Prologue: Hilbert's last problem

David Hilbert presented his famous list of open mathematical problems at the international mathematical congress in Paris in 1900. First in the list was Cantor's continuum problem, the question of the cardinality of the set of reals numbers. The second problem concerned the consistency of the arithmetic of real numbers, i.e., of analysis, and so on. These problems are generally recognized and have been at the centre of foundational research for a hundred years, but few would be able to state how Hilbert's list ended: namely with a 23rd problem about the calculus of variations – or so it was thought until some years ago, when German historian of science Rüdiger Thiele found from old archives in Göttingen some notes in Hilbert's hand that begin with:

> As a 24th problem of my Paris talk I wanted to pose the problem: criteria for the simplicity of proofs, or, to show that certain proofs are simpler than any others. In general, to develop a theory of proof methods in mathematics.

The 24th problem thus has two parts: a first part about the notion of simplicity of proofs, and a second one that calls for a theory of proofs in mathematics. Just as the problems that begin the list, what we call *Hilbert's last problem* has been at the centre of foundational studies for a long time.

When Hilbert later started to develop his *Beweistheorie* (proof theory), its aims were much more specific than the wording of the last problem suggests: he put up a programme the aim of which was to save mathematics from the threat of inconsistency, by which one would also 'solve the foundational problems for good'.

Gerhard Gentzen, a student of Paul Bernays with whom Hilbert was working, set as his objective in the early 1930s 'to study the structure of mathematical proofs as they appear in practice'. He presented the general logical structure of mathematical proofs as a system of rules of proof by which a path is built from the assumptions of a theorem to its conclusion. Earlier formalizations of logic had given a set of axioms and just two rules of inference. Another essential methodological novelty in Gentzen's work was that he presented proofs in the form of a tree instead of a linear succession from the given assumptions to the claim of a proof. Each step in a proof

determined a subtree from the assumptions that had been made down to that point, and these parts could be studied in isolation. Most importantly, such parts of the overall proof could be combined in new ways, contrary to the earlier linear style of proof. Gentzen was able to give for proofs in pure logic – that is, without any mathematical axioms – combinatorial transformations that brought these proofs into a certain direct form. Questions such as the consistency and decidability of a system of rules of proof could then be answered.

It has been generally thought that Gentzen's analysis of the structure of proofs cannot be carried through to perfection outside pure logic. This book aims at presenting a method in which mathematical axioms are converted into systems of rules of proof and the structure of mathematical proofs analyzed in the same way as Gentzen analysed proofs in pure logic. The overall aim is to gain a mastery over the combinatorial possibilities offered by a system of mathematical axioms. As a rule, such a complete mastery of the workings of an axiom system cannot perhaps be achieved. Our aim is to try to make a positive contribution to Hilbert's last problem by a gradual development of 'proof methods in mathematics', inspired by the methods of structural proof theory and illustrated by examples drawn mainly from the elementary axiomatics of algebra and geometry, and from what are known as systems of non-classical logic.

1 Introduction

We shall discuss the notion of proof and then present an introductory example of the analysis of the structure of proofs. The contents of the book are outlined in the third and last section of this chapter.

1.1 The idea of a proof

A proof in logic and mathematics is, traditionally, a deductive argument from some given assumptions to a conclusion. Proofs are meant to present conclusive evidence in the sense that the truth of the conclusion should follow necessarily from the truth of the assumptions. Proofs must be, in principle, communicable in every detail, so that their correctness can be checked. Detailed proofs are a means of presentation that need not follow in any way the steps in finding things out. Still, it would be useful if there was a natural way from the latter steps to a proof, and equally useful if proofs also suggested the way the truths behind them were discovered.

The presentation of proofs as deductive arguments began in ancient Greek axiomatic geometry. It took Gottlob Frege in 1879 to realize that mere axioms and definitions are not enough, but that also the logical steps that combine axioms into a proof have to be made, and indeed can be made, explicit. To this purpose, Frege formulated logic itself as an axiomatic discipline, completed with just two rules of inference for combining logical axioms.

Axiomatic logic of the Fregean sort was studied and developed by Bertrand Russell, and later by David Hilbert and Paul Bernays and their students, in the first three decades of the twentieth century. Gradually logic came to be seen as a formal calculus instead of a system of reasoning: the language of logic was formalized and its rules of inference taken as part of an inductive definition of the class of formally provable formulas in the calculus.

Young Gerhard Gentzen, a student of Bernays, set as his task in 1932 to develop a system of logic that is as close as possible to the actual proving of theorems in mathematics. His basic observation was that reasoning in mathematics uses assumptions from which conclusions are drawn. Some

steps of reasoning analyse the assumptions into their components, others move from these components towards a sought-for conclusion. The two-way rules of such reasoning make up a system known as **natural deduction** that has only rules of inference, but no logical axioms at all. This change from axiomatic to rule-based systems marks a break with the existing axiomatic tradition as upheld by Hilbert and Bernays. Each form of logical expression, say a conjunction $A\&B$ ('A and B') or an implication $A \supset B$ ('if A, then B'), has a rule that gives the sufficient conditions for inferring it: to infer $A\&B$, it is sufficient to have inferred the components A and B separately, and to infer $A \supset B$, it is sufficient to add A temporarily to the stock of assumptions that have been made, then to infer B. In these rules, logical reasoning proceeds from the desired result to its deductive conditions. The reverse step is, then, to reason from an assumption or previously reached conclusion to its deductive consequences: to infer A from $A\&B$, to infer B from $A\&B$, and to infer B from $A \supset B$ and A together.

Gentzen's analysis of the structure of proofs in logic was a perfect success. He was able to show that the means for proving a logical theorem can be restricted to those that concern just the logical operations that appear in the theorem. Instead of logical axioms, there are just rules of inference, separately for each logical operation such as conjunction or implication, to the said effect. Logic on the whole is seen as a method for moving from given assumptions to a conclusion. The Fregean tradition, instead, presented logic as consisting of a basic stock of logical truths, namely the axioms of logic, together with two rules by which new logical truths can be proved from the axioms.

When Gentzen's logic is applied to axiomatic systems of mathematics, the axioms take their place among the assumptions from which logical proofs can start. It is commonly thought that Gentzen's analysis of the structure of proofs does not go through in such axiomatic extensions of pure logic. We try to show that this need not be so: the topic of this book is a method that treats axiomatic systems in a way analogous to Gentzen's natural deduction for pure logic, namely through the conversion of mathematical axioms into rules of inference, and with results analogous to those obtained in the proof analysis of pure logic.

1.2 Proof analysis: an introductory example

(a) Natural deduction. Gentzen's rules of natural deduction give an inductive definition of the notion of a **derivation tree**. Such a tree begins, i.e.,

has as leaves, formulas that are called **assumptions**. Each logical rule prescribes how a derivation tree (in brief, a derivation) of the conclusion of the rule is constructed from derivations of its premisses. The letter I indicates that a formula with a specific structure is concluded or **introduced**, and the letter E indicates that such a formula is, as one says, **eliminated**. For conjunction $A\&B$ and implication $A \supset B$, Gentzen gave the following rules:

Table 1.1 Gentzen's rules for conjunction and implication

$$\frac{A \quad B}{A\&B}\,\&I \qquad \frac{A\&B}{A}\,\&E \qquad \frac{A\&B}{B}\,\&E \qquad \frac{\begin{array}{c}\overset{1}{[A]}\\ \vdots\\ B\end{array}}{A \supset B}\,\supset I,1 \qquad \frac{A \supset B \quad A}{B}\,\supset E$$

The rules, except for $\supset I$, are straightforward. In rule $\supset I$, a temporary assumption A is made, and a derivation of B from A can be turned into a derivation of $A \supset B$ by the rule. The square brackets indicate that the conclusion does not depend on the assumption A that has been **closed** or **discharged**. A **label**, usually a number, indicates which rule closes what assumptions.

Rules $\&I$ and $\supset E$ display one essential feature of Gentzen's work: they have two premisses so that derivation trees have binary branchings whenever these rules are applied. Each formula occurrence in a derivation tree determines a **subderivation** that lets us derive the formula, from precisely the assumptions it depends on. Often such subderivations can be rearranged combinatorially so that the same overall conclusion is obtained in a simpler way. Specifically, Gentzen's main result about natural deduction states that introductions followed by corresponding eliminations permit such rearrangements, with the effect that these steps of proof get removed from derivations. When no such simplifications are possible, all formulas in a derivation are parts or **subformulas** of the open assumptions or the conclusion. A brief expression is that **normal** derivations have the **subformula property**.

It is no exaggeration to say that the tree form of derivations that permits their transformation, in contrast to the earlier linear arrangement of Frege, Peano, Russell, and Hilbert and Bernays, was the key to all of Gentzen's central results: normalization in natural deduction, the corresponding method of cut elimination in sequent calculus, and the proof of the consistency of arithmetic.

Normalization consists in steps of **conversion** such as the following transformation of a part of a derivation:

$$\dfrac{\dfrac{\begin{matrix}\vdots & \vdots\\ A & B\end{matrix}}{A\,\&\,B}\,\&I}{A}\,\&E \quad\leadsto\quad \begin{matrix}\vdots\\ A\end{matrix}$$

We shall need the normalizibility of logical derivations for the separation of logical and mathematical steps of proof. Gentzen's rules of natural deduction require some small changes presented in Chapter 2, before this separation can be made completely transparent.

(b) The theory of equality. We assume given a domain $\mathcal{D}$ of individuals, objects a, b, c . . . of whatever sort, and a two-place relation $a = b$ in $\mathcal{D}$ with the following standard axioms:

Table 1.2 The axioms of an equality relation

EQ1	*Reflexivity:* $a = a$,
EQ2	*Symmetry:* $a = b \supset b = a$,
EQ3	*Transitivity:* $a = b \,\&\, b = c \supset a = c$.

These axioms can be added to a Frege–Hilbert-style axiomatization of logic. We shall instead first add them to natural deduction with the result that instances of the axioms can begin a derivation branch. Thus, when we ask whether a formula A is derivable from the collection of formulas Γ by the axioms of equality, arbitrary instances of the axioms can be added to Γ. We consider as an example a derivation of $d = a$ from the assumptions $a = b$, $c = b$, and $c = d$:

Table 1.3 A formal derivation in the axiomatic theory of equality

$$\dfrac{\dfrac{a = d \supset d = a \qquad \dfrac{\dfrac{a = c \,\&\, c = d \supset a = d \qquad \dfrac{\dfrac{\dfrac{a = b \,\&\, b = c \supset a = c \qquad \dfrac{\dfrac{c = b \supset b = c \quad c = b}{b = c}\,\supset E \qquad a = b}{a = b \,\&\, b = c}\,\&I}{a = c}\,\supset E \qquad c = d}{a = c \,\&\, c = d}\,\&I}{a = d}\,\supset E}{d = a}\,\supset E$$

Each topformula in the derivation is either one of the atomic assumptions or an instance of an equality axiom. The derivation tree looks somewhat

forbidding. The natural way to reason would be different, something like: *a is equal to b*, *b to c*, *c to d*, therefore *d is equal to a*. Here the principles are that equalities can be combined in chains and that equalities go both ways. The latter was applied to get the link *b equal to c* from *c equal to b*, and to get the conclusion *d equal to a* from *a equal to d*.

Logic in the derivation of $d = a$ from the assumptions $a = b$, $c = b$, and $c = d$ seems like some kind of a decoration necessitated by the use of logic in the writing of the axioms. We now want to say instead that $a = b$ gives at once $b = a$ and that two equalities $a = b$ and $b = c$ give at once $a = c$:

Table 1.4 Symmetry and transitivity as rules of inference

$$\frac{a = b}{b = a}\,Sym \qquad \frac{a = b \quad b = c}{a = c}\,Tr$$

Our example derivation becomes:

Table 1.5 A formal derivation by the rules for equality

$$\dfrac{\dfrac{a = b \quad \dfrac{c = b}{b = c}\,Sym}{a = c}\,Tr \qquad c = d}{\dfrac{a = d}{d = a}\,Sym}\,Tr$$

This should be contrasted with the logical derivation of Table 1.3.

To get the full theory of equality, we must add reflexivity as a **zero-premiss** rule:

Table 1.6 The rule of reflexivity

$$\frac{}{a = a}\,Ref$$

Now formal derivations start from assumptions and instances of rule *Ref*.

What about the role of logic after the addition of mathematical axioms as rules? A premiss of an equality rule can be the conclusion of a logical rule and a conclusion of an equality rule a premiss in a logical rule. It should be clear that logic itself should not be 'creative' in the sense of making equalities derivable from given equalities used as assumptions, if they were not already derivable by just the equality rules. To show that there cannot be any such creative use of logic, Gentzen's normalization theorem comes to help. No introduction rule can have as conclusions premisses of a mathematical rule,

because the latter do not have logical structure. Using a slight modification of Gentzen's elimination rules, the mathematical rules can be completely separated from the logical ones, so that in a normal derivation, the former are applied first, then the latter build up logical structure. Thus, if an equality is derivable from given equalities in natural deduction extended with the rules of equality, it is derivable by just the rules of equality. This separation of logic from mathematical axioms goes through for a large class of axiomatizations.

Assume there to be a derivation of the equality $a = c$ from given assumptions $a_1 = c_1, \ldots, a_n = c_n$ by the rules of equality. By what has been said, no logical rules need be used. Assume there to be a term b in the derivation that is neither a term in the conclusion $a = c$ nor a term in any of the assumptions. There is thus some instance of rule *Tr* that removes the **unknown** term b:

$$\frac{a = b \quad b = c}{a = c}\,Tr$$

If the premiss $a = b$ is a conclusion of rule *Tr*, we can permute up the instance of *Tr* that removes b, as follows:

$$\frac{\dfrac{a = d \quad d = b}{a = b}\,Tr \qquad b = c}{a = c}\,Tr \quad \rightsquigarrow \quad \frac{a = d \qquad \dfrac{d = b \quad b = c}{d = c}\,Tr}{a = c}\,Tr$$

A similar transformation applies if the second premiss $b = c$ has been derived by *Tr*. Thus, we may assume that neither premiss of the step of *Tr* that removes the term b has been derived by *Tr*. It can happen that both premisses have been derived by rule *Sym*. We then have a part of the derivation and its transformation:

$$\frac{\dfrac{b = a}{a = b}\,Sym \qquad \dfrac{c = b}{b = c}\,Sym}{a = c}\,Tr \quad \rightsquigarrow \quad \frac{\dfrac{c = b \quad b = a}{c = a}\,Tr}{a = c}\,Sym$$

In the end, at least one premiss of the step of *Tr* that removes the term b has an instance of rule *Ref* as one premiss, as in

$$\frac{d = b \qquad \dfrac{}{b = b}\,Ref}{d = b}\,Tr$$

Now the conclusion is equal to the other premiss, so the step of *Tr* can be deleted. Tracing up in the derivation the premiss $d = b$, the permutations can never lead to an instance of *Tr* that removes b and has an assumption as one premiss, because then b would be a term known from the assumption.

Thus, a derivation can be so transformed that it cannot have any unknown terms.

Consider next a derivation that has a 'cycle' or a 'loop', i.e., a branch with the same equality occurring twice:

$$\begin{array}{l} \vdots \\ a = b \\ \vdots \\ a = b \\ \vdots \end{array}$$

The part between the two occurrences can be cut out. This part may use some equalities as assumptions that are not otherwise used in the derivation, but their deletion just improves the result: we would get the conclusion with fewer assumptions. When no loops are permitted, all derivations of an equality $a = c$ from the assumptions $a_1 = c_1, \ldots, a_n = c_n$ have an upper bound on size, here defined as the length of the longest derivation tree branch: the number of distinct terms is at most $2n + 2$; therefore the number of distinct equalities is at most $(2n + 2)^2$, an upper bound on height.

The above permutation argument could have been cut short as follows. If the equality to be derived is not an instance of *Ref*, that rule can be left out. If a premiss of *Sym* or *Tr* has been concluded by *Ref*, a loop is produced. Therefore all terms must appear in equalities that are assumptions. Such a simple argument does not usually work. The permutation argument, instead, illustrates a type of combinatorial reasoning that is characteristic of all that follows, beginning with the first real example, namely lattice theory in Chapter 4.

1.3 Outline

(a) The four parts. The book has four parts. The first is based on natural deduction in the sense that mathematical rule systems are formulated as extensions of the logical rules of natural deduction. These rules define a constructive system of logic in which existence proofs are effective and no classical case distinctions (A or $\neg A$) are made. All elimination rules are formulated in the manner of disjunction and existence elimination. As long as an axiom system contains no essential disjunctions, ones that cannot be converted into equivalent formulas without disjunctions, the logical rules can be permuted below the mathematical ones. Therefore parts of

derivations by the latter rules can be separated from parts of derivations by the logical rules. The choice of classical or intuitionistic logic plays in this situation no role in the study of the derivations by such systems of mathematical rules.

With essentially disjunctive axioms, such as the linearity of an order relation, $a \leqslant b \vee b \leqslant a$, a classical sequent calculus formulation of logic permits the separation of logical and mathematical rules, in contrast to natural deduction. Sequent calculus was invented by Gentzen because he did not succeed in the proof analysis of classical logic formulated as a system of natural deduction. Part II of the book is based on sequent calculus in the sense that mathematical rule systems now extend the logical rules of sequent calculus.

We begin with axiomatic systems the axioms of which are universal, i.e., the axioms are quantifier-free formulas such as $a = b \,\&\, b = c \supset a = c$ in which a, b, and c are arbitrary parameters. Thus, such axioms could as well be written in the form $\forall x \forall y \forall z(x = y \;\&\; y = z \supset x = z)$. In Chapter 5 and in a general way in Part III, a much wider class of axioms is shown convertible to rules: those that are, in the terminology of category theory, geometric implications. Mathematical rules can now contain eigenvariables, which makes them behave like existential axioms, though without any visible logical structure.

Parts I–III build up gradually a method for an analysis of the structure of mathematical proofs. In each part, it is well defined to what kinds of axiomatic systems of mathematics the method can be applied. Part IV builds on all of the methods of the previous parts, but its focus is different. It occurred to the first author in 2003 that the method of proof analysis can be fruitfully applied to create systems of proof for modal logic and related non-classical logics: what is called the relational semantics of non-classical systems of logic, especially modal logic and its Kripke semantics, is formalized within the proof-theoretical calculi we use. The central new element, in comparison with Parts I–III, is the use of what are known as labelled logical calculi. Then, the properties that have been used previously on a semantical level can be represented by formulas that convert into rules just like the mathematical axioms treated in Parts I–III. It remains to be seen whether, in turn, the extension of purely logical proof systems in Part IV will find applications to more traditional mathematical structures.

(**b**) **Summary of the individual chapters.** The following is a list of the topics covered in the individual chapters, with an emphasis on new aspects that the method of proof analysis displays as well as on new results.