Zeitschr. f. math. Logik und Grundlagen d. Math.
Bd. 36, S. 29–46 (1990)

QUANTIFIED PROPOSITIONAL CALCULI AND FRAGMENTS OF BOUNDED ARITHMETIC

by JAN KRAJÍČEK and PAVEL PUDLÁK in Praha (Czechoslovakia)

§ 0. Introduction

The motivation for this paper comes from a well-known and probably very difficult problem whether Bounded Arithmetic is finitely axiomatizable. Attempts to solve this problem using the machinery of mathematical logic have failed so far. It is possible that the problem can be solved by combining logic with combinatorics. This would require a transformation onto a more combinatorial problem. The finite axiomatizability of Bounded Arithmetic seems to be tightly connected with the problem whether Polynomial Hierarchy collapses to some level Σ_i^p, but no implication relating these two problems has been proved.[1]) Here we present a different problem of a combinatorial character and prove a relation between this problem and the problem of the finite axiomatizability of Bounded Arithmetic.

COOK [4] introduced an equational theory PV of polynomial time computable functions and showed an interesting relation between PV and propositional proof system ER (Extended Resolution). He showed that (1) PV proves soundness of ER and (2) the translation of the equalities provable in PV into propositional calculus have polynomially long proofs in ER. BUSS [1] showed that S_2^1 (a fragment of the bounded arithmetic S_2) is conservative over PV; thus this relation is transferred to S_2^1.

The finite axiomatizability of S_2 is equivalent to the question whether the hierarchy S_2^i, $i = 1, 2, \ldots$, is increasing. We shall construct propositional proof systems G_i which have similar relation to S_2^{i+1} for $i \geqq 1$ as ER has to S_2^1. Then we show that the problem about the hierarchy S_2^i, $i = 1, 2, \ldots$, can be reduced to a problem about the length of proofs in proof systems G_i, $i = 1, 2, \ldots$. The systems G_i are natural extensions of a Gentzen system for the propositional logic to quantified propositional formulas with at most i quantifier alternations.

The problem about G_i's would require proving superpolynomial lower bounds to the length of proofs in these systems. This seems too difficult at present, as exponential lower bounds have been proved only for quite a weak system Resolution System (not extended) so far, cf. HAKEN [8]. However we shall show that nontrivial statements about S_2 and its fragments can be derived from this relation, in particular:

(1) For $i > j \geqq 2$ the $\forall\Sigma_j^b$-consequences if S_2^i are finitely axiomatizable (Corollary 7.1),

(2) for $i \geqq 1$, if $S_2^{i+1} \vdash$ "NP = co-NP", then G_i proves all tautologies by proofs of polynomial length (Corollary 7.3).

(WILKIE [11] proved statement (2) for S_2^1 and a Frege system with the substitution rule instead of G_0.)

---

[1]) Added in proof: Recently KRAJÍČEK, PUDLÁK and TAKEUTI proved that $T_2^i = S_2^{i+1}$ implies $\Sigma_{i+2}^p = \Pi_{i+2}^p$ for $i \geqq 1$.

After writing the first draft of this paper (January 1988) we learned about the work of MARTIN DOWD [6], [7]. In [7] he gave a full proof of COOK's theorem mentioned above and showed the same relation between the quantified propositional calculus (in our notation G) and Polynomial Space Arithmetic (PSA, an equational theory extending PV). In [6] he stated without proof a theorem which relates the fragments of S_2 to fragments of the quantified propositional calculus. He did not derive any corollaries of this theorem such as (1) and (2) above.

Throughout the paper we assume knowledge of BUSS [1], nevertheless we recall briefly some basic definitions.

§ 1. Preliminaries

The class of *quantified propositional formulas* (shortly *propositions*) is the least class of formulas containing the atoms $p_0, p_1, \ldots$, constants 0 (falsity) and 1 (truth), closed under the connectives $\wedge$, $\vee$, $\neg$ and $\supset$ and with any proposition $A(p)$ containing also propositions $\exists x A(x)$ and $\forall x A(x)$, where x substitutes for some occurrence of p in $A(p)$. The semantical meaning of $\exists x A(x)$ is $A(0) \vee A(1)$ and of $\forall x A(x)$ is $A(0) \wedge A(1)$.

We shall use the usual distinction between bounded and free atoms as is the distinction between bounded and free variables in first order logic (cf. TAKEUTI [10]).

As usual we assume that the indices i in p_i and x_i are written in the binary notation. Hence the lengths $|p_i|$ and $|x_i|$ of p_i and x_i are proportional to $\log_2(i)$.

We do not include $\equiv$ among the basic connectives but we shall occasionally use $A \equiv B$ as the abbreviation for $(A \supset B) \wedge (B \supset A)$.

Σ_i^q, Π_i^q ($i \geqq 0$) is a hierarchy of propositions defined similarly as is the arithmetical hierarchy:

$\Sigma_0^q = \Pi_0^q$ is the class of quantifier free propositions. Both Σ_i^q and Π_i^q are closed under $\wedge$, $\vee$ and the negation of a Σ_i^q-proposition is Π_i^q and vice versa. Σ_{i+1}^q contains both Σ_i^q and Π_i^q and propositions of the form $\exists x A(x)$, for $A \in \Pi_i^q$. Similarly Π_{i+1}^q contains both Σ_i^q and Π_i^q and propositions of the form $\forall x A(x)$, for $A \in \Sigma_i^q$. For A in Σ_i^q respectively B in Π_i^q the propositions $\exists x A$ and $\forall x B$ are in Σ_i^q and Π_i^q respectively, too.

Thus a proposition in a prenex form with i blocks of the like quantifiers and with the prefix beginning with the block of $\exists$'s is in Σ_i^q.

We shall consider systems of bounded arithmetic introduced by BUSS [1]. Theory S_2 is equivalent to (more precisely conservative over) $I\Delta_0 + \forall x \exists y (y = x^{\lceil \log_2(x+1) \rceil})$. The formulas in the hierarchy of formulas Σ_i^b, Π_i^b define sets which are in the corresponding levels of the polynomial hierarchy Σ_i^p, Π_i^p. The fragments S_2^i are obtained from S_2 by restricting the PIND-schema to Σ_i^b formulas. The schema PIND is

$$\varphi(0) \wedge \forall x (\varphi(\lfloor x/2 \rfloor) \supset \varphi(x)) \supset \forall x \varphi(x).$$

Thus the S_2^i is the finite set of open formulas BASIC plus Σ_i^b-PIND. The fragments T_2^i are defined similarly but with the ordinary schema of induction. The system S_2 is the union of S_2^i, $i = 1, 2, \ldots$, and is equivalent to the union of T_2^i, $i = 1, 2, \ldots$. For the details see [1].

It is well known that the syntax can be easily formalized in S_2. In particular, formulas $\mathrm{Sat}_i(A, \tau)$ and $\mathrm{Taut}_i(A)$ can be constructed in S_2, formalizing "$\Sigma_i^q \cup \Pi_i^q$-proposition A is satisfied by the truth valuation τ" and "$\Sigma_i^q \cup \Pi_i^q$-proposition A is satisfied by all truth valuations", respectively. As such constructions are quite standard (using recursion on notation) we shall only state the properties of such a formalization.

Lemma 1.1.

(i) Sat_0 *is* Δ_1^b *with respect to* S_2^1.

(ii) Sat_i *is* $\mathscr{B}(\Sigma_i^b)$ *for* $i \geqq 1$.

(iii) Taut_i *is* Π_{i+1}^b *and for* $i \geqq 1$ *also* $\forall\Sigma_i^b$.

($\mathscr{B}(X)$ denotes the class of Boolean combinations of formulas from X.)

Lemma 1.2. *For* $i \geqq 0$, S_2^1 *proves that for all propositions* A, B *of appropriate complexity and for all* k *it holds that*

(i) $\mathrm{Sat}_i(A \circ B, \tau) \equiv \mathrm{Sat}_i(A, \tau) \circ \mathrm{Sat}_i(B, \tau)$, *for* $\circ = \wedge, \vee, \supset$ *and*
$\mathrm{Sat}_i(\neg A, \tau) \equiv \neg \mathrm{Sat}_i(A, \tau)$;

(ii) $\mathrm{Sat}_i(\exists x A(x), \tau) \equiv \mathrm{Sat}_i(A(0) \vee A(1), \tau) \equiv (\exists \varepsilon \leqq 1)\, \mathrm{Sat}_i(A(p), \tau^\frown\langle p, \varepsilon\rangle)$

and analogically for $\forall$, *where* $\tau^\frown\langle p, \varepsilon\rangle$ *is the truth valuation* τ' *extending* τ *by putting* $\tau'(p) = \varepsilon$, *and* p *does not occur in* $\exists x A(x)$;

(iii) $\mathrm{Sat}_{i+1}(A, \tau) \equiv \mathrm{Sat}_i(A, \tau)$, *for* $A \in \Sigma_i^q \cup \Pi_i^q$;

(iv) $\mathrm{Sat}_i(\exists x_1 \ldots \exists x_k A(x_1, \ldots, x_k), \tau)$

$$\equiv \exists \tau'(\tau' = (\langle p_1, \varepsilon_1\rangle, \ldots, \langle p_k, \varepsilon_k\rangle) \wedge \bigwedge_j \varepsilon_j \leqq 1 \wedge \mathrm{Sat}_i(A(p_1, \ldots, p_k), \tau^\frown\tau'),$$

where $p_1, \ldots, p_k$ *do not occur in* $\exists x_1 \ldots \exists x_k A(x_1, \ldots, x_k)$, *and analogically for* $\forall$;

(v) $\mathrm{Sat}_i(\bigvee_{(\varepsilon_1, \ldots, \varepsilon_n) \in S} A(p_j/\varepsilon_j), \tau) \equiv (\exists(\varepsilon_1, \ldots, \varepsilon_k) \in S)\, \mathrm{Sat}_i(A(p_j/\varepsilon_j), \tau)$

and analogically for $\bigwedge$, *where* S *is a subset of* $\{0, 1\}^k$.

Definition. A polynomial time computable binary relation $P(x, y)$ is a *quantified propositional proof system* (shortly: *proof system*) iff $\exists d P(d, A)$ implies $A \in \bigcup_{i \geq 0} \mathrm{TAUT}_i$, where TAUT_i is the set of tautological Σ_i^q-propositions. We shall write $d: P \vdash A$ instead of $P(d, A)$ and we shall call d a *P-proof of A*.

The *length* of a formula or a proof will be denoted by $|A|$, $|d|$, respectively. We think of formulas and proofs as 0-1 sequences, thus we can use the same symbol as it is used for $\lceil \log_2(x + 1)\rceil$ in [1].

We shall often use statements about proof systems in fragments of arithmetic. In such cases we shall tacitly assume that we have a fixed arithmetical definition of P, which is Δ_1^b in S_2^1.

Definition. For P a proof system and $i \geqq 0$, i-RFN(P) is the formula

$$(\forall d, A, \tau)\, (d: P \vdash A \wedge A \in \Sigma_i^q \supset \mathrm{Sat}_i(A, \tau)).$$

Lemma 1.3. *For* $i \geqq 1$, i-RFN(P) *is an* $\forall\Sigma_i^b$-*sentence, and* 0-RFN(P) *is an* $\forall\Pi_1^b$-*sentence.*

Definition. For P, Q proof systems and $i \geqq 0$, *P i-polynomially simulates Q* iff there is a polynomial time computable function $f(x, y)$ such that for any Σ_i^q-proposition A, if $d: Q \vdash A$ then $f(d, A): P \vdash A$. $P \geqq^i Q$ will denote "P i-polynomially simulates Q" and $P \sim^i Q$ will denote the conjunction of $P \geqq^i Q$ and $Q \geqq^i P$.

This notion generalizes the notion of polynomial simulation introduced by Cook-Reckhow [5].

Finally let us recall some standard proof systems: Frege system F, extended Frege system EF, Frege system with substitution SF and extended resolution ER (cf. [5]).

§ 2. Quantified propositional calculi

Proof systems for quantified propositional calculus have been considered several times; for the history see Church [3]. We shall define a system G and its fragments G_i, for $i \geqq 0$. Our system is similar to that considered by Dowd [6, 7], which in turn is based on some earlier ones.

The *calculus* G is defined in a sequential manner analogically to the definition of LK in Takeuti [10]. The important difference is that a sequent may be a premisse of more than one inferences. Thus proof figures of G-proofs are not trees but *directed graphs*.

The calculus G works with sequents of propositions. The *rules* of the calculus G are

(a) the rule of initial sequent,

(b) structural rules,

(c) cut rule,

(d) propositional rules,

(e) quantifier rules.

Now we shall describe the rules explicitely.

(a) The initial sequents are the sequents of the form $p \to p$, $0 \to$, $\to 1$, for p a free atom.

The rules (b), (c), (d) are identical with those of Takeuti [10].

(e) Quantifier rules are

$$(\forall\colon \text{left})\ \frac{A(B), \Gamma \to \Delta}{\forall x A(x), \Gamma \to \Delta} \qquad (\forall\colon \text{right})\ \frac{\Gamma \to \Delta, A(p)}{\Gamma \to \Delta, \forall x A(x)},$$

$$(\exists\colon \text{left})\ \frac{A(p), \Gamma \to \Delta}{\exists x A(x), \Gamma \to \Delta}, \qquad (\exists\colon \text{right})\ \frac{\Gamma \to \Delta, A(B)}{\Gamma \to \Delta, \exists x A(x)},$$

with the proviso that p does not occur in the lower sequents of ($\forall$: right) and ($\exists$: left).

The G-proofs are sequences of sequents satisfying obvious conditions.

For $i \geqq 0$ *define* G_i by $d: G_i \vdash \Gamma \to \Delta$ iff $d: G \vdash \Gamma \to \Delta$ and all propositions occurring in d are in $\Sigma_i^q \cup \Pi_i^q$. In particular, $G_i \vdash A$ (i.e. $G_i \vdash \to A$) implies $A \in \Sigma_i^q \cup \Pi_i^q$.

This completes the definition of the calculi that we shall need.

Now we shall show that the substitution rule can be polynomially simulated in G and each fragment G_i, for $i \geqq 1$. We assume that only quantifier free propositions may be substituted. (This is needed for the proof of Lemma 2.1.) Clearly it is sufficient to consider only the following special case of the substitution rule

$$(*) \qquad \frac{\Gamma(p) \rightarrow \Delta(p)}{\Gamma(A) \rightarrow \Delta(A)},$$

where A is a quantifier free proposition which does not contain p and is substituted for all occurrences of p in $\Gamma(p) \rightarrow \Delta(p)$.

Lemma 2.1. *Let* SG *and* SG_i *be the systems* G *and* G_i *augmented with the substitution rule. Then for* $i \geqq 0$, $G \sim^i SG$, *and for* $i \geqq 1$, $G_i \sim^i SG_i$. *Moreover, these facts are provable in* S_2^1.

Proof. Clearly we need only to show the simulation of the substitution rule in G_i. Consider a substitution of the form $(*)$. Thus we have a proof of

$$Z_1\colon \Gamma(p) \rightarrow \Delta(p)$$

in G_i and we want to derive

$$Z_2\colon \Gamma(A) \rightarrow \Delta(A)$$

in G_i. Using the induction on the length of Γ and Δ one can show that

$$Z_3\colon p \equiv A, \Delta(p), \Gamma(A) \rightarrow \Delta(A),$$

$$Z_4\colon p \equiv A, \Gamma(A) \rightarrow \Delta(A), \Gamma(p),$$

$$Z_5\colon \rightarrow \exists x(x \equiv A)$$

are derivable in G_i by proofs whose size is polynomial in the length of Γ, Δ, A. Applying the cut-rule to Z_1, Z_4 we obtain

$$Z_6\colon p \equiv A, \Gamma(A) \rightarrow \Delta(A), \Delta(p),$$

and applying it again to Z_3 and Z_6 we obtain

$$Z_7\colon p \equiv A, \Gamma(A) \rightarrow \Delta(A).$$

Using ($\exists$: left) we get

$$Z_8\colon \exists x(x \equiv A), \Gamma(A) \rightarrow \Delta(A).$$

Thus Z_2 follows from Z_5 and Z_8 by cut. In this way the proof is increased only by an additive factor which is polynomial in the length of Γ, Δ, A. Hence it is a polynomial simulation. Since all the transformations are elementary, they can be performed in S_2^1. □

For G_0 and SG_0 it is an open problem whether G_0 polynomially simulates SG_0. We know only the following relations:

Lemma 2.2. S_2^1 *proves*

(i) $G_0 \sim^0 F$, (ii) $SG_0 \sim^0 SF \sim^0 ER \sim^0 EF$.

Proof. $G_0 \sim^0 F$, $SF \geqq^0 EF$ have been shown in [5]. $SG_0 \sim^0 SF$ and $ER \sim^0 EF$ are easy. $EF \geqq^0 SF$ has been shown in [6], [9]. □

Corollary 2.3. S_2^1 *proves* $G_1 \geqq^0 ER$. □

§ 3. Translation of bounded formulas into propositions

We define a translation of bounded formulas into propositions. The translation we use is a generalization of the translation used in COOK [3], DOWD [6], and KRAJÍČEK-PUDLÁK [9].

For $k \geqq 0$ define $k(i) = 0$ or 1, the *i-th digit of* k, by $k = \sum_{i \geqq 0} k(i) \cdot 2^i$. Observe that for $i > |k|$, $k(i) = 0$. Sometimes we shall use the following abbreviations: For a proposition B with free atoms $p_0, p_1, \ldots$ and $k \geqq 0$, we abbreviate $B(p_0/k(0), p_1/k(1), \ldots)$ by $B(\boldsymbol{p}/\boldsymbol{k})$ or simply $B(\boldsymbol{k})$.

Take a bounded formula $A(a_1, \ldots, a_k)$. As all functions in the language of S_2 are polynomial time computable, there exists a polynomial $p_A(x)$ such that for any $n_1, \ldots, n_k$ with $|n_1|, \ldots, |n_k| \leqq m$ one needs to compute only numbers with the length $\leqq p_A(m)$ in order to decide the truth value of $A(n_1, \ldots, n_k)$. This is proved by induction on the complexity of the terms occurring in A and the complexity of A.

Any polynomial $q(x)$ satisfying $\forall x(q(x) \geqq p_A(x))$ will be called a *bounding polynomial of* A.

For any bounding polynomial $q(x)$ of A we shall construct a sequence of propositions $[\![A]\!]^m_{q(m)}$, $m \geqq 0$, with the following property (we shall occasionally omit the indices $m, q(m)$, if there is no danger of confusion): If $a_1, \ldots, a_k$ are all free variables of A then the only free atoms of $[\![A]\!]^m_{q(m)}$ are $p^0_1, \ldots, p^{q(m)}_1, \ldots, p^0_k, \ldots, p^{q(m)}_k$ and for any $n_1, \ldots, n_k$ with $|n_1|, \ldots, |n_k| \leqq m$ it holds:

$$A(a_i/n_i) \text{ is true iff } [\![A]\!]^m_{q(m)}(\boldsymbol{p}_i/\boldsymbol{n}_i) \text{ is true.}$$

Moreover, we want the following properties of $[\![A]\!]$ which we state as a lemma.

Lemma 3.1. *For* $A \in \Sigma^b_i$, $i \geqq 0$, *we have:*

(1) $[\![A]\!] \in \Delta^q_1$ *with respect to* G_1 *for* $i = 0$, *and* $[\![A]\!] \in \Sigma^q_i$ *for* $i \geqq 1$;

(2) $|[\![A]\!]^m_{q(m)}| \leqq r(m)$, *for some polynomial* $r(x)$ *depending only on* A *and* $q(x)$;

(3) $[\![A \circ B]\!]$ *is* $[\![A]\!] \circ [\![B]\!]$ *for* $\circ = \wedge, \vee, \supset$, $[\![\neg A]\!]$ *is* $\neg[\![A]\!]$;

(4) $[\![(\exists x \leqq |t|)\, A(x)]\!]$ *is* $\bigvee_{\varepsilon \in S} [\![a \leqq |t| \wedge A(a)]\!]\, (p_i/\varepsilon_i)$,

where $S = \{(\varepsilon_0, \ldots, \varepsilon_{q(m)}) \in \{0, 1\}^{q(m)+1} \mid (\forall i > |q(m)|)\, \varepsilon_i = 0\}$ *and the* p_i*'s are the atoms associated to* a;

(5) $[\![(\exists x \leqq t)\, A(x)]\!]$ *is* $\exists x_0 \ldots \exists x_{q(m)} [\![a \leqq t \wedge A(a)]\!]\, (p_i/x_i)$,

where t *is a term not of the form* $|s|$;

(6) $[\![(\forall x \leqq |t|)\, A(x)]\!]$ *is* $\bigwedge_{\varepsilon \in S} [\![a \leqq |t| \supset A(a)]\!]\, (p_i/\varepsilon_i)$,

where S *is as in* (4);

(7) $[\![(\forall x \leqq t)\, A(x)]\!]$ *is* $\forall x_0 \ldots \forall x_{q(m)} [\![a \leqq t \supset A(t)]\!]\, (p_i/x_i)$,

where t *is not of the form* $|s|$;

(8) *for* $A(a) \in \Sigma^b_0$, t *a term,* a *a free variable,* $q(x)$ *a bounding polynomial of* $A(t)$,

$$S^1_2 \vdash \forall y(G_1 \vdash [\![t = a \wedge A(a)]\!]^{|y|}_{q(|y|)} \to [\![A(t)]\!]^{|y|}_{q(|y|)}).$$

It suffices to construct $[\![A]\!]^m_{q(m)}$ for A atomic, since conditions (3)–(7) determine the construction for other bounded formulas. The translation of atomic formulas

$$t(a_1, \ldots, a_k) = s(a_1, \ldots, a_k)$$

and

$$t(a_1, \ldots, a_k) \leqq s(a_1, \ldots, a_k)$$

is defined as follows: Associate with any variable a_l free atoms $p_0^{a_l}, \ldots, p_{q(m)}^{a_l}$. As any function f in the language of S_2 is polynomial time computable, there are Boolean circuits C_f^m of the size polynomial in $m, q(m)$ computing f on inputs of the length $\leqq m$. Combining these circuits one can construct circuits C_t^m computing any term t and having again size polynomial in $m, q(m)$. Circuit C_t^m has some dummy input nodes $p_j^{a_l}$, for $j > m$, and may have also some dummy output nodes q_j's if $q(m)$ is larger than the length of the output. We assume that these nodes are labelled e.g. by 0. Boolean circuit C_t^m can be easily turned to a Σ_1^q-proposition $B_t^m(\boldsymbol{p}^{a_1}, \ldots, \boldsymbol{p}^{a_k}, \boldsymbol{q})$. So for $n_1, \ldots, n_k$ with the length $\leqq m$ we have:

$$t(n_1, \ldots, n_k) = n \quad \text{iff} \quad B_t^m(\boldsymbol{n}_1, \ldots, \boldsymbol{n}_k, \boldsymbol{n}) \text{ is true.}$$

We shall occasionally say that the atoms q_j's are *associated with the term t*.

Define the translation of atomic formulas:

(a) $[\![t(a_1, \ldots, a_k) = s(a_1, \ldots, a_k)]\!]^m_{q(m)}$ is

$$\exists x_0 \ldots \exists x_{q(m)} (B_t^m(\boldsymbol{p}^{a_1}, \ldots, q_j/x_j) \wedge B_s^m(\boldsymbol{p}^{a_1}, \ldots, q_j/x_j)).$$

This can be also written in a Π_1^q-form

$$\forall \boldsymbol{x} \forall \boldsymbol{y} (B_t^m(\boldsymbol{p}^{a_1}, \ldots, \boldsymbol{x}) \wedge B_s^m(\boldsymbol{p}^{a_1}, \ldots, \boldsymbol{y}) \supset \bigwedge_{i=0}^{q(m)} x_i \equiv y_i).$$

(b) $[\![t(a_1, \ldots, a_k) \leqq s(a_1, \ldots, a_k)]\!]^m_{q(m)}$ is

$$\exists \boldsymbol{x} \exists \boldsymbol{y} (B_t^m(\boldsymbol{p}^{a_1}, \ldots, \boldsymbol{x}) \wedge B_s^m(\boldsymbol{p}^{a_1}, \ldots, \boldsymbol{y}) \wedge \bigwedge_{i=0}^{q(m)} (\bigwedge_{j=i+1}^{q(m)} x_j \equiv y_j \supset (x_i \supset y_i)).$$

Again this has a Π_1^q-form, too,

$$\forall \boldsymbol{x} \forall \boldsymbol{y} (B_t^m(\boldsymbol{p}^{a_1}, \ldots, \boldsymbol{x}) \wedge B_s^m(\boldsymbol{p}^{a_1}, \ldots, \boldsymbol{y}) \supset \bigwedge_{i=0}^{q(m)} (\bigwedge_{j=i+1}^{q(m)} x_j \equiv y_j \supset (x_i \supset y_i))).$$

Now, having $A \in \Sigma_i^b$ for $i \geqq 1$ choose such a form (Σ_1^q or Π_1^q) of the translations of the atomic subformulas of A so that $[\![A]\!] \in \Sigma_i^q$.

(8) is proved easily by induction on the length of t and A. □

Lemma 3.2. *For $A(a) \in \Sigma_i^b$, $i \geqq 1$, $A(a)$ with one free variable a, and $q(x)$ a bounding polynomial of A,*

$$S_2^1 \vdash \forall y (\mathrm{Taut}_i([\![A]\!]^{|y|}_{q(|y|)}) \equiv \forall x (|x| \leqq |y| \supset A(x))).$$

Proof. We shall prove a stronger statement by induction on the length of $A(a_1, \ldots, a_k) \in \Sigma_i^b$:

$$(*) \qquad S_2^1 \vdash \forall y \forall x_1 \ldots \forall x_k (|x_1| \leqq |y| \wedge \ldots \wedge |x_k| \leqq |y| \supset (\mathrm{Sat}_i([\![A]\!]^y_{q(|y|)}, \tau(x_1, \ldots, x_k)) \equiv A(x_1, \ldots, x_k))),$$

where $\tau(x_1, \ldots, x_k)$ is the substitution which substitutes $\boldsymbol{x}_j$ for the propositional variables corresponding to a_j, $j = 1, \ldots, k$. For A atomic one can use Σ_1^b-PIND to prove the formula in S_2^1, since $\mathrm{Sat}_i(A, \tau) \equiv \mathrm{Sat}_0(A, \tau)$ by Lemma 1.2 and Sat_0 is

3*

Δ_1^b by Lemma 1.1. If A is not atomic we can reduce the proof of $(*)$ to a simpler formula using (i), ..., (iv), (v) of Lemma 1.2 and (3)–(7) of Lemma 3.1. We shall demonstrate it on the case when A begins with $\exists$. So let A be $(\exists x \leqq t)\, B(x, z_1, \ldots, z_k)$, let τ denote $\tau(z_1, \ldots, z_k)$. Working in S_2^1 assume that $|z_1|, \ldots, |z_k| \leqq |y|$. Then by (5) of Lemma 3.1 and (i) and (iv) of Lemma 1.2 we have

$$\begin{aligned}
&\mathrm{Sat}_i([\![(\exists x \leqq t)\, B(x, z)]\!]^{|y|}, \tau)\\
&\equiv \mathrm{Sat}_i(\exists x_1 \ldots \exists x_{q(|y|)}[\![x \leqq t \wedge B(x, z)]\!]^{|y|}, \tau)\\
&\equiv \exists x(|x| = q(|y|) \wedge \mathrm{Sat}_i([\![x \leqq t]\!]^{|y|}, \tau(x, z_1, \ldots, z_k))\\
&\qquad\qquad \wedge \mathrm{Sat}_i([\![B(x, z)]\!]^{|y|}, \tau(x, z_1, \ldots, z_k))).
\end{aligned}$$

Since we have $(*)$ for atomic formulas, the first two conjuncts are equivalent to $x \leqq t(z_1, \ldots, z_k)$. By the induction assumption the last conjunct is equivalent to $B(x, z)$. Thus $(*)$ is proved. The other cases can be handled similarly. □

Lemma 3.3. *For $A \in \Sigma_i^b$, $i \geqq 1$, and $q(x)$ a bounding polynomial of A,*

$$S_2^1 \vdash i\text{-RFN}(P) \supset \forall y(P \vdash [\![A]\!]_{q(|y|)}^{|y|} \supset \forall x(|x| \leqq |y| \supset A(x))).$$

This lemma follows from Lemma 3.2. □

Lemma 3.4. (i) *For $A(a) \in \Sigma_1^b$ and $q(x)$ a bounding polynomial of A,*

$$S_2^1 \vdash A(a) \supset G_1 \vdash [\![A(\dot{a})]\!]_{q(|a|)}^{|a|}.$$

(ii) *For $i \geqq 1$ and $q(x)$ a bounding polynomial of* Taut_i,

$$S_2^1 \vdash A \in \Sigma_i^q \wedge |y| \geqq |A| \supset (G_i \vdash [\![\mathrm{Taut}_i(A)]\!]_{q(|y|)}^{|y|} \supset G_i \vdash A).$$

The same holds for $i = 0$ with G_1 instead of G_0.

Proof. Part (i) is simple: Choose the witnesses of the $\exists$-quantifiers of $A(\dot{a})$ and using their digits compute the truth value of $[\![A(\dot{a})]\!]$.

(ii) We shall prove the statement for $i = 0$. The case $i \geqq 1$ is essentially the same. $\mathrm{Taut}_0(A)$ is defined as

$$\forall \tau(|\tau| \leqq |A| \supset \mathrm{Sat}_0(A, \tau)),$$

where we have to take Sat_0 in Π_1^b-form. $\mathrm{Sat}_0(A, \tau)$ is defined by

$$\forall w\,(\text{“}w \text{ is a computation of the value of } A \text{ on } \tau\text{”} \supset \text{“the last bit of } w \text{ is } 1\text{”}).$$

Thus the translation of $\mathrm{Taut}_0(A)$ in the propositional calculus has the following form

$$\forall \boldsymbol{p} \forall \boldsymbol{q}\ \mathrm{Comp}_A(\boldsymbol{p}, \boldsymbol{q}) \supset q_r,$$

where $\boldsymbol{p}$ is a vector of atoms associated with τ, $\boldsymbol{q}$ is associated with w, q_r is the last element of $\boldsymbol{q}$, and Comp_A is the translation of “w is a computation of the value of A on τ”. We shall assume that $\boldsymbol{p}$ are just the atoms of A. In $\boldsymbol{q}$ certain atoms code the truth value of subformulas of A computed on $\boldsymbol{p}$. If the variables in $\boldsymbol{q}$ are suitably ordered, it is possible to prove (using PIND of S_2^1) that

$$G_1 \vdash \mathrm{Comp}_A(\boldsymbol{p}, \boldsymbol{q}) \supset (q_i \equiv A_i),$$

where q_i corresponds to a subformula A_i of A. In particular, we have

(1) $\quad G_1 \vdash \mathrm{Comp}_A(\boldsymbol{p}, \boldsymbol{q}) \supset (q_r \equiv A).$

Now, let $\mathrm{Comp}_A^i(\boldsymbol{p}, q_1, \ldots, q_i)$ be subformulas of $\mathrm{Comp}_A(\boldsymbol{p}, \boldsymbol{q})$ which express that there are "i steps of the computation". Again by PIND on i one can show

$$\mathrm{G}_1 \vdash \exists q_1, \ldots, \exists q_i \; \mathrm{Comp}_i^A(\boldsymbol{p}, q_1, \ldots, q_i),$$

thus in particular

(2) $$\mathrm{G}_1 \vdash \exists \boldsymbol{q} \; \mathrm{Comp}_A(\boldsymbol{p}, \boldsymbol{q}).$$

From (1) and (2) we obtain easily

$$\mathrm{G}_1 \vdash \forall \boldsymbol{p} \forall \boldsymbol{q} (\mathrm{Comp}_A(\boldsymbol{p}, \boldsymbol{q}) \supset q_r) \supset A.$$

As the above proof can be done in S_2^1, we have proved (ii). □

We have not quite specified the translation $[\![A]\!]$ for atomic formulas. If we want to be able to prove relation between weak fragments of arithmetic and weak propositional proof systems, we have to choose "natural" Boolean circuits computing the arithmetical functions in the atomic formulas. Again, we state our last condition as a lemma.

Lemma 3.5. *For A any axiom of* BASIC *and $q(x)$ any bounding polynomial of A,*

$$\mathrm{S}_2^1 \vdash \forall y (\mathrm{G}_1 \vdash [\![A]\!]^{|y|}_{q(|y|)}).$$

Proof. We shall use the construction of Cook [4]. He introduced an equational theory PV which has a function symbol for each polynomial time computable function. He defined translations of equations of PV into the propositional calculus such that the translations of equalities provable in PV have proofs of polynomial length in ER. Dowd [7] proved this simulation using EF instead of ER. The simulation can be extended to the theory PV 1 which is an extension of PV to open formulas. The proof actually gives an explicit polynomial time algorithm which, for given m, constructs an EF proof of $[\![A]\!]^m$ and, moreover, this can be formalized in S_2^1. Buss [1] has shown a close relation of PV and PV 1 to S_2^1; in particular, if we translate formulas containing $\leqq$ using Cook's function LESS, all open theorems of S_2^1 become provable in PV 1. Thus we define our translation into quantified propositional calculus by taking Cook's one for equations in the language of S_2 and by adding quantifiers to it as described above. Now the translation of atomic formulas will be different from the one described above, since we shall use equations $\mathrm{LESS}(t, s) = 0$ instead of $t \leqq s$. But one can show in S_2^1 that they are equivalent (and, moreover, it is irrelevant for this paper). Thus we obtain the condition of Lemma 3.5. □

§ 4. Relations between propositional proof systems and theories

This section develops a general connection between propositional proof systems and theories. We tacitly assume that the languages of theories discussed contain the language of S_2.

We shall write $\forall\Sigma_i^b(T)$ for the set of all $\forall\Sigma_i^b$-consequences of T.

Definition. For $i \geqq 0$, P a proof system and T a theory, P *simulates* $\forall\Sigma_i^b(T)$ iff for any $\forall \boldsymbol{x} A(\boldsymbol{x}) \in \forall\Sigma_i^b(T)$ there is a bounding polynomial $p(x)$ of A such that

$$\mathrm{S}_2^1 \vdash \forall y (P \vdash [\![A]\!]^{|y|}_{p(|y|)}).$$

Definition. For $i \geqq 0$ a proof system P is *i-regular* iff S_2^1 proves

(i) $P \geqq^1 G_1$,

(ii) $P \vdash A \supset B \wedge P \vdash A \supset P \vdash B$,

(iii) for $A \in \Sigma_i^q$, $|y| \geqq |A|$

$$P \vdash [\![\mathrm{Taut}_i(A)]\!]^{|y|}_{q(|y|)} \supset P \vdash A,$$

where $q(x)$ is a bounding polynomial of Taut_i. Observe that an i-regular proof system satisfies Lemmas 3.2, 3.4 and 3.5. This is the motivation for their definition.

Theorem 4.1. *Let $T \supseteqq S_2^1$ and P be an i-regular proof system.*

(i) *Suppose $i \geqq 2$, P simulates $\forall\Sigma_i^b(T)$ and $T \vdash i$-RFN(P). Then*

$$\forall\Sigma_i^b(T) \equiv (S_2^1 + i\text{-RFN}(P)),$$

thus $\forall\Sigma_i^b(T)$ is finitely axiomatizable.

(ii) *Suppose $i \geqq 0$, P simulates $\forall\Sigma_i^b(T)$ and $T \vdash i$-RFN(Q) for some propositional proof system Q. Then*

$$S_2^1 \vdash P \geqq^i Q.$$

(iii) *Suppose $i \geqq 0$, P simulates $\forall\Sigma_i^b(T)$ and $T \vdash \mathrm{NP} = \mathrm{coNP}$. Then there exists a polynomial $p(x)$ such that T proves*

$$(\forall A \in \mathrm{TAUT}_i)\, \exists d(d\colon P \vdash A \wedge |d| \leqq p(|A|)).$$

Statement (ii) generalizes a construction of Cook [3] using which he showed (ii) for $P = \mathrm{ER}$, $T = \mathrm{PV}$ and $j = 0$. Statement (iii) could be used to generalize a result of Wilkie [11] who proved (iii) for $T = S_2^1$ and $P = \mathrm{SF}$.

Proof. (i) S_2^1 is $\forall\Sigma_2^b$ and so $S_2^1 \subseteqq \forall\Sigma_i^b(T)$ for $i \geqq 2$. By Lemma 1.3, i-RFN$(P) \in \forall\Sigma_i^b(T)$. On the other hand, assume $\forall x A(x) \in \forall\Sigma_i^b(T)$. Then $S_2^1 \vdash \forall y(P \vdash [\![A]\!]^{|y|})$, for some bounding polynomial. By Lemma 3.3 then

$$S_2^1 + i\text{-RNF}(P) \vdash \forall y \forall x(|x| \leqq |y| \supset A(x)),$$

i.e. $S_2^1 + i\text{-RFN}(P) \vdash \forall x A(x)$.

(ii) Assume $T \vdash i$-RFN(Q), so

(1) $S_2^1 \vdash (P \vdash [\![d\colon Q \vdash A \wedge A \in \Sigma_i^q \supset \mathrm{Taut}_i(A)]\!]^{|d|+|A|})$.

By Lemma 3.4 (i), as $d\colon Q \vdash A$ and $A \in \Sigma_i^q$ are Σ_1^b-formulas and since P is i-regular we have

(2) $S_2^1 \vdash d\colon Q \vdash A \wedge A \in \Sigma_i^q \supset P \vdash [\![\mathrm{Taut}_i(A)]\!]^{|d|+|A|}$.

Since P is i-regular we can use Lemma 3.4 (ii) to deduce

(3) $S_2^1 \vdash d\colon Q \vdash A \wedge A \in \Sigma_i^q \supset P \vdash A$.

By the main theorem of Buss [1] there is a polynomial time function f such that

$$S_2^1 \vdash d\colon Q \vdash A \wedge A \in \Sigma_i^q \supset f(d, A)\colon P \vdash A.$$

(iii) Assume $T \vdash \mathrm{NP} = \mathrm{coNP}$. Then every bounded formula is equivalent to a Σ_1^b formula, thus

(1) $T \vdash \mathrm{Taut}_i(A) \equiv (\exists x \leqq t(A))\, B(x, a)$,

for some Δ_1^b-formula B. Define the proof system Q by

$$d:Q\vdash A \quad \text{iff} \quad d \leqq t(A) \wedge B(d, A).$$

From (1) then

(2) $T \vdash i\text{-RFN}(Q)$.

Using the statement (ii) then

(3) $S_2^1 \vdash P \geqq^i Q$.

Hence

$$T \vdash \mathrm{Taut}_i(A) \equiv \exists d(|d| \leqq p(|A|) \wedge d:P\vdash A),$$

where $p(x)$ is the polynomial given by the function f of (ii). □

Next corollary shows that, in principle, Theorem 4.1 can be used to show that two theories are different.

Corollary 4.2. *Assume that for* $i \geqq 0$, *theories* $T \supseteqq S_2^1$ *and* S, *and proof systems* P *and* Q *the following holds:*

(i) P *is* i*-regular,*

(ii) P *simulates* $\forall\Sigma_i^b(T)$,

(iii) $S \vdash i\text{-RFN}(Q)$,

(iv) *not* $P \geqq^i Q$.

Then $T \nvdash S$, *in particular* $T \nvdash i\text{-RFN}(Q)$.

Proof. Use Theorem 4.1 (ii). □

On the other hand, we have the following corollary:

Corollary 4.3. *Assume* $S_2^1 \subseteqq S \subseteqq T$, $i \geqq 1$, P *is* i*-regular,* P *simulates* $\forall\Sigma_i^b(T)$ *and* $T \vdash i\text{-RFN}(P)$. *Then the following statements are equivalent:*

(i) T *is* $\forall\Sigma_i^b$*-conservative over* S,

(ii) $S \vdash i\text{-RFN}(P)$.

Proof. For (i) ⇒ (ii) use Lemma 1.3. The other implication is proved as Theorem 4.1 (i). □

In the following sections we shall apply the general theorems of this section to the proof systems G_i and theories S_2^i and T_2^i.

§ 5. Provability of reflection principles

By the definition of proof systems in § 1, any formula $i\text{-RFN}(P)$ is true. In this section we are interested in the question which theory suffices to prove $i\text{-RFN}(P)$, for P a calculus of § 2.

Theorem 5.1. *For* $i \geqq 0$, $S_2^{i+1} \vdash i\text{-RFN}(G_i)$.

Proof. A sequent $\Gamma \to \Delta$ is satisfied by a truth valuation τ iff the formula $(\bigwedge\Gamma) \supset (\bigvee\Delta)$ is satisfied by τ. Analogically with Lemma 1.1, there are formulas $\mathrm{SSat}_i(Z, \tau)$ and $\mathrm{STaut}_i(Z, \tau)$ formalizing "sequent Z consisting only of $\Sigma_i^q \cup \Pi_i^q$-propositions is satisfied by truth valuation τ" and "sequent Z consisting only of $\Sigma_i^q \cup \Pi_i^q$-

Definition. For $i \geqq 0$ a proof system P is *i-regular* iff S_2^1 proves

(i) $P \geqq^1 G_1$,

(ii) $P \vdash A \supset B \wedge P \vdash A \supset P \vdash B$,

(iii) for $A \in \Sigma_i^q$, $|y| \geqq |A|$

$$P \vdash [\![\mathrm{Taut}_i(A)]\!]^{|y|}_{q(|y|)} \supset P \vdash A,$$

where $q(x)$ is a bounding polynomial of Taut_i. Observe that an i-regular proof system satisfies Lemmas 3.2, 3.4 and 3.5. This is the motivation for their definition.

Theorem 4.1. *Let $T \supseteqq S_2^1$ and P be an i-regular proof system.*

(i) *Suppose $i \geqq 2$, P simulates $\forall\Sigma_i^b(T)$ and $T \vdash i$-RFN(P). Then*

$$\forall\Sigma_i^b(T) \equiv (S_2^1 + i\text{-RFN}(P)),$$

thus $\forall\Sigma_i^b(T)$ is finitely axiomatizable.

(ii) *Suppose $i \geqq 0$, P simulates $\forall\Sigma_i^b(T)$ and $T \vdash i$-RFN(Q) for some propositional proof system Q. Then*

$$S_2^1 \vdash P \geqq^i Q.$$

(iii) *Suppose $i \geqq 0$, P simulates $\forall\Sigma_i^b(T)$ and $T \vdash \mathrm{NP} = \mathrm{coNP}$. Then there exists a polynomial $p(x)$ such that T proves*

$$(\forall A \in \mathrm{TAUT}_i)\, \exists d(d{:}P \vdash A \wedge |d| \leqq p(|A|)).$$

Statement (ii) generalizes a construction of Cook [3] using which he showed (ii) for $P = \mathrm{ER}$, $T = \mathrm{PV}$ and $j = 0$. Statement (iii) could be used to generalize a result of Wilkie [11] who proved (iii) for $T = S_2^1$ and $P = \mathrm{SF}$.

Proof. (i) S_2^1 is $\forall\Sigma_2^b$ and so $S_2^1 \subseteqq \forall\Sigma_i^b(T)$ for $i \geqq 2$. By Lemma 1.3, i-RFN$(P) \in \forall\Sigma_i^b(T)$. On the other hand, assume $\forall x A(x) \in \forall\Sigma_i^b(T)$. Then $S_2^1 \vdash \forall y(P \vdash [\![A]\!]^{|y|})$, for some bounding polynomial. By Lemma 3.3 then

$$S_2^1 + i\text{-RNF}(P) \vdash \forall y \forall x(|x| \leqq |y| \supset A(x)),$$

i.e. $S_2^1 + i$-RFN$(P) \vdash \forall x A(x)$.

(ii) Assume $T \vdash i$-RFN(Q), so

(1) $S_2^1 \vdash (P \vdash [\![d{:}Q \vdash A \wedge A \in \Sigma_i^q \supset \mathrm{Taut}_i(A)]\!]^{|d|+|A|})$.

By Lemma 3.4 (i), as $d{:}Q \vdash A$ and $A \in \Sigma_i^q$ are Σ_1^b-formulas and since P is i-regular we have

(2) $S_2^1 \vdash d{:}Q \vdash A \wedge A \in \Sigma_i^q \supset P \vdash [\![\mathrm{Taut}_i(A)]\!]^{|d|+|A|}$.

Since P is i-regular we can use Lemma 3.4 (ii) to deduce

(3) $S_2^1 \vdash d{:}Q \vdash A \wedge A \in \Sigma_i^q \supset P \vdash A$.

By the main theorem of Buss [1] there is a polynomial time function f such that

$$S_2^1 \vdash d{:}Q \vdash A \wedge A \in \Sigma_i^q \supset f(d, A){:}\ P \vdash A.$$

(iii) Assume $T \vdash \mathrm{NP} = \mathrm{coNP}$. Then every bounded formula is equivalent to a Σ_1^b formula, thus

(1) $T \vdash \mathrm{Taut}_i(A) \equiv (\exists x \leqq t(A))\, B(x, a)$,

propositions is satisfied by any truth valuation". Also it is evident that $\mathrm{SSat}_i \in \Delta^b_{i+1}$ and $\mathrm{STaut}_i \in \Pi^b_{i+1}$.

Fix $i \geqq 1$. Let $A(d)$ be the formula

$$(\forall Z \leqq d)\,(d{:}G_i \vdash Z \supset \mathrm{STaut}_i(Z)).$$

Thus $A(d)$ is a Π^b_{i+1}-formula. We shall prove $A(d)$ by induction on the number of inferences in d, i.e. using Π^b_{i+1}-PIND. As $A(0)$ is trivially true we need only to establish

$$\mathrm{S}^{i+1}_2 \vdash A(\lfloor d/2 \rfloor) \supset A(d).$$

This is proved by checking that any rule of G_i is semantically correct, i.e. that it infers a tautological sequent from tautological premisses. By Lemma 1.2 this is easily checked. (Note that it is also not hard to show the semantical correctness of the substitution rule, cf. [9].) □

Corollary 5.2. *For* $i \geqq 1$, $\mathrm{T}^i_2 \vdash i\text{-RFN}(\mathrm{G}_i)$.

Proof. By Lemma 1.3, i-RFN(G_i) is an $\forall\Sigma^b_i$-sentence. By Buss [2], $\forall\Sigma^b_{i+1}(\mathrm{S}^{i+1}_2) = \forall\Sigma^b_{i+1}(\mathrm{T}^i_2)$. Use Theorem 5.1. □

§ 6. Simulation of arithmetical proofs by propositional calculi

Theorem 6.1. *For* $i \geqq 1$, G_i *simulates* $\forall\Sigma^b_i(\mathrm{T}^i_2)$.

Proof. Assume $d{:}\mathrm{T}^i_2 \vdash A(a)$, where $A \in \Sigma^b_i$. By cut-elimination for T^i_2 (cf. Buss [1, Chapter 4]) we may assume that all formulas in d are in $\Sigma^b_i \cup \Pi^b_i$. Choose a polynomial $q(x)$ which is a bounding polynomial of all formulas occurring in d. The idea of the simulation of d is to replace any formula B in d by its translation $[\![B]\!]^m_{q(m)}$ and to fill some parts in the resulted "preproof" to obtain a G_i-proof of $[\![A]\!]^m_{q(m)}$.

To show that this can be done we shall proceed by induction on the number of inferences in d. Consider several cases according to the type of the last inference in d. We shall write $[\![\ ]\!]$ instead of $[\![\ ]\!]^m_{q(m)}$ and $[\![\Gamma]\!]$ instead of $[\![A_1]\!], \ldots, [\![A_k]\!]$ for a cedent $\Gamma = A_1, \ldots, A_k$.

(a) d is an *initial sequent*, i.e. a logical axiom, an equality axiom or an instance of an axiom of BASIC. The translations of the first two cases are easily proved in G_1. The last case is assured by Lemma 3.5.

(b) The inference is a *structural rule*, *cut-rule* or a *propositional rule*: These cases are handled by the corresponding rules of G_i.

(c) ($\forall$: right)

$$\frac{a \leqq s,\ \Gamma \to \Delta, B(a)}{\Gamma \to \Delta, (\forall x \leqq s)\, B(x)}.$$

Consider two subcases: (c1) s is not of the form $|t|$, (c2) otherwise.

(c1) By ($\supset$: right) derive

$$[\![\Gamma]\!] \to [\![\Delta]\!], [\![a \leqq s \supset B(a)]\!]$$

and using $q(m) + 1$ applications of ($\forall$: right) to the free atoms associated with a derive

$$[\![\Gamma]\!] \to [\![\Delta]\!], [\![(\forall x \leqq s)\, B(x)]\!].$$

(c2) First derive

$$Z_1: \quad \bigvee_{\varepsilon\in S} [\![a = \bar{\varepsilon} \wedge \bar{\varepsilon} \leqq |t|]\!] \to [\![a \leqq |t|]\!],$$

where $S = \{(\varepsilon_0, \ldots, \varepsilon_{q(m)}) \in \{0, 1\}^{q(m)+1} \mid (\forall i > |q(m)|)\ \varepsilon_i = 0\}$, and

$$Z_2: \quad [\![B(a)]\!] \to [\![B(a)]\!].$$

By successively applying ($\supset$: left) and ($\supset$: right) to Z_1, Z_2 get

$$Z_3: \quad [\![a \leqq |t| \supset B(a)]\!] \to \bigvee_{\varepsilon\in S} [\![a = \bar{\varepsilon} \wedge \bar{\varepsilon} \leqq |t|]\!] \supset [\![B(a)]\!].$$

Derive

$$Z_4: \quad \bigvee_{\varepsilon\in S} [\![a = \bar{\varepsilon} \wedge \bar{\varepsilon} \leqq |t|]\!] \supset [\![B(a)]\!] \to \bigwedge_{\varepsilon\in S} [\![a = \bar{\varepsilon} \wedge \bar{\varepsilon} \leqq |t| \supset B(a)]\!]$$

Applying cut-rule to Z_3, Z_4 we obtain

$$Z_5: \quad [\![a \leqq |t| \supset B(a)]\!] \to \bigwedge_{\varepsilon\in S} [\![a = \bar{\varepsilon} \wedge \bar{\varepsilon} \leqq |t| \supset B(a)]\!].$$

Now derive

$$Z_6: \quad \bigwedge_{\varepsilon\in S} [\![a = \bar{\varepsilon} \wedge \bar{\varepsilon} \leqq |t| \supset B(a)]\!] \to \bigwedge_{\varepsilon\in S} [\![a \leqq |t| \supset B(a)]\!]\,(\boldsymbol{p}/\bar{\varepsilon}),$$

and by cut from Z_5, Z_6

$$Z_7: \quad [\![a \leqq |t| \supset B(a)]\!]\,(\boldsymbol{p}) \to \bigwedge_{\varepsilon\in S} [\![a \leqq |t| \supset B(a)]\!]\,(\boldsymbol{p}/\bar{\varepsilon}).$$

Now use cut-rule to Z_7 and to the first sequent derived in the case (c1) to obtain

$$[\![\Gamma]\!] \to [\![\Delta]\!],\ \bigwedge_{\varepsilon\in S} [\![a \leqq |t| \supset B(a)]\!]\,(\boldsymbol{p}/\bar{\varepsilon}).$$

(d) ($\forall$: left)

$$\frac{B(t), \Gamma \to \Delta}{t \leqq s, (\forall x \leqq s)\, B(x), \Gamma \to \Delta}.$$

Again consider two cases: (d1) s is not of the form $|r|$, (d2) otherwise. In both cases we first derive

$$Z_0: \quad [\![t \leqq s]\!], [\![(\forall x \leqq s)\, B(x)]\!] \to [\![B(t)]\!]$$

and apply cut-rule to this sequent and to

$$[\![B(t)]\!], [\![\Gamma]\!] \to [\![\Delta]\!]$$

to get the wanted sequent

$$[\![t \leqq s]\!], [\![(\forall x \leqq s)\, B(x)]\!], [\![\Gamma]\!] \to [\![\Delta]\!].$$

(d1) First derive

$$Z_1: \quad [\![t \leqq s]\!] \to \exists \boldsymbol{x} [\![a \leqq s \wedge a = t]\!]\,(\boldsymbol{p}/\boldsymbol{x})$$

and

$$Z_2: \quad [\![(\forall x \leqq s)\, B(x)]\!], \exists \boldsymbol{x} [\![a \leqq s \wedge a = t]\!]\,(\boldsymbol{p}/\boldsymbol{x}) \to \exists \boldsymbol{x} [\![a = t \wedge B(a)]\!]\,(\boldsymbol{p}/\boldsymbol{x})$$

By cut-rule from Z_1, Z_2 it follows

$$Z_3: \quad [\![t \leqq s]\!], [\![(\forall x \leqq s)\, B(x)]\!] \to \exists \boldsymbol{x} [\![a = t \wedge B(a)]\!]\,(\boldsymbol{p}/\boldsymbol{x}).$$

We shall use the following

Claim. *If $C \in \Sigma_i^b \cup \Pi_i^b$, then for an appropriate bounding polynomial*

$$G_i \vdash [\![t = a \wedge C(a)]\!] \to [\![C(t)]\!].$$

For the proof of the Claim take the open matrix of $C(a)$ and apply Lemma 3. to it. The sequent above is easily got from this sequent in G_i.

Using Claim derive

$$Z_4: \quad \exists \boldsymbol{x}[\![a = t \wedge B(a)]\!]\,(\boldsymbol{p}/\boldsymbol{x}) \rightarrow [\![B(t)]\!]$$

and by cut from Z_3, Z_4 derive Z_0.

(d2) First derive

$$Z_1: \quad [\![t \leqq |r|]\!] \rightarrow \bigvee_{\bar{\varepsilon}\in S} [\![t = a \wedge a \leqq |r|]\!]\,(\boldsymbol{p}/\bar{\varepsilon}),$$

where the set S is the same as in (c2). Then derive

$$Z_2: \quad [\![(\forall x \leqq |r|)\, B(x)]\!], \bigvee_{\bar{\varepsilon}\in S} [\![t = a \wedge a \leqq |r|]\!]\,(\boldsymbol{p}/\bar{\varepsilon}) \rightarrow \bigvee_{\bar{\varepsilon}\in S} [\![t = a \wedge B(a)]\!]\,(\boldsymbol{p}/\bar{\varepsilon})$$

Using cut-rule obtain from Z_1, Z_2

$$Z_3: \quad [\![t \leqq |r|]\!], [\![(\forall x \leqq |r|)\, B(x)]\!] \rightarrow \bigvee [\![t = a \wedge B(a)]\!]\,(\boldsymbol{p}/\bar{\varepsilon}).$$

Using Claim deriv

$$Z_4: \quad \bigvee_{\bar{\varepsilon}\in S} [\![t = a \wedge B(a)]\!]\,(\boldsymbol{p}/\bar{\varepsilon}) \rightarrow [\![B(t)]\!]$$

and by cut from Z_3, Z_4 derive Z_0.

(e) The ($\exists$: rules) are dual to the ($\forall$: rules) and are handled similarly.

(f) Σ_i^b-IND rule:

$$\frac{B(a) \rightarrow B(a+1)}{B(0) \rightarrow B(t)}$$

We omit the side formulas. Assume that we have derived

$$Z: \quad [\![B(a)]\!] \rightarrow [\![B(a+1)]\!].$$

We assume that atoms $\boldsymbol{p}$ are associated with a and atoms $\boldsymbol{q}$ with t. We cannot replace IND by cuts as there would be exponentially many of them in m. We shall shorten the simulation essentially using the substitution rule which is provably simulable in G_1 (Lemma 2.2).

(1) We shall first derive sequents

$$W_0: \quad [\![B(a)]\!] \rightarrow [\![B(a + 2^0)]\!],$$

$$W_{q(m)}: [\![B(a)]\!] \rightarrow [\![B(a + 2^{q(m)})]\!].$$

W_0 is Z. W_{i+1} is derived from W_i as follows: Assume that atoms $\boldsymbol{p}$ are associated to a and new atoms $\boldsymbol{q}$ will be associated to the new variable b. By substitution $\boldsymbol{p} \mapsto \boldsymbol{q}$ derive from W_i

$$W_1': \quad [\![B(a)]\!]\,(\boldsymbol{p}/\boldsymbol{q}) \rightarrow [\![B(a + 2^i)]\!]\,(\boldsymbol{p}/\boldsymbol{q}).$$

Using (the translation of) equality axioms derive

$$W_2': \quad [\![a + 2^i = b]\!]\,(\boldsymbol{p}, \boldsymbol{q}), [\![B(a + 2^i)]\!]\,(\boldsymbol{p}) \rightarrow [\![B(a)]\!]\,(\boldsymbol{p}/\boldsymbol{q}).$$

Apply cut to W_i and W_2' to get

$$W_3': \quad [\![a + 2^i = b]\!]\,(\boldsymbol{p}, \boldsymbol{q}), [\![B(a)]\!]\,(\boldsymbol{p}) \rightarrow [\![B(a)]\!]\,(\boldsymbol{p}/\boldsymbol{q}).$$

Apply cut to W'_1 and W'_3 to get

$$W'_4: \quad [\![a + 2^i = b]\!]\,(\boldsymbol{p}, \boldsymbol{q}), [\![B(a)]\!]\,(\boldsymbol{p}) \to [\![B(a + 2^i)]\!]\,(\boldsymbol{p}/\boldsymbol{q}).$$

Using (the translation of) equality axioms derive

$$W'_5: \quad [\![a + 2^i = b]\!]\,(\boldsymbol{p}, \boldsymbol{q}), [\![B(a + 2^i)]\!]\,(\boldsymbol{p}/\boldsymbol{q}) \to [\![B(a + 2^{i+1})]\!]\,(\boldsymbol{p})$$

Apply cut to W'_4 and W'_5 to get

$$W'_6: \quad [\![a + 2^i = b]\!]\,(\boldsymbol{p}, \boldsymbol{q}), [\![B(a)]\!]\,(\boldsymbol{p}) \to [\![B(a + 2^{i+1})]\!]\,(\boldsymbol{p}).$$

To W'_6 apply $(q(m) + 1)$-times ($\exists$: left) with eigenvariables $\boldsymbol{q}$ to get

$$W'_7: \quad \exists \boldsymbol{x}[\![a + 2^i = b]\!]\,(\boldsymbol{p}, \boldsymbol{x}), [\![B(a)]\!]\,(\boldsymbol{p}) \to [\![B(a + 2^{i+1})]\!]\,(\boldsymbol{p}).$$

Derive

$$W'_8: \quad \to \exists \boldsymbol{x}[\![a + 2^i = b]\!]\,(\boldsymbol{p}, \boldsymbol{x})$$

and apply cut to W'_7 and W'_8 to get W_{i+1}.

(2) Now we shall derive sequents

$$Z_0: \quad [\![2^0 \geqq b]\!]\,(\boldsymbol{q}), [\![B(a)]\!]\,(\boldsymbol{p}) \to [\![B(a + b)]\!]\,(\boldsymbol{p}, \boldsymbol{q})$$

$$\vdots$$

$$Z_{q(m)}: \quad [\![2^{q(m)} \geqq b]\!]\,(\boldsymbol{q}), [\![B(a)]\!]\,(\boldsymbol{p}) \to [\![B(a + b)]\!]\,(\boldsymbol{p}, \boldsymbol{q}).$$

Now Z_0 simply follows from W_0 using

$$[\![2^0 \geqq b]\!]\,(\boldsymbol{q}) \to [\![a = b \vee a + 1 = b]\!]\,(\boldsymbol{p}, \boldsymbol{q}).$$

Z_{i+1} is derived as follows: Take new variables c, d and associate with them atoms $\boldsymbol{r}, \boldsymbol{s}$. By substitution $\boldsymbol{p} \mapsto \boldsymbol{s}$, $\boldsymbol{q} \mapsto \boldsymbol{r}$ derive from Z_i

$$Z'_1: \quad [\![2^i \geqq c]\!]\,(\boldsymbol{r}), [\![B(d)]\!]\,(s) \to [\![B(d + c)]\!]\,(\boldsymbol{s}, \boldsymbol{r}).$$

Derive from W_i

$$Z'_2: \quad [\![B(a)]\!]\,(\boldsymbol{p}), [\![a + 2^i = d]\!]\,(\boldsymbol{p}, \boldsymbol{s}) \to [\![B(d)]\!]\,(\boldsymbol{s}).$$

Apply cut to Z'_1., Z'_2 to get

$$Z'_3; \quad [\![2^i \geqq c]\!]\,(\boldsymbol{r}), [\![a + 2^i = d]\!]\,(\boldsymbol{p}, \boldsymbol{s}), [\![B(a)]\!]\,(\boldsymbol{p}) \to [\![B(d + c)]\!]\,(\boldsymbol{s}, \boldsymbol{r})$$

From Z'_3 derive

$$Z'_4: \quad [\![2^i \geqq c]\!]\,(\boldsymbol{r}), [\![b = 2^i + c]\!]\,(\boldsymbol{q}, \boldsymbol{r}), [\![B(a)]\!]\,(\boldsymbol{p}) \to [\![B(a + b)]\!]\,(\boldsymbol{p}, \boldsymbol{q}).$$

Apply to Z_i and Z'_4 ($\vee$: left) to get

$$Z'_5: \quad [\![2^i \geqq b \vee (2^i \geqq c \wedge b = 2^i + c)]\!]\,(\boldsymbol{q}, \boldsymbol{r}), [\![B(a)]\!]\,(\boldsymbol{p}) \to [\![B(a + b)]\!]\,(\boldsymbol{p}, \boldsymbol{q}).$$

Using ($\exists$: left) applied to eigenatoms $\boldsymbol{r}$ we get

$$Z'_6: \quad \exists \boldsymbol{x}[\![2^i \geqq b \vee (2^i \geqq c \wedge b = 2^i + c)]\!]\,(\boldsymbol{q}, \boldsymbol{r}/\boldsymbol{x}), [\![B(a)]\!]\,(\boldsymbol{p}) \to [\![B(a + b)]\!]\,(\boldsymbol{p}, \boldsymbol{q}).$$

Derive

$$Z'_7: \quad [\![2^{i+1} \geqq b]\!]\,(\boldsymbol{q}) \to \exists \boldsymbol{x}[\![2^i \geqq b \vee (2^i \geqq c \wedge b = 2^i + c)]\!]\,(\boldsymbol{q}, \boldsymbol{r}/\boldsymbol{x}).$$

Finally apply cut to Z'_6 and Z'_7 to get Z_{i+1}.

(3) Now we substitute to $Z_{q(m)}$ $\boldsymbol{p} \mapsto 0$, $\boldsymbol{q} \mapsto \boldsymbol{p}^t$, where $\boldsymbol{p}^t$ are atoms associated to to get

$$[\![2^{q(m)} \geqq t]\!]\,(\boldsymbol{p}^t), [\![B(0)]\!] \to [\![B(t)]\!]\,(\boldsymbol{p}^t).$$

Also is simply derived

$$\to [\![2^{q(m)} \geqq t]\!]\,(\boldsymbol{p}^t).$$

Apply cut to these two sequents to get

$$[\![B(0)]\!] \rightarrow [\![B(t)]\!].$$

This completes the proof. □

Corollary 6.2. *For* $i \geqq 1$, G_i *simulates* $\forall\Sigma_i^b(S_2^{i+1})$.

Proof. By Buss [2], $\forall\Sigma_{i+1}^b(T_2^i) = \forall\Sigma_{i+1}^b(S_2^{i+1})$, for $i \geqq 1$. Use Theorem 6.1. □

Corollary 6.3. *For* $i \geqq j \geqq 1$,

(i) G_i *simulates* $\forall\Sigma_i^b(S_2^{j+1})$ *and* $\forall\Sigma_i^b(T_2^j)$,

(ii) G *simulates* $\forall\Sigma_i^b(S_2)$. □

Consider the simulation of $\forall\Sigma_0^b$ statements. We can choose a translation of the atomic subformulas of a Σ_0^b-formula A such that $[\![A]\!]$ is Π_1^q. Denote by $*[\![A]\!]$ the proposition arising from $[\![A]\!]$ after omitting all quantifiers. So $*[\![A]\!] \in \Sigma_0^q$ and $*[\![A]\!]$ may have other free atoms then those associated to some free variable of A. Then it holds: For $|n_1|, \ldots, |n_k| \leqq m$, $A(a_i/n_i)$ is true iff $*[\![A]\!]^m(p_i^j/n_i(j))$ is tautological.

This is the translation (of Π_1^b-formulas, actually) used in Krajíček-Pudlák [9]. There it is proved, using the results of Cook [4] and Buss [1], that SG_0 simulates $\forall\Sigma_0^b(S_2^1)$ if the translation $*[\![\ ]\!]$ is used.

Observe in the next section that if we used the translation $*[\![\ ]\!]$, the theorems would extend to the case $i = 0$ too with SG_0 instead of G_0.

§ 7. Consequences for fragments S_2^i, T_2^i and for S_2

Now we shall explicitely state the consequences following from the results of § 5 and § 6 for S_2^i and T_2^i.

Corollary 7.1. *For* $i \geqq j \geqq 2$, $\forall\Sigma_j^b(S_2^{i+1}) = \forall\Sigma_j^b(T_2^i)$ *is finitely axiomatized by* $S_2^1 + j\text{-RFN}(G_i)$.

Proof. Use Theorems 4.1 (i) 5.1, 5.2 and 6.3. □

Corollary 7.2. *For* $i \geqq j \geqq 0$, $i \geqq 1$, *if* $S_2^{i+1} \vdash j\text{-RFN}(P)$ *for some proof system* P, *then* $S_2^1 \vdash G_i \geqq^j P$. *The same holds for* $i = 0$ *and* SG_0 *instead of* G_0.

Proof. Use Theorems 4.1 (ii), 6.1 for the case $i \geqq 1$. The case $i = 0$ follows from the results of Cook [4] and Buss [1], cf. Krajíček-Pudlák [9]. □

Corollary 7.3. *For* $i \geqq 1$, *if* $S_2^{i+1} \vdash NP = coNP$, *then there is a polynomial* $p(x)$ *such that*

$$(*) \qquad (\forall A \in TAUT_0)\, \exists d(|d| \leqq p(|A|) \wedge d\colon G_i \vdash A),$$

and S_2^{i+1} *proves* $(*)$. *The same holds for* $i = 0$ *with* SG_0 *instead of* G_0.

Proof. Use Theorems 4.1 (iii), 6.2 for the case $i \geqq 1$. The case $i = 0$ was proved by Wilkie [11], however it can be proved in the same way as for $i \geqq 1$, for details cf. Krajíček-Pudlák [9]. □

Some consequences mentioned above can be transferred to S_2.

Corollary 7.4.

(i) S_2 *is axiomatized by* $S_2^1 + \{i\text{-RFN}(G_i) \mid i < \omega\}$.

(ii) *If* $S_2 \vdash \mathrm{NP} = \mathrm{coNP}$, *then there is a polynomial* $p(x)$ *such that*

(*) $(\forall A \in \mathrm{TAUT}_0)\, \exists d(|d| \leqq p(|A|) \wedge d : \mathrm{G} \vdash A)$,

and S_2 *proves* (*).

(iii) *If* $S_2 \vdash$ 0-RFN(P), *for some proof system* P, *then* $S_2^1 \vdash \mathrm{G} \geqq^0 P$.

Proof. Part (i) is obvious from Corollary 7.1. Parts (ii) and (iii) are derived from Corollaries 7.2, 7.3 using a simple observation: $S_2^1 \vdash \mathrm{G} \geqq^i \mathrm{G}_i$, $i \geqq 0$. □

We shall sketch a nontrivial extension of the preceeding results with interesting corollaries.

Let A be a true $\forall\Sigma_i^b$-sentence, $i \geqq 1$. Define G_i^A to be the extension of G_i where we add initial sequents of the form

$$\to [\![A]\!]_{q(m)}^m$$

for $m = 1, 2, \ldots$ and q a bounding polynomial.

Theorem 7.5. *For* $i \geqq 1$ *and* A *a true* $\forall\Sigma_i^b$-*sentence*

(i) G_i^A *is an* i-*regular proof system,*

(ii) $S_2^{i+1} + A \vdash i$-RFN(G_i^A),

(iii) G_i^A *simulates* $\forall\Sigma_i^b(S_2^{i+1} + A)$.

Proof (sketch): (i) The only nontrivial condition of the definition of i-regular proof systems is the condition (iii). This is proved in the same way as Lemma 3.4 (ii).

(ii) The proof follows the proof of Theorem 5.1. We have only to check that it is provable in $S_2^{i+1} + A$ that initial sequents of G_i^A are tautologies. This follows from Lemma 3.2.

(iii) Here we need a modification of the proof of Theorems 6.1 and 6.2. Again the only difference is in initial sequents and again we use Lemma 3.2. It is also easily seen that the equality $\forall\Sigma_{i+1}^b(T_2^i + A) = \forall\Sigma_{i+1}^b(S_2^{i+1} + A)$ can be obtained from the proof of Buss [2]. □

Corollary 7.6. *For* $i \geqq j \geqq 2$ *and* A *a true* $\forall\Sigma_i^b$-*sentence,*

$$\forall\Sigma_j^b(S_2^{i+1} + A) = \forall\Sigma_j^b(T_2^i + A)$$

and both sets are finitely axiomatized by $S_2^1 + j$-RFN(G_i^A).

Corollary 7.7. *Suppose propositions of* TAUT_1 *have proofs of polynomial length in* G_i, $i > 1$. *Then all propositions in* TAUT_i *have proofs of polynomial length in* G_i.

Proof. Assume TAUT_1 has polynomial proofs in G_i. Thus, in particular, the following formula, denoted by A, is true: $\mathrm{Taut}_0(B) \supset \exists d(|d| \leqq p(|B|) \wedge d : \mathrm{G}_i \vdash B)$, where p is a suitable polynomial. As $S_2^{i+1} \vdash i$-RFN(G_i) we have

$$S_2^{i+1} + A \vdash \mathrm{Taut}_0(B) \equiv \exists d(|d| \leqq p(|d|) \wedge d : \mathrm{G}_i \vdash B),$$

thus $S_2^{i+1} + A$ proves NP = coNP. Hence by theorems 4.1 (iii) and 7.5 TAUT_i has polynomial proofs in G_i^A. But the formulas $[\![A]\!]_{q(m)}^m$ are in TAUT_1 (since $A \in \Sigma_1^b$), hence they have polynomial proofs in G_i. Thus G_i polynomially simulates G_i^A and the corollary follows. □

§ 8. Open problems, conclusions

In previous sections we have left open several questions. In particular, we do not know whether $S_2^1 \vdash G_i \geqq^i G_{i+1}$, whether $S_2^i \vdash i$-RFN(G_i) or whether G_{i-1} simulates $\forall\Sigma_{i-1}^b(T_2^i)$.

It follows from the next two theorems that these problems are important.

Theorem 8.1. *For $i \geqq 1$, the following statements are equivalent:*

(i) $S_2^1 \vdash G_i \geqq^i G_{i+1}$,

(ii) $S_2^{i+1} \vdash i$-RFN(G_{i+1}),

(iii) G_i *simulates* $\forall\Sigma_i^b(T_2^{i+1}) = \forall\Sigma_i^b(S_2^{i+2})$,

(iv) S_2^{i+2} *is* $\forall\Sigma_i^b$*-conservative over* S_2^{i+1}.

The same holds for $i = 0$ with SG_0 instead of G_0.

Proof. (ii) ⇒ (i): use Corollary 7.2. (i) ⇒ (iii): use Theorem 6.1. (iii) ⇒ (iv): use Theorem 5.1 and Lemma 3.3. (iv) ⇒ (ii): use Lemma 1.3 and Theorem 5.1. □

Theorem 8.2. *For $i \geqq 0$, the following statements are equivalent:*

(i) $S_2^{i+1} \vdash (i + 1)$-RFN(G_{i+1}),

(ii) S_2^{i+2} *is* $\forall\Sigma_{i+1}^b$*-conservative over* S_2^{i+1}.

Proof. (i) ⇒ (ii): use Corollary 6.2. (ii) ⇒ (i): use Lemma 1.3 and Theorem 5.1. □

References

[1] Buss, S. R., Bounded Arithmetic. Bibliopolis, Napoli 1986.
[2] Buss, S. R., Axiomatizations and conservation results for fragments of bounded arithmetic. Manuscript, Univ. of California at Berkeley, 1987, 24 p.
[3] Church, A., Introduction to Mathematical Logic, vol. I. Princeton University Press, Princeton, N.J. 1956.
[4] Cook, S. A., Feasibly constructive proofs and the propositional calculus. In: Proc. of the 7th Annual ACM Symp. on Theory of Computing (STOC) 1975, pp. 83–99.
[5] Cook, S. A., and R. A. Reckhow, The relative efficiency of propositional proof systems. J. Symbolic Logic **44** (1979), 36–50.
[6] Dowd, M., Model Theoretic Aspects of P = NP. Manuscript.
[7] Dowd, M., Propositional Representation of Arithmetic Proofs. Ph.D. dissertation, Univ. of Toronto 1979.
[8] Haken, A., The intractability of resolution. Theor. Comput. Sci. **39** (1985), 297–308.
[9] Krajíček, J., and P. Pudlák, Propositional proof systems, the consistency of first order theories and the complexity of computations. J. Symbolic Logic (to appear).
[10] Takeuti, G., Proof Theory. North-Holland Publ. Comp., Amsterdam 1975.
[11] Wilkie, A. J., Subsystems of Arithmetic and Complexity Theory. Invited talk at 8th Intern. Congress LMPS '87, Moscow 1987.

Jan Krajíček and Pavel Pudlák
Mathematical Institute
Czechoslovak Academy of Sciences
Žitná 25
11567 Praha 1
CSSR

(Eingegangen am 28. November 1988)