Synthese Library 470
Studies in Epistemology, Logic, Methodology, and Philosophy of Science

Carlo Cellucci

The Theory of Gödel

Springer

Synthese Library

Studies in Epistemology, Logic, Methodology,
and Philosophy of Science

Volume 470

The aim of *Synthese Library* is to provide a forum for the best current work in the methodology and philosophy of science and in epistemology, all broadly understood. A wide variety of different approaches have traditionally been represented in the Library, and every effort is made to maintain this variety, not for its own sake, but because we believe that there are many fruitful and illuminating approaches to the philosophy of science and related disciplines.

Special attention is paid to methodological studies which illustrate the interplay of empirical and philosophical viewpoints and to contributions to the formal (logical, set-theoretical, mathematical, information-theoretical, decision-theoretical, etc.) methodology of empirical sciences. Likewise, the applications of logical methods to epistemology as well as philosophically and methodologically relevant studies in logic are strongly encouraged. The emphasis on logic will be tempered by interest in the psychological, historical, and sociological aspects of science. In addition to monographs *Synthese Library* publishes thematically unified anthologies and edited volumes with a well-defined topical focus inside the aim and scope of the book series. The contributions in the volumes are expected to be focused and structurally organized in accordance with the central theme(s), and should be tied together by an extensive editorial introduction or set of introductions if the volume is divided into parts. An extensive bibliography and index are mandatory.

Carlo Cellucci

The Theory of Gödel

Springer

Carlo Cellucci
Department of Philosophy
Sapienza University of Rome
Rome, Italy

ISSN 0166-6991 ISSN 2542-8292 (electronic)
Synthese Library
ISBN 978-3-031-13416-6 ISBN 978-3-031-13417-3 (eBook)
https://doi.org/10.1007/978-3-031-13417-3

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This book is meant to present most of the results of mathematical logic that are relevant to the philosophy of mathematics.

Since the hard core of such results consists of Gödel's incompleteness theorems, this explains the title of the book, *The Theory of Gödel*, which is reminiscent of the subtitle of the very first book on the subject that Mostowski described as an attempt to present "the theory of Gödel," namely "the famous theory of undecidable sentences created by Kurt Gödel in 1931" (Mostowski 1952, v).

The presentation of Gödel's incompleteness theorems and other limitative results in this book is in the spirit, though not in the letter of Jeroslow (1973). Other presentations can be found in Boolos (1993), Epstein and Carnielli (2008), Felscher (2000), Fitting (2007), Girard (1987), Goldstern and Judah (1998), Grandy (1977), Halbeisen and Krapf (2020), Isaacson (2018), Kennedy (2022), Lindström (1997), Murawski (1999), Robbin (2006), Smith (2013), Smullyan (1992), Świerczkowski (2003), Tourlakis (2003), and Zach (2021).

Results are presented in the form most relevant for use in the philosophy of mathematics. Their implications for Hilbert's approach to the philosophy of mathematics are discussed in the Appendix. As to their implications for the philosophy of mathematics in general, the interested reader may refer to Cellucci (2022).

The book is self-contained, all notions being explained in full detail, but of course previous exposure to the very first rudiments of mathematical logic will help.

I am very grateful to two anonymous referees for useful remarks and suggestions. I also warmly thank Elena Griniari from Springer for her invaluable help in the editorial process.

Roma, Italy

Carlo Cellucci

References

Boolos, G. (1993). *The logic of provability*. Cambridge: Cambridge University Press.

Cellucci, C. (2022). *The making of mathematics: Heuristic philosophy of mathematics*. Cham: Springer.

Epstein, R., & Carnielli, W. (2008). *Computability: Computable functions, logic, and the foundations of mathematics*. Socorro: Advanced Reasoning Forum.

Felscher, W. (2000). *The logic of arithmetic*: *Lectures on mathematical logic* (vol. 3). New York: Gordon and Breach.

Fitting, M. (2007). *Incompleteness in the land of sets*. London: College Publications.

Girard, J.-Y. (1987). *Proof theory and logical complexity*. Naples: Bibliopolis.

Goldstern, M., & Judah, H. (1998). *The incompleteness phenomenon: A new course in mathematical logic*. Wellesley: A K Peters.

Grandy, R. (1977). *Advanced logic for applications*. Dordrecht: Reidel.

Halbeisen, L., & Krapf, R. (2020). *Gödel's theorems and Zermelo's axioms*. Cham: Springer.

Isaacson, D. (2018). *Gödel's incompleteness theorems*. Oxford: Oxford University (unpublished lecture notes). https://courses-archive.maths.ox.ac.uk/node/view_material/38136.

Jeroslow, R. (1973). Redundancies in the Hilbert-Bernays derivability conditions for Gödel's second incompleteness theorem. *The Journal of Symbolic Logic, 38*, 359–367.

Kennedy, J. (2022). *Gödel's incompleteness theorems*. Cambridge: Cambridge University Press.

Lindström, P. (1997). *Aspects of incompleteness*. Berlin: Springer.

Mostowski, A. (1952). *Sentences undecidable in formalized arithmetic: An exposition of the theory of Kurt Gödel*. Amsterdam: North-Holland.

Murawski, R. (1999). *Recursive functions and metamathematics: Problems of completeness and decidability, Gödel's Theorems*. Dordrecht: Springer.

Robbin, J. (2006). *Mathematical logic: A first course*. Mineola: Dover.

Smith, P. (2013). *An introduction to Gödel's incompleteness theorems*. Cambridge: Cambridge University Press.

Smullyan, R. (1992). *Gödel's incompleteness theorems*. Oxford: Oxford University Press.

Świerczkowski, S. (2003). Finite sets and Gödel's incompleteness theorems. *Dissertationes Mathematicae, 422*, 1–58.

Tourlakis, G. (2003). *Mathematical logic*: *Lectures in logic and set theory* (vol. 1). Cambridge: Cambridge University Press.

Zach, R. (2021). *Incompleteness and computability: An open introduction to Gödel's theorems*. Scotts Valley: CreateSpace.

Contents

Chapter 1
First-Order Logic

1.1 First-Order Languages

Summary 1.1.1 In this section we introduce a basic kind of formal languages, first-order languages.

Definition 1.1.2 The *symbols* of a first-order language **L** are:

(i) infinitely many individual variables $v_0, v_1, v_2, ...$;
(ii) any number of individual constants;
(iii) for each positive integer n, any number of n-ary function constants;
(iv) for each positive integer n, any number of n-ary predicate constants;
(v) the equality symbol $=$;
(vi) connectives $\neg, \to$;
(vii) the universal quantifier $\forall$;
(vii) parentheses (,) and comma , .

Definition 1.1.3 Individual constants, function constants and predicate constants are called the *non-logical* symbols of **L**. The equality symbol, connectives and the universal quantifier are called the *logical* symbols of **L**.

Definition 1.1.4 The *terms* of a first-order language **L** are defined as follows:

(i) any individual variable is a term;
(ii) any individual constant is a term;
(iii) if f is an n-ary function constant and $t_1, \ldots, t_n$ are terms, then $f(t_1, \ldots, t_n)$ is a term.

Proposition 1.1.5 *There is a unique set X such that:*

(i) any individual variable is in X;

C. Cellucci, *The Theory of Gödel*, Synthese Library 470,
https://doi.org/10.1007/978-3-031-13417-3_1

(ii) *any individual constant is in* X*;*
(iii) *if* f *is an* n*-ary function constant and* $t_1, \ldots, t_n$ *are in* X*, then* $f(t_1, \ldots, t_n)$ *is in* X*;*
(iv) *if* Y *is any set satisfying (i)-(iii), then* $X \subseteq Y$.

This unique set X is the set of terms of **L**.

Proof Let Φ be the set of all sets Y satisfying (i)–(iii); namely $Y \in \Phi$ iff (1) any individual variable is in Y; (2) any individual constant is in Y; (3) if f is an n-ary function constant and $t_1, \ldots, t_n$ are in Y, then $f(t_1, \ldots, t_n)$ is in Y. Then let X be the intersection of all the elements of Φ; namely $t \in X$ iff $t \in Y$ for any $Y \in \Phi$. Clearly X satisfies (i)–(iv).

Proposition 1.1.6 (Induction Principle for Terms) *Let* P *be any property. If*

(i) $P(v_i)$ *for all individual variables* v_i*;*
(ii) $P(c)$*, for all individual constants* c*;*
(iii) *if* $P(t_1), \ldots, P(t_n)$*, then* $P(f(t_1, \ldots, t_n))$*, for all* n*-ary function constants* f*;*

then $P(t)$*, for all terms of* **L**.

Proof Let X be the set of all terms t of **L** such that $P(t)$. Then X satisfies the conditions [1.1.5 (i)–(iii)] on the set of terms of **L**, so any term of **L** is an element of X, namely $P(t)$ holds for all terms of **L**.

Definition 1.1.7 We call an application of the Induction Principle for Terms [1.1.6] a *proof by induction on* t.

Definition 1.1.8 The *atomic formulas* of a first-order language **L** are all expressions of the form $P(t_1, \ldots, t_n)$ where P is an n-ary predicate constant and $t_1, \ldots, t_n$ are terms, and all expressions of the form $(t = s)$ where t, s are terms.

Definition 1.1.9 The *formulas* of a first-order language **L** are defined as follows:

(i) A is a formula, for all atomic formulas A;
(ii) if A is a formula, then $\neg A$ is a formula;
(iii) if A and B are formulas, then $(A \to B)$ is a formula;
(iv) if A is a formula and v_i is an individual variable, then $\forall v_i A$ is a formula ($i = 0, 1, 2, \ldots$).

Proposition 1.1.10 *There is a unique set* X *such that:*

(i) $A \in X$*, for all atomic formulas* A*;*
(ii) *if* $A \in X$*, then* $\neg A \in X$*;*
(iii) *if* $A \in X$ *and* $B \in X$*, then* $(A \to B) \in X$*;*
(iv) *if* $A \in X$*, then* $\forall v_i A \in X$ $(i = 0, 1, 2, \ldots)$*;*
(v) *if* Y *is any set satisfying (i)-(iv), then* $X \subseteq Y$.

Proof Similarly to the corresponding proposition for terms 1.1.5.

Proposition 1.1.11 (Induction Principle for Formulas) *Let P be any property. If*

(i) $P(A)$, *for all atomic formulas* A;
(ii) if $P(A)$, *then* $P(\neg A)$;
(iii) if $P(A)$ *and* $P(B)$, *then* $P(A \to B)$;
(iv) if $P(A)$, *then* $P(\forall v_i A)$ $(i = 0, 1, 2, ...)$;

then $P(A)$, *for all formulas* A *of* $\mathbf{L}$.

Proof Similarly to the corresponding proposition for terms 1.1.6.

Definition 1.1.12 We call an application of the Induction Principle for Formulas 1.1.11 a *proof by induction on* A.

Definition 1.1.13 A *first-order language* is a language in which symbols, terms and formulas are as described above. Thus a first-order language is completely determined by its non-logical symbols.

Definition 1.1.14 The *cardinality* of a first-order language is the cardinality of the set of its non-logical symbols. A first-order language is *finite* iff its cardinality is finite. Similarly for *denumerable*, *countable* or *uncountable*.

Assumption 1.1.15 In what follows all the basic first-order languages considered are supposed to be countable.

Definition 1.1.16 We often write:

$x, y, z, ...$ (possibly with subscripts) for individual variables,
$t, u, s, ...$ (possibly with subscripts) for terms,
$A, B, C, ...$ (possibly with subscripts) for formulas.

Moreover we write:

$(t \neq u)$ for $\neg(t = u)$,
$(A \wedge B)$ for $\neg(A \to \neg B)$,
$(A \vee B)$ for $(\neg A \to B)$,
$(A \leftrightarrow B)$ for $((A \to B) \wedge (B \to A))$,
$\bigwedge_{i=1}^{n} A_i$ for $((...(A_1 \wedge A_2) \wedge ...) \wedge A_n)$,
$\bigvee_{i=1}^{n} A_i$ for $((...(A_1 \vee A_2) \vee ...) \vee A_n)$,
$\exists x A$ for $\neg \forall x \neg A$,
$\forall x_1 ... x_n A$ for $\forall x_1 ... \forall x_n A$,
$\exists x_1 ... x_n A$ for $\exists x_1 ... \exists x_n A$.

Definition 1.1.17 We also call $\wedge$, $\vee$, $\leftrightarrow$ *connectives* and $\exists x$ *quantifier.*

Definition 1.1.18 We call $\neg A$ the *negation* of A, $(A \wedge B)$ the *conjunction* of A and B, $(A \vee B)$ the *disjunction* of A and B, $(A \rightarrow B)$ the *implication* from A to B, $(A \leftrightarrow B)$ the *equivalence* between A and B, $\forall x A$ the *universal quantification* of A, $\exists x A$ the *existential quantification* of A.

Definition 1.1.19 The *opposite* of a formula A is B if A is a negation $\neg B$, and $\neg A$ if A is not a negation.

Definition 1.1.20 We say that a formula A of a first-order language $\mathbf{L}$ is *quantifier-free* iff no quantifier occurs in A.

Notation 1.1.21 In writing terms and formulas we may omit parentheses if no ambiguity can result. Specifically, we assume that outermost parentheses can always be omitted. Thus we may write $A \wedge (B \rightarrow C)$ for $(A \wedge (B \rightarrow C))$. We assume that parentheses can always be omitted in $(t = s)$. Moreover we assume that $\wedge$ and $\vee$ bind more strongly than $\rightarrow$ and $\leftrightarrow$. Thus $A \wedge B \rightarrow A$ will stand for $((A \wedge B) \rightarrow A)$. Finally we assume that $\wedge$ and $\vee$ are associative to the left. Thus $A \wedge B \wedge C$ will stand for $((A \wedge B) \wedge C)$. This convention is not used with $\rightarrow$ and $\leftrightarrow$.

Example 1.1.22

(a) The First-Order Language of Groups $\mathbf{L_G}$ is the first-order language whose non-logical symbols are the individual constant $\dot{0}$ and the binary function constant $\dot{+}$. We write $x \dot{+} y$ for $\dot{+}(x, y)$. (For the dot notation, see Remark 2.1.6).
(b) The First-Order Language of Successor $\mathbf{L_S}$ is the first-order language whose only non-logical symbol is the binary relation constant $\dot{S}uc$.
(c) The First-Order Language of Real Numbers $\mathbf{L_R}$ is the first-order language whose non-logical symbols are the individual constants $\dot{0}$ and $\dot{1}$, the binary function constants $\dot{+}$ and $\dot{\cdot}$ and the binary relation constant $\dot{<}$. We write $x \dot{+} y$ for $\dot{+}(x, y)$, $x \dot{\cdot} y$ for $\dot{\cdot}(x, y)$ and $x \dot{<} y$ for $\dot{<}(x, y)$.

Definition 1.1.23 The *subterms* of a term t of a first-order language $\mathbf{L}$ are defined as follows:

(i) if t is an individual variable or individual constant, the only subterm of t is t;
(ii) if t is $f(t_1, \ldots, t_n)$, the subterms of t are $f(t_1, \ldots, t_n)$ and the subterms of at least one of $t_1, \ldots, t_n$.

Definition 1.1.24 The *subformulas* of a formula A of a first-order language $\mathbf{L}$ are defined as follows:

(i) if A is an atomic formula, the only subformula of A is A;
(ii) if A is $\neg B$, the subformulas of A are $\neg B$ and the subformulas of B;
(iii) if A is $B \rightarrow C$, the subformulas of A are $B \rightarrow C$ and the subformulas of B or C;
(iv) if A is $\forall x B$, the subformulas of A are $\forall x B$ and the subformulas of B.

1.2 Free Variables

Summary 1.2.1 In this section we introduce some notions concerning the occurrences of individual variables in formulas.

Definition 1.2.2 Let t be a term of a first-order language **L**. We say that all occurrences of an individual variable x in t are *free*. We say that an individual variable x *occurs free* in t iff x has some free occurrence in t.

Definition 1.2.3 Let t be a term of a first-order language **L**.We say that t is *closed* iff no individual variable occurs free in t. Thus t is closed iff no individual variable occurs in t.

Definition 1.2.4 Let A be a formula of a first-order language **L**. The *free occurrences* of an individual variable x in A are defined as follows:

(i) if A is an atomic formula, all occurrences of x in A are free;
(ii) if A is $\neg B$, the free occurrences of x in A are the free occurrences of x in B;
(iii) if A is $B \rightarrow C$, the free occurrences of x in A are the free occurrences of x in B or C;
(iv) if A is $\forall y B$, the free occurrences of x in A are the free occurrences of x in B when x is different from y, otherwise x has no free occurrences in $\forall y B$, and in that case all occurrences of x in $\forall y B$ are said to be *bound occurrences*.

Definition 1.2.5 Let A be a formula of a first-order language **L**. We say that an individual variable x *occurs free in* A, or is a *free variable of* A, iff x has some free occurrence in A. We say that A is *closed*, or a *sentence*, iff A contains no free variable.

Definition 1.2.6 Let A be a formula of a first-order language **L**. We say that A is a *universal generalization* of a formula B iff A is $\forall x_1 ... \forall x_n B$ for some individual variables $x_1, \ldots, x_n$ $(n \geq 0)$. This includes the case $n = 0$, namely any formula is a universal generalization of itself. If $x_1, \ldots, x_n$ are all the free variables of B, then we say that A is the *universal closure* of B. (We assume the order of the variables $x_1, \ldots, x_n$ to be fixed in some way). So the universal closure of a formula is a sentence.

1.3 Substitution

Summary 1.3.1 In this section we introduce the notion of substituting an individual variable by a term in a term or a formula.

Definition 1.3.2 Let u and t be terms and x an individual variable of a first-order language **L**. We define $u[x/t]$, the result of *substituting* t *for any free occurrence of* x *in* u, as follows:

(i) if u is x, then $u[x/t]$ is t;
(ii) if u is an individual variable y different from x, then $u[x/t]$ is y;
(iii) if u is an individual constant c, then $u[x/t]$is c;
(iv) if u is $f(u_1, \ldots, u_n)$, then $u[x/t]$ is $f(u_1[x/t], \ldots, u_n[x/t])$.

Definition 1.3.3 Let A be a formula, x an individual variable and t a term of a first-order language **L**. We define $A[x/t]$, the result of *substituting* t *for any free occurrence of* x *in* A, as follows:

(i) if A is $P(u_1, \ldots, u_n)$, then $A[x/t]$ is $P(u_1[x/t], \ldots, u_n[x/t])$;
(ii) if A is $u_1 = u_2$, then $A[x/t]$ is $u_1[x/t] = u_2[x/t]$;
(iii) if A is $\neg B$, then $A[x/t]$ is $\neg B[x/t]$;
(iv) if A is $B \to C$, then $A[x/t]$ is $B[x/t] \to C[x/t]$;
(v) if A is $\forall x B$, then $A[x/t]$ is A;
(vi) if A is $\forall y B$, then $A[x/t]$ is $\forall y B[x/t]$ when y is different from x, otherwise $A[x/t]$ is $\forall y B$.

Notation 1.3.4 We extend the notation $A[x/t]$ to a set of formulas Γ by writing $\Gamma[x/t]$ for $\{A[x/t] : A \in \Gamma\}$.

Remark 1.3.5 The definition of $A[x/t]$ forbids substituting x by t when x occurs bound in A. There is, however, another case were substitution should be forbidden, namely when some occurrence of an individual variable in t becomes bound after the substitution $A[x/t]$. This appears, for instance, from the fact that the substitution $(\exists y(x \neq y))[x/y]$ changes the meaning of $\exists y(x \neq y)$ in an absurd way. For, $(\exists y(x \neq y))[x/y]$ is $\exists y(y \neq y)$ and, while $\exists y(x \neq y)$ may be true, $\exists y(y \neq y)$ is always false. To avoid this problem we state the following definition.

Definition 1.3.6 Let A be a formula, x an individual variable and t a term of a first-order language **L**. We define when t is *substitutable* for x in A as follows:

(i) if A is an atomic formula, then t is always substitutable for x in A;
(ii) if A is $\neg B$, then t is substitutable for x in A iff t is substitutable for x in B;
(iii) if A is $B \to C$, then t is substitutable for x in A iff t is substitutable for x in both B and C.
(iv) if A is $\forall x B$, then t is always substitutable for x in A.
(v) if A is $\forall y B$ where y is different from x, then t is substitutable for x in A iff y does not occur free in t and t is substitutable for x in B.

Remark 1.3.7 With reference to the above example 1.3.5, y is not substitutable for x in $\exists y(x \neq y)$, while z is substitutable for x in $\exists y(x \neq y)$ where z is an individual variable distinct from y.