3.3 Hilbert System $\mathscr{H}$

In Gentzen systems there is one axiom and many rules of inference, while in a Hilbert system there are several axioms but only one rule of inference. In this section, we define the deductive system $\mathscr{H}$ and use it to prove many theorems. Actually, only one theorem (Theorem 3.10) will be proved directly from the axioms and the rule of inference; practical use of the system depends on the use of derived rules, especially the deduction rule.

Notation: Capital letters $A, B, C, \ldots$ represent arbitrary formulas in propositional logic. For example, the notation $\vdash A \rightarrow A$ means: for *any* formula A of propositional logic, the formula $A \rightarrow A$ can be proved.

Definition 3.9 (Deductive system $\mathscr{H}$) The *axioms* of $\mathscr{H}$ are:

Axiom 1 $\vdash (A \rightarrow (B \rightarrow A))$,

Axiom 2 $\vdash (A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$,

Axiom 3 $\vdash (\neg B \rightarrow \neg A) \rightarrow (A \rightarrow B)$.

The *rule of inference* is ***modus ponens*** (*MP* for short):

$$\frac{\vdash A \qquad\qquad \vdash A \rightarrow B}{\vdash B}.$$

In words: the formula B can be inferred from A and $A \rightarrow B$.

The terminology used for $\mathscr{G}$—*premises*, *conclusion*, *theorem*, *proved*— carries over to $\mathscr{H}$, as does the symbol $\vdash$ meaning that a formula is proved. ■

Theorem 3.10 $\vdash A \rightarrow A$.

Proof

1.	$\vdash (A \rightarrow ((A \rightarrow A) \rightarrow A)) \rightarrow ((A \rightarrow (A \rightarrow A)) \rightarrow (A \rightarrow A))$	Axiom 2
2.	$\vdash A \rightarrow ((A \rightarrow A) \rightarrow A)$	Axiom 1
3.	$\vdash (A \rightarrow (A \rightarrow A)) \rightarrow (A \rightarrow A)$	MP 1, 2
4.	$\vdash A \rightarrow (A \rightarrow A)$	Axiom 1
5.	$\vdash A \rightarrow A$	MP 3, 4

■

When an axiom is given as the justification, identify which formulas are substituted for the formulas A, B, C in the definition of the axioms above.

3.3.1 *Axiom Schemes and Theorem Schemes* *

As we noted above, a capital letter can be replaced by any formula of propositional logic, so, strictly speaking, $\vdash A \rightarrow (B \rightarrow A)$ is not an axiom, and similarly, $\vdash A \rightarrow A$

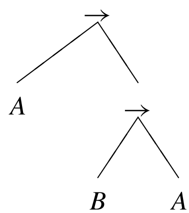
is not a theorem. A more precise terminology would be to say that $\vdash A \rightarrow (B \rightarrow A)$ is an *axiom scheme* that is a shorthand for an infinite number of axioms obtained by replacing the 'variables' A and B with actual formulas, for example:

$$\overbrace{((p \vee \neg q) \leftrightarrow r)}^{A} \rightarrow (\overbrace{\neg (q \wedge \neg r)}^{B} \rightarrow \overbrace{((p \vee \neg q) \leftrightarrow r)}^{A}).$$

Similarly, $\vdash A \rightarrow A$ is a *theorem scheme* that is a shorthand for an infinite number of theorems that can be proved in $\mathscr{H}$, including, for example:

$$\vdash ((p \vee \neg q) \leftrightarrow r) \rightarrow ((p \vee \neg q) \leftrightarrow r).$$

We will not retain this precision in our presentation because it will always clear if a given formula is an instance of a particular axiom scheme or theorem scheme. For example, a formula ϕ is an instance of Axiom 1 if it is of the form:



where there are subtrees for the formulas represented by A and B. There is a simple and efficient algorithm that checks if ϕ is of this form and if the two subtrees A are identical.

3.3.2 The Deduction Rule

The proof of Theorem 3.10 is rather complicated for such a trivial formula. In order to formalize the powerful methods of inference used in mathematics, we introduce new rules of inference called *derived rules*. The most important derived rule is the *deduction rule*. Suppose that you want to prove $A \rightarrow B$. *Assume* that A has already been proved and use it in the proof of B. This is not a proof of B unless A is an axiom or theorem that has been previously proved, in which case it can be used directly in the proof. However, we claim that the proof can be mechanically transformed into a proof of $A \rightarrow B$.

Example 3.11 The deduction rule is used frequently in mathematics. Suppose that you want to prove that *the sum of any two odd integer numbers is even*, expressed formally as:

$$odd(x) \wedge odd(y) \rightarrow even(x + y),$$

for every x and y. To prove this formula, let us *assume* the formula $odd(x) \wedge odd(y)$ as if it were an additional axiom. We have available all the theorems we have already deduced about odd numbers, in particular, the theorem that any odd number can be expressed as $2k+1$. Computing:

$$x+y=2k_1+1+2k_2+1=2(k_1+k_2+1),$$

we obtain that $x+y$ is a multiple of 2, that is, $even(x+y)$. The theorem now follows from the deduction rule which *discharges* the assumption. ■

To express the deduction rule, we extend the definition of *proof*.

Definition 3.12 Let U be a set of formulas and A a formula. The notation $U \vdash A$ means that the formulas in U are *assumptions* in the proof of A. A *proof* is a sequence of lines $U_i \vdash \phi_i$, such that for each i, $U_i \subseteq U$, and ϕ_i is an axiom, a previously proved theorem, a member of U_i or can be derived by *MP* from previous lines $U_{i'} \vdash \phi_{i'}$, $U_{i''} \vdash \phi_{i''}$, where $i', i'' < i$. ■

Rule 3.13 (Deduction rule)

$$\frac{U \cup \{A\} \vdash B}{U \vdash A \to B}.$$

We must show that this derived rule is *sound*, that is, that the use of the derived rule does not increase the set of provable theorems in $\mathcal{H}$. This is done by showing how to transform any proof using the rule into one that does not use the rule. Therefore, in principle, any proof that uses the derived rule could be transformed to one that uses only the three axioms and *MP*.

Theorem 3.14 (Deduction theorem) *The deduction rule is a sound derived rule.*

Proof We show by induction on the length n of the proof of $U \cup \{A\} \vdash B$ how to obtain a proof of $U \vdash A \to B$ that does not use the deduction rule.

For $n=1$, B is proved in one step, so B must be either an element of $U \cup \{A\}$ or an axiom of $\mathcal{H}$ or a previously proved theorem:

- If B is A, then $\vdash A \to A$ by Theorem 3.10, so certainly $U \vdash A \to A$.
- Otherwise (B is an axiom or a previously proved theorem), here is a proof of $U \vdash A \to B$ that does not use the deduction rule or the assumption A:

1.	$U \vdash B$	Axiom or theorem
2.	$U \vdash B \to (A \to B)$	Axiom 1
3.	$U \vdash A \to B$	MP 1, 2

If $n>1$, the last step in the proof of $U \cup \{A\} \vdash B$ is either a one-step inference of B or an inference of B using *MP*. In the first case, the result holds by the proof for $n=1$. Otherwise, *MP* was used, so there is a formula C and lines $i, j < n$ in the proof such that line i in the proof is $U \cup \{A\} \vdash C$ and line j is $U \cup \{A\} \vdash C \to B$. By the inductive hypothesis, $U \vdash A \to C$ and $U \vdash A \to (C \to B)$. A proof of $U \vdash A \to B$ is given by:

1.	$U \vdash A \rightarrow C$	Inductive hypothesis
2.	$U \vdash A \rightarrow (C \rightarrow B)$	Inductive hypothesis
3.	$U \vdash (A \rightarrow (C \rightarrow B)) \rightarrow ((A \rightarrow C) \rightarrow (A \rightarrow B))$	Axiom 2
4.	$U \vdash (A \rightarrow C) \rightarrow (A \rightarrow B)$	MP 2, 3
5.	$U \vdash A \rightarrow B$	MP 1, 4

■

3.4 Derived Rules in $\mathscr{H}$

The general form of a derived rule will be one of:

$$\frac{U \vdash \phi_1}{U \vdash \phi}, \qquad \frac{U \vdash \phi_1 \qquad U \vdash \phi_2}{U \vdash \phi}.$$

The first form is justified by proving the formula $U \vdash \phi_1 \rightarrow \phi$ and the second by $U \vdash \phi_1 \rightarrow (\phi_2 \rightarrow \phi)$; the formula $U \vdash \phi$ that is the conclusion of the rule follows immediately by one or two applications of *MP*. For example, from Axiom 3 we immediately have the following rule:

Rule 3.15 (Contrapositive rule)

$$\frac{U \vdash \neg B \rightarrow \neg A}{U \vdash A \rightarrow B}.$$

The contrapositive is used extensively in mathematics. We showed the completeness of the method of semantic tableaux by proving: *If a tableau is open, the formula is satisfiable*, which is the contrapositive of the theorem that we wanted to prove: *If a formula is unsatisfiable (not satisfiable), the tableau is closed (not open).*

Theorem 3.16 $\vdash (A \rightarrow B) \rightarrow [(B \rightarrow C) \rightarrow (A \rightarrow C)]$.

Proof

1.	$\{A \rightarrow B, B \rightarrow C, A\} \vdash A$	Assumption
2.	$\{A \rightarrow B, B \rightarrow C, A\} \vdash A \rightarrow B$	Assumption
3.	$\{A \rightarrow B, B \rightarrow C, A\} \vdash B$	MP 1, 2
4.	$\{A \rightarrow B, B \rightarrow C, A\} \vdash B \rightarrow C$	Assumption
5.	$\{A \rightarrow B, B \rightarrow C, A\} \vdash C$	MP 3, 4
6.	$\{A \rightarrow B, B \rightarrow C\} \vdash A \rightarrow C$	Deduction 5
7.	$\{A \rightarrow B\} \vdash [(B \rightarrow C) \rightarrow (A \rightarrow C)]$	Deduction 6
8.	$\vdash (A \rightarrow B) \rightarrow [(B \rightarrow C) \rightarrow (A \rightarrow C)]$	Deduction 7

■

Rule 3.17 (Transitivity rule)

$$\frac{U \vdash A \rightarrow B \qquad\qquad U \vdash B \rightarrow C}{U \vdash A \rightarrow C}.$$

The transitivity rule justifies the step-by-step development of a mathematical theorem $\vdash A \rightarrow C$ through a series of *lemmas*. The antecedent A of the theorem is used to prove a lemma $\vdash A \rightarrow B_1$ whose consequent is used to prove the next lemma $\vdash B_1 \rightarrow B_2$ and so on until the consequent of the theorem appears as $\vdash B_n \rightarrow C$. Repeated use of the transitivity rule enables us to deduce $\vdash A \rightarrow C$.

Theorem 3.18 $\vdash [A \rightarrow (B \rightarrow C)] \rightarrow [B \rightarrow (A \rightarrow C)]$.

Proof

1.	$\{A \rightarrow (B \rightarrow C), B, A\} \vdash A$	Assumption
2.	$\{A \rightarrow (B \rightarrow C), B, A\} \vdash A \rightarrow (B \rightarrow C)$	Assumption
3.	$\{A \rightarrow (B \rightarrow C), B, A\} \vdash B \rightarrow C$	MP 1, 2
4.	$\{A \rightarrow (B \rightarrow C), B, A\} \vdash B$	Assumption
5.	$\{A \rightarrow (B \rightarrow C), B, A\} \vdash C$	MP 3, 4
6.	$\{A \rightarrow (B \rightarrow C), B\} \vdash A \rightarrow C$	Deduction 5
7.	$\{A \rightarrow (B \rightarrow C)\} \vdash B \rightarrow (A \rightarrow C)$	Deduction 6
8.	$\vdash [A \rightarrow (B \rightarrow C)] \rightarrow [B \rightarrow (A \rightarrow C)]$	Deduction 7

■

Rule 3.19 (Exchange of antecedent rule)

$$\frac{U \vdash A \rightarrow (B \rightarrow C)}{U \vdash B \rightarrow (A \rightarrow C)}.$$

Exchanging the antecedent simply means that it doesn't matter in which order we use the lemmas necessary in a proof.

Theorem 3.20 $\vdash \neg A \rightarrow (A \rightarrow B)$.

Proof

1.	$\{\neg A\} \vdash \neg A \rightarrow (\neg B \rightarrow \neg A)$	Axiom 1
2.	$\{\neg A\} \vdash \neg A$	Assumption
3.	$\{\neg A\} \vdash \neg B \rightarrow \neg A$	MP 1, 2
4.	$\{\neg A\} \vdash (\neg B \rightarrow \neg A) \rightarrow (A \rightarrow B)$	Axiom 3
5.	$\{\neg A\} \vdash A \rightarrow B$	MP 3, 4
6.	$\vdash \neg A \rightarrow (A \rightarrow B)$	Deduction 5

■

Theorem 3.21 $\vdash A \rightarrow (\neg A \rightarrow B)$.

Proof

1.	$\vdash \neg A \rightarrow (A \rightarrow B)$	Theorem 3.20
2.	$\vdash A \rightarrow (\neg A \rightarrow B)$	Exchange 1

■

These two theorems are of major theoretical importance. They say that if you can prove some formula A and its negation $\neg A$, then you can prove *any* formula B! If you can prove any formula then there are no unprovable formulas so the concept of proof becomes meaningless.

Theorem 3.22 $\vdash \neg\neg A \rightarrow A$.

Proof

1.	$\{\neg\neg A\} \vdash \neg\neg\neg A \rightarrow (\neg\neg\neg\neg A \rightarrow \neg\neg A)$	Axiom 1
2.	$\{\neg\neg A\} \vdash \neg\neg\neg A$	Assumption
3.	$\{\neg\neg A\} \vdash \neg\neg\neg\neg A \rightarrow \neg\neg A$	MP 1, 2
4.	$\{\neg\neg A\} \vdash \neg A \rightarrow \neg\neg\neg A$	Contrapositive 3
5.	$\{\neg\neg A\} \vdash \neg\neg A \rightarrow A$	Contrapositive 4
6.	$\{\neg\neg A\} \vdash A$	MP 2, 5
7.	$\vdash \neg\neg A \rightarrow A$	Deduction 6

■

Theorem 3.23 $\vdash A \rightarrow \neg\neg A$.

Proof

1.	$\vdash \neg\neg\neg A \rightarrow \neg A$	Theorem 3.22
2.	$\vdash A \rightarrow \neg\neg A$	Contrapositive 1

■

Rule 3.24 (Double negation rule)

$$\frac{U \vdash \neg\neg A}{U \vdash A}, \qquad \frac{U \vdash A}{U \vdash \neg\neg A}.$$

Double negation is a very intuitive rule. We expect that 'it is raining' and 'it is not true that it is not raining' will have the same truth value, and that the second formula can be simplified to the first. Nevertheless, some logicians reject the rule because it is not constructive. Suppose that we can prove for some number n, 'it is not true that n is prime' which is the same as 'it is not true that n is not composite'. This double negation could be reduced by the rule to 'n is composite', but we have not actually demonstrated any factors of n.

Theorem 3.25 $\vdash (A \rightarrow B) \rightarrow (\neg B \rightarrow \neg A)$.

Proof

1.	$\{A \rightarrow B\} \vdash A \rightarrow B$	Assumption
2.	$\{A \rightarrow B\} \vdash \neg\neg A \rightarrow A$	Theorem 3.22
3.	$\{A \rightarrow B\} \vdash \neg\neg A \rightarrow B$	Transitivity 2, 1
4.	$\{A \rightarrow B\} \vdash B \rightarrow \neg\neg B$	Theorem 3.23
5.	$\{A \rightarrow B\} \vdash \neg\neg A \rightarrow \neg\neg B$	Transitivity 3, 4
6.	$\{A \rightarrow B\} \vdash \neg B \rightarrow \neg A$	Contrapositive 5
7.	$\vdash (A \rightarrow B) \rightarrow (\neg B \rightarrow \neg A)$	Deduction 6

■

Rule 3.26 (Contrapositive rule)

$$\frac{U \vdash A \rightarrow B}{U \vdash \neg B \rightarrow \neg A}.$$

This is the other direction of the contrapositive rule shown earlier.

Recall from Sect. 2.3.3 the definition of the logical constants *true* as an abbreviation for $p \vee \neg p$ and *false* as an abbreviation for $p \wedge \neg p$. These can be expressed using implication and negation alone as $p \rightarrow p$ and $\neg (p \rightarrow p)$.

Theorem 3.27

$$\vdash \mathit{true},$$
$$\vdash \neg \mathit{false}.$$

Proof $\vdash \mathit{true}$ is an instance of Theorem 3.10. $\vdash \neg \mathit{false}$, which is $\vdash \neg \neg (p \rightarrow p)$, follows by double negation. ∎

Theorem 3.28 $\vdash (\neg A \rightarrow \mathit{false}) \rightarrow A$.

Proof

1.	$\{\neg A \rightarrow \mathit{false}\} \vdash \neg A \rightarrow \mathit{false}$	Assumption
2.	$\{\neg A \rightarrow \mathit{false}\} \vdash \neg \mathit{false} \rightarrow \neg \neg A$	Contrapositive
3.	$\{\neg A \rightarrow \mathit{false}\} \vdash \neg \mathit{false}$	Theorem 3.27
4.	$\{\neg A \rightarrow \mathit{false}\} \vdash \neg \neg A$	MP 2, 3
5.	$\{\neg A \rightarrow \mathit{false}\} \vdash A$	Double negation 4
6.	$\vdash (\neg A \rightarrow \mathit{false}) \rightarrow A$	Deduction 5

∎

Rule 3.29 (Reductio ad absurdum)

$$\frac{U \vdash \neg A \rightarrow \mathit{false}}{U \vdash A}.$$

Reductio ad absurdum is a very useful rule in mathematics: Assume the negation of what you wish to prove and show that it leads to a contradiction. This rule is also controversial because proving that $\neg A$ leads to a contradiction provides no reason that directly justifies A.

Here is an example of the use of this rule:

Theorem 3.30 $\vdash (A \rightarrow \neg A) \rightarrow \neg A$.

Proof

1.	$\{A \rightarrow \neg A, \neg\neg A\} \vdash \neg\neg A$	Assumption
2.	$\{A \rightarrow \neg A, \neg\neg A\} \vdash A$	Double negation 1
3.	$\{A \rightarrow \neg A, \neg\neg A\} \vdash A \rightarrow \neg A$	Assumption
4.	$\{A \rightarrow \neg A, \neg\neg A\} \vdash \neg A$	MP 2, 3
5.	$\{A \rightarrow \neg A, \neg\neg A\} \vdash A \rightarrow (\neg A \rightarrow \mathit{false})$	Theorem 3.21
6.	$\{A \rightarrow \neg A, \neg\neg A\} \vdash \neg A \rightarrow \mathit{false}$	MP 2, 5
7.	$\{A \rightarrow \neg A, \neg\neg A\} \vdash \mathit{false}$	MP 4, 6
8.	$\{A \rightarrow \neg A\} \vdash \neg\neg A \rightarrow \mathit{false}$	Deduction 7
9.	$\{A \rightarrow \neg A\} \vdash \neg A$	Reductio ad absurdum 8
10.	$\vdash (A \rightarrow \neg A) \rightarrow \neg A$	Deduction 9

∎

We leave the proof of the following theorem as an exercise.

Theorem 3.31 $\vdash (\neg A \rightarrow A) \rightarrow A$.

These two theorems may seem strange, but they can be understood on the semantic level. For the implication of Theorem 3.31 to be false, the antecedent $\neg A \rightarrow A$ must be true and the consequent A false. But if A is false, then so is $\neg A \rightarrow A \equiv A \vee A$, so the formula is true.

3.5 Theorems for Other Operators

So far we have worked with only negation and implication as operators. These two operators are adequate for defining all others (Sect. 2.4), so we can use these definitions to prove theorems using other operators. Recall that $A \wedge B$ is defined as $\neg(A \rightarrow \neg B)$, and $A \vee B$ is defined as $\neg A \rightarrow B$.