{\boldsymbol{\Phi} \boldsymbol{\vdash} \varphi}$$

Let τ be a term and ν be a variable such that the substitution $\varphi(\nu/\tau)$ is admissible and $\nu \notin \operatorname{free}(\chi)$ for any formula $\chi \in \boldsymbol{\Phi}$ and – in the case of (E$\exists$) – $\nu \notin \operatorname{free}(\psi)$. Now we can state the corresponding introduction and elimination rules for quantifiers:

$$\text{(I}\exists\text{):}\quad \frac{\boldsymbol{\Phi} \boldsymbol{\vdash} \varphi(\tau)}{\boldsymbol{\Phi} \boldsymbol{\vdash} \exists\nu\varphi(\nu)} \qquad\qquad \text{(E}\exists\text{):}\quad \frac{\boldsymbol{\Phi} \boldsymbol{\vdash} \exists\nu\varphi(\nu) \text{ and } \boldsymbol{\Phi} + \varphi(\nu) \boldsymbol{\vdash} \psi}{\boldsymbol{\Phi} \boldsymbol{\vdash} \psi}$$

$$\text{(I}\forall\text{):}\quad \frac{\boldsymbol{\Phi} \boldsymbol{\vdash} \varphi(\nu)}{\boldsymbol{\Phi} \boldsymbol{\vdash} \forall\nu\varphi(\nu)} \qquad\qquad \text{(E}\forall\text{):}\quad \frac{\boldsymbol{\Phi} \boldsymbol{\vdash} \forall\nu\varphi(\nu)}{\boldsymbol{\Phi} \boldsymbol{\vdash} \varphi(\tau)}$$

Finally, we need to deal with equality and atomic formulae. Let τ, τ_1 and τ_2 be terms and φ an atomic formula. The following introduction and elimination rules for equality are closely related to the logical axioms L_{14}– L_{16}:

$$\text{(I}{=}\text{):}\quad \frac{}{\tau = \tau} \qquad\qquad \text{(E}{=}\text{):}\quad \frac{\boldsymbol{\Phi} \boldsymbol{\vdash} \tau_1 = \tau_2 \text{ and } \boldsymbol{\Phi} \boldsymbol{\vdash} \varphi}{\boldsymbol{\Phi} \boldsymbol{\vdash} \varphi(\tau_1/\tau_2)}$$

Formal proofs in the calculus of natural deduction are defined in a similar way as in our usual calculus: There is a formal proof of a formula φ from set of formulae $\boldsymbol{\Phi}$, denoted $\boldsymbol{\Phi} \boldsymbol{\vdash} \varphi$, if there is a FINITE sequence of of pairs $(\boldsymbol{\Phi}_0, \varphi_0), \ldots, (\boldsymbol{\Phi}_n, \varphi_n)$ such that $\boldsymbol{\Phi}_n \equiv \boldsymbol{\Phi}$, $\varphi_n \equiv \varphi$ and for each $i \leq n$, $\boldsymbol{\Phi}_i \boldsymbol{\vdash} \varphi_i$ is obtained by the application of an inference rule

$$\frac{\boldsymbol{\Phi}_{j_0} \boldsymbol{\vdash} \varphi_{j_0}, \ldots, \boldsymbol{\Phi}_{j_k} \boldsymbol{\vdash} \varphi_{j_k}}{\boldsymbol{\Phi}_i \boldsymbol{\vdash} \varphi_i}$$

with $k \leq 3$ and $j_0, \ldots, j_k < i$. Note that the the case $k = 0$ is permitted, which corresponds to an application of the INITIAL RULE.

We have now described two ways of introducing formal proofs. It is therefore natural to ask whether the two systems prove the same theorems. Fortunately, this question turns out to have a positive answer.

THEOREM 2.2. *Let $\boldsymbol{\Phi}$ be a set of formulae and let φ be a formula. Then*

$$\boldsymbol{\Phi} \vdash \varphi \Longleftrightarrow \boldsymbol{\Phi} \boldsymbol{\vdash} \varphi.$$

Proof. We need to verify that every formal proof in the usual sense can be turned into a formal proof in the calculus of natural deduction and vice versa. In order to prove that $\boldsymbol{\Phi} \vdash \varphi$ implies $\boldsymbol{\Phi} \boldsymbol{\vdash} \varphi$ for every formula φ, we need to derive all introduction and elimination rules from our logical axioms and (MP) and $(\forall)$. We focus only on some of the rules and leave the others as an exercise.

Formal proofs of the form $\boldsymbol{\Phi} \boldsymbol{\vdash} \varphi$ with $\varphi \in \boldsymbol{\Phi}$ using only (IR) obviously correspond to trivial formal proofs of the form $\boldsymbol{\Phi} \vdash \varphi$. We consider the more interesting elimination rule (E$\vee$). Suppose that $\mathbf{\Phi} \vdash \varphi \vee \psi, \mathbf{\Phi} + \varphi \vdash \chi$ and $\mathbf{\Phi} + \psi \vdash \chi$. We verify that $\mathbf{\Phi} \vdash \chi$.

φ_0:	$\varphi \to \chi$	from $\mathbf{\Phi} + \varphi \vdash \chi$ by (DT)
φ_1:	$\psi \to \chi$	from $\mathbf{\Phi} + \psi \vdash \chi$ by (DT)
φ_2:	$(\varphi \to \chi) \to ((\psi \to \chi) \to ((\varphi \vee \psi) \to \chi))$	instance of L_8
φ_3:	$(\psi \to \chi) \to ((\varphi \vee \psi) \to \chi)$	from φ_2 and φ_0 by (MP)
φ_4:	$(\varphi \vee \psi) \to \chi$	from φ_3 and φ_1 by (MP)
φ_4:	$\varphi \vee \psi$	by assumption
φ_5:	χ	from φ_4 and φ_5 by (MP)

The corresponding introduction rule (I$\vee$) follows directly from L_6 and L_7 using (DT). Note that (I$\to$) follows directly from (DT) and (E$\to$) from (MP).

We further prove the rules for negation. For (I$\neg$) suppose that $\mathbf{\Phi} + \varphi \vdash \psi \wedge \neg\psi$. It follows from (E$\wedge$) that $\mathbf{\Phi} + \varphi \vdash \psi$ and $\mathbf{\Phi} + \varphi \vdash \neg\psi$. We prove that $\mathbf{\Phi} + \varphi \vdash \neg\varphi$, since then $\mathbf{\Phi} \vdash \neg\varphi$ by (E$\vee$) and L_0. We have

φ_0:	$\neg\psi \to (\psi \to \neg\varphi)$	instance of L_9
φ_1:	$\neg\psi$	by assumption
φ_2:	$\psi \to \neg\varphi$	from φ_0 and φ_2 by (MP)
φ_3:	ψ	by assumption
φ_4:	$\neg\varphi$	from φ_2 and φ_3 by (MP)

The corresponding elimination rule (E$\neg$) follows from Example 2.1. Finally, we prove (I$\exists$) and (E$\exists$). Note that (I$\exists$) follows directly from L_{11} using (DT) and (MP). For (E$\exists$) suppose that $\mathbf{\Phi} \vdash \exists\nu\varphi(\nu)$ and $\mathbf{\Phi} + \varphi(\nu) \vdash \psi$. An application of (DT) then yields $\mathbf{\Phi} \vdash \varphi(\nu) \to \psi$. Then we have

φ_0:	$\forall\nu(\varphi(\nu) \to \psi) \to (\exists\nu\varphi(\nu) \to \psi)$	instance of L_{13}
φ_1:	$\varphi(\nu) \to \psi$	by assumption
φ_2:	$\forall\nu(\varphi(\nu) \to \psi)$	from φ_1 by ($\forall$)
φ_3:	$\exists\nu\varphi(\nu) \to \psi$	from φ_0 and φ_1 by (MP)
φ_4:	$\exists\nu\varphi(\nu)$	by assumption
φ_5:	ψ	from φ_3 and φ_4 by (MP)

This completes the proof of (E$\exists$). The verification of the other rules of the calculus of natural deduction are left to the reader (see Exercise 2.0).

Conversely, we need to check that the calculus of natural deduction proves the logical axioms L_0– L_{16} as well as the inference rules (MP) and ($\forall$). Observe that (MP) corresponds to (E$\to$) and ($\forall$) corresponds to (I$\forall$). As before, we only present the proof for some axioms and leave the others to the reader. We consider first L_9. We need to check that $\vdash \neg\varphi \to (\varphi \to \psi)$.

$$\begin{array}{rll}
\{\neg\varphi, \varphi, \neg\psi\} & \vdash \varphi & \text{by } \textsf{(IR)} \\
\{\neg\varphi, \varphi, \psi\} & \vdash \neg\varphi & \text{by } \textsf{(IR)} \\
\{\neg\varphi, \varphi, \psi\} & \vdash \varphi \wedge \neg\varphi & \text{by } \textsf{(I}\wedge\textsf{)} \\
\{\neg\varphi, \varphi\} & \vdash \neg\neg\psi & \text{by } \textsf{(I}\neg\textsf{)} \\
\{\neg\varphi, \varphi\} & \vdash \psi & \text{by } \textsf{(E}\neg\textsf{)} \\
\{\neg\varphi\} & \vdash \varphi \to \psi & \text{by } \textsf{(I}\to\textsf{)} \\
 & \vdash \neg\varphi \to (\varphi \to \psi) & \text{by } \textsf{(I}\to\textsf{)}
\end{array}$$

Secondly, we derive Axiom L$_{13}$, *i.e.* $\vdash \forall\nu(\varphi(\nu) \to \psi) \to (\exists\nu\varphi(\nu) \to \psi)$.

$$\begin{array}{rll}
\{\forall\nu(\varphi(\nu) \to \psi), \exists\nu\varphi(\nu), \varphi(\nu)\} & \vdash \varphi(\nu) & \text{by } \textsf{(IR)} \\
\{\forall\nu(\varphi(\nu) \to \psi), \exists\nu\varphi(\nu)\} & \vdash \exists\nu\varphi(\nu) & \text{by } \textsf{(IR)} \\
\{\forall\nu(\varphi(\nu) \to \psi), \exists\nu\varphi(\nu), \varphi(\nu)\} & \vdash \forall\nu(\varphi(\nu) \to \psi) & \text{by } \textsf{(IR)} \\
\{\forall\nu(\varphi(\nu) \to \psi), \exists\nu\varphi(\nu), \varphi(\nu)\} & \vdash \varphi(\nu) \to \psi & \text{by } \textsf{(E}\forall\textsf{)} \\
\{\forall\nu(\varphi(\nu) \to \psi), \exists\nu\varphi(\nu), \varphi(\nu)\} & \vdash \psi & \text{by } \textsf{(E}\to\textsf{)} \\
\{\forall\nu(\varphi(\nu) \to \psi), \exists\nu\varphi(\nu)\} & \vdash \psi & \text{by } \textsf{(E}\exists\textsf{)} \\
\{\forall\nu(\varphi(\nu) \to \psi) & \vdash \exists\nu\varphi(\nu) \to \psi & \text{by } \textsf{(I}\to\textsf{)} \\
 & \vdash \forall\nu(\varphi(\nu) \to \psi) \to (\exists\nu\varphi(\nu) \to \psi) & \text{by } \textsf{(I}\to\textsf{)}.
\end{array}$$

The other axioms can be verified in a similar way. ⊣

Methods of Proof

The inference rules of the calculus of natural deduction are very useful because they resemble methods of proof commonly used in mathematics. For example, the elimination rule (E∨) mimicks proofs by case distinction: Under the assumption that $\mathbf{\Phi} \vdash \varphi \vee \psi$, one can prove a formula χ by separately proving $\mathbf{\Phi} \cup \{\varphi\} \vdash \chi$ and $\mathbf{\Phi} \cup \{\psi\} \vdash \chi$.

In the following we list several methods of proof such as proofs by contradiction, contraposition and case distinction.

PROPOSITION 2.3 (PROOF BY CASES). *Let* $\mathbf{\Phi}$ *be a set of* $\mathscr{L}$*-formulae and let* φ, ψ, χ *be some* $\mathscr{L}$*-formulae. Then the following four statements hold:*

$$\mathbf{\Phi} \vdash \varphi \vee \psi \ \textit{ and } \ \mathbf{\Phi} + \varphi \vdash \chi \ \textit{ and } \ \mathbf{\Phi} + \psi \vdash \chi \quad \Longrightarrow \quad \mathbf{\Phi} \vdash \chi \qquad (\vee 0)$$

$$\mathbf{\Phi} + \varphi \vdash \chi \ \textit{ and } \ \mathbf{\Phi} + \neg\varphi \vdash \chi \quad \Longrightarrow \quad \mathbf{\Phi} \vdash \chi \qquad (\vee 1)$$

Proof. Note that ($\vee 0$) is exactly the statement of (E∨) and ($\vee 1$) is a special case of ($\vee 0$), since $\mathbf{\Phi} \vdash \varphi \vee \neg\varphi$ by L$_0$. ⊣

COROLLARY 2.4 (Generalised Proof by Cases). *Let* $\mathbf{\Phi}$ *be a set of* $\mathscr{L}$*-formulae and let* $\psi_0, \dots, \psi_n, \varphi$ *some* $\mathscr{L}$*-formulae. Then we have:*

$$\mathbf{\Phi} \vdash \psi_0 \vee \cdots \vee \psi_n \ \textit{ and } \ \mathbf{\Phi} + \psi_i \vdash \varphi \textit{ for all } i \leq n \quad \Longrightarrow \quad \mathbf{\Phi} \vdash \varphi.$$

Since Corollary 2.4 is just a generalization of ($\vee 0$), we will also denote all instance of this form by ($\vee 1$).

Proof of Corollary 2.4. We proceed by induction on $n \geq 1$. For $n = 1$ the statement is exactly ($\vee 0$). Now assume that $\mathbf{\Phi} \vdash \psi_0 \vee \ldots \vee \psi_n \vee \psi_{n+1}$ and $\mathbf{\Phi} + \psi_i \vdash \varphi$ for all $i \leq n+1$. Let $\mathbf{\Phi}' :\equiv \mathbf{\Phi} + \psi_0 \vee \ldots \vee \psi_n$ and observe that $\mathbf{\Phi}' \vdash \psi_0 \vee \ldots \vee \psi_n$ and $\mathbf{\Phi}' + \psi_i \vdash \varphi$, so by induction hypothesis $\mathbf{\Phi}' \vdash \varphi$. By the DEDUCTION THEOREM this implies $\mathbf{\Phi} \vdash \psi_0 \vee \ldots \vee \psi_n \to \varphi$. Moreover, by another application of (DT) we also have $\mathbf{\Phi} \vdash \psi_{n+1} \to \varphi$. Using L_8 and twice (DT) we obtain $\mathbf{\Phi} \vdash \psi_0 \vee \ldots \vee \psi_n \vee \psi_{n+1} \to \varphi$, hence (DT) yields the claim. ⊣

PROPOSITION 2.5 (EX FALSO QUODLIBET). *Let $\mathbf{\Phi}$ be a set of $\mathscr{L}$-formulae and let φ an arbitrary $\mathscr{L}$-formula. Then for every $\mathscr{L}$-formula ψ we have:*

$$\mathbf{\Phi} \vdash \varphi \wedge \neg\varphi \quad \Longrightarrow \quad \mathbf{\Phi} \vdash \psi \qquad (\text{⧉})$$

Proof. Let ψ be any $\mathscr{L}$-formula and assume that $\mathbf{\Phi} \vdash \varphi \wedge \neg\varphi$ for some $\mathscr{L}$-formula φ. By (E$\wedge$) we have $\mathbf{\Phi} \vdash \varphi$ and $\mathbf{\Phi} \vdash \neg\varphi$. Now the instance $\neg\varphi \to (\varphi \to \psi)$ of the logical axiom L_9 and two applications of MODUS PONENS imply $\mathbf{\Phi} \vdash \psi$. ⊣

Notice that PROPOSITION 2.5 implies that if we can derive a contradiction from $\mathbf{\Phi}$, we can derive *every* formula we like, even the impossible, denoted by the symbol

⧉ .

This is closely related to proofs by contradiction:

COROLLARY 2.6 (PROOF BY CONTRADICTION). *Let $\mathbf{\Phi}$ be a set of formulae, and φ be an arbitrary formula. Then the following statements hold:*

$$\mathbf{\Phi} + \neg\varphi \vdash \text{⧉} \quad \Longrightarrow \quad \mathbf{\Phi} \vdash \varphi,$$

$$\mathbf{\Phi} + \varphi \vdash \text{⧉} \quad \Longrightarrow \quad \mathbf{\Phi} \vdash \neg\varphi.$$

Proof. Note that the second statement is exactly the introduction rule (I$\neg$). For the first statment, note that by ($\vee 1$) it is enough to check $\mathbf{\Phi} + \varphi \vdash \varphi$ and $\mathbf{\Phi} + \neg\varphi \vdash \varphi$. The first condition is clearly satisfied and the second one follows directly from (I$\wedge$) and (⧉). ⊣

PROPOSITION 2.7 (PROOF BY CONTRAPOSITION). *Let $\mathbf{\Phi}$ be a set of $\mathscr{L}$-formulae and φ and ψ two arbitrary $\mathscr{L}$-formulae. Then we have:*

$$\mathbf{\Phi} + \varphi \vdash \psi \quad \Longleftrightarrow \quad \mathbf{\Phi} + \neg\psi \vdash \neg\varphi \qquad (\mathrm{CP})$$

Proof. Suppose first that $\mathbf{\Phi} + \varphi \vdash \psi$. Then by (I$\wedge$), $\mathbf{\Phi} \cup \{\neg\psi, \varphi\} \vdash \psi \wedge \neg\psi$ and hence by (I$\neg$) we obtain $\mathbf{\Phi} + \neg\psi \vdash \neg\varphi$.

Conversely, assume that $\mathbf{\Phi} + \neg\psi \vdash \neg\varphi$. A similar argument as above yields $\mathbf{\Phi} + \varphi \vdash \neg\neg\psi$. An application of (E$\neg$) completes the proof. ⊣

Note that Proposition 2.7 proves the logical equivalence

$$\varphi \to \psi \Leftrightarrow \neg\psi \to \neg\varphi.$$

THEOREM 2.8 (GENERALISED DEDUCTION THEOREM). *If $\mathbf{\Phi}$ is an arbitrary set of formulae and $\mathbf{\Phi} \cup \{\psi_1, \dots, \psi_n\} \vdash \varphi$, where in the proof of φ from $\mathbf{\Phi} \cup \{\psi_1, \dots, \psi_n\}$ the rule of* GENERALISATION *is not applied to any of the free variables of $\psi_1, \dots, \psi_n$, then $\mathbf{\Phi} \vdash (\psi_1 \wedge \dots \wedge \psi_n) \to \varphi$; and vice versa:*

$$\mathbf{\Phi} \cup \{\psi_1, \dots, \psi_n\} \vdash \varphi \quad \Longleftrightarrow \quad \mathbf{\Phi} \vdash (\psi_1 \wedge \dots \wedge \psi_n) \to \varphi \qquad \text{(GDT)}$$

Proof. Follows immediately from the DEDUCTION THEOREM and from Part (c) of DEMORGAN'S LAWS (see Exercise 2.2). ⊣

Normal forms

DNF, CNF and PNF

In many proofs it is convenient to convert an $\mathscr{L}$-formula into an equivalent formula in some normal form. The simplest normal form is the following: An $\mathscr{L}$-formula is said to be in **Negation Normal Form**, denoted NNF, if the negation symbol $\neg$ only occurs directly in front of atomic subformulae.

THEOREM 2.9. *Every $\mathscr{L}$-formula is equivalent to some $\mathscr{L}$-formula in* NNF.

Proof. We successively apply the following transformations to every non-atomic negated subformula ψ of φ, starting with the outermost negation symbols.

- If $\psi \equiv \neg\neg\psi'$ for some formula ψ', we replace $\neg\neg\psi'$ with ψ' using (F.0).
- By the DEMORGAN'S LAWS (see Chapter 2 | Exercise 2.2), we replace subformulae of the form $\neg(\psi_1 \wedge \psi_2)$ and $\neg(\psi_1 \vee \psi_2)$ respectively, with $\neg\psi_1 \vee \neg\psi_2$ and $\neg\psi_1 \wedge \neg\psi_2$ respectively.
- If $\psi \equiv \neg\exists x\psi'$ then it follows from (S.1) that $\psi \Leftrightarrow \forall x\neg\psi'$, and hence we replace ψ with $\forall x\neg\psi'$. Similarly, using (S.2), we replace subformulae of the form $\neg\forall x\psi'$ with the equivalent formula $\exists x\neg\psi'$.

⊣

An quantifier-free $\mathscr{L}$-formula φ is said to be in **Disjunctive Normal Form**, if it is a conjunction of a disjunction of atomic formulae or negated atomic formulae, *i.e.* if it is of the form

$$(\varphi_{1,1} \wedge \ldots \wedge \varphi_{1,k_1}) \vee \dots \vee (\varphi_{m,1} \wedge \dots \wedge \varphi_{m,k_m})$$

for some quantifier-free $\mathscr{L}$-formulae $\varphi_{i,j}$ which are either atomic or the negation of an atomic formula. In particular, each formula in DNF is also in NNF.

Theorem 2.10 (Disjunctive Normal Form Theorem). *Every quantifier-free $\mathscr{L}$-formula φ is equivalent to some $\mathscr{L}$-formula in DNF.*

Proof. By Theorem **??** we may assume that φ is in NNF. Starting with the outermost conjunction symbol, we successively apply the distributive laws

$$\begin{aligned} &\psi \wedge (\varphi_1 \vee \varphi_2) \Leftrightarrow (\psi \wedge \varphi_1) \vee (\psi \wedge \varphi_2) \quad \text{and} \\ &(\varphi_1 \vee \varphi_2) \wedge \psi \Leftrightarrow (\varphi_1 \wedge \psi) \vee (\varphi_2 \wedge \psi) \end{aligned}$$

until all conjunction symbols occur between atomic or negated atomic formulae. This process ends after FINITELY many steps, since there are only FINITELY many conjunction symbols. ⊣

An $\mathscr{L}$-sentence σ said to be in **Prenex Normal Form**, denoted PNF, if it is of the form

$$\exists\!\!\!\!\forall_0 \nu_0 \ldots \exists\!\!\!\!\forall_n \nu_n \tilde{\sigma},$$

where the variables $\nu_0, \ldots, \nu_n$ are pairwise distinct, each $\exists\!\!\!\!\forall_i$ stands either for "$\exists$" or for "$\forall$", and $\tilde{\sigma}$ is a quantifier-free formula. Furthermore, an $\mathscr{L}$-sentence σ is in **special prenex normal form**, denoted sPNF, if σ is in PNF and

$$\sigma \equiv \exists\!\!\!\!\forall_0 v_0 \exists\!\!\!\!\forall_1 v_1 \ldots \exists\!\!\!\!\forall_n v_n \tilde{\sigma}$$

where each $\exists\!\!\!\!\forall_m$ (for $0 \leq m \leq n$) stands for either "$\exists$" or "$\forall$", $\tilde{\sigma}$ is quantifier free, and each variable $v_0, \ldots, v_n$ appears free in $\tilde{\sigma}$.

Theorem 2.11 (Prenex Normal Form Theorem). *For every $\mathscr{L}$-sentence σ there is a semantically equivalent $\mathscr{L}$-sentence in* sPNF*.*

Proof. By Theorem 1.2 we may assume that σ does not contain the symbol $\rightarrow$. We describe an algorithm which transforms the $\mathscr{L}$-sentence σ into an $\mathscr{L}$-sentence in sPNF.

Step 1. By Theorem 2.9, we can transform σ into an equivalent $\mathscr{L}$-sentence in NNF.

Step 2. We pull all quantifiers outwards: Suppose by induction that all subformulae of φ are already in PNF. If φ is of the form $\exists x \varphi'$ or $\forall x \varphi'$, then φ is also in PNF. Hence we may assume that $\varphi \equiv \varphi_1 \circ \varphi_2$ for some formulae φ_1 and φ_2 which are in PNF, and $\circ$ is either $\wedge$ or $\vee$, then by the Variable Substitution Theorem 3.9 we may assume that φ_1 and φ_2 do not have any variables in common and that the variables occurring in φ_1 and φ_2 are $v_0, \ldots, v_n$ and $v_{n+1}, \ldots, v_m$ respectively. If

$$\begin{aligned} &\varphi_1 \equiv \exists\!\!\!\!\forall_0 v_0 \ldots \exists\!\!\!\!\forall_n v_n \varphi_1' \quad \text{and} \\ &\varphi_2 \equiv \exists\!\!\!\!\forall_{n+1} v_{n+1} \ldots \exists\!\!\!\!\forall_m v_m \varphi_2', \end{aligned}$$

where φ_1' and φ_2' are quantifier-free, then, using Tautologies (V.4), (V.5), (W.4) and (W.5), we obtain that φ is semantically equivalent to

$$\rlap{\exists}\forall_0 v_0 \ldots \rlap{\exists}\forall_m v_m (\varphi_1' \circ \varphi_2') \,.$$

Hence φ is semantically equivalent to a formula in sPNF.

Observe that after each transformation, the resulting formula remains semantically equivalent to the original one as a consequence of Theorem 1.1, and therefore, the $\mathscr{L}$-sentence σ is semantically equivalent to an $\mathscr{L}$-sentence in sPNF. ⊣

Consistency & Compactness

Let $\mathbf{\Phi}$ be a set of $\mathscr{L}$-formulae. We say that $\mathbf{\Phi}$ is **consistent**, denoted $\mathrm{Con}(\mathbf{\Phi})$, if $\mathbf{\Phi} \nvdash \text{⧉}$, *i.e.* if there is *no* $\mathscr{L}$-formula φ such that $\mathbf{\Phi} \vdash (\varphi \wedge \neg\varphi)$, otherwise $\mathbf{\Phi}$ is called **inconsistent**, denoted $\neg\,\mathrm{Con}(\mathbf{\Phi})$.

FACT 2.12. *Let $\mathbf{\Phi}$ be a set of $\mathscr{L}$-formulae.*

(a) *If $\neg\,\mathrm{Con}(\mathbf{\Phi})$, then for all $\mathscr{L}$-formulae ψ we have $\mathbf{\Phi} \vdash \psi$.*

(b) *If $\mathrm{Con}(\mathbf{\Phi})$ and $\mathbf{\Phi} \vdash \varphi$ for some $\mathscr{L}$-formula φ, then $\mathbf{\Phi} \nvdash \neg\varphi$.*

(c) *If $\neg\,\mathrm{Con}(\mathbf{\Phi} + \varphi)$, for some $\mathscr{L}$-formula φ, then $\mathbf{\Phi} \vdash \neg\varphi$.*

(d) *If $\mathbf{\Phi} \vdash \neg\varphi$, for some $\mathscr{L}$-formula φ, then $\neg\,\mathrm{Con}(\mathbf{\Phi} + \varphi)$.*

Proof. Condition (a) is just PROPOSITION 2.5. For (b), notice that if $\mathbf{\Phi} \vdash \varphi$ and $\mathbf{\Phi} \vdash \neg\varphi$, then by (I∧) we get $\mathbf{\Phi} \vdash \text{⧉}$ and thus also $\neg\,\mathrm{Con}(\mathbf{\Phi})$. Moreover, (c) coincides with the second statement of Corollary 2.6. Finally, for (d) note that if $\mathbf{\Phi} \vdash \neg\varphi$, then $\mathbf{\Phi} + \varphi \vdash \varphi \wedge \neg\varphi$ and hence $\mathbf{\Phi} + \varphi$ is inconsistent. ⊣

If we choose a set of formulae $\mathbf{\Phi}$ as the basis of a theory (*e.g.*, a set of axioms), we have to make sure that $\mathbf{\Phi}$ is consistent. However, as we shall see later, in many cases this task is impossible.

We conclude this chapter with the COMPACTNESS THEOREM, which is a powerful tool in order to construct non-standard models of Peano Arithmetic or of Set Theory. On the one hand, it is just a consequence of the fact that formal proofs are FINITE sequences of formulae. On the other hand, the COMPACTNESS THEOREM is the main tool to prove that a given set of sentences is consistent with some given set of formulae $\mathbf{\Phi}$.

THEOREM 2.13 (COMPACTNESS THEOREM). *Let $\mathbf{\Phi}$ be an arbitrary set of formulae. Then $\mathbf{\Phi}$ is consistent if and only if every finite subset $\mathbf{\Phi}'$ of $\mathbf{\Phi}$ is consistent.*

Proof. Obviously, if $\mathbf{\Phi}$ is consistent, then every finite subset $\mathbf{\Phi}'$ of $\mathbf{\Phi}$ must be consistent. On the other hand, if $\mathbf{\Phi}$ is inconsistent, then there is a formula φ such that $\mathbf{\Phi} \vdash \varphi \wedge \neg\varphi$. In other words, there is a proof of $\varphi \wedge \neg\varphi$ from $\mathbf{\Phi}$. Now, since every proof is finite, there are only finitely many formulae of $\mathbf{\Phi}$ involved in this proof, and if $\mathbf{\Phi}'$ is this finite set of formulae, then $\mathbf{\Phi}' \vdash \varphi \wedge \neg\varphi$, which shows that $\mathbf{\Phi}'$, a finite subset of $\mathbf{\Phi}$, is inconsistent. ⊣

Semi-formal Proofs

Previously we have shown that formal proofs can be simplified by applying methods of proof such as case distinctions, proofs by contradiction or contraposition. However, to make proofs even more natural, it is useful to use natural language in order to describe a proof step as in an "informal" mathematical proof.

Example 2.1 We want to prove the tautology $\vdash \varphi \rightarrow \neg\neg\varphi$. Instead of writing out the whole formal proof which is quite tedious, we can apply our methods of proof introduced above.

The first modification we make is to use (DT) to obtain the new goal

$$\{\varphi\} \vdash \neg\neg\varphi.$$

The easiest way to proceed now is to make a proof by contradiction; hence it remains to show

$$\{\varphi, \neg\varphi\} \vdash \text{⧉}$$

which by (I∧) is again a consequence of the trivial goals

$$\{\varphi, \neg\varphi\} \vdash \varphi \quad \text{and} \quad \{\varphi, \neg\varphi\} \vdash \neg\varphi.$$

To sum up, this procedure can actually be transformed back into a formal proof, so it suffices as a proof of $\vdash \varphi \rightarrow \neg\neg\varphi$. Now this is still not completely satisfactory, since we would like to write the proof in natural language. A possible translation could thus be the following:

Proof. We want to prove that φ implies $\neg\neg\varphi$. Assume φ. Suppose for a contradiction that $\neg\varphi$. But then we have φ and $\neg\varphi$. Contradiction. $\dashv$

We will now show in a systematic way how formal proofs can - in principal - be replaced by semi-formal proofs, which make use of a **controlled natural language**, *i.e.*, a limited vocabulary consisting of natural language phrases such as "assume that" which are often used in mathematical proof texts. This language is controlled in the sense that its allowed vocabulary is only a subset of the entire English vocabulary and that every word resp. phrase has a unique precisely defined interpretation. However, for the sake of a nice proof style, we will not always stick to this limited vocabulary. Moreover, this section should be considered as a hint of how formal proofs can be formulated using a controlled natural language as well as a justification for working with natural language proofs rather than formal ones.

Every statement we would like to prove formally is of the form $\mathbf{\Phi} \vdash \varphi$, where $\mathbf{\Phi}$ is a set of formulae and φ is a formula. Note that as in Example 2.1, in order to prove $\mathbf{\Phi} \vdash \varphi$—which is actually a meta-proof—we perform operations both on the set of formulae $\mathbf{\Phi}$ and on the formula to be formally proved. We call a statement of the form $\mathbf{\Phi} \vdash \varphi$ a **goal**, and denote the set of non-logical axioms $\mathbf{\Phi}$ as **premises** and the formula φ to be verified as **target**. Now instead of listing a formal proof, we can step by step reduce our current goal to a simpler one using the methods of proof from the previous section. We can follow this procedure until the target is tautological as in the case of Example 2.1.

Methods of proof are in that sense simply operations on the premises and the targets. The proof by contraposition for example adds the negation of the target to the premises and replaces the original target by the negation of the premise from which it shall be derived:

If we want to show

$$\mathbf{\Phi} + \psi \vdash \varphi,$$

we can prove instead

$$\mathbf{\Phi} + \neg\varphi \vdash \neg\psi.$$

A slightly different example is the proof of a conjunction

$$\mathbf{\Phi} \vdash \varphi \wedge \psi,$$

which is usually split into the two goals given by

$$\mathbf{\Phi} \vdash \varphi \text{ and } \mathbf{\Phi} \vdash \psi.$$

Thus we have to revise our first attempt and interpret methods of proof as operations on FINITE lists of goals consisting of premises and targets.

We distinguish between two types of operations on goals. **Backward reasoning** means performing operations on targets, whereas **forward reasoning** denotes operations on the premises. We give some examples of both backward and forward reasoning and indicate how such proofs can be phrased in a semi-formal way.

Backward reasoning

- Targets are often of the universal conditional form $\forall\nu(\varphi(\nu) \rightarrow \psi(\nu))$. In particular, this pattern includes the purely universal formulae $\forall\nu\psi(\nu)$ by taking φ to be a tautology as well as simple conditionals of the form $\varphi \rightarrow \psi$. Now the usual procedure is to reduce $\mathbf{\Phi} \vdash \forall\nu(\varphi(\nu) \rightarrow \psi(\nu))$ to $\mathbf{\Phi} + \varphi(\nu) \vdash \psi(\nu)$ using $(\forall)$ and (DT). This can be rephrased as

  Assume $\varphi(\nu)$. *Then ... This shows* $\psi(\nu)$.

- As already mentioned above, if the target is a conjunction $\varphi \wedge \psi$, one can show them separately using (I∧). This step is usually executed without mentioning it explicitly.
- If the target is a negation $\neg\varphi$, one often uses a proof by contradiction or by contraposition: In the first case we transform $\mathbf{\Phi} \vdash \neg\varphi$ to $\mathbf{\Phi} + \varphi \vdash$ ⧉ and use the natural language notation

Suppose for a contradiction that φ. Then . . . Contradiction.

In the latter case, we want to go from $\mathbf{\Phi} + \neg\psi \vdash \neg\varphi$ to $\mathbf{\Phi} + \varphi \vdash \psi$ resp. in its positive version from $\mathbf{\Phi} + \psi \vdash \neg\varphi$ to $\mathbf{\Phi} + \varphi \vdash \neg\psi$. In both cases we can mark this with the keyword *contraposition*, e.g. as

We proceed by *contraposition. . . This shows* $\neg\varphi$.

Forwards reasoning

- By (E∧), conjunctive premises $\varphi \wedge \psi$ can be split into two premises φ, ψ; i.e. $\mathbf{\Phi} + \varphi \wedge \psi \vdash \chi$ can be reduced to $\mathbf{\Phi} \cup \{\varphi, \psi\} \vdash \chi$. This is usually performed automatically.
- Disjunctive premises are used for proofs by case distinction: If Given a goal of the form $\mathbf{\Phi} + \varphi \vee \psi \vdash \chi$, we reduce it to the new goals $\mathbf{\Phi} + \varphi \vdash \chi$ and $\mathbf{\Phi} + \psi \vdash \chi$. We can write this in a semi-formal way as

  Case 1: Assume φ. . . . This proves χ.
  Case 2: Assume ψ. . . . This proves χ.

- Intermediate proof steps: Often we want to prove first some intermediate statement which shall then be applied to resolve the target. Formally this means that we want to show $\mathbf{\Phi} \vdash \varphi$ by showing first $\mathbf{\Phi} \vdash \psi$ and then we add ψ to the list of premises and check $\mathbf{\Phi} + \psi \vdash \varphi$. Clearly, if we have $\mathbf{\Phi} \vdash \psi$ and $\mathbf{\Phi} + \psi \vdash \varphi$, using (DT) and (MP) we obtain that $\mathbf{\Phi} \vdash \varphi$. In a semi-formal proof this can be described by

  We show *first ψ... This proves ψ.*

  Note that it is important to mark when the proof of the intermediate statement ψ ends, because from this point on, ψ can be used as a new premise.

Observe that in any case, once a goal $\mathbf{\Phi} \vdash \varphi$ is reduced to a tautology, it can be removed from the list of goals. This should be marked by a phrase like

This shows/proves φ.

so that it is clear that we go on to the next goal. The proof is complete as soon as no unresolved goals remain.

What is the use of such a formalized natural proof language? First of all, it increases readibility. Secondly, by giving some of the common natural language phrases appearing in proof texts a precise formal definition, we show how – in principal – one could write formal proofs with a controlled natural language input. This input could then be parsed into a formal proof and then be verified by a proof checking system.

We would like to emphasize that this section should only be considered a motivation rather than a precise description of how formal proofs can be translated into semi-formal ones and vice versa. Nevertheless, it suffices to understand how this can theoretically be achieved. Therefore, in subsubsequent chapters, especially in Chapters **??** and **??**, we will often present semi-formal proofs rather than formal ones.

Notes

Give some references.

Exercises

2.0 Complete the proof of Theorem 2.2.

2.1 Formalize the method of proof by counterexample and prove that it works.

2.2 Let $\varphi_0, \dots, \varphi_n$ be formulae. Prove the DEMORGAN'S LAWS:

(a) $\neg(\varphi_0 \wedge \cdots \wedge \varphi_n) \Leftrightarrow (\neg\varphi_1 \vee \cdots \vee \neg\varphi_n)$
(b) $\neg(\varphi_0 \vee \cdots \vee \varphi_n) \Leftrightarrow (\neg\varphi_1 \wedge \cdots \wedge \neg\varphi_n)$
(c) $\varphi_0 \to \big(\varphi_1 \to (\cdots \to \varphi_n)\cdots\big) \Leftrightarrow \neg(\varphi_0 \wedge \cdots \wedge \varphi_n)$

2.3 Show that $\vdash (\varphi \to \psi) \to \big((\varphi \to \neg\psi) \to \neg\varphi\big)$.

2.4 Prove the following generalization of L_{15}: to an arbitrary formula φ:

$$\vdash (\tau_1 = \tau_1' \wedge \cdots \wedge \tau_n = \tau_n') \to (\varphi(\tau_1, \dots, \tau_n) \to \varphi(\tau_1', \dots, \tau_n')),$$

where $\tau, \tau_1, \dots, \tau_n, \tau_1', \dots, \tau_n'$ are terms and φ is a formula with n free variables.