Highlights in Proof Theory

Solomon Feferman*

Abstract

This is a survey of some of the principal developments in proof theory from its inception in the 1920s, at the hands of David Hilbert, up to the 1960s. Hilbert's aim was to use this as a tool in his finitary consistency program to eliminate the "actual infinite" in mathematics from proofs of purely finitary statements. One of the main approaches that turned out to be the most useful in pursuit of this program was that due to Gerhard Gentzen, in the 1930s, via his calculi of "sequents" and his Cut-Elimination Theorem for them. Following that we trace how and why prima facie infinitary concepts, such as ordinals, and infinitary methods, such as the use of infinitely long proofs, gradually came to dominate proof-theoretical developments.

In this first lecture I will give an overview of the developments in proof theory since Hilbert's initiative in establishing the subject in the 1920s. For this purpose I am following the first part of a series of expository lectures that I gave for the Logic Colloquium '94 held in Clermont-Ferrand 21-23 July 1994, but haven't published. The theme of my lectures there was that although Hilbert established his theory of proofs as a part of his foundational program and, for philosophical reasons which we shall get into, aimed to have it developed in a completely finitistic way, the actual work in proof theory

*This is the first of three lectures that I delivered at the conference, Proof Theory: History and Philosophical Significance, held at the University of Roskilde, Denmark, Oct. 31–Nov. 1, 1997. I wish to thank the organizers, Prof. Stig Andur Pedersen and Prof. Vincent F. Hendricks for inviting me to present these lectures, and especially for their substantial work in preparing them for publication with the assistance of Mr. Klaus Frovin Jorgenson, from their audio tapes and my transparencies. I have edited the resulting articles for coherence and readability, but otherwise have maintained their character as lectures, as originally given.

has moved steadily away from that towards rather infinitary methods. So the question posed in the title of my Clermont-Ferrand lectures was "how did this happen, how is it that finitary proof theory became infinitary?". We will get some idea of this. But what I want to concentrate on here are some of the technical aspects of that work, to try to give you an idea of what are the notions involved in the actual development of proof theory, and what are some of the main results, at least up to the 1960s. The references below should be consulted for more detailed technical expositions that bring the subject up to the present, and for a more complete history of its development.

1. Review of Hilbert's Program and finitary proof theory. Let us start with Hilbert's program itself and his conception of it. He was very concerned foundationally about the problems of the infinite in mathematics. Those were of two kinds. On the one hand you had explicit uses of the infinite in Cantorian set theory, that featured in some sense the completed infinite, the transfinite. Then he also saw as a problem implicit uses of the infinite, already in classical number theory with the use of first order predicate calculus and ordinary classical logic where, by the law of the excluded middle, one assumes the natural numbers as in some sense a completed totality. For, one would have to be able to decide between alternatives of the form: either all integers have a property R or there is some integer which fails to have the property R, i.e.

$$\forall x R(x) \vee \exists x \neg R(x). \tag{1}$$

In general you would not be able to decide this. That kind of reasoning is essential in various number-theoretical arguments. In order to eliminate the infinite in mathematics, Hilbert's idea was that somehow the actual infinite should only function as an ideal element, and is to be eliminated in favor of the potential infinite. So finitism, in his conception of it, is to be the mathematics of the potential infinite.

We have a contrast here between finitary (or finitistic) methods and infinitary (or infinitistic) methods. Paradigmatic for the former is the establishment of universal propositions of the form: that each natural number, x, has an effectively decidable property $R(x)$ (or even more basically, a primitive recursive property). A finitistic proof of this has the character that for each individual natural number x you can establish R at x, just by running though a finite portion of the integers. In the proof itself one might have to go far beyond x in order to establish that, but, so to speak, the finitary character of the statement that R holds of all natural numbers x is that at each individual natural number you will only need a finite portion of the natural numbers to verify the statement. If A on the other hand were a proposition of the form $\forall x R(x)$, i.e. that all integers have a certain property, there would be no way in principle in which that could be verified by finite means as a whole.

How then was proof theory to be used as a tool in Hilbert's program? How was proof theory to be used in order to eliminate the actual infinite in systems like number theory, as reflected in propositions like (1)—to reduce proofs in systems which involve implicitly or explicitly the actual infinite, to proofs of finitary, or apparently finitary, propositions, and to reduce these proofs to actual finitary verifications? The idea was: you would show finitistically that if you have a formal system S in which a body of mathematics is represented, and if S proves a finitarily meaningful proposition $R(x)$, then you would want to be able to show that at each natural number, x, $R(x)$ holds. Formally that takes the form: If u is a proof in S of $R(x)$ then $R(x)$ holds, i.e.

$$Proof_S(u, \ulcorner R(x) \urcorner) \rightarrow R(x) \tag{2}$$

In order to do this it turns out that for a system S which is able to establish elementary facts about primitive recursive functions and relations, it is sufficient to establish the consistency of S in a finitistic way. For suppose you have a proof of $R(x)$ in S (with free variable x) but $R(x)$ does not hold (for a specific x). Then within S you could prove that $R(x)$ does not hold (at that x) and therefore you would have a contradiction; if you assumed consistency this could not happen. That is why Hilbert's program concentrated on the consistency problem. The essence of the program was just this so-called reflection principle (2) for finitary statements: that if you have a proof of $R(x)$ in the system then in fact $R(x)$ holds. The elimination of the actual or potential infinite in S would be accomplished if the reflection principle could be established and, for that, if the consistency of S, $Con(S)$, could be proved finitistically. In free variable form, $Con(S)$ is the formula

$$\neg Proof_S(u, \ulcorner 0 = 1 \urcorner) \tag{3}$$

and it is thus a candidate for a finitistic proof.

There are a number of methods which were developed in proof theory in order to be able to establish this program for the elimination of the actual infinite (explicit or implicit) in mathematics, and I just want to mention several before turning to what has come to be the dominant method. I divide these into two groups, one group having a certain kind of functional character and a second group having a more syntactic character.

Functional Methods. The ones having functional character start with Hilbert's ε-calculus. 'ε' there had nothing to do with the membership relation, but is simply a symbol he used to write $\varepsilon x A(x)$, which is interpreted informally as *an* x such that the property $A(x)$ holds, if there is any such x at all. So you would say there is an x such that $A(x)$ holds if A holds of $\varepsilon x A(x)$, that is, if $A(\varepsilon x A(x))$. You have here a formal elimination of quantifiers in favor of such ε-terms. Then you want to show that ε-terms can be eliminated, and thence reduce predicate calculus to propositional calculus, which is unproblematic from a finitary point of view. I think of these ε-terms as functions, because what they really do is to provide choices of an x, such that $A(x)$ holds; they are functions of the other arguments in $A(x)$. Hilbert himself initiated this approach; he proposed some theorems that ought to be established. Wilhelm Ackermann, his collaborator, continued the work, and it went a certain distance, but then after the 1930s not much was done until William Tait took it up again in the late 1950s and early 1960s. In more recent years it is my colleague Grigori Mints who has really pursued this quite systematically and has extended the method to rather strong systems. That is work which is currently in progress.

Jacques Herbrand had in some respects a related approach, and again the idea was to reduce validity in the predicate calculus in some way to validity in the propositional calculus. Basically what you do is that you introduce functions which are in a sense dual to Skolem functions, which are objects that are appropriate to eliminate quantifiers for satisfiability. Herbrand used functions which are appropriate for eliminating quantifiers for validity. Formally you can reduce questions of validity of arbitrary formulas in the predicate calculus to questions of validity of purely existential formulas by introducing many new function symbols in them. Then Herbrand showed that if an existential formula is provable then some finite disjunction of instances of that formula is propositionally provable; moreover you can derive the original formula, once you have such a finite substitution instance. This approach was initiated by Herbrand around 1930; there were technical problems in Herbrand's own work, which were dealt with in the 1950s by Burton Dreben and his (then) student Warren Goldfarb, among others.

Related work that continues this into number theory was carried on by

Georg Kreisel, under the rubric of "the no-counter-example interpretation". That designation comes from the fact that if a formula is not valid, Skolem functions would give an example for the negation of the formula. So if you say that the formula is derivable, there is no counter-example, and that leads you to the idea of the "no-counter-example interpretation".

Kurt Gödel did something quite different where functionals enter very clearly, in the use of the so-called "Dialectica" (functional) interpretation. He formally reduced classical systems to intuitionistic systems, and then showed that proofs of intuitionistic propositions have a natural functional interpretation which he provided.

Syntactical Methods. Now I want to turn to what are more syntactic or purely logical approaches. In the early 1930s Gerhard Gentzen introduced two kinds of calculi: First, the Calculi of Natural Deduction, which are reasonably close in formal terms to the actual way in which we carry out proofs. He tried to use this kind of calculus in order to carry out Hilbert's program. But there were some technical problems there, and he put it aside in favor of what are called Calculi of Sequents, which I will describe in more detail. However, in later years the Natural Deduction Calculi were taken up again by Dag Prawitz, and these have become very important, both in proof theory and in applications of proof theory to computer science.

Gentzen introduced two kinds of Sequent Calculi for the first order predicate calculus: **LJ** (intuitionistic predicate calculus) and **LK** (classical predicate calculus). **L** stands for pure predicate logic, **J** for intuitionistic, **K** for classical. The calculi have rules which are specific to each logical operation, and they separate the use of implication from its use in inferences in a way that we will see in a moment.

2. Results of finitary proof theory via Gentzen's L-Calculi.

Let $\Gamma, \Delta, \cdots$ be finite sequences (or multi-sets, or sets—different people take different choices) of formulas. Instead of using set notation we write $A_1, \cdots, A_n$ for the sequence (multi-set or set), i.e. $\Gamma = A_1, \cdots, A_n$. Γ, A simply means adjoin A to the sequence Γ, i.e. $\Gamma, A = A_1, \cdots, A_n, A$. Gentzen considers derivations of formal expressions of the form $\Gamma \vdash A$ where the sequence Γ represents the hypotheses or premises and A the conclusion of an argument. An expression of the form $\Gamma \vdash A$ is called a sequent. Gentzen also allows that there may be an empty conclusion, that is, expressions of the form $\Gamma \vdash$. Instead of thinking of $\Gamma \vdash$ as an argument without conclusion, you can think of it as an argument with a contradiction as conclusion by, for instance, putting as conclusion a formula for a contradiction, like $0 = 1$ or $A \wedge \neg A$ for some specific formula A. The character of these rules is that they tell you how something is to be introduced in a conclusion and they also tell you how something is to be used as a hypothesis. So in each of these rules you have a left rule and a right rule. They either introduce a formula with a specific principal logical operator as conclusion—those are the right rules—or they introduce in a similar way a formula in the hypotheses (or antecedent of the sequent)—those are the left rules.

The calculus **LJ** has the following formal rules

- Rules for logical operations

	Right	Left
$\neg$	$\dfrac{\Gamma, A \vdash}{\Gamma \vdash \neg A}$	$\dfrac{\Gamma \vdash A}{\Gamma, \neg A \vdash}$
$\rightarrow$	$\dfrac{\Gamma, A \vdash B}{\Gamma \vdash A \rightarrow B}$	$\dfrac{\Gamma \vdash A \quad \Gamma, B \vdash C}{\Gamma, A \rightarrow B \vdash C}$
$\vee$	$\dfrac{\Gamma \vdash A}{\Gamma \vdash A \vee B}$ $\dfrac{\Gamma \vdash B}{\Gamma \vdash A \vee B}$	$\dfrac{\Gamma, A \vdash C \quad \Gamma, B \vdash C}{\Gamma, A \vee B \vdash C}$
$\wedge$	$\dfrac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B}$	$\dfrac{\Gamma, A \vdash C}{\Gamma, A \wedge B \vdash C}$ $\dfrac{\Gamma, B \vdash C}{\Gamma, A \wedge B \vdash C}$
$\exists$	$\dfrac{\Gamma, A(x) \vdash B}{\Gamma, \exists x A(x) \vdash B}$ restriction on x	$\dfrac{\Gamma \vdash A(t)}{\Gamma \vdash \exists x A(x)}$
$\forall$	$\dfrac{\Gamma, A(t) \vdash B}{\Gamma, \forall x A(x) \vdash B}$	$\dfrac{\Gamma \vdash A(x)}{\Gamma \vdash \forall x A(x)}$ restriction on x

- Cut rule

$$\frac{\Gamma \vdash A \qquad \Gamma', A \vdash B}{\Gamma, \Gamma' \vdash B}$$

- Structural rules

$$\frac{\Gamma \vdash A}{\Gamma' \vdash A}, \quad \Gamma \subseteq \Gamma'$$

Let's first of all look at disjunction, which is very characteristic. You can infer $A \vee B$ as the conclusion if you have either inferred A or you have inferred B. To use $A \vee B$ you argue as follows: if I knew A and I obtain C, that would be one way, if I knew B and I obtain C that would be another way. I do not know which of A or B holds, but if I have $A \vee B$, either A then C or B then C, and therefore we can conclude C. The rules for conjunction are dual to those for disjunction.

With negation it goes as follows. To infer that Γ establishes $\neg A$ you need to know that Γ together with A reaches a contradiction. On the other hand to establish that Γ together with $\neg A$ reaches a contradiction you establish that A follows from Γ.

Now, perhaps the most natural rule, although they all are quite natural, is the one with implication-introduction on the right. In order to infer $A \rightarrow B$ from Γ you simply say: let us take A as an additional hypothesis with Γ, use it and then infer B.

For quantification, let us just look at existential quantification: in order to infer $\exists x A(x)$ you infer that Γ establishes $A(t)$ for a specific t. In order to establish that Γ together with $\exists x A(x)$ has as consequence B you say: suppose $\exists x A(x)$, let x be anything with that property and show that B is a consequence of Γ together with $A(x)$. We need x there to be a fresh variable that has not been named otherwise in the sequent involved; that is the restriction on x in the quantifier rules.

There are some axioms and rules besides these that are not specific to the connectives. Axioms are simply of the form $A \vdash A$, i.e. that from A, A follows; that is clear. The structural rules say that the set of premises in an argument can be expanded: if from a set of hypotheses Γ you have obtained A and you take any larger set Γ' that also gives A.

The most significant rule here is the Cut Rule which says: if from Γ you have inferred A and from another Γ' with A you have inferred B then from both Γ and Γ' you can infer B. This rule has some of the character of Modus Ponens, but you can also think of it as a rule which corresponds to the use of lemmas or theorems in the process of a derivation: in that process you establish some particular lemma or theorem A as an intermediate step and then by use of A you establish the final conclusion B.

The Cut Rule differs from the other rules in an important respect. With the rules for introduction of a connective on the left or the right, you see that every formula that occurs above the line occurs below the line either directly, or as a sub-formula of a formula below the line, and that is also true for the structural rules. (We count $A(t)$ as a subformula, in a slightly extended sense, of both $\exists x A(x)$ and $\forall x A(x)$.) But in the case of the Cut Rule, the cut formula A disappears. So you have a kind of detour, perhaps, through a formula which may be more complicated than formulas which occur in the final sequent.

Now, a crucial question would be: under what conditions can we establish a direct derivation from a set of hypotheses, Γ, to a conclusion, A, without use of the Cut Rule, because the Cut Rule means that we are making some detours?

Before we get into this; what we have described so far is the Gentzen calculus for intuitionistic logic. The Gentzen calculus for classical logic is formally similar, except that on the right hand side we may also have a finite set or sequence of formulas $\Delta (= B_1, \ldots, B_n)$, and the rules look exactly the same as they did with the intuitionistic calculus allowing these sets of side formulas, Δ, throughout.

- The calculus LK has the following formal rules:
- Rules for logical operations

	Right	Left
$\neg$	$\dfrac{\Gamma, A \vdash \Delta}{\Gamma \vdash \neg A, \Delta}$	$\dfrac{\Gamma \vdash A, \Delta}{\Gamma, \neg A \vdash \Delta}$
$\rightarrow$	$\dfrac{\Gamma, A \vdash B, \Delta}{\Gamma \vdash A \rightarrow B, \Delta}$	$\dfrac{\Gamma \vdash A, \Delta \quad \Lambda, B \vdash C, \Theta}{\Gamma, \Lambda, A \rightarrow B \vdash C, \Delta, \Theta}$
$\vee$	$\dfrac{\Gamma \vdash A, \Delta}{\Gamma \vdash A \vee B, \Delta}$ $\dfrac{\Gamma \vdash B, \Delta}{\Gamma \vdash A \vee B, \Delta}$	$\dfrac{\Gamma, A \vdash C, \Delta \quad \Gamma, B \vdash C, \Delta}{\Gamma, A \vee B \vdash C, \Delta}$
$\wedge$	$\dfrac{\Gamma \vdash A, \Delta \quad \Gamma \vdash B, \Delta}{\Gamma \vdash A \wedge B, \Delta}$	$\dfrac{\Gamma, A \vdash C, \Delta}{\Gamma, A \wedge B \vdash C, \Delta}$ $\dfrac{\Gamma, B \vdash C, \Delta}{\Gamma, A \wedge B \vdash C, \Delta}$
$\exists$	$\dfrac{\Gamma, A(x) \vdash B, \Delta}{\Gamma, \exists x A(x) \vdash B, \Delta}$ restriction on x	$\dfrac{\Gamma \vdash A(t), \Delta}{\Gamma \vdash \exists x A(x), \Delta}$
$\forall$	$\dfrac{\Gamma, A(t) \vdash B, \Delta}{\Gamma, \forall x A(x) \vdash B, \Delta}$	$\dfrac{\Gamma \vdash A(x), \Delta}{\Gamma \vdash \forall x A(x), \Delta}$ restriction on x

- Cut rule

$$\frac{\Gamma \vdash A, \Delta \qquad \Gamma', A \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta'}$$

- Structural rules

$$\frac{\Gamma \vdash \Delta}{\Gamma' \vdash \Delta'}, \quad \Gamma \subseteq \Gamma', \Delta \subseteq \Delta'$$

In the classical case, Cut informally takes the form: if from Γ you have obtained Δ, A and from Γ', A you get Δ' then from Γ, Γ' you obtain Δ, Δ'. In a moment we will look at an interpretation of **LK** in more usual Hilbert style terms, but to see how this form of the calculus gives us something that the intuitionistic calculus does not give, let us just look at a proof of the Law of the Excluded Middle in **LK**. A formal proof of this law takes the following form:

$$\begin{array}{ll} A & \vdash A \\ & \vdash A, \neg A \\ & \vdash A \vee \neg A, \neg A \\ & \vdash A \vee \neg A, A \vee \neg A \\ & \vdash A \vee \neg A \end{array}$$

From the axiom $A \vdash A$, you can bring negation over to the right hand side and then, using the disjunction rule applied to the first formula A on the right hand side you get $A \vee \neg A$, and then using the disjunction rule once more for introducing disjunction on the right you again obtain $A \vee \neg A$. Now the set of formulas to the right of $\vdash$ just consists of two exemplars of $A \vee \neg A$ which simply collapses to the set $A \vee \neg A$. From the way this argument goes we do not know why $A \vee \neg A$ holds, i.e. we do not know which of $A, \neg A$ is true, but by the mechanics of this form of Gentzen's classical calculus we are able to derive it in that form. Note that we are blocked at the very outset from carrying out this derivation in **LJ**, where at most one formula appears on the right hand side of a sequent. In **LJ** if $\vdash A \vee B$ then $\vdash A$ or $\vdash B$, but this doesn't hold in **LK**. In particular $\vdash A \vee \neg A$ in **LK** even when neither $\vdash A$ nor $\vdash \neg A$.

The interpretation of **LK** is simply that a sequent $\Gamma \vdash \Delta$ is derivable in that calculus just in case the conjunction of formulas in Γ implies the disjunction of formulas in Δ, i.e.

$$A_1 \wedge \cdots \wedge A_n \rightarrow B_1 \vee \cdots \vee B_m \tag{4}$$

is valid. If you look back at the rules with this interpretation in mind you see that each of the rules preserves validity in that sense.

So the Cut Rule is a kind of generalization of Modus Ponens and takes over its roles. The principal theorem that Gentzen obtained is the Cut-Elimination Theorem, and this works both for the classical and the intuitionistic calculi. It says that for each derivation d of a sequent $\Gamma \vdash \Delta$ we can associate a cut-free derivation d^* of the same sequent, i.e. one which has no applications of the cut rule in it. Everything that is used in it is either an axiom, a structural rule or one of the rules for introducing one of the connectives either on the left or on the right. There is a price, though, that you have to pay for this, and that is that the length of the cut-free derivation associated with the original derivation is much longer in the worst case, and it can grow at a hyper-exponential rate, which can be measured as follows: Associated with each derivation is its cut-rank, which is the maximum of the complexities of the cut-formulas A which are eliminated by the use of the cut rule. You have a natural measure of the complexity of those formulas and you also have a measure of the length of the derivation $|d|$ and the length of the new derivation $|d^*|$. The length of the cut-free derivation, compared to the length of the original derivation, is bounded by a stack of 2's above which the highest exponent is $|d|$, and the length of the stack is the cut-rank of the original derivation. In symbols: $|d^*| \leq 2_r(|d|)$ where $r = \text{cut-rank}(d)$, $2_0(a) = a$ and $2_{n+1}(a) = 2^{2_n(a)}$. Basically each reduction of the cut-rank by 1 corresponds to an increase in the length of the modified derivation, by exponent to the base 2. By suitable examples this is in general the best possible; you can't do much better than this.

Though not feasible operations in general, because of the possible hyper-exponential rate of growth, these are effective transformations of the original derivations into new cut-free derivations. The cut-free derivations are significant, as I say, because they have the sub-formula property. In particular, if you ask whether you can derive the empty sequent, which would simply be a derivation of a contradiction, the answer is no, because by the sub-formula property everything in such a derivation would have to be a sub-formula of this eventual conclusion, but you have to start with axioms so that can't be possible. So this proves that the classical predicate calculus is consistent, which is not surprising—but this is a simple example of how the Cut-Elimination Theorem might be useful in establishing consistency of stronger systems S.

Applications. As with the disjunction property for **LJ**, that if $\vdash A \vee B$ then $\vdash A$ or $\vdash B$, we have an existential instantiation property for **LJ**: if $\vdash \exists x A(x)$, then for some t, $\vdash A(t)$. This cannot be done in **LK**, even for quantifier-free A. Herbrand's Theorem, which I have already mentioned, tells us the next best thing we can do in **LK**. It runs as follows:

> If R is a quantifier-free formula and $\vdash \exists x R(x)$ in **LK** then there exist terms $t_1, \cdots, t_n$ such that $\vdash R(t_1) \vee \cdots \vee R(t_n)$. And, more generally, when Γ is purely universal, if in **LK**, $\Gamma \vdash \exists x R(x)$ then $\Gamma \vdash R(t_1) \vee \cdots \vee R(t_n)$ for some $t_1, \ldots, t_n$.

A proof of Herbrand's Theorem using **LK** goes along the following lines: If you have a derivation d of an existential statement, $\vdash \exists x R(x)$, where R is quantifier free, then by the Cut-Elimination Theorem we can transform it into a cut-free derivation d^* of $\vdash \exists x R(x)$. Now in d^*, this final sequent would have had to have been established by the right rule for existential quantification, which would mean that you had a substitution instance $R(t)$ which brought that in. This $\exists x R(x)$ might still have been there, because by the structural rules the set of formulas would collapse if I proved $\exists x R(x)$, $R(t)$ and then established $\exists x R(x)$, $\exists x R(x)$, and since the set of those are the same as $\exists x R(x)$ I would still have that formula $\exists x R(x)$ one step back. So we would continue up the tree of that derivation in that way, each time having perhaps some new substitution instance. But eventually we have to stop with that, and we will simply have a bunch of formulas on the right of the form: $R(t_1), R(t_2), \ldots, R(t_n)$, and of course that is the same as having a proof of the disjunction $R(t_1) \vee R(t_2) \vee \ldots \vee R(t_n)$.

What Herbrand's Theorem tells us—and what comes out of the proof in the Gentzen calculus—is that: classically, if you have a proof of an existential statement, you do not necessarily know of one specific instance which realizes that statement, but you will always have a finite set of instances, by means of which you can say at least one of those realizes the statement. More generally, if you have a purely universal set of hypotheses and that proves an existential statement then there will be a finite set of witnesses which proves the corresponding disjunction. One way of seeing that is: take the universal statements in Γ, use the negation rule to bring them over to the right side, they then become existential statements, and we can apply Herbrand's Theorem and then go back and we will have the conclusion.

The reason this is useful for Hilbert's Program is that some formal systems of interest to us have particularly simple axioms which are purely universal.

In particular, in Peano Arithmetic the axioms for 0, successor, addition, multiplication, and perhaps other primitive recursive functions are all universal. The only thing which might not be universal is the Axiom of Induction, or the scheme of induction with various formulas. Let us take the simplest form of the scheme of induction, namely, Quantifier Free Induction Axiom (**QF-IA**)

$$R(0) \wedge \forall x[R(x) \rightarrow R(x')] \rightarrow \forall x R(x) \tag{5}$$

In this, $R(x)$ is a quantifier-free formula, with perhaps additional free variables. It expresses that if $R(0)$ holds and if $R(x)$ implies $R(x')$ for any x, then $\forall x R(x)$ holds.

Assuming you have some elementary properties of primitive recursive functions and relations—in particular the "less than" relation, then you can show that (5) is equivalent to the statement

$$\forall x[R(0) \wedge \forall y < x[R(y) \rightarrow R(y')] \rightarrow R(x)] \tag{6}$$

which is purely universal with primitive recursive body. To see this, it is sufficient to know that the induction hypothesis—that R transmits from y to y'—holds for numbers under x in order to get up to x itself. In primitive recursive arithmetic the bounded quantifier formula $\forall y < x[R(y) \rightarrow R(y')]$ is equivalent to a quantifier-free formula. So, the universal closure of (6) now becomes a purely universal statement. Therefore, if you are looking at consequences of that particular subsystem of Peano Arithmetic which just uses **QF-IA**, by Gentzen's Cut-Elimination Theorem (or if one prefers, Herbrand's Theorem, or suitable theorems for Hilbert's ε-calculus), you are able to establish that if you prove an existential statement in that system you will have a disjunction of a finite number of instances provable there. In particular, the consistency of this system is an immediate consequence.

That result for **QF-IA** was established by Ackermann as the first contribution to Hilbert's consistency program for a system of any mathematical interest. Though there is not very much you can do mathematically within that system, it is non-trivial at the same time. Much later there was an extension of this result to the instances of the induction axiom scheme which use what are called Σ_1^0 formulas denoted ($\mathbf{\Sigma_1^0 - IA}$):

$$A(0) \wedge \forall x[A(x) \rightarrow A(x')] \rightarrow \forall x A(x), \tag{7}$$

where A is a Σ_1^0 formula, i.e. it is of the form $\exists x R(x)$, where R itself is quantifier free. (The super '0' simply means you have just numerical quantification, the sub '1' means one quantifier and the Σ means that it is existential.) The result for $\mathbf{\Sigma_1^0 - IA}$ was first proved by Charles Parsons using an adaptation of Gödel's functional interpretation. Wilfried Sieg later gave a new proof, using Herbrand-Gentzen style methods, which happens to be useful for other things.

It turns out from the work of Parsons that the system based on the Σ_1^0 Induction-Axiom is conservative over the system of quantifier-free Primitive Recursive Arithmetic, **PRA**. Being conservative means that any formula formulated in the language of **PRA**, which is provable with Σ_1^0 induction ($\mathbf{\Sigma_1^0 - IA}$), is already provable in **PRA**.

PRA itself is a system based on entirely quantifier-free axioms and rules for primitive recursive functions, including a Rule of Induction rather than an Axiom of Induction. It has been argued by Tait, and is generally agreed, that everything that is obtained in **PRA** is finitistically justifiable, at least in principle (Tait has also argued the converse). Assuming this, Parsons' result is, therefore, again a contribution to Hilbert's program: you eliminate the use of the infinite as reflected in the use of classical predicate calculus, together with the axioms for this fragment of arithmetic $\mathbf{\Sigma_1^0 - IA}$ in favor of purely finitary proofs, as represented in **PRA**.

If you are a radical finitist you only talk about things that are feasibly computable, and this result would not cover that because the primitive recursive functions, beginning with exponentiation, go far beyond the feasibly computable functions. But if you are finitist "in principle" then conservativity over **PRA** should certainly satisfy you.

3. Shifting paradigms. How far can Hilbert's program be carried out? Gödel's second incompleteness theorem told us that one would not be able to prove the consistency of full first order arithmetic, **PA** (whose induction axiom scheme applies to arbitrary formulas) by means which could be formalized within **PA** itself. Although Hilbert never made precise what exactly finitist arguments were to consist in, all the finitary arguments that had been carried out up to Gödel's incompleteness theorem were evidently formalizable in very weak subsystems of **PA**, and in fact in **PRA** itself. Thus, one had to rethink the Hilbert program at that point, and ask whether there is some reasonable modification of it which could establish the consistency of **PA** and yet stronger systems. For example, instead of reducing infinitary systems to finitary systems, one could try to reduce non-constructive systems to constructive systems, or, as we will see, even seek other kinds of reductions.

If one accepts intuitionistic logic as being a formal expression of constructive ideas, then one could say: suppose we replace classical logic in **PA**, which is behind the first implicit use of the actual infinite (as in, e.g. $\forall x R(x) \vee \exists x \neg R(x)$, which cannot be inferred in intuitionistic logic). Suppose we leave out the law of the excluded middle; then we obtain a system which is called Heyting Arithmetic, **HA**, which looks just like **PA** with the difference that it is based on intuitionistic rather than classical logic. But there is a simple translation of classical logic into intuitionistic logic and of classical arithmetic into intuitionistic arithmetic—obtained independently by Gödel and Gentzen. It is not clear that Hilbert would have been satisfied with this reduction in favor of a system where there is no implicit appeal to the actual infinite. But this is evidence of a kind of thing that could be done if you replaced Hilbert's program by this modification where you say: let us just see what can be obtained by reducing non-constructive formal systems involving appearances of the infinite into constructive formal systems which do not contain such appearances.

But to carry out something like Hilbert's original program for the full system of arithmetic, **PA**, more must be done if you are going to try to push a proof of its consistency as far down to finitist arguments as possible. That again was accomplished by Gentzen, who brought elements of the transfinite into the picture, with the use of ordinals. The ordinals here are those which are below Cantor's ordinal ε_0. That is the limit of the sequence of ordinals

$$\omega,\ \omega^{\omega},\ \omega^{\omega^{\omega}}\ \cdots$$

These ordinals can be represented in finite form, in what is called Cantor-normal form to base ω, and when represented symbolically in that form, these simply look like certain finite symbolic configurations

$$\omega^{\alpha_m} + \cdots + \omega^{\alpha_0}, \tag{8}$$

where the exponents are again of such forms. The ordering relation between these finite symbolic configurations turns out to be primitive recursive, in a very easy way: you can decide whether one configuration, representing an ordinal less than ε_0, is less than another in a primitive recursive way. Consequently, you could say: though I am talking about certain transfinite objects here—transfinite ordinals—I am representing them in a finitary way, using this primitive recursive relation.

What Gentzen did was to associate with each derivation in elementary number theory a derivation of sequents using the induction rule (**Ind-Rule)**

$$\frac{\Gamma \vdash A(0), \Delta \qquad \Gamma, A(x) \vdash \Delta, A(x')}{\Gamma \vdash A(x), \Delta} \tag{9}$$

(A arbitrary) to supplement the logical rules of **LK**, to see whether something like the Cut-Elimination Theorem could hold.

It turns out that we do not have full cut-elimination for this extension of **LK**. But what we do have by Gentzen's work is that if you have a derivation, d, which is a possible derivation of the empty sequent, then with each such derivation you can assign an ordinal $ord(d)$ less than ε_0, such that certain reductions like cut-elimination can be applied to that derivation to lower the complexity. If a derivation were a derivation of the empty sequent you would be able to successively lower its complexity. That is, to each derivation d of the empty sequent is associated another one d', such that $ord(d') < ord(d) < \varepsilon_0$. But if that were the case, then you would have a descending sequence in the above-mentioned order relation of order type ε_0. So, if you assume that that is a well-founded relation, or equivalently that you can apply transfinite induction up to the ordinal ε_0, then you can verify that there cannot be any such descending sequence of ordinals, and, therefore, there cannot be such a reduction sequence: so you cannot have a derivation of the empty sequent. Consequently, we have a consistency proof of the full first order Peano Arithmetic, **PA**[1].

What principle of transfinite induction do we need here? It turns out that we only need transfinite induction applied to quantifier-free formulas, and that the rest of the argument can be carried out with just things that are purely within primitive recursive arithmetic **PRA**. The statement that **PA** itself is consistent is a statement of purely universal character, it says no derivation d is a proof of $0 \neq 1$. All of this can be done in a purely, so to speak, finitary way except for the assumption of quantifier free transfinite induction up to ε_0. Quantifier-free transfinite induction up to ε_0 proves the consistency of **PA**,

$$\mathbf{QF\text{–}TI}(\varepsilon_0) \vdash \mathbf{Con}(\mathbf{PA}) \tag{10}$$

finitistically (and certainly over **PRA**). Gentzen showed that this was the best possible in the following sense. For each ordinal α less than ε_0 you can

[1]Gentzen's work actually establishes something stronger, namely the reflection principle for Σ^0_1-formulas $\exists x R(x)$ in **PA**.

prove transfinite induction up to α, $\mathbf{TI}(\alpha)$, in $\mathbf{PA}$:

$$\mathbf{PA} \vdash \mathbf{TI}(\alpha) \quad \text{for each } \alpha < \varepsilon_0. \tag{11}$$

In some sense, ε_0 was thus attached as the ordinal of Peano arithmetic.

Further proposed modifications of Hilbert's program. That is one way in which Hilbert's program was extended. Much further work in this direction was carried on in the 1950s by the Munich school of Kurt Schütte and in the school of Gaisi Takeuti. Though the work of these schools had substantial technical differences, they agreed on a general extension of Hilbert's consistency program. The principal aim of people within this program has been to prove the consistency of stronger and stronger formal systems, S. And the way that is to be done is to associate an ordinal, α, which can be represented primitive recursively, with S, and then prove two things. First, that by just using finitary methods and transfinite induction up to α you are able to prove the consistency of S; and, second, that this is best possible in the sense that for each ordinal smaller than α, you can prove the transfinite induction principle up to β, $\mathbf{TI}(\beta)$, for $\beta < \alpha$, in S itself. But also the transfinite induction principle up to α itself must somehow be recognized in some constructive way. So you have a kind of curious combination here of getting larger and larger ordinals attached to formal systems; you try to be as finitary as possible, but there is this one non-finitary element, the transfinite induction up to the ordinal associated with the system. You may have to use very strong constructive methods in order to establish that transfinite induction principle.

A quite different way of looking at what proof theory ought to do, was proposed by Kreisel and continued by me in the article "Hilbert's program relativized" (see the References). Instead of saying, as in the initial Hilbert's program, that what we are trying to do is reduce infinitary systems to finitary systems, let us say: well, one thing we can do is to perform various other kinds of reductions. For instance, reduce

- non-constructive systems to constructive systems
- impredicative systems to predicative systems,

and so on, and then obtain related conservation results. There the reduction does not mention ordinals at all, but the proof of the reduction might actually have to use the techniques of Gentzen and the extensions by Schütte, Takeuti and others behind the scenes in order to establish such conservation results. That is still another direction that one might hope to pursue with extensions of proof theory. This is something I will talk more about in the following lectures.

Kreisel himself also promoted what you might call a non-ideological attempt at using proof theory as a kind of extension of Herbrand's Theorem, dealing with the question: if you prove some existential statements, what kinds of information can you obtain about witnesses to those existential statements? In the case of number theory typically statements of interest are of the form

$$\forall x \exists y A(x, y) \tag{12}$$

that is, for every natural number there exists another natural number having a certain property, A. When proving such statements non-constructively you would like to extract the constructive or recursive content, i.e., you would like to know how y is obtained as a function of x. In the case that A is primitive recursive, y is obtained as a recursive function of x. Then you talk about what are called the provably recursive functions of the system S. One way of thinking of that is that you may have a non-constructive proof of the existence of a recursive many-one relation, but the question is actually to produce that recursive relation, y as a function of x, or as a program which realizes that function. Again, that can be carried out. So, for example, in the case of arithmetic people have produced what are called hierarchies of recursive functions, or sub-recursive hierarchies, starting with the primitive recursive functions, going a little farther with the Ackermann function, which uses a certain bit of ordinal recursion, and then farther up to the ordinal ε_0. Kreisel characterized the provably recursive functions of **PA** in terms of a sub-recursive hierarchy up to ε_0. For stronger systems the associated ordinal also leads you to hierarchies of provably recursive functions. That is another way in which proof theory is used to obtain, in principle useful, mathematical information, although in fact it does not go much beyond producing such hierarchies.

There have been different ways that people have talked about the ordinal of a formal system. One is the Gentzen-Schütte-Takeuti way, the least ordinal that you can use to prove the consistency of the system,

$$\text{the least } \alpha : \mathbf{PRA} + (\mathbf{QF{-}TI}(\alpha)) \vdash Con(S). \tag{13}$$

Another is that it is the least provably recursive well-ordering of the system, i.e.

$$\alpha = sup\{|\prec|: \prec \text{ is a provably recursive well-ordering of } S\}. \qquad (14)$$

Still another is that it is the ordinal of the least hierarchy which cannot be established in the system

$$\alpha = sup\{\beta : \text{ effective transfinite recursion up to } \beta \text{ is justified in } S\}. \quad (15)$$

There are no very good robust definitions of these concepts although they agree in practice. Unfortunately, we do not have a theory that tells us exactly what we are doing when we obtain the ordinal of a formal system, though it is clear that we are doing something of interest.

4. Countably infinitary methods (getting the most out of logic). In the 1950s there was a shift in techniques to the use of infinitary formal and semi-formal systems[2]. Various people were involved in this but the person who promoted it most vigorously was Kurt Schütte in Munich. He had a different way of getting the ordinal associated with Peano Arithmetic out of a direct extension of logic to an infinitary system where, instead of the usual right universal quantifier rule, we use the ω-rule. It says that if you proved $A(\bar{n})$ for each numeral $\bar{n}$, then you are allowed to prove $\forall x A(x)$, where A here is an arbitrary formula. In sequent form it takes the following form

$$\frac{\cdots \Gamma \vdash \Delta, A(\bar{n}) \cdots (n = 0, 1, 2, \ldots)}{\Gamma \vdash \Delta, \forall x A(x)} \qquad (16)$$

The system obtained by adding this rule to **LK** is denoted **LK**(ω).

Derivations here are in general going to be infinite, since you have infinite branching at the ω-rule. But going up along any branch you will eventually come to an axiom. If you invert the derivation trees you can see that these are well-founded because, going down, every branch comes to an end. Being well-founded infinite trees in this sense, they have a natural associated ordinal length, namely: the height of the tree as an ordinal. The ordinal associated with an ω-inference is the sup of the ordinals associated with the hypotheses plus 1, that is, it is the least ordinal greater than all the ordinals associated with the sub-derivations of each instance $\Gamma \vdash \Delta, A(\bar{n})$.

[2]The history here is incomplete. In particular, P. S. Novikov developed cut-elimination for infinitary derivations with ordinal bounds already in the late 1930s-early 1940s, but this was largely unknown in the West for some time; cf. Mints (1991), 387-389.

You can replace the use of the induction axiom in Peano arithmetic by applications of the ω-rule. There is an immediate map of derivations in **PA** into derivations, d, in $\mathbf{LK}(\omega)$ preserving cut-rank. So these derivations in $\mathbf{LK}(\omega)$ have cut-rank less than ω, but, due to the ω-rule, the length of d is now going to be infinite, but not too large, namely, less than $\omega + \omega$.

Now, lowering the cut-rank in d by 1 gives us an exponential cost of $\mid d \mid$ to base ω. That is,

$$d \longmapsto d', \quad \mid d' \mid \leq \omega^{|d|} \tag{17}$$

By repeating this r times, where r is the cut-rank of d, $r = rnk(d)$, we end up with a cut-free derivation d^* of the same sequent, whose length is at most a stack of ω's r times up to the original length of the derivation d, i.e.

$$d \longmapsto d^*, \quad \mid d^* \mid \leq \omega_r(|d|), \tag{18}$$

where $\omega_0(\alpha) = \alpha$ and $w_{n+1}(\alpha) = \omega^{\omega_n(\alpha)}$.

If we do things right you can take base 2 instead of base ω, just as we did in the finitary case, but when we are in ordinals the effect of that is essentially the same, since $2^\omega = \omega$. The first ordinal which is bigger than all those ordinals that we get via (18) is simply ε_0. So, ε_0 falls out of this directly, and now this is a *full* Cut-Elimination Theorem in ω-logic, of derivations with finite cut-rank, whereas in the Gentzen-treatment of **PA** you only had a *partial* Cut-Elimination Theorem of derivations of Σ^0_1-sequents.

Here we have what people regarded as an essential methodological or conceptual improvement, and that is that ordinals here were associated in a canonical way with infinite derivations in contrast to the original work of Gentzen, where the association of ordinals looked rather ad hoc. It turns out in recent work by Wilfried Buchholz that in fact there is a much closer intrinsic connection between the way that Gentzen assigned the ordinal ε_0 to derivations in **PA** and the way that Schütte assigned ordinals to the associated derivations in **LK** (ω), than was realized for many years. In fact, this same connection had been pretty much established back in the mid 1970s by Mints, but it was not well-known. What Buchholz has done is essentially to incorporate the work of Mints and to simplify it in a way now that it can be extended to much stronger systems, so that it turns out that there is a much closer intrinsic connection than had been previously suspected between ordinal assignments in Schütte-style proof theory using infinite derivations and Takeuti-style proof theory using finitary derivations.

The ordinal ε_0 came out as a natural stopping point, because it is the first ordinal closed under exponentiation. If you take the function $\phi_0(\beta) = \omega^\beta$, and if you define ϕ_1 as the function which enumerates the fixed-points of ϕ_0, and continue this procedure into the transfinite, then you get a hierarchy of ordinal functions which is called the Veblen hierarchy ϕ_α of critical functions, defined by:

$$\begin{aligned} \phi_0(\beta) &= \omega^\beta \\ \phi_{\alpha+1} &\text{ enumerates}\{\beta \mid \phi_\alpha(\beta) = \beta\} \\ \phi_\lambda &\text{ enumerates} \bigcap_{\alpha<\lambda} Rng(\phi_\alpha), \text{ when } \lambda \text{ is a limit ordinal.} \end{aligned} \tag{19}$$

So, e.g. ϕ_1 enumerates the ε-numbers, i.e., $\phi_1(\beta) = \varepsilon_\beta$. The ordinals $\phi_\alpha(\beta)$ appear in further extensions into infinitary proof theory, as we shall see next.

The calculus $\mathbf{LK}_{\omega_1,\omega}$. The basic simplification here is due to Tait who said instead of just using the ω-rule—which seems rather special to arithmetic—let us use countably infinite conjunctions and disjunctions[3]. There are natural rules which generalize those for finite disjunctions and conjunctions, and which look exactly like the ones we had before in **LK**, namely

- countable infinite conjunctions $\bigwedge_n A_n$ with the rules

$$\frac{\cdots \Gamma \vdash \Delta, A_n \cdots (n < \omega)}{\Gamma \vdash \Delta, \bigwedge_n A_n} \qquad \frac{\Gamma, A_k \vdash \Delta}{\Gamma, \bigwedge_n A_n \vdash \Delta} \tag{20}$$

- countable infinite disjunctions $\bigvee_n A_n$ with the rules

$$\frac{\Gamma \vdash \Delta, A_k}{\Gamma \vdash \Delta, \bigvee_n A_n} \qquad \frac{\cdots \Gamma, A_n \vdash \Delta \cdots (n < \omega)}{\Gamma, \bigvee_n A_n \vdash \Delta} \tag{21}$$

Let $\mathbf{LK}_{\omega_1,\omega}$ be Gentzen's calculus **LK** extended with these infinitary rules. (The sub 'ω_1' here means that all conjunctions and disjunctions have length less than the first uncountable ordinal ω_1, while the sub 'ω' means that we only allow finite strings of quantifiers at each occurrence.) Tait established a Cut-Elimination Theorem for this language. Derivations might now have

[3] Again, this was anticipated by Novikov; cf. Mints (1991), pp. 387-389.

transfinite cut-rank instead of finite cut-rank; roughly speaking, you can bound the length of a cut-free derivation obtained from an original derivation of cut-rank α and length β in the Veblen hierarchy by iterating α places out in the ϕ hierarchy. That is, for a derivation d with arbitrary cut-rank α and length β, where $\alpha, \beta < \omega_1$, there exists a cut-free derivation d^* of the same conclusion with

$$| d^* | \leq \phi_\alpha(\beta) \tag{22}$$

Actually this is not best possible but this is pretty much the order of ordinal complexity of what we get.

Applications. This gets applied to Ramified Analysis, which is of very great significance in the proof theory of predicativity, and which is something I am going to be talking about a fair amount in my next lecture. Predicative analysis, or ramified analysis, is essentially Gödel's notion of constructibility restricted to sets of natural numbers, where you have sets indexed by ordinals at different levels and basically each level corresponds to sets which are definable using only quantification over sets of lower levels. By $\mathbf{RA}_\alpha$ one means the system using levels up to α, and in $\mathbf{RA}_{<\alpha}$, you use variables only of levels less than α.

Schütte established that the proof theoretical ordinal of $\mathbf{RA}_{<\alpha}$ is $\phi_\alpha(0)$ when $\omega^\alpha = \alpha$. That can be proved quite directly as Tait did by cut-elimination in the infinitary system $\mathbf{LK}_{\omega_1,\omega}$ that I described above. Now, Kreisel had proposed that to explicate the notion of predicativity we use Ramified Analysis, not through arbitrary ordinals, but only through those ordinals which are accessible from below. That is, we only consider autonomously generated $\mathbf{RA}_\alpha$ which means that there is a β, $\beta < \alpha$, in a previously obtained $\mathbf{RA}_\beta$ with

$$\mathbf{RA}_\beta \vdash WO(\prec_\alpha), \tag{23}$$

where $\prec_\alpha$ is a primitive recursive ordering of order type α and $\mathbf{WO}(\prec_\alpha)$ expresses that it is a well ordering. So that is what we call a boot-strapping or autonomy procedure.

Schütte and I, independently, established in 1964 that the least impredicative ordinal is the least ordinal not obtained through the Veblen process. That is called Γ_0 and it is the least ordinal with the property

$$\alpha, \beta < \Gamma_0 \Rightarrow \phi_\alpha(\beta) < \Gamma_0 \tag{24}$$

By (24), one has complete cut-elimination in the infinitary system up to Γ_0.

This is basically where we came to in the 1960s, and is the beginning of what is called predicative proof theory; further extensions make use of prima facie uncountable infinitary derivations and much more complicated ordinal notation systems. For an introduction to predicative proof theory, see the monograph by Wolfram Pohlers in the References below. For a survey of many further extensions, see also, among other sources, his article "Subsystems of set theory and second order number theory" in the *Handbook of Mathematical Logic* edited by Samuel R. Buss. This material gets extremely complicated technically, and is beyond the scope of this lecture to try to explain, even in the roughest terms.

References

TEXTS:

Girard, Jean-Yves, *Proof Theory and Logical Complexity.* Naples, 1987.

Hilbert, David and Bernays, Paul, *Grundlagen der Mathematik*, Vol. I. 2nd edn. Berlin, 1968. *Ibid.* Vol. II, 2nd edn. Berlin, 1970.

Prawitz, Dag, *Natural Deduction.* Stockholm, 1965.

Schütte, Kurt, *Beweistheorie.* Berlin, 1960.

Schütte, Kurt, *Proof Theory.* Berlin, 1977.

Takeuti, Gaisi, *Proof Theory*, 2nd edn. Amsterdam, 1987.

Troelstra, A. S. and Schwichtenberg, Helmut, *Basic Proof Theory.* Cambridge, 1996.

COLLECTIONS:

Aczel, Peter; Simmons, Harold; and Wainer, Stanley (eds.) *Proof Theory*, Cambridge, 1992.

Buss, Samuel R., (ed.) *Handbook of Proof Theory*, Amsterdam, 1998.

Buchholz, Wilfried; Feferman, Solomon; Pohlers, Wolfram; and Sieg, Wilfried, *Iterated Inductive Definitions and Subsytems of Analysis. Recent Proof-theoretical Studies*, Lecture Notes in Mathematics 897 (1981).

Gentzen, Gerhard, *The Collected Papers of Gerhard Gentzen* (M. Szabo, ed.). Amsterdam, 1969.

Mints, Grigori E., *Selected Papers in Proof Theory.* Naples, 1992.

Kino, Akiko; Myhill, John; and Vesley, Richard E. (eds.), *Intuitionism and Proof Theory.* Amsterdam, 1970.

MONOGRAPHS:

Buchholz, Wilfried and Schütte, Kurt, *Proof Theory of Impredicative Systems.* Naples, 1988.

Jäger, Gerhard, *Theories for Admissible Sets. A Unifying approach to Proof Theory.* Naples, 1986.

Pohlers, Wolfram, *Proof Theory. An Introduction.* Lecture Notes in Mathematics 1407, 1989.

RESEARCH AND EXPOSITORY ARTICLES:

Avigad, Jeremy and Feferman, Solomon, "Gödel's functional ('Dialectica') interpretation", in Buss (1998) pp. 337-405.

Buchholz, Wilfried, "Explaining Gentzen's consistency proof within infinitary proof theory", in G. Gottlob et al. (eds.), *Computational Logic and Proof Theory, KGC '97*, Lecture Notes in Computer Science 1289 (1997).

Buss, Samuel R., "An introduction to proof theory", in Buss (1998), 1-78.

Buss, Samuel R., "First-order proof theory of arithmetic", in Buss (1998), 79-147.

Dreben, Burton and Denton, John, "Herbrand-style consistency proofs", in Myhill et al (1970), 419-433.

Fairtlough, Matt and Wainer, Stanley S., "Hierarchies of provably recursive functions", in Buss (1998), 149-207.

Feferman, Solomon, "Proof theory. A personal report", Appendix to Takeuti (1987), 447-485.

Feferman, Solomon, "Hilbert's program relativized. Proof-theoretical and foundational reductions". *J. Symbolic Logic* 53 (1988), 364-384.

Kreisel, Georg, "A survey of proof theory", *J. Symbolic Logic* 33 (1965) 321-388.

Mints, Grigori E., "Proof theory in the USSR 1925-1969", *J. Symbolic Logic* 56 (1991), 385-423.

Mints, Grigori E., "Gentzen-type systems and Hilbert's epsilon substitution method. I", in *Logic, Methodology and Philosophy of Science IX* (D. Prawitz et al, eds., 1994), 91-122.

Pohlers, Wolfram, "Contributions of the Schütte school in Munich to proof theory", Appendix to Takeuti (1987), 406-431.

Pohlers, Wolfram, "Subsystems of set theory and second order number theory", in Buss (1998), 209-335.

Rathjen, Michael, "Recent advances in ordinal analysis: $\prod_2^1$-CA and related systems", *Bulletin of Symbolic Logic* 1 (1995), 468-485.

Schwichtenberg, Helmut, "Proof theory: some applications of cut-elimination", in *Handbook of Mathematical Logic* (J. Barwise, ed., 1997) 867-895.

Sieg, Wilfried, "Hilbert's program sixty years later", *J. Symbolic Logic* 53 (1998) 338-348.

Simpson, Steven G., "Partial realizations of Hilbert's program", *J. Symbolic Logic* 53 (1988) 349-363.

Tait, William, "The substitution method", *J. Symbolic Logic* 30 (1965) 175-192.

Tait, William, "Normal derivability in classical logic", in *The Syntax and Semantics of Infinitary Languages* (J. Barwise, ed.), Lecture Notes in Mathematics 72 (1968) 204-236.