On Gödel's Theorems on Lengths of Proofs I: Number of Lines and Speedup for Arithmetics

Samuel R. Buss*

Department of Mathematics

University of California, San Diego

La Jolla, CA 92093-0112, USA

sbuss@ucsd.edu

October 1, 1992

Abstract

This paper discusses lower bounds for proof length, especially as measured by number of steps (inferences). We give the first publicly known proof of Gödel's claim that there is superrecursive (in fact, unbounded) proof speedup of $(i+1)$-st order arithmetic over i-th order arithmetic, where arithmetic is formalized in Hilbert-style calculi with $+$ and $\cdot$ as function symbols or with the language of PRA. The same results are established for any weakly schematic formalization of higher-order logic; this allows all tautologies as axioms and allows all generalizations of axioms as axioms.

Our first proof of Gödel's claim is based on self-referential sentences; we give a second proof that avoids the use of self-reference based loosely on a method of Statman's.

1 Introduction

This paper presents the first publicly known proof of a theorem of Gödel on lengths of proofs; a sequel to this paper presents a proof of another of

*Supported in part by NSF grants DMS-8902480 and DMS-9205181.

Gödel's claims on lengths of proofs. Gödel's first theorem on lengths of proofs appeared in his 1936 paper and states that there is no recursive function bounding the proof speedup of $(i+1)$-st order arithmetic over i-th order arithmetic. Gödel's second second theorem on lengths of proofs appears in a recently discovered letter to von Neumann; this letter contains the claim that a Turing machine which, given a formula and an integer n determines if the formula has a proof in first-order logic of $\leq n$ symbols, must have runtime $> \epsilon n$ steps infinitely often, for some constant ϵ.

Our proofs of these two theorems of Gödel's use self-reference and are presumably similar to Gödel's own proof methods. In this paper, we also give a proof of the first theorem that avoids the use of self-reference, based partly on methods of Statman. The proof of the second theorem is in the sequel [2] to this paper.

We begin by giving the basic definitions and an overview of this paper. The i-th order theory of arithmetic is denoted Z_{i-1} — so Z_0 is first-order Peano arithmetic, Z_1 is second-order arithmetic, etc. Z_i has as non-logical symbols 0, S, $+$, $\cdot$ and optionally has function symbols for all primitive recursive functions.[2] The usual axioms of Z_i consist of: (1) the axioms of Robinson's theory Q which define the basic properties of 0, S, $+$ and $\cdot$, and (2) universal axioms defining the primitive recursive functions, if they are in the language of Z_i, and (3) induction for all formulas of Z_i, and (4) if $i > 0$, comprehension axioms

$$(\exists X)(\forall x)(x \in X \leftrightarrow \phi(x))$$

for all formulas ϕ and variables x and X of orders $j-1$ and $j \leq i$, respectively. For the purposes of this paper, Z_i must be formalized in a Hilbert-style system, not in the Gentzen-style G^1LC of Takeuti [16].[3] If the rules of inference for Z_i are given by a finite set of schemata in the sense of Parikh [11] then we have a *schematic* axiomatization for Z_i. For a schematic axiomatization, the valid rules of inference are specified by schemata into which any formulas, terms and variables may be substituted to obtain valid rules of inference, subject to syntactic conditions on variable names.

[2] We are using $+$ and $\cdot$ as function symbols but our methods work equally well if they are three-place relation symbols.

[3] The difficulty with G^1LC is that it allows the use of substitution of formulas for variables. This is logically equivalent to the use of the comprehension axiom; however, we are currently unable to establish the analogue of Theorem 6 for G^1LC.

Schematic axioms are specified by schemata with no premises. However, there are non-schematic ways to specify axioms as well: two notable examples are that all (instances of) tautologies may be allowed as axioms and that any universal quantification (generalization) of an axiom may be taken as an axiom. If Z_i is formalized in a Hilbert-style system with all rules of inference given by a finite set of schemata except for axioms which may also be given as tautologies and as obtained by arbitrary generalization, then we say Z_i is *weakly schematic*. All the methods of this paper apply to Z_i as either schematic or weakly schematic. See section 2.3 below for the complete definitions of schematic and weakly schematic proof systems.

Because we allow weakly schematic proof systems, the effect is that the work of this paper applies to Z_i under almost any Hilbert-style system, for instance, the proof systems found in textbooks by Enderton, Mendelsohn, Kleene and Shoenfield.

Definition The *symbol-length* of a proof is the total number of occurrences of symbols in the proof (including symbols occurring in the subscripts of variables in, say, base two notation).

The *step-length* of a proof is equal to the total number of formulas in the proof. The step-length is obviously also equal to the total number of uses of axioms and inferences in the proof.

Gödel [6] stated the following theorem:

Theorem 1 *For $i \geq 0$ and h any recursive function, there is an infinite family $\mathcal{F}$ of Π^0_2-formulas such that for any sentence $\phi \in \mathcal{F}$, ϕ is provable in Z_i and if k is the step-length of the shortest Z_{i+1}-proof of ϕ then the shortest Z_i-proof of ϕ has step-length greater than $h(k)$.*

In section 2.1 we shall prove the following "symbol-length" analogue of Gödel's theorem using a self-referential formula that says "This formula is not provable in $\leq f(x)$ symbols". This theorem can be found in Mostowski [10] and in Ehrenfeucht-Mycielski [3].

Theorem 2 *Let $i \geq 0$ and f be any recursive function. Then there is an infinite family $\mathcal{F}$ of Π^0_1-formulas such that*

(1) *for all $\phi \in \mathcal{F}$, $Z_i \vdash \phi$, and*

(2) *for* $\phi \in \mathcal{F}$*, if* k *is the least integer such that* $Z_{i+1} \vdash\!\frac{k\ symbols}{}\, \phi$ *then it is not the case that* $Z_i \vdash\!\frac{f(k)\ symbols}{}\, \phi$*.*

Although Theorem 2 is well-known, we include a proof here, partly since it is a nice application of H. Friedman/Pudlák lower bounds on proof lengths of partial consistency statements, and partly to introduce the general method for proving Theorems 1 and 3. The Friedman/Pudlák results we exploit consist of lower bounds on the symbol length of proofs of partial consistency statements $Con_T(f(n))$ which state that theory T has no proof with symbol-length $\leq f(n)$ of $0 = 1$.

In section 2.2 we prove a strengthened version of Gödel's theorem giving an arbitrary (instead of recursive) speedup for Z_{i+1} over Z_i for proof length measured in steps:

Theorem 3 *Let* $i \geq 0$*. Then there is an infinite family* $\mathcal{F}$ *of* Π^0_1*-formulas such that*

(1) *for all* $\phi \in \mathcal{F}$*,* $Z_i \vdash \phi$*, and*

(2) *there is a fixed* $k \in \mathbb{N}$ *such that for all* $\phi \in \mathcal{F}$*,* $Z_{i+1} \vdash_{\overline{k\ steps}} \phi$*, and*

(3) *there is no fixed* $k \in Z_i$ *such that for all* $\phi \in \mathcal{F}$*,* $Z_i \vdash_{\overline{k\ steps}} \phi$*.*

Theorem 3 obviously implies that Theorem 1 holds for an arbitrary function h. The fact that Theorem 3 produces a family of Π^0_1-formulas instead of Π^0_2-formulas is based on Matijasevič's theorem. The proof in section 2.2 of Theorem 3 uses a self-referential statement that says "This formula is not provable in $\leq x$ steps". However, unlike the symbol-length case of section 2.1, we are unable to obtain any lower bounds on the step-length of partial consistency statements. In order to prove Theorem 3 when Z_i is formalized as a weakly schematic system (especially when all tautologies are axioms of Z_i), we must, in section 3, extend results of Parikh and Krajíček bounding the number of logical complexity of formulas in a proof in terms of the step-length of the proof.

Statman [15] gave a similar unbounded proof-speedup for first-order schematic theories with infinite models. In section 3, we rework his proof dropping the restriction that the theory have infinite models, and allowing weakly schematic theories, and also allowing higher-order logic. This provides an alternative proof of Theorem 3 that avoids the use of self-reference.

Although this latter proof is probably not what Gödel had in mind, it still provides interesting insight into the structure of proofs.[4]

Our proofs of Theorem 3 give the first known proofs (except presumably to Gödel) of Gödel's theorem (Theorem 1) for so-called Hilbert-style axiomatizations of higher-order arithmetic where addition and multiplication are taken to be function symbols. In the case where Z_i has at most one unary function symbol and has no nonunary function symbols (so addition and multiplication are three-place relation symbols), Theorem 3 has already been proved for $i = 0$ by Parikh [11] and for $i > 0$ by Krajíček [7]. Parikh's proof used a precursor of Theorem 6 below and Kreisel's conjecture; Krajíček used a sharpened version of Parikh's precusor of Theorem 6 and a diagonalization argument. The methods of Kreisel-Wang [9], based on self-referential statements, can be easily extended to prove the analogue of Theorem 3 for ϵ-calculus formalizations of higher-order arithmetic. More discussion of prior work on this problem can be found in [12].

I'd like to thank Jan Krajíček for helpful suggestions and comments. In addition, I thank Daniel Leivant for extensive discussions.

2 Proof speedup of Z_{i+1} over Z_i

In section 2.1 we prove Theorem 2, which is the analogue of Gödel's lengths of proofs theorem with proof length measured by the number of symbols in the proof. The development in section 2.1 is carried out by using lower bounds of Pudlák and H. Friedman for the lengths of partial self-consistency proofs. This yields a very satisfactory proof of Theorem 2. In section 2.2, Theorem 3 is proved, which concerns proof length measured by the number of steps. It is more difficult to handle this case where proof length is measured by number of steps and we do not know if the corresponding lower bounds to lengths of partial consistency proofs hold for proof length measured by number of steps. Instead, Theorem 3 is proved by using a self-referential formula directly. To complete the proof of Theorem 3, we must obtain upper bounds on the logical complexity of formulas in a proof in terms of the number of steps in the proof; this is carried out for weakly schematic proof systems in section 2.3.

[4]Our best guess is that Gödel had in mind a proof which used both self-referential statements and numeralwise representability of functions.

Definition Let T be a theory and ϕ a formula. We write $T \vdash_{\overline{k \text{ steps}}} \phi$, or sometimes just $T \vdash_{\overline{k}} \phi$, to mean that there is a T-proof of ϕ of $\leq k$ steps. We write $T \vdash^{\underline{k \text{ symbols}}} \phi$, or sometimes just $T \vdash^{\underline{k}} \phi$, to mean that there is a T-proof of ϕ of $\leq k$ symbols.

Definition For ϕ a formula or t a term, $|\phi|$ or $|t|$ denotes the number of symbols in ϕ or t. For P a proof or w the Gödel number of a proof, $|P|$ or $|w|$ denotes the number of symbols in the proof. We also write $|n|$ to denote the length of the binary representation of an integer. (This creates two definitions for $|w|$ where w is a Gödel number; by efficient coding techniques the definitions are polynomially related, and in any event, context will serve to indicate which definition is intended.)

2.1 Symbol-Length Speedup

We shall henceforth let T denote an axiomatizable, consistent theory of arithmetic with language containing $+$ (plus) and $\cdot$ (times) as function symbols. The language may also contain other function symbols such as successor S and possibly symbols for primitive recursive functions. We shall assume that T is strong enough to arithmetize metamathematics efficiently, e.g., we assume that (an extension by definitions of) T contains one of the bounded arithmetic theories S^1_2 or $I\Delta_0{+}\Omega_1$ (see Buss [1] and Wilkie-Paris [17] for descriptions of these bounded arithmetics). For $n \geq 0$, $\underline{n}$ denotes a closed term with value n. Since $+$ and $\cdot$ are functions symbols, we can ensure that $|\underline{n}|$ is $O(\log n)$ by defining $\underline{2n}$ to be $(\underline{2}\cdot\underline{n})$ and defining $\underline{2n+1}$ to be $(\underline{2}\cdot\underline{n}+1)$.

Since T is axiomatizable, it has an axiomatization for which the axioms are recognizable in polynomial time on a multitape Turing machine. We fix some such axiomatization for T. For example, T's axioms might be given by a finite set of axiom schemata.

Definition $Con_T(x)$ is a first-order formula expressing the partial consistency statement for T that there is no T-proof of $0 = 1$ with symbol-length $\leq x$:

$$\neg\exists w[w \text{ codes a } T\text{-proof of } 0 = 1 \text{ of symbol-length} \leq x].$$

Theorem 4 *(Pudlák [13, 14], H. Friedman) Let T be a theory of arithmetic as above. Then there is an $\epsilon > 0$, such that, for all $n > 0$,*

$$\neg\, T \vdash^{\underline{n^\epsilon \textit{ symbols}}} Con_T(\underline{n})$$

Proof (Sketch) By Gödel's diagonal lemma, there is a formula ϕ such that

$$T \vdash \Big(\phi(x) \leftrightarrow (\neg T \overset{x}{\vdash} \phi(\underline{x}))\Big). \tag{1}$$

Actually, the usual statement of Gödel's diagonal lemma applies only to sentences, but it is easily seen that the diagonal lemma applies also to formulas with free variables x. Now $\phi(x)$ is of the form $\neg(\exists w)\psi(w,x)$ where ψ is a Δ_1^b-formula (and hence a Σ_1^b-formula) which expresses the condition that w is the Gödel number of a proof of the formula $\phi(\underline{x})$ of symbol-length $\leq x$. By Theorem 7.4 of Buss [1] or Theorem 6.4 of Wilkie-Paris [17] there is a polynomial $p(n)$ such that

$$T \vdash \Big(\psi(w,x) \rightarrow T \overset{p(|w|)}{\vdash} \psi(\underline{w},\underline{x})\Big);$$

the essential idea is that if $\psi(w,x)$ is true, then T can prove $\psi(w,x)$ by a direct proof which is of polynomial size. (Note that if $\psi(w,x)$ holds, then $|w| \geq |\phi(\underline{x})| > |x|$; thus the length bound $p(|w|)$ does not need to depend on $|x|$.) Since sequences are presumed to be efficiently coded in T, T also proves that $\psi(w,x) \rightarrow |w| \leq k \cdot x$ for some constant k. Thus there is a constant $b > 0$ such that

$$T \vdash \Big(\neg\phi(x) \rightarrow T \overset{x^b}{\vdash} \neg\phi(\underline{x})\Big). \tag{2}$$

On the other hand, (1) implies

$$T \vdash \Big(\neg\phi(x) \rightarrow T \overset{x}{\vdash} \phi(\underline{x})\Big). \tag{3}$$

Combining (2) and (3) we get that

$$T \vdash \forall x\Big(\neg\phi(x) \rightarrow \neg Con_T(x^c)\Big) \tag{4}$$

for some constant $c \in \mathbb{N}$. Take $\epsilon < 1/c$. Let $m > 0$ and $n = m^c$ and suppose, for sake of a contradiction, that

$$T \overset{n^\epsilon}{\vdash} Con(\underline{n}). \tag{5}$$

Now, for all $m \in \mathbb{N}$,

$$T \overset{|m|^{O(1)}}{\vdash} (\underline{m})^c = \underline{m^c} \tag{6}$$

by Lemma 7.5 of [1] or Lemma 6.3 of [17]. Intuitively, (6) is proved by noting that T can prove facts like $S(\underline{m}) = \underline{m+1}$, $\underline{m}+\underline{r} = \underline{m+r}$, and $\underline{m}\cdot\underline{r} = \underline{m\cdot r}$ with polynomial size proofs using the axioms defining the symbols S, $+$ and $\cdot$; formally one establishes this first for S, then for $+$ and then for $\cdot$ by induction on the length of m.

From (6) it follows trivially that

$$T \vdash^{|m|^{O(1)}} Con_T((\underline{m})^c) \leftrightarrow Con_T(\underline{m^c}). \tag{7}$$

Thus, by (4), (5) and (7),

$$T \vdash^{m^{c\epsilon}+|m|^{O(1)}} \phi(\underline{m}).$$

For m sufficiently large, $m^{c\epsilon} + |m|^{O(1)} \le m$ so that $T \vdash^{m} \phi(m)$. But then T also proves $\neg\phi(m)$ by the definition of ϕ, which contradicts the consistency of T.

We have shown that for m sufficiently large, it is not that case that $T \vdash^{n^\epsilon \text{ symbols}} Con(n)$ with $n = m^c$ for some $m \in \mathbb{N}$. To establish this for arbitrary n not of the form m^c, we note that $n \le \left(\lceil n^{1/c}\rceil\right)^c \le 2n$ for n sufficiently large. From this and the fact that for $n' \ge n$, $T \vdash^{|n'|^{O(1)}} Con_T(\underline{n'}) \to Con_T(\underline{n})$, we get that $\neg T \vdash^{n^{\epsilon'}} Con_T(\underline{n})$ for *all* sufficiently large n and some constant $\epsilon' \le \epsilon$. By taking ϵ' smaller if necessary, this holds also for all $n > 1$.
Q.E.D. Theorem 4

Theorem 4 can be used give exponential speedup between Z_{i+1} and Z_i; for establishing arbitrary recursive speedup, we generalize the partial consistency statement Con_T as follows.

Definition A partial recursive function f is *time constructible* if and only if there is a Turing machine M_f, such that, for all x,

$$f(x) = y \quad\Leftrightarrow\quad M_f(x) \text{ halts in exactly } y \text{ steps.}$$

We define $f(x) \ge y$ to mean that $M_f(x)$ does not halt in $< y$ steps. Note that $f(x) \ge y$ will hold whenever $f(x)$ does not converge. By convention, every time-constructible function f satisfies $f(x) \ge |x|$; the Turing machine M_f will be explicitly constructed to read its entire input, which forces $f(x) \ge |x|$ since x input in binary notation has length $|x|$.

Obviously, every partial recursive function is dominated by a time-constructible functions. Note that $f(x) \geq y$ is a recursive predicate and, in fact, $f(x) \geq |y|$ is a polynomial time predicate of x and y.

Definition $Con_T(f(x))$ is the Π^0_1-formula with free variable x expressing

$$\neg \exists m \left(\left(T \vdash^{\underline{\ m \text{ symbols}\ }} 0 = 1 \right) \wedge f(x) \geq m \right)$$

Theorem 4 can be generalized to partial consistency statements involving $f(x)$:

Theorem 5 *Let T be a theory of arithmetic as above. Then there are constants $\epsilon > 0$ and $c \in \mathbb{N}$ so that, if f is a time constructible function, then for all sufficiently large n,*

$$\neg T \vdash^{\underline{\ f(n)^\epsilon \textit{ symbols}\ }} Con_T((f(\underline{n}))^c).$$

When $f(n)$ is undefined, T does not prove $Con_T(f(\underline{n})^c)$ at all.

Here $Con_T((f(\underline{n}))^c)$ should be interpreted as the formula $Con_T(g(\underline{n}))$ where g is the time-constructible function $g(x) = f(x)^c$. For many theories T, c can be a small integer like 2 or 3 and ϵ can be any constant < 1.

Proof The proof of Theorem 5 is very similar to the proof of Theorem 4, so we indicate only the major differences and leave the details to the reader. First, let $\phi(x)$ now be the Gödel diagonal formula such that

$$T \vdash \left(\phi(x) \leftrightarrow \neg \exists y \left(\left(T \vdash^{\underline{\ y \text{ symbols}\ }} \phi(\underline{x}) \right) \wedge f(x) \geq y \right) \right). \tag{1'}$$

Now, there is a polynomial p so that

$$T \vdash \left(f(x) \geq y \rightarrow T \vdash^{\underline{\ p(|x| + y) \text{ symbols}\ }} f(\underline{x}) \geq \underline{y} \right). \tag{8}$$

The polynomial p depends on f (or rather, on the Turing machine which time-constructs f); however, the degree of p is independent of f. Hence, taking $a \in \mathbb{N}$ greater then the degree of p, we have that $p(|x| + y)$ in (8) can be replaced by $(|x| + y)^a$ for sufficiently large x, y; i.e., there is a constant C so that

$$T \vdash \forall x, y \geq C \left(f(x) \geq y \rightarrow T \vdash^{\underline{\ (|x| + y)^a \text{ symbols}\ }} f(\underline{x}) \geq \underline{y} \right). \tag{9}$$

Combining this with the reasoning from the previous proof shows that

$$T \vdash (\forall x \geq C)\left(\neg\phi(x) \to T \overset{(f(x))^b}{\vdash} \neg\phi(x)\right) \tag{2'}$$

for some constant b which, like a, is independent of f. Thus there is a constant $c \in \mathbb{N}$, independent of f, so that

$$T \vdash \forall x \geq C\,(\neg\phi(x) \to \neg Con_T((f(\underline{x})^c))). \tag{4'}$$

On the other hand, we have, for some $\epsilon < 1$ and for all sufficiently large n, that if

$$T \overset{f(n)^\epsilon}{\vdash} Con_T((f(\underline{n}))^c),$$

then T would prove $\phi(\underline{n})$ in $\leq f(n)$ symbols. As before, this is a contradiction, since then also $T \vdash \neg\phi(\underline{n})$.
Q.E.D. Theorem 5

We are now ready to prove the analogue of Gödel's theorem involving symbol-length of proofs in place of step-length of proofs.

Theorem 2 *Let $i \geq 0$ and f be any recursive function. Then there is an infinite family $\mathcal{F}$ of Π^0_1-formulas such that*

(1) *for all $\phi \in \mathcal{F}$, $Z_i \vdash \phi$, and*

(2) *for $\phi \in \mathcal{F}$, if k is the least integer such that $Z_{i+1} \overset{k \text{ symbols}}{\vdash} \phi$ then it is not the case that $Z_i \overset{f(k) \text{ symbols}}{\vdash} \phi$.*

Proof Recall that Z_i and Z_{i+1} may be taken to be either schematic theories or weakly schematic theories (see section 3 for the full definitions). It is easy to see that the schematic versions of these theories have at most exponential speedup over the non-schematic versions; hence it will suffice to prove Theorem 2 for the schematic versions of Z_i and Z_{i+1}. This is fortunate since the previous two theorems depended on T having a polynomial time recognizable set of axioms.

Let c and ϵ be the constants from Theorem 5 that apply to the (schematic) theory $T = Z_i$. Without loss of generality, $1/\epsilon$ is an integer and $f(x) \geq x$ for all x and f is a time-constructible function. Let h be the time-constructible function $h(x) = (f(x))^{1/\epsilon}$. Let $\mathcal{F}$ be the set of formulas

$$Con_{Z_i}((h(\underline{n}))^c)$$

for $n \geq 1$. We must prove three things to establish the theorem:

(1) For all n, $Z_i \vdash Con_{Z_i}((h(\underline{n}))^c)$.
For a fixed $n \geq 1$, Z_i can prove $h(\underline{n}) = \underline{h(n)}$ by explicitly verifying the execution of the Turing machine which time constructs h. Then Z_i can examine all of the finitely many Z_i-proofs of $\leq h(n)$ symbols to verify that none of them is a valid Z_i-proof of $0 = 1$.

(2) For all sufficiently large n, it is not the case that $Z_i \overset{f(n)}{\vdash} Con_{Z_i}((h(\underline{n}))^c)$.
This is immediate from Theorem 5. We may discard from $\mathcal{F}$ the statements $Con_{Z_i}((h(\underline{n}))^c)$ where n is not sufficiently large.

(3) For all n, $Z_{i+1} \overset{O(\log n)}{\vdash} Con_{Z_i}((h(\underline{n}))^c)$.
Using a truth-definition for $(i+1)$-st order formulas, Z_{i+1} can prove the consistency of Z_i, i.e., Z_{i+1} can prove $(\forall x)Con_{Z_i}(x)$. From this, Z_{i+1} easily infers $(\forall x)Con_{Z_i}((h(x))^c)$. Now $|\underline{n}| = O(\log n)$, so by using a $\forall$-instantiation inference, $Z_{i+1} \overset{O(\log n)}{\vdash} Con_{Z_i}((h(\underline{n}))^c)$.

Theorem 2 follows immediately from (1), (2) and (3). Q.E.D.

2.2 Step-Length Speedup

We next prove Gödel's theorem as originally stated with proof lengths measured in terms of number of steps instead of number of symbols. Using symbol-length for proofs was convenient since there are only a finite number of proofs with a given symbol length and it is easy to enumerate all of them. However, the situation is more difficult when using step-length since there may well be an infinite number of proofs of a given step-length. This complicating fact does provide one advantage however: we shall improve Gödel's stated theorem by obtaining an *unbounded* speedup for Z_{i+1} over Z_i.

Since we can not bound the number of proofs of a given step-length, it will be important for us to instead bound the logical complexity of formulas occurring in a proof in terms of the step-length of the proof. For this, we extend a construction of Parikh to apply to theories which are merely weakly schematic.

Definition If ϕ is a formula, the *logical depth*, $dp(\phi)$, of ϕ is the maximum depth of nesting of the logical connectives in ϕ. For ϕ atomic, $dp(\phi)$ is 0; for ϕ of the form $(Qx)\psi$ or $\neg\psi$, $dp(\phi) = dp(\phi) + 1$; and for $\phi = (\psi \odot \chi)$ with $\odot$ a binary connective, $dp(\psi)$ is $1 + \max\{dp(\psi), dp(\chi)\}$.

The *quantifier depth* or *q-depth* of ϕ is denoted *q-dp*(ϕ) and is defined to be the maximum depth of nesting of quantifiers in ϕ.

The *quantifier block depth* or *qb-depth* of ψ is denoted *qb-dp*(ψ) and is defined to be the maximum depth of nesting of blocks of **like** quantifiers in ψ. Boolean connectives do not contribute to qb-depth, and if ϕ is $(Q\vec{x})\psi$ where $(Q\vec{x})$ denotes a block of quantifiers which either all existential or all universal and where ψ does not start with a quantifier of the same type, then *qb-dp*$(\phi) = 1 +$ *qb-dp*(ψ).

Theorem 6 *Let T be a weakly schematic theory and suppose $T \vdash_{\overline{k\ steps}} \phi$. Then there is a T-proof P of ϕ with step-length $\leq k$ such that every formula in P has quantifier-block depth $\leq$ qb-dp$(\phi) + O(k)$.*

The constant implicit in the $O(k)$ depends on the theory T, of course. The proof of Theorem 6 is postponed to section 2.3; actually, we prove a strengthening of this theorem there. By examining the proof of Theorem 6 it will be obvious that $I\Delta_0 + exp$, and hence Z_i, for $i \geq 0$, can prove the arithmetized version of this theorem — this will be important for our proof of the following strengthened form of Gödel's theorem:

Theorem 3 *Let $i \geq 0$. Then there is an infinite family $\mathcal{F}$ of Π^0_1-formulas such that*

(1) *for all $\phi \in \mathcal{F}$, $Z_i \vdash \phi$, and*

(2) *there is a fixed $k \in \mathbb{N}$ such that for all $\phi \in \mathcal{F}$, $Z_{i+1} \vdash_{\overline{k\ steps}} \phi$, and*

(3) *there is no fixed $k \in Z_i$ such that for all $\phi \in \mathcal{F}$, $Z_i \vdash_{\overline{k\ steps}} \phi$.*

Proof Let $i \geq 0$. Let $\phi(x)$ be the formula obtained by Gödel diagonalization such that Z_0 proves

$$\phi(x) \leftrightarrow \left(\neg Z_i \vdash_{\overline{x \text{ steps}}} \phi(\underline{x})\right)$$

Let $\mathcal{F} = \{\phi(\underline{n}) : n \geq 0\}$. We must show three facts:

(1) For all $n \geq 0$, it is not the case that $Z_i \vdash_{\overline{n}} \phi(\underline{n})$.
 This is immediate from the consistency of Z_i, since, if $Z_i \vdash_{\overline{n}} \phi(\underline{n})$, then Z_i proves this fact and thus $Z_i \vdash \neg\phi(\underline{n})$.

(2) For $n \geq 0$, $Z_i \vdash \phi(\underline{n})$.

To prove this, fix $n \geq 1$ and reason informally in Z_i. First, suppose $\neg\phi(n)$; then $Z_i \vdash_{\overline{n}} \phi(\underline{n})$. By Theorem 6 (which is formalizable in Z_i), it follows that there is a Z_i-proof of $\phi(\underline{n})$ in which every formula has qb-depth $\leq d \cdot n$ for d some constant. Now Z_i has a truth-definition predicate $Tr^i_{d \cdot n}$ for formulas of qb-depth $\leq d \cdot n$ (of necessity, the qb-depth of $Tr^i_{d \cdot n}$ is $> d \cdot n$). This truth predicate $Tr^i_{d \cdot n}(\ulcorner \psi \urcorner, x^i)$ takes the Gödel number of a formula ψ with qb-depth $\leq d \cdot n$ and an element x^i which is of order $i+1$ and construes x^i as encoding values for all the variables appearing free in ψ, and gives the truth value of ψ for these values. Thus validity of ψ can be expressed as

$$Valid^i_{d \cdot n}(\ulcorner \psi \urcorner) \quad \equiv \quad \forall x^i Tr^i_{d \cdot n}(\ulcorner \psi \urcorner, x^i).$$

Furthermore, as is usual for partial truth predicates, Z_i can prove that all axioms of Z_i of qb-depth $\leq d \cdot n$ are valid in the sense of $Tr^i_{d \cdot n}$ and that the rules of inference preserve validity. Hence, Z_i proves that if $Z_i \vdash_{\overline{n}} \phi(\underline{n})$, then $Valid^i_{d \cdot n}(\ulcorner \phi(\underline{n}) \urcorner)$. Furthermore, Z_i proves that $Valid^i_{d \cdot n}(\ulcorner \phi(\underline{n}) \urcorner)$ implies $\phi(n)$ holds (by induction on the complexity of ϕ).

Thus, we have shown that Z_i proves $\neg\phi(n)$ implies $\phi(n)$; i.e., $Z_i \vdash \phi(\underline{n})$, as desired.

(3) There is a $k \geq 0$ such that, for all $n \geq 0$,

$$Z_{i+1} \vdash_{\overline{k \text{ steps}}} \phi(\underline{n}).$$

To establish this, it will suffice to show that $Z_{i+1} \vdash (\forall x)\phi(x)$; since from $(\forall x)\phi(x)$ the formula $\phi(\underline{n})$ can be derived in a constant number of inferences independent of n.

The argument that $Z_{i+1} \vdash (\forall x)\phi(x)$ is like the argument for (2) above; however, Theorem 6 is no longer necessary since Z_{i+1} can define a truth predicate for all formulas of the language of Z_i.

Q.E.D. Theorem 3

The proof of Theorem 3 was similar in spirit to the proof of Theorem 2; however, one important difference is that the proof of Theorem 3 used a Gödel sentence ϕ expressing its own unprovability in a bounded number of steps, whereas the proof of Theorem 2 used a partial consistency statement asserting the unprovability of $0 = 1$ in a bounded number of of symbols. It would be interesting to know if Theorem 3 can also be proved with a partial consistency statement. This leads to the following question which is quite interesting in its own right. Let $Con_{Z_i}(k \text{ steps})$ be the formula $\neg Z_i \vdash_{\overline{k \text{ steps}}} 0 = 1$:

Question Is it the case that, for all constants $0 < \epsilon < 1$,

$$Z_i \vdash_{\overline{n^\epsilon \text{ steps}}} Con_{Z_i}(\underline{n} \text{ steps})$$

holds for all sufficiently large n? More generally, what is the least step-length of Z_i-proofs of $Con_{Z_i}(\underline{n} \text{ steps})$? Even, is there a constant k such that for all n

$$Z_i \vdash_{\overline{k \text{ steps}}} Con_{Z_i}(\underline{n} \text{ steps})?$$

This last question must have answer 'No' if Kreisel's conjecture holds for the particular formalization used for Z_i; however, it would be desirable to have self-referential disproof of this (assuming it's false, of course). By 'Kreisel's conjecture' we mean the statement that, for all formulas $A(x)$, if there is a fixed k such that for all n, $Z_i \vdash_{\overline{k \text{ steps}}} A(\underline{n})$, then $Z_i \vdash \forall x A(x)$.[5] The reason that Kreisel's conjecture implies the last question has answer 'No' is that otherwise Z_i would prove its own consistency.

On the other hand, we have the following example of a theory that does not satisfy Kreisel's conjecture and does not have $o(n)$ step-length proofs of its partial selfconsistency statements:[6] Let $\widehat{PA}$ be the first-order theory of Peano arithmetic with the usual axioms plus all true statements of the form $\underline{m} + \underline{n} = \underline{m+n}$ and $\underline{m} \cdot \underline{n} = \underline{m \cdot n}$. If the formulas $Con_{\widehat{PA}}(\underline{n} \text{ symbols})$ are formulated in the form $(\exists \vec{x})(p(n, \vec{x}) = q(n, \vec{x}))$ with p and q polynomials, then they have constant length $\widehat{PA}$-proofs for all integers n. Thus Kreisel's conjecture does not hold for $\widehat{PA}$, since otherwise $\widehat{PA}$ would prove its own consistency. On the other hand, choose a diagonal formula $\phi(x)$ so that

$$\widehat{PA} \vdash \Big(\phi(x) \leftrightarrow \neg(\widehat{PA} \vdash_{\overline{x \text{ steps}}} \phi(\underline{x})) \Big)$$

[5]Kreisel's conjecture is usually stated with the term $S^n(0)$ instead of $\underline{n}$. See Parikh [11] or Kreisel's footnote on page 400 of [16] for more information on Kreisel's conjecture.

[6]This example is a strengthening of a construction due to J. Krajíček (private communication).

and let $\neg\phi(x)$ be equivalent to a formula of the form $\exists\vec{y}(p(x,\vec{y}) = q(x,\vec{y}))$ with p and q polynomials. By the definition of ϕ, $\widehat{PA}$ proves

$$\neg\phi(x) \rightarrow (\widehat{PA} \vdash_{\overline{x}} \phi(\underline{x}))$$

and, by the form of $\neg\phi$, there is a constant ℓ such that $\widehat{PA}$ proves

$$\neg\phi(x) \rightarrow (\widehat{PA} \vdash_{\overline{\ell}} \neg\phi(\underline{x})).$$

So $\widehat{PA}$ proves $Con_{\widehat{PA}}(x + \ell' \text{ steps}) \rightarrow \phi(x)$ for some constant ℓ'. Now if $\widehat{PA}$ could prove its partial selfconsistency statements with $o(n)$ step-length proofs, then it would follow that $\widehat{PA}$ could prove each formula $\phi(\underline{m})$ in $o(m)$ many steps — which is a contradiction. The same reasoning shows that $\widehat{Z}_i$ also can not prove its partial selfconsistency statements $Con_{\widehat{Z}_i}(\underline{n} \text{ steps})$ in $o(n)$ steps.

The first question above is motivated by work of Pudlák [13, 14] where $O(n^{\epsilon})$ lower bounds are shown for the symbol-length of proofs in PA of $Con_{PA}(\underline{n} \text{ symbols})$. In analogy, we conjecture the answer to this question is 'No' but current techniques do not appear to be adequate to show this. Pudlák also shows nearly linear upper bounds on the symbol-length of these proofs too. As pointed out by J. Krajíček, these same methods also suffice to show that Z_i has proofs of step-length $O(n)$ of the selfconsistency statements $Con_{Z_i}(n \text{ steps})$. To prove this, note that Z_i can formalize Theorem 6 and prove

$$\neg Con_{Z_i}(x \text{ steps}) \rightarrow \neg Con_{I\Sigma^i_{cx+c}}(x \text{ steps})$$

for some constant c, where $I\Sigma^i_{cx+c}$ means the theory Z_i restricted to induction on Σ^i_{cx+c}-formulas. Now, for $m \geq 0$, Z_i can give a partial truth definition for Σ^i_m-formulas and prove that it satisfies the usual properties in $O(m)$ lines. Then Z_i can prove $Con_{I\Sigma^i_{cn+c}}(\underline{n} \text{ steps})$, and hence $Con_{Z_i}(\underline{n} \text{ steps})$, in $O(n)$ steps.

One final observation, which Krajíček has also pointed out to us, concerns formalizations of Z_i which have successor function S as the only function symbol (allowing $+$, $\cdot$, etc as relation symbols). In this case, if $Z_i \vdash_{\overline{k}} Con_{Z_i}(\underline{n} \text{ steps})$, then $k \geq d \log\log n$ for some constant d which depends on the formalization of Z_i. This is proved by using Theorem 3.1 of [7] which implies that if $Z_i \vdash_{\overline{k}} Con_{Z_i}(\underline{n} \text{ steps})$ then $Z_i \vdash^{\underline{(\log n)2^{2^{O(k)}}}} Con_{Z_i}(\underline{n} \text{ steps})$ and hence that $Z_i \vdash^{\underline{(\log n)2^{2^{O(k)}}}} Con_{Z_i}(\underline{n} \text{ symbols})$. And the Pudlák-Friedman lower bound then implies that $(\log n)2^{2^{O(k)}} > n^\epsilon$ for some $\epsilon > 0$; from whence $k = \Omega(\log\log n)$ follows.

Another open question is whether a "nonspeedup" result holds for Z_{i+1} over Z_i. Namely, is there an infinite set $\{\phi_j\}_j$ of consequences of Z_{i+1} with a moderately growing function g, say $g(j) = 2^j$, such that, if k_j is the minimum step-length of a Z_{i+1}-proof of ϕ_j, then the following hold: (1) $k_j \geq g(|\phi_j|)$ and (2) $Z_i \vdash_{\overline{k_j}} \phi_j$ also? The corresponding question using symbol-length is also open.

2.3 Bounding Formula Complexity via Step-Length

It is entirely possible for a formula ϕ to have a proof P of step-length k, but with formulas in P having arbitrarily large logical complexity. There are various reasons for this, most obviously, it may be that P has axioms, say tautologies of the form $A \to (B \to A)$, containing subformulas of arbitrarily high logical complexity. We shall see below, however, that such high logical complexity can be avoided by showing that whenever $T \vdash_{\overline{k \text{ steps}}} \phi$ where T is a weakly schematic theory then there is also a T-proof of ϕ in which every formula has its logical complexity (qb-depth) bounded by the complexity of ϕ plus $O(k)$.

First, we define formally what schematic and weakly schematic theories are. Parikh [11] was the first to introduce schematic theories.

Notation The language of i-th order logic contains variables x^j of order $j+1 \leq i$ (often denoted just x without the superscript denoting the order), constant symbols (we assume 0 is a constant symbol), function symbols (such as S, $+$, $\cdot$), various relation symbols (including $=$ and $\in$), and logical connectives consisting of Boolean connectives and quantifiers $\forall x^j$ and $\exists x^j$. Function symbols must take first-order arguments and produce first-order values. Relation symbols other than $=$ and $\in$ take only first-order arguments.

We say an object is of *type* j if and only if it is of order $j+1$; thus the superscript in x^j refers to the type of the variable.

To specify axioms and rules of inference schematically, we need metanotation for the syntax of i-th order logic. Accordingly, there are

(i) metavariables α^j of type j, which range over variables of type j,

(ii) term variables $\sigma, \tau, \ldots$ which range over terms (of type 0), and

(iii) formula variables $A(x_1), B(x_1, x_2), \ldots$ of various arities ranging over formulas.

Metaterms are formed from variables, meta-variables, constants, function symbols, and term variables. An *atomic metaformula* consists of either a formula variable applied to metaterms or a predicate symbol applied to metaterms. *Metaformulas* are formed from atomic metaformulas using Boolean connectives and quantifiers; we frequently use capital Greek letters $\Phi, \Psi, \ldots$ to denote metaformulas.

A *substitution* S consists of a mapping from metavariables to metavariables, from term variables to terms and from formula variables to formulas. We write κS for the value that S maps κ to. A substitution S can be extended to a mapping from metaterms and metaformulas to terms and formulas. If Φ is a metaformula and S is a substitution, then the formula ΦS is the result of replacing each metasymbol κ in Φ by κS. This is readily defined for κ a metavariable or a term variable; for κ a formula variable, $\kappa = A(u_1, \ldots, u_k)$, we define $(A(s_1, \ldots, s_k))S$ to be the formula obtained by substituting the terms $s_1 S, \ldots, s_k S$ for all free occurrences of the variables $u_1, \ldots, u_k$ in the formula κS (here, the s_i's are metaterms). This last definition applies even if $s_i S$ is not freely substitutable for u_i in κS; however, it is possible to formulate weakly schematic systems to avoid the case where the substitutions are not free (in fact, weakly schematic systems for higher-order arithmetic will always have "side conditions" that explicitly rule out the possibility of a substitution not being free). The variables $u_1, u_2, \ldots$ may be any distinguished list of type 0 variables. See Farmer [5] for a more comprehensive definition of metaformulas and substitutions.

Formally, a substitution S must be total, in that S assigns values to all metasymbols; but obviously, the formula ΦS depends only the values κS where κ is a metasymbol occurring in Φ. Thus it will be useful to utilize partial substitutions. An even more general notion is that of metasubstitutions: *metasubstitutions* are defined similarly to substitutions except that a metasubstitution maps metavariables to metavariables and variables, maps term variables to metaterms and maps formula variables to metaformulas. If S is a metasubstitution and Φ is a metaformula, then ΦS is defined in the obvious way.

The composition of two metasubstitutions, denoted $S \circ S'$ or just SS', is computed by applying first S and then S'.

Definition A theory T is *schematic* if and only if the following conditions hold:

- proofs in T consist of sequences of formulas where each formula is derived from earlier formulas by inference rules, and
- inference rules are given by a finite set of schemata

$$\Phi_1, \ldots, \Phi_k / \Psi \quad (R)$$

where $\Phi_1, \ldots, \Phi_k, \Psi$ are metaformulas and R is a finite set of *side conditions* (see below for allowable side conditions). A valid inference is of the form

$$\frac{\Phi_1 S, \ldots, \Phi_k S}{\Psi S}$$

where S is a substitution such that the conditions RS hold.

The side conditions in R may be of the following four forms:

α^j (or x^j) is not free in Φ

s is freely substitutable for α^0 (or x^0) in Φ

where α^j, x^j, s may be any metavariable of type j, variable of type j or metaterm (respectively), and where Φ is a metaformula.

Many, but by no means all, formal systems for i-th order logic are schematic. Examples of schematic inference rules are:

Modus Ponens:	$\Phi_1, \Phi_1 \to \Phi_2 / \Phi_2$ (no side conditions)
Universal Instantiation Rule:	$\forall \alpha^j A(\alpha^j) / A(\tau^j)$ provided τ^j is freely substitutable for α^j in $A(\alpha^j)$
Universal Instantiation Axiom:	$/ \forall \alpha^j A(\alpha^j) \to A(\tau^j)$ Same side condition
Comprehension Axiom:	$/ \exists \alpha^{j+1} \forall \beta^j (\beta^j \in \alpha^{j+1} \leftrightarrow A)$ provided α^{j+1} is not free in A

In the last two examples, there are no hypotheses to the inference (i.e., they're axioms). Recall that for our proof of Gödel's theorem (or Theorem 3), we used the assumption that a theory has some schematic form of the universal instantiation rule or axiom.

It turns out however, that many theories are not schematic in that they have axioms or rules that can not be captured by a finite list of schemata. Two common ways that a theory can fail to be schematic are that the theory may admit all tautologies as axioms and that the theory may allow all generalizations of axioms as axioms. (For an example of the latter, see Enderton [4].) By *generalization* of a formula ϕ, we mean any formula that can be obtained by prefixing ϕ with universal quantifiers.

Definition A theorem T is *weakly schematic* if T-proofs consist of sequences of formulas and if valid inference rules are specified by:

(1) a finite set of schemata as for schematic theories (inferences with zero hypotheses are axioms),

(2) all tautologies are axioms, and

(3) all generalizations of axioms are axioms.

However, it is possible to omit having all tautologies as axioms or to have only selected axioms have all generalizations also axioms; in these cases we still call the theory weakly schematic. A weakly schematic theory in which *none* of the axioms are selected to have all generalizations as axioms is called *generalization-free*.

We can now state the strengthened form of Theorem 6; part (a) is essentially due to Parikh [11] and is proved by Krajíček [7] in the form stated here:

Theorem 7 (a) *Let T be a schematic theory and suppose $T \vdash_{\overline{k\ steps}} \phi$. Then there is a T-proof P of ϕ with step-length $\le k$ such that every formula in P has depth $\le dp(\phi) + O(k)$.*

(b) *Let T be a generalization-free, weakly schematic theory and suppose $T \vdash_{\overline{k\ steps}} \phi$. Then there is a T-proof P of ϕ with step-length $\le k$ such that every formula in P has quantifier depth $\le q\text{-}dp(\phi) + O(k)$.*

(c) *Let T be a weakly schematic theory and suppose $T \vdash_{\overline{k\ steps}} \phi$. Then there is a T-proof P of ϕ with step-length $\leq k$ such that every formula in P has quantifier-block depth $\leq qb\text{-}dp(\phi) + O(k)$.*

Our proof of Theorem 7 is based on Parikh's notion of proof skeleton and unification. A *proof skeleton* is a partial description of a potential proof; a proof skeleton tells how each formula in the proof is to be derived by specifying which inference rule or axiom schema is used and which lines are used as hypotheses and which variables were substituted for any metavariables in the schema. If an axiom is justified as being a tautology then the proof skeleton gives the propositional tautology of which the axiom is an instance. If the axiom is a generalization of a tautology or an axiom schema, then the proof skeleton specifies not only the tautology or schema but also specifies the entire block of universal quantifiers. However, the proof skeleton does not specify the instantiations of term variables and formula variables in the inference schemata (this is in contrast to the fact that the proof skeleton does specify what variables are substituted for metavariables).

Obviously every proof has a proof skeleton; however, obviously not every proof skeleton corresponds to a valid proof. In addition, distinct proofs may have the same proof skeleton. By renaming term variables and formula variables, we can ensure that the proof skeleton uses distinct term variables and formula variables for each inference. Such a proof skeleton corresponds to a proof of the formula ϕ if and only if there is a substitution S which makes the skeleton into a proof of ϕ; specifically, (a) for the conclusion Φ of the final inference in the skeleton, $\Phi S = \phi$, and (b) S satisfies the side conditions of each inference, and (c) whenever a metaformula Ψ in the proof skeleton represents the conclusion of an inference and Ψ' in the proof skeleton represents the same formula used as a hypothesis of an inference, then $\Psi S = \Psi' S$. In other words, a proof skeleton represents a unification problem which has as solutions precisely the substitutions which transform the skeleton into a proof. See [11, 8, 7, 5] for more discussion on unification and proof skeletons.

Suppose T is weakly schematic. Given a T-proof P of ϕ of k steps, let P_S be its skeleton; Theorem 7(c) will be established by showing that there is a T-proof P^* with the same skeleton such that every formula has qb-depth $\leq qb\text{-}dp(\phi) + O(k)$.

First, we can assume w.l.o.g. that every variable appearing in the proof P is mentioned in the skeleton P_S, since, otherwise if y is a variable not occurring in P, we can erase every quantifier Qy appearing in P and replace the rest of the (now free) occurrences of y by the constant 0.[7]

The proof skeleton P_S corresponds to a unification problem consisting of a set of equations of the form $\Phi = \Psi$ plus a set of restrictions (side conditions). We must prove that there is a substitution S satisfying the equations and restrictions, so that S induces a proof of ϕ with all formulas in the proof have qb-depth bounded by $qb\text{-}dp(\phi) + O(k)$. We shall not be concerned with limiting the size of terms in the proof so we need only consider bounding the size of formulas that S assigns to metaformulas; for this, we examine only the part of the unification problem corresponding to logical connectives — this is a so-called first-order unification problem and can be readily solved using standard techniques:

Claim The unification problem for P_S has a *most general formula solution (mgfs)*; i.e., there is a metasubstitution S_{mgfs} such that S_{mgfs} maps formula variables to metaformulas but maps term variables to themselves and such that any other solution S to the unification problem can be expressed as $S_{mgfs} \circ S'$ for some (meta)substitution S'.

Furthermore, let D equal the total number of quantifier blocks occurring in all the formulas in P_S. Then for all metaformulas Φ appearing in P_S, the qb-depth of ΦS_{mgfs} is $\leq D$.

To prove the Claim, we assume w.l.o.g. that every equation in D is of the form $A = \Phi$ where A is a 0-ary formula variable (we can ensure this by replacing each equation $\Phi = \Psi$ in D by the two equations $A = \Phi$ and $A = \Psi$ where A is a new formula variable). We will produce S_{mgfs} by iteratively modifying the set D of equations, creating a new set E of equations. Initially, E is the empty set. For A and B formula variables occurring in D, we let $A \sim B$ mean that an equation $A = B(\vec{t})$ is in D. Let $\approx$ be the reflexive, transitive, symmetric closure of $\sim$. We write $A \succ_0 B$ to mean that there is an equation of the form $A = \odot(\vec{\Psi})$ in D with B occurring in $\vec{\Psi}$, where $\odot$ represents a Boolean connective or a quantifier. We write $A \succeq B$ to denote the transitive closure of the relation $(A \approx B \vee A \succ_0 B)$. It is important to stress that the $\succeq$ relation is redefined after each step of the iterative process modifying D. It is clear that there must be at least one variable which is maximal with respect to the $\succeq$-ordering; since D has a solution.

[7] If the language does not have a constant symbol, the same effect can be achieved by using a fixed variable not mentioned in the proof skeleton.

The iterative process for modifying D and E is as follows: Pick an arbitrary A which is $\succeq$-maximal in D such there is an equation of the form $A = \odot(\Psi_1, \ldots, \Psi_s)$ in D, where either $\odot$ represents a Boolean connective (so $s = 1$ or $s = 2$), or $\odot$ is a predicate symbol ($s \geq 1$), or $\odot$ is a maximal length block of like quantifiers. (In the last case, $s = 1$ and Ψ_1 does not begin with a quantifier of the same type. Also, by 'maximal length' we mean maximal over all choices for the equation $A = \odot(\vec{\Psi})$.) Note that $\Psi_1, \ldots, \Psi_s$ are metaformulas except in the case where $\odot$ is a predicate symbol; in the this case $\Psi_1, \ldots, \Psi_s$ are metaterms. Let $\vec{u}$ be an infinite sequence of new type 0 variables; when C is k-ary, we write $C(\vec{u})$ for $C(u_1, \ldots, u_k)$. Now, for each 0-ary formula variable $B \approx A$, do the following:

(a) Choose $B_1, \ldots, B_s$ to be new formula variables, unless $\odot$ is a predicate symbol, in which case, choose $B_1, \ldots, B_s$ to be new term variables.

(b) Find every equation in D of the form $B = \Psi$. Remove this equation from D. Replace every occurrence of the formula variable B in E with the equation $\odot(B_1, \ldots, B_s)$. Then add the formula $B = \odot(B_1, \ldots, B_s)$ to E.

(c) For each equation $B = \Psi$ found in (b) such that Ψ is equal to $C(\vec{t})$ with C a formula variable and $\vec{t}$ metaterms, let $C_1, \ldots, C_s$ be new formula variables of the same arity as C, and add the equations $B_i = C_i(\vec{t})$ to D (for $i = 1, \ldots, s$). Find every occurrence of $C(\vec{u})$ in E and replace it with $\odot(C_1(\vec{u}), \ldots, C_s(\vec{u}))$; then add the equation $C(\vec{u}) = \odot(C_1(\vec{u}), \ldots, C_s(\vec{u}))$ to E.

(d) Now consider each equation $B = \Psi$ found in (b) such that Ψ is $\otimes(\Theta_1, \ldots, \Theta_{s'})$ with $\otimes$ a Boolean connective or a predicate symbol: If $\otimes$ is not equal to $\odot$, then the unification problem has no solution, contradicting the fact that the proof P exists and provides a solution. Since $\otimes$ and $\odot$ are equal, $s' = s$. Add the s equations $B_i = \Theta_i$ to D.

(e) Now consider each equation $B = \Psi$ found in (b) such that Ψ starts with a quantifier block $Q\vec{x}$. Since $\odot$ was chosen to be a maximal length and since there exists a a solution to the unification problem, $\odot$ consists of $Q\vec{x}$ followed by a (possibly empty) list of similar quantifiers $Q\vec{y}$. In the case where $Q\vec{y}$ is the empty list and thus Ψ is of the form $\odot\Theta$, add the equation $B_1 = \Theta$ to D. Otherwise $Q\vec{y}$ is not empty and we have

Ψ of the form $Q\vec{x}C(\vec{t})$ where C is a formula variable; in this case, add $C' = C(\vec{t})$ and $C' = Q\vec{y}B_1$ to D, where C' is a new formula variable.

This iterative process stops when all the equations in D are of the form $A = C(\vec{s})$ where A and C are formula variables. The process must eventually halt because each iteration of the process reduces the total number of logical connectives in D. In fact, each iteration removes at least one Boolean connective or quantifier block from D; and adds at most one (the same one) connective or quantifier block to the depth of formulas in E. From this it follows that the qb-depth of formulas in E is bounded by the total number of quantifier blocks in the proof skeleton P_S. In addition, since one step of the process is performed on an A that is $\succeq$-maximal, it is easy to see that after that step, the new set D of equations does not contain any occurrence of A or any $B \approx A$.

Finally, S_{mgfs} can be defined as follows: Let E be the set of equations obtained after the final iteration of the above process. If $A(\vec{u}) = \Phi$ is an equation in E, then AS_{mgfs} is defined to be Φ. If κ does not appear on the lefthand side of an equation in E, then κS is just κ. It is easily checked that S_{mgfs} is well-defined; it is also easily verified that any substitution S which satisfies the original set of equations and restrictions from the proof skeleton must be expressible as $S = S_{mgfs} \circ S'$ for some metasubstitution S'. (The last fact is proved by induction on the number of iterations in the process creating E.)

That completes the proof of the Claim. Theorem 7 can now be proved relatively easily. Recall that P was assumed to be a T-proof of ϕ of k steps and that P_S is its proof skeleton. From P_S we formed the unification problem D with side conditions. D has, w.l.o.g., one equation of the form $A = \phi$ and has a total of $O(k)$ equations for the k steps. The number of quantifier blocks occurring in each of the $O(k)$ equations is bounded by a constant since there is only a finite number of inference and axiom schemata.[8] Let S_{mgfs} be the most general formula substitution from the Claim. Since D has $\mathit{qb\text{-}dp}(\phi) + O(k)$ quantifier blocks, S_{mgfs} maps formula variables to metaformulas of qb-depth $\leq \mathit{qb\text{-}dp}(A) + O(k)$. The proof P is obtained by applying some substitution S to P_S and, by the Claim, $S = S_{mgfs} \circ S'$ for

[8] However, there is not a constant bound on the number of logical connectives in each equation since all tautologies are axioms. We do not know if it is possible to bound the number of connectives in axioms in terms of the number k of proof steps.

some substitution S'. The problem is that S' may introduce new quantifier blocks, giving formulas of high qb-depth in P. However, we can replace S' by another substitution S'' which does not introduce *any* new quantifier blocks (or new Boolean connectives either). S'' is defined as follows: for each formula variable AS' is a formula, let $P(s_1, \ldots, s_\ell)$ be the first (i.e., leftmost) atomic subformula of AS' and suppose $P(\vec{s})$ is in the scope of quantifiers $Q_1 y_1, \ldots, Q_r y_r$ in AS'. Define AS'' to be the formula obtained from $P(\vec{s})$ by replacing every occurrence of $y_1, \ldots, y_r$ in $P(\vec{s})$ by the constant symbol 0. For all κ's which are not formula variables, let $\kappa S''$ equal $\kappa S'$. Let $P' = P_S S_{mgfs} S''$ be the proof obtained by applying the substitution $S_{mgfs} S''$ to the skeleton P_S. Clearly every formula in P' has qb-depth $\leq$ *qb-dp*$(\phi) + O(k)$. It remains to show that $S_{mgfs} S''$ satisfies all the side conditions since then P' is a valid proof. The side conditions become of the form

x does not occur free in Φ

and

τ is freely substitutable for x in ϕ.

It is easy to check that, since these conditions hold in $P = P_S S$, they also hold in P'; namely, if x does not occur free in $\Phi S = \Phi S_{mgfs} S'$, then it certainly does not occur free in $\Phi S_{mgfs} S''$, and, if $\tau S = \tau S' = \tau S''$ is freely substitutable for x in $\Phi S = \Phi S_{mgfs} S'$, then it is also freely substitutable for x in $\Phi S_{mgfs} S''$.
Q.E.D. Theorem 7(c).

Parts (a) and (b) of Theorem 7 are proved similarly. One merely replaces the qb-depth by 'depth' and 'q-depth', respectively, and all the arguments go through with minor modifications.

3 Proof Speedups without Self-Reference

In this section we give an alternative method of proof for Gödel's theorem and for Theorem 3 that avoids the use of self-reference; this proof is almost certainly not the kind of proof that Gödel envisioned and, unlike proofs using self-reference, is somewhat disappointing from a philosophical (foundational) point of view. Nonetheless, it provides important information about the limitations of proof systems.

The basic idea for this section is due to Statman, to whom the next theorem is due. Our proof will be based on Theorem 7.

Theorem 8 *(Statman [15]) Let T be a schematic theory and let ϕ be a formula independent of T. Further suppose $T \cup \{\neg\phi\}$ has an infinite model. Then there is a number m such that, for each $k > 0$, there is a tautology ψ_k such that*

(1) $T \cup \{\phi\} \vdash_{\overline{m\ steps}} \psi_k$,

(2) $T \vdash \psi_k$, *and*

(3) *It is not the case that* $T \vdash_{\overline{k\ steps}} \psi_k$.

We shall prove a somewhat stronger version of Statman's theorem which holds for weakly schematic theories; although, of course, in this case, ψ_k can no longer be a tautology since a weakly schematic theory may have all tautologies as axioms. We also omit the unnecessary assumption that $T \cup \{\neg\phi\}$ has an infinite model.

Theorem 9

(a) *Let T be a weakly schematic theory and ϕ be a formula not provable by T. Then there is a number m such that, for all $k > 0$, there is a valid formula ψ_k so that*

  (1) $T \cup \{\phi\} \vdash_{\overline{m\ steps}} \psi_k$,

  (2) $T \vdash \psi_k$, *and*

  (3) *It is not the case that* $T \vdash_{\overline{k\ steps}} \psi_k$.

(b) *The same holds for schematic T with the additional condition that ψ_k is a tautology.*

Note that (b) is basically Theorem 8 with out the infinite model assumption.

The idea of our proof of Theorem 9 is very easy to explain: to prove (b), take ψ_k to be a formula $\theta_k \vee \phi$ where θ_k is the formula

$$\bot \vee (\bot \vee (\bot \vee \cdots \vee (\bot \vee \top) \cdots))$$

where there are k many $\vee$'s and where $\top$ is some valid sentence such as $\forall x(x = x)$ and $\bot$ is $\neg\top$. Then ψ_k is clearly a tautology and provable in T; furthermore, $T \cup \{\phi\}$ proves ψ_k in a constant number of lines. On the other hand, we will show that if ψ_k has a proof of fewer than ϵk lines in some schematic theory S (where the constant ϵ depends on S), then $S \vdash \phi$. The intuitive idea behind this last assertion is that since a schematic system can work with only a finite number of nested $\vee$'s in a single proof step, any proof of ψ_k with fewer that ϵk lines must proceed by proving ϕ. More formally, we shall show that any proof of ψ_k of less than ϵk lines can be transformed into a proof of $\theta_k^\bot \vee \phi$ where $\theta_k^\bot$ is the formula θ_k with the $\top$ replaced by $\bot$. But $\theta_k^\bot \vee \phi \to \phi$ is clearly provable, which gives a contradiction.

The idea of the proof of (a) is similar but now take θ_k to be the formula

$$\exists x(\bot \vee \exists x(\bot \vee \exists x(\bot \vee \cdots \vee \exists x(\bot \vee \exists x(\top)) \cdots))$$

where, again, there are k many $\vee$'s.

To make this proof idea formal, we first need a lemma. We say that ϕ and ϕ' are *logically similar* if they both have the same logical structure (i.e., they differ only in choice of atomic subformulas).

Lemma 10 (a) *Let $\mathbb{R}$ be a schematic inference rule and suppose every metaformula appearing in $\mathbb{R}$ has logical depth $\le D$. Let S be a substitution such that $\mathbb{R}S$ is a valid instance of the rule. Let ϕ be a formula and suppose that all occurrences in $\mathbb{R}S$ of subformulas logically similar to ϕ are at depth $> D$. Then, if every occurrence of every subformula in $\mathbb{R}S$ logically similar to ϕ is replaced by an arbitrary fixed sentence ψ (such as $\top$ or $\bot$), then another valid instance of the rule is obtained.*

(b) *Suppose $\mathbb{R}$ is a weakly schematically specified axiom given by a metaformula, possibly with side conditions, which may be arbitrarily universally quantified. Suppose that every metaformula of $\mathbb{R}$ and its side conditions has logical depth $\le D$. Let substitution S and a block of universal quantifiers specify a valid instance of $\mathbb{R}$. Let ϕ be a formula that does not begin with a universal quantifier and suppose every occurrence in $\mathbb{R}S$ of a subformula logically similar to ϕ is at depth $> D$. Then by replacing every occurrence in $\mathbb{R}S$ of every subformula logically similar to ϕ by an arbitrary fixed sentence ψ, one obtains another valid instance of the rule.*

Proof (a) Obtain S' from S by modifying S by letting $\kappa S' = \kappa S$ for κ not a formula variable and letting $\kappa S'$ be κS except with every subformula ϕ^* logically similar to ϕ replaced by ψ. Then $\mathbb{R}S'$ is a valid instance of the schema R as the side conditions are easily seen to be still satisfied since ψ is a sentence and has no free occurrences of variables.

Part (b) is proved similarly. $\Box$

We can now prove Theorem 9(a). Since T is a weakly schematic theory, its rules are either schematic or are generalizations of schematic axioms or are generalizations of tautologies. Let D be the maximum logical depth of the metaformulas occurring in the rules of T and let c be the maximum number of metaformulas occurring in any weakly schematic rule of S. Since all logical connectives have arity ≤ 2, a formula can have at most $2^{D+1}-1$ subformulas occurring at depth $\leq D$. Thus in a T-proof of k steps, there are fewer than $k \cdot c \cdot 2^{D+1}$ formulas which occur as a subformula of depth $\leq D$ in a formula in P.

Let θ_0 be $\exists x(\top)$ and θ_{i+1} be $\exists x(\bot \vee \theta_i)$. Let ψ_k be $\theta_{kc2^{D+1}} \vee \phi$. Now suppose P is a T-proof of ψ_k of $\leq k$ steps. By the pigeonhole principle, there is some $\ell < kc2^{D+1}$ such that there is no occurrence of ψ_ℓ or any formula logically similar to ψ_ℓ as a subformula of depth $\leq D$ in any formula in P. Thus, by Lemma 10, if all occurrences of subformulas in P that are logically similar to ψ_ℓ are replaced by $\bot$, then a valid T-proof P' is obtained. Now P' is a proof of $\theta^{\bot}_{kc2^{D+1}-\ell} \vee \phi$; since the first disjunct is disprovable, it follows that $T \vdash \phi$.

We have shown that if $T \vdash_{\overline{k \text{ steps}}} \psi_k$, then $T \vdash \phi$. So, if ϕ is not provable by T, then it is impossible for T to prove ψ_k in k steps. However, $T \vdash \psi_k$ since ψ_k is valid. Also, it is obvious that $T \cup \{\phi\} \vdash_{\overline{m \text{ steps}}} \psi_k$ for some constant m independent of k since ψ_k contains ϕ as a disjunct. That proves Theorem 9(a).

Part (b) is proved similarly but $\theta_{kc2^{D+1}}$ may now be the tautology

$$\bot \vee (\bot \vee (\bot \vee \cdots \vee (\bot \vee \top) \cdots))$$

with $kc2^{D+1}$ many $\vee$'s.

References

[1] S. R. Buss, *Bounded Arithmetic*, Bibliopolis, 1986. Revision of 1985

Princeton University Ph.D. thesis.

[2] ——, *On Gödel's theorems on lengths of proofs II: Lower bounds for recognizing k symbol provability*, in Feasible Mathematics II, P. Clote and J. Remmel, eds., Birkhäuser-Boston, 1995, pp. 57–90.

[3] A. Ehrenfeucht and J. Mycielski, *Abbreviating proofs by adding new axioms*, Bulletin of the American Mathematical Society, 77 (1971), pp. 366–367.

[4] H. Enderton, *A Mathematical Introduction to Logic*, Academic Press, 1972.

[5] W. M. Farmer, *A unification-theoretic method for investigating the k-provability problem*, Annals of Pure and Applied Logic, 51 (1991), pp. 173–214.

[6] K. Gödel, *Über die Länge von Beweisen*, Ergebnisse eines Mathematischen Kolloquiums, (1936), pp. 23–24. English translation in *Kurt Gödel: Collected Works, Volume 1, pages 396-399, Oxford University Press, 1986*.

[7] J. Krajíček, *On the number of steps in proofs*, Annals of Pure and Applied Logic, 41 (1989), pp. 153–178.

[8] J. Krajíček and P. Pudlák, *The number of proof lines and the size of proofs in first-order logic*, Archive for Mathematical Logic, 27 (1988), pp. 69–84.

[9] G. Kreisel and H. Wang, *Some applications of formalized consistency proofs*, Fundamenta Mathematicae, 42 (1955), pp. 101–110.

[10] A. Mostowski, *Sentences Undecidable in Formalized Arithmetic: An Exposition of the Theory of Kurt Gödel*, North-Holland, 1952.

[11] R. J. Parikh, *Some results on the lengths of proofs*, Transactions of the American Mathematical Society, 177 (1973), pp. 29–36.

[12] ——, *Introductory note to 1936(a)*, in Kurt Gödel, Collected Works, Volume 1, Oxford University Press, 1986, pp. 394–397.

[13] P. Pudlák, *On the lengths of proofs of finitistic consistency statements in first order theories*, in Logic Colloquium '84, North-Holland, 1986, pp. 165–196.

[14] ——, *Improved bounds to the lengths of proofs of finitistic consistency statements*, in Logic and Combinatorics, vol. 65 of Contemporary Mathematics, American Mathematical Society, 1987, pp. 309–331.

[15] R. Statman, *Speed-up by theories with infinite models*, Proceedings of the American Mathematical Society, 81 (1981), pp. 465–469.

[16] G. Takeuti, *Proof Theory*, North-Holland, Amsterdam, 2nd ed., 1987.

[17] A. J. Wilkie and J. B. Paris, *On the scheme of induction for bounded arithmetic formulas*, Annals of Pure and Applied Logic, 35 (1987), pp. 261–302.