Fragments of First and Second Order Arithmetic and Length of Proofs

By

Aleksandar Djordje[Ignjatovic]

Graduate (University of Belgrade) 1981
Graduate (University of Belgrade) 1984

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

LOGIC AND THE METHODOLOGY OF SCIENCE

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA at BERKELEY

Approved:

Jack Silver	4/17/90
Chair	Date
Charles D. Parsons	4/16/90
Robert Solovay	4/3/90

****

Fragments of First and Second Order Arithmetic and Length of Proofs

by

Aleksandar Djordje Ignjatović

Fragments of First and Second Order Arithmetic and Length of Proofs

by

Aleksandar Ignjatović

Abstract

The main objective of this thesis is to give an analysis of certain forms of the view in the Philosophy of Mathematics usually called Mathematical Instrumentalism. For this purpose we obtain the following technical results relevant for our philosophical arguments.

(1) The proofs of variable free equations involving primitive recursive functions in $I\Sigma_1$ are much shorter than the proofs of the same equations in the Primitive Recursive Arithmetic. More precisely, we show that $I\Sigma_1$ has a non-elementary speed-up over the Primitive Recursive Arithmetic.

(2) RCA_0 has no significant speed-up over $I\Sigma_1$.

These results are relevant to how useful some mathematical instruments are, and we use them to show that certain forms of the instrumentalist view are not tenable.

The second part of the thesis deals with the main concern of an instrumentalist, the consistency proof. We show that a proposal for proving the consistency of theories put forward by M. Detlefsen is not viable as originally formulated, but that a variant of it leads to a partial program, not justifiable on purely finitistic grounds. We also discuss a suggestion of Gödel to replace the consistency proofs by "practical reductions". We show that while his original proposal is related to some open questions in Complexity Theory, a philosophically more acceptable form of it has a very limited potential. The relevant technical results for the second part of the thesis are the following ones.

(3) If T is a theory extending PRA, then $PRA \vdash Con(T) \leftrightarrow Con(PRA)$ iff there exists a primitive recursive function f such that $PRA \vdash \forall x Prf_{PRA}(f(x), \ulcorner \neg Prf_T(\underline{x}, \bot) \urcorner)$.

(4) $PRA \vdash \forall x (\exists y < hyperexp(x)) Prf_{PRA}(y, \ulcorner Con_{WKL}(\underline{x}) \urcorner)$.

(5) For all natural numbers n the following incompleteness results holds:

$$PRA \nvdash \forall x \exists y < \underbrace{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}}_{x} Prf_{PRA}(y, \ulcorner Con_{I\Sigma_1}(\underline{x}) \urcorner).$$

Contents

Preface and Acknowledgements

This thesis is an essay on the view in the Philosophy of Mathematics which is usually called Mathematical Instrumentalism. The main goal of this thesis is to show how the tools of Mathematical Logic, and in particular the tools used in some areas of the Complexity Theory, can be fruitfully used to give an informative analysis of some particular forms of Mathematical Instrumentalism. I believe that the technical results obtained in this thesis can be interpreted in a philosophically informative way. Thus, besides pointing to a general strategy for an informative analysis of such views which goes beyond what can be obtained by pure philosophical arguments, I hope that this thesis will shed more light on some very particular problems.

Before elaborating, I want to explain why I am interested in Mathematical Instrumentalism. When I came to Berkeley to study Mathematical Logic and the Philosophy of Mathematics I was a militant Platonist. But I soon discovered that embracing this view had some embarrassing consequences. Postulating the existence of non-spatial and extemporal objects leaves a Platonist with the problem of explaining how our mathematical intuition is connected with the truth of mathematical propositions on the Platonic universe. Also, this view could be either supported or challenged with purely philosophical arguments, and did not seem to be amenable to a rigorous analysis using our most powerful tool – Mathematics[1]. On the other hand, even though its original form was seriously shaken by Gödel's Second Incompleteness Theorem, Hilbert's instrumentalism impressed me with several of its features. First of all it does not make any dubious ontological assumptions; it has a very appealing clarity, and finally, it is amenable to a rigorous mathematical investigation. This motivated me to try to make a more elaborate analysis of a particular instrumentalist view with finitist mathematics (in the form of PRA) as the fully meaningful part of mathematics and some subsystems of Analysis (provided by the Reverse Mathematics Program) as various instruments. This Thesis is the outcome of this project. Again, the purpose of

[1]Some Platonists would sharply disagree claiming that the whole mathematical practice gives a support for Platonist's view – an argument that never seemed very persuasive to me.

it is not to vindicate or discredit anyone's particular view, but to provide a strategy for analyzing the instrumentalist view in an informative way.

***

I want to thank my advisor Professor Jack Silver for his support and the many interesting discussions which I had with him during my five years at Berkeley; they ranged from Logic and Philosophy to Molecular Biology and Politics. His scepticism about logic which is only concerned with sophisticated technical manipulations and is not motivated by metamathematical and philosophical issues influenced me to search for problems which are firmly philosophically grounded. Professor Charles Chihara spent an enormous amount of time and good will teaching me Philosophy, and encouraging me, especially in difficult moments. Thus, I have to forgive him for ousting me from the Platonist's Paradise. Professor William Craig provided many stimulating ideas and very helpful criticism during the project. He also provided me with many references that still keep me busy reading. I also took great pleasure and benefited from talking with Professors Leo Harrington, Leon Henkin and Robert Solovay. On several occasions I spent time in San Diego, and there I had the opportunity to talk to Professor Samuel Buss. Not only did he teach me Proof Theory on the blackboard, but his constant interest in my work and willingness to discuss it made the the whole project possible. In preparing to see him I would make real progress in my project, and he never hesitated to encourage me to go further. He also called my attention to Pudlák's work, and this was crucial for my further interest in the subject. The starting point of this project was my critical reading of Professor Detlefsen's paper "On Interpreting Gödel's Second Theorem", [2]; Professor Resnik's Book "Frege and the Philosophy of Mathematics", [25], also stimulated my interest in the subject. I would also like to thank all other Professors and students at UC Berkeley with whom I shared the joy of doing Logic and Philosophy.

***

In San Diego I stayed with my second family – the Bishops. The central technical result of this Thesis (The Main Lemma) was proved in the studio of the late Professor Errett Bishop, the founder of the Constructive Analysis. The encouragements and love of Jane Bishop, Tom and Janny, Ed and Loki, and Rosemary and Richard was truly invaluable,

and I express my deepest love for them here. I am grateful to my parents Irina and Djordje who brought me up to love knowledge, to my grandmother Nenka who was "in charge of me" for the first six years of my life and who is the best philosopher I know, living with no fear of death and with a perfect understanding of her world and people who inhabit it. Invaluable support came from my sister Jelena whose advice I always take even though she is ten years younger than I. Finally my thanks and love go to Carel who stayed around when it was the hardest to be around me.

This thesis is dedicated to all of them.

Aleksandar Ignjatović

Chapter 1

Introduction

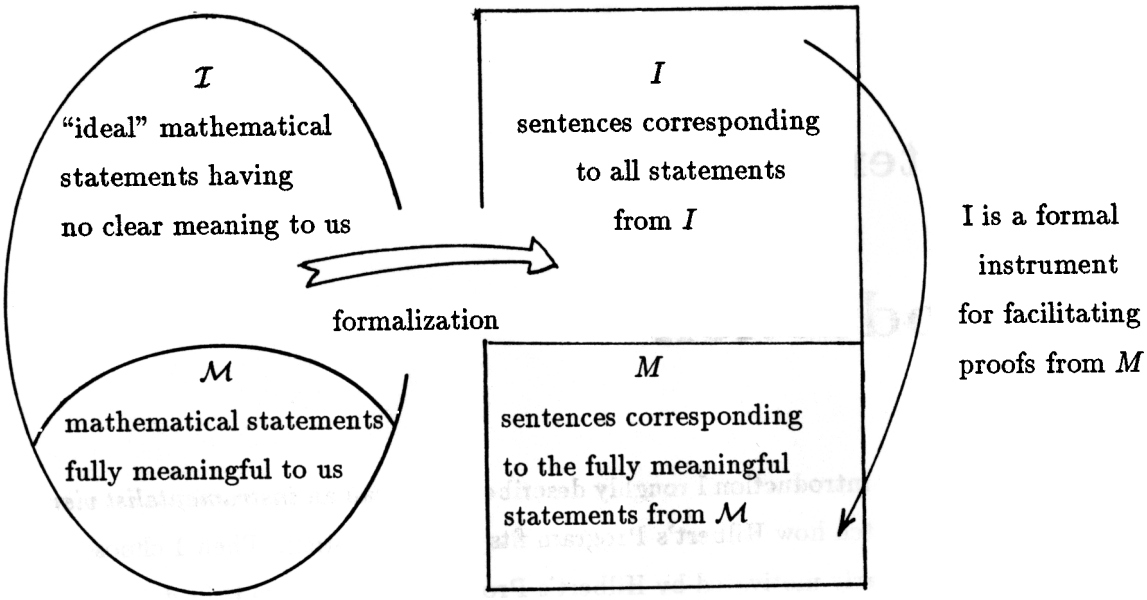
In this introduction I roughly describe what I call an *instrumentalist view of mathematics* and sketch how Hilbert's Program fits this description. Then I choose a very particular view, clearly motivated by Hilbert's Program, and analyze it in detail in the rest of this thesis. The summary of this analysis is given in this introduction.

1.1 The Instrumentalist View and Hilbert's Program

Roughly speaking, an instrumentalist takes only a limited part of mathematics $\mathcal{M}$ and claims that this is the only part of mathematics that has clear, real meaning to us[1]. Then he adjoins another part $\mathcal{I}$ of mathematics, $\mathcal{I} \supset \mathcal{M}$, for which he claims that it has no real meaning. This implies that he cannot apply real, material deduction to the propositions from $\mathcal{I}$; consequently he must formalize the part of mathematics $\mathcal{I}$ into a formal calculus I and replace material deduction in $\mathcal{I}$ by a formal proof in I. Since $\mathcal{M} \subset \mathcal{I}$, there is a subset M of I which consists of all sentences that correspond to the fully meaningful mathematical statements from $\mathcal{M}$. An instrumentalist now claims that he uses formalization I of $\mathcal{I}$ only as a technical device for *facilitating proofs* of sentences from M. Consequently, for as long as he can show that I is a sound instrument in proving sentences from M, he does not have to argue about the meaning of the part of mathematics formalized by I or about the

[1] Sometimes a weaker claim is made about $\mathcal{M}$; for example one can claim that $\mathcal{M}$ is special because we have a clear intuitive representation of the objects that $\mathcal{M}$ seems "to talk about" (e.g. finite sequences of signs) or just that $\mathcal{M}$ is "a minimal kind of reasoning presupposed by all nontrivial mathematical reasoning about numbers" (see Tait [36]). We will consider here primarily the most radical instrumentalist view formulated above.

existence of objects that I seems to "talk about". Thus we have the following picture.



The most important example of such a view is the instrumentalism that was put forward in Hilbert's Program, initiated by him in his paper "On the Foundation of Logic and Arithmetic" [13]. He formulated it more precisely in his paper "On the Infinite" [14], and then further elaborated on it in "On the Foundations of Mathematics" [15].

For Hilbert the object of consideration of number theory are the numerals I, II, III, ... which are signs that "have no meaning at all in themselves" (see [15], p.377)[2]. Genuinely meaningful statements about numbers (which he calls "finitary meaningful statements") are of two kinds. Statements of the form $II + III = IIIII$ or $IIII > II$ express that if we concatenate symbol II to symbol III we get symbol $IIIII$, and that symbol II is an initial segment of symbol $IIII$. Such statements which contain only numerals and symbols for primitive recursive "manipulations" of these numerals are not only fully meaningful to us, but they can be meaningfully negated and logical operations can be freely applied to them. Moreover, they can be ether proved or refuted by material deduction and the process of their verification is epistemologically basic and immediately evident to those who apply it. The other kind of statements which Hilbert claimed to be finitistically meaningful are more problematic; they involve numerals, symbols for primitive recursive manipulations of numerals and *schematic letters*. Schematic letters are not variables but rather they stand

[2]Here we talk about late work of Hilbert; in his earlier writings he claimed that these signs represented "taught objects" (see [13], p.131).There is no doubt that this clarification of his view was partly a result of the polemic correspondence he had with Frege.

for arbitrary numerals. Thus, for example[3], $\alpha + I > \alpha$ means that whenever α is replaced by a concrete numeral, the numeral on the right-hand side of the symbol $>$ is an initial segment of the numeral on the left-hand side of $>$. Statements involving schematic letters are only *hypothetical judgements* which assert something only when schematic letters are substituted by concrete numerals, and are not in general capable of finitary meaningful negation. Existential statements are not in general finitistically meaningful unless the existential quantifier can be bounded by a concrete numeral and consequently replaced by a finite disjunction – see [14] p. 377-8 and [15], p. 470. Yet, some of the statements which involve schematic letters can also be proved through material deduction. Since not all finitistically meaningful statements are capable of finitary meaningful negation and since we cannot always apply meaningfully the $\exists$-introduction rule, the domain of real statements is neither closed under the application of logical operations, nor under the rules of inference of classical logic. In order to avoid being forced to abandon the simple laws of Aristotelian logic which are so inherent to our thinking (and consequently a necessary feature of any successful mathematical theory) we adjoin the ideal propositions to the finitary ones[4]. To achieve this we use the "pre-established harmony" as Hilbert says; the *logical calculus* had already been developed, even though for a different purpose.

> To be sure it [logical calculus] was originally created in an entirely different context, and, accordingly, its signs were initially introduced for purposes of communication only; but we will be consistent in our course if we now divest the logical signs, too, of all meaning, just as we did the mathematical ones, and declare that the formulas of the logical calculus do not mean anything in themselves either, but are ideal propositions. In the logical calculus we possess a sign language that is capable of representing mathematical propositions in formulas and of expressing logical inference through formal processes. In a way that exactly corresponds to the the transition from contentual number theory to formal algebra we regard the signs and operation symbols of the logical calculus as detached from their contentual meaning. In this way we now finally obtain, in place of the contentual mathematical science that is communicated by means of ordinary language, an inventory of formulas that are formed from mathematical and logical signs and follow each other according to definite rules ([14], p. 381).

Thus, to summarize, in place of *statements about numbers* we have *formulas* which are themselves the object of study, and in place of *number theoretical material deduction* we

[3]Hilbert used Gothic letters for schematic letters.

[4]Hilbert compares this situation with the introduction of the imaginary unit in algebra ntersection point of parallel lines in geometry which were introduced to preserve "the laws of

have *formal derivation* according to some determinate rules. Only some formulas correspond to the meaningful statements about numerals, while formulas that correspond to the ideal mathematical propositions signify nothing in themselves and are the ideal structures of our theory.

Did Hilbert really think that the ideal formulas were totally meaningless? In his paper "The foundations of mathematics" [15], he hints otherwise:

> This formula game [proof theory of a formal system] enables us *to express the entire thought-content of the science of mathematics in a uniform manner and develop it in a such a way that, at the same time, the interconnections between the individual propositions and facts become clear.* To make it a universal requirement that each individual formula then be interpretable by itself is by no means reasonable; on the contrary, a theory by its very nature is such that we do not need to fall back upon intuition or meaning in the midst of some argument. What the physicist demands precisely of a theory is that particular propositions be derived from laws of nature or hypotheses solely by inferences, hence on the basis of a pure formula game, without extraneous considerations being adduced. Only certain combinations and consequences of the physical laws can be checked by experiment - just as in my proof theory only the real propositions are directly capable of verification ([15], p. 475, my emphasis).

Thus, for Hilbert, even though ideal sentences have no meaning realizable in intuition in themselves when taken in isolation, they do have a meaning as a system as a whole. Hilbert never claimed that the ideal sentences had literal meaning and that logical operations and inferences could be materially applied to them, but he did attach importance to "the formula game" that goes far beyond the importance of a totally meaningless calculus and beyond the importance of a device for a mere facilitation of proofs of sentences that correspond to the finitistically meaningful statements about numbers.

The main condition he attached to the use of ideal elements is *the consistency proof* for the system in question

> But in our joy that we have, in general, been so successful and that, in particular, we found ready-made that indispensable tool, the logical calculus, we must nevertheless not forget the essential prerequisite of our procedure. For there is a condition, a single but absolutely necessary one, to which the use of the method of ideal elements is subject, and that is the *proof of consistency*; for, extension by the addition of ideals is legitimate only if no contradiction is thereby brought about in the old, narrower domain, that is, if the relations that result for the old objects whenever the ideal objects are eliminated are valid in the old domain ([14], p. 383).

He called for a consistency proof of arithmetic (which must be absolute and not reductive like the one he gave for geometry) and other important theories of mathematics which would be finitary and thus provide a safe ground for a belief in a paradox-free future of mathematics. He (unlike Frege) clearly understood that the consistency of a formal system is a purely combinatorial question about the (un)derivability of a particular formula using definite rules of the formal calculus in question from some particular formulas chosen as axioms of the system. Thus, he realized that the consistency problem is perfectly amenable to mathematical treatment, but he mistakenly believed that such a combinatorial question must be solvable using restricted finitary means. He also thought that beyond the consistency proof of a formal system no proof of the existence of objects that the formal system seems to "talk about" was needed or possible[5].

Hilbert attached only one further condition that a theory had to satisfy:

> The final test of every new theory is its success in answering preexistent questions that the theory was not specifically created to answer ([14], p. 384).

This (in another sense) instrumentalist attitude is present among Platonists as well; Gödel in his "What is Cantor's continuum problem" claims that the introduction of new axioms of set theory can be justified by their success in resolving old problems in a satisfactory way.

The problem of supporting the consistency of a formal theory taken as an ideal instrument for facilitating derivations, either by a proof or by other far less persuasive means, as well as the question of how successful the theory is, are the two main issues and guidelines in our analysis of the particular instrumentalist view we will investigate.

1.2 The instrumentalist view we want to investigate

We now want to specify what the fully meaningful part of mathematics is for the view we want to investigate, as well as various instrumental theories which we will adjoin to the fully meaningful part. We want our view to be as close as possible to Hilbert's original view, but there are two main obstacles in this regard. First of all, Hilbert never specified what exactly the finitistic, fully meaningful part of mathematics was, mostly because he seemed to believe that if a finitistic consistency proof of, say, analysis was achieved, everyone

[5]This view was one of the main points of con

would recognize the means used in the proof as clearly finitistically valid. Second, his project is too ambitious for us: Hilbert incorporated the full induction for the natural numbers into his formal system (see [14], p. 382), claiming that a finitary consistency proof for this system was possible. It is reasonable to expect that in such a strong system we can prove sentences corresponding to the finitistically meaningful statements that are not provable by purely finitistic means.

As far as the first problem is concerned, there is still a great deal of controversy over the question of what constitutes a valid finitistic proof. Tait has given, in his paper "Finitism", an interesting analysis of finitistic mathematics which seems to indicate the following.

1 A function is defined in a finitistically acceptable manner just in case it is given by an explicit definition by primitive recursion.

2. Finitistically meaningful sentences about numbers are Π_1 sentences of the Primitive Recursive Arithmetic (PRA).

Finitistically valid means of proof are exactly those formalized by PRA [6].

Thus, according to Tait, finitistic truths correspond just to the Π_1-theorems of PRA. Even though his arguments are quite persuasive, some doubt still remains, as it is always the case with the arguments concerning whether a formalization adequately captures an informal notion or not. In any case, at this moment, Tait's Thesis seems to be the most acceptable working assumption for making a strict delimitation of finitism, and this is the main reason we accept it.

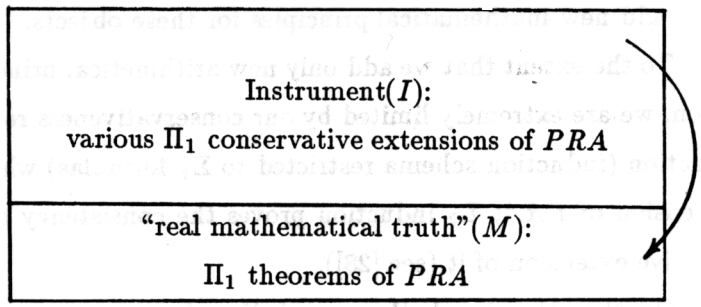
Regarding the second problem, we will consider a more radical form of instrumentalism that sees the formal system only as an instrument for facilitating proofs of the sentences that correspond to the finitistically meaningful statements, and does not attach any other meaning to the instrument[7]. Thus, it seems reasonable to require that our instrumental theories are conservative extensions of the the base theory formalizing contentual derivation (in our case PRA) for all sentences corresponding to the meaningful mathematical statements (in our case Π_1 sentences of PRA). After all, one can argue that an

[6] In the further text we call this assumptions Tait's Thesis.

[7] This is not in the spirit of Hilbert's original view, and in fact we will indicate that it does not seem to be a very tenable one. Nevertheless, it is *prima facie* a plausible view worth investigating (our technical results will be relevant to other forms of instrumentalism as well). We take it again as a convenient working

instrument should not provide us with new truths but only facilitate the proofs obtained by formalizing contentual derivations. This is again merely a useful working assumption with reasonable philosophical justification; it is obviously open to objections.

Thus, our working "prototype" of a pair of theories such that one of them formalizes the meaningful part of mathematics and the other is an instrument for facilitating proofs which are formalizations of contentual derivations has the following form:



We can ask two kind of questions about the relationship of the theories M and I motivated by Hilbert's consistency and "successfulness" concerns:

1. Since by Gödel's Second Incompleteness theorem we cannot prove the consistency of I using only tools available in M, can we find some other (obviously much weaker) support in the belief in the consistency of I on the basis of M?

2. Is the instrument I useful at all? In particular:

   - Are the I proofs of the sentences from M of much shorter length than their M proofs?
   - Are the I proofs of sentences from M *conceptually much clearer* than their M proofs?

If we want our considerations to have any philosophical importance and not to be based on artificial examples of no real interest, we must choose our instruments I such that they formalize important theories of mathematical practice. Fortunately, such theories have been developed and studied in a great deal of detail in the course of the Program of Reverse Mathematics, most notably by Friedman who initiated the program, and by Simpson and his students who did the bulk of the work.

1.3 Conservative extensions of PRA as finitist's instruments

Building Π_1 conservative extensions of PRA of mathematical importance can be done in two directions. We can add stronger arithmetical principles (which seems to amount to adding more induction - see Isaacson's [20] for an argument in support of this), or we can first expand the language with variables for new type of objects (like sets, for example) and then add new mathematical principles for these objects.

To the extent that we add only new arithmetical principles in the form of stronger induction, we are extremely limited by our conservativeness requirement: all we can add is Σ_1 induction (induction schema restricted to Σ_1 formulas) which is indeed a Π_1 conservative extension of PRA; Σ_2 induction proves the consistency of PRA and thus it is not a conservative extension of it (see [28]).

On the other hand, the second direction leaves us much more space, for several theories that formalize significant portions of mathematical practice are Π_1 conservative extensions of PRA. The weakest one is called the Recursive Comprehension (RCA_0) and it brings the ability to define second order objects (i. e. sets of numbers) if they are simply definable (in a Δ_1^0 way) as well as the power of induction for Σ_1^0 formulas with free second order variables as parameters. The Weak König's Lemma (WKL_0) adds to this the power of the compactness principle which is one of the most fundamental mathematical tools[8]. While in RCA_0 we can mainly define basic notions of analysis in a correct way, in WKL_0 we prove all theorems needed for, say, classical mechanics. Good examples are:

- any real function continuous on a closed bounded interval is uniformly continuous, Rieman integrable and attains a maximum value on that interval;
- Heine-Borel covering theorem;
- the local existence theorem for solutions of ordinary differential equations (for more examples see [30]).

Further strengthenings of such theories have been obtained, for example, by adding to WKL_0 a formalization of Baire's Category Theorem (WKL_0^+ of Simpson and Brown, see [1] and [30]). Thus, we have the following picture.

$$PRA \subset I\Sigma_1 \subset RCA_0 \subset WKL_0 \subset WKL_0^+ \subset \ldots$$

Here PRA is the only part of mathematics of "real interest" to us (of course only in our discussion of mathematical instrumentalism); $I\Sigma_1$ is the only purely arithmetical instrument, while the other theories formalize the set theoretical principles.

One could expect that theories formalizing powerful mathematical tools such as compactness (WKL_0) would greatly speed-up proofs of various universal propositions about numbers. On the other hand, even though the consistency of such theories is not provable finitistically, their importance and "relative modesty" should motivate us to look for alternative ways of supporting their consistency by appealing to some forms of finitistic evidence. In this thesis we deal both with the successfulness and with the consistency problems associated with a particular instrumentalist view. In this view the meaningful part of mathematics is represented by Π_1 theorems of PRA and in which instruments are the theories mentioned above. In the next section, motivated by Gödel's "successfulness" test that each theory has to pass, we deal with the question of how useful some particular instrumental theories are.

1.4 Usefulness of a mathematical instrument

Unless otherwise specified we will work in the standard first order Hilbert type proof system (see [3]). Let p be a proof in a first order theory. We will denote by $|p|$ the total length of the proof p, counting also the length of the formulas of the proof (i.e. the total number of symbols in p; see [23] for the details). We will assume that all formulas and proofs are written in a two letter alphabet, say $\{0,1\}$, just to simplify length estimates. Obviously lengths of formulas and proofs in any other finite alphabet differ insignificantly from the point of view that we will adopt (see below); namely they differ linearly, and so our assumption is unproblematic.

We denote by 2^n_m the stack of m two's ending with n as the last exponent:

$$2^n_m = \underbrace{2^{2^{\cdot^{\cdot^{2^n}}}}}_{m}$$

More formally, $2^n_0 = n$, $2^n_{m+1} = 2^{2^n_m}$. A function has *Kalmar elementary growth rate* if there exists a natural number m such that $f(x)$ is majorized by 2^x_m. We say that a function f has a *roughly hyperexponential growth rate* if its growth rate is not Kalmar elementary but for some polynomial $P(x)$ with natural coefficients, $f(x)$ it is dominated by $P(2^x_x)$ (i.e. the function obtained by replacing the variable x in $P(x)$ by 2^x_x; the function 2^x_x is called the

hyperexponential function and is the first function in the Wainer hierarchy which dominates all the elementary functions). A function has a polynomial growth rate if it is dominated by a polynomial with natural coefficients.

In the next two definitions let S and T be two theories such that $S \subset T$, and let Φ be a subset of the set of theorems of S.

Definition 1.1 *We say that theory T has a* **roughly hyperexponential speed-up over the theory S with respect to the set Φ** *if*

1. *there exists a sequence $\{\varphi_i\}_{i\in\omega}$ of formulas from Φ such that if p_n^S and p_n^T are the shortest proofs of φ_n in S and T respectively, then for no elementary function f is it true that for all n, $|p_n^S| < f(|p_n^T|)$ holds;*

2. *there is a function f with a roughly hyperexponential growth rate such that for any formula $\varphi \in \Phi$, if p_φ^S and p_φ^T are the shortest proofs of φ in S and T respectively, then $|p_\varphi^S| < f(|p_\varphi^T|)$.*

Definition 1.2 *We say that theory T has at most a polynomial speed-up over the theory S if there exists a polynomial $P(x)$ with natural coefficients such that for any theorem φ of both S and T, if p_φ^S and p_φ^T are the shortest proofs of φ in S and T respectively, then $|p_\varphi^S| < P(|p_\varphi^T|)$.*

We can now formulate the main questions we associated with the instrumentalist point of view (see page 7) more precisely and specifically for the instrumental theories we want to consider.

Question 1 *Are $I\Sigma_1$, RCA_0, WKL_0 and similar Π_1 conservative extensions of PRA useful and efficient instruments at all? More precisely:*

1. *Are the $I\Sigma_1$ proofs of Π_1 theorems of PRA much shorter than the PRA proofs with the same conclusions? (Here we take PRA as a first order theory rather than as an equational theory to eliminate the advantage that $I\Sigma_1$ could have purely from its first order logic with the cut-rule; see also footnote 9). Are RCA_0 (WKL_0) proofs shorter than $I\Sigma_1$ (RCA_0 respectively) proofs with the same conclusions?*

2. *Are the proofs that use the set theoretical concepts and principles (RCA_0, WKL_0 and WKL_0^+ in our case) conceptually clearer and easier to understand than the (basically*

We will show that $I\Sigma_1$ indeed has a huge, in fact what we call a roughly hyperexponential speed-up over PRA. We will also show that RCA_0 has an insignificant, at most polynomial speed-up over $I\Sigma_1$ (quadratic or linear depending on the particular proof system). We also conjecture that WKL_0 (as well as WKL_0^+) can also have only a polynomial speed-up over RCA_0, proposing a strategy to prove this conjecture[9].

Note that in the above considerations the speed-up is measured only in terms of lengths of proofs; an instrumental theory can have a conceptual advantage over the base theory (i.e. proofs which use concepts and principles that make the proof easier to grasp), regardless of whether there is a significant speed-up in terms of length of formal proofs or not. We will address this question later.

We now have to make a decision about which speed-up we consider significant and which we do not. One could argue that we should find (a finite number of) mathematically important Π_1 theorems about numbers and compare the lengths of their PRA proofs with, for example, the lengths of their WKL_0 proofs. But if we want to draw conclusions of any philosophical importance, we must consider mathematics as an (at least potentially) infinite collection of truths, and in this case only the growth rate of the size of proofs matters. Extensive research in Complexity Theory strongly indicates that it is natural to consider two procedures equally efficient (in a more theoretical sense) if the number of steps to execute any one of them is smaller than the value of a polynomial evaluated at the number of steps needed to execute the other one for the same input. We accept this for provability in formal theories, and consequently if a theory T has only a polynomial speed-up over a theory S, we consider them as equally efficient instruments for deriving truths[10]. This convention justifies our conclusions that $I\Sigma_1$ has a significant speed up over PRA and that RCA_0 does not have a significant speed up over $I\Sigma_1$.

Evaluating the conceptual benefits that an instrumental theory provides is a much more difficult task and must be done on the "empirical basis", by providing examples

[9]One can argue that PRA is not the right theory for representing finitistic reasoning since its proof system involves formulas of arbitrary complexity, and according to Hilbert's view, they are ideal objects of our system. A better theory would then be $eqPRA$, which is a version of the primitive recursive arithmetic formulated as an equational theory in a language without quantifiers, and consequently whose entire proof system involves only finitistically meaningful formulas. In this case PRA as a first order theory could be seen as a *logical instrument* over the base theory $eqPRA$. I conjecture that PRA has also a non-elementary speed-up over $eqPRA$. Statman [35] proved that the usual first order logic has a non-elementary speed-up over the cut-free proofs, and while PRA has a proof system allowing arbitrary cuts, $eqPRA$ can only have cuts on quantifier free formulas only. Thus the above conjecture seems reasonable and most likely it can be proved by relativizing Statman's proof to PRA.

of reasonably interesting theorems whose proofs which use set theoretical concepts and principles are conceptually clearer than their purely arithmetical proofs. We believe that our technical results indicate that the set theoretical concepts and principles do provide a great deal of conceptual facilitation, even in the case of the weakest set theory we consider, namely RCA_0. The definition of a cut $J_{I\Sigma_1}$ in $I\Sigma_1$ (see below for the definitions) and the $I\Sigma_1$ proof of the consistency of PRA on this cut (i.e. that $I\Sigma_1 \vdash \forall x(J_{I\Sigma_1}(x) \rightarrow \neg Prf_{PRA}(x, \underline{1=0}))$) is conceptually more difficult to understand than the proof of the *same* proposition in RCA_0: the cut $J_{I\Sigma_1}$ is included in a cut in RCA_0 whose definition provides an immediate understanding of the nature of the cut, and the proof itself is conceptually much clearer than the one in $I\Sigma_1$, since it involves much less coding. It seems obvious to me that using recursively defined sets is as natural to us as using coding is artificial. Certainly this is only a hint for an argument supporting the claim that theories which involve set theoretic concepts and principles do provide conceptual facilitation; perhaps better examples can be found with WKL_0, since Weak König's Lemma provides us with compactness and consequently avails model theoretic arguments as well as application of compactness in the form used in mathematical analysis. The last seems the most promising source of conceptual facilitation which WKL_0 might bring; one must only recall proofs of some theorems in number theory which use complex integration and similar other techniques from analysis.

Thus, if it turns out that indeed WKL_0 has no significant speed-up over RCA_0, and WKL_0^+ has no significant speed-up over WKL_0, this would seem to support the following two principles:

1. If we want a conservative instrument that makes proofs of arithmetical proposition *shorter in length*, we need stronger *arithmetical* principles (i.e. induction) embodied in the instrument;

2. Conservative instruments that formalize stronger mathematical principles that are *not* of an arithmetical nature (e.g. set theoretical principles) can produce only *conceptual facilitation*, i.e., they can make proofs conceptually clearer and easier to grasp, without making formal proofs shorter in length.

We are now in a position to demonstrate how our technical result can be used to derive some philosophical conclusions about our understanding of the meaning of certain mathematical concepts and principles and the instrumentalist view. Consider, for example,

View 1 *We have a clear understanding of the concept of numbers satisfying the principle of Σ_1 induction, but we have no equally clear understanding of the concept of sets satisfying the principle of recursive comprehension. The later concept is used (in the form of the theory RCA_0) merely as a formal instrument for facilitating proofs which use only Σ_1 induction.*

We argue that the above view is not tenable. Even though we indicated that it might well be the case that RCA_0 does provide conceptual facilitation over $I\Sigma_1$, since we do not have enough "empirical" evidence to support this claim, our argument will not make use of it. Rather, we argue by distinguishing the following two cases.

Assume first that there is no significant body of arithmetical truths[11] whose RCA_0 proofs are conceptually clearer than their $I\Sigma_1$ proofs. By our result mentioned above RCA_0 has no significant speed up in terms of the lengths of formal proofs over $I\Sigma_1$. Thus, since we have neither significantly shorter nor conceptually clearer proofs in RCA_0 than in $I\Sigma_1$, the instrument RCA_0 is useless, for it is hard to think of any other benefit besides the above two that an instrument might offer. Thus, in this case the instrumentalist view in question would collapse.

In the second case, we assume that there is a significant body of arithmetical truths whose RCA_0 proofs are indeed conceptually clearer than their $I\Sigma_1$ proofs. In this case, since the benefit from our instrument is *conceptual clarification* of proofs, the part of mathematics formalized in the instrument RCA_0 must have equally clear meaning to us as the part of mathematics formalized in $I\Sigma_1$, otherwise the RCA_0 proofs could not have any conceptual advantage over the $I\Sigma_1$ proofs. This again conflicts with the instrumentalist view as stated above. Thus, in either case, the above view does not seem to be tenable.

It is now clear that a proof of our conjecture that WKL_0 and WKL_0^+ have only a polynomial speed up over RCA_0 (and consequently $I\Sigma_1$) would have similar consequences as the above argument, implying that if these theories are useful in deriving arithmetical truths, then the concepts and principles formalized in these theories must have equally clear meaning to us as the concept of numbers satisfying Σ_1 induction.

One might argue that the sequence of the theories which we have considered is a very particular one and so our considerations cannot be important enough either to be philosophically relevant or to have any bearing on mathematical instrumentalism in general. Even though we did claim that this project was designed more to indicate a general strategy

[11] RCA_0 is conservative over $I\Sigma_1$ for all arithmetical sentences; this is a consequence of the fact that RCA_0 is interpretable in $I\Sigma_1$ with an unchanged arithmetical part of the theory (see our Theorem 2.7).

of how to analyze instrumentalist view than to give decisive conclusions about anyone's particular view, we still claim that it does also provide a philosophically relevant analysis of instrumentalism. First of all, even though it is true that this indeed is a particular case of instrumentalism, it is a very important one since the base theory is finitism and instrumentalist theories formalize important parts of mathematical analysis. It is difficult to offer any other choice of the base and instrumental theories that clearly formalize more important parts of mathematics than the theories we consider. On the other hand, I believe that one can provide a similar analysis for other important cases, such as the predicative analysis and its conservative extensions. If a significant number of important cases are analyzed in the spirit of our analysis, I think it is clear that philosophically important consequences must follow naturally.

The second and third chapter of this thesis respectively are devoted to the proofs of the results mentioned above:

(Theorem 2.9) $I\Sigma_1$ has Kalmar non-elementary speed up over PRA even for the variable free sentences of PRA.

2. (Theorem 3.5) RCA_0 has at most polynomial speed up (quadratic or linear, depending on the details of the proof system) over $I\Sigma_1$ for all formulas of the language of $I\Sigma_1$.

1.5 Consistency problem

We now turn to the the other, more important of Hilbert's concerns, namely the problem of the consistency proof. If we accept Tait's Thesis and assume that all finitary arguments are formalizable in PRA, then by Gödel's Second Incompleteness Theorem we cannot give finitistic consistency proofs for theories extending PRA. On the other hand, it is very reasonable to look for some other (much weaker) "evidence" for the consistency of our instrumental theories $I\Sigma_1$, RCA_0, WKL_0 etc. Such evidence can be provided, for example, by a consistency proof which uses transfinite induction up to a "small" (ω^ω) but nevertheless transfinite ordinal. One could try to obtain another piece of such evidence by following an idea of Gödel which takes into account the lengths of proofs, but before elaborating this idea we will briefly consider a more radical attempt by Michael Detlefsen[12] [2], which will eventually lead us to Gödel's proposal.

Detlefsen notes that the consistency statements are of the form $\forall x F(x)$, where $F(x)$ is a primitive recursive predicate. Thus, he claims, in order to give a finitistic proof of such a statement, we must construe the universal quantifier *finitistically* and *not classically.* This implies, he argues, that in order to prove finitistically the consistency of a theory T one only has to exhibit a finitistic procedure (what he calls a "manual") which when applied to each particular numeral n produces a finitistic proof of the fact that n is not a code of an inconsistency in T. To support his view he gives a quotation of Herbrand (see Chapter 4).

As we will see in the fourth chapter, for every consistent theory T it is possible to define by primitive recursion a function[13] f which for each n produces a finitistic proof of the statement "n is not a code of a proof of an inconsistency in T". Thus, Detlefsen's claim is unreasonable as formulated. To make such such a claim plausible at all, we must at least assume that a finitist *can prove* that f indeed has the above property. By Tait's Thesis this is the case if and only if for a primitive recursive function f

$$PRA \vdash \forall x Prf_{PRA}(f(x), \ulcorner \neg Prf_T(\underline{x}, \underline{1 = 0}) \urcorner). \tag{1.1}$$

As we will see later, the above is true for all theories T we consider up to (and including) WKL_0, and it and can also be seen as a premise of a very restricted ω rule. Not only do we require that each instance $\neg Prf_T(\underline{n}, \underline{1 = 0})$ is finitistically provable and that these proofs are generated uniformly by a finitistically acceptable function f, but we also require that this property of the function f must itself be finitistically verifiable. Thus, we can accept that the consistency of a theory T is proved whenever 1.1 is true if we accept the following rule of inference for arbitrary primitive recursive predicate $F(x)$:

Restricted ω-rule 1 *Conclude $\forall x F(x)$ whenever possessing an explicit definition of a primitive recursive function f for which $PRA \vdash \forall x Prf_{PRA}(f(x), \ulcorner F(\underline{x}) \urcorner)$.*

In the fourth chapter we will indicate why this kind of ω rule is not in general a finitistically valid rule of inference. We show that the theories whose consistency can be proved using the above rule are exactly those for which one can prove in PRA that they are equiconsistent to PRA. We also characterize theories whose consistency can be proved using, in some sense, finite iterations of such a rule.

[13]Recall that according to Tait's Thesis the functions defined by primitive recursion are exactly the nitistically acceptable ones.

Let $Con_T(\underline{n})$ be a formalization of "there is no proof of an inconsistency in T of length at most n". It is easy to see that 1.1 implies that for $T = PRA, I\Sigma_1, RCA_0$ or WKL_0 there is a primitive recursive function f such that

$$PRA \vdash \forall x \exists y(|y| < f(x) \wedge Prf_{PRA}(y, \lceil Con_T(\underline{x})\rceil)). \tag{1.2}$$

We now relate the above result with the following idea of Gödel. We quote Kreisel [21], marginal comments on page 241.

> ...one may modify Hilbert's (generalized) programme by taking into account the *lengths of proofs*. Thus given a formal system F and an area of evidence $\mathcal{P}$, let $\psi(x)$ be the length of the shortest proof in $\mathcal{P}$ of the consistency of F restricted to the proofs of length n. (If the number of proofs of bounded length in F is finite and F is consistent there will always be such a proof of their consistency.) If $\psi(n)$ is of the same order of magnitude as n, we should have a 'practical' reduction of F to $\mathcal{P}$. I first heard this interesting suggestion in a conversation with Gödel.

In our analysis of this idea, we will take $\mathcal{P}$ to be the finitistic area of evidence since this is the only area of evidence inherent in Hilbert's program. Also, we have to make the informal description "$\psi(n)$ is of the same order of magnitude as n" and the informal notion of a "practical reduction" precise. We will again accept standards from complexity theory: $\psi(n)$ will be considered to be of the same order of magnitude as n if its is bounded by a polynomial in n i.e., if for some polynomial $P(x)$ with natural coefficients the following holds for all n:

$$\psi(n) \leq P(n).$$

If $d(n)$ are proofs of length $\psi(n)$, and the above inequality holds, we will say that proofs $d(n)$ are *feasible in* n. The above stipulation is standard and is discussed in length in the literature. Our second assumption will be more delicate and we defend it on the same grounds we used in criticizing Detlefsen's suggestion about the finitistic provability of universal sentences. In order to have a "practical" reduction of F to $\mathcal{P}$ that is *satisfactory from the standpoint of the area of evidence* $\mathcal{P}$, not only do we have to have $\mathcal{P}$-proofs $d(n)$ of $Con_F(\underline{n})$ whose length $\psi(n)$ is smaller or equal than $P(n)$, but *this fact itself must be verifiable by tools that all belong to the area of evidence* $\mathcal{P}$. In other words, one should be able to give a justification on the basis of the area of evidence $\mathcal{P}$ of the statement:

Here $\overline{\mathcal{P}}$ stands for a formal system corresponding to $\mathcal{P}$; it is necessary to formalize the area of evidence $\mathcal{P}$ in order to make the above statement expressible in such a way that it makes sense to ask if it is justifiable within the area of evidence $\mathcal{P}$ or not. This is because the areas of evidence one wants to consider in this context are various areas of mathematical evidence and so informal proofs or justifications do not fall within their scope. It may seem that this limits what areas of evidence $\mathcal{P}$ we can consider, since these areas must have their formal counterparts. Since we want to consider the finitistic area of evidence which is formalizable as PRA (see [36],[18]), this is not a problem for us. But more importantly, the requirement that $\mathcal{P}$ must be formalizable is built into the problem. For in order to be able to to measure the lengths $\psi(n)$ of the proofs $d(n)$ belonging to the area of evidence $\mathcal{P}$, these proofs must be formalizable as sequences of symbols whose length is counted, and so our assumption is not restrictive at all. The above considerations motivate us to introduce the following Metadefinition, which will replace the informal notion of a "practical reduction".

Metadefinition 1 *We say that the consistency of a theory T is feasibly reducible to the area of evidence $\mathcal{P}$ if one can give a $\mathcal{P}$-proof of the statement "for all n there is a feasible $\overline{\mathcal{P}}$-proof $d(n)$ of the assertion that there is no proof of an inconsistency in T of length $\leq n$".*

Bearing in mind that we identify finitistic provability with provability in PRA, we have the following Metatheorem.

Metatheorem 1 *The consistency of a theory T is feasibly reducible to the finitistic area of evidence if for a polynomial $P(x)$ we have:*

$$PRA \vdash \forall x \exists y(|y| < P(x) \wedge Prf_{PRA}(y, \lceil Con_T(\underline{x})\rceil)).$$

Obviously 1.2 from page 16 shows that any of the theories $PRA, I\Sigma_1, RCA_0$ and WKL_0 have a primitive recursive reduction to the finitistic area of evidence. Can we strengthen this result and get feasible reductions? Formalizing an argument of Pudlák (with a suitable choice of coding), we will show that in the case when T is PRA, we can indeed take for f in formula 1.2 a polynomial. More precisely, by our Theorem 5.1 there is a polynomial $P(x)$ with natural coefficients such that

$$PRA \vdash \forall x \exists y(|y| < P(x) \wedge Prf_{PRA}(y, \lceil Con_{PRA}(\underline{x})\rceil)). \tag{1.3}$$

The above result obviously means that PRA is feasibly reducible to the finitistic area of evidence (in our, stronger sense, as defined in Metadefinition 1). It is now natural to ask

Question 2 *Is there a consistent theory T in which one can formalize a significant portion of mathematical practice and whose consistency is feasibly reducible to the finitistic area of evidence?*

What would be the benefit of such a reduction? At the first sight it seems that in order to prove that there are no proofs of inconsistencies in T of length at most n, one can use the following procedure. There are $2^{n+1}-1$ sequences of 0's and 1's of length at most n, and since T is consistent, none of these sequences is a proof of an inconsistency in T.[14] Since it is easy to determine if a sequence of symbols is a proof, for each of these sequences one can give a proof of length polynomial in n of the fact that this sequence is not a proof of an inconsistency in T. Thus, putting all these proofs together, it seems that one can always give a proof of the "consistency of T for proofs of length up to n" which is of length exponential in n, while in the above question we are looking for feasible proofs of the same statement (i.e. proofs of length polynomial in n). But this is not the case at all; the benefit could be far bigger than just replacing proofs of exponential lengths with proofs of polynomial lengths, because the procedure described above, which gives proofs of exponential lengths, cannot be carried out in PRA, and this would be necessary in order to get a reduction which is itself verifiable in PRA. For the correctness of the above procedure rests upon the assumption that T is consistent and the consistency of any reasonably strong theory T cannot be proved in PRA.

One might hope that a way out for some theories like WKL_0 is to use a result of Sieg's which gives (see [28]) a primitive recursive procedure $g(x)$ for replacing a WKL_0 proof p of any Π_2 sentence of PRA with a PRA proof $g(p)$ of the same sentence, and then combine it with our version of Pudlák's result (Theorem 5.1). In this way, (see Theorem 5.5), we do get a procedure which is formalizable in PRA, but since Sieg's result applies to cut-free proofs only, we must first use a cut elimination procedure which drasticly increases the sizes of proofs. Thus, proofs of $Con_T(\underline{n})$ obtained in this way are of roughly hyperexponential size in n. Interestingly enough[15] this method is the best possible. The following result which we are going to prove in Chapter 5, shows not only that no theory T containing $I\Sigma_1$ is feasibly reducible to the finitistic area of evidence but also that for every such theory T the size of proofs of $Con_T(\underline{n})$ that PRA "accepts" grows non-Kalmar elementary.

[14] Recalls that we code proofs as sequences of 0's and 1's.

[15] And to my disappointment, because I was hoping to find mathematically interesting theories which are asibly reducible to the finitistic area of evidence.

(Theorem 5.4) *For every natural number* n,

$$I\Sigma_1 \nvdash \forall x \exists y(|y| < 2_n^x \wedge Prf_{PRA}(y, \lceil Con_{I\Sigma_1}(\underline{x})\rceil)).$$

As it is shown by Simpson and Smith, Σ_1 induction is necessary (over a very weak base theory) to prove some of the most elementary mathematical theorems, like theorems about factorization of polynomials into irreducible factors (see [32]). Thus, it is safe to say that no significant theory from the point of view of mathematical practice is "practically reducible" to the area of finitistic evidence with a finitistic justification of such a reduction.

With the incompleteness Theorem 5.4 in mind, it is natural to ask what the "least" function f such that $\omega \models \forall x \exists y(|y| < f(x) \wedge Prf_{PRA}(f(x), \lceil Con_{WKL_0}(\underline{x})\rceil))$ is. Obviously f can be taken of exponential growth rate; the previously mentioned method with checking all the proofs of length at most x now works. Can we take a polynomial for f? This obviously brings us back to Gödel's original proposal. Our additional requirement to Gödel's original idea, namely that the existence of a feasible reduction of F to the area of evidence $\mathcal{P}$ must itself be provable in $\mathcal{P}$, is not only philosophically justified, but it also made the problem manageable. If we only require that we have polynomial size proofs of $Con_F(\underline{n})$ from the area of evidence $\mathcal{P}$, without requiring that the existence of such proofs be provable in $\mathcal{P}$, then the problem of the existence of such proofs is still open and is closely related to some open questions in Complexity Theory. It is well known (and easy to prove) that if a (reasonably strong) theory S proves that the classes NP and $coNP$ coincide (for the definitions see [10]), then for every theory T which is consistent, S has proofs of polynomial length in n of $Con_T(\underline{n})$. Thus, if PRA proves that $NP = coNP$ (which the majority of complexity theorists do not believe is true, let alone provable in a theory like PRA), then there are PRA proofs of polynomial length in n of $Con_T(\underline{n})$ for all consistent theories T, thus including WKL_0. Even though $PRA \vdash NP = coNP$ is a very unlikely assumption, we cannot disprove it either, and so we do not know if PRA has proofs of polynomial length in n of $Con_{WKL_0}(\underline{n})$.

One might hope that the above incompleteness result could be used to distinguish among equiconsistent theories of different mathematical strength on a basis closely related to their consistency strength. As is well known[16], for theories T like RCA_0 or WKL_0 we

[16]This is a consequence of Sieg's results from [28].

have:

$$PRA \vdash Con(T) \leftrightarrow Con(PRA).$$

Yet, such theories formalize stronger and stronger mathematical principles, and so it is natural to ask the following question.

Question 3 *Can we develop a finer scale for measuring the consistency strength of theories on which theories of different mathematical strength would "weigh" differently even if they are provably equiconsistent to PRA?*

Recall (1.2 on page 16) that there are primitive recursive functions f_{PRA}, f_{RCA_0} and f_{WKL_0}, such that

$$PRA \vdash \forall x \exists y(|y| < f_T(x) \wedge Prf_{PRA}(y, \lceil Con_T(\underline{x})\rceil))$$

for $T = PRA, RCA_0, WKL_0$. Is it true that for mathematically stronger theories T the least f_T for which the above holds grows faster than $f_{T'}$ corresponding to a weaker theory T'? After all, our version of Pudlák's result (Theorem 5.1) shows that f_{PRA} grows polynomially, while Theorem 5.4 shows that $f_{I\Sigma_1}$ grows roughly hyperexponentially. Unfortunately, the following result which we prove in Chapter 5 together with Theorem 5.4, shows that $f_{I\Sigma_1}$, f_{RCA_0} and f_{WKL_0} all grow roughly hyperexponentially.

(**Theorem 5.5**). *There is a function f with a roughly hyperexponential growth rate such that*

$$PRA \vdash \forall x \exists y(|y| < f(x) \wedge Prf_{PRA}(y, \lceil Con_{WKL_0}(\underline{x})\rceil).$$

Thus, a significant difference in the growth rate of the functions $f_{PRA}, f_{I\Sigma_1}, f_{RCA_0}$ and f_{WKL_0} occurs only between f_{PRA} and $f_{I\Sigma_1}$, while other functions have approximately the same growth rate as $f_{I\Sigma_1}$.[17] I believe that this phenomenon is related to the rate of speed up which such theories have over $I\Sigma_1$, and in fact I put forward the following conjecture. It could be called **"the Second Instrumentalist Incompleteness Conjecture"**, because it claims that the speed up which an instrumental theory T can have over a base theory S is offset by how difficult it is to prove the consistency of T for proofs of length up to n on the basis of S.

Conjecture 1 *Let T and S be a pair of (reasonable) theories such that $S \subset T$, and such that one can prove in S that T is a Π_1 conservative extension of S. Then for every function f definable in S for which*

$$S \vdash \forall x Prf_S(f(x), \lceil Con_T(\underline{x}) \rceil)$$

holds[18] *the following is true: for every Π_1 formula φ, if p_φ^S and p_φ^T are the shortest proofs of φ in S and T respectively, then $|p_\varphi^S| \leq f(|p_\varphi^T|)$.*

The content of Chapters 2 and 3 also appears in my unpublished paper [19]. In the next chapter we develop the main technical tools and prove that $I\Sigma_1$ has a Kalmar non-elementary speed up over PRA.

[18] As we will see in Chapter 4 our assumptions about S and T imply that there are functions with such a

Chapter 2

$I\Sigma_1$ versus PRA

In this thesis, Z is the elementary number theory (which is obtained from PRA by adding the induction schema for all formulas of the language of PRA) and $I\Sigma_n$ is the fragment of Z which contains the induction schema restricted to Σ_n formulas only[1]. RCA_0 is a fragment of second order arithmetic extending PRA by the addition of, besides a few usual basic axioms, only the Σ_1^0 induction and the Δ_1^0 comprehension[2]. Let $\mathcal{L}_{PRA}(y)$ be a formula formalizing *"y is a (Gödel code of a) functional symbol[3] of PRA"*, and let $\mathcal{L}_{PRA}(x,y)$ be $\mathcal{L}_{PRA}(y) \wedge y < x$; thus $\mathcal{L}_{PRA}(x,y)$ formalizes *"y is a functional symbol of PRA whose code is smaller than x"*. $\mathcal{T}_{PRA}(x,y)$ is a formula formalizing *"y is a closed term of PRA containing only functional symbols $< x$"*, while $Ax_{PRA}(x)$ formalizes *"x is an axiom of PRA"*. Notice that $\mathcal{T}_{PRA}(x,y)$ restricts only the language and not the size of the term y – an important feature for the cut-elimination procedure to be used later. In all standard coding procedures these formulas are at most Δ_1^0. We will use special notation for primitive recursive functions which operate on the (formalized) syntactical objects of PRA. $Ar(f)$ equals the arity of the functional symbol f. $\mathcal{G}(f) =< 0, g_f, h_f >$ if f is defined by primitive

[1]Even though $I\Sigma_1$ formulated on the language $\{+,\cdot,=,<,0\}$ suffices to introduce all primitive recursive functions in an extension by definitions, we prefer that the language of our instruments extends the language of our base theory PRA. This has also an advantage that all formulas with only bounded quantifiers are equivalent to quantifier-free formulas by replacing bounded quantifiers with their primitive recursive definitions.

[2]The superscript 0 indicates that we can have free second order variables in the formulas in these schemas but not second order quantifiers. Recall that the Δ_1^0 comprehension is the schema

$$(\Delta_1^0 - CA): \quad \forall x(\varphi(x) \leftrightarrow \psi(x)) \rightarrow \exists X \forall x(x \in X \leftrightarrow \varphi(x))$$

whenever φ is a Σ_1 formula and ψ a Π_1 formula.

[3]In the sequel we will identify functional symbols, terms,formulas and proofs with their Gödel codes and

recursion from g_f and h_f (i. e. if $Ax_{PRA}[f(0,\vec{x}) = g_f(\vec{x}) \wedge f(y+1,\vec{x}) = h_f(y, f(y,\vec{x}),\vec{x})]$), and $\mathcal{G}(f) =< 1, h_f, g_1^f, ..., g_k^f >$ if f is defined by composition from $h_f, g_1^f, ..., g_k^f$. The usual codings have the property that $(\mathcal{G})_i < f$. $\underline{x}$ denotes the primitive recursive function producing the (Gödel code of the) x^{th} numeral.

By a cut in a theory T we mean a formula J with only one free variable x s. t.

$$T \vdash J(0) \wedge \forall x(J(x) \rightarrow (J(x+1) \wedge \forall z < \underline{x}\ J(z))).$$

We will extensively use the following result of Solovay from [34] (see also [24], Lemma 1.1):

Theorem 2.1 *(Solovay) Let $J(x)$ be a cut in a theory T, then*

1. *There is a cut $I(x)$ such that $T \vdash \forall x(I(x) \rightarrow (J(x) \wedge I(x^2))$. (Note that this implies that the cut $I(x)$ is closed for addition and multiplication.)*
2. *For every natural number n there is a cut $K_n(x)$ such that $T \vdash \forall x(K_n(x) \rightarrow J(2_n^x))$.*

In our proofs we use the facts that Σ_1^0 induction, Π_1^0 induction, Σ_1^0 least number principle and Π_1^0 least number principle (all with second order free variables) are all equivalent over a weak base theory; see [28].

After having set these basic definitions, we can proceed to prove that $I\Sigma_1$ has a large speed-up over PRA. The basis for our proof is the following result of Pudlák, which we state in a general form rather than for ZF and GB specifically, as Pudlák did (see Theorem 4.2 of his paper [23]).

Theorem 2.2 *(Pudlák) Let S and T be two consistent theories with sufficiently strong arithmetic*[4] *and such that the axioms of S can be accepted by a nondeterministic polynomial time Turing machine. Let $Prf_S(x,y)$ be a suitable formalization*[5] *of "x is a proof of y in S", and $Con_S(x)$ a suitable formalization of "there is no proof of an inconsistency in S of length at most x". Assume that for a cut $J(x)$ in T, $T \vdash \forall x(J(x) \rightarrow \neg Prf_S(x, \underline{1=0})$. Then:*

[4]While the theories we deal with contain PRA, Pudlák assumes just that they contain (or even only interpret) Robinson's Q. In case S and T contain PRA most of the proofs of lemmas needed to get this theorem can be significantly simplified.

[5]The reader should consult [23] to see how to formalize the syntax such that this theorem holds, and that in fact the usual formalizations which are in use today suffice.

1. *There is a polynomial $P(x)$ such that for all n there is a proof of $Con_S(2^0_n)$ in T of length at most $P(n)$;*

2. *There is a positive real number ε such that for all n there is <u>no</u> proof of $Con_S(2^0_n)$ in S of length at most $(2^0_n)^\varepsilon$.*

Theorem 2.2 obviously implies the following important corollary.

Corollary 2.3 *Whenever S and T are as above, T has a non-Kalmar elementary speed-up over S.*

To prove part (1) of Theorem 2.2 we start with the cut $J(x)$ such that $T \vdash \forall x(J(x) \rightarrow \neg Prf_S(x, \underline{1=0}))$ and use Solovay's cut shortening technique (see Theorem 2.1) to get a cut $I(x)$ such that $T \vdash \forall x(I(x) \rightarrow J(2^x))$. Then obviously $T \vdash \forall x(I(x) \rightarrow Con_S(x))$. We now use the fact that for any cut $I(x)$ there are polynomial size proofs of $I(2^0_n)$ (see [23] for the details); this clearly implies part (1) of the Theorem.

Part (2) of Theorem 2.2 is proved via a modified Gödel-type argument diagonalizing "$\varphi(x)$ is provable in S with a proof of length $\leq x$" rather than just "φ is provable in S" which we diagonalize in the standard Gödel's argument. Again for the details the reader should consult [23].

Pudlák proved the above theorem with GB and ZF in place of our S and T and used it to show that GB has a non-Kalmar elementary speed-up over ZF. In order to use Theorem 2.2 to show that $I\Sigma_1$ has a non-Kalmar elementary speed-up over PRA, we have to find a cut $J_{I\Sigma_1}(x)$ in $I\Sigma_1$ such that $I\Sigma_1 \vdash \forall x(J_{I\Sigma_1}(x) \rightarrow \neg Prf_{PRA}(x, \underline{1=0}))$. As we mentioned in the Introduction, instead of doing this, we will first find a cut J_{RCA_0} in RCA_0 such that $RCA_0 \vdash \forall x(J_{RCA_0}(x) \rightarrow \neg Prf_{PRA}(x, \underline{1=0})$, and then we will use an interpretability argument to get the analogue result for $I\Sigma_1$. The advantage of working in RCA_0 rather than directly in $I\Sigma_1$ is the conceptual facilitation that RCA_0 brings; the proof for RCA_0 involves much less coding and is consequently conceptually much clearer than a direct proof in $I\Sigma_1$.

Before taking this up, we will briefly sketch the usual proof of the corresponding fact for GB and ZF, just in order to be able to see what kind of difficulties we have to overcome in our proof for the case of RCA_0 and PRA. As we will see, we have to use totally different techniques, with proof theory in the place of model theoretic arguments. The usual proof of the fact that there is a cut J_{GB} such that $GB \vdash \forall x(J_{GB}(x) \rightarrow \neg Prf_{ZF}(x, \underline{1=0}))$

proceeds as follows. Let $J(x)$ be the formula of GB formalizing "there exists a satisfaction class satisfying Tarski's inductive truth properties for all Π_x formulas of the language of ZF". It is easy to see that $J(x)$ defines a cut in GB, because if one has a satisfaction class for all Π_x formulas of ZF, then one can define from it explicitly a satisfaction class for all Π_{x+1} formulas of ZF. Then one uses a formalized version of Montague's Reflection Theorem in which a satisfaction class is used rather than the real truth in the set-universe to obtain a set model of the Π_x part of ZF whenever $J_{GB}(x)$ holds. We then apply the usual inductive soundness argument for all proofs that belong to the cut J_{GB}, and show that the conclusion formula of such a proof p must be true in the constructed set model of the Π_p part of ZF. This is possible since the complexity of each formula in the proof is smaller than the (Gödel code of) the proof itself.

Obviously the above kind of proof cannot be applied to the pair RCA_0 and PRA, since in RCA_0 we cannot prove the existence of "satisfaction sets" for quantified sentences of the language of PRA because the comprehension in RCA_0 is too weak to handle quantified formulas[6]. But the cut J_{GB} just defines the collection of numbers that correspond to the complexities for which there is a satisfaction class satisfying Tarski's inductive truth properties, and it seems *prima facie* that we are left without the very tool needed to define the cut with necessary properties.

The above also shows that if we want to do some kind of a soundness argument, we can only hope to be able to find a very partial satisfaction set. This implies that we must use a cut-elimination procedure to keep formulas of the proof (on which we apply the soundness argument) simple enough to be under the scope of the partial satisfication set we can obtain. This imposes another, at the first sight serious problem. Namely, even if we had a "good hold" of proofs belonging to a cut $I(x)$, after applying the cut-elimination procedure the resulting proofs will "jump out" of this cut, since such a procedure increases lengths of proofs hyperexponentially. It is easy to see that Corollary 3.7. of [24] implies that there is a cut $J(x)$ in RCA_0 such that there is no shortening $I(x)$ of it for which $RCA_0 \vdash \forall x(I(x) \rightarrow J(2_x^x))$ holds[7]. Thus, since we do not know in general if we can shorten

[6] In fact, we cannot prove the existence of a satisfaction set for variable-free sentences of PRA which contains (codes of) closed instances of all axioms of primitive recursive definitions in PRA (including the nonstandard part of the language of PRA); otherwise, as it will be clear later, we would be able to prove the consistency of PRA in RCA_0 which is impossible since RCA_0 is a Π_1 conservative extension of PRA.

[7] As we will see later, there is a finitely axiomatizable theory such that RCA_0 is an extension by definitions of this theory. This enables us to use Corollary 3.7. of [24]; our observation then follows immediately from the fact that 2_x^x dominates 2_n^x for all $x > n$.

our cuts sufficiently, we need to find a way of dealing with proofs that are outside of any cut we might construct. Fortunately, using the right techniques, both of the above problems can be surmounted, as the next Lemma shows.

Lemma 2.4 *(The Main Lemma) There is a cut J_{RCA} in RCA s. t.*

$$RCA_0 \vdash \forall x(J_{RCA_0}(x) \rightarrow \neg Prf_{PRA}(x, \underline{1=0})).$$

Proof: The proof is quite involved and is broken into several claims. The idea underlying it is to define a kind of a truth predicate for variable-free sentences of some restrictions of the language of PRA and then prove the soundness of PRA restricted to such languages with respect to the variable-free sentences of these restrictions. To simplify our formalism, we will use nearly the same notation for codes of formulas as for the formulas themselves; what we mean will be clear from the rest of the expression. Thus in

$$S(\underline{x}) = \underline{y} \in E \leftrightarrow S(x) = y$$

$S(\underline{x}) = \underline{y}$ denotes a number that is the code of the formula built from the x^{th} and y^{th} numeral and the symbols for the successor function and equality, while $S(x) = y$ is just a formula of the language of PRA. In the rest of this chapter $\vec{x}$ stands for $(x)_0, ...(x)_{lh(x)\dot{-}1}$. Consider the following formulas[8] of RCA_0:

$$\begin{aligned}
\Theta_1(n,E) &\equiv \forall t(\mathcal{T}_{PRA}(n,t) \rightarrow \exists! y(t = \underline{y} \in E)) \\
\Theta_2(n,E) &\equiv \forall x \forall y[(S(\underline{x}) = \underline{y} \in E \leftrightarrow S(x) = y) \wedge (\underline{x} = \underline{y} \in E \leftrightarrow x = y)] \\
\Theta_3(n,E) &\equiv \forall f[\mathcal{L}_{PRA}(n,f) \rightarrow [(\mathcal{G}(f))_0 = 0 \rightarrow \forall x,y,z(Ar(f) = lh(x)+1 \\
&\quad (f(\underline{0},\vec{\underline{x}}) = \underline{z} \in E \leftrightarrow g_f(\vec{\underline{x}}) = \underline{z} \in E) \wedge (f(\underline{y+1},\vec{\underline{x}}) = \underline{z} \in E \\
&\quad h_f(\underline{y}, f(\underline{y},\vec{\underline{x}}),\vec{\underline{x}}) = \underline{z} \in E))] \wedge [(G(f))_0 = 1 \\
&\quad \forall x,y(f(\vec{\underline{x}}) = \underline{y} \in E \leftrightarrow \exists z(lh(z) = Ar(h_f) \wedge h_f(\vec{\underline{z}}) = \underline{y} \\
&\wedge \quad \forall i < lh(z)(g_i^f(\vec{\underline{x}}) = \underline{(z)_i} \in E)]] \\
\Theta_4(n,E) &\equiv \forall t \forall z[\mathcal{T}_{PRA}(n,t) \rightarrow (t = \underline{z} \in E \leftrightarrow \exists f,y,v(\mathcal{L}_{PRA}(n,f) \\
&\quad lh(y) = lh(v) = Ar(f) \wedge \forall i < lh(y)(\mathcal{T}_{PRA}(n,(y)_i) \\
&\wedge \quad (y)_i = \underline{(v)_i} \in E \wedge t = f(\vec{y}) \wedge f(\vec{v}) = \underline{z} \in E)].
\end{aligned}$$

[8] Here n is just a number variable of RCA_0 and not necessarily a standard numeral.

Thus, Θ_1 asserts that E codes a unique evaluation of each closed term of the language of PRA restricted to the functional symbols smaller than n. Θ_2 asserts that E "respects" the successor function and equality while Θ_3 says that E respects the rules of primitive recursion and composition that define new functional symbols from the previous ones. The above properties of E will imply that E correctly evaluates all standard primitive recursive functions. Θ_4 asserts that E evaluates terms inductively according to how they are built, which will imply that E evaluates correctly all standard terms. Finally let $\Theta(n,E) \equiv \bigwedge_{i<4}\Theta_i(n,E)$ and $J_{RCA_0}(n) \equiv \exists E\Theta(n,E)$.

Claim 1 *$J_{RCA_0}(n)$ defines a cut in RCA_0.*

Proof: Let the set E_0 consist of the codes of formulas of the form $S(\underline{x}) = \underline{y}$ for all x and y such that $S(x) = y$ together with the codes of all formulas of the form $\underline{x} = \underline{x}$. It is easy to see that this set is Δ_1^0 definable and thus its existence can be proved in RCA_0. Assume now that there is an E_n such that $\Theta(n,E_n)$ holds for some n; we want to show that then there is an E_{n+1} such that $\Theta(n+1,E_{n+1})$ also holds. This is obviously enough to show that $J_{RCA_0}(n)$ is a cut since the other condition is trivially satisfied. If n is not a functional symbol of $\mathcal{L}_{PRA}$ then set $E_n = E_{n+1}$; if it is a functional symbol, in order to avoid possible notational confusion let us denote it by a more standard letter f and consider the following two cases.

(1) f is defined by primitive recursion from some g and $h < f$.

(2)f is defined by composition from some $h, g_1, \ldots g_k$, such that $Ar(h) = k$ and $h, g_i < f$ (here k can be nonstandard).

Case 1. Define first

$$\begin{aligned}\overline{E_{n+1}} \;=\; & E_n \cup \{f(\underline{y},\vec{\underline{x}}) = \underline{z} : (y = 0 \wedge g(\vec{\underline{x}}) = \underline{z} \in E_n) \vee \exists v(g(\vec{\underline{x}}) = \underline{(v)_0} \in E_n \\ \wedge\; & (\forall i < y)h(\underline{i},\underline{(v)_i},\vec{\underline{x}}) = \underline{(v)_{i+1}} \in E_n \wedge (v)_y = z)\}.\end{aligned}$$

$\overline{E_{n+1}}$ is obviously definable via a Σ_1^0 formula ψ that defines the set on the right hand side of the above equation. To show that $\overline{E_{n+1}}$ has also a Π_1^0 definition in RCA_0 we first show in RCA_0 that if $\theta(n,E_n)$ holds, then

$$\forall f,y,x\exists z(\mathcal{L}_{PRA}(n+1,f) \rightarrow \psi(f(\underline{y},\vec{\underline{x}}) = \underline{z})).$$

Since by our assumption $\Theta(n,E_n)$ holds, from the definition of E_{n+1} it follows

that it also holds if $f = n$. The proof follows easily via Σ_1^0 induction on y fixing f, x, E_n as parameters, and using the properties of E_n given by $\Theta(n, E_n)$. In a similar way we can further show that

$$\forall f, y, x \exists! z(\mathcal{L}_{PRA}(n+1, f) \rightarrow \psi(f(\underline{y}, \vec{\underline{x}}) = \underline{z}));$$

here the Σ_1^0 least number principle and the properties of E_n suffice. Now, to get a Π_1^0 definition of $\overline{E_{n+1}}$ just replace $\exists v$ with $\forall v$ and the last $\wedge$ by $\rightarrow$. Thus, by Δ_1^0 Comprehension axiom $\overline{E_{n+1}}$ is a set in RCA_0. We now define E_{n+1} as the set of all formulas of the form $t = \underline{k}$ such that there are v, h, w for which $lh(v) = lh(w)$, v codes the sequence of terms from which t is inductively built and $(v)_{lh(v) \dot{-} 1} = t$; h codes the sequence of functional symbols used to build higher complexity subterms of t from some of the simpler ones and w codes a sequence of numerals that represent the values of all the subterms of t; for all $i < lh(v)$, $(v)_i$ is built from some $(v)_{j_1}, ..., (v)_{j_s}, (j_1, ..., j_s < i)$ and a functional symbol $f_i = (h)_i$ of arity s; whenever $(v)_i$ is a numeral then $(w)_i = (v)_i$; if $(v)_i = f((\vec{v})_j)$ then $(w)_i = f((\vec{w})_j) \in \overline{E_{n+1}}$; finally $(w)_{lh(w) \dot{-} 1} = \underline{k}$).

Exactly as before, E_{n+1} is Δ_1^0 definable and consequently a set in RCA_0. It is easy to check that if $\Theta(n, E_n)$ holds, then so does $\Theta(n+1, E_{n+1})$.

Case 2 is handled similarly. QED.

Claim 2 *For all standard primitive recursive functions F we have*

$$RCA_0 \vdash \forall x, y, E, k[k > \lceil F \rceil \wedge \Theta(k, E) \rightarrow (\lceil F \rceil(\vec{\underline{x}}) = \underline{y} \in E_k \leftrightarrow F(\vec{x}) = y)].$$

Also, for every standard term t of the language of PRA and any sufficiently large k

$$RCA_0 \vdash \forall E(\Theta(k, E) \rightarrow \forall x(t = \underline{x} \in E \leftrightarrow t = x)).$$

The proof is an easy mathematical induction on the complexity of F combined with an application of the Σ_1^0 least number principle and the properties of E given by $\Theta(k, E)$. More precisely, take the first primitive recursive function such that

$$\exists x, y \neg(\lceil F \rceil(\underline{x}) = \underline{y} \in E_n \leftrightarrow F(x) = y),$$

and the least x witnessing the first existential quantifier in the above formula. Now we just use the properties of E and the primitive recursive definition of F to derive a contradiction. The claim about the evaluation of terms is proved from the above by induction on complexity

It is also easy to verify that if $m > n$ and E_m and E_n are such that $\Theta(m, E_m)$ and $\Theta(n, E_n)$ then E_m and E_n "agree" about the evaluation of terms t such that $\mathcal{L}_{PRA}(n, t)$.

Claim 3 *RCA_0 proves that for all t, k, E if $\Theta(k, E)$ holds, and t is a term of PRA all of whose free variables are among $x_1, ...x_s$ and all functional symbols are $< k$, and $t_1, ..., t_s$ is a sequence of closed terms such that $\forall i < s \mathcal{T}_{PRA}(k, t_i)$, then for all $m, n_1, ...n_s$ if $(\forall i < s)t_i = \underline{n_i} \in E$ then $t(\underline{n_1}, ..., \underline{n_s}) = \underline{m} \in E$ iff $t(t_1, ..., t_s) = \underline{m} \in E$.*

Proof: An easy Π^0_1 induction on the complexity of t, using the properties of E given by $\Theta(k, E)$. QED.

We can now define a truth predicate for all *variable-free* formulas whose all functional symbols are smaller than k, providing k belongs to the cut J_{RCA_0}. Let L be the primitive recursive function mapping variable-free formulas into variable-free terms satisfying:

$$\begin{aligned} L(t_1 = t_2) &= (t_1 \dot{-} t_2) + (t_2 \dot{-} t_1) \\ L(\psi \wedge \theta) &= L(\psi) + L(\theta) \\ L(\psi \vee \theta) &= L(\psi) \cdot L(\theta) \\ & \quad \underline{1} \dot{-} L(\psi) \\ & \quad \underline{2} \end{aligned}$$

whenever x is not a variable-free formula of PRA[9]. Let $\mathcal{F}_{PRA}(n, \varphi)$ be a formalization of "φ is a variable-free sentence of PRA such that all functional symbols appearing in it are $< n$."

Definition 2.5

$$\Omega(n, E, T) \leftrightarrow (\Theta(n, E) \wedge \forall \varphi(\varphi \in T \leftrightarrow (\mathcal{F}_{PRA}(n, \varphi) \wedge L(\varphi) = \underline{0} \in E))).$$

$$\overline{\Omega}(n, T) \leftrightarrow \exists E \Omega(n, E, T)$$

Remark: Note that whenever $\Theta(n, E)$ holds, then for T defined by $T_n = \{\varphi : \mathcal{F}_{PRA}(n, \varphi) \wedge L(\varphi) = \underline{0} \in E_n\}$ we have $\Omega(n, E, T)$. Thus, whenever $J_{RCA_0}(n)$, there exists a T such that $\overline{\Omega}(n, T)$ holds.

[9] In the above equations $\dot{-}$, $+$, etc. are no

Claim 4

$$RCA_0 \vdash \forall n, T, E, \varphi, \psi [\mathcal{F}_{PRA}(n, \varphi) \wedge \mathcal{F}_{PRA}(n, \psi) \wedge \Omega(n, E, T) \rightarrow$$

$$(\varphi \wedge \psi \in T \leftrightarrow \varphi \in T \wedge \psi \in T)) \wedge (\varphi \vee \psi \in T \leftrightarrow (\varphi \in T \vee \psi \in T)) \wedge$$

$$(\neg\varphi \in T \leftrightarrow \varphi \notin T) \wedge (t_1 = t_2 \in T \leftrightarrow \exists n(t_1 = \underline{n} \in E \wedge t_2 = \underline{n} \in E)).$$

Proof of Claim 4 is an easy Π^0_1 induction using Claim 2 and Claim 3. In particular as a corollary of Claim 4, we get

Claim 5 *For each standard variable-free sentence* φ

$$RCA_0 \vdash \forall n, T(\overline{\Omega}(n, T) \wedge n > \lceil \varphi \rceil \rightarrow (\varphi \in T \leftrightarrow \varphi)).$$

Claim 6 *RCA_0 proves that for all n, T, t, m, φ such that $\overline{\Omega}(n, T)$ and φ is a quantifier-free formula whose all free variables are among $x_1, ..., x_s$ and $\forall i < s(\mathcal{T}_{PRA}(n, (t)_i) \wedge t_i = \underline{(m)_i} \in T \rightarrow (\varphi(\vec{t}) \in T \leftrightarrow \varphi(\underline{(m)_i}) \in T)$.*

Proof is an easy Π^0_1 induction on the complexity of φ using the definition of T and Claim 4.

Claim 7 *RCA_0 proves: for all n, E, T, t, θ such that $\Omega(n, E, T)$ if $\theta(\vec{x})$ is an equality axiom or an axiom of PRA which is not an induction axiom and whose all functional symbols are smaller than n and if $\forall i < lh(t)\mathcal{T}_{PRA}(n, (t)_i)$, then $\theta(\vec{t}) \in T$.*

Proof follows easily from Claim 3 and the properties of T and E given by $\Omega(n, E, T)$ and $\Theta(n, E)$.

For the next step of the proof we must switch from the more usual Hilbert-type proof system to a Gentzen-type proof system (see [37]) because we want to use a partial cut-elimination theorem. The equivalence of these two proof systems is provable in $I\Sigma_1$, and the proof transformation converting a Hilbert-style proof into a Gentzen-style proof does not change the language of the formulas of the proof. We use the following fact whose proof is effective and consequently can be formalized in $I\Sigma_1$ (PRA in fact). For a proof of it see [37], p. 116.

Theorem 2.6 *(Gentzen) There is a primitive recursive procedure for transforming any proof p in PRA of a sequent $\Gamma \Rightarrow \Delta$ into a proof p^* of the same sequent such that p^* has the following properties:*

(i) the language of p^* contains only functional symbols of the language of p;

(ii) all initial sequents of of p^* are of the form $\varphi(\vec{t})$ where $\varphi(\vec{x})$ is either a logical axiom or an equality axiom or an (open) axiom of PRA which is not an induction axiom and $\vec{t}$ is a sequence of terms of the appropriate length;

(ii) instead of the induction axioms of PRA we have the induction rule in the form

$$\frac{\Gamma, \varphi(a) \Rightarrow \varphi(a+1), \Delta}{\Gamma, \varphi(0) \Rightarrow \varphi(t), \Delta}$$

where φ is a quantifier-free formula and t is a term; a cannot appear in $\Gamma, \varphi(0), \Delta$ or t. (As it is usual in proof-theory, we use the first few letters of the alphabet to denote free variables while the last few denote the bound ones.)

(iii) all cuts of p^* are on quantifier-free formulas only.

Such proofs are called *free cut-free proofs.*

The free cut-free proofs can be much lengthier than the proofs involving arbitrary cuts, but, as we noted above, the process of partial cut elimination does not change the *language* of such proofs. Let $\mathcal{FP}_{PRA}(n, p)$ be a formalization of "p is a free cut-free proof of a quantifier-free sequent such that all formulas of the proof contain only functional symbols smaller than n".

Working within RCA_0, consider a proof p^* such that $\mathcal{FP}_{PRA}(n, p^*)$ holds for some n and let T be such that $\overline{\Omega}(n, T)$ holds. We want to associate to each such a proof p^* and T a tree of sequents $\overline{p}$ which we call a *a proof transform of p^* relative to T*, and which will contain only variable-free sentences of PRA. To do so, we move backwards through the proof p^* (i. e. from the conclusions towards the axioms). We first replace all the free variables of the conclusion of the proof by 0's throughout the proof. Assume we are at a height i. If the inference from a sequent at height $i+1$ to a sequent at height i is by a propositional rule we move one step upwards without changing anything. Since all cuts are on quantifier-free formulas and the conclusion of the proof contains no formulas with quantifiers, no quantifier rules are used in the proof. If the inference is a cut on a quantifier-free formula $\varphi(\vec{x})$, replace $\vec{x}$ by $0, 0, ..., 0$ in the entire part of the proof above and including

the cut. If it is an application of the induction rule of the form

$$\frac{\Gamma, \varphi(a) \Rightarrow \varphi(a+1), \Delta}{\Gamma, \varphi(0) \Rightarrow \varphi(t), \Delta}$$

then all variables in t must have been previously replaced by numerals and so we can assume that t is closed. Let k be the unique number such that $t = \underline{k} \in T$. Replace a by the numeral $\underline{m \dot{-} 1}$ such that[10]

$$m = \mu x \leq k(\varphi(\underline{x}) \notin T).$$

In RCA_0 such an m always exists by the Σ_1^0 least number principle applied to the formula $\varphi(\underline{x}) \notin T \vee x = k+1$.

It is easy to see that the statement "$\overline{p}$ is a proof transform of a proof p such that $\mathcal{FP}_{PRA}(n,p)$ holds" can be formalized by a Δ_1^0 formula.

The following Claim can easily be proved by Σ_1 induction on the height of the proof p.

Claim 8 *RCA_0 proves that for all p, n, T such that $\mathcal{FP}_{PRA}(n,p)$ and $\Omega(n,T)$ hold, there exists $\overline{p}$ such that $\overline{p}$ is a proof transform of p with respect to T.*

Finally we can prove our Lemma, i. e. we can show that

$$RCA_0 \vdash \forall x(J_{RCA_0}(x) \rightarrow \neg Prf_{PRA}(x, \underline{1=0})).$$

We work inside RCA_0; assume p is a proof of $1 = 0$ in PRA such that $J_{RCA_0}(p)$ holds. Let T_p be such that $\Omega(p, T_p)$ holds; such a T_p exists by our remark on page 29. We first transform this proof into a Gentzen-style proof of $\Rightarrow 1 = 0$ and then, using Theorem 2.6, we get a free cut-free proof p^* of $\Rightarrow 1 = 0$ of the *same* language as p. Thus, all functional symbols of p^* are smaller than p, and consequently $\mathcal{FP}_{PRA}(p, p^*)$ holds.

From Claim 8 it follows that there exists a proof transform $\overline{p}$ of p^* with respect to T_p. But from Claim 7 it follows that all the initial segments of $\overline{p}$ belong to T_p since they are of the form $\Theta(\vec{t})$ where $\Theta(\vec{x})$ is a logical axiom, an equality axiom or an axiom of PRA which is not an induction axiom and t is a closed term. An easy induction on height $i \leq height(\overline{p})$ shows that if all the formulas in Γ of a sequent $\Gamma \Rightarrow \Delta$ on the height i belong to T_p then there is a formula in Δ that also belongs to T_p. The induction step in proving

this claim is easy; in the case of a propositional or cut rule it follows immediately from Claim 4 while in the case of the induction rule it follows from our choice of m in defining the proof transform of a proof. More precisely, consider

$$\frac{\Gamma, \varphi(\underline{m \dot{-} 1}) \Rightarrow \varphi(\underline{m}), \Delta}{\Gamma, \varphi(0) \Rightarrow \varphi(t), \Delta}$$

Assume that all formulas in Γ and $\varphi(0)$ belong to T_p but neither $\varphi(t)$ nor any formula in Δ belong to T_p. But this implies that there is a least $x \leq t$ such that $\varphi(\underline{x}) \notin T_p$ and it must be m by our choice. Since $\varphi(0) \in T_p$ then $m \dot{-} 1 \neq m$ and $\varphi(\underline{m \dot{-} 1}) \in T_p$. But then all formulas in Γ and $\varphi(\underline{m \dot{-} 1})$ belong to T_p and no formula in Δ or $\varphi(\underline{m})$ belong to T_p which is a contradiction with the inductive hypothesis. Thus, since the final sequent is $\Rightarrow 1 = 0$, we conclude that $1 = 0$ belongs to T_p, which is a contradiction with Claim 5. This finishes our proof of The Main Lemma 2.4.QED.

We now want to show that the presence of second order objects (set) and Δ_1^0 comprehension are not essential for the above Lemma; even though they greatly contribute to the clarity of the proof, it is possible to carry out a proof of the analogous claim for the first order theory $I\Sigma_1$, just by using codes for recursive sets rather than the sets themselves. Instead of taking this approach, we use some simple interpretability results to prove the analogous result for $I\Sigma_1$.

First of all, we can extend the standard Σ_1 truth predicate of $I\Sigma_1$ (a predicate having Tarski's truth properties for Σ_1 formulas; see Theorem 6.6 of [33]) to a Σ_1^0 truth predicate of RCA_0 for formulas of the same quantifier complexity but which can have free set variables. In such formulas $x \in X$ is regarded as an atomic formula and all we have to do to obtain a Σ_1^0 truth predicate is to add to the basic clauses of the truth definition a clause for atomic formulas of this sort. Thus, for a Σ_1^0 formula[11] $Tr_{\Sigma_1^0}(e, x, y, Y)$ and a set F consisting of the axioms of PRA and a finite subset of the axioms of RCA_0 which suffices to prove that $Tr_{\Sigma_1^0}(e, x, y, Y)$ has Tarski's truth properties for Σ_1^0 formulas of RCA_0, we have that for all Σ_1^0 formulas of RCA_0

$$F \vdash \forall x, y, y_1, ..., y_k, Y, Y_1, ..., Y_s (y =< y_1, ..., y_k > \wedge Y =< Y_1, Y_2, ..., Y_s > \\ (\varphi(x, y_1, ..., y_k, Y_1, ..., Y_s) \leftrightarrow Tr_{\Sigma_1^0}(\lceil \varphi \rceil, x, y, Y))). \qquad (2.1)$$

[11] For practical purposes, instead of contracting all numerical variables of a formula into one (using coding of finite sequences), we will keep separate a distinguished number variable x and contract other variables, which will play the role of parameters, into one; we also code finite sequences of sets via $< X_1, X_2, ..., X_s >=$ [illegible]

Now adding to F the following two axioms that are special instances of the comprehension axiom and the induction axiom

$$\begin{aligned}\forall e,e',y,Y(\forall x(Tr_{\Sigma_1^0}(e,x,y,Y) &\leftrightarrow \neg Tr_{\Sigma_1^0}(e',x,y,Y)) \rightarrow \\ \exists X\forall x(x\in X &\leftrightarrow Tr_{\Sigma_1^0}(e,x,y,Y))),\end{aligned}$$

$$\begin{aligned}\forall e,y,Y(Tr_{\Sigma_1^0}(e,0,y,Y) &\wedge \forall x[Tr_{\Sigma_1^0}(e,x,y,Y)\rightarrow Tr_{\Sigma_1^0}(e,x+1,y,Y)] \\ &\rightarrow \forall x Tr_{\Sigma_1^0}(e,x,y,Y),\end{aligned}$$

we obtain an axiomatization of RCA_0 which is a finite extension of PRA. It is easy to see that the usual axiomatization or RCA_0 has proofs that are not much shorter than its version which contains only finitely many axioms besides the axiom of PRA, since for any formula φ there is a proof of 2.1 polynomial in length of φ.

Theorem 2.7 *RCA_0 is interpretable in $I\Sigma_1$ with the same domain of numbers and same primitive recursive functions.*

Proof: The predicate defining numbers will be $x = x$, all arithmetical operations and relations remain the same; the domain of sets defined by[12] $\Sigma(w) \equiv \forall y(Tr_{\Sigma_1}((w)_0,y) \leftrightarrow \neg Tr_{\Sigma_1}((w)_1,y))$ while $y \in X$ is interpreted as $Tr_{\Sigma_1}((w_X)_0,y)$. It is easy to see that under this interpretation all the basic axioms are satisfied while the Δ_1^0 comprehension becomes

$$\begin{aligned}\forall e,e',y,w(\forall x(Tr_{\Sigma_1}((w)_0,x) &\leftrightarrow \neg Tr_{\Sigma_1}((w)_1,x)) \rightarrow ((T^I(e,x,y,w) \leftrightarrow \\ \neg T^I(e',x,y,w)) &\rightarrow \exists v(\forall x(Tr_{\Sigma_1}((v)_0,x) \leftrightarrow \neg Tr_{\Sigma_1}((v)_1,x)) \wedge \\ \forall x(Tr_{\Sigma_1}((v)_0,x) &\leftrightarrow T^I(e,x,y,w)),\end{aligned} \tag{2.2}$$

where T^I is the interpretation of the formula $Tr_{\Sigma_1^0}(e,x,y,Y)$. But this is a theorem of $I\Sigma_1$ for the following reason. $Tr_{\Sigma_1^0}(e,x,y,Y)$ is a Σ_1^0 formula; thus if we push all the negations in T^I inside until we reach either a formula of the form $Tr_{\Sigma_1}((w)_0,x)$ obtained by replacing $x \in X$ or otherwise an atomic formula (whichever comes first), and then replace all formulas $\neg Tr_{\Sigma_1}((w)_0,x)$ by $Tr_{\Sigma_1}((w)_1,x)$ we get a formula Ψ that is in $I\Sigma_1$ equivalent

[12] We will use v, w for variables obtained by replacing set variables in the interpretations of second order formulas, indexing them (when needed) by the second order variables which they interpret.

to a Σ_1 formula[13]. On the other hand, if we push all negations in $\neg T^I$ inside in the above manner,and then replace all formulas $Tr_{\Sigma_1}((w)_0, x)$ that do not have a negation in front of them by $\neg Tr_{\Sigma_1}((w)_1, x)$ and then further push negations inside, we get a formula equivalent to a Π_1 formula Θ, for similar reasons as before. If the first two equivalences in 2.2 are true, then $\Psi \leftrightarrow \Theta$. Replace now in Ψ and Θ e, e', y and w by the corresponding (possibly non-standard) numerals to get (Gödel codes of possibly non-standard) sentences $\overline{\Psi}$, $\overline{\Theta}$. Then we can take $v = <\overline{\Psi}, \overline{\Theta}>$.

To prove that $I\Sigma_1$ proves the interpretation of $I\Sigma_1^0$ induction, again push inside the negation in T^I (as above) and replace $\neg Tr_{\Sigma_1}((w)_0, x)$ by $Tr_{\Sigma_1}((w)_1, x)$ to get a Σ_1 formula equivalent to T^I whenever $\forall x(Tr_{\Sigma_1}((w)_0, x) \leftrightarrow \neg Tr_{\Sigma_1}((w)_1, x))$.

Corollary 2.8 *There is a cut $J_{I\Sigma_1}$ in $I\Sigma_1$ such that*

$$I\Sigma_1 \vdash \forall x(J_{I\Sigma_1}(x) \rightarrow \neg Prf_{PRA}(x, \underline{1=0})).$$

Proof: By Lemma 2.4 there is a cut J_{RCA_0} in RCA_0 s.t.

$$RCA_0 \vdash \forall x(J_{RCA_0}(x) \rightarrow \neg Prf_{PRA}(x, \underline{1=0}))$$

Thus,

$$I\Sigma_1 \vdash \forall x(J^I_{RCA_0} \rightarrow \neg Prf_{PRA}(x, \underline{1=0})).$$

It is easy to see that the fact that J_{RCA_0} is a cut in RCA_0 implies that $J^I_{RCA_0}$ is a cut in $I\Sigma_1$; just use the fact that the interpretation I neither changes numbers nor arithmetical operations and relations. Thus, we can take $J_{I\Sigma_1} = J^I_{RCA_0}$. QED.

In fact, there is an RCA_0 proof of

$$\forall x(J_{I\Sigma_1}(x) \rightarrow \neg Prf_{PRA}(x, \underline{1=0}) \tag{2.3}$$

which is *conceptually* much clearer than the $I\Sigma_1$ proof of the same sentence; it is based on the fact (which is easy to prove) that $RCA_0 \vdash \forall x(J_{I\Sigma_1}(x) \rightarrow J_{RCA_0}(x))$ and is conceptually clearer because the RCA_0 proof of $\forall x(J_{RCA_0}(x) \rightarrow \neg Prf_{PRA}(x, \underline{1=0})$ uses naturally

[13]The reason for this is that all bounded quantifiers can be exchanged with the $\exists$ quantifier using the following bounding principle for Δ_0^0 formulas φ:

$$\forall x < a \exists y \varphi(x, y) \rightarrow \exists b \forall x < a \exists y < b \varphi(x, y)$$

defined sets rather than cumbersome coding used in the $I\Sigma_1$ proof. The sentence 2.3 is the sentence we referred to on page 35 of the Introduction. Of course this sentence is not a Π_1 sentence, but the above fact gives us reasonable grounds to conjecture that RCA_0 can indeed bring conceptual facilitation for proofs of finitistically meaningful sentences as well (i.e. Π_1 sentences of PRA). As a Corollary we get the main theorem of this Chapter.

Theorem 2.9 *$I\Sigma_1$ has a speed-up over PRA for variable-free sentences of PRA which is not Kalmar-elementary.*

Proof: Immediate from Lemma 2.4, Corollary 2.3 and the fact that $Con_{PRA}(2^0_n)$ can be formalized as a quantifier-free sentence of the language of PRA by the usual replacement of bounded quantifiers with their primitive recursive substitutes. QED.

Chapter 3

RCA_0 versus $I\Sigma_1$

We now want to show that RCA_0 has at most polynomial speed-up over PRA. In fact, we will show that with the proof system we chose the speed-up is at most linear. The exact speed-up seems to depend on the particular proof system one uses, and for this reason we must go into the details of proof systems. On the other hand these differences in the speed-up can only be minor since the same method of proof that we will use with our proof system can be used with minor modifications to show that for any other Hilbert-type proof system the speed-up is at most quadratic.

To make our proof a little simpler we choose a Hilbert-style proof system slightly different than the one which is most commonly used. It is more efficient than the standard one since it allows us to use the rule of universal generalization for several variables simultaneously and its axioms allow simultaneous instantiations of universal formulas by a sequence of terms. It is convenient to change our notation and use $\forall\vec{x}\varphi(\vec{x})$ for $\forall x_1\forall x_2...\forall x_k\varphi(x_1,x_2,...,x_k)$ with the assumption $\vec{x}=x_1,x_2,...,x_k$. The logical axioms for the proof system we use for $I\Sigma_1$ are:

(1) all tautologies (i.e. all first order formulas obtained from a tautology by replacing propositional letters with some first order formulas corresponding to each letter);

(2) $\forall\vec{x}\alpha(\vec{x})\rightarrow\alpha_{\vec{x}}(\vec{t})$, where each t_i of $\vec{t}$ is is substitutable for x_i in α;

(3) $\forall\vec{x}(\alpha\rightarrow\beta(\vec{x}))\rightarrow(\alpha\rightarrow\forall\vec{x}\beta(\vec{x}))$, where none of x_i of $\vec{x}$ is free in α;

(4) $x=x$;

(5) $x=y\rightarrow(\alpha\rightarrow\alpha')$, where α is atomic and α' is obtained from α by replacing some free occurrences of x in α by y.

the logical axioms of the proof system that we use for $I\Sigma_1$ in the following way.

In the tautologies propositional letters can be replaced by the second order formulas as well (i.e. formulas involving second order variables and second order quantifiers). In the axioms of group 2 not only that α can be an arbitrary formula of the language of RCA_0 but we also have the second order version of this axiom, i.e.

$$\forall \vec{X} \alpha(\vec{X}) \rightarrow \alpha_{\vec{X}}(\vec{Y})$$

where each Y_i of $\vec{Y}$ is substitutable for X_i of $\vec{X}$ in α; note that besides the second order variables we have no other second order terms. The axioms of group (3) are changed in a similar way also adding versions with the sequence of second order universal quantifiers. To the groups (4) and (5) for the first order logic we add the corresponding second order additions:

$$X = X$$

$$X = Y \rightarrow (x \in X \rightarrow x \in Y)$$

$$X \quad Y \rightarrow (X = Z \rightarrow Y = Z)$$

$$x = y \rightarrow (x \in X \rightarrow y \in X)$$

In both theories we have as a rule of inference Modus Ponens (MP):

$$\text{from } \alpha \text{ and } \alpha \rightarrow \beta \text{ infer } \beta.$$

In $I\Sigma_1$ we have another rule of inference called Generalization (GEN):

$$\text{from } \alpha(\vec{x}) \text{ infer } \forall \vec{x} \alpha(\vec{x}),$$

while in RCA_0 besides this rule which we will call the *First Order Generalization Rule* we also have its second order version which we will call the *Second Order Generalization Rule*:

$$\text{from } \varphi(\vec{X}) \text{ infer } \forall \vec{X} \varphi(\vec{X}).$$

Proofs are sequences of formulas such that every formula of the sequence is either an axiom or is obtained from some previous ones by one of the rules of inference. Note that we do not require that axioms must be sentences and that we do not treat $\exists x$ as a primitive logical symbol; it is rather an abbreviation for $\neg \forall x \neg$.

Given an interpretation of the language of a formula φ in a theory T, an interpretation of the formula $\varphi(x_1, ..., x_k)$ (where $x_1, ..., x_k$ are exactly all free variables of φ

in the alphabetical order) is usually defined to be the formula $\varphi^I \equiv U(x_1) \to (U(x_2) \to (...(U(x_k) \to \varphi^*)...)$; here φ^* stands for the formula obtained from φ by replacing the predicate letters of φ by the corresponding formulas given by the interpretation of the language of φ and by relativizing quantifiers to the domain of interpretation $U(x)$, respecting the logical connectives. More precisely, φ^* is defined by induction on the complexity of φ. If φ is of the form $\theta_1 \wedge \theta_2, \theta_1 \vee \theta_2$ or $\neg\theta$, then φ^* is $\theta_1^* \wedge \theta_2^*, \theta_1^* \vee \theta_2^*$ or $\neg\theta^*$ respectively. If θ is of the form $\forall x \theta_1(x)$, then θ^* is $\forall x(U(x) \to \theta_1^*(x))$ (see [3]).

In order to make our proof simpler, we will slightly change this definition of interpretation. Given an interpretation of a language $\mathcal{L}$ in a theory T, the new definition of interpretation associates to each formula φ of $\mathcal{L}$ a formula $\varphi^{\mathcal{I}}$ for which it is easy to see that it is equivalent to the standard interpretation φ^I, but is more convenient for our purpose.

Definition 3.1 *Let $\mathcal{L}$ be a language, T a theory and $\overline{I}$ an interpretation of the language $\mathcal{L}$ in T. For each formula φ we define a formula $\varphi^{\mathcal{I}}$ to be the formula $\bigwedge_{i<k} U(x_i) \to \varphi^{\#}$ where $x_1, x_2, ..., x_k$ are exactly the free variables of $\varphi^{\#}$ and where $\varphi^{\#}$ is defined inductively as follows:*

1. *if φ is an atomic formula then $\varphi^{\#}$ is obtained in the same way as φ^*, by replacing the predicate letters of φ with the corresponding formulas given by the interpretation of the language $\mathcal{L}$ in T;*
2. *if φ is of the form $\theta_1 \wedge \theta_2, \theta_1 \vee \theta_2$ or $\neg\theta$, then $\theta^{\#}$ is $\theta_1^{\#} \wedge \theta_2^{\#}, \theta_1^{\#} \vee \theta_2^{\#}$ or $\neg\theta_1^{\#}$ respectively.*
3. *if φ is of the form $\forall \vec{x} \theta(\vec{x})$ where θ does not begin with a universal quantifier, then $\varphi^{\#}$ is $\forall \vec{x}(\bigwedge_{x_i \in \vec{x}} U(x_i) \to \theta^{\#}(\vec{x}))$.*

Lemma 3.2 *There is a natural number c such that the interpretation $\varphi^{\mathcal{I}}$ of any logical axiom φ of RCA_0 given by the interpretation of the language of RCA_0 in the theory $I\Sigma_1$ from the Theorem 2.7, has a proof from the logical axioms of $I\Sigma_1$ of length bounded by $c|\varphi|$.*

Proof: Assume $\Psi(\vec{x}, \vec{X})$ is a tautology; then $\Psi^{\mathcal{I}}$ is of the form $\bigwedge_{i<k} \Sigma(w_i) \to \Psi^*(\vec{x}, \vec{w})$ which is again a tautology; its length is linear in the length of Ψ because the number of free variables of Ψ is bounded by the length of Ψ. If Ψ is an axiom of group 2 then either its interpretation is again an axiom of group 2 in the case when the universal quantifier is a number quantifier, or, if it is of the form $\forall \vec{X} \alpha(\vec{X}) \to \alpha_{\vec{Y}}(\vec{Y})$, then its interpretation is of the form $\bigwedge_{w_i \in \vec{w_Y} \cup \vec{w_Z}} \Sigma(w_i) \to [\forall \vec{w_X}(\bigwedge_{w_j \in \vec{w_X}} \Sigma(w_j) \to \alpha^{\#}(\vec{w_X})) \to \alpha^{\#}_{\vec{w_X}}(\vec{w_Y})]$, where

the sequence of variables $\vec{w_Z}$ corresponds exactly to all free second order variables $\vec{Z}$ of α not among $\vec{Y}$. But it is easy to see that this is a consequence of the axiom of type 2 $\forall \vec{w_X}(\bigwedge_{w_j \in \vec{w_X}} \Sigma(w_j) \rightarrow \alpha^{\#}(\vec{w_X})) \rightarrow (\bigwedge_{w_k \in \vec{w_Y}} \Sigma(w_k) \rightarrow \alpha^{\#}(\vec{w_Y}))$, with a proof of length linear in $|\forall \vec{X} \alpha(\vec{X}) \rightarrow \alpha_{\vec{X}}(\vec{Y})|$. In a similar way one can show that the claim holds for the axioms of other groups as well. QED.

Definition 3.3 *Let T be a theory; we call a sequence of formulas $< \varphi_1, ..., \varphi_n >$* a proof of φ_n in T from the hypotheses $\theta_1, ..., \theta_m$ *if $< \theta_1, ..., \theta_m, \varphi_1, ..., \varphi_n >$ is a proof of φ_n in $T \cup \{\theta_1, ..., \theta_m\}$.*

Lemma 3.4 *There is a natural number d such that for any proof $p =< \varphi_1, ..., \varphi_n >$ in RCA_0 there is a proof $\overline{p}$ in $I\Sigma_1$ such that:*

1. *$\overline{p}$ is of the form $< \underbrace{...., \varphi_1^{\mathcal{I}}}_{b_1}, \underbrace{...., \varphi_2^{\mathcal{I}}}_{b_2}, ..., \underbrace{...., \varphi_{n-1}^{\mathcal{I}}}_{b_{n-1}}, \underbrace{...., \varphi_n^{\mathcal{I}}}_{b_n} >$, where all blocks b_i are disjoint and each block b_i ends with the formula $\varphi_i^{\mathcal{I}}$.*
2. *If φ_i is an axiom of RCA_0 (logical or one of the finitely many nonlogical ones) then b_i is a proof of $\varphi_i^{\mathcal{I}}$ in $I\Sigma_1$ and $|b_i| \leq d|\varphi_i|$.*
3. *If φ_i is obtained by the Modus Ponens from some φ_j and some φ_k which is of the form $\varphi_j \rightarrow \varphi_i$, $(j, k < i)$, then b_i is a proof of $\varphi_i^{\mathcal{I}}$ from the hypotheses $\varphi_i^{\mathcal{I}}$ and $\varphi_k^{\mathcal{I}}$, and $|b_i| \leq d|\varphi_k|$.*
4. *If φ_i is a formula of the form $\forall \vec{x} \varphi_j(\vec{x}, \vec{y}, \vec{Y})$ obtained by the First Order Generalization Rule from the formula $\varphi_j(\vec{x}, \vec{y}, \vec{Y})$, or a formula of the form $\forall \vec{X} \varphi_j(\vec{y}, \vec{X}, \vec{Y})$ obtained by the Second Order Generalization Rule from the formula $\varphi_j(\vec{y}, \vec{X}, \vec{Y})$, then b_i is a proof of $\varphi_i^{\mathcal{I}}$ from the hypothesis $\varphi_j^{\mathcal{I}}$ and $|b_i| \leq d|\varphi_i|$.*

Proof: From our proof it will be clear how to choose the number d. We proceed by induction on the number of formulas in the proof p; case $n = 0$ is trivial. Assume that our claim holds for all proofs with n formulas; let $p =< \varphi_1, ..., \varphi_n, \varphi_{n+1} >$ and denote $< \varphi_1, ..., \varphi_n, >$ by p'. By our inductive hypothesis there is a proof $\overline{p'} =< \underbrace{...., \varphi_1}_{b_1}, ..., \underbrace{...., \varphi_n}_{b_n} >$ satisfying 1-4.

Case 1. φ_{n+1} is an axiom. Then let $\overline{p}$ be the proof obtained by adding to the proof $\overline{p'}$ block b_{n+1} which is a proof of $\varphi_{n+1}^{\mathcal{I}}$ in $I\Sigma_1$. Then if d is chosen larger than c from the previous Lemma, $|b_{n+1}| \leq d|\varphi_{n+1}|$.

Case 2. φ_{n+1} is obtained from some φ_i and φ_k of the form $\varphi_i \to \varphi_{n+1}$. Let $\vec{X}$ and $\vec{Y}$ be the sequences of all free second order variables of φ_i and φ_{n+1} respectively; let $\vec{Z}$ be the sequence of all free second order variables that appear in either $\vec{X}$ or $\vec{Y}$ (these two sequences can overlap). We assume that all sequences are arranged in the alphabetical order. Then $\varphi_i^{\mathcal{I}}$ and $\varphi_k^{\mathcal{I}}$ are respectively of the form

$$\bigwedge_{w_i \in \vec{w}_X} \Sigma(w_i) \to \varphi_i^{\#}(\vec{w}_X) \tag{3.1}$$

$$\bigwedge_{w_i \in \vec{w}_Z} \Sigma(w_i) \to (\varphi_i^{\#}(\vec{w_X}) \to \varphi_{n+1}^{\#}(\vec{w_Y})) \tag{3.2}$$

We describe informally block b_{n+1}. It starts with the tautology needed to derive

$$\bigwedge_{w_i \in \vec{Z}} \Sigma(w_i) \to \varphi_i^{\#}(\vec{w}_X) \tag{3.3}$$

from formula 3.1 using one application of MP. After formula 3.3 a tautology follows needed to derive

$$(\bigwedge_{w_i \in \vec{w}_Z} \Sigma(w_i) \to \varphi_i^{\#}(\vec{w}_X)) \to (\bigwedge_{w_i \in \vec{w}_Z} \Sigma(w_i) \to \varphi_{n+1}^{\#}(\vec{w}_Y)) \tag{3.4}$$

from 3.2. After the formula 3.4 in b_{n+1} we have

$$\bigwedge_{w_i \in \vec{Z}} \Sigma(w_i) \to \varphi_{n+1}^{\#}(\vec{w}_Y)$$

which can be obtained by MP from the previous formula and formula 3.3. Then a couple of formulas follow needed to derive the following tautological consequence of the last formula:

$$(\bigwedge_{w_i \in \vec{Z} \setminus \vec{Y}} \Sigma(w_i)) \to (\bigwedge_{w_i \in \vec{w}_Y} \Sigma(w_i) \to \varphi_{n+1}^{\#}(\vec{w}_Y))$$

Let $\vec{V} = \vec{Z} \setminus \vec{Y}$; the next formula of b_{n+1} is

$$\forall \vec{w}_V [(\bigwedge_{w_i \in \vec{V}} \Sigma(w_i)) \to (\bigwedge_{w_i \in \vec{w}_Y} \Sigma(w_i) \to \varphi_{n+1}^{\#}(\vec{w}_Y))]$$

and it can be obtained from the previous formula using *one single* application of the Generalization Rule of *our system*[1]. Let e be a code for the empty set in $I\Sigma_1$; then there is

[1]This is the point at which we need our Generalization rule that allows simultaneous quantification over several variables and our axioms which allow simultaneous treatment of several quantifiers. Without it, in the standard system, we would have to repeat applications of the Generalization Rule and axioms of type 2 and 3 several times depending on the number of variables of formulas involved. This is why we get only quadratic bound for the speed-up if we use the standard system. It is easy to see that the same technique gives a linear bound if we write proof in the form of trees or if we have an extra rule of the form $\frac{\varphi(\vec{x})}{\varphi(\vec{t})}$, where $\vec{t}$ is a sequence of closed terms.

a proof in $I\Sigma_1$ of $\Sigma(e)$. Using an instance of an axiom of group 2 and an application of Modus Ponens we can get

$$(\bigwedge_{w_i \in \vec{Z} \setminus \vec{Y}} \Sigma(e)) \rightarrow (\bigwedge_{w_i \in \vec{w}_Y} \Sigma(w_i) \rightarrow \varphi^{\#}_{n+1}(\vec{w}_Y))$$

In a few more steps, using only tautologies and MP, we can collapse all identical conjuncts $\Sigma(e)$ into one, add the proof of $\Sigma(e)$ in $I\Sigma_1$ and get $\varphi^{\mathcal{I}}_{n+1}$. Note that the number of formulas in b_{n+1} does not depend on particular formulas involved but only on the fact that b_{n+1} was obtained by Modus Ponens, and that the length of each formula in b_{n+1} is bounded by $s|\varphi_k|$, for a sufficiently large number s. Thus, if d is chosen large enough, $|b_{n+1}| \leq d|\varphi_k|$.

Case 3: φ_{n+1} is obtained from some φ_i using one of the two generalization rules. In this case the proof proceeds in a similar manner as in the previous case, using Generalization Rule and axioms of the groups 2 and 3 to distribute quantifiers and subformulas of the form $\Sigma(w_i)$ in the appropriate way. We omit the details, which are easy to supply. QED.

Theorem 3.5 *RCA_0 has at most a linear speed-up over $I\Sigma_1$ with respect to the arithmetical formulas and with the above proof systems.*

Proof: Assume $p =< \varphi_1, ..., \varphi_n >$ is a proof in RCA_0 of an arithmetical formula $\varphi \leq \varphi_n$ such that no proper subsequence of p is still a proof of φ. In the previous Lemma we defined for each formula φ_i of p a block b_i of formulas of $I\Sigma_1$ such that $|b_i| \leq d|\varphi_{b_i}|$ and such that

(i) if φ_i is an axiom or if it is obtained by either of our two Generalizations Rules, then φ_{b_i} is φ_i;

(ii) if φ_i is obtained by Modus Ponens from φ_j and $\varphi_k \equiv \varphi_j \rightarrow \varphi_i$, then $\varphi_{b_i} = \varphi_k$. Obviously for any formula φ_k of the form $\varphi_j \rightarrow \varphi_i$ there is at most one application of Modus Ponens in which φ_k could have been used as the implicative premise of the this rule[2]. Thus, any such formula φ_k could have been assigned to at most two blocks: φ_k could only be φ_{b_k} or φ_{b_i}, for i as above. In both cases every formula of the proof p was associated to at most two blocks, and for each block b_i , $|b_i| \leq d|\varphi_{b_i}|$. This obviously implies that for $\overline{p} =< b_1, ..., b_n >$, $|\overline{p}| \leq 2d|p|$. QED.

To get a similar result for WKL_0 versus RCA_0[3] one could try to adapt Harring-

[2]This is not true of φ_i, but this does not matter here.

[3]This we need in order to be able to give a philosophical argument about our understanding of the concept of sets satisfying compactness principle similar to the one we gave on page 13 about our understanding the concept of sets satisfying recursive comprehension.

ton's forcing proof of the conservativity of WKL_0 over RCA_0 for arithmetical sentences such that the proof produces in fact an interpretation of WKL_0 in RCA_0 which does not change the domain of numbers. We do know that WKL_0 is interpretable in RCA_0 but with a changed domain of numbers which prevents us from using it to get the corresponding linear speed-up result. This proof is based on a formalization of a proof-theoretic procedure of Sieg's in $I\Delta_0 + exp$ and a result of Friedman's relating interpretability with the cut-free consistency of theories. We hope that an application of Boolean valued models will help get the right kind of interpretation, but the work is still under way.

Chapter 4

The consistency problem and the ω rule

We now turn to the consistency problem. As we noted in the Introduction, the consistency of any of the theories we consider is not provable by purely finitistic means (we again accept Tait's Thesis). In order to overcome this difficulty, one could accept more powerful methods in consistency proofs as perhaps non-finitistic, but in some sense still constructive: transfinite induction up to ω^ω applied only to primitive recursive predicates suffices to prove the consistency of PRA (see [37], p. 116), while transfinite induction up to ε_0 suffices to prove the consistency of PA. Unfortunately, the epistemological value of such a consistency proof seems more doubtful than the value of a purely finitistic consistency proof; transfinite induction is, after all, a transfinite method that does not seem justifiable without a reference to some completed infinite totalities.

Another way out would be to accept more general forms of logical inference in consistency proofs as legitimate. Hilbert himself took such a path in [16] with another aim: obtaining a complete and consistent system of arithmetic, most likely in response to Gödel's First Incompleteness Theorem (see [5] for a detailed discussion). In this paper Hilbert added to a standard form of the first order arithmetic Z an informal rule of the following form [1]:

Restricted ω-rule 2 *If $A(x)$ is a quantifier free formula for which the following is finitarily shown:* **$A(z)$ is a correct numerical formula for each particular numerical instance** *z, then its universal generalization can be taken as a new premise in all further proofs.*

[1] See [5]; for the original formulation in German see [16], p. 491.

He denoted this semi-formal system by Z^* and went on to show certain completeness properties of Z^* with respect to Π_1 sentences.

As Feferman mentions in [5], the system Z^* is not at all in the spirit of Hilbert's original ideas: instead of having a precisely formalized system in which rules of inference are purely combinatorially described, here a vaguely formulated rule of inference is used *within the system* Z^*. The vagueness is because the rule depends on what is accepted as a finitistically valid proof.

Nevertheless, since the ultimate, finitistic consistency proof of any reasonably strong theory is not possible, it is worth considering a rule of this sort *as long as we can make it precise and then find some justification for it.* Thus, in the rest of this section we consider various rules obtained from the Hilbert's informal rule by interpreting differently what the word "finitarily" in its description means. We investigate grounds on which one can justify such a rule and what the theories are whose consistency can be proven using it.

As we noted before, if a proposition of the above form is provable using only finitistic means[2], then, assuming Tait's Thesis, there is a proof $p(x)$ in PRA of $F(x)$. This implies that all instances $F(\underline{n})$ of $F(x)$ have uniform proofs: there exists a finite skeleton proof $p(x)$ such that for each n, $p(\underline{n})$ is a finitistic proof of $F(\underline{n})$. (Here, of course, the finiteness of the skeleton comes from the fact that the proof is formalizable in a first order theory). One can argue that it is not necessary to require this uniformity. Perhaps just having an effective, explicitly given construction[3] producing for each individual n a finitistic proof of $F(\underline{n})$ would suffice to accept $\forall x F(x)$ as finitistically proven on the basis of an ω-rule in which the only restriction to make it "finitary" is that the finitistic proofs of each instance are produced using an effective construction. In particular, according to this view, to prove the consistency of a theory T it would be enough to exhibit an effective construction which produces for each n a finitistic proof of the fact that n does not code a proof of an inconsistency in T (with the standard first order logic). The last claim seems to be supported in the following quotation from Detlefsen [2]:

> "In this section I would like to sketch an argument against the claim that G2 implies the failure of the Hilbert's Program for finding a finitistic consistency

[2]Here we assume that a finitist manipulates only numbers (actually token numbers like strokes). As noted in [36], since the syntax of the first order language is primitive recursively encodable, to prove finitistically that formulas or proofs of a primitive recursively axiomatized theory T have a (finitistically meaningful) property amounts to proving finitistically that the primitive recursive predicate corresponding to that property is true of the corresponding codes.

[3]This term is used in Tait [36]; Detlefsen [2] uses the term "manual" to denote the same thing.

proof for the various theories of classical mathematics. The central claim of the argument is that $Con(T)$, the consistency formula shown to be unprovable by G2, does not really 'express' consistency in the sense of that term germane to an evaluation of Hilbert's Program. In order for a consistency formula to 'express' consistency in the appropriate sense the quantifiers and operators in it must be construed finitistically, and *not* classically, since it is the finitistic consistency of a classical system that is at issue. But a finitistic interpretation of the universal quantifier would seem to differ drastically from a classical interpretation of it, as is clear from the following remark of Herbrand."

At this point Detlefsen gives the following quotation of Herbrand taken from Goldfarb's [12], pp. 288-9, footnote 5:

"...when we say that an argument (or theorem) is true for all (these) x, we mean that, for each x taken by itself, it is possible to repeat the general argument in question, which should be considered to be merely the prototype of these particular arguments."

Then Detlefsen continues:

"And, again, he says that a proof of a universal claim is merely a description or manual of operations which are to be executed in each particular case ([12], pp. 49-51). This view of the universal quantifier would seem to sponsor the following restricted ω-rule: if I have an effective procedure P (i.e., a manual of operations P) for showing of each individual n that '$F(\underline{n})$' is finitistically provable, then '$\forall x F(x)$' is also finitistically provable. Indeed in a 1930 paper ([16], pp. 49-51.), Hilbert stated a rule something like this. And at that time it was apparent to finitists that the rule did not give one the power to go beyond the means of some methods that had already been accepted as finitistic ([12], p. 297). Now one would not, in general, want to add the abovementioned ω-rule to a scheme designed to serve as the finitistic proof theory of the classical theory T, since that rule does not constitute a truth of the finitistic proof theory of the classical T! Still, certain instances of the rule would seem to be called for; in particular the one producing $Con(T)$ from its instances. This addition made, $Con(Z)$ becomes provable in Z_{ω^*} (= Z plus the above-mentioned instance of the restricted ω-rule)."

We now investigate the claims made above. They are essentially the following ones.

Detlefsen's Claim 1 *Since it <u>does not</u> constitute a truth of the finitistic proof theory of the classical T, we <u>cannot</u> accept the following ω-rule: any formula of the form $\forall x F(x)$ can be derived whenever we have an effective procedure for producing a finitistic proof of $F(\underline{n})$ for each n.*

Detlefsen's Claim 2 *Nevertheless, we can deduce $Con(T)$ whenever we have an effective procedure for producing a finitistic proof of $\neg Prf_T(\underline{n}, \bot)$ for each n.*

As Shoenfield [27] shows, using the rule mentioned in claim 1, if we do not put any restrictions on the complexity of $F(x)$, any true formula of arithmetic becomes provable; hence, it is hard to believe that one can find a finitistic justification for applying such a rule. But I find Detlefsen's reason for refuting it unclear; for what does it mean that the above rule "does not constitute a truth of the finitistic proof theory of the classical T"? If we start with true axioms for T (meaning true on the natural numbers), then obviously the above rule always yields true conclusions from these true premises. On the other hand, it cannot be that we refuse it because it is not conservative over the standard first order logic, since the rule accepted in the second claim is also not conservative. It seems to me that the ω-rule from the first claim should be acceptable on the same grounds as the ω-rule from the second claim as long as $F(x)$ is a primitive recursive predicate. For it is hard to think of any special property of $\neg Prf_T(x, \bot)$ that would justify an application of the ω-rule on it, and that is not shared by any other primitive recursive predicate.

We will now argue that the second claim is unacceptable as formulated, and point out a modification of it that leads to a partial program which encompasses all the theories we mentioned in the Introduction. We also show that under the same kind of restriction the restricted ω-rule mentioned in Claim 1 is indeed acceptable on the same grounds and is actually of the same power as the restricted version of the ω-rule from Claim 2, as long as consistency proofs are considered.

To refute the second claim, consider, for example, the theory ZF and the following explicitly given primitive recursive function f:

$$f(n) = \text{the least code of a finitistic proof that } n \text{ does not code a proof of an inconsistency in } ZF \text{ with the standard first order logic.}$$

Assuming that ZF is consistent, for every sequence $\varphi_0, \varphi_1, ..., \varphi_k$ of formulas of ZF such that φ_k is an inconsistency, one can find a formula φ_i in that sequence that is neither an axiom of ZF nor is obtained from the previous ones using a first order rule of inference. For every such sequence this fact itself has a "short" finitistic proof because ZF is primitive recursively

axiomatized and so it is easy to verify that φ_i is not an axiom of ZF, while to show that φ_i is not derived from some of the previous formulas one has to check only a couple of rules of inference when applied only to some of $\varphi_0, ..., \varphi_{i-1}$. Thus, $f(x)$ is total and the μ-operator in its definition can be bounded by a primitive recursive function. Consequently, $f(x)$ is primitive recursive itself, and it satisfies Detlefsen's requirement for proving the consistency of ZF. But no finitist can take this as a satisfactory proof of the consistency of ZF because he even cannot realize that the "manual" f indeed produces for each n the necessary proof: in the above discussion we *assumed* that ZF is consistent. Hence, Detlefsen's second claim is untenable as formulated. In order to accept the consistency of a theory T on grounds that can be *reasonably* viewed as "close" to finitistic grounds, we have at least to be able to give a *finitistic proof* that our "manual" indeed has the above property. This leads us to introduce the following metadefinitions and to consider the following two questions.

Metadefinition 2 *We say that the consistency of a primitive recursively axiomatized theory T is provable* **almost finitistically** *if there exists a finitistically acceptable function f of which it can be finitistically proven that for each natural number n, $f(n)$ is a (Gödel code of a) finitistically acceptable proof that n is not a (code of a) proof of a contradiction from the axioms of T with the standard first order logic.*

Thus, one can see the consistency of such a theory as being derived from all the instances $\neg Prf_T(\underline{n}, \bot)$ using an ω-rule that is in certain sense *finitistically warranted.*

Question 4 *What are the theories whose consistency can be proven almost finitistically?*

There is no reason to restrict the above kind of ω-rule only to the consistency formulas; with equal justification we can apply such a rule to any sentence of the form $\forall x \varphi(x)$, for φ a primitive recursive formula, and so we introduce the following metadefinition.

Metadefinition 3 *A sentence of the form $\forall x \varphi(x)$, where φ is primitive recursive*[4]*, is provable* **almost finitistically** *if there exists a finitistically acceptable function f of which it can be finitistically proven that for all n, $f(n)$ is a finitistically acceptable proof of $\varphi(\underline{n})$.*

Definition 4.1 *We denote by S the set of all Π_1 sentences which are provable almost finitistically.*

[4]Detlefsen does not explicitly impose this complexity restriction; nevertheless this restriction follows from what Hilbert accepted as finitistically meaningful sentences and is present in Hilbert's paper [16] that Detlefsen quotes.

Again, one can see S as the set of all sentences provable using one application of a finitistically warranted ω-rule. This rule can be seen as a strengthened combination of the rules mentioned in [5][5], pages 212 and 213: not only do we restrict the complexity of the formula on which the rule is applied and require that the fact *"every instance of the formula is finitistically provable"* must be itself finitistically provable, but we also require that the proofs of all these instances must be generated by a finitistically acceptable function. These restrictions eliminate the vagueness from Hilbert's description of the informal ω-rule. We can adjoin to this form of the ω-rule the standard first order logic to obtain a well defined formal system. Further, the form of such an ω rule permits us to replace it with just the set S of its conclusions and get a purely first order theory axiomatized by S.

Note that induction axioms of PRA can be written in the form $\psi(0) \wedge (\forall y < x)(\psi(y) \rightarrow \psi(y+1)) \rightarrow (\forall y \leq x)\psi(y)$, for ψ a primitive recursive formula, which is itself a primitive recursive formula. All other axioms of PRA obviously are primitive recursive formulas. Since for any primitive recursive formula $\varphi(x)$ provable in PRA with a proof $p(x)$, the function $f(x)$ given by $f(x) = \lceil p(\underline{x}) \rceil$ satisfies the condition from the definition of the set S, the following Lemma immediately follows.

Lemma 4.2 $PRA \subseteq S$.

Thus, S extends the standard finitist reasoning; we want to know if S is substantially stronger than PRA in proving the consistency of theories.

Question 5 *What are the theories whose consistency can be proven on the basis of S within the standard first order reasoning?*

In the next sections of this chapter we answer the above two questions and consider some possible generalizations.

4.1 Theories whose consistency can be proved almost finitistically

We now want to characterize theories whose consistency can be proven almost finitistically. Tait's Thesis implies that one can give such a consistency proof for T if and

[5] I am grateful to professor Solomon Feferman for sending me this reference in a reply to my paper [18].

only if there exists a function f defined by primitive recursion such that

$$PRA \vdash \forall x Prf_{PRA}(f(x), \lceil \neg Prf_T(\underline{x}, \bot) \rceil).$$

As is well known, using the standard cut-elimination procedure one can show that the provably recursive functions in PRA are precisely functions definable by primitive recursion (the procedure explicitly produces a definition of a primitive recursive function). Thus we get

Metatheorem 2 *Let T be a primitive recursively axiomatized theory, then the consistency of T can be proved almost finitistically if and only if*

$$PRA \vdash \forall x \exists y Prf_{PRA}(y, \lceil \neg Prf_T(\underline{x}, \bot) \rceil).$$

Thus, we introduce the following definition.

Definition 4.3 $\overline{Con}(T) \equiv \forall x \exists y Prf_{PRA}(y, \lceil \neg Prf_T(\underline{x}, \bot) \rceil)$.

It is easy to see that there are indeed theories whose consistency cannot be proven finitistically, but can be proven almost finitistically; PRA is such an example. There are much more powerful theories than PRA whose consistency can be proven almost finitistically but in order to show this we need further technical results.

Convention: 1 *Throughout the rest of this chapter T denotes a consistent, primitive recursively axiomatized theory whose axioms are presented in such a way that T provably extends PRA, i.e. such that*

$$PRA \vdash \forall x (Ax_{PRA}(x) \rightarrow Ax_T(x)). \tag{4.1}$$

Proposition 4.4 *Let T be as above, then*

$$(i) \quad PRA \vdash Con(T) \leftrightarrow (\overline{Con}(T) \wedge Con(PRA));$$
$$(ii) \quad PRA \vdash \overline{Con}(T) \leftrightarrow (Con(PRA) \rightarrow Con(T)).$$

Proof: (i) We argue model-theoretically. Let M be an arbitrary model of PRA. We can assume $M \models Con(PRA)$ since otherwise by 4.1, $M \models \neg Con(T)$, and in this case (i) is obviously true. Now, if $M \models Con(T)$ then $M \models \neg Prf_T(\underline{a}, \bot)$ for all $a \in M$ and so by

demonstrable Σ_1 completeness[6] of PRA, $M \models \exists y Prf_{PRA}(y, \lceil \neg Prf_T(\underline{a}, \bot) \rceil)$, i.e. $M \models \overline{Con}(T)$. On the other hand, if $M \models \neg Con(T)$ then for some $c \in M$, $M \models Prf_T(\underline{c}, \bot)$ and so, as before, $M \models \exists y Prf_{PRA}(y, \lceil Prf_T(\underline{c}, \bot) \rceil)$. Since by our assumption $M \models Con(PRA)$, we have $M \models \neg \exists y Prf_{PRA}(y, \lceil \neg Prf_T(\underline{c}, \bot) \rceil)$. Thus $M \models \neg \overline{Con}(T)$, which implies our claim.

(ii) This follows directly from (i) and the fact that

$$PRA \vdash \neg Con(PRA) \rightarrow \overline{Con}(T). \quad \square$$

Corollary 4.5 $PRA \vdash \overline{Con}(PRA)$.

Combining this with Metatheorem 2 we get a proof of the fact that the consistency of PRA is provable almost finitistically.

Well-known proof theoretic results about Π_1 conservative extensions of PRA show that there are serious theories whose consistency can be proved almost finitistically. The most important example of such an extension is WKL_0, which was shown by Sieg to be Π_2 conservative over PRA (see Sieg's Theorem 5.7 of [28]). Since Sieg's methods are finitary, his proof in fact shows how to define a primitive recursive function g such that

$$PRA \vdash \forall \varphi \in \Pi_1 \forall p (Prf_{WKL_0}(p, \varphi) \rightarrow Prf_{PRA}(g(p), \varphi)). \tag{4.2}$$

Definition 4.6 *(We adopt the terminology and notation of [6], p. 683) We say that a consistent theory T extending PRA is proof-theoretically reducible to PRA conservatively for Π_1 sentences of PRA if for some primitive recursive function $g(x)$*

$$PRA \vdash \forall \varphi \in \Pi_1 \forall p (Prf_T(p, \varphi) \rightarrow Prf_{PRA}(g(p), \varphi)); \tag{4.3}$$

we denote this fact by

$$T \leq PRA[\Pi_1].$$

Thus, 4.2 means that $WKL_0 \leq PRA[\Pi_1]$, and so all we need in order to show that the consistency of WKL_0 is provable almost finitistically is the following proposition.

Proposition 4.7 *Whenever T is such that $T \leq PRA[\Pi_1]$ then the consistency of T is provable almost finitistically.*

[6]This means that whenever $\varphi(\vec{x})$ is a Σ_1 formula, $PRA \vdash \forall \vec{x}(\varphi(\vec{x}) \rightarrow Thm_{PRA}(\varphi(\underline{\vec{x}}))$; for more details see Smorynski's [33], Proposition 6.22 on page 61. Smorynski denotes by PRA what we denote by $I\Sigma_1$, nevertheless, since $I\Sigma_1$ is (provably in PRA) Π_2 conservative over PRA, his result also applies for PRA.

Proof: Just take $\varphi \equiv \bot$, to get $PRA \vdash Con(PRA) \rightarrow Con(T)$, and so by Proposition 4.4(ii), $PRA \vdash \overline{Con}(T)$. QED.

Thus, we get the following corollary.

Corollary 4.8 *The consistency of WKL_0 can be proven almost finitistically, and yet, one can develop in it a great deal of the classical mathematics needed for empirical sciences.*

Thus, finding theories in which a lot of standard mathematics can be done, and whose consistency can be shown almost finitistically can be seen as a kind of a partial program, derived from Detlefsen's proposal by further restriction of the ω-rule that he suggests.

To find an "upper limit" for the strength of the theories whose consistency is provable almost finitistically, we prove the following theorem.

Theorem 4.9 *If the consistency of a theory T extending PRA is provable almost finitistically then T is proof-theoretically reducible to the theory $PRA + Con(PRA)$, conservatively for Π_1 sentences of PRA, i.e. $T \leq (PRA + Con(PRA))[\Pi_1]$.*

Proof: We have to show that there is a primitive recursive function h such that

$$PRA \vdash \forall \varphi \in \Pi_1 \forall p(Prf_T(p, \varphi) \rightarrow Prf_{PRA+Con(PRA)}(h(p), \varphi)). \tag{4.4}$$

We start by examining the proof of demonstrable Σ_1-completeness of PRA (see for example Theorem 6.22 of [33]).

Theorem 4.10 *(Essentially Gödel) Let $\varphi(\vec{x})$ be an arbitrary Σ_1 formula of PRA; then*

$$PRA \vdash \forall \vec{x}(\varphi(\vec{x}) \rightarrow Thm_{PRA}(\ulcorner \varphi(\underline{\vec{x}}) \urcorner)).$$

The proof of the above Theorem is by mathematical induction on the complexity of the formula φ and is purely elementary and uniform in φ. Thus, there is a primitive recursive function $s(x)$ such that

$$PRA \vdash \forall \varphi \in \Sigma_1 Prf_{PRA}(s(\varphi), \ulcorner \forall x(\varphi(x) \rightarrow Thm_{PRA}(\ulcorner \varphi(\underline{\vec{x}}) \urcorner)) \urcorner)$$

because the mathematical induction on the complexity of the formula φ can be formalized in PRA and so the above fact can be proved using an instance of the formal induction of PRA on φ. It is easy to see that the above implies

$$PRA \vdash \forall \varphi \in \Pi_1 Thm_{PRA}(\ulcorner \neg Thm_{PRA}(\neg \varphi) \rightarrow \varphi \urcorner). \tag{4.5}$$

Assume now that $PRA \vdash \overline{Con}(T)$ i.e. that $PRA \vdash Con(PRA) \rightarrow Con(T)$; since obviously $PRA \vdash \forall\varphi(Thm_T(\varphi) \wedge Con(T) \rightarrow \neg Thm_T(\neg\varphi))$ combining the last two facts we get

$$PRA \vdash \forall\varphi(Thm_T(\varphi) \wedge Con(PRA) \rightarrow \neg Thm_T(\neg\varphi)).$$

This clearly implies that

$$PRA \vdash \forall\varphi(Thm_{PRA}(\lceil Thm_T(\underline{\varphi}) \wedge Con(PRA) \rightarrow \neg Thm_T(\underline{\neg\varphi})\rceil). \tag{4.6}$$

Since by the demonstrable Σ_1 completeness

$$PRA \vdash \forall\varphi(Thm_T(\varphi) \rightarrow Thm_{PRA}(\lceil Thm_T(\underline{\varphi})\rceil)),$$

we get from 4.6 that

$$PRA \vdash \forall\varphi(Thm_T(\varphi) \rightarrow Thm_{PRA}(\lceil Con(PRA) \rightarrow \neg Thm_T(\neg\varphi)\rceil)).$$

This, together with 4.5 implies

$$PRA \vdash \forall\varphi \in \Pi_1(Thm_T(\varphi) \rightarrow Thm_{PRA}(\lceil Con(PRA) \rightarrow \varphi\rceil)),$$

and so

$$PRA \vdash \forall\varphi \in \Pi_1 \forall p \exists p'(Prf_T(p,\varphi) \rightarrow Prf_{PRA+Con(PRA)}(p',\varphi)),$$

which for the similar reasons as before implies the claim of our Theorem. QED.

Corollary 4.11 *If the consistency of T is provable almost finitistically and $T \vdash Con(S)$, then the consistency of S is also provable almost finitistically.*

Proof: Since $Con(S)$ is a Π_1 sentence of PRA provable in T, by the previous Theorem it is also provable in $PRA + Con(PRA)$ and so $PRA \vdash Con(PRA) \rightarrow Con(S)$ which by Proposition 4.4(ii) and Metatheorem 2 implies our claim.

Combining 4.7 and 4.9 we get:

$$\{T : T \leq PRA[\Pi_1]\} \subset \{T : PRA \vdash \overline{Con}(T)\} \subset \{T : T \leq PRA + Con(PRA)[\Pi_1]\}.$$

Here *both inclusions are proper*; to see this for the first inclusion note that the property of a theory of having an almost finitistic consistency proof is even weaker than the Π_1 conservativeness of T. Let for example φ be a Rosser sentence for PRA and let T be the theory $PRA + \varphi$. Then we have (see for example [33]) $PRA \vdash Con(PRA) \rightarrow \neg Thm_{PRA}(\neg\varphi)$,

i.e. $PRA \vdash Con(PRA) \rightarrow Con(PRA + \varphi)$, but $PRA + \varphi$ is obviously not Π_1 conservative over PRA because $PRA \nvdash \varphi$. Thus, the fact that the consistency of a theory T is provable almost finitistically does not necessarily imply 4.2[7].

To see that the second implication is proper, just take $T = PRA + Con(PRA)$. Then T obviously belongs to the third set but not to the second: $PRA \vdash \overline{Con}(PRA + Con(PRA))$ implies by 4.4(ii) $PRA + Con(PRA) \vdash Con(PRA + Con(PRA))$ which is impossible by Gödel's Second Theorem.

Obviously, our characterization of the theories whose consistency can be proven almost finitistically implies that the consistency of theories which contain stronger induction than $I\Sigma_1$ induction (and consequently prove the consistency of PRA) is not provable almost finitistically. This can be formulated in the form of a second incompleteness-type theorem.

Proposition 4.12 *Let T be a consistent primitively recursively axiomatized theory provably extending PRA. If T is strong enough to prove the consistency of the finitistic reasoning about numbers, i.e. if $T \vdash Con(PRA)$, then $T \nvdash \overline{Con}(T)$.*

Proof: If the consistency of T is provable almost finitistically then $PRA \vdash Con(PRA) \rightarrow Con(T)$; thus, the assumption that T proves $Con(PRA)$ implies $PRA \vdash Con(T)$ which is impossible.

One can also show an analogue of the First Incompleteness Theorem, for a bit odd provability predicate corresponding to $\overline{Con}(T)$.

4.2 A bizarre provability predicate

Definition 4.13 *Let T be as before, then $\overline{Thm}_T(\varphi) \equiv \neg\overline{Con}(T + \neg\varphi)$.*

The proof of the following proposition is routine, and consequently we omit it.

Proposition 4.14

$$PRA \vdash \overline{Thm}_T(\lceil\varphi\rceil) \leftrightarrow Thm_T(\lceil\varphi\rceil) \wedge Con(PRA),$$
$$PRA \vdash \overline{Thm}_T(\lceil\varphi\rceil) \rightarrow Thm_T(\lceil\varphi\rceil),$$
$$PRA \nvdash Thm_T(\lceil\varphi\rceil) \rightarrow \overline{Thm}_T(\lceil\varphi\rceil);$$

[7] It would be interesting to see if there are *mathematically significant* theories whose consistency can be almost finitistically proven, and which are not Π_1 conservative over PRA.

$$\vdash Con(T) \to \overline{Con}(T), \tag{4.10}$$

$$\nvdash \overline{Con}(T) \to Con(T); \tag{4.11}$$

$$\overline{Thm}_T(\lceil\varphi\rceil) \iff T \vdash \varphi, \tag{4.12}$$

$$\omega \models \quad \overline{Con}(T) \iff T \text{ is consistent;} \tag{4.13}$$

$$PRA \quad \nvdash \overline{Thm}_T(\lceil\varphi\rceil) \text{ , even for } \varphi \equiv \top; \tag{4.14}$$

$$PRA \quad \vdash \overline{Con}(T) \leftrightarrow \neg\overline{Thm}_T(\bot); \tag{4.15}$$

$$PRA \quad \vdash \overline{Thm}_{T+\varphi}(\lceil\psi\rceil) \leftrightarrow \overline{Thm}_T(\lceil\varphi \to \psi\rceil). \tag{4.16}$$

Despite having some unusual properties, $\overline{Thm}_T(\lceil\varphi\rceil)$ still behaves in many respects as a provability predicate, as the following analogue of the First Incompleteness Theorem shows.

Proposition 4.15 *Let T be a theory as before and let φ be a sentence asserting its own "bar-unprovability" in T, i.e.*

$$PRA \vdash \varphi \leftrightarrow \neg\overline{Thm}_T(\lceil\varphi\rceil),$$

then $T \vdash \varphi$ if and only if $T \vdash \neg Con(PRA)$, and $T \vdash \neg\varphi$ if and only if $T \vdash \neg\overline{Con}(T)$. Thus, if T is a true theory (i. e. $\omega \models T$), then T neither proves nor refutes φ.

Proof: If $T \vdash \varphi$ then $PRA \vdash Thm_T(\lceil\varphi\rceil)$. Thus, since by 4.7 of the previous Proposition and the definition of φ we have

$$\begin{aligned} \varphi &\leftrightarrow \neg(Thm_T(\lceil\varphi\rceil) \wedge Con(PRA)) \\ &\leftrightarrow \neg Thm_T(\lceil\varphi\rceil) \vee \neg Con(PRA), \end{aligned} \tag{4.17}$$

we get $PRA \vdash \varphi \leftrightarrow \neg Con(PRA)$ and so $T \vdash \neg Con(PRA)$. Conversely, if $T \vdash \neg Con(PRA)$ then from 4.7 $T \vdash \neg\overline{Thm}_T(\lceil\varphi\rceil)$ and so $T \vdash \varphi$.

If $T \vdash \neg\varphi$ then, from 4.7 $T \vdash Thm_T(\lceil\varphi\rceil) \wedge Con(PRA)$; hence, since $PRA \vdash Thm_T(\lceil\neg\varphi\rceil)$, we get $T \vdash \neg Con(T) \wedge Con(PRA)$ i. e. $T \vdash \neg\overline{Con}(T)$. Conversely, if $T \vdash \neg\overline{Con}(T)$ then $T \vdash \neg Con(T) \wedge Con(PRA)$ and so $T \vdash Thm_T(\lceil\varphi\rceil) \wedge Con(PRA)$. Thus, again by 4.7 we have $T \vdash \overline{Thm}_T(\lceil\varphi\rceil)$ and so $T \vdash \neg\varphi$.

If T is a true theory then $T \nvdash \neg Con(PRA)$ and $T \nvdash \neg\overline{Con}(T)$, since both statements are false on ω. Thus $T \nvdash \varphi$ and $T \nvdash \neg\varphi$. QED.

The proposition that would correspond to Löb's theorem, i. e.

$$T \vdash \overline{Thm_T}(\lceil\varphi\rceil) \to \varphi \quad \Longleftrightarrow \quad T \vdash \varphi$$

is true for theories having enough strength to prove $Con(PRA)$, since $T \vdash Con(PRA)$ implies $T \vdash \overline{Thm_T}(\lceil\varphi\rceil) \leftrightarrow Thm_T(\lceil\varphi\rceil)$, and so

$$T \vdash \overline{Thm_T}(\lceil\varphi\rceil) \to \varphi \iff T \vdash Thm_T(\lceil\varphi\rceil) \to \varphi \iff T \vdash \varphi.$$

For weaker theories this need not be true; moreover we have the following "Π_1 bar-soundness" of theories whose consistency can be proven almost finitistically.

Proposition 4.16 *Let T be as before; then the consistency of T is provable almost finitistically if and only if for all Π_1 sentences of PRA*

$$PRA \vdash \overline{Thm_T}(\lceil\varphi\rceil) \to \varphi. \tag{4.18}$$

Proof: One direction is obvious: assuming 4.18 just take $\varphi \equiv \bot$ to get $PRA \vdash \overline{Con}(T)$. For the other one assume $PRA \vdash \overline{Con}(T)$, i.e. $PRA \vdash Con(PRA) \to Con(T)$. By demonstrable Σ_1 completeness of PRA we have for all Π_1 sentences of PRA that $PRA \vdash \neg\varphi \to Thm_T(\lceil\neg\varphi\rceil)$, and so $PRA \vdash \neg Thm(\lceil\neg\varphi\rceil) \to \varphi$. Thus,

$$\begin{aligned} PRA \vdash \overline{Thm_T}(\lceil\varphi\rceil) \quad \to \quad & Thm_T(\lceil\varphi\rceil) \wedge Con(PRA) \\ & Thm_T(\lceil\varphi\rceil) \wedge Con(T) \\ & \neg Thm_T(\lceil\neg\varphi\rceil) \\ & \varphi. \qquad \text{QED.} \end{aligned}$$

Taking for φ a Π_1 sentence independent of T (e.g. $Con(T)$) we get a counterexample to the corresponding version of Löb's Theorem.

It is possible to show that $\overline{Thm_T}$ has further properties resembling ones of the standard provability predicate Thm_T; for example some versions of derivability conditions, formalized incompleteness theorems or formalized Löb's Theorem are true of it.

4.3 S and the consistency proofs

From Tait's Thesis it follows that the set S (of what we called almost finitistically provable sentences) consists of all the sentences of the form $\forall x\varphi(x)$, where $\varphi(x)$ is

primitive recursive, for which there exists a primitive recursive function $f(x)$ such that $PRA \vdash \forall x Prf_{PRA}(f(x), \lceil\varphi(\underline{x})\rceil)$.

One might wonder why the sentences φ satisfying the above condition cannot be considered finitistically proved. If $PRA \vdash \forall x Prf_{PRA}(f(x), \lceil\varphi(\underline{x})\rceil)$, this does implies that a finitist can realize that for every n there is a PRA proof of $\varphi(\underline{n})$, and he can "read" each of the individual proofs and obtain finitistic proofs for an arbitrary finite number of numerical instances of $\varphi(x)$, but he *cannot* know that *every* PRA proof corresponds to a finitistically acceptable proof since he does not understand the general concept of a finitary proof; informal finitary contentual derivations, even though he performs them, are not objects of his considerations.

We want to show that the theories whose consistency can be proven on the basis[8] of S are exactly theories whose consistency can be proven almost finitistically. By the same argument as before, there is a primitive recursive function f such that $PRA \vdash \forall x Prf_{PRA}(f(x), \lceil\varphi(\underline{x})\rceil)$ if and only if $PRA \vdash \forall x \exists y Prf_{PRA}(y, \lceil\varphi(\underline{x})\rceil)$, and so we introduce

Definition 4.17 *Let φ be of the form $\forall x \psi(x)$, where $\psi(x)$ is primitive recursive. Then*[9]

$$AFThm(\lceil\varphi\rceil) \equiv \forall x \exists y Prf_{PRA}(y, \lceil\varphi(\underline{x})\rceil).$$

Thus, $\varphi \in S$ if and only if $PRA \vdash AFThm(\lceil\varphi\rceil)$.

Proposition 4.18 *Let φ be a Π_1 sentence; then*

$$AFThm(\varphi) \leftrightarrow (Con(PRA) \rightarrow \varphi).$$

Proof: Similar to the proof of Proposition 4.4. QED.

From Proposition 4.16, taking $T = PRA$, we get that for an arbitrary Π_1 sentence φ, $PRA \vdash \overline{Thm}_{PRA}(\varphi) \rightarrow \varphi$; this, together with the previous proposition, implies that for all Π_1 sentences φ

$$PRA \vdash \overline{Thm}_{PRA}(\varphi) \rightarrow AFThm(\varphi).$$

Also, just from the definitions, we have

$$\vdash \quad \overline{Con}(T) \leftrightarrow AFThm(Con(T)). \tag{4.19}$$

To answer Question 5 on page 49 we first prove the following proposition.

[8] Recall that by Lemma 4.2 $PRA \subseteq S$.

[9] AFThm stands for "Almost Finitistic Theorem".

Proposition 4.19 *S and $PRA + Con(PRA)$ have the same set of theorems.*

Proof: Since by Corollary 4.5 $PRA \vdash \overline{Con}(PRA)$, 4.19 implies $PRA \vdash AFThm(Con(PRA))$. Thus, $Con(PRA) \in S$ and so, since by Lemma 4.2 $PRA \subseteq S$, we get $PRA + Con(PRA) \subseteq S$. Conversely, let $\varphi \in S$; then φ is Π_1 and $PRA \vdash AFThm(\varphi)$; thus, by Proposition 4.18 we have $PRA \vdash AFThm(\varphi) \rightarrow (Con(PRA) \rightarrow \varphi)$, and so $PRA \vdash Con(PRA) \rightarrow \varphi$. Hence $PRA + Con(PRA) \vdash \varphi$, which shows that all sentences from S are provable in $PRA + Con(PRA)$. Thus, $PRA + Con(PRA)$ and S have the same set of theorems. QED.

Corollary 4.20 *Let T be as before, then $S \vdash Con(T)$ if and only if the consistency of T can be proven almost finitistically.*

Proof: By the previous Proposition $S \vdash Con(T)$ if and only if $PRA + Con(PRA) \vdash Con(T)$, which is the case if and only if $PRA \vdash \overline{Con}(T)$. Now, our Corollary follows from Metatheorem 2. QED.

Thus, if we restrict the ω-rule from Detlefsen's first claim in the same manner (and with the same justification) as we restricted the ω-rule from Detlefsen's second claim, it turns out that this, at first sight the stronger rule, does not provide us with more power in proving consistency of theories than the previous one.

4.4 A generalization

One could argue that once we accept the above finitistically warranted ω-rule as a means of proof closely related to the finitistic proof, then there is no reason why we could not iterate it, i.e. use it to justify some stronger ω-rules. Thus, we could build a chain of theories S^i starting from PRA by adding to S^{i+1} only those primitive recursive instances of the ω-rule applied to only finitistically meaningful formulas that are already "warranted" by S^i.

Definition 4.21 *Let $S^0 = PRA$, and let S^{i+1} be the collection of all sentences of the form $\forall x \varphi(x)$, for $\varphi(x)$ primitive recursive, for which there exists a primitive recursive function $f(x)$ such that $S^i \vdash \forall x Prf_{S^i}(f(x), \lceil \varphi(\underline{x}) \rceil)$. Then we set $S^\omega = \bigcup_{i \in \omega} S^i$.*

Note that the above definition allows us to remain committed to the same class of sentences as meaningful, and the same class of functions as acceptable.

Now one could argue that the resulting theory S^ω is quite close to the standard finitistic reasoning, with an argument as follows. Assume $S^\omega \vdash \theta$. To prove θ, a finitist could start with some axioms of PRA and then add in stages some instances of the ω-rule mentioned in Definition 4.21, always justifying in the already obtained system any new instance he wants to add. After finitely many steps, he gains enough power to prove θ[10].

But what are the theories whose consistency is provable in S^ω? As before, it is easy to see that $S^i \subseteq S^{i+1}$ and that $PRA \vdash \overline{Con}^{S^i}(S^i)$. Thus, again using the fact that the provably recursive functions of PRA are exactly the primitive recursive ones, we get that for some primitive recursive function $f(x)$, $PRA \vdash Prf_{S^i}(f(x), \ulcorner \neg Prf_{S^i}(\underline{x}, \bot) \urcorner)$. Thus, $Con(S^i) \in S^{i+1}$. On the other hand, in a way similar to the proof of Proposition 4.19 it is easy to see that for any Π_1 sentence $\varphi \in S^{i+1}, S^i + Con(S^i) \vdash \varphi$. Hence, we have the following theorem.

Theorem 4.22 *S^{i+1} and $S^i + Con(S^i)$ have the same set of theorems.*

This obviously implies that $PRA \vdash Con(S^{i+1}) \rightarrow Con(S^i)$ and so for all i, j such that $i < j$ we have $PRA \vdash Con(S^i) \rightarrow Con(S^j)$. Thus we get the following Corollary.

Corollary 4.23 *S^{i+1} and $PRA + Con(S^i)$ have the same set of theorems.*

Obviously in the theory $S^\omega \equiv \bigcup_{i \in \omega} S^i$ one can prove consistency of more theories than in S, but these theories do not seem to be mathematically much more interesting than theories whose consistency is already provable in S: for example, even though $S^\omega \vdash Con(I\Sigma_1)$, the

[10] After reading an early version of a paper of mine [18] on which this chapter is based, David Libert called my attention to Feferman's paper [4]. In this paper he considers transfinite progressions of theories obtained by addition of reflection principles with the purpose of bridging the gap between the r.e. theories for which the incompleteness theorem holds and theories with non-constructive set of axioms (e.g. all true sentences of arithmetic). One of them (2.16(iv)) is similar to the way the S^i's are built, except that the proofs need not be produced by primitive recursive functions and that there are no complexity restrictions put on formulas (i.e. we add to a theory A all formulas of the form $\forall x \varphi(x)$ for *all* φ such that $A \vdash \forall x \exists y Prf_A(y, \ulcorner \varphi(\underline{x}) \urcorner)$). While the first restriction is inessential, the complexity restriction is important: without it, starting with the PRA, in the very next step we get full PA (of course our complexity restriction is imposed by what we accept as finitistically meaningful sentences). The iteration is transfinite along a path through Kleene's O, and the theory obtained as the union of all these iterations is the complete theory of the structure of natural numbers. In [7] Feferman and Spector showed that there are paths through O (all Π^1_1 ones) along which the same procedure as above gives a theory incomplete even for Π_1 sentences. A relevant reference is also Rosser's paper [26] which, according to [5], contains the first published discussion of the "provable" ω-rule with no restriction on the complexity of formulas.

strength of induction available in a theory whose consistency is provable in S^ω is below full Σ_2-induction. To show this we need the following Lemma.

Lemma 4.24 *For any natural number n, $I\Sigma_2 \vdash Con(S^n)$*

Proof: We proceed by induction on n, using the following fact which is a consequence of 1.6(ii) and 3.1 of Sieg's paper [28]:

$$I\Sigma_2 \vdash \forall \varphi \in \Sigma_1 (Thm_{PRA}(\varphi) \rightarrow Tr_{\Sigma_1}(\varphi)). \tag{4.20}$$

For $n = 0$ we have $S^0 \equiv PRA$ and so 4.20 implies (taking $\varphi \equiv \bot$) that $I\Sigma_2 \vdash Con(S^0)$. Assume $I\Sigma_2 \vdash Con(S^n)$; by the formalized deduction theorem, Corollary 4.23 and 4.20 we have

$$\begin{aligned} I\Sigma_2 \vdash \forall \varphi \in \Sigma_1 (Thm_{S^{n+1}}(\varphi) \quad \rightarrow \quad & Thm_{PRA}(Con(S^n) \rightarrow \varphi)) \\ & Tr_{\Sigma_1}(Con(S^n) \rightarrow \varphi)) \\ & (Con(S^n) \rightarrow Tr_{\Sigma_1}(\varphi)), \end{aligned}$$

and so by the inductive hypothesis $I\Sigma_2 \vdash \forall \varphi \in \Sigma_1 (Thm_{S^{n+1}}(\varphi) \rightarrow Tr_{\Sigma_1}(\varphi))$, which implies $I\Sigma_2 \vdash Con(S^{n+1})$. QED.

Proposition 4.25 *Let T be a theory extending PRA. If $S^\omega \vdash Con(T)$ then there is an instance of Σ_2 induction which is not provable in T.*

Proof: Assume $S^\omega \vdash Con(T)$ and that $T \vdash I\Sigma_2$. Let k be such that $S^k \vdash Con(T)$; by the above Lemma $T \vdash Con(S^k)$. We now use the following proposition known as Kreisel's Π_1 conservativeness theorem: for any Π_1 sentence φ of PRA and any consistent extension T of PRA, if $T \vdash \varphi$, then $PRA + Con(T) \vdash \varphi$ [11]. Thus, $T \vdash Con(S^k)$ implies $PRA + Con(T) \vdash Con(S^k)$, and so since $S^k \vdash Con(T)$ we get $S^k \vdash Con(S^k)$ which is impossible. QED.

Consequently, it seems that by adding the ω-rules as above, we lose finitistic grounds faster than we gain power in proving consistency of theories.

[11] To prove this proposition we basically formalize Hilbert's argument from [15], page 474. Let φ be a Π_1 sentence and assume that $T \vdash \varphi$. This obviously implies $PRA \vdash Thm_T(\varphi)$. We now use the provable Σ_1 completeness of PRA to get $PRA \vdash \neg\varphi \rightarrow Thm_{PRA}(\neg\varphi)$, which together with the previous fact implies $PRA \vdash \neg\varphi \rightarrow \neg Con(T)$. This implies $PRA + Con(T) \vdash \varphi$.

Chapter 5

Feasible reductions

In this Chapter we first sketch how to adapt Pudlák's original argument to show that PRA is feasibly reducible to the finitistic area of evidence. We then prove the results used in the Introduction to show that there are no mathematically important theories which are feasibly reducible to the finitistic area of evidence, as well as that f_{WKL_0} has approximately the same growth rate as $f_{I\Sigma_1}$[1].

Theorem 5.1 *There is a polynomial $P(x)$ with natural coefficients such that*

$$PRA \vdash \forall x \exists y(|y| < P(x) \wedge Prf_{PRA}(y, \lceil Con_{PRA}(\underline{x})\rceil)).$$

Sketch of the Proof: We first state the original Pudlák's result.

Theorem 5.2 *(Pudlák, Theorem 5.5 of [23]) Let A be a schema of the form*

$$\forall y \Phi(\varphi(y, z))$$

where $\varphi(y, z)$ is an arbitrary formula with free variables y and z, such that A is an axiomatization of a sequential theory[2] of a language which contains only finitely many predicate symbols. Then for some polynomial $P(x)$ and every $n \in \omega$ there is a proof from A of $Con_A(\underline{n})$ whose length smaller than $P(n)$.

While PRA is a sequential theory, it is neither on a finite language nor is it axiomatized by a schema, since besides the schema of induction it also contains axioms for primitive recursive definitions. Nevertheless, the conclusion of the above theorem holds for PRA as well if we choose a suitable coding.

[1] For the definitions see the Introduction.

[2] The reader should consult [23] for definitions and the details of the proof.

Claim 9 *Let the syntax of PRA be coded in such a way that there are at most n functional symbols whose code has length at most n (we again assume that the syntax of PRA is coded by the two letter alphabet $\{0,1\}$). Then there is a polynomial $P(x)$ such that for every $n \in \omega$ there is a proof from PRA of $Con_{PRA}(\underline{n})$ whose length is smaller than $P(n)$.*

To prove the above claim, one has to make only minor changes in the original Pudlák's proof from [23]. First of all, in Lemma 5.1 of [23] we replace the atomic clause in (1) with one that is a conjunction of the corresponding clauses for every functional symbol whose code is of length at most n. Since by our assumption there are at most n such functions, all lengths that Pudlák estimates in the proof of Lemma 5.1 are still bounded by a polynomial in n, and his proof goes through. Proofs of Lemmas 5.2 to 5.4 remain the same. The proof of Theorem 5.5 is essentially the same, since we have only what Pudlák calls a *sparse* set of axioms besides the schema of induction. This means that for each n we have at most polynomially many in n axioms which are not instances of the induction schema (axioms for the primitive recursive definitions). This is again a consequence of the fact that for each n there are at most n functions whose code is of length at most n. Thus, if we take the conjunction of all primitive recursive definitions for functions whose code is at most n and the formula β_n of the proof of Theorem 5.5 we can use it instead of just β_n in his proof. This increases the length of β_n only polynomially in n, and his proof is still valid (this essentially follows from the remark on the bottom of page 189 of [23]). An inspection of the proof of the above claim shows that it is purely proof-theoretical and uniform; thus it is formalizable in PRA and we get a proof of Theorem 5.1.

We now turn to to the proof of the incompleteness theorem mentioned in the introduction. We need the following result of Pudlák.

Theorem 5.3 *(Pudlák, Theorem 2.1. of [24]) Let T contain Q, be consistent and finitely axiomatizable. Let J be a cut in T. Then*

$$T \not\vdash \forall x(J(x) \rightarrow \neg Prf_T(x, \bot)).$$

We want to apply the above theorem to $I\Sigma_1$ which is not finitely axiomatized. Nevertheless, $I\Sigma_1$ is interpretable in its fragment obtained by restricting its language to the language of Peano Arithmetic $\mathcal{L}_{PA} = \{+, \cdot, <, 0\}$. This fragment, which we denote by $I\Sigma_1^*$, is indeed finitely axiomatizable. To prove this fact we can use the method from the

page 33; the only difference is that here we can use a simpler Σ_1 truth predicate since we do not have all primitive recursive functions in the language.

To interpret $I\Sigma_1$ in $I\Sigma_1^*$, we keep $+, \cdot, S, \leq, 0$ and $=$ the same and replace all other primitive recursive functions by their Δ_1^0 graphs. It is easy to see that for every primitive recursive function f there is a Δ_1^0 formula Θ_f of $I\Sigma_1^*$ (thus on the language $\mathcal{L}_{PA}$), such that

$$I\Sigma_1 \vdash \forall \vec{x} \forall y (f(\vec{x}) = y \leftrightarrow \Theta_f(\vec{x}, y)).$$

With the interpretation of the language of $I\Sigma_1$ in $I\Sigma_1^*$ given by $f(\vec{x}) = y \overset{I}{\mapsto} \Theta_f(\vec{x}, y)$ it is easy to see that $I\Sigma_1$ proves the interpretation of all axioms of $I\Sigma_1$. Thus, if there were a cut J such that

$$I\Sigma_1 \vdash \forall x (J(x) \rightarrow \neg Prf_{I\Sigma_1}(x, \underline{1 = 0})),$$

then, since under the interpretation I the $\mathcal{L}_{PA}$ part of the language remains the same,

$$I\Sigma_1^* \vdash \forall x (J^I(x) \rightarrow \neg Prf_{I\Sigma_1}(x, \underline{1 = 0})),$$

and since $I\Sigma_1^*$ (provably) extends $I\Sigma_1$, the last formula would contradict Theorem 5.3, applied to $I\Sigma_1^*$. Thus, the same claim is true of $I\Sigma_1$ even though it is not finitely axiomatizable.

Theorem 5.4 *For any natural number n*

$$I\Sigma_1 \nvdash \forall x \exists y (|y| < 2_n^x \wedge Prf_{PRA}(y, \lceil Con_{I\Sigma_1}(\underline{x}) \rceil)).$$

Proof: Assume the opposite, and let n be a natural number such that

$$I\Sigma_1 \vdash \forall x \exists y (|y| < 2_n^x \wedge Prf_{PRA}(y, \lceil Con_{I\Sigma_1}(\underline{x}) \rceil)).$$

let $J_{I\Sigma_1}$ be the cut defined in the proof of the Corollary 2.8, such that

$$I\Sigma_1 \vdash \forall x (J_{I\Sigma_1}(x) \rightarrow \neg Prf_{PRA}(x, \underline{1 = 0})).$$

Using Theorem 2.1(1) we first obtain a shortening $I(x)$ of the cut $J_{\Sigma_1}(x)$ such that $I(x)$ is closed under $+$ and $\cdot$ and consequently under any polynomial with natural coefficients. We now apply Theorem 2.1(2) to get a shortening $K_n(x)$ of the cut $I(x)$ with the property

$$I\Sigma_1 \vdash (K_n(x) \text{ is a cut contained in } I) \wedge \forall x (K_n(x) \rightarrow I(2_n^x)).$$

By Pudlák's Theorem 5.3 there exists a model $\mathcal{A}$ of $I\Sigma_1$ containing in K_n an $\mathcal{A}$-proof p of a contradiction from $I\Sigma_1$, i.e. such that $\mathcal{A} \models K_n(p) \wedge Prf_{I\Sigma_1}(p, \underline{1=0})$[3]. Since we can check in PRA the syntax of a sequence of formulas and determine whether it is a correct proof in $I\Sigma_1$, and since this can be done in polynomially many steps in the length of the sequence, there is a proof $p^* \in \mathcal{A}$ of length polynomial in length of p, such that

$$\mathcal{A} \models Prf_{PRA}(p^*, \lceil Prf_{I\Sigma_1}(p, \underline{1=0})\rceil)$$

Thus, for some polynomial $P(x)$ with natural coefficients and some p' obtained from p^* in the obvious way we get

$$\mathcal{A} \models Prf_{PRA}(p', \lceil \neg Con_{I\Sigma_1}(|p|)\rceil) \wedge |p'| \leq P(|p|).$$

On the other hand, by our assumption, for some $\overline{p} \in \mathcal{A}$,

$$\mathcal{A} \models Prf_{PRA}(\overline{p}, \lceil Con_{I\Sigma_1}(|p|)\rceil) \wedge |\overline{p}| < 2_n^{|p|}.$$

Since $K_n(p)$ and $|p| < p$ imply $K_n(|p|)$, and since $|\overline{p}| < 2_n^{|p|}$ then by the definition of K_n we get that $I(|\overline{p}|)$.

Combining p' and $\overline{p}$ we can clearly get a proof $p^{\#}$ of an inconsistency in PRA whose length is polynomial in the lengths of $\overline{p}$ and p and so by our assumption about closure properties of the cut I, $I(|p^{\#}|)$ which is a contradiction since I is a shortening of K and $I\Sigma_1 \vdash \forall x(K(x) \rightarrow Con_{PRA}(x))$. QED.

We now prove the other theorem we appealed to in the Introduction.

Theorem 5.5 *There is a function g_{WKL_0} with a roughly hyperexponential growth rate such that*

$$PRA \vdash \forall x \exists y(|y| < f(x) \wedge Prf_{PRA}(y, \lceil Con_{WKL_0}(\underline{x})\rceil)).$$

Proof: By the result of Sieg's we mentioned on page 51 (Corollary 5.7 of [28]) there is a primitive recursive function g such that

$$PRA \vdash \forall \varphi \forall p(\Pi_2(\varphi) \rightarrow (Prf_{WKL_0}(p, \varphi) \rightarrow Prf_{PRA}(g(p), \varphi)). \tag{5.2}$$

Such g formalizes a proof transformation procedure that first eliminates cuts and then performs some other proof transformations that do not increase lengths of proofs significantly

[3]The technique which consists of shortening a cut and then injecting an inconsistency in it is due to Pudlák; see Corollary 4.4 of [24].

comparing to the roughly hyperexponential growth rate of the cut-elimination procedure. Thus, g is dominated by a monotone function h of roughly hyperexponential growth rate and consequently we have

$$PRA \vdash \forall x(Con_{PRA}(h(x)) \rightarrow Con_{WKL_0}(x)).$$

As a consequence of Theorem 5.1 we have

$$PRA \vdash \forall x \exists y(|y| < (h(x))^k \wedge Prf_{PRA}(y, Con_{PRA}(\underline{h(x)}))).$$

But from 5.2 we have that for a standard natural number c

$$PRA \vdash Prf_{PRA}(\underline{c}, \ulcorner \forall x(Con_{PRA}(h(x)) \rightarrow Con_{WKL_0}(x)) \urcorner),$$

which obviously implies that there is a function f with a roughly hyperexponential growth rate such that

$$PRA \vdash \forall x \exists y(|y| < f(x) \wedge Prf_{PRA}(y, Con_{WKL_0}(x))).$$

Since we have $PRA \subset I\Sigma_1 \subset RCA_0 \subset WKL_0$, (and provably so in PRA for the corresponding representation of the axioms of these theories) Theorems 5.4 and 5.5 imply that $f_{I\Sigma_1}$, f_{RCA_0} and f_{WKL_0} all have a roughly superexponential growth rate.

Bibliography

[1] D.K.Brown: *Subsystems of Second Order Arithmetic*, Ph.D. Thesis, Pennsylvania State University, May 1987.

[2] M. Detlefsen: *On interpreting Gödel's Second Theorem*, **Journal of Philosophical Logic**, vol 8 (1979), pp. 297-313.

[3] H. Enderton: A Mathematical Introduction to Logic, Academic Press (1972).

[4] S. Feferman: *Transfinite recursive progressions of axiomatic theories*, **Journal of Symbolic Logic**, vol. 27 number 3 (1962), pp. 259-316.

S. Feferman: Introductory note to 1931c, in **Kurt Gödel: Collected Works vol.I**, (S. Feferman editor-in-chief), Oxford University Press (1986).

[6] S. Feferman: *Hilbert's Program Relativized: Proof-theoretical and foundational reductions*, **Journal of Symbolic Logic**, vol. 53 (1988).

S. Feferman, C. Spector: Incompleteness along paths in progressions of theories, **Journal of Symbolic Logic**, vol. 27 (1962).

[8] H. Friedman: *Systems of second order arithmetic with restricted induction.* I, II (abstracts), **Journal of Symbolic Logic**, vol. 41 (1976), pp. 557-559.

[9] H. Friedman,S.G.Simpson,R.L.Smith: *Countable algebra and set existence axioms*, **Annals of Pure and Applied Logic**, vol.25 (1983), pp. 141-181; addendum, vol 28 (1985), pp. 319-320.

[10] M. R. Garey, D. S. Johnston: Computers and Intractability, Freeman and company (1979).

[11] P. Hájek: On interpretability of theories containing arithmetic, II, **Commentationes Mathematicae Universitatis Carolinae**, vol. 22 (1981), pp. 667-688.

[12] *Jacques Herbrand:* Logical Writings, (W. Goldfarb, editor) Harvard University Press (1971), pp. 288-9.

[13] D. Hilbert: *Über die Grundlagen der Logik und der Arithmetik, Verhandlungen des Dritten Internationalen Mathematiker-Kongresses in Heilderberg vom 8. bis 13. August 1904* (Teubner, Leipzig, 1905), pp. 174- 185, translated in [38].

[14] D. Hilbert: *Über das Unendliche*, **Mathematische Annalen**, vol. 95 pp. 161-190; translated in Philosophy of Mathematics, selected readings, (P. Benacerraf and H. Putnam, editors), Cambridge University Press (1983), also in [38].

[15] D.Dilbert: *Die Grundlagen der Mathematik,* **Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universität**, 6 (1928), pp. 65-85, translated in [38].

[16] D. Hilbert: *Die Grundlegung der elementaren Zahlenlehre*, **Mathematische Annalen**, (1931), 104, pp. 485-494.

[17] D. Hilbert and P. Bernays: Grundlagen der Mathematik I, II; 2. Auflage, Springer, Berlin (1968/70).

[18] A. Ignjatović: *Hilbert's Program and the ω-rule*, unpublished manuscript, December 1988.

[19] A. Ignjatović: *On Mathematical Instrumentalism*, unpublished manuscript, July 1989.

[20] D. Isaacson: *Arithmetical truth and hidden higher-order concepts*, in **Logic Colloquium 85**, edited by the Paris Logic Group, North Holland, 1987.

[21] G. Kreisel: *What can be done for Mathematical Logic*, in Bertrand Russell, the philosopher of the century, Essays in his honour, edited by Ralph Shoenman, Little, Brown and company, Boston (1967).

[22] J. Paris and C. Dimitracopulos, *A note on undefinability of cuts*, **Journal of Symbolic Logic**, vol. 48 (1983) pp. 564-569.

[23] Pavel Pudlák: *On the length of proofs of finitistic consistency statements in first-order theories*, Logic Colloquium '84, J.B.Paris, A.J.Wilkie and G.M.Wilmers (Editors) Elsevier Science Publishers, North-Holland (1986).

[24] P. Pudlák: *Cuts, Consistency Statements and Interpretations*, **Journal of Symbolic Logic**, vol. 50 (1985) pp.423-441.

[25] M.D. Resnik: Frege and the Philosophy of Mathematics, Cornell University Press (1980).

[26] B. Rosser: *Gödel Theorems for non-constructive logics* **Journal of Symbolic Logic**, vol. 2 (1937), pp. 129-137.

[27] J.R. Shoenfield: *On a Restricted ω-rule*, **Bulletin De L'Academie Polonaise Des Sciences**, vol. VII, No. 7 (1959).

[28] W. Sieg: *Fragments of arithmetic*, **Annals of Pure and Applied Logic**, vol. 28 (1985), pp. 33-71.

[29] S.G. Simpson: *Which set existence axioms are needed to prove the Cauchy-Peano theorem for ordinary differential equations?* **Journal of Symbolic Logic**, vol. 49 (1984), pp. 783-802.

[30] S.G. Simpson: *Partial realization of Hilbert's Program*, **Journal of Symbolic Logic**, vol. 53 (1988), pp. 349-363.

[31] S.G. Simpson: *Subsystems of Second order arithmetic* (in preparation).

[32] S.G. Simpson and R.L. Smith: *Factorization of polynomials and Σ^0_1-induction*, **Annals of Pure and Applied Logic**, 31 (1986) pp. 289-306.

[33] C. Smorynski: *The Incompleteness Theorems*, in **Handbook of Mathematical Logic** (J.Barwise, editor), North-Holland (1977), pp. 821-895.

[34] R. Solovay: *Letter to P. Hájek;* see also [11] and [22].

[35] R. Statman: *Bounds for proof search and speed-up in the predicate calculus*, **Annals of Mathematical Logic**, vol 15 (1978), pp. 225-287.

[36] W.W. Tait: *Finitism*, **Journal of Philosophy**, vol. 78 (1981), pp. 524-546.

[37] G. Takeuti: Proof Theory, second edition, North-Holland (1987).

[38] J. Van Heijenoort: From Frege to Gödel; a source book in mathematical logic 1879-1931. Cambridge, Harvard University Press (1967).