Part II
Gödel's Completeness Theorem

In this part of the book we shall prove Gödel's COMPLETENESS THEOREM and show several consequences.

Gödel proved his famous theorem in his doctoral dissertation *Über die Vollständigkeit des Logikkalküls* [10] which was completed in 1929. In 1930, he published the same material as in the doctoral dissertation in a rewritten and shortened form in [11]. However, instead of presenting Gödel's original proof we decided to follow Henkin's construction, which can be found in [17] (see also [19]), since it fits better in the logical framework developed in Part I. Even though Henkin's construction works also for uncountable signatures, we shall prove in Chapter **??** the general COMPLETENESS THEOREM with an ultraproduct construction, using ŁOŠ'S THEOREM.

Etwas sagen zur Konstruktion: Warum braucht man maximal konsistente Erweiterungen und warum brauchen wir Zeugen?

We would like to mention that in our proof of the COMPLETENESS THEOREM for countable signatures (carried out in Chapters 4 & **??**), one only has to assume the existence of *potentially infinite* sets but no instance of an *actually infinite* set is required (see also Chapter 0).

Chapter 4
Maximally Consistent Extensions

Throughout this chapter we require that all formulae are written in Polish notation and that the variables are among $v_0, v_1, v_2, \ldots$ Notice that the former requirement is just another notation which does not involve brackets and that by the VARIABLE SUBSTITUTION THEOREM 3.9, the latter requirement gives us semantically equivalent formulae.

Maximally Consistent Theories

Let $\mathscr{L}$ be an arbitrary signature and let T be an $\mathscr{L}$-theory (*i.e.*, a set of $\mathscr{L}$-sentences). We say that T is **maximally consistent** if T is consistent and for every $\mathscr{L}$-sentence σ we have *either* $\sigma \in \mathsf{T}$ *or* $\neg\operatorname{Con}(\mathsf{T} + \sigma)$. In other words, a consistent theory T is maximally consistent if no proper extension of T is consistent.

The following fact is just a reformulation of the definition.

FACT 4.1. *Let $\mathscr{L}$ be a signature and let T be a consistent $\mathscr{L}$-theory. Then T is maximally consistent iff for every $\mathscr{L}$-sentence σ, either $\sigma \in \mathsf{T}$ or $\mathsf{T} \vdash \neg\sigma$.*

Proof. By FACT 2.12.(c) & (d) we have:

$$\neg\operatorname{Con}(\mathsf{T} + \sigma) \iff \mathsf{T} \vdash \neg\sigma$$

Hence, an $\mathscr{L}$-theory is maximally consistent *iff* for every $\mathscr{L}$-sentence σ, either $\sigma \in \mathsf{T}$ or $\mathsf{T} \vdash \neg\sigma$. ⊣

As a consequence of FACT 4.1 we get

LEMMA 4.2. *Let $\mathscr{L}$ be a signature and let T be a consistent $\mathscr{L}$-theory. Then T is maximally consistent iff for every $\mathscr{L}$-sentence σ, either $\sigma \in \mathsf{T}$ or $\neg\sigma \in \mathsf{T}$.*

Proof. We have to show that the following equivalence holds:

$$\forall\sigma\big(\sigma \in \mathsf{T} \ \textit{ or } \ \mathsf{T} \vdash \neg\sigma\big) \quad \Longleftrightarrow \quad \forall\sigma\big(\sigma \in \mathsf{T} \ \textit{ or } \ \neg\sigma \in \mathsf{T}\big)$$

($\Rightarrow$) Assume that for every $\mathscr{L}$-sentence σ we have $\sigma \in \mathsf{T}$ or $\mathsf{T} \vdash \neg\sigma$. If $\sigma \in \mathsf{T}$, then the implication obviously holds. If $\sigma \notin \mathsf{T}$, then $\mathsf{T} \vdash \neg\sigma$, and since T is consistent, this implies $\mathsf{T} \nvdash \sigma$. Now, by TAUTOLOGY (F), this implies $\mathsf{T} \nvdash \neg\neg\sigma$ and by our assumption we finally get $\neg\sigma \in \mathsf{T}$.

($\Leftarrow$) Assume that for every $\mathscr{L}$-sentence σ we have $\sigma \in \mathsf{T}$ or $\neg\sigma \in \mathsf{T}$. If $\sigma \in \mathsf{T}$, then the implication obviously holds. Now, if $\sigma \notin \mathsf{T}$, then by our assumption we have $\neg\sigma \in \mathsf{T}$, which obviously implies $\mathsf{T} \vdash \neg\sigma$. ⊣

Maximally consistent theories have similar features as complete theories: Recall that an $\mathscr{L}$-theory T is complete if for every $\mathscr{L}$-sentence σ we have *either* $\mathsf{T} \vdash \sigma$ *or* $\mathsf{T} \vdash \neg\sigma$.

As an immediate consequence of the definitions we get

FACT 4.3. *Let $\mathscr{L}$ be a signature, let T be a consistent $\mathscr{L}$-theory, and let $\mathbf{Th}(\mathsf{T})$ be the set of all $\mathscr{L}$-sentences which are provable from T.*

(a) *If T is complete, then $\mathbf{Th}(\mathsf{T})$ is maximally consistent.*
(b) *If T is maximally consistent, then $\mathbf{Th}(\mathsf{T})$ is the same as T.*

The next result gives a condition under which a theory can be extended to maximally consistent theory, and is in fact just a reformulation of PROPOSITION 3.10.

FACT 4.4. *If an $\mathscr{L}$-theory T has a model, then T has a maximally consistent extension.*

Proof. Let $\mathbf{M}$ be a model of the $\mathscr{L}$-theory T and let $\mathsf{T}_{\mathbf{M}}$ be the set of $\mathscr{L}$-sentences σ such that $\mathbf{M} \vDash \sigma$. Then $\mathsf{T}_{\mathbf{M}}$ is obviously a maximally consistent theory which contains T. ⊣

Later we shall see that every consistent theory has a model. For this, we first show how a consistent theory can be extended to a maximally consistent theory.

Universal List of Sentences

Let $\mathscr{L}$ be an arbitrary but fixed countable signature, where by "countable" we mean that the symbols in $\mathscr{L}$ can be listed in a FINITE or POTENTIALLY INFINITE list $L_{\mathscr{L}}$.

First, we encode the symbols of $\mathscr{L}$ corresponding to the order in which they appear in the list $L_{\mathscr{L}}$: The first symbol is encoded with "2", the second with "22", the third with "222", and so on. For every symbol $\zeta \in L_{\mathscr{L}}$ let $\#\zeta$ denote the code of ζ. So, the code of a symbol of $\mathscr{L}$ is just a sequence of 2's.

Furthermore, we encode the logical symbols as follows:

Symbol ζ	Code $\#\zeta$
$=$	11
$\neg$	1111
$\wedge$	111111
$\vee$	11111111
$\rightarrow$	1111111111
$\exists$	111111111111
$\forall$	11111111111111
v_0	1
v_1	111
$\vdots$	$\vdots$
v_n	$\underbrace{1111 \ \ldots \ 11111}_{(2n+1)\ 1\text{'s}}$

In the next step, we encode strings of symbols: Let $\bar{\zeta} \equiv \zeta_0\zeta_1\zeta_2 \ldots \zeta_n$ be a finite string of symbols, then

$$\#\bar{\zeta} := \#\zeta_0 \mathsf{0} \#\zeta_1 \mathsf{0} \#\zeta_2 \ldots \mathsf{0} \#\zeta_n$$

For a string $\#\zeta$ (*i.e.*, a string of 0's, 1's, and 2's) let $|\#\zeta|$ be the length of $\#\zeta$ (*i.e.*, the number of 0's, 1's, and 2's which appear in $\#\zeta$).

Now, we order the codes of strings of symbols by their length and strings of the same length lexicographically, where $\mathsf{0} < \mathsf{1} < \mathsf{2}$. If, with respect to this ordering, $\#\zeta$ is less than $\#\zeta'$, we write $\zeta \prec \zeta'$.

Finally, let

$$\Lambda_{\mathscr{L}} := [\sigma_1, \sigma_2, \ldots]$$

be the potentially infinite list of all $\mathscr{L}$-sentences, where we require

$$\#\sigma_i \prec \#\sigma_j \iff i < j \,.$$

We call $\Lambda_{\mathscr{L}}$ the **universal list of $\mathscr{L}$-sentences**.

Lindenbaum's Lemma

In this section we show that every consistent set of $\mathscr{L}$-sentences T can be extended to a maximally consistent set of $\mathscr{L}$-sentences $\overline{\mathsf{T}}$. Since the universal list of $\mathscr{L}$-sentences contains all possible $\mathscr{L}$-sentences, every set T of $\mathscr{L}$-sentences can be listed in a finite or potentially infinite list.

LINDENBAUM'S LEMMA 4.5. *Let $\mathscr{L}$ be a countable signature and let* T *be a consistent set of $\mathscr{L}$-sentences. Furthermore, let σ_0 be an $\mathscr{L}$-sentence which cannot be proved from* T, *i.e.*, $\mathsf{T} \nvdash \sigma_0$. *Then there exists a maximally consistent set* $\overline{\mathsf{T}}$ *of $\mathscr{L}$-sentences which contains $\neg\sigma_0$ as well as all the sentences of* T.

Proof. Let $\Lambda_{\mathscr{L}} = [\sigma_1, \sigma_2, \ldots]$ be the universal list of all $\mathscr{L}$-sentences. First we extend $\Lambda_{\mathscr{L}}$ with the $\mathscr{L}$-sentence $\neg\sigma_0$; let $\Lambda^0_{\mathscr{L}} = [\neg\sigma_0, \sigma_1, \sigma_2, \ldots]$.

Now, we go through the list $\Lambda^0_{\mathscr{L}}$ and define step by step a list $\overline{\mathsf{T}}$ of $\mathscr{L}$-sentences: For this, we define T_1 as the list which contains just $\neg\sigma_0$, *i.e.*, $T_1 := [\neg\sigma_0]$. If T_n is already defined, then

$$T_{n+1} := \begin{cases} T_n + [\sigma_n] & \text{if } \operatorname{Con}(T + T_n + \sigma_n), \\ T_n & \text{otherwise.} \end{cases}$$

Let $\overline{\mathsf{T}} = [\neg\sigma_0, \sigma_{i_1}, \ldots]$ be the resulting list, *i.e.*, $\overline{\mathsf{T}}$ is the union of all the T_n's. Notice that the construction only works if we assume the metamathematical LAW OF EXCLUDED MIDDLE or a similar principle like the WEAK KÖNIG'S LEMMA (see EXERCISE 4.1): Even in the case when we cannot decide whether $T + T_n + \sigma_n$ is consistent or not, we assume, from a metamathematical point of view, that *either* $T + T_n + \sigma_n$ is consistent *or* $T + T_n + \sigma_n$ is inconsistent (and *neither* both, *nor* none).

CLAIM. $\overline{\mathsf{T}}$ *is a maximally consistent set of $\mathscr{L}$-sentences which contains $\neg\sigma_0$ as well as all the sentences of* T.

Proof of Claim. First we show that $\neg\sigma_0$ belongs to $\overline{\mathsf{T}}$, then we show that $\mathsf{T} + \overline{\mathsf{T}}$ is consistent (which implies that $\overline{\mathsf{T}}$ is consistent), in a third step we show that $\overline{\mathsf{T}}$ contains T, and finally we show that for every $\mathscr{L}$-sentence σ we have either $\sigma \in \overline{\mathsf{T}}$ or $\neg\operatorname{Con}(\overline{\mathsf{T}} + \sigma)$.

$\neg\sigma_0$ *belongs to* $\overline{\mathsf{T}}$: By definition, $T_1 = [\neg\sigma_0]$, and since T_1 is an initial segment of the list $\overline{\mathsf{T}}$, $\neg\sigma_0$ belongs to $\overline{\mathsf{T}}$.

$\mathsf{T} + \overline{\mathsf{T}}$ *is consistent*: By the COMPACTNESS THEOREM 2.13 it is enough to show that every finite subset of $\mathsf{T} + \overline{\mathsf{T}}$ is consistent. So, let $\mathsf{T}' + T_k$ be a finite subset of $\mathsf{T} + \overline{\mathsf{T}}$, where T' is a finite subset of T and T_k is some finite initial segment of the list $\overline{\mathsf{T}}$. Since $\mathsf{T} \nvdash \sigma_0$, by FACT 2.12 (c) we have $\operatorname{Con}(\mathsf{T} + \neg\sigma_0)$, and since $T_1 = [\neg\sigma_0]$, we obtain $\operatorname{Con}(\mathsf{T} + T_1)$. Thus, if $T_k = [\neg\sigma_0]$, then $\mathsf{T}' + T_k$ is consistent. Otherwise, if $T_k = [\ldots, \sigma_n]$ for some $n \geq 1$, then $T_k = T_n + [\sigma_n]$, *i.e.*, $T_k = T_{n+1}$. So, by construction we have $\operatorname{Con}(\mathsf{T} + T_n + \sigma_n)$, which implies the consistency of $\mathsf{T}' + T_k$.

$\overline{\mathsf{T}}$ *contains all sentences of* T: We already know that $\neg\sigma_0$ belongs to $\overline{\mathsf{T}}$. Now, for every $\sigma \in \mathsf{T} - [\neg\sigma_0]$ there is a $\sigma_n \in \Lambda^0_{\mathscr{L}}$ such that $\sigma \equiv \sigma_n$. If $\operatorname{Con}(\mathsf{T} + T_n + \sigma_n)$, then $\sigma_n \in T_{n+1}$ and therefore $\sigma_n \in \overline{\mathsf{T}}$. Otherwise, if $\neg\operatorname{Con}(\mathsf{T} + T_n + \sigma_n)$, then, since $\sigma_n \in \mathsf{T}$, we have $\neg\operatorname{Con}(\mathsf{T} + T_n)$, and for $T_n = [\ldots, \sigma_m]$ (with $m \leq n$) we get $\neg\operatorname{Con}(\mathsf{T} + T_m + \sigma_m)$, which contradicts our construction.

For every σ, either $\sigma \in$ $\overline{\mathsf{T}}$ *or* $\neg\operatorname{Con}(\overline{\mathsf{T}} + \sigma)$: For every $\mathscr{L}$-sentence σ there is a $\sigma_n \in \Lambda^0_{\mathscr{L}}$ such that $\sigma \equiv \sigma_n$. By the law of excluded middle, we have *either*

$\mathrm{Con}(\mathsf{T}+T_n+\sigma_n)$ *or* $\neg\,\mathrm{Con}(\mathsf{T}+T_n+\sigma_n)$. In the former case we obtain $\sigma_n \in T_{n+1}$ which implies $\sigma \in \overline{\mathsf{T}}$, in the latter case we obtain $\neg\,\mathrm{Con}(\overline{\mathsf{T}}+\sigma_n)$, which is the same as $\neg\,\mathrm{Con}(\overline{\mathsf{T}}+\sigma)$. $\dashv_{\text{Claim}}$

Thus, the list $\overline{\mathsf{T}}$ has all the required properties, which completes the proof. $\dashv$

The following fact summarises the main properties of $\overline{\mathsf{T}}$.

FACT 4.6. *Let* $\mathsf{T}, \overline{\mathsf{T}}$, *and* σ_0 *be as above, and let* σ *and* σ' *be any* $\mathscr{L}$*-sentences.*

(a) $\neg\sigma_0 \in \overline{\mathsf{T}}$.

(b) *Either* $\sigma \in \overline{\mathsf{T}}$ *or* $\neg\sigma \in \overline{\mathsf{T}}$.

(c) *If* $\mathsf{T} \vdash \sigma$, *then* $\sigma \in \overline{\mathsf{T}}$.

(d) $\overline{\mathsf{T}} \vdash \sigma$ *iff* $\sigma \in \overline{\mathsf{T}}$.

(e) *If* $\sigma \Leftrightarrow \sigma'$, *then* $\sigma \in \overline{\mathsf{T}}$ *iff* $\sigma' \in \overline{\mathsf{T}}$.

Proof. (a) follows by construction of $\overline{\mathsf{T}}$.

Since $\overline{\mathsf{T}}$ is maximally consistent, (b) follows by LEMMA 4.2.

For (c), notice that $\mathsf{T} \vdash \sigma$ implies $\neg\,\mathrm{Con}(\mathsf{T}+\neg\sigma)$, hence $\neg\sigma \notin \overline{\mathsf{T}}$ and by (b) we get $\sigma \in \overline{\mathsf{T}}$.

For (d), let us first assume $\overline{\mathsf{T}} \vdash \sigma$, where $\sigma \equiv \sigma_n$. This implies $\mathrm{Con}(\overline{\mathsf{T}}+\sigma)$, hence $\mathrm{Con}(\mathsf{T}+T_n+\sigma_n)$, and by construction of $\overline{\mathsf{T}}$ we get $\sigma_n \in \overline{\mathsf{T}}$. On the other hand, if $\sigma \in \overline{\mathsf{T}}$, then we obviously have $\overline{\mathsf{T}} \vdash \sigma$.

For (e), recall that $\sigma \Leftrightarrow \sigma'$ is just an abbreviation for $\vdash \sigma \leftrightarrow \sigma'$. Thus, (e) follows immediately from (d). $\dashv$

FACT 4.6 shows that the $\mathscr{L}$-sentences in $\overline{\mathsf{T}}$ "behave" like valid sentences in a model, which is indeed the case—as the following proposition shows.

PROPOSITION 4.7. *Let* $\overline{\mathsf{T}}$ *be as above, and let* $\sigma, \sigma_1, \sigma_2$ *be any* $\mathscr{L}$*-sentences in Polish notation.*

$$\begin{array}{lrcl}
\text{(a)} & \neg\sigma \in \overline{\mathsf{T}} & \Longleftrightarrow & \text{NOT } \sigma \in \overline{\mathsf{T}} \\
\text{(b)} & \wedge\sigma_1\sigma_2 \in \overline{\mathsf{T}} & \Longleftrightarrow & \sigma_1 \in \overline{\mathsf{T}} \text{ AND } \sigma_2 \in \overline{\mathsf{T}} \\
\text{(c)} & \vee\sigma_1\sigma_2 \in \overline{\mathsf{T}} & \Longleftrightarrow & \sigma_1 \in \overline{\mathsf{T}} \text{ OR } \sigma_2 \in \overline{\mathsf{T}} \\
\text{(d)} & \to\sigma_1\sigma_2 \in \overline{\mathsf{T}} & \Longleftrightarrow & \text{IF } \sigma_1 \in \overline{\mathsf{T}} \text{ THEN } \sigma_2 \in \overline{\mathsf{T}}
\end{array}$$

Proof. (a) Follows immediately from FACT 4.6.(b).

(b) First notice that by FACT 4.6.(d), $\wedge\sigma_1\sigma_2 \in \overline{\mathsf{T}}$ *iff* $\overline{\mathsf{T}} \vdash \wedge\sigma_1\sigma_2$. Thus, by L_3 and L_4 and (MP) we get $\overline{\mathsf{T}} \vdash \sigma_1$ and $\overline{\mathsf{T}} \vdash \sigma_2$. Thus, by FACT 4.6.(d), we get $\sigma_1 \in \overline{\mathsf{T}}$ AND $\sigma_2 \in \overline{\mathsf{T}}$. On the other hand, if $\sigma_1 \in \overline{\mathsf{T}}$ AND $\sigma_2 \in \overline{\mathsf{T}}$, then, by FACT 4.6.(d), we get $\overline{\mathsf{T}} \vdash \sigma_1$ and $\overline{\mathsf{T}} \vdash \sigma_2$. Now, by TAUTOLOGY (B), this implies $\overline{\mathsf{T}} \vdash \wedge\sigma_1\sigma_2$, and by by FACT 4.6.(d) we finally get $\wedge\sigma_1\sigma_2 \in \overline{\mathsf{T}}$.

(c) and (d) follow from FACT 4.6.(e) and from the 3-SYMBOLS THEOREM 1.2 which states that for each formula σ there is an equivalent formula σ' which contains neither "$\vee$" nor "$\to$". $\dashv$

EXERCISES

4.0 Show that all the logical axioms of propositional logic (*i.e.*, L_0–L_9) were used in the proofs of FACT 4.1, LEMMA 4.2, FACT 4.6, and PROPOSITION 4.7. Notice that in the proof of FACT 4.1, we used FACT 2.12.(c) & (d).

4.1 The WEAK KÖNIG'S LEMMA is a very weak choice principle. It states that

every infinite 0-1-*tree contains an infinite branch.*

In other words, for every potentially infinite set S of finite 0-1-sequences with the property that for every sequences in S, S contains all its initial sub-sequences, then there exists a potentially infinite sequence $\bar{s}$, such that all its finite initial sub-sequences belong to S.

Show that in the proof of LINDENBAUM'S LEMMA 4.5, the LAW OF EXCLUDED MIDDLE can be replaces with the metamathematical WEAK KÖNIG'S LEMMA.

Hint: Firstly, consider the set Λ of finite lists $\lambda = [\sigma_{i_0}, \dots, \sigma_{i_n}]$ of $\mathscr{L}$-sentences, where $k < l$ implies $i_k < i_l$. Secondly, construct the tree T consisting of the lists $\lambda \in \Lambda$, such that there is no formal inconsistency proof of $\mathsf{T} + \lambda$ of length $|\lambda|$, where $|\lambda|$ denotes the length of λ. Then T corresponds to an infinite binary tree, where each infinite branch through T corresponds to a maximally consistent set of $\mathscr{L}$-sentences.