Proofs as objects

Wilfried Sieg

Abstract. The rigor of mathematics lies in its systematic organization that supports *conclusive proofs* of assertions on the basis of assumed principles. *Proofs* are constructed through thinking, but they can also be taken as objects of mathematical thought. That was the insight prompting Hilbert's call for a "theory of the specifically mathematical proof" in 1917. This pivotal idea was rooted in revolutionary developments in mathematics and logic during the second half of the 19-th century; it also shaped the new field of *mathematical logic* and grounded, in particular, Hilbert's *proof theory*. The derivations in logical calculi were taken as "formal images" of proofs and thus, through the *formalization of mathematics*, as tools for developing a theory of mathematical proofs. These initial ideas for proof theory have been reawakened by a confluence of investigations in the tradition of Gentzen's work on natural reasoning, interactive verifications of theorems, and implementations of mechanisms that search for proofs. At this intersection of proof theory, interactive theorem proving, and automated proof search one finds a promising avenue for exploring the structure of mathematical thought. I will detail steps down this avenue: the formal representation of proofs in appropriate mathematical frames is akin to the representation of physical phenomena in mathematical theories; an important dynamic aspect is captured through the articulation of bi-directional and strategically guided procedures for constructing proofs.

The objects of proof theory
shall be the proofs carried
out in mathematics proper.[1]

0. Introduction. In late December of 1933, Gödel delivered an invited lecture at the meeting of the Mathematical Association of America in Cambridge (Massachusetts); its title was, *The present situation in the foundations of mathematics*. Understanding by mathematics "the totality of the methods of proof actually used by mathematicians", Gödel articulated two problems for any foundation of mathematics. The first problem asks to state the methods of proof as precisely as possible and to reduce them to a minimum; the second seeks to justify the axioms involved and to provide "a theoretical foundation of the fact that they lead to results agreeing with each other and with empirical facts". Then Gödel

[1] This is a deeply programmatic remark in (Gentzen 1936, 499), Gentzen's classical paper in which he proved the consistency of elementary arithmetic by transfinite induction up to ε_0. It fully coheres with remarks by Hilbert, as we will see.

asserted that the first problem "has been solved in a perfectly satisfactory way" and that the solution consists "in the so-called *formalization* of mathematics," which means that a perfectly precise language has been invented, by which it is possible to express any mathematical proposition by a formula. Some of these formulas are taken as axioms, and then certain rules of inference are laid down which allow one to pass from the axioms to new formulas and thus deduce more and more propositions, the outstanding feature of the rules of inference being that they are purely formal, i.e., refer only to the outward structure of the formulas, not to their meaning, so that *they could be applied by someone who knew nothing about mathematics, or by a machine*. [My emphasis; cf. also Poincaré's remarks in Note 12.]

Today, these observations seem to be almost quaint and too obvious to be worth quoting. They do describe, however, the endpoint of a remarkable evolution away from an influential philosophical perspective that required proofs to be accompanied by the intuition of the mathematical objects they presumably deal with.[2]

How is it that the mathematical activity of proving theorems can be freed from that requirement and be represented as formal manipulation of symbolic configurations in accord with fixed rules? I will argue that the steps toward formalization are grounded in the radical transformation of mathematics in the second half of the 19-th century and the contemporaneous dramatic expansion of logic. In the Preface to the first edition of his essay *Was sind und was sollen die Zahlen?*, Dedekind asserts, "This essay can be understood by anyone who possesses, what is called common sense; …" Dedekind emphasizes then immediately, and articulates in the spirit of the last sentence of Gödel's remark, that philosophical or mathematical "school knowledge" is not in the least needed for understanding it.[3]

What is needed, however, is the capacity to *use the characteristic conditions* of precisely defined concepts in the *stepwise construction* of arguments, i.e., of proofs. The steps in proofs are *not* to appeal to "inner intuition" and, according to Dedekind, that condition is imposed by the nature of our *Treppenverstand*. Proofs, understood in this way, play a central role for Dedekind as witnessed by

---

[2] That is a requirement of Kant's. It is discussed in detail and with reference to the logicist tendencies of Dedekind and Hilbert in (Sieg 2016).

[3] (Dedekind 1932, 336) In German, "Diese Schrift kann jeder verstehen, welcher das besitzt, was man den gesunden Menschenverstand nennt; philosophische oder mathematische Schulkenntnisse sind dazu nicht im geringsten erforderlich."

his demand in the very first sentence of the Preface, "Was beweisbar ist, soll in der Wissenschaft nicht ohne Beweis geglaubt werden."[4] Two features of proofs are crucial for making plausible Dedekind's initial assertion concerning the understanding of his essay, namely, (1) the exclusive focus on characteristic conditions of concepts as starting-points and (2) the exclusive use of elementary steps for obtaining logical consequences. I point out that Hilbert called the characteristic conditions of fundamental concepts "axioms".

The first feature is examined in section 1 under the heading *Mathematical context: structural concepts*. It will be illustrated by pertinent examples from Dedekind's foundational essays and Hilbert's early work. The second feature is described in section 2, *Logical analysis: natural formal proofs*. Following Whitehead and Russell, Hilbert started in 1917 to systematically use and mathematically sharpen the formal logical tools of *Principia Mathematica*. However, the idea of using these tools to articulate difficult epistemological problems was pivotal. The problems included the decision problem, but also the call for a "theory of the specifically mathematical proof".[5] That call began to be answered only in the winter term 1921-22. That semester saw the invention of proof theory, the first steps in pursuit of the finitist consistency program, and the formulation of a new logical calculus that aimed for a more direct representation of elementary logical steps. About ten years later, this new *axiomatic calculus* of Hilbert and Bernays was transformed by Gentzen into a *rule-based natural deduction calculus* and that, in turn, is expanded into the *normal intercalation calculus*. The latter is used for the *bi-directional* construction of proofs.[6]

---

[4] In English, "What is provable should not be believed in science without proof." An underlying principled separation of *analysis* (leading to fundamental concepts) and *synthesis* (using those concepts as the sole starting-points for the development of a subject) is articulated for elementary number theory most clearly in Dedekind's letter to Keferstein (Dedekind 1890). The significance of creating new concepts is dramatically pointed out in the Preface to (Dedekind 1888, 339). For the broader methodological context, Dedekind points there to his *Habilitationsrede* (Dedekind 1854).

[5] Hilbert viewed the decision problem, i.e., the problem of deciding a mathematical question in finitely many steps, as the "best-known and the most discussed" question. This issue, he asserts, "goes to the essence of mathematical thought". (Hilbert 1918, 1113)

[6] Bernays started to work as Hilbert's assistant in the fall of 1917. He was intimately involved in every aspect of Hilbert's foundational work. In his 1922-paper, the integration of structural and formal axiomatics is expressed very clearly. He views, fully aligned with Hilbert's perspective, the representation of mathematical proofs in formalisms as a tool for their investigation not as a way of characterizing mathematics as a formal game. About the logical calculus of "Peano, Frege, and Russell" he writes: these three logicians expanded the calculus in such a way "that the thought-inferences of mathematical proofs can be completely reproduced by symbolic operations." (p. 98)

Section 3, *Natural formalization: CBT*, sketches the formal verification of the Cantor-Bernstein Theorem from the axioms of ZF.[7] It broadens the rule-based approach from logical connectives to mathematical definitions and insists that rule applications are always goal-directed and strategically guided. Thus, definitions are systematically incorporated in the bi-directional construction of proofs. Crucial is also a hierarchical conceptual organization of the material for obtaining a *mathematical frame* that facilitates the use of lemmas-as-rules. The totality of these tools reflect central aspects of ordinary mathematical practice in a completely natural way. Section 4, *Beyond formal verification*, examines the efficacy of these tools in the automated search for humanly intelligible proofs. As a case of going "beyond", I will discuss the search for proofs of Gödel's incompleteness theorems.

Sections 3 and 4 bring to life both features of Dedekind's perspective on proofs and enrich them in a modern way with computational experimentation. As far as the first feature is concerned, Saunders MacLane suggested in his (1934) that structural concepts are obtained by detailed *analyses of proofs*. As far as the second aspect is concerned, Dedekind himself hoped that his essay might stimulate other mathematicians to reduce the "long sequences of inferences to a more moderate, more pleasant size". (Dedekind 1888, 338) Clearly, changes of mathematical frames as well as of proof strategies can be implemented and set to work for experiments in a computational environment.

Ever since Gödel's second incompleteness theorem revealed the profound difficulty of the finitist consistency problem and Gentzen's calculi offered tools to address it, proof theorists have been preoccupied with the investigation of *formal deductions* in theories for arithmetic, subsystems of analysis or parts of set theory. In order to make their work relevant for mathematical practice, it has of course been important that mathematical proofs *can be formalized* in those theories, but questions like "*How* are they formalized?", "Do they have *structural features of practical significance*?", "Are these structural features of *broader methodological interest*?", and "Do they reveal *crucial aspects of mathematical cognition*?" – such questions have not been topics of detailed investigations.

---

[7] The full verification is presented in (Sieg & Walsh 2019) and discussed with a particular focus on "proof identity" in (Sieg 2019); see corresponding remarks in section 4 below.

In this paper, I am taking tentative steps toward finding ways to answer such questions. If mathematics is, as Gödel suggested in 1933, "the totality of the methods of proof actually used by mathematicians", then a third problem has to be raised – in addition to the two problems Gödel articulated for any foundation. An adequate representation of proofs cannot be achieved by using the minimum of formal methods. Rather, it has to be given in appropriate formal frames akin to the representation of physical phenomena in mathematical theories; important dynamic aspects are captured through the articulation of systematic procedures for constructing proofs.

1. Mathematical context: structural concepts. *Mathematical proofs* are constructed through thinking and can be objects of mathematical thought. That was Hilbert's idea already in 1900 when he formulated the *24-th Problem* for his Paris talk; see (Thiele 2003). Ultimately, he may have viewed the problem as too open-ended, since it is not included in his final list of *Mathematical Problems*. The problem requested:

Develop a theory of the method of proof in mathematics in general.

The request is followed by a remark on the simplicity of proofs that claims, "under a given set of conditions there can be but one simplest proof". Hilbert continues,

Quite generally, if there are two proofs for a theorem, you must keep going until you have derived each from the other, or until it becomes quite evident what variant conditions (and aids) have been used in the two proofs.

The development of an investigation of proofs began only after Hilbert had addressed the Swiss Mathematical Society in Zürich on 11 September 1917 and demanded in his talk *Axiomatisches Denken*,

… we must - that is my conviction - turn the concept of the specifically mathematical proof into an object of investigation, just as the astronomer considers the movement of his position, the physicist studies the theory of his apparatus, and the philosopher criticizes reason itself.

Hilbert admitted immediately that "the execution of this program is at present, … , still an unsolved task".[8]

[8] The call for such an investigation was not a whim for Hilbert. After all, we just saw that he had already formulated in 1900 the 24-th problem concerning "a theory of the method of proof in mathematics".

At the time of this talk, Hilbert was 55 years old and one of the world's most distinguished mathematicians. He had done groundbreaking work in core areas of mathematics during the 1890s that culminated in his *Zahlbericht*; he had pursued work on the foundations of geometry and analysis that was published in *Grundlagen der Geometrie* and in *Über den Zahlbegriff*; he had lectured since the mid-1890s on various parts of physics and in 1915 he had contributed to the theory of general relativity. This remarkable experience in mathematics and its "applications" is reflected in his Zürich talk.

Hilbert begins his talk by discussing the relations of mathematics to the two "great realms" of physics and epistemology. He claims that the essence of these relations becomes especially clear when describing the *general research method*, "which seems to have become more and more prominent in the newer mathematics", the *axiomatic method*. Why is this method becoming prominent in the "newer" mathematics? Isn't its *classical* paradigm already found in Euclid's *Elements*? Does Hilbert's *Grundlagen der Geometrie* not follow that very method and present the mathematical core of the *Elements* without the "flaws" that had been discovered in the 19-th century?

"Yes", Hilbert's *Festschrift* does take into account the corrections that had been proposed for a logically flawless foundation, but also most definitely "No": it is not an improved version of the *Elements* as an exposition of geometry. It is rather a penetrating reflection on the conceptual organization of the field and on meta-mathematical issues like independence and relative consistency. The axiomatic system is set up as a *structural definition* articulating the properties of the fundamental geometric notions that characterize relations between three unspecified systems of things:

> We think three different systems of things: we call the things of the first system *points* …; we call the things of the second system *lines* …; we call the things of the third system *planes* …; We think the points, lines, planes in certain mutual relations ..; the precise and complete description of these relations is obtained by the *axioms of geometry*.

Hilbert's axioms do not express a priori truths but define the abstract notion of a *Euclidean Space* in the same way in which the axioms for groups or fields define the concept of a *group* or of a *field*.[9]

---

[9] See my paper (Sieg 2014), in particular, the analysis of the Frege-Hilbert correspondence.

Hilbert followed in the footsteps of Dedekind who had introduced in *Was sind und was sollen die Zahlen?* the notion of a *simply infinite system* and considered as *natural numbers* the elements of any system falling under that notion. Here is Dedekind's formulation:

A system N is *simply infinite* if and only if there is an element 1 and a mapping ϕ, such that the characteristic conditions (α) – (δ) hold for them.

The characteristic conditions of the concept are: (α) ϕ[N] is contained in N, (β) N is the chain of the system {1} with respect to ϕ, (γ) 1 is not an element of ϕ[N], and (d) ϕ is injective. - Dedekind formulated *meta-mathematical problems* and achieved significant *results*: (1) He proved the "consistency" of the notion via the logically (and problematically) defined system ***N*** or, as he put it in his famous letter to Keferstein, he showed that the notion simply infinite system does not contain an "internal contradiction". (2) He proved a *representation theorem* showing that every simply infinite system is isomorphic to ***N*** and directly implying that the notion is *categorical*. (3) He argued for the *proof theoretic equivalence* of different simply infinite systems.[10]

The argument for the last result exploits the two features of Dedekind's approach I emphasized in the Introduction, namely, the requirement that in proofs *only* the characteristic conditions of the fundamental concepts are appealed to as starting points, and the conviction that the elementary logical steps are the same in every area of mathematics. This approach is prefigured in his earlier essay *Stetigkeit und irrationale Zahlen* for the concept of a complete ordered field and executed for real numbers in Hilbert's *Über den Zahlbegriff*. For both Dedekind and Hilbert, the nature of the objects in "models" does not enter proofs. (Here, a "model" is just any system that falls under the structural concept; when arguing that a particular system falls under the structural concept, the nature of the objects plays of course a role.) In a different but complementary way, Hilbert made the same point in *Grundlagen der Geometrie* by giving an analytic model for geometry and a geometric model for analysis.

[10] The notion of proof theoretic equivalence was introduced in (Sieg & Morris 2018) and is based on a close reading of #73 and #134 in (Dedekind 1888), as well as the study of the penultimate version of that essay.

This perspective on mathematics is an absolutely radical conception of the subject. Dedekind and Hilbert locate their work in *logic*, broadly conceived. For Dedekind, mathematics *is* part of logic: the work in his two foundational essays, (1872) and (1888), gives a logical analysis of the number concept and develops the theory thereof systematically. For Hilbert, *Grundlagen der Geometrie* gives a "logical analysis of our spatial intuition" and investigates the question of what can be proved from what – without appealing to the intuition that is being analyzed. The issue to be addressed next is, what inferential principles can be used in proofs?

2. Logical analysis: natural formal proofs. On account of the fact that Dedekind had not given an explicit list of inferential steps, Frege severely and polemically criticized the former's essay (1888); that critique clearly extends to Hilbert's later *Grundlagen der Geometrie*. In the Preface of his *Grundgesetze der Arithmetik*, Frege claimed that the brevity of Dedekind's development of arithmetic in (1888) is only possible, "because much of it is not really proved at all". He continued:

> … nowhere is there a statement of the logical or other laws on which he builds, and, even if there were, we could not possibly find out whether really no others were used – for to make that possible the proof must be not merely indicated but completely carried out.[11]

Apart from making the logical principles explicit there is an additional aspect that is hinted at in Frege's critique and detailed below.[12] Frege thinks of his own work as standing in the Euclidean tradition but going beyond it, because it lists not only the axioms in advance, but also the inferential principles. Both are articulated in a precise and expressive artificial language. Furthermore, the

---

[11] *Grundgesstze der Arithmetik*, p. 139 of *Translations from the philosophical writings of Gottlob Frege*, Peter Geach and Max Black (eds.), Oxford 1977.

[12] Poincaré's 1902-review of Hilbert's *Grundlagen der Geometrie* brings out this additional aspect in a quite vivid way, namely, through the idea of formalization as machine executability: "M. Hilbert has tried, so-to-speak, putting the axioms in such a form that they could be applied by someone who doesn't understand their meaning, because he has not ever seen either a point, or a line, or a plane. It must be possible, according to him, to reduce reasoning to purely mechanical rules." Indeed, Poincaré suggests giving the axioms to a reasoning machine, like Jevons' logical piano, and observing whether all of geometry would be obtained. Such formalization might seem "artificial and childish", were it not for the important question of completeness: "Is the list of axioms complete, or have some of them escaped us, namely those we use unconsciously? … One has to find out whether geometry is a logical consequence of the explicitly stated axioms, or in other words, whether the axioms, when given to the reasoning machine, will make it possible to obtain the sequence of all theorems as output [of the machine]."

inferential principles have to be applied in a rule-bound manner.[13] Frege asserted that in his logical system "inference is conducted like a calculation" and observed:

> I do not mean this in a narrow sense, as if it were subject to an algorithm the same as … ordinary addition or multiplication, but only in the sense that there is an algorithm at all, i.e., a totality of rules which governs the transition from one sentence or from two sentences to a new one in such a way that nothing happens except in conformity with these rules. (Frege 1984, 237)

That points to one crucial element of modern logic – developed by Frege himself and, among others, Peano, Whitehead and Russell – that is the basis for *formally* presenting parts of mathematics. The question of what might be understood by a formal presentation or a formal theory led in the 1930s to the introduction of rigorous mathematical notions to characterize "formality", "algorithm" or "computation". The most familiar mathematical notions are Gödel's *general recursiveness*, Church's *λ-definability* and Turing's *machine computability*.[14]

Hilbert had argued already in his (1905) for the joint development of logic and mathematics in pursuit of the meta-mathematical goal of a "direct" proof for consistency. He did not get very far, as his view of symbolic logic in the modern sense was extremely limited. He began studying *Principia Mathematica* in 1913. That provided the background for taking up such a joint development in 1917 with the *full logical toolbox* of Whitehead and Russell's work. The latter was radically reshaped and extended in his lectures during the winter term 1917-18. These lectures make for dramatic reading, as one can see in them the invention and presentation of a new subject that Hilbert called *mathematical logic*. One goal of this mathematical logic was the *full formalization* of mathematical practice. In a rough and ready way, the lectures achieved that goal for parts of number theory and analysis. Hilbert recognized, however, that a more adequate representation of proofs required moving from the awkward calculus of *Principia Mathematica* toward a more *direct* method of formalization.

---

[13] It is a normative requirement (to insure intersubjectivity on a minimal cognitive basis), but it is also a practical one: if an inferential step required a proof that its premises imply its conclusion, we would circle into an infinite regress. Frege pursued also the particular philosophical goal of gaining "a basis for deciding the epistemological nature of the law that is proved." (Grundgesetze, 118)

[14] The adequacy of these notions for capturing the informal concepts is still discussed today under the special headings, Church's Thesis, or Turing's Thesis or the Church-Turing Thesis. Tn my (Sieg 2018), I gave a structural-axiomatic characterization of the *concept* of computation. That turns the "adequacy problem" into a standard problem any mathematical concept has to face when confronted with the phenomena it is purportedly capturing.

Hilbert and Bernays took a first step in that direction via a novel *axiomatic calculus* they introduced in early 1922. It was used in all their subsequent proof theoretic work extending to 1939, when the second volume of *Grundlagen der Mathematik* was published. The reasons for introducing this calculus and abandoning that of *Principia Mathematica* are partly methodological and partly pragmatic. Methodologically, the organization of the calculus into groups of axioms for each individual logical connective was to parallel that of the axiomatic system in *Grundlagen der Geometrie*, where groups of axioms were formulated for the basic concepts of geometry. Here are the axioms for conjunction and disjunction:

Conjunction I:	$A \to (B \to A\&B)$	
Conjunction E:	$A\&B \to A$	$A\&B \to B$
Disjunction I:	$A \to A \vee B$	$B \to A \vee B$
Disjunction E:	$A \vee B \to ((A \to C) \to ((B \to C) \to C)))$	

Pragmatically, the new calculus was to provide a more natural and direct way of formalizing proofs by bringing out the meaning of the logical connectives in logical inference steps. For example, having proved the conjunction A&B, the step to A would be mediated in this way: the proof of A&B is expanded by the axiom $A\&B \to A$, then A is inferred via the rule of modus ponens. (They consider a linear presentation of proofs; if A&B has been obtained in a proof, then A&B can be repeated at the end of the given proof.)[15] To reach a *natural formalization* was not only Hilbert's goal, but a few years later also Gentzen's.

Gentzen is the proof theorists' proof theorist. His investigations and insights concern prima facie formal proofs. However, when he specifies in his 1936-paper the concept of a *deduction* he adds in parentheses *formal image of a proof*, i.e., deductions are viewed as formal images of mathematical proofs and

[15] Proofs officially are sequences of formulas, but for the proof theoretic investigations, Hilbert and Bernays turn them into *proof trees* by a process they call "Auflösung in Beweisfäden".

are obtained by formalizing the latter. The process of formalization is explained as follows:

> The words of ordinary language are replaced by particular *signs*, the logical inference steps [are replaced by] rules that form new formally presented statements from already proved ones.

Only in this way, Gentzen claims, is it possible to obtain a "rigorous treatment of proofs". As if to make that point crystal clear, he presents in minute detail an informal proof of the infinity of primes using as inferential principles only steps that turn out to be rules of his natural deduction calculus.[16] Returning to the remark on deductions as formal images of mathematical proofs, he emphasizes, "*The objects of proof theory shall be the proofs carried out in mathematics proper*." (Gentzen 1936, 499)

The calculus of Gentzen's (1936) is a natural deduction one and goes back to his thesis. It is a version of Hilbert and Bernays' calculus in which the axioms have been transformed into I(ntroduction)- and E(limination)-rules for the logical connectives and to which one crucial new feature has been added: *making and discharging assumptions*. This 1936-paper used, in Gentzen's own terminology, a *sequent formulation of natural deduction*, i.e., at each node of a proof tree there is a sequent configuration $\Gamma \supset \phi$. In the top-down, forward construction of proofs all the directly logical actions happen on the r.h.s. of the sequent symbol $\supset$; the finite set Γ of formulas keeps only track of the assumptions on which the proof of ϕ depends. Notice that in this way of building proofs, one can make "detours" by inferring a formula by an I-rule and immediately applying the corresponding E-rule. Proofs without detours are called *normal*.[17]

Let me emphasize most strongly that Gentzen viewed making and discharging assumptions as *the* characteristic feature for his natural deduction calculi and as reflecting an *absolutely fundamental aspect of* proof construction in *mathematical practice*. Another general move in the construction of informal

---

[16] It is unfortunate that "Kalkül des natürlichen Schließens" has been translated as "natural deduction calculus". "Calculus of natural reasoning" would express better that it is a calculus reflecting in a formal way patterns of natural, informal argumentation.

[17] Gentzen's *Urdissertation* (Gentzen 1932-33) contains a formulation of a natural deduction calculus for intuitionist logic. Gentzen proved for that logic a normalization theorem; see (von Plato 2008) and (Sieg 2009). The systematic investigations of both classical and intuitionist natural deduction calculi was taken up in (Prawitz 1965); Prawizt established normalization theorems and discovered important structural features of normal proofs.

proofs is, however, not reflected in the syntactic configurations Gentzen considered. The additional general move is *taking backward steps*, for example, in proofs of universally quantified assertions or in indirect arguments. So, the question is, how can backward steps be joined with forward steps in a *single syntactic configuration*?

For the purpose of obtaining a single syntactic configuration, Gentzen's *sequent presentation* of natural deduction is a convenient starting-point. It can be modified to an *IC calculus*[18] that allows bi-directional reasoning through *intercalating* formulas and *using structured sequents* of the form $\Gamma; \Delta \supset \phi$. Γ still has the role of containing the assumptions on which ϕ depends, whereas Δ consists of formulas that are obtained by E-rules, (successively) applied to elements of Γ. The formula ϕ on the r.h.s. of the sequent symbol $\supset$ is called *goal*. Forward steps toward the goal are taken via E-rules on the l.h.s., whereas backward steps from the goal are reflected by applications of inverted I-rules on the r.h.s.. Notice that the E-rules yield as consequences only (strictly) positive subformulas of their premises. In sum, the forward steps make it possible to *extract* more specific information from assumptions and backward steps lead by *inversion* to less complex new proof obligations.

A modification of the IC calculus, its normal *NIC* version, restricts forward steps to be *goal-directed* in the following sense: elimination rules are applied only if the goal formula ϕ is a (strictly) positive subformula of an assumption in Γ. Here is a very simple proof that involves only conjunction but allows the reader, nevertheless, to discover the fundamental ideas.

$$\dfrac{\dfrac{\dfrac{\underline{(A\&B)\&C;\ A\&B,\ A \supset A}}{\underline{(A\&B)\&C;\ A\&B \supset A}} \qquad \underline{(A\&B)\&C;\ C \supset C}}{\underline{(A\&B)\&C \supset A} \qquad\qquad \underline{(A\&B)\&C \supset C}}}{(A\&B)\&C \supset A\&C}$$

[18] "Intercalate" does not only mean "interpolate an intercalary period in a calendar" but also "insert something between layers in a crystal lattice or other structure". So, I apply it to the insertion of formulas between layers in logical proof structures. ("Interpolate" could not be used in this logical context for obvious reasons.)

Remark. The evolution of natural deduction calculi and their diagrammatic presentation is described in great detail in section 1 of (Sieg and Derakhshan). That description includes, in addition, an important discussion of the diagrams that were used by Jaskowski and Fitch. Indeed, the representation of bi-directionally constructed natural deduction proofs is given by *restricted Fitch-diagrams*. The above proof construction would begin with the configuration

(A&B)&C	Premise
.	
A&C	Goal

The task is to bridge the gap between Premise and Goal. The elimination strategy would not be used, as the Goal is not a positive subformula of the Premise. So, the inverted &-I would yield the next configuration:

(A&B)&C	Premise
.	
A	New Goal
.	
C	New Goal
A&C	Goal

The reader can undoubtedly take the next steps in this bi-directional proof construction to obtain this completed proof:

1. (A&B)&C	Premise
2. A&B	&El: 1
3. A	&El: 2
4. C	&Er: 1
5. A&C	&I: 3, 4

This is a representation that is obviously much more economical and genuinely suitable for a computer screen, in particular, once subderivations are being considered in a quasi-linear presentation of "boxes" within "boxes" as it is done via Fitch diagrams. For human construction on paper it is most inconvenient, whereas the computer's capacities can be effectively exploited for this naturally bi-directional construction. One can find in (Sieg & Walsh 2019) complex examples. Perhaps it should be emphasized that the finished proof can be presented with annotations that reflect the order of the steps taken for obtaining this proof. **End of Remark**

NIC proofs are easily translated into normal natural deduction proofs. Their *construction* is strategically guided to close the gap between assumptions and goals by intercalating steps. The strategies reflect, on the one hand, guidance for informal proof construction and, on the other hand, rely on structural features of normal proofs.[19] They are very efficient for the construction of proofs in logic, but how can their use be expanded to formally represent proof construction in (parts of) mathematics?

3. Natural formalization: CBT. For the expansion of the basic approach of bi-directional argumentation to mathematics, the logical toolbox has to be enriched by new components:
(1) the specification of *definitions* by E- and I-rules extending the rule-based approach from logical connectives to mathematical operations and notions;
(2) a well-structured *conceptual organization* its *formal frame*;
(3) the integration of proofs with the formal frame by using *lemmas-as-rules*.[20]
These are the characteristic components of *natural* formalization; "natural" is used here in analogy to how it is used in *natural* deduction. They are used informally and formally.

---

[19] For the NIC calculus one can prove a *strengthened completeness theorem* for both classical and intuitionist logic: for any ϕ and Γ, there is either a *normal* proof of ϕ from Γ or a counterexample to Γ, ϕ. The central considerations in the completeness proof have been organized into efficient logical strategies for automated search and have been implemented in the system AProS. The completeness proofs are found in (Sieg and Byrnes 1998) and (Sieg and Cittadini 2005).

[20] Lemmas have been used as rules in classical texts but, of course, also for the standard structuring of mathematical expositions. Classical examples are found in Book I of Euclid's *Elements* (for example, in the proof of the Pythagorean Theorem the earlier, very important theorem I.4 is being used in just that way), but also in the development of the theory of systems and mappings in (Dedekind 1888).

Such an expansion was implemented in the system AProS. In (Sieg & Walsh 2019), AProS was used as a proof checker to formally verify in ZF a classical result of set theory, the Cantor-Bernstein Theorem CBT:

If $f \in \text{inj}(a,b)$ and $g \in \text{inj}(b,a)$, then $a \approx b$.

Here, $f \in \text{inj}(a,b)$ expresses that f is an injection from a to b and $g \in \text{inj}(b,a)$ expresses similarly that g is an injection from b to a; the consequent of the conditional states that a and b are equinumerous, i.e., there is a bijection between a and b.

CBT is actually equivalent to the assertion I have come to call *Dedekind's Fundamental Lemma*:

If $e \subseteq d \subseteq a$ and $a \approx e$, then $a \approx d$.

The proof of the Fundamental Lemma using CBT is direct. To obtain CBT from this Fundamental Lemma, we have to specify subsets of a that satisfy the premises of the Lemma (using the premises of CBT). That is done for lines 3, 4 and 5 by appealing to very straightforward lemmas that are indicated in the proof of CBT in the AProS interface:

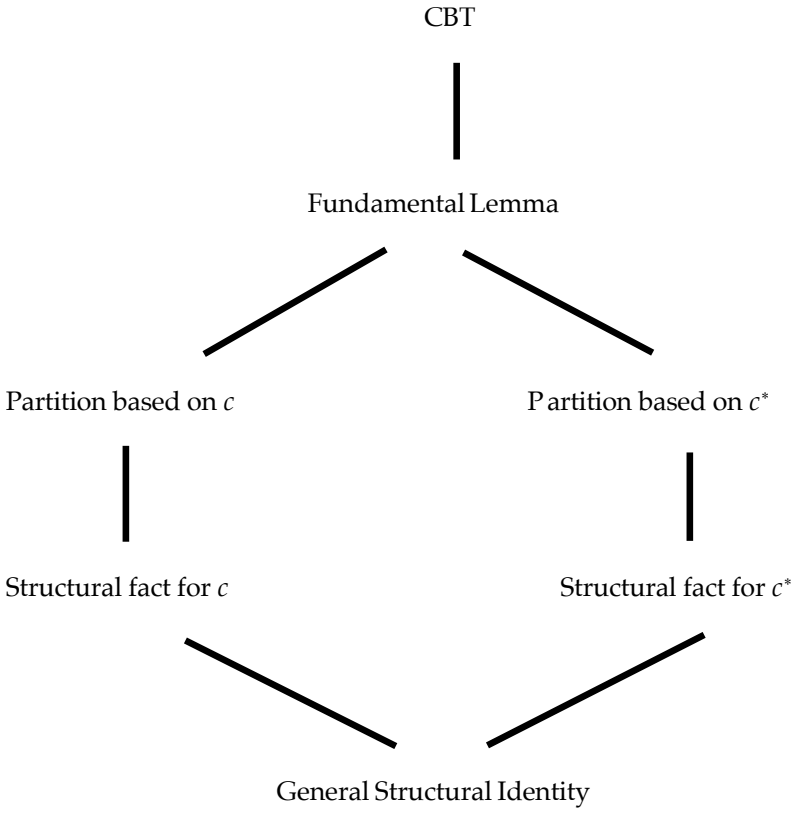
1. $f \in \text{inj}(a, b)$	*Premise*
2. $g \in \text{inj}(b, a)$	*Premise*
3. $1(g \circ f) \in \text{bij}(a, g \circ f[a])$	*Theorem (Core12)*: 1, 2
4. $g[b] \subseteq a$	*Theorem (Func17)*: 2
5. $g \circ f[a] \subseteq g[b]$	*Theorem (Comp11)*: 1, 2
6. $a \approx g[b]$	*Theorem (Fundamental Lemma)*: 3, 4, 5
7. $b \approx g[b]$	*Theorem (Equi4)*: 2
8. $a \approx b$	*Theorem (Equi8)*: 6, 7

Proof of CBT in the AProS interface.

The proof crucially appeals to Dedekind's Fundamental Lemma as a rule and also to a few straightforward and individually meaningful lemmas of elementary set theory, e.g., line 8 is obtained from lines 6 and 7 by appealing to the transitivity of equinumerosity. This proof is based on a theoretical analysis of mathematical proofs and gives, as MacLane would suggest (1934, 57), "an adequate logical explanation of the fact, that genuine mathematical proofs can be given through short descriptions".

Dedekind proved the Fundamental Lemma and, from it, CBT in 1887; he never published the proofs. Zermelo proved CBT in 1908 without knowing of Dedekind's proof, but being quite familiar with Dedekind's notion of the *chain* of a system a w.r.t. a mapping f.[21] Such chains k satisfy a central *structural identity*, namely, $k = a \cup f[k]$. Dedekind and Zermelo considered in their respective proofs two different chains c and c*. Instantiated with c and c^*, the structural identities are used to define two partitions of a and d, as well as bijections between their respective components; the unions of these bijections are then bijections between a and d.

The overall organization of the proofs is depicted in the diagram below. In (Sieg & Walsh 2019), everything is proved in ZF from the ground up and is literally identical except for the parts of the proofs represented by the diamond in the diagram. I consider it as a paradigmatic case of natural formalization that includes, of course, the detailed analysis of mathematical proofs as a crucial component.



[21] It makes explicit the informal notion of *objects obtained from elements in a by finitely iterating f* ; it is defined as the intersection of all sets that contain a as a subset and are closed under f.

This is only one direction for investigating mathematical proofs. Other directions can be pursued. Recall Hilbert's request when he articulated his 24-th problem:

> Quite generally, if there are two proofs for a theorem, you must keep going until you have derived each from the other, or until it becomes quite evident what variant conditions (and aids) have been used in the two proofs.

Concerning Dedekind's and Zermelo's proofs the following question can be raised, *Are they at all different* or *are they identical*? In 1932, when these proofs could be compared for the first time, Emmy Noether viewed them as "exactly the same", whereas Zermelo thought of them as "inessentially different". In Hilbert's spirit, we can gain a mathematical insight from the comparison and see very clearly "what variant conditions (and aids) have been used in the two proofs": in Dedekind's proof, the chain c is the smallest fixed-point of the general structural identity $k = a \cup f[k]$, whereas in Zermelo's proof c^* is the largest.

Proof identity has been and will continue to be a central topic for any theory of mathematical proofs. In the case of CBT, I am convinced that all its known proofs are either of the Dedekind or Zermelo variety.[22] So, we have gained material that can inform a theory of mathematical proofs. For example, it seems clear that proof identity – tentatively taken as literal syntactic identity of their formalizations – must be a notion that depends both on the mathematical frame and on the strategies for proof construction. In the next section, I will introduce another source of complementary material, the automated search for *humanly intelligible formal proofs*.

4. Beyond formal verification. The focus on human-centered automated proof search has a long tradition in computer science. I am thinking, in particular, of Woody Bledsoe's work in the 1970s and 1980s. This tradition seems, at the moment, to be dormant in computer science. However, the mathematician Gowers has expressed a strong interest in what he calls "extreme human-centered automatic theorem proving"; see the *Interview with Sir Timothy Gowers*,

---

[22] That point is at the center of my paper (Sieg 2019). Aries Hinkis' book *Proofs of the Cantor-Bernstein Theorem: A mathematical excursion* is a comprehensive discussion of proofs and their history. I have examined most of the proofs in that mathematical excursion, with special care given to the proofs of König and Banach. They all fall into one of these categories.

(Diaz-Lopez 2016). A strategically guided automatic theorem prover, let's call it the *G&G-prover*, is presented in his joint paper with Ganesalingam (2017) and its earlier version (Ganesalingam and Gowers 2013). Their *central goal* is to generate, with the help of the automatic prover, a proof in English; the generated proof text is programmatically to be indistinguishable from excellent mathematical writing. That's feasible only, they rightly argue, if the automatic prover uses strategies that are also used by human mathematicians as every formal step taken by the prover is *paraphrased directly* in English. The latter is a decision on their *basic method* that is supposed to reflect the simple "rhetorical structure" of mathematical arguments, constituted by linearly organized forward steps.

Their central goal, basic method and observation on human-like strategies limit the moves their prover can make. Indeed, Ganesalingam and Gowers state that their prover, on account of the constraints on moves, is solving only "routine problems". Here are a few of the restrictions imposed: no backtracking moves are allowed, as they would interfere with the linear step-by-forward-step writing of proofs; the first-order language they are using does not contain negation; the treatment of disjunction is described as "work in progress" and will definitely require backtracking. In (Sieg & Derakhshan 2020) the G&G-prover is analyzed and compared with AProS. It is shown that the G&G-prover's mathematical work can easily be mimicked in a very weak fragment of AProS. Nevertheless, I want to emphasize that there is a common perspective, namely, that mechanisms for human-centered automated proof search should "think" like mathematicians.[23]

I will turn now to a second example of a detailed analysis and significant reshaping of a complex proof that, ultimately, provided the tools for a successful automated search. During the last two years of my graduate studies at Stanford, I was working as a research associate at Patrick Suppes' IMSSS (Institute of Mathematical Studies in the Social Sciences). The Institute was then a hotbed of *computer assisted instruction*, CAI. The Suppes team had developed a computer-

[23] Gowers' and my work actually have a common motivating ambition: a pedagogical one. Gowers' is articulated in (Diaz-Lopez 2016); mine has propelled the development of a fully web-based course *Logic & Proofs*; see (Sieg 2007). The AProS website has this URL http://www.phil.cmu.edu/projects/apros/. AProS forms the basis of a dynamic, interactive *proof tutor* in this introduction to logic. The course has as its primary goal teaching students strategic proof construction. *Logic & Proofs* has been completed by more than 12,000 students for credit at their home institution.

based course in ZF set theory. My task was to formalize the proofs of Gödel's incompleteness theorems and closely related ones, all for ZF. How could one accommodate on a small screen, I asked myself at the very beginning, the sheer notational complexity of dealing with the formal system ZF, its meta-mathematical description, the arithmetization of that description and, finally, its representation within ZF? – My answer was, in the end: "Omit arithmetization and, instead, *represent directly* the central syntactic notions!" The latter notions (including Substitution, Formula and Proof in ZF) are, after all, precisely given by elementary inductive definitions and structural recursions. With this strategic simplification I succeeded in formally verifying the incompleteness theorems and related ones, for example, Löb's Theorem.[24]

Many years later, around 2003, AProS was endowed with the two levels of argumentation used for my Gödel proofs, in the object-, respectively meta-theorey. I formulated suitable principles connecting these two levels, ProvI and ProvE. The first principle of *provability introduction* allows the step from an object theoretic ZF-proof of ϕ to the meta-theoretic assertion that ϕ is provable in ZF; the second principle of *provability elimination* allows the complementary step from the meta-theoretic assertion that ϕ is provable in ZF to using ϕ as a theorem in an object theoretic proof. With these principles, the semi-representability of the theorem predicate and the defining bi-conditional of the Gödel sentence, AProS very quickly finds a proof of the first theorem that is absolutely canonical. Indeed, it does so also for the second incompleteness theorem and related theorems, in particular, Löb's Theorem; see (Sieg & Field 2005).

I consider detailed, unifying proof analyses and natural formalizations as necessary steps toward an investigation of the notion of *mathematical proof*. We have to exploit the rich body of mathematical knowledge that *is systematic*, but that is also *structured for human intelligibility and discovery*. In this way, we can isolate creative elements in proofs and formulate suitable heuristics. Such work gets us closer to uncovering techniques of our thinking and to realizing Hilbert's "fundamental idea" for his proof theory, namely, "to describe the activity of our

[24] See (Sieg 1978) and (W. Sieg, I. Lindstrom and S. Lindstrom 1981). Ironically, it is quite clear that Gödel himself used such a direct representation in his original argument for a version of the first incompleteness theorem; see Gödel's description of his discovery in (Wang 1981, 654).

understanding, to make a protocol of the rules according to which our thinking actually proceeds." (Hilbert 1927)

The above work on natural formalization and automated proof search has given us insights: the particular representation of proofs as restricted Fitch-diagrams allows the systematic bi-directional construction of *normal* proofs and reflects mathematical practice; for more advanced parts of mathematics, an embedding in a conceptually organized mathematical frame is necessary. Going beyond the analysis and formalization of proofs, we can implement search procedures and run computer experiments, observing how strategically motivated modifications affect search. This way of proceeding is also a way of building cognitive models of proof construction that are deeply informed by the practice of mathematics and may reveal sophisticated capacities as well as real limitations of the human mathematical mind. In any event, at the intersection of proof theory, interactive theorem proving and automated proof search, we find ways for exploring the structure of the mathematical mind.

Let me end by listening to another voice that connects the rich past with this programmatic future: Saunders MacLane was one of the last logic students in Göttingen and a good friend of Gentzen's. He completed his thesis *Abgekürzte Beweise im Logikkalkul* in 1934. A year later, he published a summary in which he pointed out that proofs are not "mere collections of atomic processes, but are rather complex combinations with a highly rational structure". He reflected in 1979 on his early work in logic and remarked, "There remains the real question of the actual structure of mathematical proofs and their strategy. It is a topic long given up by mathematical logicians, but one which still – properly handled – might give us some real insight." Obviously, I share MacLane's hope. I also think that the work described above already has given us some insights; whether they are to be called "real" or not, I leave to the reader's judgment.[25]

[25] Implicitly, I argued against an artificial opposition of informal and formal proofs. By incorporating the strategic, dynamic aspect of interactive proof construction into a fully automated search procedure, one obtains the means for exploring structural features of proofs and their construction.

References

Bernays, P.
1922 Hilberts Bedeutung für die Philosophie der Mathematik; *Die Naturwissenschaften* 4, 93-99.

Bledsoe, W.
1977 Non-resolution theorem proving; *Artificial Intelligence* 9, 1-35.
1983 The UT natural-deduction prover; Technical Report, Department of Computer Science, University of Texas, Austin.

Dedekind, R.
1854 Über die Einführung neuer Funktionen in der Mathematik; Habilitationsvortrag; In (Dedekind 1932, 428-438) and translated in (Ewald 1996, 754-762).
1872 *Stetigkeit und irrational Zahlen*; Vieweg. Translated in (Ewald 1996, 765-779.
1888 *Was sind und was sollen die Zahlen?*; Vieweg. Translated in (Ewald, 1996, 787–833).
1890 Letter to H. Keferstein, Cod. Ms. Dedekind III, I, IV. Translated in van Heijenoort (editor), *From Frege to Gödel*, Harvard University Press 1967, 98-103.
1932 *Gesammelte Mathematische Werke*, Vol. 3; Vieweg.

Diaz-Lopez, A.
2016 Interview with Sir Timothy Gowers, Notices of the American Mathematical Society 63 (9), 1026-1028.

Ewald, W. B. (editor)
1996 *From Kant to Hilbert: Readings in the Foundations of Mathematics*; Oxford University Press.

Fitch, F.
1952 *Symbolic Logic: An Introduction*; Ronald Press, N.Y.

Frege, G.
1893 *Grundgesetze der Arithmetik, begriffsschriftlich abgeleit80*; Pohle Verlag.
1980 *Gottlob Freges Briefwechsel*; edited by G. Gabriel, F. Kambartel, and C. Thiel, Meiner.
1984 *Collected Papers on Mathematics, Logic, and Philosophy*; edited by B. McGuinness. Oxford University Press.

Ganesalingam, M. and Gowers, T.
2013 A fully automatic problem solver with human-style output. arXiv:1309.4501.
2017 A fully automatic theorem prover with human-style output; *J. Automated Reasoning* 58, 253-291.

Gentzen, G.
1932/3 "Urdissertation", Wissenschaftshistorische Sammlung, Eidgenössische Technische Hochschule, Zürich, Bernays Nachlass, Ms. ULS.
1934/5 Untersuchungen über das logische Schließen I, II, *Mathematische Zeitschrift* 39, 176–210, 405–431
1936 Die Widerspruchsfreiheit der reinen Zahlentheorie. *Mathematische Annalen* 112 (1), 493–565.

Gödel, K.
1933 The present situation in the foundations of mathematics; in S. Feferman e.a. (editors), *Gödel's Collected Works* vol. III, Oxford University Press, 1995, 36-53.

Hilbert, D.
1899 *Grundlagen der Geometrie*; Teubner.
1900 Über den Zahlbegriff; *Jahresbericht der DMV* 8, 180-183.
1900a Mathematische Probleme; *Nachrichten der Königlichen Gesellschaft der Wissenschaften*, 253-297.

1905 Über die Grundlagen der Logik und der Arithmetik; in *Verhandlungen des Dritten Internationalen Mathematiker Kongresses*, Teubner, 174-185.
1918 Axiomatisches Denken; *Mathematische Annalen* 78, 405-415.
1927 Die Grundlagen der Mathematik; *Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universität* (6), 65–85.

Hilbert, D. and Bernays, P.
1917-18 *Prinzipien der Mathematik*; in W. B. Ewald and W. Sieg (editors) *David Hilbert's Lectures on the Foundations of Mathematics and Physics, 1917-1933*, Springer, 2013, 64-214.
1921-22 *Grundlagen der Mathematik*; in W. B. Ewald and W. Sieg (editors) *David Hilbert's Lectures on the Foundations of Mathematics and Physics, 1917–1933* Springer, 2013, 431-518.
1939 *Grundlagen der Mathematik*, volume II; Springer.

Jaskowski, S.
1934 On the rules of suppositions in formal logic; *Studia Logica* 1, 4-32.

MacLane, S.
1934 *Abgekürzte Beweise im Logikkalkul*; Dissertation, Göttingen.
1935 A logical analysis of mathematical structure; *The Monist* 45, 118-130.
1979 A late return to a thesis in logic. In I. Kaplansky (editor), *Saunders MacLane — Selected Papers*, Springer, 63–66.

Poincaré, H.
1902 Review of (Hilbert 1899), *Bulletin des sciences mathématiques* 26, 249–272.

Prawitz, D.
1965 *Natural deduction – A proof-theoretical study*; Almqvist & Wiksell.

Sieg, W.
1978 *Elementary proof theory*; Technical Report 297, Institute for Mathematical Studies in the Social Sciences, Stanford 1978, 104 pp.
1992 *Mechanisms and Search – Aspects of Proof Theory*, Vol. 14. Associazione Italiana di Logica e sue Applicazioni.
2007 The AProS Project: Strategic thinking & computational logic; *Logic Journal of the IGPL* 15 (4), 359-368.
2009 Review of (von Plato 2008), *Mathematical Reviews* 2413304.
2010 Searching for proofs (and uncovering capacities of the mathematical mind); In S. Feferman and W. Sieg (editors), *Proof, Categories and Computations*, College Publications:189-215.
2014 The ways of Hilbert's axiomatics: structural and formal; *Perspectives on Science* 22(1), 133-157.
2016 On Tait on Kant and Finitism. *Journal of Philosophy* 112(5– 6), 274– 285.
2018 What is the *concept* of computation? In F. Manea, R.G. Millere, and D. Nowotka (editors), *Sailing routes in the world of computation*, CiE 2018, Lecture Notes in Theoretical Computer Science and General Issues 10936, Springer, 386-396.
2019 The Cantor–Bernstein theorem: How many proofs?; *Philosophical Transactions of the Royal Society A*, 377 (2140), 20180031.

Sieg, W. and Byrnes, J.
1998 Normal natural deduction proofs (in classical logic); *Studia Logica* 60, 67-106.

Sieg, W. and Cittadini, S.
2005 Normal natural deduction proofs (in non-classical logics); in D. Hutter and W. Stephan (editors) *Mechanizing mathematical reasoning* Lecture Notes in Computer Science 2605, Springer, 169-191.

Sieg, W. and Derakhshan, F.
2020 Human-centered automated proof search; manuscript, submitted, 46 pp.

Sieg, W. and Field, C.
2005 Automated search for Gödel's proofs, *Annals of Pure and Applied Mathematics* 133, 319-338.

Sieg, W., Lindstrom, I. and Lindstrom, S.
1981 Gödel's incompleteness theorems - a computer-based course in elementary proof theory; in P. Suppes (editor) *University-Level Computer-Assisted Instruction at Stanford 1968-80*, Stanford, 183–193

Sieg, W. and Morris, R.
2018 Dedekind's structuralism: creating concepts and deriving theorems; in E. Reck (editor), *Logic, Philosophy of Mathematics, and their History: Essays in Honor of W.W. Tait*, College Publication, 251-301.

Sieg, W. and P. Walsh
2019 Natural formalization: deriving the Cantor-Bernstein Theorem in ZF; *Review of Symbolic Logic*, 1-35; doi:10.1017/S175502031900056X.

Thiele, R.
2003 Hilbert's twenty-fourth problem; *American Mathematical Monthly* 110, 1-24.

Von Plato, J.
2008 Gentzen's proof of normalization for natural deduction; *Bulletin of Symbolic Logic* 14, 240-257.

Wang, H.
1981 Some fact s about Kurt Gödel; *Journal of Symbolic Logic* 46, 653-659.

Zermelo, E.
1908 Untersuchungen über die Grundlagen der Mengenlehre I; *Mathematische Annalen* 65 (2), 261–281.