On the lengths of proofs of consistency

a survey of results

Pavel Pudlák

Mathematical Institute
Academy of Sciences
Prague, Czech Republic

Preface. *This article is essentially a part of my thesis for the degree DrSc (Doctor of Sciences). Therefore it mainly surveys my articles [41, 42, 43, 28, 29, 44, 22], and it is structured according to the requirements for such theses. I made only minor changes in the original text and added a few further references. Since Gödel's main achievement concerns the problem of consistency and some of the problems that I am going to describe had been considered by him, I think that it is appropriate to publish this article in Gödel Society.*

1 Historical remarks

The question that we are going to consider in is interesting *per se* and is related to some more practical questions, especially in complexity theory, but the original motivation for it comes from foundational studies. Among the variety of streams in foundations of mathematics, the one which had the biggest influence and which very much determined later development of mathematical logic was Hilbert's *formalism.* It was an attempt to reduce mathematical theories to *logical theories*, which in turn are defined using a formal logical calculus. In this way the obscure concepts of *intuition, meaning, etc.* are eliminated. But then the consistency of a theory, which is a condition *sine qua non*, looses any support, and hence *must be proved.*

Apparently Hilbert hoped that eventually it will be possible to prove consistencies of all mathematical theories (as they can be formalized in set theory, it would suffice to do it for set theory). This program received a serious blow, before there was any significant progress in proving such consistencies.[1] The result, which in fact essentially destroyed it, is what is nowadays referred to as *Second Gödel's Incompleteness Theorem*, which says that the consistency of a theory T cannot be proved in T itself. We shall denote the sentence saying that T is consistent by Con_T. If we restrict the class of theories to some natural systems, say theories which are subsets of the true sentences in the standards model of arithmetic $\mathbf{N}$, then the relation $T \vdash Con_S$ (T proves Con_S) is a strict partial ordering. Hilbert's idea was to use some *finitistic* arguments for the proofs of consistency, i.e. to use simple combinatorial arguments. However, at the same time he assumed that any mathematical argument can be a part of some theory.

[1] The famous proof of consistency of the first order arithmetic of Gentzen [17] was published several years after Gödel's [18].

This implies that we can prove Con_T only in some "stronger" theory; thus if we doubt in the consistency of T, we can only reduce it to the consistency even more doubtful. Of course, for some T, we can reduce the consistency of T, which uses seemingly strong statements, to a theory for which we have more intuitive reasons to believe to be consistent. But then we are back at the beginning – we refer to intuition.

A possible way out is to give up absolute consistency and consider only practical consistency, i.e. the question whether we can ever encounter a contradiction, if we use only small proofs. One can argue that the size of the proofs that men will ever consider is bounded or that the universe is finite etc. At first glance it seems that the restriction to proofs of a size bounded by some constant n makes a big difference. In any case, it is clear that now we do not need a strong theory to prove the consistency up to n, since the procedure of checking all such proofs can be described in a weak theory. The problem is, however, that this may be exponentially longer than n; put otherwise the question is now, to find a *feasible* proof that all feasible proofs are consistent. Kreisel mentioned a discussion with Gödel, where Gödel talked about this problem.

The only attempt of actually proving such a consistency result that I know of, is a paper by Esenin-Volpin [12]. He claimed that he proved the consistency of a set theory up to some large number n using a finitistic argument. Unfortunately his arguments were obscure and I have not met anybody who could explain me, what actually Esenin-Volpin proved.

These questions influenced several directions of research. We shall mention some results related to the consistency problem.

The consistency problem was an initiating idea for studying inconsistent theories. The basic result about "useful" inconsistent theories is due to Parikh [37], who constructed theories which are inconsistent, but where any proof of contradiction is extremely long, and showed some nice properties of such theories. For further results see [11, 16, 34].

Another problem connected with the result of Gödel was the question, for how weak theories the second incompleteness theorem is true. Bezboruah and Shepherdson [5] proved the second incompleteness theorem in $\mathbf{Q}$, which is one of the weakest arithmetical theories. The result was in certain sense problematic: if the theory is so weak, does the particular formalization of $Con_{\mathbf{Q}}$ really mean what was intended? A solution, which will be presented below, is to define an initial segment J of the numbers in $\mathbf{Q}$, which is an inner model of a stronger theory T and prove that it is consistent to assume that a proof of contradiction from axioms of $\mathbf{Q}$ is encoded by a number which is already in J. Since T is strong, the meaning of $Con_{\mathbf{Q}}$ is not so ambiguous.

Let us observe that this naturally leads to questions about the *lengths* of proofs of contradictions and this in turn can be investigated by computing the lengths of proofs of statements about finite consistency, as we shall see below.

The finite consistency of T is the statement, that *there are no proofs of contradiction of length n in T,* where n is some positive integer. We shall denote this sentence by $Con_T(n)$. The *length of a proof* will always mean the length of the sequence which represents the proof. Clearly, if n is a concrete (standard) integer and T is reasonably strong (in fact already $\mathbf{Q}$ suffices), then $T \vdash Con_T(n)$. The question *what is the minimal length of a proof of* $Con_T(n)$ has been considered for the first time by Harvey Friedman [14]. He proved a lower bound of the form n^ε, for a constant $\varepsilon > 0$. His paper was almost unknown, since it existed only in a preprint form so I rediscovered this result several years later [42].[2]

Around 1980 a new field in logic become popular – weak fragments of arithmetic, nowadays called *Bounded Arithmetic*.[3] The first such system was $I\Delta_0$ which is the usual first order arithmetic (Peano Arithmetic) where induction is postulated only for bounded formulas; later a whole variety of systems has been introduced. This line of research was pursued at the beginning mainly by the Manchester school, Paris, Wilkie and others, later it was very influenced by Buss's book [6]. Though it has a different motivation, this topic is related to the consistency proofs too. Firstly, new more efficient ways of formalizing syntax have been developed. The bounds on the lengths of proofs of finite consistencies $Con_T(n)$ are very much based on such techniques. Secondly, Paris and Wilkie realized that restricted forms of consistency are useful for separations of weak fragments of arithmetic [55]. (They will be defined below.)

In spite of the fact that the importance of the problems about the lengths of proofs has been realized very early, see e.g. Gödel's [19], the literature on this field was not very rich until recently. Kreisel was probably the main proponent of this field. His student Statman [19] proved that the size of proofs after cut-elimination cannot be bounded by an elementary function of the original size. His famous conjecture about generalizations of proofs in Peano Arithmetic has been proved only in some special cases (see e.g. [3]). Another line of research has been started by Cook [9] who was the first to realize the importance of the lower bounds on the lengths of proofs in propositional calculus and a relation of this problem to equational and first order theories. In this field there was quite a significant progress recently; to name just a few [10, 30, 1, 4, 2, 27].

After the famous result of Paris and Harrington [40], who proved independence of a concrete mathematical statement from Peano Arithmetic, we started to study fragments of arithmetic with an intention to prove eventually independence of some simpler (from the point of view of quantifier complexity)[4] sentences, in particular we hoped to prove independence of some problems in complexity theory. After more than 15 years of experience it has become clear that it will be necessary to develop the theory much further in order to obtain such results. Presently we are able to show only so called *oracle results*, which means that we can show independence of a certain combinatorial principle for

---

[2] I have presented it on an invited talk at *Logic Colloquium'84* in Manchester, but nobody told me about Friedman's paper there.

[3] Parikh had proposed to study such systems already in 1971 [37], but no new results had been published till late seventies.

[4] The sentences of Paris and Harrington are Π_2, the only "explicit" Π_1 sentences that we know of are Con_T.

an extra uninterpreted predicate (see e.g. [31]). It is still possible that interesting independence results will be obtained by reduction from certain consistency statements, though very little has been done in this direction recently. After all, Paris-Harrington's statement is equivalent to a reflection principle, which is a strengthening of the consistency statement.

2 A brief overview

In this section we shall briefly state the results which we shall discuss below in more detail.

I arrived to the problem of finite consistencies through interpretability of theories. The problem was to characterize interpretability of theories containing a fragment of arithmetic using provability of restricted consistencies. This had been done long before for *reflexive theories* [35, 36, 20]; the main examples of reflexive theories are Peano Arithmetic and Zermello-Fraenkel Set Theory. The property of being reflexive narrows the class of theories too much, e.g. any such theory is not finitely axiomatizable. The solution is to take the restricted consistency statements of Paris and Wilkie [54] and furthermore restrict them to *cuts*. It turns out that these results have other unexpected corollaries, e.g. one can derive the result of Statman mentioned above.

As the condition that the inconsistency proof should lie in a cut is essentially a restriction on the length of such a proof, the next natural step was to investigate the lengths of proofs of finite consistencies $Con_T(n)$ and their versions with restrictions of quantifier complexity. This, in particular, gives more precise bounds in the same applications. Moreover it answers some foundational questions, e.g. proves that the finite modification of Hilbert's program is hardly feasible – for proving $Con_T(n)$ in T one needs a proof of length about n. We have both a lower bound and an upper bound on the lengths of proofs of $Con_T(n)$ in T, and, a bit surprisingly, they are quite close to each other and just around n.

Though these results belong mainly to proof theory, some corollaries can be derived also for the model theory of arithmetic. Namely, the question of how much is the structure of a model M beyond an element a determined by the initial segment $M|_a$, can be solved using knowledge about the lengths of proofs of $Con_T(n)$.

The question whether there is a (say, finitely axiomatized) theory S which proves $Con_T(n)$ for *any* (finitely axiomatized) theory T by a proof of length polynomial in n remains open. So it is possible that this very weak version of Hilbert's program could be realized. It turns out that this question is equivalent to a problem about existence of the most efficient proof system for propositional calculus.

It is, maybe, questionable whether logic can help to solve the persistent deep problems in complexity theory, but if it can, then the following seems to be the most promising way. There are fragments of arithmetic which roughly correspond to various complexity classes. If one could separate (= prove that they are not equal) a pair of such theories, it would be a strong justification for our belief, that the corresponding complexity classes are different. There is essentially only one known method for separating theories which contain some fragment of arithmetic or set theory: the second Gödel's incompleteness theorem. Namely the method is based on proving that one theory proves the consistency of the other. This fails for weak fragments of arithmetic, since they are all equiconsistent modulo a very weak theory. Still it is possible that a suitable weakening of the concept of consistency will do. Another application of the bounds on the lengths of proofs of consistency shows that a seemingly natural concept of *bounded consistency* is still too strong for this purpose.

The last result that we will survey is another application of these bounds which gives lower bounds on the lengths of proofs of a little more "mathematical" sentence in case of the theory $I\Delta_0 + \Omega_1$.

3 Basic tools

We shall give some definitions and list a few specific tools on which the proofs are based.

We shall assume in the whole paper that a *theory* is recursively axiomatizable and contains Robinson's arithmetic **Q**. Whenever we talk about the length of a proof in a theory T the particular axiomatization becomes essential. So we rather talk about particular sets of axioms, not just deductively closed sets. In such cases we shall moreover assume that the set of axioms is defined effectively, namely we shall assume that it is a set in $\mathcal{P}$. (Note that $\mathcal{NP}$ would suffice in most cases.)

The Gödel number of a formula φ will be denoted by $\ulcorner\varphi\urcorner$, the length of a formula or a proof will be denoted by $|\ldots|$.

The following is the main lemma which goes back to Gödel [18]. Gödel constructed a concrete selfreferential formula, the general lemma is due to Carnap [8]; in the parametric form it was proved in [33]. (For the proof see e.g. [23].)

Lemma 1 Diagonal Lemma, Fixpoint Theorem. *Let T be a theory and let $\psi(x, y, z)$ be a formula. Then there exists a formula $\varphi(y, z)$such that*

$$T \vdash \varphi(y, z) \equiv \psi(\ulcorner\varphi(y, \underline{z})\urcorner, y, z),$$

where $\underline{z}$ denotes the z-th numeral, (see below for definition).

Let T be a theory. Then a formula $\varphi(x)$ is called *a cut in T*, if T proves the following sentences:

1. $\varphi(0)$;
2. $\varphi(x) \rightarrow \varphi(x + 1)$;
3. $\varphi(x) \ \wedge \ y \leq x \ \rightarrow \ \varphi(y)$.

The following are the basic techniques of "cut shortening", which enable to construct cuts with additional properties.

Lemma 2 Solovay, unpublished, see e.g. [23]. *1. For every cut $\varphi(x)$ in T, there exists a cut $\psi(x)$ in T, such that $\psi(x)$ is contained in $\varphi(x)$ and $\psi(x)$ is closed under addition and multiplication (provably in T).*

2. For every cut $\varphi(x)$ in T, there exists a cut $\psi(x)$ in T, such that $\psi(x)$ is contained in $\varphi(x)$ and

$$T \vdash \forall x(\psi(x) \rightarrow \varphi(2^x)),$$

provided the function 2^x is suitably formalized.

Let us note, that in general it is not possible to find a shortening closed under exponentiation (the condition in 2 is weaker).

We shall use the following theories, fragments of Peano Arithmetic: $I\Delta_0$, $I\Delta_0+\Omega$, $I\Delta_0+Exp$, where $I\Delta_0$ is the theory axiomatized by $\mathbf{Q}$ and the induction schema for all bounded formulas, Ω denotes the sentence $\forall x \exists y\ (y = x^{\lceil \log x \rceil})$, and Exp denotes the sentence $\forall x \exists y\ (y = 2^x)$. Since the exponential function is not in the language, we use a Δ_0 definition of the graph of this function (see e.g. [23]).

Here is another important result about cuts.

Theorem 3 Wilkie [54]. *For every cut $\varphi(x)$ in a theory[5] T, there exists a cut $\psi(x)$ in T such that $\psi(x)$ is contained in $\varphi(x)$ and such that T proves $I\Delta_0 + \Omega$ restricted to ψ, (i.e. there is an interpretation of $I\Delta_0 + \Omega$ with absolute $+$ and $\cdot$ and the domain contained in φ).*

Futhermore we need a technique of constructing short formulas defined by an iterative schema. Let $\varphi_0(\mathbf{x})$ and $\Phi(X, \mathbf{x})$ be given and suppose we want to construct a sequence of formulas $\varphi_i(\mathbf{x})$, $i = 1, 2, \ldots$, such that

$$\varphi_{i+1}(\mathbf{x}) \ \equiv \ \Phi(\varphi_i(\mathbf{x}), \mathbf{x}).$$

Then we can simply substitute $\varphi_0(\mathbf{x})$ in $\Phi(X, \mathbf{x})$ to get $\varphi_1(\mathbf{x})$, then substitute $\varphi_1(\mathbf{x})$ in $\Phi(X, \mathbf{x})$ to get $\varphi_2(\mathbf{x})$ etc. But if X occurs at least twice in $\Phi(X, \mathbf{x})$, the size of such formulas $\varphi_i(\mathbf{x})$ grows exponentially. Ferrante and Rackoff [13] and Solovay [unpublished] have devised techniques which produce formulas $\varphi_i(\mathbf{x})$ of polynomial size.

4 Results

4.1

For a formula $J(x)$ and a sentence φ, we shall denote by φ^J the sentence obtained by restricting the quantifiers to the domain $J(x)$. The following is a strengthening of the second Gödel's incompleteness theorem.

Theorem 4 [41] – Theorem 2.1. *For every consistent theory T and every cut $J(x)$ in T, T does not prove Con_T^J.*

[5] recall that we assume that $\mathbf{Q} \subseteq T$

Put otherwise, there exists a model M of T in which there is a contradiction in T where we, moreover, have the contradiction in J^M. The original proof was an extension of the classical proof of Gödel Theorem. We shall see that this, as well as the next result, are also easy consequences of bounds on the lengths of proofs of finite consistencies.

A corollary of this result is that the second Gödel's incompleteness theorem holds in weak theories, in particular in $\mathbf{Q}$, without any doubts about what Con_T really means there. This is because by Theorem 3 there is a cut in $\mathbf{Q}$ which is a model of $I\Delta_0 + \Omega$. In such a cut all reasonable definitions of Con_T are equivalent.

This result suggests a natural question: what additional conditions can we impose on the proofs of contradiction in the second Gödel's incompleteness theorem? The general question has not been answered; the following is an example of a different condition, though also based on restricting the size.

Theorem 5 [41] – Theorem 2.5, [42] – Theorem 4.3. *For every consistent theory T and every Δ_0-formula $I(x)$, there exists a constant k such that, if*

$$T \vdash I(0), I(1), I(2), \ldots,$$

then

$$T + \exists x \ (I(x) \wedge \neg Con_T(x^k))$$

is consistent.

For the following results we have to introduce a restricted provability predicate. The *quantifier complexity* of a formula φ is the maximal number of nesting of quantifiers. *A k-proof* is a proof in which all formulas have quantifier complexity at most k. Note that each first order tautology φ has a cut-free proof and a cut-free proof has the *subformula property.* Hence, if k is at least the quantifier complexity of φ, then φ has a k-proof. Let us note that originally Paris and Wilkie considered the consistency with respect to proofs in which the complexity of every formula is bounded by a constant, where the complexity of terms is disregarded. This definition can also be applied below.

Let $Con_{T,k}$ denote the consistency statement restricted to k-proofs.

A theory is called *sequential*, if there is a formula for coding sequences of elements from the universe which satisfies two axioms: 1. the empty sequence has a code, 2. every sequence s can be prolonged by adding an arbitrary element x. (For a precise definition see [23, page 151].)

The theorem below is a further strengthening of the second Gödel's incompleteness theorem, since we use a weaker statement $Con_{T,k}$ instead of Con_T. Moreover the converse shows that it is, in a sense, sharp.

Theorem 6 [41] – Theorem 2.7, Corollary 3.2. *(i) Let T be a finitely axiomatized consistent theory. Then for every cut J in T, there exists a number k such that $T \nvdash Con^J_{T,k}$.*

(ii) Let T be a sequential theory, then for every k, there exists a cut J in T such that $T \vdash Con^J_{T,k}$.

The two statements suggest a natural conjecture: *If T is a finite consistent sequential theory, then there exists a model M of T in which the shortest k-proofs of contradiction, for $k = 1, 2, \ldots$, are cofinal downwards with definable cuts.* This does not follow simply by compactness and requires an additional trick. The conjecture has been proved by Krajíček [26]. For further extensions of these results see Visser [53] and Verbrugge [52]; [53] surveys and gives precise proofs of some previous results whose proofs were only sketched.

The restricted consistency statements can be used to characterize interpretability of a sentence φ in a theory T. An *interpretation* ι of φ in a theory T is given by a translation of the relation and function symbols into formulas of T with the corresponding number of free variables, which inductively defines the translation φ^ι, and by a proof of φ^ι in T. For a more precise definition see [41] or [23]; an interpretation can use parameters etc.

Further concepts that we need are *Herbrand consistency* and *cut-free consistency.* Herbrand's theorem characterizes provability by stating that a sentence φ is provable iff there exists a certain finite object, namely a Herbrand disjunction for φ. We can think of this object as a proof of φ. A theory T is Herbrand consistent, if there is no such proof of the negation of a conjunction of axioms of T. We shall denote Herbrand consistency of T by $HCon_T$. Of course, if a theory S is sufficiently strong, say it is Peano Arithmetic, then $S \vdash HCon_T \equiv Con_T$. So the distinction appears only if we consider weak theories, e.g. $I\Delta_0 + Exp$, or we consider restriction on the lengths of such proofs. The cut-free consistency is a closely related concept. It is based on cut-free proofs in sequent calculi. For more information about this relation see [23, Chapter 5, Section 5].

Theorem 7 [41] – Theorem 3.1. *Let T be a sequential theory and φ a sentence. Then the following are equivalent:*

(i) there is an interpretation of φ in T;

(ii) for every k, there exists a cut J in T such that $T \vdash Con^J_{\varphi,k}$;

(iii) there exists a cut J in T such that $T \vdash HCon^J_\varphi$.

Note that in particular, if T is finitely axiomatized and sequential, then $T \vdash HCon^J_T$ on some cut J in T.

If a theory S has an interpretation in T, then the consistency of T implies the consistency of S, (this is the main application of the concept of interpretability). Does the converse in some modified way hold too? The above theorem gives some reduction to provability of consistencies, but it can hardly be considered as such a converse. This has been achieved in a beautiful theorem of Friedman:[6]

Theorem 8 Friedman [15]. *Let S, T be finite sequential theories. Then the following are equivalent:*

(i) there exists an interpretation of S in T;

[6] This paper also has not been published; the theorem is presented in an article of Smoriński in [45]. Had I known the result, I would not publish my characterization.

(ii) for some m,

$$I\Delta_0 + Exp \vdash \ Con_{T,m} \rightarrow Con_{S,k},$$

where k is the number of quantifiers in S;

(iii)

$$I\Delta_0 + Exp \vdash \ HCon_T \rightarrow HCon_S.$$

Thus S has an interpretation in T iff a certain week form of relative consistency of S with respect to T is provable in $I\Delta_0 + Exp$.

There is a result of Hájek and Švejdar for Gödel-Bernays set theory GB, which is very much related to Theorem 7. ZF denotes Zermello-Fraenkel set theory, a *set sentence* is a sentence which talks only about sets, i.e. does not talk about classes.

Theorem 9 Hájek and Švejdar [49]. *For every two set sentences φ, ψ, $GB+\varphi$ has an interpretation in $GB+\psi$ iff there exists a cut J in $GB+\psi$ such that*

$$GB+\psi \ \vdash \ Con^J_{ZF+\varphi}.$$

Thus provability in ZF corresponds to Herbrand, or cut-free, provability in GB. There is a result of Friedman in [15] which characterizes interpretability of predicative extensions of theories using relative consistency of the theories, in a way similar to Theorem 8,.

We shall give some applications of the results above. First we show that in general cuts cannot be shortened to cuts closed under exponentiation. Originally this was proved by Paris and Dimitracopoulos [39].

Theorem 10 [39], [41] – Corollary 3.7. *There is a cut J in $\mathbf{Q}$ which cannot be shortened to a cut I closed under exponentiation. The same holds for $I\Delta_0$ and $I\Delta_0 + \Omega$.*

Proof. Take a cut J which is a model of $I\Delta_0$ and such that $\mathbf{Q} \vdash HCon^J_{\mathbf{Q}}$. (We do not know if $\mathbf{Q}$ is sequential, but there is a finite theory T which is sequential and which is interpretable in $\mathbf{Q}$.) Suppose there is a cut I in $\mathbf{Q}$ closed under exponentiation. For any fixed k we can transform k-proofs into Herbrand proofs whose size is bounded by a fixed time iterated exponential function. Thus we get all $Con_{\mathbf{Q},k}$ on a single cut I which is impossible by Theorem 6. □

A situation, where in one system the proofs are much shorter than in the other one, is usually called *speed-up*. Next application proofs that there is no elementary bound on the speed up of the ordinary proofs with respect to the Herbrand or cut-free proofs. This is a weaker version of Statman's result [48]. We denote by 2^x_k the k-times iterated 2^x.

Theorem 11 [41] – **Corollary 3.8.** *There is* <u>*no*</u> *k such that for every n the following holds: if there is a proof of some φ of length $\leq n$, then there is a Herbrand (or cut-free) proof of φ of length $\leq 2^n_k$.*

Proof. Suppose the contrary was true. Let T be a finite sequential fragment of arithmetic in which it is provable. Take a cut J which is a model of $I\Delta_0$ and such that $T \vdash HCon^J_T$. Take I such that

$$T \ \vdash \ I(x) \rightarrow J(2^x_k).$$

Then clearly $T \vdash Con^I_T$ which is a contradiction with Theorem 4. □

Now we consider the speed-up of GB over ZF. I want to use this opportunity to correct a mistake in the paper [41]. There I stated the following theorem (Theorem 4.1):

Theorem 12. *There exists a polynomial $p(x)$ such that, for every proof d of a set sentence φ in GB, there exists a proof $d' \leq 2^0_{p(d)}$ of φ in ZF.*

No proof was given there, only an idea of a proof. It has turned out that this idea does not work so easily. Fortunately the theorem is true, such a bound was proved recently by Solovay [47]. Moreover he proved the theorem with a much more precise bound $2^{|\varphi|}_{O(\sqrt{|d|})}$.

Using the above results it is possible to prove a similar nonelementary speed-up of GB over ZF with respect to set sentences [41, Corollary 4.5] (i.e. the common language of the two theories).

Such a speed up holds for other pairs of theories, where one is the predicative extension of the other; e.g. PA and its predicative extension ACA_0. Ignjatović [25] considered the pair PRA (Primitive Recursive Arithmetic) and $I\Sigma_0$. It is a classical result that $I\Sigma_0$ is conservative over PRA, if restricted to the common language, but $I\Sigma_0$ is not a predicative extension of PRA. Though, as Ignjatović proved, there is a similar speed-up of $I\Sigma_0$ over PRA.

4.2

We now turn to the question about the lengths of proofs of finite consistencies $Con_T(n)$. This has a philosophical motivation, as discussed above, but also the bounds on the lengths of proofs of $Con_T(n)$ enable us to get better estimates for speed-up theorems, which we have mentioned above.

We need to introduce some notation. The numeral corresponding to a number n will be denoted by $\underline{n}$ and it is the term

$$a_0 + ((\underline{1} + \underline{1}) \cdot (a_1 + ((\underline{1} + \underline{1}) \cdot (a_2 + \ldots)))),$$

where $a_0, \ldots, a_k \in \{0, 1\}$ and $n = \sum_0^k 2^i a_i$. The syntactical objects are first represented as 0-1 sequences in a natural way and then coded by numbers with a given diadic representation.

The statement that a sentence φ has a proof of length $\leq n$ in a theory T is usually written as $T \vdash^n \varphi$. However it is more convenient to use the notation where $\|\varphi\|_T$ denotes the length of a shortest proof in T (and ∞ if it is not provable).

Theorem 13 Friedman [14], Pudlák [42, 43]. [7]

(i) Let T be a consistent, $\mathbf{Q} \subseteq T$, T axiomatized by a $\mathcal{P}$ set of axioms. Then for some $\varepsilon > 0$ and every n

$$\|Con_T(\underline{n})\|_T \geq n^{\varepsilon}.$$

(ii) If T is sequential and finitely axiomatized, then

$$\|Con_T(\underline{n})\|_T = O(n).$$

First note that $\underline{n}$, hence also $Con_T(\underline{n})$ has length logarithmic in n, thus the theorem gives nontrivial information. Let us also note that (ii) can be proved also for theories axiomatized by a suitable schema, in particular it holds for PA and ZF, however the bound is only polynomial, see [42, Theorem 5.5].

The proof of the lower bound is in a sense a finitization of the proof of second Gödel's incompleteness theorem. We take a diagonal sentence $\delta(\underline{n})$ such that

$$T \vdash \delta(\underline{n}) \equiv (\|\ulcorner\delta(\underline{n})\urcorner\| > \underline{n})$$

and show that it does not have a proof in T of length $\leq n$. Then we show that $Con_T(\underline{m})$ implies $\delta(\underline{n})$ by a polynomial size proof with m also polynomially bounded by n.

It turns out that the critical part of the proof is the estimate of the length of a proof of the sentence $\|\ulcorner\varphi\urcorner\|_T \leq \underline{n}$, i.e. how long a proof do we need for proving that φ has a proof of length $\leq n$. The smaller bound we get, the larger bound we get for $\|Con_T(\underline{n})\|_T$, more precisely we get a lower bound which is asymptotically equal to the inverse of the upper bound, see [43, Lemma 2.1].

To see why there is a polynomial upper bound on $\|\ulcorner\varphi\urcorner\|_T \leq \underline{n}$, we shall introduce the following concept, cf. [29, §1]. We say that $\rho(x_1, \ldots, x_k)$ *$\mathcal{P}$-numerates* a relation $R \subseteq N^k$ in a theory T, if

$$R(n_1, \ldots, n_k) \;\equiv\; T \vdash \rho(\underline{n}_1, \ldots, \underline{n}_k),$$

and there exists a Turing machine M running in polynomial time such that M produces a proof of $\rho(\underline{n}_1, \ldots, \underline{n}_k)$ on input $(\underline{n}_1, \ldots, \underline{n}_k)$, whenever $R(n_1, \ldots, n_k)$ is true. Thus, in particular

$$R(n_1, \ldots, n_k) \;\rightarrow\; \|\rho(\underline{n}_1, \ldots, \underline{n}_k)\|_T \leq p(\log n_1, \ldots, \log n_k),$$

for some polynomial $p(x_1, \ldots, x_k)$. The definition of *polynomial numeration* in [42] is different (it is essentially the *$\mathcal{NP}$-numeration* of [29]) but the theorem below holds also with the stronger concept of $\mathcal{P}$-numeration.

[7] As noted above, Friedman proved only part (i).

Theorem 14 [42] – Theorem 3.2, [29] – Lemma 1.1. *Let T be a consistent, $\mathbf{Q} \subseteq T$, T axiomatized by a $\mathcal{P}$ set of axioms. Then the following are equivalent*

(1) R is in $\mathcal{P}$;

(2) there exists a polynomial numeration of R in T.

The basic property of every proof system is that the proofs can be checked in polynomial time. Also the length of a proof can be computed in polynomial time. Thus we get a polynomial bound on $\|\ulcorner\varphi\urcorner\|_T \leq \underline{n}$, hence the bound (i) of Theorem 13.

It would be interesting to know the minimal length of a proof of $Con_T(\underline{n})$ exactly, or at least asymptotically. We do not see how to reduce the gap in the ordinary calculus. If we however include an additional rule, we can get a lower bound very close to n. This rule is called *Rule C*. This rule enables to name an object whose existence is proved, i.e. formally

$$\frac{\exists x\ \varphi(x)}{\varphi(c)},$$

where c is a new constant. Using the constant c, we do not have to repeat its property $\varphi(x)$ in the proof, thus the proof becomes shorter. (It is, however, necessary to code the constants efficiently, we cannot use Hilbert's ε-terms.) In this calculus we get a better upper bound on $\|\ulcorner\varphi\urcorner\|_T \leq \underline{n}$, which gives, see [43, Theorem 2.9],

$$\|Con_T(\underline{n})\|_T \ = \ \Omega\left(\frac{n}{(\log n)^2}\right),$$

($\Omega(\ldots)$, as usual, denotes a lower bound $\varepsilon \cdot (\ldots)$ for some constant $\varepsilon > 0$).

The upper bound of Theorem 13 is obtained as follows. We define a partial satisfaction relations for formulas of quantifier complexity m. This is done inductively and the technique of Ferrante–Rackoff–Solovay produces such formulas of linear size. To prove their properties we need a proof of size m^2. This gives the following bound:

$$\|Con_{T,m}\|_T = O(m^2).$$

Furthermore we use:

Lemma 15 [43] – Lemma 3.2. *An optimal size proof of length n contains only formulas of quantifier complexity $O(\sqrt{n})$.*

Thus we need the satisfaction relation only for formulas of quantifier complexity $m = O(\sqrt{n})$, which gives the linear upper bound. □

In order to see that the lower bound on $\|Con_T(\underline{n})\|_T$ is very useful, let us derive Theorem 4 from Theorem 13(i). We need the following lemma.

Lemma 16 [42] – Lemma 2.2. *Let J be a cut in T, then the sentences $J(\underline{n})$ have proofs of length $p(\log n)$, for some polynomial p.*

To prove the lemma, take a cut I contained in J and closed under plus and times. Since $\underline{n}$ is a term with $O(\log n)$ arithmetic operations, we get a proof of $I(\underline{n})$ of length $p(\log n)$. But J contains I, so we get immediately $J(\underline{n})$.

To finish the proof of Theorem 4, observe that Con_T^J is equivalent to $\forall x(J(x) \rightarrow Con_T(x))$. So if this sentence were provable in T, we would get proofs of $Con_T(\underline{n})$ of length $p(\log n) + O(\log n)$, which is asymptotically less than n^ε. $\square$

The formulas defining partial satisfaction relation have also been used for a proof of *the small reflection principle* in $I\Delta_0 + \Omega$ by Verbrugge [52].

We shall now consider some speed-up consequences of the lower bound on $\|Con_T(\underline{n})\|_T$. First we observe that we can prove a similar bound for other expressions instead of the numeral $\underline{n}$, we only need to have polynomial proofs that these expressions evaluate to their numerical value. Thus we get

$$\|Con_T(2^0_{\underline{n}})\|_T \;\geq\; \left(2^0_n\right)^\varepsilon ,$$

for some $\varepsilon > 0$. Let us note, by passing, that this is an example of an explicit formula with a very long proof. The same proof can be performed with restricted provability, for a sufficiently large k, as the restriction. Thus we get the first part of the following theorem. The second part is proved using cuts, see [43]. ($\|\varphi\|_{T,k}$ denotes the shortest k-proof of φ in T.)

Theorem 17 [43] – Propositions 4.1 and 4.4. *(i)* $\|Con_T^k(2^0_{\underline{n}})\|_{T,k} \;\geq\; \left(2^0_n\right)^\varepsilon$, *where* $\varepsilon > 0$ *and* k *are constants,* k *sufficiently large;*

(ii) $\|Con_T^k(2^0_{\underline{n}})\|_T \;=\; O(n^2)$.

If T is finite and $k = |T|$, then a cut-free proof in T is essentially a k-proof. Thus this result gives a speed-up n to $2^0_{\varepsilon\sqrt{n}}$ for ordinary proofs and cut-free proofs. Up to the ε this is the best possible. This follows from the well-known upper bound $2^n_{O(d)}$ on cut-elimination for a proof of size n and quantifier complexity d, and Lemma 15.

Similar methods give the same speed up n to $2^0_{\varepsilon\sqrt{n}}$ for proving set formulas in GB and ZF. This was suggested in [43] and proved by Solovay [47]; he also showed that this speed up is optimal up to the ε.

4.3

As another application of the lower bound on the lengths of proofs of finite consistency statements, we consider a seemingly unrelated problem from model theory. Let M be a model of a fragment of arithmetic. For an $a \in M$, we denote by $M|_a$ the structure obtained by restricting M to the interval $[0, a]$, where we replace functions by relations. The question is, how much does the structure of $M|_a$ determine the structure of $M|_b$ for some $b \in M$, $b > a$. This means that for a definable function such that $f(a) = b$, we want to know, if there exists a Δ_0 formula $\varphi(x)$ such that $M \models \exists c < f(a)\ \varphi(c)$ while for some other model

$K \not\models \exists c < f(a)\ \varphi(c)$, or vice versa, where we require $M|_a = K|_a$. (Clearly, f must grow faster than any term.)

There is a series of papers on this subject, which gradually decreased the bound f: Paris and Dimitracopoulos [38], Hájek [21], Solovay [46]. We present the bound from [28], which is the best. We shall continue to assume that a theory contains a fragment of arithmetic and it is axiomatized by a $\mathcal{P}$ set of axioms, however we need a little more than $\mathbf{Q} \subseteq T$, namely $I\Delta_0 + Exp \subseteq T$.

Theorem 18 [28] – Theorem 2.1. *Let M be a countable nonstandard model of T and a, c nonstandard elements of M, $c \leq a$. Then there exists a countable model K of T such that $a \in K$ and*

1. $M|_a = K|_a$;
2. $M|_{2^a} \subseteq K$;
3. $K \models \neg Con_T(a^c)$.

The proof of this theorem uses the lower bound of Theorem 13, it is essential for getting better bounds than the previous ones. Furthermore it uses the Omitting Type Theorem.

Verbrugge [52] sharpened the theorem by proving that assuming $M \models Con_T(a^k)$, for all $k \in \mathbf{N}$, we can find a K with the properties above and for which $M \models Con_T(a^k)$, for all $k \in \mathbf{N}$.

Such results are interesting because they are often connected with open problems in complexity theory. Unfortunately, nobody has succeeded in solving such an open problem using these methods so far. We shall state one corollary of the above theorem which is a result of this type.

Corollary 19 [28] – Proposition 2.6. *Suppose that T proves that every Δ_0-definable set is in $\mathcal{NP}$. Then for every countable model M of T and nonstandard elements $c \leq a \in M$, there exists a model K of T such that $a \in K$ and*

1. $M|_a = K|_a$;
2. $K \models \neg Con_T((\log a)^c)$.

Thus under the assumption that T proves that every Δ_0-definable set is in $\mathcal{NP}$, we can find a K where the difference is much closer to the common segment.[8]

4.4

Theorem 13 settles the problem about the lengths of proofs of $Con_T(\underline{n})$ in T (up to a polynomial gap). If S is essentially stronger than T, namely if $S \vdash Con_T$, then S proves $Con_T(\underline{n})$ by a proof of length $O(\log n)$ (just derive $Con_T(\underline{n})$ from $\forall x\ Con_T(x)$ by substitution). However if S is *weaker* than T, the problem

[8] Note that proofs of length $(\log a)^c$ have Gödel numbers of size about $K^{(\log a)^c}$, K a constant.

remains open. A natural conjecture is e.g. that $\|Con_T(\underline{n})\|_S$ is exponential, if $T \vdash Con_S$. From the foundational point of view the most interesting question is if there is an S which proves all sentences $Con_T(\underline{n})$ by polynomial size proofs. These problems turn out to be closely related to open problems in complexity theory and propositional calculus; we are going to consider these relations in this section.

A proof system for propositional calculus P is any sound and complete system where the proofs can be checked in polynomial time. Formally it means that we have a polynomial time Turing machine M which accepts an input (t, d) only if t is a propositional tautology and for every propositional tautology t there exists a string d such that M accepts the input (t, d). The question whether there exists a propositional proof system in which all tautologies have polynomial size proofs is equivalent to the open problem $\mathcal{NP} = co\mathcal{NP}$?. Our present knowledge does not rule out that such a system is just the usual proof system based on a finite number of axiom schemas and *Modus Ponens*. We introduce a quasiordering relation on propositional proof systems by defining

> $P \leq Q$ if there exists a polynomial p such that for every tautology t, if t has a proof of length n in Q, then t has a proof of length $\leq p(n)$ in P.

In words it means that, up to a polynomial increase, the system P is at least as efficient as the system Q. Even if $\mathcal{NP} \neq co\mathcal{NP}$ it is possible that there exists a least element in the quasiordering.

To simplify the matter, we shall assume throughout this section that *all theories are finite*. Let us first note that if $S \vdash \mathcal{NP} = co\mathcal{NP}$, (it suffices $S \vdash \mathcal{NEXP} = co\mathcal{NEXP}$), then S proves all sentences $Con_T(\underline{n})$ by polynomial size proofs. Roughly speaking this is because the condition of "checking all proofs of a given length" can be replaced by "checking a single one".

A set of strings X is called *sparse*, if for every length n, there are only polynomially many elements of this length in X.

The following theorem relates important problems from three different fields.

Theorem 20 [29] – Theorem 2.1. *The following are equivalent:*

(1) There exists a finite fragment S of arithmetic such that for every finitely axiomatized consistent theory T there exists a polynomial p such that $\|Con_T(\underline{n})\|_S \leq p(n)$, for every $n \in N$.

(2) There exists a propositional proof system which is a least element in the quasiordering, (i.e. which is the most efficient with respect to the lengths of proofs).

(3) For every set $A \in co\mathcal{NP}$ there exists a nondeterministic turing machine M which accepts A and such that for every $X \subseteq A$, $X \in \mathcal{P}$, X sparse, M accepts all elements of X in polynomial time, (i.e. M is efficient on all "easy" subsets).

To give an idea about the proof, we shall show a construction of a propositional proof system from a theory and a construction of a theory from a propositional proof system. Let $Taut(x)$ be a suitable formalization of the concept of a propositional tautology, (namely we take the formula $\forall y\ Sat(x,y)$, where $Sat(t,a)$ is a $\mathcal{P}$-numeration of the predicate t *is satisfied by an assignment* a).

Let T be a theory. Then we define a propositional proof system $P(T)$ by saying that d is a proof of tautology t, iff d is a proof of the sentence $Taut(\lceil t\rceil)$ in T.

Now, let P be a propositional proof system. Let $Prf(t,d)$ be a $\mathcal{P}$-numeration of the proof predicate of P. Then we construct a theory $S(P)$ by taking a suitable basic fragment of arithmetic T_0 and adding the reflection principle

$$\forall x, y(Prf(x,y) \rightarrow Taut(x)).$$

Furthermore we need to translate some first order sentences into propositional formulas. Thus in particular $Con_T(\underline{n})$ can be translated onto a propositional tautology $t_{Con_T(\underline{n})}$ whose length is polynomial in n.

Very little progress has been done in solving problems in complexity theory such as $\mathcal{NP} = co\mathcal{NP}$? or the problem in Theorem 20. The only successful approach is based on showing that the problems cannot be solved using certain restricted means. The idea is to use the so called *relativized problems*, which means that we take a suitable set A, called *an oracle*, and allow the Turing machines mentioned in the problems to use information about this set for free. Thus we obtain a relativized class $\mathcal{X}^A$ for every complexity class $\mathcal{X}$. E.g. one can construct A and B such that $\mathcal{NP}^A = co\mathcal{NP}^A$ and $\mathcal{NP}^B \neq co\mathcal{NP}^B$ which proves that this problem cannot be solved using techniques which relativize. In [29, Theorem 4.1] we show that there is an oracle C such that the statement (3) of Theorem 20 is false, if relativized to C. Together with the relativization $\mathcal{NP}^A = co\mathcal{NP}^A$ this gives a kind of independence for the problems in Theorem 20.

Also using relativizations, Verbickij [51] showed that (3) is weaker than $\mathcal{NEXP} = co\mathcal{NEXP}$, thus it seem really hard to refute the statements in Theorem 20.

4.5

We consider two more applications of the lower bound on finite consistency statements.

The first one concerns the problem of separation of fragments of arithmetic. Buss [6] introduced bounded arithmetic S_2 and its fragments S_2^i. The language of these theories contains the usual language of arithmetic plus the function symbols $\lfloor 1/2x \rfloor$, $|x|$, and $x\#y$ whose interpretation is

$$|x| = \lceil \log_2(x+1) \rceil \text{ and } x\#y = 2^{|x|\cdot|y|}.$$

S_2 is axiomatized by a finite set of basic axioms and the following schema of induction

$$\varphi(0) \wedge \forall x(\varphi(\lfloor 1/2x \rfloor) \rightarrow \varphi(x)) \rightarrow \forall \varphi(x),$$

for all bounded formulas. This theory is a conservative extension of $I\Delta_0+\Omega$. The fragments S_2^i are defined by restricting the schema of induction to Σ_i^b-formulas, where the classes Σ_i^b are defined so that the formulas define just the sets in the Polynomial Hierarchy Σ_i^p. In particular Σ_1^b defines the $\mathcal{NP}$ sets, so the theory S_2^1 is closely related to the class $\mathcal{NP}$.

Similarly as it is an open problem if the Polynomial Hierarchy is increasing, it is an open problem if the hierarchy of theories S_2^i is increasing. The subtheories of Peano Arithmetic $I\Sigma_i$ can be separated by sentences $Con_{I\Sigma_i}$ which are provable in $I\Sigma_{i+1}$, but not in $I\Sigma_i$. This is not possible for S_2^i's, since the whole S_2 does not prove even $Con_{\mathbf{Q}}$. In fact it is even worse:

Theorem 21 Wilkie and Paris [55].

$$I\Delta_0 + Exp \nvdash Con_{\mathbf{Q}}.$$

A natural modification of the concept of consistency has been considered, where one talks only about proofs of contradiction which use only *bounded formulas*. Such a restricted statement of consistency of a theory T is denoted by $BDCon_T$. Though it looks like the right concept for Bounded Arithmetic, it turns out that it is still too strong for such a separation:

Theorem 22 [44].

$$S_2 \nvdash BDCon_{S_2^1}.$$

Thus we cannot separate in such a way even S_2 from S_2^1. This improves a former result of Buss [6], where he showed that $S_2^{i+1} \vdash BDCon_{S_2^i}$ holds for at most one i. The result has been further improved by Takeuti [50] and Buss and Ignjatović [7]. Let us note that Krajíček and Pudlák [30], and Krajíček and Takeuti [32] found sentences that separate such fragments, if the fragments are different, however, since these sentences are rather "combinatorial" than "logical", it seems impossible to use diagonalization arguments for them.

The proof of the theorem is based on a modification of the lower bound on the lengths of proofs of $Con_T(\underline{n})$, in which we consider the Gödel numbers of the proofs instead of their lengths. Let $\gamma(\varphi)$ denote the minimal Gödel number of a proof of the sentence φ in S_2^1. Let $BDCon^{\#}_{S_2^1}(x)$ denote that there is no proof of contradiction in S_1^2 whose Gödel number is $\leq x$; it is a bounded formula in the language of S_2. Using a similar proof as above, we get:

Lemma 23 [44] – Lemma 1. *For every term $s(x)$, there exists a term $r(x)$ such that for all but finitely many n*

$$\gamma\left(BDCon^{\#}_{S_2^1}(r(\underline{n}))\right) > s(n).$$

The proof of Theorem 22 uses another lemma, which states that for every bounded formula φ

$$\gamma(\varphi(\underline{n})) \leq p(n\#n),$$

for some polynomial p, if $S_2 \vdash \forall x\ \varphi(x)$. The proof is finished by taking $s(x) = x\#x\#x$, which grows faster than $p(n\#n)$ for any polynomial p, and $\varphi(x) = BDCon^{\#}_{S^1_2}(x)$. $\square$

The last application is to the speed-up obtained by *proving the provability of* φ instead of proving φ. Parikh [37] has shown that for Peano arithmetic this speed-up cannot be bounded by any recursive function. We observe that we can get concrete examples for such a speed up from consistency statements. E.g. we have the following upper bound:

Proposition 24 [22] – Theorem 1,(b). *Let $Prov_T$ be a suitable formalization of the provability in T. Then*

$$\|Prov_T\left(\ulcorner Con_T(2^0_{\underline{n}})\urcorner\right)\|_T \ \leq\ p(n),$$

for some polynomial p.

With the bound

$$\|Con_T(2^0_{\underline{n}})\|_T \ \geq\ \left(2^0_{\underline{n}}\right)^{\varepsilon},$$

$\varepsilon > 0$, which we have already considered, we get a more than elementary speed-up for an explicit sequence of sentences. It would be interesting to prove such speed-ups for more natural mathematical sentences. Presently this can be done only for sentences expressing existence of fast growing functions. We conclude our survey with such a result.

Theorem 25 [22] – Theorem 3. *(a) There exists a polynomial p such that for all n*

$$\|Prov_{I\Delta_0+\Omega}\left(\ulcorner \exists y(y = 2^0_{2^0_{\underline{n}}})\urcorner\right)\|_{I\Delta_0+\Omega} \ \leq\ p(n).$$

(b) For every sufficiently large n,

$$\|\exists y(y = 2^0_{2^0_{\underline{n}}})\|_{I\Delta_0+\Omega} \ \geq 2^0_n/4.$$

5 Conclusions

We have shown that bounds on finite consistency statements are useful for several purposes. They can be used to show speed-up for different logical calculi, speed-up between some theories and speed-up when we prove the existence of proofs instead of constructing them. They are also explicit examples of formulas with long proofs.

The problem about the lengths of proofs of $Con_T(\underline{n})$ has been solved only for T itself and stronger theories.[9] If S is weaker than T, then $\|Con_T(\underline{n})\|_S$ can

[9] To be quite precise, if S proves Con_T, then, trivially $Con_T(\underline{n})$ has proofs of length $O(\log n)$, for $S = T$ it is about n, but for S stronger than T which does not prove Con_T, we also do not know; we only know that it is between these values.

be exponential in n. I think this is a fundamental problem. The connection with problems in complexity theory shows that it is a very difficult problem.

Suppose we proved that for every S, there exists T such that $\|Con_T(\underline{n})\|_S$ is exponential. Then this would solve a lot of problems in complexity theory, in particular it would imply $\mathcal{P} \neq \mathcal{NP}$. It does not seem to be possible to use some higher order argument from logic to solve the problems in complexity theory, without really going into deep combinatorial problems. But still, maybe we are just missing some simple clever idea. I imagine that Hilbert had similar feelings about the consistency problem before Gödel solved it.[10] The problem of consistency is a statement about application of finite rules finitely many times, so why anything else than finite combinatorics is needed? Though Gödel had to do some technical work, his basic trick was the selfreference, (essentially the paradox of liar, which was known already to ancient Greeks), which is a higher order argument which has little to do with finite combinatorics.

Let us conclude by repeating a conjecture that we stated in [41]:

Conjecture *If T is sufficiently strong, then $\|Con_{T+Con_T}(\underline{n})\|_T$ is not bounded by a polynomial.*

The intuition is that $T + Con_T$ is essentially stronger than T, since it proves $Con_T(\underline{n})$ by proofs of logarithmic length, while T needs n^ε, therefore T should not have short proofs of $Con_{T+Con_T}(\underline{n})$. As stated above, the positive answer would solve fundamental problems in complexity theory.

Maybe the solution of $\mathcal{P} = \mathcal{NP}$? etc. will be as surprising as was Gödel's ...

References

1. Miklos Ajtai. The complexity of the pigeonhole principle. In *Proceedings of the 29th Annual Symposium on Foundations of Computer Science*, pages 346–355, Piscataway, New Jersey, 1988. IEEE Computer Society.
2. Miklos Ajtai. The independence of the modulo p counting principles. In *Proceedings of the 26th Annual ACM Symposium on the Theory of Computing*, pages 402–411, New York, 1994. Association for Computing Machinery.
3. Matthias Baaz and Pavel Pudlák. Kreisel'sKreisel, G. conjecture for $L\exists_1$. In Peter Clote and Jan Krajíček, editors, *Arithmetic Proof Theory and Computational Complexity*, pages 30–39. Oxford University Press, 1993.
4. Paul Beame, Russell Impagliazzo, Jan Krajíček, Toniann Pitassi, Pavel Pudlák, and Alan Woods. Exponential lower bounds for the pigeonhole principle. In *Proceedings of the 24th Annual ACM Symposium on the Theory of Computing*, pages 200–221, New York, 1992. Association for Computing Machinery.

[10] Hilbert hoped that his consistency program will be realized using the new theory that he invented and named *Beweistheorie* [24], (nowadays his theory is called *ϵ-calculus*).

5. A. Bezboruah and J. Shepherdson. Gödel'sGödel, K. second incompleteness theorem for Q. *Journal of Symbolic Logic*, pages 503–512, 1976.
6. Samuel R. Buss. *Bounded Arithmetic*. Bibliopolis, Napoli, 1986. Revision of 1985 Princeton University Ph.D. thesis.
7. Samuel R. Buss and Aleksandar Ignjatović. Unprovability of consistency statements in fragments of bounded arithmetic. In preparation, 1993.
8. R. Carnap. *Logische Syntax der Sprache*. Springer-Verlag, Berlin, 1934.
9. Stephen A. Cook. Feasibly constructive proofs and the propositional calculus. In *Proceedings of the Seventh Annual ACM Symposium on the Theory of Computing*, pages 83–97, New York, 1975. Association for Computing Machinery.
10. Martin Dowd. *Propositional Representation of Arithmetic Proofs*. PhD thesis, University of Toronto, 1979.
11. A. G. Dragalin. Correctness of inconsistent theories with notions of feasibility. In A. Skowron, editor, *Computation Theory, Fifth Symposium Proceedings*, volume 108 of *Lecture Notes in Computer Science #208*, pages 58–79, Berlin, 1985. Springer-Verlag.
12. A. Esenine-Volpin. Le programme ultra-intuitionniste des des fondements des mathematiques. In *Infinitistic Methods, Proceedings of the Symposium on Foundations of Mathematics*, pages 201–223, New York, 1961. Pergamon Press.
13. J. Ferrante and Ch. W. Rackoff. *The Computational Complexity of Logical Theories*. Lecture Notes in Mathematics #718. Springer-Verlag, Berlin, 1979.
14. H. Friedman. On the consistency, completeness, and correctness problems. Ohio State University, unpublished, 1979.
15. H. Friedman. Translatability and relative consistency II. Ohio State University, unpublished, 1979.
16. Yu. V. Gavrilenko. Monotone theories of feasible numbers. *Doklady Akademii Nauk SSSR*, 276(1):18–22, 1984.
17. Gerhard Gentzen. Die Widerspruchsfreiheit der reinen Zahlentheorie. *Mathematische Annalen*, 112:493–565, 1936. English translation in [?], pp. 132-213.
18. K. Gödel. Über formal unentscheitbare Sätze der Principia Mathematica und vewandter Systeme I. *Monatshefte für Mathematik und Physik*, 38:173–198, 1931.
19. Kurt Gödel. Über die Länge von Beweisen. *Ergebnisse eines Mathematischen Kolloquiums*, pages 23–24, 1936. English translation in *Kurt Gödel: Collected Works, Volume 1*, pages 396-399, Oxford University Press, 1986.
20. P. Hájek. On interpretability in set theories II. *Commentationes Mathematicae Universitatis Carolinae*, 13:445–455, 1972.
21. Petr Hájek. On a new notion of partial conservativity. In *Logic Colloquium'83*, pages 217–232, Berlin, 1983. Springer-Verlag.
22. Petr Hájek, Franco Montagna, and Pavel Pudlák. Abbreviating proofs using metamathematical rules. In Peter Clote and Jan Krajíček, editors, *Arithmetic Proof Theory and Computational Complexity*, pages 197–221. Oxford University Press, 1993.
23. Petr Hájek and Pavel Pudlák. *Metamathematics of First-order Arithmetic*. Springer-Verlag, Berlin, 1993.
24. D. Hilbert. *Die Grundlagen der Mathematik*, volume 5 of *Hamburger Mathematische Einzelschriften*. Teubner, 1934. A lecture presented in Hamburg in July 1927.
25. A. Ignjatović. *Fragments of First and Second Order Arithmetic and Length of Proofs*. PhD thesis, University of California at Berkeley, 1990.

26. Jan Krajíček. A note on proofs of falsehood. *Archiv für Mathematische Logic und Grandlagen der Mathematik*, 26:169–179, 1987.
27. Jan Krajíček. Lower bounds to the size of constant-depth propositional proofs. *Journal of Symbolic Logic*, 59(1):73–86, 1994.
28. Jan Krajíček and Pavel Pudlák. On the structure of initial segments of models of arithmetic. *Archive for Mathematical Logic*, 28:91–98, 1989.
29. Jan Krajíček and Pavel Pudlák. Propositional proof systems, the consistency of first-order theories and the complexity of computations. *Journal of Symbolic Logic*, 54:1063–1079, 1989.
30. Jan Krajíček and Pavel Pudlák. Quantified propositional calculi and fragments of bounded arithmetic. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 36:29–46, 1990.
31. Jan Krajíček, Pavel Pudlák, and Gaisi Takeuti. Bounded arithmetic and the polynomial hierarchy. *Annals of Pure and Applied Logic*, 52:143–154, 1991.
32. Jan Krajíček and Gaisi Takeuti. On induction-free provability. *Annals of Mathematics and Artificial Intelligence*, 6:107–126, 1992.
33. R. Montague. Theories incomparable with respect to relative interpretability. *Journal of Symbolic Logic*, 27:195–211, 1962.
34. V. P. Orevkov. Correctness of short proofs in theory with notions of feasibility. In P. Martin-Löf and G. E. Mints, editors, *COLOG-88: International Conference on Computer Logic, Tallinn, USSR, Dec. 1988, Proceedings*, Lecture Notes in Computer Science #417, pages 242–245, Berlin, 1990. Springer-Verlag.
35. S. Orey. Relative interpretations. *Journal of Symbolic Logic*, 24:281–282, 1959.
36. S. Orey. Relative interpretations. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 7:146–153, 1961.
37. R. Parikh. Existence and feasibility in arithmetic. *Journal of Symbolic Logic*, 36:494–508, 1971.
38. J. B. Paris and C. Dimitracopoulos. Truth definitions for formulas. In *Logic and Algorithmic: An International Symposium held in honour of Ernst SpeckerSpecker, E.*, Monographie #30 de L'Enseignement Mathématique, pages 318–329, 1982.
39. J. B. Paris and C. Dimitracopoulos. A note on the undefinability of cuts. *Journal of Symbolic Logic*, 48:564–569, 1983.
40. J. B. Paris and L. Harrington. A mathematical incompleteness in Peano Arithmetic. In *Handbook of Mathematical Logic*, pages 1133–1142, Amsterdam, 1977. North-Holland.
41. Pavel Pudlák. Cuts, consistency statements and interpretation. *Journal of Symbolic Logic*, 50:423–441, 1985.
42. Pavel Pudlák. On the lengths of proofs of finitistic consistency statements in first order theories. In *Logic Colloquium '84*, pages 165–196, Amsterdam, 1986. North-Holland.
43. Pavel Pudlák. Improved bounds to the lengths of proofs of finitistic consistency statements. In S. J. Simpson, editor, *Logic and Combinatorics*, volume 65 of *Contemporary Mathematics*, pages 309–331, Providence, Rhode Island, 1987. American Mathematical Society.
44. Pavel Pudlák. A note on bounded arithmetic. *Fundamenta Mathematicae*, 136:85–89, 1990.
45. C. Smoriński. Nonstandard models and related developments. In *Harvey Friedman'sFriedman, H. Research on the Foundations of Mathematics*, pages 179–229, Amsterdam, 1985. North Holland.

46. R. M. Solovay. Injecting inconsistencies into models of PA. *Annals of Pure and Applied Logic*, 44:101–132, 1989.
47. R. M. Solovay. Upper bounds on the speedup of GB over ZF. preprint, 1990.
48. R. Statman. Proof search and speed-up in the predicate calculus. *Annals of Mathematical Logic*, 15:225–287, 1978.
49. V. Švejdar. Modal analysis of generalized RosserRosser, J. B. sentences. *Journal of Symbolic Logic*, 48:986–999, 1983.
50. G. Takeuti. Some relations among systems for bounded arithmetic. *Annals of Pure and Applied Logic*, 39:75–104, 1988.
51. O. V. Verbitsky. Optimal algorithms for coNP-sets and the problem EXP=?NEXP. *Matematicheskie zametki*, 50(2):37–46, 1991. English translation in: *Mathematical Notes* 50,1-2 (1991), pp. 798-801.
52. L. Ch. Verbrugge. *Efficient Mathematics*. PhD thesis, Universiteit van Amsterdam, 1993.
53. A. Visser. The unprovability of small consistency. *Archive for Mathematical Logic*, 32:275–298, 1993.
54. A. Wilkie. On sentences interpretable in systems of arithmetic. In *Logic Colloquium '84*, pages 329–342, Amsterdam, 1986. North-Holland.
55. A. J. Wilkie and J. B. Paris. On the schema of induction for bounded arithmetical formulas. *Annals of Pure and Applied Logic*, 35:261–302, 1987.