1998

A Realizability Interpretation for Classical Arithmetic

Jeremy Avigad
Carnegie Mellon University, avigad@cmu.edu

A Realizability Interpretation for Classical Arithmetic

Jeremy Avigad [*]

Carnegie Mellon University, Pittsburgh, PA 15231, USA

Summary. A constructive realizablity interpretation for classical arithmetic is presented, enabling one to extract witnessing terms from proofs of Σ_1 sentences. The interpretation is shown to coincide with modified realizability, under a novel translation of classical logic to intuitionistic logic, followed by the Friedman-Dragalin translation. On the other hand, a natural set of reductions for classical arithmetic is shown to be compatible with the normalization of the realizing term, implying that certain strategies for eliminating cuts and extracting a witness from the proof of a Σ_1 sentence are insensitive to the order in which reductions are applied.

1 Introduction

Even though, as is well known, the classical and intuitionistic versions of first-order arithmetic prove the same Π_2 sentences, the two theories are very different in nature. In particular, the intuitionistic version has a constructive interpretation which seems to be lacking in its classical counterpart.

Heyting arithmetic, which is based on intuitionistic logic, is perhaps best represented in a system of natural deduction. In this framework, proofs can be associated with (or construed as) realizing terms, which come equipped with a natural set of reductions. These reductions are strongly normalizing, which is to say that any procedure that applies them repeatedly will eventually reach a proof that cannot be reduced any further; and they satisfy the Church-Rosser property, which implies that any two such procedures yield the same normal proof. Given a normal proof of a Σ_1 sentence, one can easily read off a numerical witness to the existential quantifier.

In contrast, Peano arithmetic is naturally represented in a classical sequent calculus, to which one can apply cut-elimination to extract numerical witnesses. Cut-elimination procedures tend to seem less canonical than intuitionistic normalization, and the particular witness extracted from a given proof may depend on the technical details of the implementation. Alternatively, one can use a double-negation interpretation followed by the Friedman-Dragalin translation to translate the proof to one in HA, though, of course, the resulting witness will depend on the details of the translation.

In this paper I will try to show that, from a constructive point of view, classical logic is not as bad as it may seem. After some preliminaries in Section 2, I present a realizability interpretation for classical arithmetic in Section 3; by this

I mean that I provide a procedure for assigning a computational term to each sequent in the proof of a Σ_1 formula, in such a way that the term assigned to the final sequent yields a witness to the existential quantifier. In Section 4, I show that this realizability interpretation amounts to nothing more than Kreisel's familiar modified realizability, under a novel translation of classical arithmetic to intuitionistic arithmetic, together with the Friedman-Dragalin translation.

In Section 5, I present set of reductions that one can use to transform any PA-proof of a Σ_1 sentence to a cut-free one. I will show that these reductions are compatible with the normalization of the classical realizing term, which implies that the final result is insensitive to the order in which the reductions are applied. Furthermore, the process of reading off a numerical witness corresponds, in a sense, to the Friedman-Dragalin translation. With a suitable restriction, the reductions are also strongly normalizing.

In Section 6, I show that these reductions can be used to give Peano arithmetic a finitary ordinal analysis. The argument is a variation of Gentzen's original cut-elimination procedure [11, 25], recast here in a Tait-style setting and with some additional simplifications. In Section 7, I use a technique due to Mints [16] to extract reduction sequences of this form from an infinitary cut-elimination procedure as well. As a corollary of the foregoing analysis, both methods of extracting numerical witness from proofs of Σ_1 sentences yield the same result as the normalization of the classical realizing term. Section 8 closes with some final remarks.

A brief discussion of the literature might help put the results reported here in context. In [17], Mints shows that a number of methods of extracting constructive information from proofs in intuitionistic arithmetic yield the same result. Much of this paper can be seen as an extension of these results to classical arithmetic, with the surprising twist that the Friedman-Dragalin translation is implicit in certain cut-elimination arguments for the sequent calculus.

In [4, 5], Buchholz presents a finitary reduction procedure for classical arithmetic, obtained using notations for infinitary derivations and Mints' continuous cut-elimination operators. In particular, in [5] one finds another method of extracting finitary reduction sequences from infinitary cut-elimination arguments, and an analysis of the relationship between these reductions and the ones used in Gentzen's original procedure.

Murthy [19] and Parigot [21] show that versions of the double-negation translation together with the Friedman-Dragalin translation correspond to a certain computational semantics for various systems of classical deduction. Though the work here was carried out without knowledge of these two papers, some of the parallels are striking.

The translation described in Section 4 shows that one can interpret classical sequent calculi in intuitionistic systems of natural deduction, in such a way that proofs of sequents in the former correspond to proofs of a contradiction from suitable hypotheses in the latter. This interpretation may be seen as a contribution

to study of the relationship between sequent calculi and natural deduction in general, which is also addressed in [10, 23, 30].

Finally, the referee has pointed out that another realizability interpretation for arithmetic, somewhat different in form and purpose, is given in [20].

I am grateful to Wilfried Buchholz for the observations noted at the end of Section 6, to Samuel Buss for providing me with the macros that typset the derivations in this paper, and to the anonymous referee for an extremely thorough and helpful review.

2 Preliminaries

The primitive recursive functionals were introduced by Gödel in [13], but were already to some extent implicit in earlier work by Hilbert and Weyl. For a more detailed discussion than the one which follows, see [2, 15, 27, 26].

I will take the finite types to consist of the smallest set containing a symbol $\mathbb{N}$, and closed under an operation which takes types σ and τ to a new type $\sigma \to \tau$. In the intended ("full") interpretation, $\mathbb{N}$ denotes the set of natural numbers, and $\sigma \to \tau$ denotes the set of all functions from σ to τ. A set of terms, PR^ω, is defined inductively as follows:

1. For each type σ, there is a stock of variables $x, y, z, \ldots$ of type σ.
2. 0 is a term of type $\mathbb{N}$.
3. S (successor) is a term of type $\mathbb{N} \to \mathbb{N}$.
4. if s is a term of type $\tau \to \sigma$ and t is a term of type τ, then $s(t)$ is a term of type σ.
5. if s is a term of type σ and x is a variable of type τ, then $\lambda x \, s$ is a term of type $\tau \to \sigma$.
6. If s is a term of type σ, and t is a term of type $\mathbb{N} \to (\sigma \to \sigma)$, then R_{st} is a term of type $\mathbb{N} \to \sigma$.

Intuitively, $s(t)$ denotes the result of applying s to t, $\lambda x \, s$ denotes the function which takes any value of x to s, and R_{st} denotes the function defined from s and t by primitive recursion, with

$$R_{st}(0) = s$$
$$R_{st}(x') = t(x, R_{st}(x)).$$

In this last equation, I have adopted the convention of writing x' instead of $S(x)$ and $t(r, s)$ instead of $(t(r))(s)$. Using these schemata one can define any primitive recursive function, as well as functions, like Ackermann's, that are not primitive recursive. Below when I refer to a primitive recursive relation R, this should be taken as an implicit reference to its characteristic function χ_R.

It will be convenient below to augment the finite types with products $\sigma \times \tau$, associated pairing operations $\langle \cdot, \cdot \rangle$, and projections $(\cdot)_0$ and $(\cdot)_1$. Product types

can eliminated in the usual way by currying and replacing terms t with sequences of terms t_i.

It will also be convenient to have disjoint union types $\sigma + \tau$, an element of which is either an element of σ or an element of τ, tagged to indicate which is the case. That is, for each such type we have insertion operations, inl and inr, which convert elements of type σ and τ respectively to an element of type $\sigma + \tau$; predicates isleft(a) and isright(a), which indicate whether a is tagged to be of type σ or τ; and functions left(a) and right(a), which interpret a as an element of type σ and τ, respectively. Using primitive recursion we can define functions

$$\text{cases}(a, f, g) = \begin{cases} f(a) \text{ if isleft}(a) \\ g(a) \text{ otherwise} \end{cases}$$

where f is of type $\sigma \rightarrow \rho$ for some ρ, and g is of type $\tau \rightarrow \rho$. References to such sum types can be eliminated by taking $\sigma + \tau$ to be $\mathbb{N} \times \sigma \times \tau$, defining $\text{inl}(a) = \langle 0, a, 0^\tau \rangle$, defining $\text{inr}(a) = \langle 1, 0^\sigma, a \rangle$, where 0^σ and 0^τ are constant zero functionals of type σ, τ respectively, and so on.

Finally, to simplify exposition in many of the inductive definitions and proofs below, it will be convenient to have a trivial type Null with a single element nil. Ultimately none of the quantifier-free formulae or terms I define will depend on nil or any variable of type Null, so if we interpret Null $\times \sigma$ and Null $\rightarrow \sigma$ as σ, $\sigma \rightarrow$ Null as Null, and $\forall x \in$ Null φ as φ, references to nil and Null can be eliminated as well.

The realizability interpretation of the next section is designed for classical arithmetic formulated in a Tait-style calculus. In this system formulae are built up from atomic formulae A and their negations $\bar{A}$ using the connectives $\wedge$, $\vee$, $\forall$, and $\exists$. If φ is such a formula, $\neg\varphi$ denotes the formula obtained by putting the negation of φ in negation-normal form, i.e. replacing every $\wedge$, $\forall$, and A in φ respectively with $\vee$, $\exists$, and $\bar{A}$, and conversely. Other connectives are introduced via their classical definitions.

Sequents Γ consist of sets of formulae $\{\varphi_1, \ldots, \varphi_k\}$ and are intended to denote the assertion that at least one of the formulae φ_i is true. As usual, I will use Γ, φ to denote $\Gamma \cup \{\varphi\}$, and if $\varphi(y)$ has been introduced as a formula with a distinguished variable y, $\varphi(t)$ denotes the result of substituting t for y (after renaming bound variables of φ to prevent clashes).

The calculus below is essentially the one found in [24]. In connection with these rules I will refer to the formulae in Γ as the side formulae of the inference, and the other formulae in the hypotheses and conclusions as the main premises and principal formulae respectively. As in [4, 5], a rule of the form "From $\ldots \Gamma_i \ldots$ conclude Γ" should be read "From subsets of $\ldots \Gamma_i \ldots$ conclude Γ," which is to say, implicit weakenings of the hypotheses are allowed at each stage. For technical reasons, I need to assume that in the *Exists* rule below the main premise $\varphi(t)$ is always present. The main premises of a cut rule are also called the cut formulae.

1. Propositional rules
   a) *Atomic excluded middle:* $\Gamma, A, \bar{A}$
   b) *And:* From Γ, φ and Γ, ψ conclude $\Gamma, \varphi \wedge \psi$
   c) *Or:* From either Γ, φ or Γ, ψ conclude $\Gamma, \varphi \vee \psi$
   d) *Cut:* From Γ, φ and $\Gamma, \neg\varphi$ conclude Γ
2. Quantifier rules
   a) *For all:* From $\Gamma, \varphi(y)$ conclude $\Gamma, \forall x \; \varphi(x)$, provided that y is not free in any formula of Γ or $\forall x \; \varphi(x)$
   b) *Exists:* If $\varphi(y)$ is any formula, from $\Gamma, \varphi(t)$ conclude $\Gamma, \exists x \; \varphi(x)$
3. Equality rules

The equality rules consist of quantifier-free sequents asserting the reflexivity, symmetry, and transitivity of equality, and the fact that equality acts as a congruence relation relative to all the functions and relations in the language.

In the language of arithmetic, the only atomic formulae are of the form $t_1 = t_2$. One obtains the rules for classical (Peano) arithmetic, *PA*, by adding the following:

1. Quantifier-free defining equations
2. *Induction:* From $\Gamma, \varphi(0)$ and $\Gamma, \neg\varphi(y), \varphi(y')$ conclude $\Gamma, \forall x \; \varphi(x)$, as long as y is not free in any formula of Γ or $\forall x \; \varphi(x)$

Note that the induction axiom,

$$\Gamma, \varphi(0) \wedge (\forall x \; (\varphi(x) \rightarrow \varphi(x'))) \rightarrow \forall x \; \varphi(x),$$

follows easily from the rule above. As to the "quantifier-free defining equations," we could limit ourselves to weakenings of the sequents $\{x' \neq 0\}, \{x \neq y, x' = y'\}$, and the recursive definitions of addition and multiplication; but it is convenient to include the definitions of arbitrary primitive recursive functions and relations as well.

A variation of the above calculus is obtained if one uses sequents that are *sequences* of formulae rather than sets, and makes structural rules (weakening, exchange, and contraction) explicit. The realizability interpretation in the next section works equally well in this setting. One can simulate the rules of Gentzen's two-sided sequent calculus in this system if one identifies each formula with its negation-normal form, and replaces sequents of the form $\varphi_1, \ldots, \varphi_k \Rightarrow \psi_1, \ldots, \psi_l$ with the corresponding sequent $\neg\varphi_1, \ldots, \neg\varphi_k, \psi_1, \ldots, \psi_l$.

3 Realizability for classical arithmetic

Assuming one can prove a formula φ in *PA*, using a cut one can also prove the empty sequent from the hypothesis $\neg\varphi$. In particular, from a proof of $\exists x \; A(x)$ for a primitive recursive relation A, we obtain a proof of the empty sequent from the hypothesis $\forall x \; \bar{A}(x)$. Our goal is to extract from such a proof a term t in

PR^ω having type $\mathbb{N}$ which satisfies A. In general I will allow A to have parameters (free variables) other than x, in which case t is a function of these parameters. By renaming them if necessary, we can assume that none of these parameters occur in the main premise of a *For all* rule.

Let us then fix A and consider the proof system PA augmented by the hypothesis $\forall x\ \bar{A}(x)$. Having added this hypothesis, each sequent Γ appearing in a proof can no longer be interpreted as the assertion that some formula in Γ is true; instead, we can read it as the assertion that if every formula in Γ is false then the hypothesis is false, and hence $\exists x\ A(x)$ is true. Our strategy will be to analyze the proof of Γ for information that will enable us to compute an element satisfying A from constructive witnesses to the "falsity" of every formula in Γ.

To every formula φ the clauses below inductively associate a type, Type_φ, and a predicate, $\text{Realizes}_\varphi(f)$, for PR^ω-terms f of this type whose free variables are among those of φ. For f of type $\text{Type}_\varphi \to \mathbb{N}$ the predicate $\text{Refutes}_\varphi(f)$ is defined simultaneously, via the equivalence

$$\text{Refutes}_\varphi(f) \equiv \forall g\ (\text{Realizes}_\varphi(g) \to A(f(g))).$$

In words, f maps any g realizing φ to an element satisfying A. Note that if φ has free variables then the predicate $\text{Realizes}_\varphi(f)$ depends on them, in which case the clauses below should be read as asserting that the corresponding equivalences hold for every assignment of natural numbers to these variables.

- $\text{Type}_\varphi = \text{Null},$ if φ is an atomic formula or its negation
  $\text{Realizes}_\varphi(f) \equiv \varphi$

- $\text{Type}_{\varphi \vee \psi} = \text{Type}_\varphi + \text{Type}_\psi$
  $\text{Realizes}_{(\varphi \vee \psi)}(f) \equiv (\text{isleft}(f) \wedge \text{Realizes}_\varphi(\text{left}(f))) \vee$
  $\quad (\text{isright}(f) \wedge \text{Realizes}_\psi(\text{right}(f)))$

- $\text{Type}_{\varphi \wedge \psi} = \text{Type}_{\neg\varphi \vee \neg\psi} \to \mathbb{N}$
  $\text{Realizes}_{\varphi \wedge \psi}(f) \equiv \text{Refutes}_{\neg\varphi \vee \neg\psi}(f)$

- $\text{Type}_{\exists x\ \varphi(x)} = \mathbb{N} \times \text{Type}_{\phi(x)}$
  $\text{Realizes}_{\exists x\ \varphi(x)}(f) \equiv \text{Realizes}_{\varphi(f_0)}(f_1)$

- $\text{Type}_{\forall x\ \varphi(x)} = \text{Type}_{\exists x\ \neg\varphi(x)} \to \mathbb{N}$
  $\text{Realizes}_{\forall x\ \varphi(x)}(f) \equiv \text{Refutes}_{\exists x\ \neg\varphi(x)}(f)$

It is easy to verify the following

Lemma 1. *If $\varphi(x)$ is any formula and t is any term in the language of PA, then $\text{Type}_{\varphi(x)}$ is equal to $\text{Type}_{\varphi(t)}$. Furthermore, if $F(x)$ is any PR^ω-term of this type, then $F(x)$ realizes $\varphi(x)$ if and only if $F(t)$ realizes $\varphi(t)$.*

If φ is a formula, it will be convenient to use α_φ to denote a variable of type Type$_\varphi$. Below, to every proof d of a sequent $\Gamma = \{\varphi_1, \ldots, \varphi_k\}$, I will associate a term F_d of PR$^\omega$, with the following properties:

1. the free variables of F_d are among $\{\alpha_{\neg\varphi_1}, \ldots, \alpha_{\neg\varphi_k}\}$ and the free number variables of Γ and $\exists x\ A(x)$; and
2. on the assumption that each $\alpha_{\neg\varphi_i}$ realizes $\neg\varphi_i$, $A(F_d)$ holds.

In short, F_d takes elements realizing the negations of formulae in Γ to a witness for $\exists x\ A(x)$. In particular, if Γ is the empty sequent then F_d depends only on the free variables of $\exists x\ A(x)$, and for each assignment to these variables, $A(F_d)$ holds.

Before defining the assignment, we need a few lemmata. In Section 5 we will see that the first two are analogous to $\wedge$- and $\forall$-inversion lemmata which are typically found in cut-elimination arguments.

Lemma 2. *Suppose a term f realizes $\varphi \wedge \psi$. Then*

$$\lambda\alpha_{\neg\varphi}\ f\,(\mathrm{inl}(\alpha_{\neg\varphi}))$$

refutes $\neg\varphi$, and

$$\lambda\alpha_{\neg\psi}\ f\,(\mathrm{inr}(\alpha_{\neg\psi}))$$

refutes $\neg\psi$.

Proof. If f realizes $\varphi \wedge \psi$, then it refutes $\neg\varphi \vee \neg\psi$. Assuming $\alpha_{\neg\varphi}$ realizes $\neg\varphi$, $\mathrm{inl}(\alpha_{\neg\varphi})$ realizes $\neg\varphi \vee \neg\psi$, and hence $A(f\,(\mathrm{inl}(\alpha_{\neg\varphi})))$ is true. So $\lambda\alpha_{\neg\varphi}\ f\,(\mathrm{inl}(\alpha_{\neg\varphi}))$ refutes $\neg\varphi$.

The argument for $\lambda\alpha_{\neg\psi}\ f\,(\mathrm{inr}(\alpha_{\neg\psi}))$ is similar. $\square$

In much the same way we can prove

Lemma 3. *Suppose a term f realizes $\forall x\ \varphi(x)$. Then for any t,*

$$\lambda\alpha_{\neg\varphi(t)}\ f\,(\langle t, \alpha_{\neg\varphi(t)}\rangle)$$

refutes $\varphi(t)$.

Lemma 4. *Let φ be any formula. If f and g are terms such that f refutes $\neg\varphi$ and g refutes φ, then there is a term $\mathrm{Cut}_\varphi(f, g)$ satisfying A.*

Proof. If φ is atomic or the negation of an atomic formula, "c refutes φ" means "if φ, then $A(c)$." In that case, use primitive recursion to define

$$\mathrm{Cut}_\varphi(f, g) = \begin{cases} g \text{ if } \varphi \\ f \text{ if } \neg\varphi. \end{cases}$$

If f refutes $\neg(\varphi \wedge \psi)$ then it realizes $\varphi \wedge \psi$, and if f refutes $\neg\forall x\,\varphi$ then it realizes $\forall x\,\varphi$. So define

$$\mathrm{Cut}_{\varphi \wedge \psi}(f, g) = g(f)$$
$$\mathrm{Cut}_{\varphi \vee \psi}(f, g) = f(g)$$
$$\mathrm{Cut}_{\forall x\,\varphi}(f, g) = g(f)$$
$$\mathrm{Cut}_{\exists x\,\varphi}(f, g) = f(g). \qquad\qquad \square$$

The assignment of terms F_d to proofs d is done inductively, according to the last rule of d. In each of the cases below, verifying that F_d satisfies properties 1 and 2 above is straightforward. If Γ is the set $\{\psi_1, \dots, \psi_k\}$, I will use $\alpha_{\neg\Gamma}$ to denote a sequence of variables $\alpha_{\neg\psi_1}, \dots, \alpha_{\neg\psi_k}$. In that case, requirement 2 above is equivalent to the assertion that whenever d is a proof of a sequent Γ, φ and $\alpha_{\neg\Gamma}$ realizes the negations of the elements of Γ, then $\lambda\alpha_{\neg\varphi}\,F_d$ refutes $\neg\varphi$.

Quantifier free axioms: Suppose d consists of a single sequent $\{\varphi_1, \dots, \varphi_k\}$, corresponding to an equality axiom, an instance of the atomic law of the excluded middle, or a quantifier-free axiom of arithmetic. Then at least one of the formulae φ_i is true, and hence it is impossible to realize all the $\neg\varphi_i$'s simultaneously. As a result, we can simply take

$$F = 0.$$

Or: Suppose d is a proof of $\Gamma, \varphi \vee \psi$ obtained by applying the *Or* rule to a proof d' of Γ, φ. Let F' the term inductively assigned to d'. Assuming $\alpha_{\neg\Gamma}$ realizes the negations of the elements of Γ and $\alpha_{\neg\varphi \wedge \neg\psi}$ realizes $\neg\varphi \wedge \neg\psi$, we have that $\lambda\alpha_{\neg\varphi}\,F'$ refutes $\neg\varphi$ and $\lambda\alpha_{\varphi}\,\alpha_{\neg\varphi \wedge \neg\psi}(\mathrm{inl}(\alpha_\varphi))$ refutes φ. Define

$$F = \mathrm{Cut}_\varphi(\lambda\alpha_{\neg\varphi}\,F', \lambda\alpha_\varphi\,\alpha_{\neg\varphi \wedge \neg\psi}(\mathrm{inl}(\alpha_\varphi))).$$

To handle the case where d' is instead a proof of Γ, ψ, replace φ by ψ and inl by inr.

And: Suppose d results by applying the *And* rule to proofs d_0 of Γ, φ and d_1 of Γ, ψ. Let F_0 and F_1 be the terms assigned inductively to d_0 and d_1 respectively. Whenever $\alpha_{\neg\Gamma}$ realizes the negations of the elements of Γ, then $\lambda\alpha_{\neg\varphi}\,F_0$ refutes $\neg\varphi$ and $\lambda\alpha_{\neg\psi}\,F_1$ refutes $\neg\psi$. Define

$$F = \mathrm{cases}(\alpha_{\neg\varphi \vee \neg\psi}, \lambda\alpha_{\neg\varphi}\,F_0, \lambda\alpha_{\neg\psi}\,F_1).$$

Exists: Suppose d consists of an application of the *Exists* rule to a proof d' of $\Gamma, \varphi(t)$. If t has any free variables that do not occur in the conclusion, replace the corresponding free variables of F' with 0. Then, as in the *Or* rule, set

$$F = \mathrm{Cut}_{\varphi(t)}(\lambda\alpha_{\neg\varphi(t)}\,F', \lambda\alpha_{\varphi(t)}\,\alpha_{\forall x\,\neg\varphi(x)}(\langle t, \alpha_{\varphi(t)}\rangle)).$$

For all: If d is obtained by applying the *For all* rule to a proof d' of $\Gamma, \varphi(y)$, then F' may have y free and whenever $\alpha_{\neg\Gamma}$ realizes the negations of the formulae in Γ, then $\lambda\alpha_{\neg\varphi(y)}\,F'$ refutes $\neg\varphi(y)$. Define

$$F = (\lambda y, \alpha_{\neg\varphi(y)} \, F')((\alpha_{\exists x \, \varphi(x)})_0, (\alpha_{\exists x \, \varphi(y)})_1).$$

Cut: Given a proof d_0 of Γ, φ and a proof d_1 of $\Gamma, \neg\varphi$, let F_0 and F_1 be the corresponding terms. If there are any free variables in φ that are not free in Γ, replace these variables by 0 in F_0 and F_1. Define

$$F = \mathrm{Cut}_\varphi(\lambda\alpha_{\neg\varphi} \, F_0, \lambda\alpha_\varphi \, F_1)$$

Induction: Suppose d_0 and d_1 are proofs of $\Gamma, \varphi(0)$ and $\Gamma, \neg\varphi(y), \varphi(y')$ respectively, and let F_0 and F_1 be the corresponding terms. Use primitive recursion to define a function h such that

$$h(0) = \lambda\alpha_{\neg\varphi(0)} \, F_0$$
$$h(y') = \lambda\alpha_{\neg\varphi(y')} \, \mathrm{Cut}_{\varphi(y)}(h(y), \lambda\alpha_{\varphi(y)} \, F_1).$$

Note that by Lemma 1 $\alpha_{\neg\varphi(y)}$ and $\alpha_{\neg\varphi(0)}$ have the same type, so this definition is reasonable. Assuming $\alpha_{\neg\Gamma}$ realizes the negations of the formulae in Γ, one can show by induction that for every natural number n, $h(n)$ refutes $\neg\varphi(n)$. Define

$$F = h((\alpha_{\exists x \, \neg\varphi(x)})_0, (\alpha_{\exists x \, \neg\varphi(x)})_1).$$

Recall that by the conventions introduced in Section 2, this is an abbreviation for $(h((\alpha_{\exists x \, \neg\varphi(x)})_0))((\alpha_{\exists x \, \neg\varphi(x)})_1).$

This takes care of the axioms and rules of *PA*. Now, given a proof of $\{\exists x \, \varphi(x)\}$, cutting with the following hypothesis gives us the desired proof of the empty sequent.

Hypothesis: To the sequent $\{\forall x \, \bar{A}(x)\}$ assign the term

$$F = \alpha_{\exists x \, A(x)}.$$

By definition an element realizing $\exists x \, A(x)$ is just a natural number satisfying A, so if $\alpha_{\exists x \, A(x)}$ realizes $\exists x \, A(x)$, $A(F)$ holds.

Putting this all together yields

Theorem 1. *Suppose PA proves a formula $\exists x \, A(x)$. Then there is a PR^ω-term t of type $\mathbb{N}$, with the same parameters, such that $A(t)$ holds.*

I will use the phrase "provable function" as an abbreviation for "provably total recursive function," which I take to be defined in the usual way.

Corollary 1. *Every provable function of PA is given by a type 1 primitive recursive functional.*

Proof. If *PA* proves $\forall y \, \exists x \, A(y, x)$ then it also proves $\exists x \, A(y, x)$. Let t be a term such that $A(y, t)$ holds, and let $f = \lambda y \, t$. □

4 The M-translation

In this section I will describe a variation of the Gödel-Gentzen double-negation translation that is implicit in the realizability interpretation we have just seen. In contrast to classical logic, I will take the basic intuitionistic connectives to consist of the symbols $\forall$, $\exists$, $\wedge$, $\vee$, $\rightarrow$, and $\bot$; $\sim\varphi$ is defined to be $\varphi \rightarrow \bot$. I will take intuitionistic logic to be given by a system of natural deduction, as in [29], and take the sequent $\Gamma \Rightarrow \varphi$ to mean that φ follows from the hypotheses in Γ. For Heyting arithmetic one adds the quantifier-free defining equations for the basic functions, and an induction rule. In general the law of the excluded middle $\varphi \vee \sim\varphi$ is not assumed to hold, although one can use induction in HA to prove that it holds for atomic formulae.

The double-negation translation maps classical formulae φ to intuitionistic formulae φ^N, and is defined inductively as follows: $(\psi \vee \theta)^N$ is $\sim(\sim\psi^N \wedge \sim\theta^N)$ (or, which is equivalent over intuitionistic logic, $\sim\sim(\psi^N \vee \theta^N)$); $(\exists x\ \psi)^N$ is $\sim\forall x\ \sim\psi^N$ (or, equivalently, $\sim\sim\exists x\ \psi^N$); ψ^N is $\sim\sim\psi$ for atomic formulae ψ; and the translation commutes with the connectives $\wedge$ and $\forall$. Let us consider an alternative translation, which I will call the M-translation, given inductively by the following clauses:

$$
\begin{aligned}
A^M &= A \quad \text{for atomic formulae } A \\
\bar{A}^M &= \sim A \\
(\varphi \vee \psi)^M &= \varphi^M \vee \psi^M \\
(\varphi \wedge \psi)^M &= \sim(\neg\varphi \vee \neg\psi)^M \\
(\exists x\ \varphi)^M &= \exists x\ \varphi^M \\
(\forall x\ \varphi)^M &= \sim(\exists x\ \neg\varphi)^M.
\end{aligned}
$$

The M-translation is not quite equivalent to the N-translation, but one can show by induction that the following relationship holds:

Proposition 1. *For any formula φ in the language of Tait's sequent calculus, one can prove*

$$\sim\varphi^M \leftrightarrow \sim\varphi^N.$$

in intuitionistic logic; *in fact, in minimal logic.*

By minimal logic I mean the fragment of intuitionistic logic which omits the rule "ex falso sequitur quod libet," i.e. "from $\bot$ anything follows." One can think of the M-translation as a version of the double-negation translation which is parsimonious with negations, doling them out only when they are absolutely necessary. In fact, Proposition 1 would still hold if we had defined $(\varphi \wedge \psi)^M$ to be $\varphi^M \wedge \psi^M$, in which case the translation would be trivial on formulae without universal quantifiers.

As a corollary to Proposition 1 we have

Theorem 2. *Suppose $\{\varphi_1, \ldots, \varphi_k\}$ is provable classically, then*

$$(\neg\varphi_1)^M, \ldots, (\neg\varphi_k)^M \Rightarrow \bot$$

is provable over minimal logic.

In fact, a direct translation of proofs is implicit in the realizability interpretation of the previous section. For example, writing $(\neg\Gamma)^M$ to denote the M-translations of the negations of the formulae in Γ, the cut rule,

$$\frac{\Gamma, \varphi \qquad \Gamma, \neg\varphi}{\Gamma}$$

translates to

$$\frac{\dfrac{(\neg\Gamma)^M, (\neg\varphi)^M \Rightarrow \bot}{(\neg\Gamma)^M \Rightarrow {\sim}(\neg\varphi)^M} \qquad \dfrac{(\neg\Gamma)^M, \varphi^M \Rightarrow \bot}{(\neg\Gamma)^M \Rightarrow {\sim}\varphi^M}}{(\neg\Gamma)^M \Rightarrow \bot}$$

This relies on the symmetry of the M-translation, which ensures that either $(\neg\varphi)^M$ is the negation of φ^M, or vice-versa. The $\wedge$ rule,

$$\frac{\Gamma, \varphi \qquad \Gamma, \psi}{\Gamma, \varphi \wedge \psi}$$

translates to

$$\frac{(\neg\Gamma)^M, (\neg\varphi)^M \Rightarrow \bot \qquad (\neg\Gamma)^M, (\neg\psi)^M \Rightarrow \bot}{(\neg\Gamma)^M, (\neg\varphi)^M \vee (\neg\psi)^M \Rightarrow \bot}$$

and the $\vee$ rule,

$$\frac{\Gamma, \varphi}{\Gamma, \varphi \vee \psi}$$

translates to

$$\frac{\dfrac{(\neg\Gamma)^M, (\neg\varphi)^M \Rightarrow \bot}{(\neg\Gamma)^M \Rightarrow {\sim}(\neg\varphi)^M} \qquad {\sim}(\varphi^M \vee \psi^M) \Rightarrow {\sim}\varphi^M}{(\neg\Gamma)^M, {\sim}(\varphi^M \vee \psi^M) \Rightarrow \bot}$$

where in the last derivation one uses a "canonical" proof of ${\sim}(\varphi^M \vee \psi^M) \Rightarrow {\sim}\varphi^M$. The quantifier rules are treated similarly. The translation works equally well for *PA* and *HA*, so we have

Theorem 3. *If $\{\varphi_1, \ldots, \varphi_k\}$ is provable in PA, then*

$$(\neg\varphi_1)^M, \ldots, (\neg\varphi_k)^M \Rightarrow \bot$$

is provable in arithmetic over minimal logic.

In fact, since in arithmetic every atomic formula has an atomic "complement," here we can define $\bar{A}^M$ to be $\bar{A}$.

We can now employ a trick, due to Friedman [9] and Dragalin independently. Given a proof in minimal logic, if one replaces $\perp$ by an arbitrary formula, then every rule in the original proof remains valid. (If the original proof is in intuitionistic logic, one must first replace atomic formulae θ by $\theta \vee \perp$.) Now, starting with a proof of $\exists x\ A(x)$ in classical arithmetic, Theorem 3 yields a proof of $\perp$ from $\exists x\ A(x) \to \perp$. Replacing $\perp$ by $\exists x\ A(x)$ yields a proof of $\exists x\ A(x)$ from $\exists x\ A(x) \to \exists x\ A(x)$, and hence a proof of $\exists x\ A(x)$ outright. This yields

Theorem 4. *If PA proves $\exists x\ A(x)$, then so does HA.*

Finally, one can apply Kreisel's modified form of Kleene's realizability (see, for example, [28]) to the proof in HA. Roughly speaking, one declares that an atomic formula is realized if and only if it's true; a realizer for $\varphi \wedge \psi$ is a realizer for φ paired with a realizer for ψ; a realizer for $\varphi \vee \psi$ is either a realizer for φ or a realizer for ψ, with a tag to indicate which is the case; a realizer for $\varphi \to \psi$ is a function(al) mapping any realizer for φ to a realizer for ψ; a realizer for $\forall x\ \varphi(x)$ is a function mapping any $a \in \mathbb{N}$ to a realizer for $\varphi(a)$; and a realizer for $\exists x\ \varphi(x)$ is an element $a \in \mathbb{N}$ paired with a realizer for $\varphi(a)$. The result is that from the proof of $\exists x\ A(x)$ in HA one obtains a witnessing term, providing an alternative proof of Theorem 1. Or rather, the same proof in disguise: the reader can now check that the classical realizability interpretation given in the last section is nothing more than the modified realizability interpretation coupled with the translation of the classical sequent calculus to natural deduction given by Theorem 3, once we replace $\perp$ by $\exists x\ A(x)$. Note that cutting with the hypothesis $\forall x\ \bar{A}(x)$ on the classical side corresponds to utilizing the canonical proof of $\exists x\ A(x) \to \exists x\ A(x)$ on the intuitionistic side. In sum, we have the following

Theorem 5. *Given a proof of $\exists x\ A(x)$ in PA, the witnessing term extracted using the classical realizability of Section 3 is the same as the witnessing term obtained by translating the classical proof to an intuitionistic one using the M-translation, applying the Friedman-Dragalin translation, and then applying modified realizability.*

At this point it seems worthwhile to mention another interpretation, that is a little closer to the N-interpretation:

$$A^L = \sim\sim A \quad \text{for atomic formulae } A$$
$$\bar{A}^L = \sim A$$
$$(\varphi \wedge \psi)^L = \varphi^L \wedge \psi^L$$
$$(\varphi \vee \psi)^L = \sim(\neg\varphi \wedge \neg\psi)^L$$
$$(\forall x\ \varphi)^L = \forall x\ \varphi^L$$
$$(\exists x\ \varphi)^L = \sim(\forall x\ \neg\varphi)^L.$$

For this translation we have the following

Proposition 2. *For any formula φ in negation-normal form, one can prove*

$$\varphi^L \leftrightarrow \varphi^N.$$

in intuitionistic logic.

Had we used the L-translation in the last section, we would have been able to translate every proof of $\{\varphi_1, \ldots, \varphi_k\}$ in PA to a proof of $\bot$ from assumptions $\sim\varphi_1^L, \ldots \sim\varphi_k^L$ in HA; but dealing with the extra negation symbol would have been unwieldy.

5 Reductions for Classical Arithmetic

In this section I will present a set of reductions for proofs in a slight extension of the system PA defined above. We will see that these reductions are compatible with the normalization of the term extracted by realizability, in the following sense: if one reduces a proof d to d', then F_d and $F_{d'}$ convert to the same term. In the next two sections I will show that these reductions can be used to eliminate cuts from proofs of Σ_1 sentences in PA and extract a witness for the conclusion.

To begin, let us augment the proof system PA in the following two ways:

1. Close the set of axioms under substitution of terms for free variables, and cut.
2. Add the following "inversion" rules:
   a) *Atomic inversion*: If φ is a true closed atomic formula, from $\Gamma, \bar{\varphi}$ conclude Γ.
   b) *And inversion*: From $\Gamma, \varphi \wedge \psi$ conclude either Γ, φ or Γ, ψ
   c) *For all inversion*: From $\Gamma, \forall x\, \varphi(x)$ conclude $\varphi(n)$ for any numeral n.

If the reader is concerned with the loss of transparency due to clause 1, he or she can insist that these axioms be labeled with a derivation from the original ones. Since weakenings are allowed, an axiom may contain complex side formulae; I will take the principal formulae in an axiom to be those that are either atomic or negation atomic. Though it is a slight abuse of notation I will henceforth use PA to refer to the system of Section 2 with these modified axioms, and I will use PA^+ to refer to the system with the additional inversion rules.

The realizability interpretation of Section 3 can easily be extended to handle the new rules. If d is one of the new axioms, we can still take $F_d = 0$. If d is obtained by an application of the *Atomic inversion* rule to another proof d', we can just take $F_d = F_{d'}$. If d is obtained from d' by an application of *And inversion*, as in the proof of Lemma 2 we can take F_d to be either $(\lambda\alpha_{\neg\varphi \vee \neg\psi}\, F_{d'})(\text{inl}(\alpha_{\neg\varphi}))$ or $(\lambda\alpha_{\neg\varphi \vee \neg\psi}\, F_{d'})(\text{inr}(\alpha_{\neg\varphi}))$ as necessary. Finally, if d is obtained by applying *For all inversion* to d', we can take F_d to be $(\lambda\alpha_{\exists x\, \neg\varphi(x)}F_{d'})(\langle n, \alpha_{\neg\varphi(n)}\rangle)$.

Terms in PR^ω come equipped with a natural reducibility relation, under which $(\lambda x\; t)(s)$ reduces to $t[s/x]$, $R_{st}(0)$ reduces to s, and $R_{st}(u')$ reduces to

$t(u, R_{st}(u))$. I will write $s \rightsquigarrow t$ if t can be obtained by iteratively applying such reductions to subterms of s, and $t \leftrightsquigarrow s$ if t and s reduce to a common term, up to a renaming of bound variables. A term is said to be in normal form if it cannot be reduced any further. More precise definitions can be found in [15].

It is well known that the reductions just described are *strongly normalizing* (i.e. there are no infinite sequences of one-step reductions, so that arbitrary reduction procedures are guaranteed to terminate) and *confluent* (i.e. the relation $\leftrightsquigarrow$ is an equivalence relation). Taken together these imply that every term t reduces to a unique term in normal form. It is easy to check that if t is a closed term of type $\mathbb{N}$ then the normal form of t must be a numeral, for which I will write n_t. Note that if C is a primitive recursive relation and $t_1, t_2, \ldots, t_k$ are closed terms, then $\chi_C(t_1, t_2, \ldots, t_k)$ reduces to either 0 or 1.

Under the conventions described in Section 2, we have the following reductions:

1. a) $Cut_C(a, b) \rightsquigarrow a$ if C is a false closed atomic formula
    b) $Cut_C(a, b) \rightsquigarrow b$ if C is a true closed atomic formula
2. a) $\text{cases}(\text{inl}(a), f, g) \rightsquigarrow f(a)$
    b) $\text{cases}(\text{inr}(a), f, g) \rightsquigarrow g(a)$
3. a) $(\langle a, b \rangle)_0 \rightsquigarrow a$
    b) $(\langle a, b \rangle)_1 \rightsquigarrow b$

If d is a PA-proof of a sequent Γ with free variable x, let $d[t/x]$ denote the proof of $\Gamma[t/x]$ obtained by inductively replacing x by t in d (barring instances of x that are not "related" to those in the final sequent). Similarly, if t is a closed term of type $\mathbb{N}$, let $d[t \rightsquigarrow n_t]$ denote the result of replacing t by n_t. One can verify that $d[t/x]$ and $d[t \rightsquigarrow n_t]$ are again proofs in PA, and we have the following

Lemma 5. $F_{d[t/x]} = F_d[t/x]$ and $F_d \rightsquigarrow F_{d[t \rightsquigarrow n]}$.

I will now define a set of transformations with which a proof d in PA^+ can be converted to another proof d' having the same conclusion. I will say that d *reduces to d' in one step*, written $d \rightsquigarrow_1 d'$, if d' can be obtained by applying one of these transformations to a subproof of d; and d *reduces to d'*, written $d \rightsquigarrow d'$, if d' can be obtained from d by a series of one-step reductions. Finally, say that d is in *normal form* if no further one-step reduction is possible.

In illustrating the transformations below, I have chosen instances in which the hypotheses in the proof on the left are as "large" as possible. But only the rules are important, which is to say that the same transformations are allowed when the hypotheses in the proofs on the left are subsets of the ones shown.

The reductions on proofs d are organized according to the last rule of d, as follows:

1. *Principal cuts:* Cuts in which the cut formula that is atomic, of the form $\varphi \vee \psi$, or of the form $\exists x \, \varphi(x)$, is principal in the previous inference.

a) *Axiom:* if φ closed and atomic and the principal formula of the preceding inference, then that inference must be an axiom.

$$\frac{\dfrac{d_0}{\Gamma,\varphi} \quad \Gamma,\bar{\varphi}}{\Gamma} \quad \text{reduces to} \quad \Gamma \quad \text{or} \quad \frac{\dfrac{d_0}{\Gamma,\bar{\varphi}}}{\Gamma}$$

If φ is false then $\{\bar{\varphi}\}$ is an axiom, and so Γ is also an axiom. If φ is true we can apply atomic inversion to d_0.

b) *Or:*

$$\frac{\dfrac{\dfrac{d_0}{\Gamma,\varphi \vee \psi,\varphi}}{\Gamma,\varphi \vee \psi} \quad \dfrac{d_1}{\Gamma,\neg\varphi \wedge \neg\psi}}{\Gamma}$$

reduces to

$$\frac{\dfrac{\dfrac{d_0}{\Gamma,\varphi \vee \psi,\varphi} \quad \dfrac{d_1}{\Gamma,\neg\varphi \wedge \neg\psi}}{\Gamma,\varphi} \quad \dfrac{\dfrac{d_1}{\Gamma,\neg\varphi \wedge \neg\psi}}{\Gamma,\neg\varphi}}{\Gamma}$$

First, cut d_0 against d_1; then invert d_1 and cut again on a smaller formula. The case where the *Or* rule is applied on the right is similar.

c) *Exists:* if t is a closed term,

$$\frac{\dfrac{\dfrac{d_0}{\Gamma,\exists x\, \varphi(x),\varphi(t)}}{\Gamma,\exists x\, \varphi(x)} \quad \dfrac{d_1}{\Gamma,\neg\forall x\, \varphi(x)}}{\Gamma}$$

reduces to

$$\frac{\dfrac{\dfrac{d_0[t \rightsquigarrow n_t]}{\Gamma,\exists x\, \varphi(x),\varphi(n_t)} \quad \dfrac{d_1}{\Gamma,\forall x\, \varphi(x)}}{\Gamma,\varphi(n_t)} \quad \dfrac{\dfrac{d_1}{\Gamma,\forall x\, \varphi(x)}}{\Gamma,\varphi(n_t)}}{\Gamma}$$

This is similar to the last reduction. Here one converts t to a numeral n_t, and uses n_t in the *For all inversion* on the right.

2. *Nonprincipal cuts:* cuts in which the cut formula that is atomic, of the form $\varphi \vee \psi$, or of the form $\exists x\, \varphi$, is not principal in the preceeding inference.

   a) If the preceding rule is *And*, *Or*, *Exists*, *For all*, or an axiom, pass the cut "through" this rule. For example,

$$\frac{\dfrac{\cdots \quad \dfrac{d_i}{\Gamma_i,\exists x\, \varphi(x)} \quad \cdots}{\Gamma,\exists x\, \varphi(x)} \quad \dfrac{\hat{d}}{\Gamma,\forall x\, \neg\varphi(x)}}{\Gamma}$$

reduces to

$$
\frac{
\begin{array}{cc}
d_i & \hat{d} \\
\Gamma_i, \exists x \; \varphi(x) & \Gamma, \forall x \; \neg\varphi(x) \\
\end{array}
}{
\cdots \quad \dfrac{\Gamma, \Gamma_i}{\Gamma} \quad \cdots
}
$$

Note that if $\Gamma, \exists x \; \varphi(x)$ is an axiom then so is Γ, and hence Γ, Γ_i as well.

b) If the preceding rule is again a cut rule, pass the lower cut through the one above. For example,

$$
\frac{
\dfrac{
\begin{array}{cc}
d_0 & d_1 \\
\Gamma, \exists x \; \varphi, \theta & \Gamma, \exists x \; \varphi, \neg\theta
\end{array}
}{\Gamma, \exists x \; \varphi}
\qquad
\begin{array}{c}
d_2 \\
\Gamma, \forall x \; \neg\varphi
\end{array}
}{\Gamma}
$$

reduces to

$$
\frac{
\dfrac{
\begin{array}{cc}
d_0 & d_2 \\
\Gamma, \exists x \; \varphi, \theta & \Gamma, \forall x \; \neg\varphi
\end{array}
}{\Gamma, \theta}
\qquad
\dfrac{
\begin{array}{cc}
d_1 & d_2 \\
\Gamma, \exists x \; \varphi, \neg\theta & \Gamma, \forall x \; \neg\varphi
\end{array}
}{\Gamma, \neg\theta}
}{\Gamma}
$$

3. *Elimination of unnecessary free variables:* If the last inference of d is either *Cut* or *Exists*, and there is a free variable in one of the hypotheses that is not free in the conclusion, replace that free variable with 0. For example,

$$
\begin{array}{ccc}
\dfrac{
\begin{array}{c}
d_0 \\
\Gamma, \exists x \; \varphi(x), \varphi(t(y))
\end{array}
}{\Gamma, \exists x \; \varphi(x)}
& \text{reduces to} &
\dfrac{
\begin{array}{c}
d_0[0/y] \\
\Gamma, \exists x \; \varphi(x), \varphi(t(0))
\end{array}
}{\Gamma, \exists x \; \varphi(x)}
\end{array}
$$

4. *Principal inversions:* These are inversions whose main formula is the principal formula of the preceding inference.

   a) *Atomic inversion:*

$$
\frac{\Gamma, \bar{\varphi}}{\Gamma} \quad \text{reduces to} \quad \Gamma
$$

   If $\bar{\varphi}$ is principal in the previous inference, then that inference must be an axiom. If φ is a true atomic formula, then $\{\bar{\varphi}\}$ is an axiom, and hence so is Γ.

   b) *And inversion:*

$$
\begin{array}{ccc}
\dfrac{
\dfrac{
\begin{array}{cc}
d_0 & d_1 \\
\Gamma, \varphi \wedge \psi, \varphi & \Gamma, \varphi \wedge \psi, \psi
\end{array}
}{\Gamma, \varphi \wedge \psi}
}{\Gamma, \varphi}
& \text{reduces to} &
\dfrac{
\begin{array}{c}
d_0 \\
\Gamma, \varphi \wedge \psi, \varphi
\end{array}
}{\Gamma, \varphi}
\end{array}
$$

   In words, one simply implies the inversion to the subproof d_0. The case where the inversion is on the right formula ψ is similar.

c) *For all inversion,* where the principal formula is the conclusion of a *For all* rule:

$$
\cfrac{\cfrac{\begin{array}{c}d_0\\[2pt]\Gamma, \forall x\ \varphi(x), \varphi(y)\end{array}}{\Gamma, \forall x\ \varphi(x)}}{\Gamma, \varphi(n)}
\qquad \text{reduces to} \qquad
\cfrac{\begin{array}{c}d_0[n/y]\\[2pt]\Gamma, \forall x\ \varphi(x), \varphi(n)\end{array}}{\Gamma, \varphi(n)}
$$

In this case, one applies the inversion to $d_0[n/y]$.

d) *For all inversion,* where the principal formula is the conclusion of the induction rule:

$$
\cfrac{\cfrac{\begin{array}{c}d_0\\[2pt]\Gamma, \forall x\ \varphi(x), \varphi(0)\end{array} \qquad \begin{array}{c}d_1\\[2pt]\Gamma, \forall x\ \varphi(x), \neg\varphi(y), \varphi(y')\end{array}}{\Gamma, \forall x\ \varphi(x)}}{\Gamma, \varphi(n)}
$$

reduces to

$$
\cfrac{\cfrac{\cfrac{\begin{array}{c}d_0\\[2pt]\Gamma, \forall x\ \varphi(x), \varphi(0)\end{array} \quad \begin{array}{c}d_1[0/y]\\[2pt]\Gamma, \forall x\ \varphi(x), \neg\varphi(0), \varphi(1)\end{array}}{\Gamma, \forall x\ \varphi(x), \varphi(1)} \quad \begin{array}{c}d_1[1/y]\\[2pt]\Gamma, \forall x\ \varphi(x), \neg\varphi(1), \varphi(2)\end{array}}{\begin{array}{c}\Gamma, \forall x\ \varphi(x), \varphi(2)\\[4pt]\vdots\\[2pt]\Gamma, \forall x\ \varphi(x), \varphi(n)\end{array}}}{\Gamma, \varphi(n)}
$$

In words, one applies n cuts to $d_0, d_1[0/y], \dots, d_1[n-1/y]$, and then inverts the result.

5. *Nonprincipal inversions:* if the last rule is an inversion applied to a formula that is not principal in the preceding inference, carry out the inversion on the hypotheses of the preceding rule instead. For example,

$$
\cfrac{\cfrac{\dots \quad \begin{array}{c}d_i\\[2pt]\Gamma_i, \forall x\ \varphi(x)\end{array} \quad \dots}{\Gamma, \forall x\ \varphi(x)}}{\Gamma, \varphi(n)}
$$

reduces to

$$
\cfrac{\dots \quad \cfrac{\begin{array}{c}d_i\\[2pt]\Gamma_i, \forall x\ \varphi(x)\end{array}}{\Gamma_i, \varphi(n)} \quad \dots}{\Gamma, \varphi(n)}
$$

The reductions for the other inversion rules are similar.

Lemma 7 below shows that a proof of a Σ_1 sentence is normal (irreducible) if and only if it is cut-free (and this remains true even if reductions of the form 2b are disallowed). In the next two sections I will present both finitary and infinitary cut-elimination procedures that use these reductions to transform such proofs to ones in normal form. This shows that these reductions are *weakly normalizing* when applied to proofs of Σ_1 sentences. Note that in the presence of rule 2b they are *not* strongly normalizing, since this rule makes it possible to pass two cuts through each other ad infinitum; but Theorem 10 shows that without 2b, the resulting set of reductions is, in fact, strongly normalizing.

Leaving the issue of normalizability aside for the moment, I will now show that the reductions above are compatible with the reduction of the realizing term extracted in Section 3.

Lemma 6. *If a proof d is reduced to another proof d' according to any of the transformations listed above, then $F_d \leftrightsquigarrow F_{d'}$.*

Proof. The proof requires checking each reduction. In each case, I will use d to denote the proof on the left, d' to denote the proof on the right, and d_0, d_1, etc. as in the diagrams above. Similarly, I will write F instead of F_d, F' instead of $F_{d'}$, and so on.

When it comes to the principal cuts, the atomic case 1a is easy. Consider case 1b, in which the cut is on a formula of the form $\varphi \vee \psi$. In this case, F is given by

$$\left(\lambda\alpha_{\neg\varphi\wedge\neg\psi} \; \mathrm{Cut}_\varphi \left(\lambda\alpha_{\neg\varphi} \; F_0, \lambda\alpha_\varphi \; \alpha_{\neg\varphi\wedge\neg\psi}(\mathrm{inl}(\alpha_\varphi)) \right) \right) (\lambda\alpha_{\varphi\vee\psi} \; F_1).$$

F' is given by

$$\mathrm{Cut}_\varphi \left(\lambda\alpha_{\neg\varphi} \; ((\lambda\alpha_{\neg\varphi\wedge\neg\psi} \; F_0)(\lambda\alpha_{\varphi\vee\psi} \; F_1)), \lambda\alpha_\varphi \; ((\lambda\alpha_{\varphi\vee\psi} \; F_1)(\mathrm{inl}(\alpha_{\neg\varphi}))) \right).$$

Both terms reduce to

$$\mathrm{Cut}_\varphi \left(\lambda\alpha_{\neg\varphi} \; F_0[(\lambda\alpha_{\varphi\vee\psi} \; F_1)/\alpha_{\neg\varphi\wedge\neg\psi}], \lambda\alpha_\varphi \; ((\lambda\alpha_{\varphi\vee\psi} \; F_1)(\mathrm{inl}(\alpha_\varphi))) \right).$$

Case 1c is handled similarly.

To handle the nonprincipal cuts, note that in the example shown for 2a, F is of the form

$$\left(\lambda\alpha_{\forall x \; \neg\varphi(x)} \; t(\ldots F_i \ldots) \right) (\lambda\alpha_{\exists x \; \varphi(x)} \; \hat{F})$$

where t is a term and the only free occurrences of $\alpha_{\forall x \; \neg\varphi(x)}$ occurring in t are those that occur in the terms F_i. F' is then given by

$$t\left(\ldots \left(\lambda\alpha_{\forall x \; \neg\varphi(x)} \; F_i \right) (\lambda\alpha_{\exists x \; \varphi(x)} \; \hat{F}) \ldots \right).$$

Both terms reduce to

$$t(\ldots F_i[\lambda\alpha_{\exists x\ \varphi(x)}\ \hat{F}/\alpha_{\forall x\ \neg\varphi(x)}]\ldots).$$

Case 2b is similar.

Since we substituted 0 for extraneous free variables in the defining clauses of Section 3, the application of a reduction according to 3 leaves the realizing term unchanged.

Checking that the claim holds of *Atomic inversion*, 4a, is easy. To handle the instance of *And inversion* shown in 4b, note that F is given by

$$\Big(\lambda\alpha_{\neg\varphi\vee\neg\psi}\ \mathrm{cases}(\alpha_{\neg\varphi\wedge\neg\psi},\lambda\alpha_{\neg\varphi}\ F_0,\lambda\alpha_{\neg\psi}\ F_1)\Big)(\mathrm{inl}(\alpha_{\neg\varphi})),$$

and that in this expression $\alpha_{\neg\varphi}$ occurs both free and bound; to reduce this term, one must temporarily replace $\lambda\alpha_{\neg\varphi}\ F_0$ by $\lambda\beta_{\neg\varphi}\ F_0[\beta_{\neg\varphi}/\alpha_{\neg\varphi}]$. F' is given by

$$\big(\lambda\alpha_{\neg\varphi\vee\neg\psi}\ F_0\big)(\mathrm{inl}(\alpha_{\neg\varphi})).$$

Both F and F' reduce to the term

$$F_0[\mathrm{inl}(\alpha_{\neg\varphi})/\alpha_{\neg\varphi\vee\neg\psi}].$$

For all inversion following a *For all rule* as in 4c is handled similarly. Finally, in case 4d, where *For all inversion* follows *Induction*, F is given by

$$\Big(\lambda\alpha_{\exists x\ \neg\varphi(x)}\ h((\alpha_{\exists x\ \neg\varphi(x)})_0,(\alpha_{\exists x\ \neg\varphi(x)})_1)\Big)(\langle n,\alpha_{\neg\varphi(n)}\rangle)$$

where h is the function defined in the *Induction* clause of Section 3. Note that in general $\alpha_{\exists x\ \neg\varphi(x)}$ will be free in F_0 and F_1, and $\alpha_{\varphi(y)}$ and y will be free in F_1. After renaming some of the variables for clarity, F' is given by

$$\Big(\lambda\alpha_{\exists x\ \neg\varphi(x)}\ \mathrm{Cut}_{\varphi(n-1)}\Big(\lambda\alpha_{\neg\varphi(n-1)}\ \mathrm{Cut}_{\varphi(n-2)}(\ldots$$
$$\lambda\alpha_{\neg\varphi(1)}\ \mathrm{Cut}_{\varphi(0)}(\lambda\alpha_{\neg\varphi(0)}\ F_0,\lambda\alpha_{\varphi(0)}\ F_1[0/y,\alpha_{\varphi(0)}/\alpha_{\varphi(y)}])\ldots),$$
$$\lambda\alpha_{\varphi(n-1)}\ F_1[n-1/y,\alpha_{\varphi(n-1)}/\alpha_{\varphi(y)}]\Big)\Big)(\langle n,\alpha_{\neg\varphi(n)}\rangle).$$

Both terms reduce to

$$\mathrm{Cut}_{\varphi(n-1)}\Big(\lambda\alpha_{\neg\varphi(n-1)}\ \mathrm{Cut}_{\varphi(n-2)}(\ldots$$
$$\lambda\alpha_{\neg\varphi(1)}\ \mathrm{Cut}_{\varphi(0)}(\lambda\alpha_{\neg\varphi(0)}\ F_0[\langle n,\alpha_{\neg\varphi(n)}\rangle/\alpha_{\exists x\ \neg\varphi(x)}],$$
$$\lambda\alpha_{\varphi(0)}\ F_1[0/y,\alpha_{\varphi(0)}/\alpha_{\varphi(y)}\langle n,\alpha_{\neg\varphi(n)}\rangle/\alpha_{\exists x\ \neg\varphi(x)}])\ldots),$$
$$\lambda\alpha_{\varphi(n-1)}\ F_1[n-1/y,\alpha_{\varphi(n-1)}/\alpha_{\varphi(y)},\langle n,\alpha_{\neg\varphi(n)}\rangle/\alpha_{\exists x\ \neg\varphi(x)}]\Big).$$

The nonprincipal inversions in case 5 are handled in a manner similar to the one we used to deal with the nonprincipal cuts.   □

Corollary 2. *If $d \rightsquigarrow d'$, then $F_d \leftrightsquigarrow F_{d'}$.*

Before we proceed let me introduce some terminology.

Definition 1. 1. *In any cut, the cut formula that is either atomic, of the form $\varphi \vee \psi$, or of the form $\exists x \; \varphi$ is called the* active premise *of the cut. The corresponding sequent is called the* active sequent.
  2. *A cut is* movable *if it is subject to a reduction of the form 2a. In other words, a cut is movable if it is nonprincipal, and the active sequent is an axiom or the conclusion of an* And, Or, Exists, *or* For all *inference.*
  3. *A formula is Σ if it contains no universal quantifiers. A sequent is Σ (resp. closed) if every formula in it is Σ (resp. closed).*

Consider a cut whose conclusion is a closed sequent. Using the terminology above we can observe that if this cut is neither principal nor movable, then the active sequent is the conclusion of either an inversion, an induction inference, or another cut. If the conclusion of the original cut happens to be Σ, then the active sequent is Σ and the second possibility is eliminated.

The next lemma shows that normal proofs of closed Σ sequents are cut-free (even if one omits reductions of the form 2b).

Lemma 7. 1. *If d is a PA^+-proof with an instance of an inversion rule, then a reduction of type 4 or 5 applies to d. Hence any normal PA^+ proof is a PA-proof.*
  2. *If d is a PA-proof of a closed Σ sequent whose last inference is a* Cut, *then a reduction of type 1, 2a, or 3 applies to d.*
  3. *If d is a normal PA-proof of a closed Σ sequent, then d is cut-free and every sequent is Σ and closed.*

Proof. In each case, the proof proceeds by induction on d. One obtains the first clause by noting that a reduction of type 4 or 5 can be applied to any inversion rule, unless the preceding rule is again an inversion.

To handle the second clause, suppose that no reduction of type 1, 2a, or 3 applies to the last cut; in other words this cut is neither principal nor movable, and there are no unnecessary free variables. Then the active sequent of this cut is again a closed Σ sequent, and the hypotheses imply that this sequent can only be the conclusion of another cut.

To obtain the final clause, note that by (2) the last inference of d cannot be a cut, and there are no unnecessary free variables. It follows that the immediate subproofs of d are again normal proofs of closed Σ sequents. $\square$

Corollary 3. *Let d be a proof of a sentence $\exists x \; A(x)$ in normal form. Then d is of the form*

$$\frac{A(t_1), A(t_2), \dots, A(t_{k-1}), A(t_k), \exists x \ A(x)}{A(t_1), A(t_2), \dots, A(t_{k-1}), \exists x \ A(x)}$$

$$\vdots$$

$$\frac{A(t_1), A(t_2), \exists x \ A(x)}{\dfrac{A(t_1), \exists x \ A(x)}{\exists x \ A(x)}}$$

for some $k \geq 1$. *That is, d consists of an axiom* $\{A(t_1), A(t_2), \dots, A(t_k)\}$ *in which each term t_i is closed, with or without the formula* $\exists x \ A(x)$; *followed by k applications of the* Exists *rule.*

Lemma 8. *Let $\hat{d}$ be obtained by combining the proof above with a cut on the hypothesis* $\forall x \ \bar{A}(x)$. *Then $F_{\hat{d}}$ reduces to $\hat{n}$, where*

$$\hat{n} = \begin{cases} n_{t_1} \ \text{if } A(n_{t_1}) \\ n_{t_2} \ \text{if } \bar{A}(n_{t_1}) \ \text{and } A(n_{t_2}) \\ n_{t_3} \ \text{if } \bar{A}(n_{t_1}), \ \bar{A}(n_{t_2}), \ \text{and } A(n_{t_3}) \\ \vdots \\ n_{t_k} \ \text{otherwise} \end{cases}$$

Proof. Note that if d is a proof of whose last inference is an application of the *Exists* rule to a proof d' of $\Gamma, A(t)$, then F_d is given by

$$\text{Cut}_{\bar{A}(t)}(F_{d'}, \alpha_{\forall x \ \bar{A}(x)}(t)).$$

This reduces to $\alpha_{\forall x \ \bar{A}(x)}(t)$ if $A(t)$ holds, and $F_{d'}$ otherwise. Cutting with the hypothesis involves replacing $\alpha_{\forall x \ \bar{A}(x)}$ by the identity function. The result follows easily by induction on k. □

From Corollary 2 we obtain

Theorem 6. *Let d be a proof of a Σ_1 sentence $\exists x \ A(x)$ in PA. Let n_1 be the value obtained by extracting a realizing term as in Section 3 and normalizing it. Let n_2 be any value obtained by normalizing the proof using the reductions above and then extracting a witness as in Lemma 8. Then $n_1 = n_2$.*

6 Finitary cut elimination

Following Gentzen's approach [11, 25], one obtains an ordinal analysis of *PA* by assigning ordinals to proofs, and then defining a procedure which transforms any proof containing at least one cut to another proof with a smaller ordinal. Iterating this procedure is then guaranteed to yield a cut-free proof. In this section I will

define such an assignment of ordinals, together with an iterative procedure that relies on the reductions defined in the previous section. The ordinal assignment and reduction procedure are essentially an adaptation of Gentzen's (but see the discussion in Section 8).

The ordinal ε_0 is defined to be the limit of $\omega_0, \omega_1, \omega_2, \ldots$, where $\omega_0 = 1$ and $\omega_{n+1} = \omega^{\omega_n}$. As usual, 2^α denotes ordinal exponentiation to the base 2, and 2_n^α denotes the iterated exponential defined by $2_0^\alpha = \alpha$ and $2_{n+1}^\alpha = 2^{2_n^\alpha}$. Any ordinal α can be written uniquely in the form $\sum_{i<k} 2^{\alpha_i}$, where $\alpha_k > \alpha_{k-1} > \ldots > \alpha_0$. If α is of this form and, similarly, β is equal to $\sum_{i<l} 2^{\beta_i}$, then the symmetric sum of α and β, denoted $\alpha\#\beta$, is given by $\sum_{i<k+l} 2^{\gamma_i}$, where the sequence $\langle \gamma_i \rangle$ lists the elements of $\langle \alpha_i \rangle$ and $\langle \beta_i \rangle$ in decreasing order.

The essential property of the symmetric sum is that it is monotone in both arguments, and the essential property of base 2 exponentiation is that whenever α_1 and α_2 are less than α, $2^{\alpha_1}\#2^{\alpha_2}$ is less than 2^α. In fact, since $\alpha_1, \alpha_2 < \alpha$ implies $2_{l-1}^{\alpha_1}, 2_{l-1}^{\alpha_2} < 2_{l-1}^\alpha$, we have that $2_l^{\alpha_1}\#2_l^{\alpha_2} < 2_l^\alpha$ for every l greater than or equal to one.

Define the rank of a formula to be the number of logical connectives it contains; that is, $\mathrm{rank}(B) = \mathrm{rank}(\bar{B}) = 0$, $\mathrm{rank}(\exists x\ \varphi) = \mathrm{rank}(\forall x\ \varphi) = \mathrm{rank}(\varphi)+1$, and $\mathrm{rank}(\varphi \vee \psi) = \mathrm{rank}(\varphi \wedge \psi) = \mathrm{rank}(\varphi) + \mathrm{rank}(\psi) + 1$. The rank of a *Cut* inference is defined to be the rank of its cut formulae, and the rank of an *Induction* inference is the rank of the induction formula $\varphi(y)$.

If we imagine each derivation D to be given by a tree growing upwards from the conclusion to the axioms, the *height* of a node σ in D, denoted $\mathrm{height}(\sigma, D)$, is defined to be the supremum of the ranks of all the cut and induction inferences encountered as one travels along the path from σ down to the root of D. (This terminology is due to Gentzen, and should not be confused with the height of the proof tree.)

I will now assign an ordinal $\mathrm{ord}(D)$ to each *PA*-proof D, by inductively assigning a value $\mathrm{ord}(\sigma, D)$ to each node σ of D and then taking $\mathrm{ord}(D)$ to be the value assigned to the root. Using D_σ to denote the subproof of D rooted at σ, the assignment is defined as follows:

1. If D_σ is a proof consisting of a single axiom, then $\mathrm{ord}(\sigma, D) = 1$.
2. If D_σ is obtained from subproofs D_{τ_i} using an inference of the form *And, Or, For all*, or *Exists*, then $\mathrm{ord}(\sigma, D) = \sup(\{\mathrm{ord}(\tau_i, D)\}) + 1$.
3. If D_σ is obtained by applying a *Cut* rule to subproofs D_τ and D_ρ, let $l = \mathrm{height}(\sigma, D)$ and let $m = \mathrm{height}(\tau, D) = \mathrm{height}(\rho, D)$. Then

$$\mathrm{ord}(\sigma, D) = 2_{m \dot{-} l}^{\mathrm{ord}(\tau, D)\#\mathrm{ord}(\rho, D)}$$

   where "$\dot{-}$" denotes truncated subtraction.
4. If D_σ is obtained by applying an *Induction* rule to subproofs D_τ and D_ρ, then

$$\mathrm{ord}(\sigma, D) = 2_{m \dot{-} l}^{\mathrm{ord}(\tau, D)\#\mathrm{ord}(\rho, D)\cdot\omega}$$

   where l and m are as above.

One can extend this to PA^+ proofs by simply ignoring the inversion rules:

5. If D_σ is obtained by applying an *Inversion* inference to D_τ, set $\text{ord}(\sigma, D)$
$= \text{ord}(\tau, D)$.

In the end, $\text{ord}(\sigma, D)$ depends only on the subproof D_σ rooted at σ and the height of σ in D. In other words, one can define another function $o(d, l)$ such that for any proof d and node σ,

$$\text{ord}(\sigma, D) = o(D_\sigma, \text{height}(\sigma, D)).$$

In particular, $\text{ord}(D) = o(D, 0)$. The function $o(d, l)$ is defined for every l simultaneously, by the following clauses:

1. If d is an axiom, $o(d, l) = 1$.
2. If d is obtained from subproofs d_i using an rule of the form *And, Or, For all,* or *Exists*, $o(d, l) = \sup(\{o(d_i, l)\}) + 1$.
3. If d is obtained by applying a *Cut* of rank m to subproofs d_0 and d_1, then

$$o(d, l) = 2^{o(d_0, \max(m,l)) \# o(d_1, \max(m,l))}_{m \dotminus l}.$$

4. If d is obtained by applying an *Induction* of rank m to subproofs d_0 and d_1, then

$$o(d, l) = 2^{o(d_0, \max(m,l)) \# \text{ord}(d_1, \max(m,l)) \cdot \omega}_{m \dotminus l}.$$

5. If d is obtained by applying an *Inversion* to d_0, $o(d, l) = o(d_0, l)$.

The reduction procedure for PA-proofs can now be described in rough terms, as follows:

1. Keep applying reductions of the form 2a, until there are no more movable cuts.
2. Keep applying reductions of the form 3, until there are no unnecessary free variables.
3. Apply a reduction of the form 1 to a suitable principal cut, and then apply certain reductions of the form 2b to guarantee that the ordinal of the resulting proof decreases.
4. If the previous step introduced an *Inversion* inference, apply reductions of the form 4 and 5 to remove it.

The following lemmata spell out the details. The first shows that the ordinal of a proof D decreases if we apply an appropriate reduction to a subproof D_σ.

Lemma 9. *Let D be any proof in PA^+ and let σ be a node of D. Let D' be a proof obtained by replacing D_σ with another proof, D'_σ, of the same sequent. If $\text{ord}(\sigma, D') = \text{ord}(\sigma, D)$ then $\text{ord}(D') = \text{ord}(D)$, and if $\text{ord}(\sigma, D') < \text{ord}(\sigma, D)$ then $\text{ord}(D') < \text{ord}(D)$.*

Proof. By induction on the length of the path from σ to the root of D. $\qquad\square$

The next three lemmata allow us to carry out the first two steps of the procedure described above, eliminating movable cuts and unnecessary free variables.

Lemma 10. *Let d be a PA-proof in which there are no movable cuts except at the last inference. Then there is a PA-proof d' of the same sequent, such that d' has no movable cuts and for every l, $o(d',l) \leq \mathrm{ord}(d,l)$.*

Proof. By induction on the height of the proof tree. If the active sequent of the cut is the conclusion of subproofs d_i, and the non-active sequent of the cut is the conclusion of subproof $\hat{d}$, apply reduction 2a to pass the cut through this inference, and then inductively apply the lemma to any subproof in which the cut is still movable. Then we have

$$\begin{aligned}
o(d',l) &\leq \sup(\{o(d_i,l)\#o(\hat{d},l)\}) + 1 \\
&= \sup(\{o(d_i,l)+1\})\#o(\hat{d},l) \\
&= o(d,l). \qquad\square
\end{aligned}$$

Lemma 11. *Let d be any PA-proof. Then there is a PA-proof d' of the same sequent, such that d' has no movable cuts and for every l, $o(d,l) \leq o(d',l)$.*
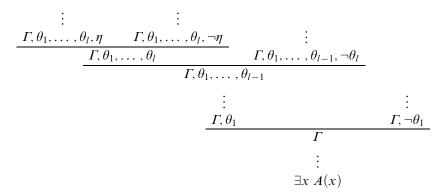
Proof. By induction on d; keep applying Lemma 10. $\qquad\square$

Say that a proof d' "has no unnecessary free variables" if reduction 3 cannot be applied to d'.

Lemma 12. *Let d be any PA-proof. Then there is a PA-proof d' of the same sequent, such that d' has no unnecessary free variables and for every l, $o(d,l) \leq \mathrm{ord}(d',l)$. Furthermore, if d has no movable cuts, then neither does d'.*

Proof. Inductively apply reduction 3 to the last inference of d, if necessary, and then the immediate subproofs. $\qquad\square$

Suppose that D is a *PA* proof of a sentence $\exists x \ A(x)$, and that D has no movable cuts or unnecessary free variables. If D is cut-free, we are done. Otherwise, if the last inference of D is not a cut, then it can only be an instance of the *Exists* rule. In the latter case, if the preceeding inference is also not a cut, then it is again an instance of the *Exists* rule, and so on. In other words, for some j, D ends with j applications of the *Exists* rule preceded by a cut. Now, if this cut is not principal, then its active sequent is again the conclusion of a cut; and again this argument can be iterated. In short, D must have the following form:

$$\frac{\begin{array}{c}\vdots \\ \Gamma, \theta_1, \ldots, \theta_l, \eta\end{array} \qquad \begin{array}{c}\vdots \\ \Gamma, \theta_1, \ldots, \theta_l, \neg\eta\end{array}}{\Gamma, \theta_1, \ldots, \theta_l}$$

In this derivation every sequent is closed, Γ is of the form

$$\{A(t_1), \ldots, A(t_j), \exists x\ A(x)\}$$

(or, possibly, of this form but without the formula $\exists x\ A(x)$), the cut on η is principal, and each θ_i is closed and either atomic, of the form $\alpha \vee \beta$, or of the form $\exists x\ \alpha$.
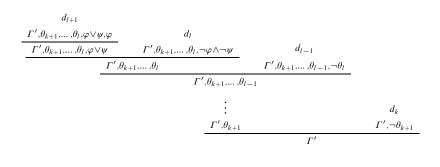
Lemma 13. *Let D be a PA-proof of the above form. Then there is a PA$^+$ proof D′ of the same sequent, such that* $\operatorname{ord}(D')$ *is less than* $\operatorname{ord}(D)$.

Proof. If η is atomic or negation atomic, apply a reduction of the form 1a. It is not difficult to show that the resulting ordinal decreases.

Suppose η is of the form $\varphi \vee \psi$ or $\varphi \wedge \psi$. Without loss of generality, we can assume the former, since otherwise we can consider $\neg\eta$ instead.

Note that $\Gamma, \theta_1, \ldots, \theta_l, \eta$ and $\Gamma, \theta_1, \ldots, \theta_l, \neg\eta$ have the same height; call this height n. Note also that the height of Γ is 0, and the heights of the sequents $\Gamma, \theta_1, \ldots, \theta_i$ are nondecreasing in i. Choose k such that $\Gamma, \theta_1, \ldots, \theta_k$ is the first sequent below $\Gamma, \theta_1, \ldots, \theta_l, \eta$ whose height is strictly less than n; if this sequent is Γ, take k to be 0.

Writing Γ' for $\Gamma, \theta_1, \ldots, \theta_k$, we have the following proof d as a subproof of D:

In this diagram the height of Γ' in D is some value m that is strictly less than n, and the height of every other sequent shown is n.

Let e_0 be the proof of Γ', φ obtained by omitting the top *Or* inference, and cutting d_{l+1} with $d_l, \dots, d_k$, as follows:

$$
\begin{array}{c}
\dfrac{
\dfrac{
\dfrac{d_{l+1}}{\Gamma',\theta_{k+1},\dots,\theta_l,\varphi\vee\psi,\varphi} \quad \dfrac{d_l}{\Gamma',\theta_{k+1},\dots,\theta_l,\neg\varphi\wedge\neg\psi}
}{
\Gamma',\theta_{k+1},\dots,\theta_l,\varphi
} \quad \dfrac{d_{l-1}}{\Gamma',\theta_{k+1},\dots,\theta_{l-1},\neg\theta_l}
}{
\begin{array}{c}
\Gamma',\theta_{k+1},\dots,\theta_{l-1},\varphi \\
\vdots \\
\dfrac{\Gamma',\theta_{k+1},\varphi \qquad\qquad \dfrac{d_k}{\Gamma',\neg\theta_{k+1}}}{\Gamma',\varphi}
\end{array}
}
\end{array}
$$

Let e_1 be the proof of $\Gamma', \neg\varphi$ obtained by inverting d_l, and then cutting it with $d_{l-1}, \dots, d_k$, as follows:

$$
\dfrac{
\dfrac{
\dfrac{d_l}{\Gamma',\theta_{k+1},\dots,\theta_l,\neg\varphi\wedge\neg\psi}}{\Gamma',\theta_{k+1},\dots,\theta_l,\neg\varphi} \quad \dfrac{d_{l-1}}{\Gamma',\theta_{k+1},\dots,\theta_{l-1},\neg\theta_l}
}{
\begin{array}{c}
\Gamma',\theta_{k+1},\dots,\theta_{l-1},\neg\varphi \\
\vdots \\
\dfrac{\Gamma',\theta_{k+1},\neg\varphi \qquad\qquad \dfrac{d_k}{\Gamma',\neg\theta_{k+1}}}{\Gamma',\neg\varphi}
\end{array}
}
$$

Finally let d' be the proof obtained by cutting e_0 with e_1:

$$
\dfrac{\dfrac{e_0}{\Gamma',\varphi} \qquad \dfrac{e_1}{\Gamma',\neg\varphi}}{\Gamma'}
$$

Note that d' is exactly the proof obtained from d by applying reduction 1b to the principal cut on $\varphi \vee \theta$, and then applying reduction 2b l times to bring each of the cuts on $\theta_l, \dots, \theta_{k+1}$ through the resulting cut on φ.

Let D' be the proof that results from replacing d by d' in D. Since the height of Γ' is m in both D and D', I only need to show that $\mathrm{o}(d', m)$ is strictly less than $\mathrm{o}(d, m)$.

Since the height of every sequent in d other than the bottom one is n, we have

$$
\mathrm{o}(d, m) = 2^{(\sum_{i=k}^{l+2} \mathrm{o}(d_i, n)) + 1}_{n-m}
$$

where Σ denotes a summation using the symmetric sum. Notice that the "+1" comes from the application of the *Or* rule. Let p be the height of Γ', φ in D'. Since the rank of φ is strictly less than n, we have $m \leq p < n$. Calculating the ordinal assigned to d', we find that

$$\mathrm{o}(d',m) = 2^{2^{\sum_{i=k}^{l+2} \mathrm{o}(d_i,n)}}_{\substack{n-p \\ p-m}} \# 2^{\sum_{i=0}^{l+1} \mathrm{o}(d_i,n)}_{n-p} .$$

Since $n - p$ greater than or equal to 1 we have

$$2^{\sum_{i=k}^{l+2} \mathrm{o}(d_i,n)}_{n-p} \# 2^{\sum_{i=0}^{l+1} \mathrm{o}(d_i,n)}_{n-p} < 2^{(\sum_{i=k}^{l+2} \mathrm{o}(d_i,n))+1}_{n-m}$$

The desired inequality comes from applying $2^{\cdot}_{p-m}$ to both sides.

This completes the case in which η is of the form $\varphi \vee \psi$ or $\varphi \wedge \psi$. The case in which η is of the form $\exists x\ \varphi$ or $\forall x\ \varphi$ is handled similarly. $\square$

Finally, we have

Lemma 14. *Let d be a PA^+-proof with no inversion rules except at the last inference. Then there is a PA-proof d' of the same sequent, such that for every l, $\mathrm{o}(d',l) \leq \mathrm{o}(d,l)$.*

Proof. An easy induction on d; keep applying reduction 4 and 5. $\square$

We now have the desired cut-elimination procedure: given a proof D that is not cut-free, apply Lemma 11 to eliminate movable cuts, apply Lemma 12 to remove unnecessary free variables, apply Lemma 13 to reduce one principal cut and lower the ordinal rank, and, finally, apply Lemma 14 to remove the inversion that the previous step has introduced.

With a reasonable arithmetization of proofs and ordinal notations, steps 1–4 above are clearly primitive recursive. This yields the usual results of an ordinal analysis:

Theorem 7. *The proof-theoretic ordinal of PA is bounded by ε_0, in all of the following senses:*

1. *Every provable function of PA is $<\varepsilon_0$-recursive.*
2. *Primitive recursive arithmetic, together with the assertion that every descending sequence of notations below ε_0 terminates, proves the 1-consistency of PA.*
3. *Let $PA(f)$ denote the theory obtained by adding a new function symbol to PA and allowing it to appear in induction inferences. Suppose $\prec$ is any primitive recursive relation that $PA(f)$ proves to be a well-ordering, in the sense that it proves $\exists m\ (f(m') \not\prec f(m))$. Then the order type of $\prec$, in the standard model, is less than ε_0.*

Proof. The first two are standard; details and definitions can be found in, say, [4, 22]. To obtain the third clause, suppose that PA proves that $\prec$ is well-ordered in the above sense. From the procedure above we can obtain a $<\varepsilon_0$-recursive functional F such that for every function f,

$$f(F(f)') \not\prec f(F(f)).$$

If the order-type of $\prec$ is greater than ε_0, then there is an isomorphism g of notations below ε_0 with an initial segment of $\prec$. We can use this to obtain a $<\varepsilon_0$ recursive functional $\hat{F}(g, f)$ such that for every f,

$$f(\hat{F}(g, f)') \not\prec f(\hat{F}(g, f)),$$

where "$<$" refers to the ordering on ε_0-notations. Using the normal-form theorems of [6], this can be used to obtain a universal function for ε_0-computable functions relativized to g. But now an easy diagonalization yields a contradiction. □

Of course, corresponding "lower bounds" can be obtained by developing a theory of ordinals below ε_0 within PA.

As a corollary to Theorem 6 we obtain

Theorem 8. *Let d be a proof of a Σ_1 sentence $\exists x\ A(x)$ in PA. Let n_1 be the value obtained by extracting a realizing term as in Section 3 and normalizing it. Let n_2 be the witness obtained from the finitary cut-elimination procedure described above. Then $n_1 = n_2$.*

Wilfried Buchholz has pointed out to me that one can extract additional information from the proof above. First, using modifications similar to the ones described in [25, pp. 116–117], one can obtain sharp ordinal bounds on the strength of the fragments of arithmetic $I\Sigma_k$. More specifically, one has to alter the induction rule so that it reads "from $\Gamma, \varphi(0)$ and $\Gamma, \neg\varphi(x), \varphi(x')$ conclude $\Gamma, \varphi(t)$ for arbitrary terms t." Then one has to augment *For all* inversion to accomodate arbitrary terms t, and modify the reduction rules accordingly. Given a proof d in $I\Sigma_k$, one carries out a finitary "partial" cut elimination, so that in any cut one of the cut formulae is the principal conclusion of an instance of the induction rule. The one uses the ordinal assignment above, but redefining the rank of a formula to be number of logical connectives in the formula *minus one*. The treatment of cuts that previously had rank one has to be suitably modified; for example, given an inference

$$
\cfrac{\cfrac{\cfrac{d_0}{\Gamma, \exists x\ B(x), B(t)}}{\Gamma, \exists x\ B(x)} \qquad \cfrac{d_1}{\Gamma, \forall x\ \bar{B}(x)}}{\Gamma}
$$

in which $B(t)$ is closed and atomic, one cuts d_0 against d_1 and then applies atomic inversion, if $B(t)$ is false; or one applies two inversions on d_1, if $B(t)$ is true. When all is said and done, we have

Theorem 9. *The proof-theoretic ordinal of $I\Sigma_k$ is ω_{k+1}.*

A second observation is that with a more liberal assignment of ordinals one can obtain a stronger normalization theorem. Let $\varphi_0, \varphi_1, \varphi_2, \ldots$ denote the first

ω functions in the Veblen hierarchy (see [22]), so that $\varphi_0(\alpha) = \omega^\alpha$ and φ_{n+1} enumerates the fixed points of φ_n. Modify the assignment of ordinals above by replacing 3 and 4, as follows:

3. If D_σ is obtained by applying a *Cut* of rank m to D_τ and D_ρ, then

$$\mathrm{ord}(\sigma, D) = \varphi_m(\mathrm{ord}(\tau, D)\#\mathrm{ord}(\rho, D)).$$

4. If D_σ is obtained by applying an *Induction* rule of rank m to subproofs D_τ and D_ρ, then

$$\mathrm{ord}(\sigma, D) = \varphi_m(\mathrm{ord}(\tau, D)\#\mathrm{ord}(\rho, D) \cdot \omega).$$

Now if one applies an arbitrary reduction from Section 5 that is not of the form 2b, the ordinal assigned to the proof decreases. In the case of a principal *For all* inversion of the form 4d, if $o(d_0) = \alpha$, $o(d_1) = \beta$, and φ has rank m, the initial proof has ordinal $\varphi_m(\alpha\#\beta \cdot \omega)$, and the final proof has ordinal

$$\varphi_m(\beta\#\varphi_m(\beta\# \ldots \varphi_m(\beta\#\alpha))),$$

where there are n appearances of φ_m. Using the fact that in general $\varphi_m(\gamma)\#\delta \le \varphi_m(\gamma\#\delta)$, the last term is bounded by

$$\varphi_m^n(\beta\# \ldots \#\beta\#\alpha) \le \varphi_m(\beta\# \ldots \beta\#\alpha + 1)$$
$$< \varphi_m(\alpha\#\beta \cdot \omega),$$

as required. As a result, we have

Theorem 10. *Without 2b, the set of reductions in Section 5 is strongly normalizing.*

I suspect that this theorem remains true if one allows reductions of the form 2b, with the restriction that the lower cut must have a strictly higher rank than the one above.

7 Infinitary cut elimination

Schütte's introduction of infinitary logic to the field of ordinal analysis marked a significant step in the development of the subject: recast in the infinitary setting, cut-elimination arguments often become clearer and more natural. In this section I will use a technique due to Mints [16] to analyze one such argument, and, in particular, to extract sequences of finitary reductions from an infinitary cut-elimination procedure. By Theorem 6, this implies that witnesses extracted from proofs of Σ_1 sentences using the infinitary methods also agree with the ones extracted by classical realizability.

Let us convert PA to an infinitary proof system PA_∞ as follows:

1. Replace the *For all* and *Induction* rules by the single infinitary ω rule:
   From $\Gamma, \varphi(n)$ for each numeral n, conclude $\Gamma, \forall x\ \varphi(x)$.
   Here all sequents are assumed to consist of *closed* formulae only.
2. Restrict the other axioms and rules so that the hypotheses and conclusion are required to be closed sequents as well.

The infinitary cut-elimination procedure that I am about to describe is by now standard, and I will only sketch the details. The presentation differs from that of [24] in the form of the ω rule, and in the fact that all sequents of our PA_∞ are closed; it differs from that of [4] in that here we have no need for the "repetition" rule. The presentation differs from the ones in both sources just mentioned in that here the inversion and reduction lemmata are phrased in terms of removing a final inversion or cut rule; but this difference is inessential.

I will use PA_∞^+ to denote PA_∞ with the additional inversion rules; but, as in the finitary cut-elimination procedure, instances of these rules will be removed as soon as they appear. Also as in the finitary case, with each of the lemmata below the corresponding proof provides an implicit (infinitary) procedure for obtaining d' from d.

Note that in this section the word "height" refers to the height of the proof tree, and not the function $\text{height}(\sigma, D)$ of Section 6.

Lemma 15 (Embedding). *If d is a proof of a closed sequent in PA, there is a proof d' of the same sequent in PA_∞.*

Proof. The proof is by a straightforward induction on d. Instances of the *For all* rule applied to a proof d_0 are replaced by an ω inference applied to proofs corresponding to $d_0[n/y]$. Instances of the *Induction* rule applied to proofs d_0 and d_1 are replaced by an ω rule applied to proofs corresponding to substitutions and cuts on d_0 and d_1, as in reduction 4d of Section 6. □

Lemma 16 (Evaluation). *If d is a proof of a sequent $\Gamma(t)$ in PA_∞, then there is a proof d' of $\Gamma(n_t)$ in PA_∞, whose height is less than or equal to that of d.*

Lemma 17 (Inversion). *Suppose d is a proof in PA_∞^+ whose last inference is an inversion, and there are no other inversions in d. Then there is a proof d' of the same sequent in PA_∞.*

Proof. By induction on the height of d. If the last inversion is not principal, one moves the inversion to the immediate subproofs and applies the lemma inductively. If the last inversion is applied to an axiom or an instance of the *And* rule, one applies a reduction of the form 4a or 4b, and applies the lemma inductively. If the last inversion is applied to the conclusion of an ω rule, one moves the inversion to the appropriate hypothesis and again applies the lemma inductively. □

Lemma 18 (Reduction). *Suppose d is a proof in PA_∞ whose last inference is a cut of rank n, and suppose every other cut in d has rank less than n. Then there is a proof d' in PA_∞ in which every cut has rank less than n.*

Proof. By induction on the sum of the heights of the subproofs rooted at d. If the cut is nonprincipal, apply a reduction of the form 2 (or the analogue for the ω rule). If the cut is principal, apply a reduction of the form 1. Then apply the reduction lemma inductively, as well as the inversion lemma if necessary. $\square$

Define the cut rank of a proof to be one more than the supremum of the ranks of its cut inferences.

Lemma 19 (Elimination). *For any $n \geq 0$, if d is a proof of cut rank $n + 1$ in PA_∞, then there is a proof d' of cut rank n.*

Proof. By induction on the height of d. If the last inference is a cut of rank n, apply the lemma inductively to its immediate subproofs, and then apply the reduction lemma. $\square$

Theorem 11. *Let d be a proof of a closed sequent in PA. Then there is a cut-free proof d' of the same sequent in PA_∞.*

Proof. Use the embedding lemma to translate d into a proof in PA_∞, let n be the cut rank of the resulting proof, and then apply the elimination lemma n times. $\square$

In fact, the height of d' can be shown to be less than ε_0. This fact provides an alternate route to an ordinal analysis of PA, but will not be needed below.

Now suppose that in the statement of Theorem 11, d is a proof of a Σ_1 sentence. Then the resulting d' is of the form described by Lemma 7, which is to say, in this case d' is finite; so the net result of the detour through the infinite is the reduction of d to a cut-free PA-proof d'. Our goal now is to keep track of enough information during this detour so that, in the end, we also have a finite reduction sequence from d to d'.

To that end, if d is a proof in PA_∞^+, define an *enrichment* of d to be an assignment of a sequence of PA^+ proofs

$$\text{start}(d_\sigma) \rightsquigarrow_1 \ldots \rightsquigarrow_1 \text{end}(d_\sigma)$$

to every subproof d_σ of d, satisfying the following conditions for each σ:

1. $\text{start}(d_\sigma)$ and $\text{end}(d_\sigma)$ are proofs of the same sequent as d_σ.
2. If the last inference of d_σ is an instance of anything other than the ω rule, and is applied to subproofs $\ldots d_{\tau_i} \ldots$, then $\text{end}(d_\sigma)$ consists of the same rule applied to $\ldots \text{start}(d_{\tau_i}) \ldots$.
3. If the last inference of d_σ is an instance of the ω rule applied to subproofs $d_{\tau_0}, d_{\tau_1}, d_{\tau_2}, \ldots$, then either
   a) the last inference of $\text{end}(d_\sigma)$ is the *For all* rule applied to a proof $e[y]$, and for each i, $\text{start}(d_{\tau_i}) = e[i/y]$; or
   b) the last inference of $\text{end}(d_\sigma)$ is the *Induction* rule applied to proofs e_0 and $e_1[y]$, and for each i, $\text{start}(d_{\tau_i})$ is the result of cutting e_0 with $e_1[0/y], \ldots, e_1[i-1/y]$.

An *enriched* proof is simply one that is equipped with an enrichment.

Suppose d is an enriched cut-free proof of a closed Σ_1 sentence. By the remarks above d is a finitary *PA* proof, which is to say that, in particular, it contains no instances of the ω rule. The enrichment provides a recipe for reducing $\text{start}(d)$ to d, as follows: first, reduce $\text{start}(d)$ to $\text{end}(d)$; then reduce each of the immediate subproofs $\text{start}(d_{\tau_i})$ of $\text{end}(d)$ to the corresponding proofs $\text{end}(d_{\tau_i})$; then reduce the subproofs of those, and so on. In other words, the following lemma provides the reduction sequence we are looking for.

Lemma 20. *The procedures implicit in the proofs of Lemmata* 15–19 *can be modified, so that the following hold*:

1. *In the embedding lemma, d' is an enriched proof, and* $\text{start}(d') = d$.
2. *In the evaluation, inversion, reduction, and elimination lemmata, if d is an enriched proof, then so is d' and* $\text{start}(d') = \text{start}(d)$.

Proof. Straightforward. Enrichments can be extracted from any "one-step" reduction of an enriched proof in PA_∞, and enrichments obtained by applying the lemmata (inductively) to subproofs can be cobbled together to yield an enrichment for the entire result.   □

This yields

Theorem 12. *Let d be a proof of a Σ_1 sentence $\exists x\ A(x)$ in PA. Let n_1 be the value obtained by extracting a realizing term as in Section 3 and normalizing it. Let n_2 be the witness obtained from the infinitary cut-elimination procedure described above. Then $n_1 = n_2$.*
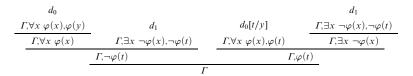
8 Final comments

One should note that the M-translation, as a variation of the double-negation translation, is not efficient on fragments of arithmetic. Although it is known that $I\Sigma_1$ and its intuitionistic version $I\Sigma_1^i$ have the same provably total recursive functions, induction on a Σ_1 sentence translates to induction on the negation of a Π_1 sentence, and applying the Friedman-Dragalin translation yields an instance of induction that is even more complex. Thierry Coquand has recently discovered a method of interpreting $I\Sigma_1$ that avoids this difficulty; it is discussed and extended to additional theories in both [8] and [1].

One should also keep in mind that Theorems 8 and 12, which assert that finitary cut elimination, infinitary cut elimination, and classical realizability produce the same result, depend on both the particular translation of *PA* to *HA* and the particular cut-elimination procedures I have chosen. In contrast, infinitary cut-elimination arguments in [22] and [5] do not use an inversion lemma; instead, cuts are moved upwards until *both* cut formulae are principal in the preceding inferences, and a "symmetric" reduction is applied. For example, when the cut formula is of the form $\forall x\ \varphi(x)$,

$$\frac{\dfrac{d_0}{\Gamma, \forall x\ \varphi(x), \varphi(y)}}{\Gamma, \forall x\ \varphi(x)} \qquad \frac{\dfrac{d_1}{\Gamma, \exists x\ \neg\varphi(x), \neg\varphi(t)}}{\Gamma, \exists x\ \neg\varphi(x)}$$
$$\Gamma$$

reduces to

$$\frac{\dfrac{\dfrac{d_0}{\Gamma, \forall x\ \varphi(x), \varphi(y)}}{\Gamma, \forall x\ \varphi(x)} \quad \dfrac{d_1}{\Gamma, \exists x\ \neg\varphi(x), \neg\varphi(t)}}{\Gamma, \neg\varphi(t)} \qquad \frac{\dfrac{d_0[t/y]}{\Gamma, \forall x\ \varphi(x), \varphi(t)} \quad \dfrac{d_1}{\Gamma, \exists x\ \neg\varphi(x), \neg\varphi(t)}}{\Gamma, \varphi(t)}$$
$$\Gamma$$

These reductions, which are closer in form to the ones used by Gentzen, are *not* compatible with the normalization of the realizing term. Can one find an interesting form of realizability that has this property? It would also be nice to have a better sense of how sensitive Theorems 8 and 12 are to variations in the cut-elimination procedure. To take a simple example: cutting a proof of Γ, φ with a "canonical" proof of $\neg\varphi, \varphi$ can change the witnessing term.

References

1. Jeremy Avigad. Intepreting classical theories in constructive ones. Submitted.
2. Jeremy Avigad and Solomon Feferman. Gödel's functional (Dialectica) interpretation. In Buss [7].
3. Jon Barwise, editor. *The Handbook of Mathematical Logic*. North-Holland, Amsterdam, 1977.
4. Wilfried Buchholz. Notation systems for infinitary derivations. *Archive for Mathematical Logic*, 30:277–296, 1991.
5. Wilfried Buchholz. Explaining the Gentzen-Takeuti reduction steps. Preliminary version, 1997.
6. Samuel R. Buss. The witness function method and provably recursive functions of Peano arithmetic. In D. Westerstahl D. Prawitz, B. Skyrms, editor, *Proceedings of the Ninth International Congress on Logic, Methodology, and Philosophy of Science, Uppsala, Sweden, 1991*, pages 29–68. Elsevier North-Holland, Amsterdam, 1994.
7. Samuel R. Buss, editor. *The Handbook of Proof Theory*. North-Holland, To appear.
8. Thierry Coquand and Martin Hofmann. Applications of a representation theory to subsystems of arithmetic. Preprint, 1998.
9. Harvey M. Friedman. Classically and intuitionistically provable functions. In H. Müller and D. Scott, editors, *Higher Set Theory*, Lecture Notes in Mathematics 669, pages 21–27. Springer-Verlag, Berlin, 1978.
10. Gerhard Gentzen. Untersuchungen über das logische Schliessen. *Mathematische Zeitschrift*, 39:176–210, 405–431, 1935. Translated as "Investigations into Logical Deduction" in [12], pages 68–131.
11. Gerhard Gentzen. Neue Fassung des Widerspruchsfreiheit für die reine Zahlentheorie. *Forschungen zur Logik und zur Grundlegung der exakten Wissenschaften*, New Series 4:19–44, 1938. Translated as "New Version of the Consistency Proof for Elementary Number Theory" in [12], pages 252–286.

12. Gerhard Gentzen. *Collected Works*. North-Holland, Amsterdam, 1969. Edited by M. E. Szabo.
13. Kurt Gödel. Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *Dialectica*, 12:280–287, 1958. Reproduced with English translation in [14], pp. 241–251.
14. Kurt Gödel. *Collected Works*, volume II. Oxford University Press, New York, 1990. Solomon Feferman et al. eds.
15. J. Roger Hindley and Jonathan P. Seldin. *Introduction to Combinators and λ-calculus*. Cambridge University Press, Cambridge, 1986.
16. Grigorii E. Mints. Normalization of finite terms and derivations via infinite ones. [18], pages 73–76.
17. Grigorii E. Mints. On E-theorems. [18], pages 105–115.
18. Grigorii E. Mints. *Selected Papers in Proof Theory*. Bibliopolis / North-Holland, Naples / Amsterdam, 1992.
19. Chetan R. Murthy. An evaluation semantics for classical proofs. In *Proceedings, Sixth Annual IEEE Symposium on Logic in Computer Science*, pages 96–107, Amsterdam, 1991.
20. N. M. Nagornyĭ. On the consistency of classical formal arithmetic. *Russian Acad. Sci. Dokl. Math.*, 48, 1994. See also Mathematical Reviews, 98j:03081.
21. Michel Parigot. Proofs of strong normalization for second order classical natural deduction. *Journal of Symbolic Logic*, 62(4):1461–1479, 1997.
22. Wolfram Pohlers. *Proof Theory: An Introduction*. Lecture Notes in Mathematics #1407. Springer-Verlag, Berlin, 1989.
23. Dag Prawitz. *Natural Deduction: A Proof-theoretical Study*. Almquist and Wiksell, Stockholm, 1965.
24. Helmut Schwichtenberg. Proof theory: Some aspects of cut-elimination. In Barwise [3], pages 867–895.
25. Gaisi Takeuti. *Proof Theory*. North-Holland, Amsterdam, second edition, 1987.
26. A. S. Troelstra. Introductory note to 1958 and 1972. In [14], pages 217–241.
27. A. S. Troelstra. Aspects of constructive mathematics. In Barwise [3], pages 973–1052.
28. A. S. Troelstra. Realizability. In Buss [7].
29. Dirk van Dalen. *Logic and Structure*. Springer, Berlin, third edition, 1997.
30. J. I. Zucker. The correspondence between cut-elimination and normalization I, II. *Annals of Mathematical Logic*, 7:1–156, 1974.