Proof Mining in Subsystems of Analysis

Paulo Oliva

PhD Dissertation

Department of Computer Science
University of Aarhus
Denmark

Proof Mining in Subsystems of Analysis

A Dissertation
Presented to the Faculty of Science
of the University of Aarhus
in Partial Fulfilment of the Requirements for the
PhD Degree

by
Paulo Oliva
July 8, 2003

Abstract

This dissertation studies the use of methods of proof theory in extracting new information from proofs in subsystems of classical analysis. We focus mainly on ineffective proofs, i.e. proofs which make use of ineffective principles ranging from weak König's lemma to full comprehension. The main contributions of the dissertation can be divided into four parts:

(1) A discussion of how monotone functional interpretation provides the right notion of "numerical implication" (cf. [21]) in analysis. We show among other things that monotone functional interpretation naturally creates well-studied moduli when applied to various classes of statements (e.g. uniqueness, convexity, contractivity, continuity and monotonicity) and that the interpretation of implications between those statements corresponds to translations between the different moduli.

(2) A case study in L_1-approximation, in which we analyze Cheney's proof of Jackson's theorem, concerning uniqueness of the best approximation, w.r.t. L_1-norm, of continuous functions $f \in C[0,1]$ by polynomials of bounded degree. The result of our analysis provides the first effective modulus of uniqueness for L_1-approximation. Moreover, using this modulus we give the first complexity analysis of the sequence of best L_1-approximations for polynomial-time computable functions $f \in C[0,1]$.

(3) A comparison between three different forms of bar recursion, in which we show among other things that the type structure of strongly majorizable functionals is a model of modified bar recursion, that modified bar recursion is not S1-S9 computable over the type structure of total continuous functions and that modified bar recursion defines (primitive recursively, in the sense of Kleene) Spector's bar recursion.

(4) An adaptation of functional interpretation to handle ineffective proofs in feasible analysis, which provides the first modular procedure for extracting polynomial-time realizers from ineffective proofs (i.e. proofs involving weak König's lemma) of Π^0_2-theorems in feasible analysis.

Acknowledgements

First of all, I would to express my deep appreciation for the guidance and assistance my supervisor, Ulrich Kohlenbach, has provided me during the last four years. His passion for mathematical logic and his care for details are two of the many things that will certainly have an impact on my future research.

I am very much indebted to Ulrich Berger for making my four-months-visit to Swansea very enjoyable and productive. I am also most grateful to Katia Guimarães, for leading me through the first years of my scientific carrier, and for her constant support afterwards.

Many thanks to Dag Normann and Thierry Coquand for having accepted to evaluate this dissertation. Having them as examiners is certainly a great honor for me.

I would like to thank all my friends at BRICS (Anders, Bartek, Bernd, Branimir, Claus, Dan, Daniele, Federico, Frank, Giuseppe, Jesus, Jiri, Kirill, Mads, Maciej, Mikkel, Oliver, Pawel, Philipp, Riko, Saurabh and Sunil) for the wonderful time I have had in those four year. In particular, I would like to thank Mario Caccamo, for making me feel at home when I first arrived in Denmark four year ago.

I would like to express my appreciation for the excellent working environment at BRICS, University of Aarhus. Special thanks goes to Janne Christensen, Lene Kjeldsteen, Uffe Engberg, Karen Møller, Hanne Jensen and Ingrid Larsen. I am also extremely grateful to Mogens Nielsen and Olivier Danvy, for their encouraging manner and continuous support. I thank the Danish National Research Foundation for funding my PhD research in BRICS.

Last but not least, I would like also to thank my wife, Emanuela, whose love and support were fundamental to me during my PhD, my parents for being my first and best teachers, and my sisters for the support and care they have always had towards me.

Paulo Oliva,
Århus, July 8, 2003.

Contents

Part I

Context

Chapter 1

Introduction

This dissertation studies the use of methods of proof theory in extracting new information from proofs in subsystems of classical analysis. The focus shall be mainly on ineffective proofs, i.e. proofs which make use of ineffective principles such as König's lemma for binary trees, so-called weak König's lemma, and comprehension. The main contributions of the dissertation can be divided into four parts: a discussion of the monotone functional interpretation of numerical implications in analysis, a case study in L_1-approximation, a comparison between three different forms of bar recursion and an adaptation of functional interpretation to handle ineffective proofs in feasible analysis. In this introductory chapter we shall briefly discuss the state of the art in each of these areas and give a general description of the contributions. This work belongs to the general program of proof mining.

1.1 Proof Mining

The program of proof mining originated in ideas of Georg Kreisel who throughout his scientific carrier advocated the application of methods of proof theory, which were initially developed to prove the consistency of formal systems, to core mathematics. What Kreisel noticed was that mathematical proofs in various occasions carry more information than just the truth of their associate theorems, and moreover, that, even when this new information apparently cannot be read-off directly from the given proof, that it can in many cases be extracted in a systematic way with the help of logic.

When faced with such a general program, various questions immediately come up.

(*i*) First of all, which kind of new information can possibly be extracted from a proof? What are the *potential areas of application* of such a theory?

(*ii*) And once we are convinced that such applications exist, there comes the question of which *tools and techniques* to use for doing the extraction. Or which general properties should those tools have? For instance, since proofs are highly modular objects, one might wish the tools to respect modularity, i.e. the whole information to be extracted is likely to be composed of partial information obtained from sub-proofs of the original proof.

(iii) Finally, which is perhaps the most important point for a person who wants to contribute to the project, is the question of how one can *identify proofs* from which one might expect to obtain new information. And moreover, based on the principles used in the proof, which *a-priori knowledge* about the information to be extracted can one have? In other words, can a mathematician or a computer scientist with none or very little knowledge of proof theory contribute to the program?

The aim of this introductory chapter is to discuss those questions, at the same time that the main results of the dissertation are put into perspective.

1.1.1 Areas of Applications

Let us begin by tackling question (i). As one might expect, the new information to be obtained from a proof $\mathcal{P}$ must somehow be related to $\mathcal{P}$ itself and consequently to its associated theorem A. As we shall see, this will be in general a proof $\mathcal{P}'$ of a *stronger* theorem A'.

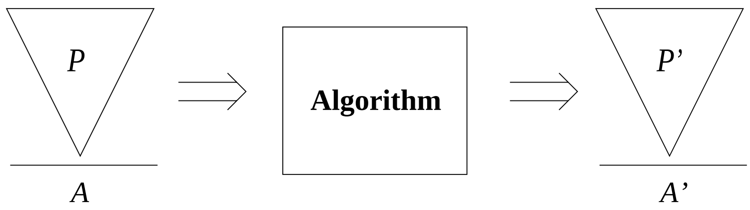


Figure 1.1: The general program of proof mining

For instance if A is a theorem of the form $B \vee C$ then the theorems $A' \equiv B$ and $A' \equiv C$ are both possible strengthenings of A. A strengthening of an existential statement $A \equiv \exists x B(x)$ could be e.g. a theorem $A' \equiv B(c)$ for some fixed object c in the domain of quantification, or simply a collection of possible witnesses $A' \equiv B(c_1) \vee \ldots \vee B(c_n)$. Perhaps the most cited class of examples, usually associated with the label "programs from proofs", is when A has the form $\forall x \exists y B(x, y)$. In this case, a possible strengthening of A would be $A' \equiv \forall x B(x, p(x))$, where p is a program producing for each given x an element $y = p(x)$ satisfying $B(x, y)$.

One might also be interested in finding a *better* proof $\mathcal{P}'$ of the same theorem A. For instance, transforming a proof $\mathcal{P}$ of A which makes use of some principle Δ into a new proof $\mathcal{P}'$ which uses only special cases of Δ, or even avoids Δ altogether, can be viewed as a strengthening of the proof of A. We can even combine both approaches and obtain better proofs of stronger theorems.

In order to give the reader an idea of the potential applications of the program, it is perhaps more pragmatic to list some of the proofs in various branches of mathematics which in the past have been analyzed and in various cases yielded previously unknown stronger versions of the original theorems.

Hilbert's 17th Problem

As the person to launch the program, Kreisel was the first to use techniques of proof theory (an adaptation of Hilbert's ε-substitution method) in order to obtain new information from proofs. His first case study was Artin's solution of Hilbert's 17th problem. Making heavy use of classical logic and König's lemma, Artin proved that every rational function $f \in \mathbb{Q}(x_1, \ldots, x_n)$ which is non-negative everywhere where it is defined can be written as the sum of squares of rational functions $f_1, \ldots, f_n$. Being ineffective, Artin's proof does not provide a procedure for producing $f_1, \ldots, f_n$ out of a given function f. In fact, for many years not even a bound on n nor on the degrees of the functions f_i were known. By analyzing Artin's proof Kreisel obtained the first primitive recursive bounds [40, 126]. For further information on that and other case studies by Kreisel see [47, 113, 114].

Roth's Theorem

Another application of proof theory (based on Herbrand's theorem, so-called "Herbrand analysis") to number theory was carried out by Luckhardt [125]. Roth's theorem states that for every algebraic real number x and positive real number ε the inequality

$$|x - \frac{p}{q}| < \frac{1}{q^{2+\varepsilon}},$$

has *finitely many solutions*, i.e. there are finitely many $q \in \mathbb{N}$ such that $R(q)$ holds, where

$$R(q) :\equiv q > 1 \wedge \exists! p \in \mathbb{Z}\big((p, q) = 1 \wedge |x - \frac{p}{q}| < \frac{1}{q^{2+\varepsilon}}\big).$$

Roth's original proof is also ineffective, it gives bounds neither on the number of solutions nor on the size of the solutions. The first bound on the number of solutions was obtained by Davenport and Roth. Via a proof analysis Luckhardt [125] improved the *exponential bound* b obtained by Davenport and Roth to $\sqrt[4]{b}$. Luckhardt obtained subsequently, by analyzing a different proof of Roth's theorem given by Esnault and Viehweg, the first *polynomial bound* on the number of solutions (see also [24]), which shows that the quality of the information obtained from analyzing a proof depends mostly on the proof itself, and therefore, different proofs might yield significantly numerically different results. The potential application of Herbrand analysis to finiteness theorems was first observed by Kreisel in [117].

Infinite Ramsey Theorem

The fact that a set $A \subseteq \mathbb{N}$ is infinite can be expressed as $\forall x \exists y \geq x A(y)$. A proof analysis applied to such theorems can lead to functions effectively producing subsets of A of arbitrary finite cardinality. This was applied by Bellin [10] to the Infinite Ramsey Theorem (IRT). What he then obtained was a parametrized version of IRT which implies both IRT and the Finite Ramsey Theorem.

Normalization Theorems

A normalization theorem states that every term of some rewriting system has a normal form, i.e. $\forall r \exists s \mathsf{N}(r,s)$, where the predicate $\mathsf{N}(r,s)$ states that s is the normal form of r. Ulrich Berger [13] has applied proof theoretic techniques (Kreisel's modified realizability) in order to obtain, from a proof of strong normalization for the typed λ-calculus plus uniqueness of the long beta-normal form, a normalization algorithm. Even though the original proof is constructive, and therefore has a clear computational meaning, the obtained algorithm has interesting new features, such as a detour through higher types and the use of the compiler evaluation procedure to obtain the normal form (so-called *normalization by evaluation*).

Uniqueness Theorems

Ulrich Kohlenbach observed in [88] that various classes of uniqueness proofs in analysis could be analyzed to yield so-called moduli of uniqueness, a concept which generalizes the well-known notion of strong unicity in approximation theory. In general, from a proof that a function $f : Y \to \mathbb{R}$, Y a Polish space (a complete separable metric space), has at most one root,

$$\forall y_1, y_2 \in Y(\bigwedge_{i=1}^{2} f(y_i) = 0 \to y_1 = y_2),$$

one can in various cases extract a function Φ (depending on the representation of elements y_1 and y_2) satisfying

$$\forall y_1, y_2 \in Y \forall \varepsilon \in \mathbb{Q}^*_+(\bigwedge_{i=1}^{2} |f(y_i)| < \Phi(y_1, y_2, \varepsilon) \to d_Y(y_1, y_2) < \varepsilon),$$

which is a much stronger statement than just the fact that f has at most one root on Y. The interesting fact observed by Kohlenbach is that the functional Φ is always independent of elements ranging over compact spaces. For example, if one proves that a continuous $f : K \to \mathbb{R}$, K compact, has at most one root, then a function Φ (independent of y_1 and y_2) can be extracted from such proof satisfying

$$\forall y_1, y_2 \in K \forall \varepsilon \in \mathbb{Q}^*_+(\bigwedge_{i=1}^{2} |f(x, y_i)| < \Phi(\varepsilon) \to d_K(y_1, y_2) < \varepsilon).$$

This feature was exploited by Kohlenbach [88, 90] in the analysis of various proofs of uniqueness for the best Chebycheff approximation of $f \in C[0,1]$ by elements of Haar spaces, obtaining better moduli of uniqueness than the previously known.

One of the main results of the present dissertation consists of an analysis (coauthored by U. Kohlenbach) of Cheney's proof [33,34] of Jackson's theorem [70], which concerns uniqueness in L_1-approximation. In this case, no effective (in all parameters) modulus of uniqueness was known before. In fact, the modulus obtained led to the first complexity analysis of the problem of L_1-approximation by the author. The interesting thing is that both Cheney's proof of uniqueness and the existence proof make use of the ineffective principle of attainment of the infimum by continuous functions on closed intervals, which in logical terms is equivalent to weak König's lemma. Further below, in Section 1.2.1, the reader can find a more detailed discussion of the use of weak König's lemma in the problem of L_1-approximation. The full analysis of Cheney's proof can be found in Chapter 6. The complexity analysis of the problem of L_1-approximation is given in Chapter 7.

Fixed Point Theorems

Nowadays, the fixed point theory of *nonexpansive mappings* constitutes one of the main areas in nonlinear functional analysis. Different from the cases of contractions (Banach fixed point theorem) or of contractive mappings (Edelstein's fixed point theorem), in the case of nonexpansive mappings fixed points in general do not exist, when they do exist they are not necessarily unique, and even when unique, the usual Banach-Picard iteration ($x_{n+1} = f(x_n)$) will in general not converge to the fixed point. In 1965, Browder, Göhde and Kirk showed independently that a nonexpansive mapping $f : C \to C$ on a convex, closed and bounded subset C of a uniformly convex Banach space $(X, \|\cdot\|)$ has a fixed point. Although fixed points in general do not exist without the condition of uniform convexity on $(X, \|\cdot\|)$, Ishikawa [69] showed that a property, called *asymptotic regularity*, $\|x_n - f(x_n)\| \stackrel{n\to\infty}{\longrightarrow} 0$, holds in arbitrary normed spaces using the so-called Krasnoselski-Mann iteration $(x_n)_{n\in\mathbb{N}}$ of f. By similar techniques to those applied in the uniqueness theorems, described above, Kohlenbach carried out the extraction of *rates of asymptotic regularity* from the proofs of Groetsch [62], Ishikawa [69] and Borwein-Reich-Shafrir [25] theorems (see [100, 101, 103]), obtaining even new qualitative uniformity results. Together with L. Leuştean, this work was generalized to hyperbolic spaces [105]. In Chapter 5 one can find among others a more comprehensive discussion of the applications of proof mining to functional analysis.

Although in some of the case studies listed above the parts of the proof which needed to be analyzed involve only simple mathematical notions, they are all logically quite intricate. This is mainly due to the use of classical logic and non-computational principles such as the attainment of the infimum for continuous functions on closed intervals or the principle of convergence of monotone bounded sequences of real numbers.

Hopefully, the applications illustrated above are enough to convince the reader of the importance of logical tools for guiding the process of proofs analysis. The next section shall be devoted to discuss the second question, namely, which methods can one use for extracting information from proofs?

1.1.2 Tools and Techniques

The techniques used for doing proof mining were, interestingly enough, developed with the quite different goal of proving relative consistency between mathematical systems, which was motivated by Hilbert's consistency program. The first techniques developed in the realm of the program were Herbrand's theorem, which provided a proof of consistency for predicate logic, and the ε-substitution method, which allows quantifiers to be systematically eliminated from proofs at the cost of the addition of new special constants. After successfully applying the ε-substitution method to prove the consistency of open first order arithmetical theories, the natural next step would be to prove, by finitistic means, the consistency of classical first order arithmetic with full induction, also called Peano arithmetic PA. Gödel showed, however, that to be an impossible task in his second incompleteness theorem. According to the theorem, not even Peano arithmetic itself could prove its own consistency. People immediately started looking for "minimal" abstract notions which need to be added to PA in order to prove PA's consistency. Gentzen [54] showed that one such a notion is transfinite induction up to ε_0, namely he proved the consistency of Peano arithmetic via a combination of transfinite induction up to ε_0 with cut elimination. Ackermann [1] then showed that by means of this form of transfinite induction ε-substitution also applies for Peano arithmetic.

A different approach was taken by Gödel [58] around 1943 (but published only in 1958). Instead of transfinite induction, Gödel extended primitive recursive arithmetic to all finite types. The abstract notion in this case is the scheme of primitive recursion for each finite type, as had been anticipated by Hilbert [64] already in 1926. The resulting system is normally called Gödel's T. He then showed that the consistency of T implies the consistency of intuitionistic first order arithmetic, and via the negative translation[1] also of classical arithmetic. That was done by means of a proof interpretation translating formulas and proofs of intuitionistic arithmetic (so-called Heyting arithmetic HA) and PA into formulas and proofs of the calculus T. This interpretation is the so-called *functional interpretation* (also known as *Dialectica interpretation* or *D-interpretation*). In 1961, the functional interpretation of classical arithmetic was extended to classical analysis by Spector [154], who gave an interpretation of comprehension via a new scheme of recursion called *bar recursion* (cf. Section 1.2.2).

In [77] Kleene developed an interpretation, so-called *realizability*, based on the Brouwer-Heyting-Kolmogorov interpretation of constructive logic, in order to prove results about intuitionistic arithmetic, such as the disjunction and the numerical existence properties. A modification of Kleene's realizability, so-called *modified realizability*, was defined and used by Kreisel in [115]. As functional interpretation, modified realizability is based on a language of functionals of finite type. In fact, modified realizability can in a precise sense be viewed as a simpler form of functional interpretation.

Therefore, the situation around the middle of the 20th century was that a

[1]The negative translation was independently discovered by Kolmogorov [108], Gödel [57] and Gentzen (see Section 3.1).

great number of proof-theoretic techniques had been developed, among others ε-substitution, cut elimination, realizability, negative translation and functional interpretation, mainly with the goal of proving relative consistency. But at that time the question of consistency of mathematics was of limited interest to the mathematical community. Therefore, it was necessary for proof theory to have a new perspective.

Kreisel's Unwinding of Proofs

In [114] Kreisel asked for a shift of emphasis in proof-theoretic research and launched a program, which later came to be called *unwinding of proofs* (cf. [47, 126]). In his own words:

> *"There is a different general program which does not seem to suffer the defects of [Hilbert's] consistency program: To determine the constructive (recursive) content or the constructive equivalent of the non-constructive concepts and theorems used in mathematics, particularly in arithmetic and analysis."*

The idea is that, instead of studying abstract general properties of formal mathematical system such as consistency, one should focus on particular concrete proofs. For instance, as we saw above, Hilbert's 17th problem states that every rational function f having certain properties can be written as the sum of squares of rational functions $f_1, \ldots, f_n$. Artin's proof of Hilbert's 17th problem, however, at first sight does not provide any procedure which given f produces some information about the functions $f_1, \ldots, f_n$. An unwinding of Artin's proof in order to obtain such procedure would be an interesting mathematical result, which would certainly draw the attention of the mathematical community to the field of mathematical logic and proof theory. That was Kreisel's point.

Already in 1952, Kreisel [111] had described general properties of the tools, which he called *interpretations*, to be used for analysing proofs. An interpretation of a theory $\mathcal{T}$ is a function $\mathcal{I}$ associating formulas A in the language of $\mathcal{T}$ to sequences of formulas $(A_n)_{n\in\mathbb{N}}$, satisfying the following four conditions[2]

(α) $(A_n)_{n\in\mathbb{N}}$ are formulas in the language of $\mathcal{T}$ which are considered *empty of information*,

(β) if A is proved in $\mathcal{T}$, from the proof we find a numeral n (which we shall call a realizer of A) such that the formula A_n is provable,

(γ) if $\neg A$ is proved in $\mathcal{T}$, for each n we find a substitution π for the (individual and function) variables of A_n such that $\neg A_n$ is provable,

[2] In the original definition of Kreisel, he requires $(A_n)_{n\in\mathbb{N}}$ to be a sequence of *quantifier-free formulas*. We relax this condition here and leave it open what "empty of information" might mean. Consequently, provability is used as a way of verifying a given formula. In the subsequent paper [114], Kreisel also changed significantly the conditions above leading to a different notion of interpretation. For a discussion on the relation between the two definitions of Kreisel see [97].

(δ) if B is proved from assumption A in $\mathcal{T}$ then there exists a function $g : \mathbb{N} \to \mathbb{N}$ such that $B_{g(n)}$ is provable whenever A_n is provable.

Kreisel was assuming here an enumeration of some class of functionals $\mathcal{F}$ so that n above ranges over the codes of such functionals. Therefore, the goal of an interpretation is to separate the content of an arbitrary formula A into two parts: the "computational content" of A, represented by n (the code of a functional in $\mathcal{F}$), and its "computationally trivial part", represented by the sequence $(A_n)_{n\in\mathbb{N}}$. The two conditions (β) and (γ) make sure that the interpretation of a formula A remains strongly related to A itself. The condition (δ) requires an interpretation to have a certain modular behaviour, i.e. since B is proved, by a simple application of modus ponens, from $A \wedge (A \to B)$, the last condition implies the existence of a set of functions g, for each modus ponens in the proof, taking hypothetical realizers n_1 and n_2 for A and $A \to B$ into a potential realizer n_3 for B, such that B_{n_3} is provable whenever A_{n_1} and $(A \to B)_{n_2}$ are. This implies that, given a proof of a theorem A, the realizer n for the interpretation of A shall be obtained by recursion on the proof tree, assuming that realizers for the interpretations of the axioms are given.

The first example of an interpretation given by Kreisel [111] was Herbrand's theorem, but he also developed his own interpretation [111, 112], the so-called *no-counterexample interpretation*, or *n.c.i.* for short. Notice that, according to Kreisel's definition, an interpretation $\mathcal{I}$ only says *what* kind of information one should expect from a given theorem A, but it does not say *how* that information should be obtained, i.e. how to obtain the number n in condition (β), the substitution π in condition (γ) and the function g in (δ). Those procedures will in general be provided by a constructive proof that a given $\mathcal{I}$ satisfies conditions (α)–(δ). In the case of the n.c.i., Kreisel gave a constructive proof based on Hilbert's ε-substitution method. For a given enumeration $\mathcal{F}$ of a certain class of functionals, we shall say that a procedure $\mathcal{A}$ is an *algorithm for the interpretation* $\mathcal{I}$, if $\mathcal{A}$ provides constructions which, from given proofs of A, $\neg A$ or a proof of B from A, produce n, π or g, respectively, according to conditions (α)–(δ). In order to discuss properties of different combinations of interpretations and their associated constructions Kreisel's definition of an interpretation shall be extended in the following way.

Definition 1.1 *Let $\mathcal{T}$ be a fixed system, $\mathcal{F}$ an enumeration of a class of functionals, $\mathcal{I}$ an interpretation in the sense of* (α)–(δ) *and $\mathcal{A}$ an algorithm for $\mathcal{I}$. The triple* $(\mathcal{I}, \mathcal{F}, \mathcal{A})$ *shall be called a* proof interpretation for $\mathcal{T}$.

As we shall see in Section 3.3, functional interpretation combined with the negative translation (having its standard proof as the associated algorithm $\mathcal{A}$) also provides an example of a proof interpretation for PA, via the functionals in Gödel's T. One can even use functional interpretation or cut elimination for giving a constructive proof of the n.c.i.. As we shall discuss in Chapter 3, the n.c.i. has poor scalability behaviour (compared to functional interpretation) in the sense that e.g. it cannot be used for interpreting subsystems of PA by fragments of T, in a way satisfying condition (δ) (cf. Section 3.3.1 and [97]).

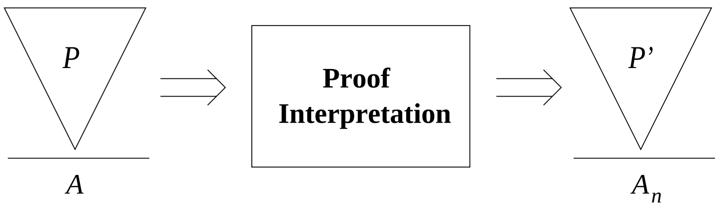


Figure 1.2: The unwinding of proof program

Since the actual goal of this dissertation is to study subsystems of analysis, the main tool we shall use is going to be functional interpretation.

When speaking about unwinding of proofs, Kreisel always made two important points. The first one is that proofs of universal lemmas do not carry any information and can be simply taken as axioms. This means that in practice large parts of the proof need not to be analyzed in order to obtain the required information. The second point is that when unwinding a proof, one should naturally formalize the given informal proof in as constructive a manner as possible, appealing to the machinery only when the most intricate parts of the proof are considered ([114], pg. 171). This can be seen in our treatment of L_1-approximation, in Chapter 6. When formalizing Cheney's proof we have tried to make it as constructive as possible, facilitating the process of extracting the modulus of uniqueness.

We also observe here that ordinary mathematical proofs are in general "partial proofs", in which various lemmas are taken as true facts. Obviously, depending on the choice of proofs for those lemmas one might get different numerical information. The case study on Roth's theorem by Luckhardt, mentioned above, shows that *significant* improvements on the information obtained, however, usually need a totally different *proof idea*.

For more information on Kreisel's unwinding program the reader can consult e.g. [40, 47, 126].

Proof Mining

Due to the lack of more substantial applications, the program of unwinding proofs was considered by some leading proof theories as rather unsuccessful. Only recently, due to the development of new techniques (e.g. monotone functional interpretation), the automatization and refinement of well-known proof interpretations (cf. [16, 17]), and a number of applications in approximation theory and fixed point theory, the idea of analyzing proofs in search for new information reborn. After a suggestion of Dana Scott, Kohlenbach started using the label *proof mining* for this "reborn" program of analyzing proofs in mathematics, mainly ineffective proofs in numerical analysis.

The work in proof mining can be clearly divided into two parts which we shall call *groundwork* and *applications*. By groundwork we mean the development, analysis and study of formal systems about which metatheorems are proven for guaranteeing that certain information can be extracted from some classes of proofs in the system. Moreover, groundwork includes the development of the techniques and algorithms used for extracting such information. These algorithms shall normally be given in the constructive proofs of the metatheorems. The applications include partially formalizing actual proofs from mathematics and extracting interesting new information from it. This dissertation contains contributions in both areas. While our case study in L_1-approximation falls into the range of applications, our study of different forms of bar recursion and our treatment of weak König's lemma in feasible analysis constitute significant groundwork improvements. Moreover, in Chapter 5 the reader can find a general discussion about the program of proof mining, including a survey of current and potential future applications of monotone functional interpretation in numerical analysis.

1.1.3 Metatheorems

We address in this section the last question raised at the beginning of this chapter, namely, how to put the machinery of proof mining into work? In other words, can a mathematician or a computer scientist with some basic knowledge of proof theory do proof mining? The idea is first of all to have general guidelines for recognizing potential applications by observing (based on the language of the formal system and on the representation of the complex concepts) the logical form of the theorem as well as the mathematical principles used in the proof. This can be attained for instance via metatheorems, i.e. theorem of the form:

Theorem 1.1 (General Structure of a Metatheorem) *If a theorem A has a certain restricted form, and if its proof $\mathcal{P}$ can be formalized in a certain restricted system $\mathcal{F}$, then one can effectively produce a new proof $\mathcal{P}'$ of the stronger theorem A'.*

Moreover, a constructive proof of the metatheorem provides a procedure for obtaining $\mathcal{P}'$ and A'. The hope is that, having such metatheorem at hand, a mathematician can match theorems (with their associated proofs) from core mathematics against the pattern described by the metatheorem. And in the case of a positive match, it should be possible to carry out the extraction of the information guaranteed by the metatheorem. For instance, the metatheorem which guided our unwinding of Cheney's proof of Jackson's theorem was the following:

Theorem 1.2 ([90], Theorem 4.1) *Let $\mathcal{T}$ denote a basic classical theory for analysis, $\mathcal{T}_i$ its intuitionistic counterpart and $\mathcal{T}^*$ an extension of $\mathcal{T}$ with principles such as the attainment of the infimum for continuous functions on compact spaces*[3]*. Moreover, let X, K be $\mathcal{T}$-definable Polish spaces, K compact. Consider a sentence A which can be written (when formalized in the language of $\mathcal{T}$) in the form*

$$A :\equiv \forall n \in \mathbb{N} \forall x \in X \forall y \in K \exists k \in \mathbb{N}\, A_1(n, x, y, k),$$

[3] For the precise statement of the theorem see Chapter 6.

where A_1 is a purely existential. If the theory $\mathcal{T}^*$ proves A then one can extract a $\mathcal{T}$-definable functional Φ s.t. $\mathcal{T}_i$ proves

$$\forall n \in \mathbb{N} \forall x \in X \forall y \in K \exists k \leq \Phi(n, \boldsymbol{x})\, A_1(n, x, y, k),$$

where Φ depends on a representative $\boldsymbol{x} \in \mathbb{N} \to \mathbb{N}$ of x in the formal system (see the discussion below on the so-called standard representation of Polish spaces).

One of the crucial points for applying such metatheorems is that the formal system $\mathcal{T}$ has a restricted language in which e.g. only equality on natural numbers is taken as a primitive concept. Consider a simple statement that a real number x is strictly positive, $x >_{\mathbb{R}} 0$. If real numbers and inequality between real numbers are taken as primitive notions this is a computationally empty statement. On the other hand, if rational numbers are represented as pairs of natural numbers, and real numbers as Cauchy sequences of rational numbers with a fixed rate of convergence, then equality between real numbers becomes a defined notion. The Cauchy sequences x and y represent the same real number if *at every point n* the corresponding rational approximations of both sequences are close to each other by 2^{-n}, i.e. equality between real numbers becomes a universal statement. Therefore, inequality between real numbers is an existential statement which can potentially be witnessed. Polish spaces in general shall be represented by number theoretic functions. For instance, elements f of the space of all continuous functions on the unit interval $C[0,1]$, with the uniform metric, are represented via pairs of functions (f_r, ω_f), where f_r is the rational restriction of f and ω_f is a modulus of uniform continuity of f on $[0,1]$. Therefore, a basic knowledge of how mathematical objects are usually represented in restricted formal systems and a good practise in distinguishing the quantifiers hiding in the representation are definite advantages for recognizing potential applications of proof mining. For further discussion on representation see Chapter 5 and e.g. [9, 161].

The second important point is the logical form of the theorem. It is, among others, the restricted logical form of the theorem which guarantees the extraction of additional information. Therefore, after the quantifiers hidden in the representation are presented, one must identify mainly the alternation of quantifiers when the theorem is put in prenex normal form. The extractability of information is normally guaranteed when the theorem can be prenexed into the $\forall\exists$-form, with further restrictions on the type of the variables and on the logical form of the matrix. As we discuss in Chapter 5, those will originally mainly be found in the form of an implication $\forall \to \forall$ or $\exists \to \exists$. In fact, when the whole proof involves only statement of those simple forms the extraction is in general very simple. We shall be mainly interested in situations in which mathematical principles having a $\forall\exists\forall$-form are used in proofs of $\forall\exists$-theorems. In such cases various of those logically more complex principles can either be eliminated from the proof or can be analyzed by more intricate forms of recursion.

In order to ease the job of identifying proofs which can be formalized in the restricted system $\mathcal{T}$, this system should be mathematically as strong as possible. As mentioned above, universal statements do not contribute at all to the unwinding of a proof. Hence, it is often convenient to stipulate that $\mathcal{T}$ contains all true purely universal statements over the underlying language.

Finally, it is important to identify general classes of theorems to which the meta-theory applies, such that one can for some time concentrate the efforts in finding potential applications in some specific area. In Chapter 5 we investigate various classes of statements (e.g uniqueness, convexity, contractivity, monotone convergence and asymptotic regularity) which have the restricted logical form to which the Metatheorem 1.2 applies. We show that when applied to such concepts, monotone functional interpretation creates well-studied notions (e.g. moduli of uniqueness, uniform convexity and so on). Moreover, the treatment given by m.f.i. to the implication between such statements naturally corresponds to translations between the different moduli (cf. the discussion on the functional interpretation of implication in [21]).

Therefore, some basic knowledge of representations, the main metatheorems and potential classes of applications are very helpful in finding theorems and proofs from which extra information can be obtained. The actual extraction of such information will furthermore require at least some "high level" understanding of the tools, i.e. proof interpretations, involved in proof mining.

1.2 Subsystems of Analysis

As already mentioned, this thesis concerns the study and unwinding of proofs which can be formalized in fragments of classical analysis. We present, therefore, a short overview of the main subsystems of classical analysis and our contributions to the study of those.

Let us first, however, discuss a bit about subsystems of arithmetic. Classical first order arithmetic PA contains defining axioms for the basic symbols[4] plus the *induction axiom*

$$\mathsf{IND} \quad : \quad A(0) \wedge \forall n(A(n) \to A(n+1)) \to \forall n A(n).$$

The systematic study of fragments of arithmetic was partially motivated by the correspondence between the logical form of the induction used in the proof of a Π^0_2-theorem and the computational complexity of the program realizing the theorem. In other words, the class of provably recursive functions of an arithmetical theory is closely related to the allowed induction scheme. Kreisel showed that the class of functions provably recursive in PA consists precisely of the *transfinite recursive functions on (standard) well-orderings of ordinals less than* ε_0. Various people have then shown that by restricting induction to Σ^0_1-formulas (we call this system simply Σ^0_1-IND) this class reduces to contain only the *primitive recursive functions*. The further restriction of induction to bounded existential formulas Σ^b_1 – i.e. formulas of the form $\exists x \leq t A_0(x)$, with A_0 a polynomial-time predicate[5] – gives rise to a system in which only

[4]One usually takes here all the primitive recursive functions together with their defining equations.

[5]The class of formulas Σ^b_1 correspond to the NP predicates, i.e. predicates decidable in polynomial-time on a non-deterministic Turing machine.

polynomial-time computable functions are provably recursive. Examples of such systems are Buss' S^1_2 [31] and Cook and Urquhart's CPV [38].

Subsystems of first order arithmetic			
Induction	Σ^b_1	Σ^0_1	full
	$\mathsf{S}^1_2, \mathsf{CPV}$	Σ^0_1-IND	PA

Due to their flexibility and expressive power, we shall normally work in higher type extensions of those theories in the language of functionals of finite type. For instance, by adding to classical first order arithmetic variables and quantifiers for functionals of each finite types, and further axioms for those, one obtains the theory PA^ω. The same was done in [38] leading to the feasible arithmetical theory CPV^ω in the language of finite types.

Second order arithmetic Z^2, also called classical analysis, can be obtained by extending first order arithmetic with variables for sets of numbers and the axiom of *comprehension*

$$\mathsf{CA} \;:\; \exists X \forall n (n \in X \leftrightarrow A(n)),$$

for arbitrary formulas A, possibly containing set parameters. One should notice, however, that in the presence of CA the induction schema can be stated as a single second order axiom:

$$\mathsf{IND} \;:\; 0 \in X \wedge \forall n(n \in X \to n+1 \in X) \to \forall n(n \in X),$$

since formulas of arbitrary logical complexity have associated sets in the system. One of the first persons to systematically study the formalizability of mathematics in weak subsystems of classical analysis was H. Friedman [50]. He developed, among others, the systems $\mathsf{RCA}_0, \mathsf{WKL}_0, \mathsf{ACA}_0$, which form the backbone of Reverse Mathematics [153]. The subsystem of second order arithmetic RCA_0 was first defined in [50], it contains the usual axioms for successor, addition and multiplication; induction restricted to Σ^0_1-formulas (with set parameters) and comprehension for Δ^0_1 sets. The systems WKL_0 and ACA_0 are obtained by extending RCA_0 with weak König's lemma and arithmetical comprehension respectively. Friedman then showed that RCA_0 is Π^0_2-conservative over Σ^0_1-IND, and hence the provably recursive functions of RCA_0 are also the primitive recursive ones. Prompted by a question of Sieg about the development of a relevant subsystem of analysis whose provably recursive functions consist only of the *feasible ones*, Fernando Ferreira [48] developed the system BTFA, by taking feasibility to mean *polynomial-time computability.* Like S^1_2 and CPV, the system BTFA contains induction for Σ^b_1-formulas only.

Subsystems of second order arithmetic			
Comp\Ind	Σ^b_1	Σ^0_1	full (Σ^1_∞)
Δ^0_1	BTFA	RCA_0	RCA
WKL	$\mathsf{BTFA}+\mathsf{WKL}$	WKL_0	WKL
arithmetical	$\times$	ACA_0	ACA
full	$\times$	$\times$	Z^2

As mentioned above, we shall mostly work in systems based on functionals of finite types. In this setting comprehension states the existence of the characteristic function of any given set definable by a formula $A(n)$ in the language

$$\mathsf{CA} \ : \ \exists f \forall n (f(n) = 0 \leftrightarrow A(n)).$$

Notice that, in the presence of classical logic, comprehension can be obtained by countable choice for numbers

$$\mathsf{AC} \ : \ \forall n \exists m B(n,m) \to \exists f \forall n B(n, f(n)),$$

applied to the formula $B(n,m) \equiv m = 0 \leftrightarrow A(n)$. Since we work with systems based on classical logic, restrictions on comprehension in the second order systems shall correspond to restrictions in the axiom of choice in the finite type systems.

The main work associated with Reverse Mathematics [153] consists in showing that not only great part of analysis can be carried out in those few subsystems, but various mathematical theorems are actually equivalent, over the basic theory RCA_0, to those. In the next two sections we comment further on the two logical principles: weak König's lemma and comprehension.

1.2.1 Weak König's Lemma

Weak König's lemma WKL states that every infinite binary branching tree has an infinite path. This principle can be viewed as the logical abstraction of numerous ineffective mathematical principles, such as (cf. [153] for a more comprehensive list)

- a continuous function $f : [0,1] \to \mathbb{R}$ has an infimum and attains it,
- the Cauchy-Peano theorem for differential equations,
- the Heine/Borel covering lemma in sequential form for $[0,1]$.

In fact, as shown in Reverse Mathematics [153], over the basic analytical system RCA_0, weak König's lemma is provably equivalent to those principles, even when $f \in C[0,1]$ is given together with a modulus of uniform continuity, as in our case study in L_1-approximation.

WKL is non-computational in the sense that there are *(primitive) recursive* infinite binary trees which have only *non-computable* infinite branches. This is in correspondence with the fact that there are computable (and even polytime [83]) continuous functions $f \in C[0,1]$ which attain their infimum only at non-computable reals. Therefore, WKL and all the mathematical principles listed above are not valid in a model where all the functions/real numbers are computable.

Nevertheless, Harvey Friedman showed that in proofs of Π^0_2-theorems over the basic system RCA_0 uses of WKL can be eliminated. This means that the system $\mathsf{RCA}_0 + \mathsf{WKL}$ (called WKL_0) has precisely the same Π^0_2-theorems as the basic system RCA_0. Moreover, it also implies that the provably recursive functions of $\mathsf{RCA}_0 + \mathsf{WKL}$ are exactly the primitive recursive ones. This result was later extended by Harrington, who showed that $\mathsf{RCA}_0 + \mathsf{WKL}$ is actually Π^1_1-conservative over RCA_0. Both Friedman's and Harrington's proofs involve non-constructive model theoretic arguments, and hence do not provide at first sight any procedure for extracting primitive recursive programs realizing Π^0_2-theorems of WKL_0. The first effective version of Friedman's result was given by Sieg [151] using cut-elimination, a Herbrand analysis and a simple form of Howard's majorizability for primitive recursive terms. In [89], a combination of Gödel's functional interpretation with Howard's hereditary majorizability for functionals in all finite types is developed to extract uniform bounds for $\forall\exists$-theorems in analysis from proofs based on various analytical principles including WKL. In particular, [89] yields effective forms of extensions of Friedman's WKL-conservation result to higher types (cf. also [5], Theorem 7.1.1). In the *feasible setting*, Ferreira also showed that weak König's lemma does not have any computational impact in the class of provably recursive functions, i.e. Π^0_2-theorems of the system $\mathsf{BTFA}+\mathsf{WKL}$ still have polynomial-time computable realizers. As Friedman's and Harrington's proofs, Ferreira's proof is also non-constructive, and therefore does not provide a procedure for extracting polynomial-time computable programs realizing Π^0_2-theorems of $\mathsf{BTFA} + \mathsf{WKL}$.

Weak König's Lemma in Cheney's Proof

As mentioned above, one of the main contributions of this dissertation is the logical analysis of Cheney's proof of Jackson's theorem, and the extraction of the first fully effective modulus of uniqueness for L_1-approximation. The problem of L_1-approximation is defined as follows: for a given number n and a continuous real valued function $f \in C[0,1]$ find a polynomial of degree bounded by n (we let P_n denote that space) such that p best approximates f w.r.t. the L_1-norm, where the L_1-norm of a function $f \in C[0,1]$ is defined as

$$\|f\|_1 := \int_0^1 |f(x)|dx.$$

Jackson's theorem states that the solution for this problem is unique, i.e. for fixed n and $f \in C[0,1]$, there exists a unique polynomial $p \in P_n$ such that

$$\|f - p\|_1 = \inf_{p' \in P_n} \|f - p'\|_1.$$

The interesting fact about L_1-approximation is that, as for Chebycheff approximation but in contrast to L_p-approximation for $1 < p < \infty$, both the existence and the uniqueness proof make use of weak König's Lemma. First we note that, without loss of generality the L_1-approximation problem can be restricted to the *compact space* $K_{f,n} :\equiv \{p \in P_n : \|p\|_1 \leq 2\|f\|_1\}$. The proof of existence is very simple. Define a function $\Psi(f,p) := \|f - p\|_1$, which takes a continuous function and a polynomial as input and returns a real number. By the continuity of the L_1-norm the function Ψ is also continuous. Therefore, on the compact space $K_{f,n}$ it attains its infimum, and that is a best approximation of f w.r.t. the L_1-norm. The use of WKL in this proof seems unavoidable, since the existence statement

$$\forall f \in C[0,1], n \in \mathbb{N} \exists p \in K_{f,n}(\Psi(f,p) = dist_1(f, K_{f,n})),$$

has a logical form as complicated as the statement of the attainment of the infimum itself. The uniqueness statement,

$$\forall f \in C[0,1], n \in \mathbb{N}, p_1, p_2 \in K_{f,n}(\bigwedge_{i=1}^{2}(\Psi(f,p_i) = dist_1(f, K_{f,n})) \to p_1 = p_2),$$

on the other hand, has a logical form to which WKL-conservation applies, since it has the form A in Metatheorem 1.2. This can be seen by the fact that both $\Psi(f,p_i) = dist_1(f,K_{f,n})$ and $p_1 = p_2$ are purely universal statements which when prenexed lead to a $\forall\exists$-form. Notice, moreover, that $dist_1(f,K_{f,n})$ is computable on the representation of f as an element of $C[0,1]$ (cf. Chapter 6).

Therefore, although the uniqueness proof given by Cheney [33] also makes use of WKL (in the form *the infimum of a strictly positive continuous function on a closed interval is strictly positive*) we are guaranteed by the metatheorem that a functional $\Phi(f,n,\varepsilon) \in \mathbb{Q}^*_+$, independent of p_1 and p_2, can be extracted from such a proof satisfying,

$$\begin{cases} \forall f \in C[0,1], n \in \mathbb{N}, \varepsilon \in \mathbb{Q}^*_+ \forall p_1, p_2 \in K_{f,n} \\ (\bigwedge_{i=1}^{2}(|\Psi(f,p_i) - dist_1(f,K_{f,n})| \le \Phi(n,f,\varepsilon)) \to \|p_1 - p_2\|_1 \le \varepsilon). \end{cases}$$

As we mentioned in Section 1.1.3, this functional will depend on the representation of $f \in C[0,1]$ in our formal system, i.e. on the pair (f_r, w_f) of rational values of f and the modulus of uniform continuity of f on $[0,1]$. Moreover, since the space $\mathcal{K}_{\omega,M}$ of all functions $f \in C[0,1]$ which have the common modulus of uniform continuity ω and the common bound $\|f\|_\infty \le M$ is (pre-)compact (w.r.t. $\|\cdot\|_\infty$), the same logical meta-theorem guarantees the extractability of a modulus of uniqueness Φ which only depends on $\mathcal{K}_{\omega,M}$ i.e. on ω, M (in addition to n,ε). And even the M-dependency can be a-priori eliminated as the approximation problem for f can be reduced to that for $\tilde{f}(x) := f(x) - f(0)$ so that only a bound $N \ge \sup_{x\in[0,1]} |f(x) - f(0)|$ is required. This bound can easily be computed from ω, e.g take $N := \lceil \frac{1}{\omega(1)} \rceil$.

Theorem 1.3 (Chapter 6, [107]) *The functional*

$$\Phi(n,\omega,\varepsilon) := \min\{\tfrac{c_n\varepsilon}{8(n+1)^2}, \tfrac{c_n\varepsilon}{2}\omega_n(\tfrac{c_n\varepsilon}{2})\},$$

where

$$c_n := \tfrac{\lfloor n/2 \rfloor! \lceil n/2 \rceil!}{2^{4n+3}(n+1)^{3n+1}} \quad \text{and} \quad \omega_n(\varepsilon) := \min\{\omega(\tfrac{\varepsilon}{4}), \tfrac{\varepsilon}{40(n+1)^4 \lceil \frac{1}{\omega(1)} \rceil}\},$$

is a uniform modulus of uniqueness for the best L_1-approximation from P_n of any function f in $C[0,1]$ having modulus of uniform continuity ω.

Interestingly enough, having such a modulus of uniqueness Φ one can produce an algorithm for computing the best L_1-approximation of a given function f, leading to an *effective existence proof*, as follows. Given a function f and number n, take p_1 in the above equation to be the best L_1-approximation p_b of

f and p_2 to be an arbitrary polynomial in P_n. Since $|\Psi(f,p_b) - dist_1(f, K_{f,n})|$ is actually zero, we can omit that premise and we get

$$\forall \varepsilon \in \mathbb{Q}^*_+ \forall p \in K_{f,n}(|\Psi(f,p) - dist_1(f, K_{f,n})| \leq \Phi(n,f,\varepsilon) \to \|p_b - p\|_1 \leq \varepsilon).$$

This means that, in order to find a polynomial p ε-close to the solution p_b we just need to find a polynomial p satisfying

$$|\Psi(f,p) - dist_1(f, K_{f,n})| \leq \Phi(n,f,\varepsilon)$$

which can be done by doing a search on the $\Phi(n,f,\varepsilon)$-net of the compact space $K_{f,n}$. Notice that for this argument it is fundamental that the modulus Φ does not depend on the polynomials p_1 and p_2. The complexity analysis leads to the following result:

Theorem 1.4 (Chapter 7, [134]) *Let $f \in C[0,1]$ be polynomial-time computable and A_f and B_f oracles computing general left cuts of $dist_1(f, P_n)$ and of the integral of f on $[0,1]$, respectively, then the sequence $(p_n)_{n\in\mathbb{N}}$ of best L_1-approximations of f from P_n is*

- *strongly* $\mathbf{NP}[A_f,\, B_f]$ *computable,*
- *strongly* $\mathbf{NP}$ *computable in* $\mathbf{NP}[B_f]$.

In Section 2.3.1 we include, for completeness, Cheney's proof of Jackson's theorem, and we indicate exactly where in the proof WKL is used. The analysis of Cheney's proof is carried out in detail in Chapter 6. In Chapter 7 we use this modulus in order to give the first complexity analysis of the problem of L_1-approximation.

Weak König's Lemma in Feasible Analysis

In this dissertation we also study the computational impact of WKL over basic systems of feasible analysis. As we have mentioned above, F. Ferreira was the first to develop a relevant feasible subsystem of analysis, BTFA. Moreover, he showed that by adding an appropriate form of WKL to BTFA one does not get any new Π^0_2-theorems. Therefore, the Π^0_2-theorems of the system $\mathsf{BTFA} + \mathsf{WKL}$ still have polynomial-time computable realizers. The proof of this conservation result, however, is non-constructive and therefore does not provide a procedure for extracting from a proof of a Π^0_2-theorem in $\mathsf{BTFA} + \mathsf{WKL}$ a polynomial-time computable realizer for the theorem.

In Chapter 9 we show how the second order theory BTFA can be recast and extended to the setting of finite types. We use for that purpose Cook and Urquhart's [38] system CPV^ω extended with quantifier-free choice. This system can be viewed as an extension to higher types of BTFA without bounded collection. In order to demonstrate the potential of the system $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ as a basis for feasible analysis we indicate in Chapter 9 how to formalize the proof of the sequential form of Heine/Borel covering lemma for the unit interval in this system. The motivation for using this finite type system was to make use of functional interpretation as a tool for extracting polynomial-time realizers from proofs of Π^0_2-theorems in feasible analysis, given that $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ has, via negative translation, a Gödel functional interpretation in IPV^ω, the intuitionistic counterpart of CPV^ω [38]. The functional interpretation of WKL, however, cannot be realized by functionals of IPV^ω. Therefore, our interpretation of $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} + \mathsf{WKL}$ involves also a novel form of binary bar recursion. Nevertheless, we show that when used to build type one programs in the feasible setting this binary bar recursion can always be eliminated in favour of the allowed limited recursion on notation. This provides an effective procedure for extracting polynomial-time realizers from proofs of Π^0_2-theorems in the system $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} + \mathsf{WKL}$. It should be noted, however, that while our procedure handles only WKL for trees defined by formulas of the kind $\forall x T(w, x)$, where $T(w, x)$ is a quantifier-free formula, Ferreira's conservation result holds for arbitrary bounded formulas $T(w, x)$.

An interpretation of WKL via binary bar recursion had already been presented in [66] for stronger arithmetic settings, in which bounded quantifiers can be absorbed by quantifier-free matrices and need not be considered by functional interpretation. In Chapter 9 we also discuss how our treatment differs from [66].

The author has been recently informed by J. Avigad that Sieg's proof of WKL elimination was successfully translated to the feasible setting by B. Kauffmann [72], using an extension of S^1_2 with 0-1 function variables and quantifier-free choice for those functions. In fact, by making use of a version of Parikh's lemma due to S. Buss, Kauffmann obtains an effective WKL elimination for trees defined by arbitrary bounded formulas. This provides another effective procedure for extracting polynomial-time realizers from WKL-proofs in feasible analysis, via an initial elimination of WKL from the proof. It should be noticed, however, that our algorithm produces a polynomial-time realizer *directly* from the original WKL-proof, without having to initially go through the elimination procedure. Moreover, our approach has a finite type theory as basis and enjoys the *full modularity* of the negative translation and functional interpretation, whereas Sieg's proof, as well as Parikh's lemma, are based on cut-elimination.

1.2.2 Full Comprehension

As mentioned above, Spector [154] extended Gödel's functional interpretation of arithmetic to analysis by interpreting comprehension (and even dependent choice) via a new scheme of recursion, so-called *bar recursion.* The idea is that, given a well-founded tree T, one initially assigns values to the leaves of the tree, and from those values one computes the values of the internal nodes, i.e.

$$\Phi(T, G, H, s) =_\tau \begin{cases} G(s) & \text{if } s \notin T \\ H(s, \lambda x^\rho.\Phi(T, G, H, s * x)) & \text{otherwise,} \end{cases}$$

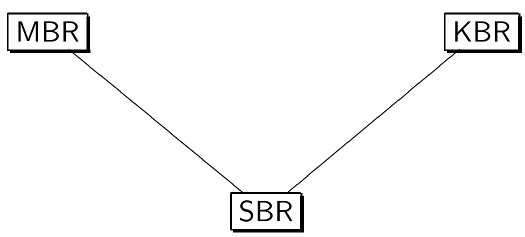
where $s * x$ denotes the extension of the finite sequence $s : \rho^*$ with a new element $x : \rho$.

Note that the value of the recursion Φ at an internal node s might depend on the values of all successor nodes $s * x$. The well-founded tree used by Spector to interpret comprehension is the tree obtained from a continuous functional $Y : (\mathbb{N} \to \rho) \to \mathbb{N}$, by looking at the point of continuity of Y on each infinite sequence $\alpha : \mathbb{N} \to \rho$. In fact, only a weaker property than continuity is required. Spector's tree is defined as $s \in T :\equiv Y(\hat{s}) > |s|$, where $\hat{s}$ is the infinite continuation of s with the zero object of type ρ. It is easy to see that, if Y is continuous, as the sequence s gets longer and longer the point of continuity is eventually reached for some s, so that the value of $Y(\hat{t})$ is constant for all extensions t of s. Therefore, the condition $Y(\hat{s}) \leq |s|$ is bound to be satisfied as s gets longer, and T as defined above is well-founded. Although the argument above relies on continuity, Spector's condition is not restricted to continuity. This can be seen by the fact that Spector's bar recursion holds in the model of strongly majorizable functionals, as defined by Howard [65] and Bezem [18], which contains discontinuous functions.

Recently, U. Berger [15] showed that comprehension can also be interpreted via modified realizability combined with negative translation and Friedman's A-translation. This interpretation, however, seems to require a different form of bar recursion, so-called *modified bar recursion*, MBR, and a continuity argument for the verification. We have studied this new scheme of bar recursion and together with U. Berger obtained the following results:

- the type structure of strongly majorizable functionals is also a model of MBR,
- Spector's bar recursion is primitive recursively (in the sense of Kleene) definable in MBR,
- MBR together with a version of bar recursion KBR due to Kohlenbach [88] defines primitive recursively the minimal fan functional[6]. Therefore, since the fan functional is not S1-S9 computable (and KBR is) we obtain that MBR is not S1-S9 computable.

Moreover, since Spector's bar recursion is S1-S9 computable, it follows that modified bar recursion is strictly stronger than Spector's bar recursion. Kohlenbach also shows in [88] that the structure of strongly majorizable functionals is not a model of KBR nor of the fan functional. Therefore, the results above implies that MBR and KBR are incomparable and that none of those alone define the fan functional. Those results are presented in Chapters 4 and 8.



In Chapter 5 one can find a general description of how proofs involving comprehension can be treated via the use of bar recursion. We illustrate this

[6] A type three functional $\Phi(Y)$ is called a *fan functional* if it computes points of uniform continuity of given type two functionals $Y : (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$ on the Cantor space.

procedure in Chapter 4, where we analyze a proof of a $\forall\exists$-statement which uses comprehension. The proof is included in Section 2.3.2.

1.3 Organization

This dissertation is divided into two parts. The first part can be viewed as presenting the background necessary for understanding the second part. There are, however, new results in part one as well, mainly in Chapter 4.

Part One – Context

Chapter 2. Here we introduce the formal systems of arithmetic used in the dissertation. Those are systems based on finite types which mainly differ by the treatment of higher type equality. Moreover, we define the majorizability relation and the type structure of strongly majorizable functionals, which shall be used in Chapters 4 and 8. We also show how those systems can be extended to cover well-known subsystems of analysis. For completeness we include here Cheney's proof, which is based on weak König's lemma and is studied in further detail in Chapter 6, and a proof based on arithmetical comprehension due to J. Avigad [4], which we use in Chapter 4 to illustrate the use of bar recursion.

Chapter 3. Here we introduce the main interpretations and translations which we use in this dissertation. Those are negative translation, functional interpretation, monotone functional interpretation, modified realizability and A-translation. Furthermore, we discuss the relation between functional interpretation and the no-counterexample interpretation of Kreisel. Monotone functional interpretation is further discussed in Chapter 5.

Chapter 4. In this chapter we describe the interpretation of comprehension via bar recursion. We also include new results concerning the relation between modified and Spector's original bar recursion. Those results complement the ones presented in Chapter 8.

Part Two – Papers

Chapter 5. This chapter contains a survey on proof mining, with applications of monotone functional interpretation to various mathematical principles. We show that in every case the result of the interpretation corresponds to well-known concepts. We argue that the monotone functional interpretation of an implication has a very natural numerical meaning, which in many special cases has been studied in mathematics before. We also describe in further detail how to analyze proofs which involve weak König's lemma and fixed instances of Π^0_1 comprehension.

Chapter 6. Here we present a case study of L_1-approximation where, from the proof of uniqueness, the first fully effective (in all parameters) modulus of uniqueness is extracted. The analyzed proof makes use of weak König's lemma, in the form "the infimum of a strictly positive function $f \in C[0,1]$ is strictly positive".

Chapter 7. We use the modulus of uniqueness presented in Chapter 6 to give the first complexity analysis of the sequence of polynomials (from P_n) best L_1-approximating polynomial-time computable functions $f \in C[0,1]$.

Chapter 8. In this chapter we present the new form of bar recursion, *modified bar recursion* and we show how it can be used to interpret comprehension, via modified realizability and A-translation. We also show that any set theoretic functional which satisfies the equation for MBR can be interpreted in the structure of strongly majorizable functionals.

Chapter 9. Here we present the first modular procedure for extracting polynomial-time realizers from ineffective proofs of Π^0_2-theorems in feasible analysis. By "ineffective proof" we mean a proof involving, in addition to classical logic, weak König's lemma. The procedure makes use of functional interpretation for the feasible setting [38] plus a new form of binary bar recursion.

1.4 Contributions

Most of the material of this dissertation has been previously reported in the following papers:

Refereed articles

[15] **Modified bar recursion and classical dependent choice**
with U. Berger, to appear in: *Lecture Notes in Logic*, 19 page. (Chapter 8)

[106] **Proof mining: A systematic way of analyzing proofs in mathematics**
with U. Kohlenbach, to appear in: *Proc. Steklov Inst. Math.*, 33 pages. (Chapter 5)

[107] **Proof mining in L_1-approximation**
with U. Kohlenbach, *Annals of Pure and Applied Logic*, 121:1–38, 2003. (Chapter 6)

[134] **On the computational complexity of L_1-approximation**
Mathematical Logic Quarterly, 48.s1:66–77, 2002. (Chapter 7)

[135] **Polynomial-time algorithms from ineffective proofs**
Proc. of the Eighteenth Annual IEEE *Symposium on Logic in Computer Science* LICS'03, pages 128–137, IEEE Press, 2003. (Chapter 9)

Technical reports (not covered by the above publications)

[14] **Modified bar recursion**
with U. Berger. BRICS *Report Series*, RS-02-14, 23 pages, 2002. (Chapter 4)

Published abstracts

- **Effective bounds on strong unicity in L_1-approximation**
  with U. Kohlenbach. *Bull. Symbolic Logic*, 8, pg. 143, 2002.

- **Proof mining in L_1-approximation**
  with U. Kohlenbach. *Proc. ICC 2001*, BRICS Notes Series NS-01-3, 117–122, 2001.

- **On the extraction of polynomial-time algorithms from ineffective proofs in feasible analysis**
  to appear in: *Bull. Symbolic Logic.*

Chapter 2

Formal Systems of Arithmetic and Analysis

In this section we shall introduce the formal systems to be used in this dissertation. We start by presenting the intuitionistic and classical first order systems of arithmetic: Heyting and Peano arithmetic, respectively. Then we define four extensions of those systems to all finite types, giving different treatments to the higher type equality. Subsystems of analysis are then obtained via the addition of weak König's lemma and various forms of countable choice for numbers. The description of the systems is based on [87, 124, 160].

2.1 First Order Arithmetic

Most of the systems used in this dissertation are based on first order intuitionistic arithmetic, so-called Heyting arithmetic HA. The language $\mathcal{L}$ of HA contains the logical constants $\forall, \exists, \rightarrow, \wedge$ and $\vee$, non-logical symbols 0 (constant zero), S (successor function)[1] and function symbols for all primitive recursive functions. Moreover, HA contains a single binary predicate symbol $=$ (equality between numbers). The *terms* and *formulas* of HA are defined out of the above symbols as usual.

The *logical axioms* and *rules* of HA consist of those of intuitionistic first order predicate logic. Heyting arithmetic contains also the following *non-logical axioms*:

(i) basic equality axioms: $x = x$, $x = y \wedge z = y \rightarrow x = z$,

(ii) extra equality axioms: for each n-ary function symbol f,

$\bigwedge_{i=1}^{n}(x_i = y_i) \rightarrow f(\underline{x}) = f(\underline{y})$,

(iii) successor axioms: $Sx \neq 0$, $Sx = Sy \rightarrow x = y$,

(iv) induction scheme: $A(0) \wedge \forall x(A(x) \rightarrow A(Sx)) \rightarrow \forall x A(x)$,

(v) the defining equation of each primitive recursive function.

The classical counterpart of Heyting arithmetic, so-called Peano arithmetic PA, is obtained from HA by adding all instances of the law of excluded middle

[1]The negation of a formula A (denoted by $\neg A$) is defined as $A \rightarrow 0 = S0$.

$\mathsf{LEM} \; : \; A \vee \neg A.$

2.2 Arithmetic in All Finite Types

The set of *finite types* $\mathbf{T}$ is defined inductively as follows[2]: $\mathbb{N} \in \mathbf{T}$, and if $\rho, \tau \in \mathbf{T}$ then $\rho \to \tau \in \mathbf{T}$. The set $\mathbf{P} \subset \mathbf{T}$ of *pure finite types* is defined by $\mathbb{N} \in \mathbf{P}$, and if $\rho \in \mathbf{P}$ then $\rho \to \mathbb{N} \in \mathbf{P}$. The *level* of a finite type is defined by

- $\mathsf{level}(\mathbb{N}) = 0$,
- $\mathsf{level}(\rho \to \tau) = \max(\mathsf{level}(\rho) + 1, \mathsf{level}(\tau))$.

We shall now introduce four extensions of HA to all finite types, which mainly differ in the treatment of higher type equality $=_\sigma$, $\mathsf{level}(\sigma) > 0$. All the theories are based on many-sorted predicate logic. Ideally, objects and predicates should be treated, as in ordinary mathematics, as extensional entities, in the sense that

(I) objects with the same extensional behaviour are identified (say $\sigma \equiv \rho \to \tau$)

$$z_1 =_\sigma z_2 \leftrightarrow \forall x^\rho (z_1 x =_\tau z_2 x),$$

(II) objects do not distinguish extensionally equal objects (z of type $\rho \to \tau$)

$$x_1 =_\rho x_2 \to z x_1 =_\tau z x_2.$$

The amount of extensionality present in the four systems we shall consider can be summarized as follows:

- $\mathsf{N\text{-}HA}^\omega$ contains equality between objects of higher types as a primitive relation, and those not necessarily respect (I). Hence, even though (II) is included, the system is called *not extensional* or *neutral*.
- HA^ω is a subsystem of $\mathsf{N\text{-}HA}^\omega$ in which equality between objects of type $\mathbb{N}$ is the only primitive relation. Moreover, HA^ω contains only a weak form of (II) for the type $\mathbb{N}$ only. The system is also called *not extensional* or *neutral*, and serves as a basis for the next two systems.
- $\mathsf{WE\text{-}HA}^\omega$ is an extension of HA^ω in which a quantifier-free rule version of (II) is included, by taking (I) as the definition of higher type equality. The rule says that if $A_0 \to s =_\rho t$ is proven in the system, A_0 quantifier-free, then one can also infer $A_0 \to r[s] =_\sigma r[t]$. The equality between higher type objects is a mere abbreviation as given in (I). The system is called *weakly extensional*,
- $\mathsf{E\text{-}HA}^\omega$ extends $\mathsf{WE\text{-}HA}^\omega$ by taking (II) as an axiom. The system is then called *fully extensional.*

[2]In this dissertation we shall confuse the basic finite type with the set of natural numbers, since in all the concrete models we use the basic type is indeed populated by the natural numbers.

Notation 2.1 *We shall use the following conventions in the dissertation:*

- ρ^ω *abbreviates the finite type* $\mathbb{N} \to \rho$,
- 1 *for any type* ρ *such that* $\mathsf{level}(\rho) = 1$.
- $t : \rho$ *and* t^ρ *stand for "term* t *has type* ρ*"*,
- $A(x)$ *denotes that a variable* x *is free in the formula* A*, and* $A(s)$ *denotes the formula obtained by replacing all the occurrences of* x *in* A *by the term* s,
- $t[x]$ *denotes that a variable* x *is free in the term* t*, and* $t[s]$ *denotes the term obtained by replacing all the occurrences of* x *in* t *by* s,
- *sequences of variables* $x_1, \ldots, x_n$ *shall be abbreviated by* $\underline{x}$,
- *An* atomic formula *is a formula of the form* $P(\underline{x})$*, where* P *is predicate,*
- *A formula* A_0 *is said to be* quantifier-free *if it is built from atomic formulas by the use of* $\wedge$, $\vee$ *and* $\to$ *only.* $A_0, B_0, \ldots$ *shall be used to denote quantifier-free formulas*[3],
- *A formula* A *is* universal *(or* purely universal*) if it has the form* $\forall x A_0(x)$*, where* $A_0(x)$ *is quantifier-free,*
- *A formula* A_{ef} *is called* $\exists$-free *if it is built from atomic formulas by the use of* $\wedge, \to$ *and* $\forall$ *only.*

Neutral Arithmetic in All Finite Types

The first system we present shall be denoted by $\mathsf{N\text{-}HA}^\omega$ (following nomenclature of [160]). The *language* $\mathcal{L}^\omega_{\mathsf{h}}$ of $\mathsf{N\text{-}HA}^\omega$ contains a countable amount of variables for each finite type σ (indicated by $x^\sigma, y^\sigma, \ldots$) and the following constants, for all finite types ρ, σ and τ,

- 0 of type $\mathbb{N}$,
- S of type $\mathbb{N} \to \mathbb{N}$,
- $\Pi_{\sigma,\tau}$ of type $\sigma \to \tau \to \sigma$,
- $\Sigma_{\rho,\sigma,\tau}$ of type $(\rho \to \sigma \to \tau) \to (\rho \to \sigma) \to \rho \to \tau$,
- R_σ of type $\sigma \to (\rho \to \mathbb{N} \to \rho) \to \mathbb{N} \to \rho$.

[3] In the systems with decidable atomic formulas which we shall present, for each quantifier-free formulas $A_0(x)$ one can build a term $t(x)$ such that $A_0(x) \leftrightarrow t(x) = 0$ is provable in the theory. Therefore, the class of quantifier-free formulas coincide with the class of atomic formulas. Moreover, for those systems the class of quantifier-free formulas is included in the class of $\exists$-free formulas.

The *predicate symbols* of $\mathsf{N\text{-}HA}^\omega$ consists of equality $=_\sigma$ for each finite type σ. Whenever clear from the context, we shall omit the typing subscript in the equality predicates. The language of $\mathsf{N\text{-}HA}^\omega$ contains also quantification over the newly introduced variables of finite type. With each *term* of $\mathsf{N\text{-}HA}^\omega$ we associate a finite type so that variables and constants of type σ are terms of type σ, and if t is a term of type $\rho \to \tau$ and s is a term of type ρ then (ts) is a term of type τ. We shall often abbreviate (ts) by ts or $t(s)$, and, in general, $(\ldots(t_1 t_2)\ldots t_n)$ by either $t_1 t_2 \ldots t_n$ or $t_1(t_2, \ldots, t_n)$. The *formulas* of $\mathsf{N\text{-}HA}^\omega$ consist of the closure of the atomic formulas $t^\sigma =_\sigma s^\sigma$ over the logical symbols in the usual way. This concludes the description of the language of $\mathsf{N\text{-}HA}^\omega$. The *non-logical axioms* consist of

(i) basic equality axioms: $x^\sigma = x^\sigma, \;\; x^\sigma = y^\sigma \wedge z^\sigma = y^\sigma \to x^\sigma = z^\sigma$,

(ii) extra equality axioms:

$x^\sigma = y^\sigma \to z^{\sigma\to\rho}x^\sigma = z^{\sigma\to\rho}y^\sigma, \;\; x^{\sigma\to\rho} = y^{\sigma\to\rho} \to xz^\sigma = yz^\sigma$,

(iii) successor axioms: $Sx^{\mathbb{N}} \neq 0, \;\; Sx = Sy \to x^{\mathbb{N}} = y^{\mathbb{N}}$,

(iv) the induction scheme: $A(0) \wedge \forall x(A(x) \to A(Sx)) \to \forall x A(x)$,

where A is an arbitrary formula in the language of $\mathsf{N\text{-}HA}^\omega$,

(v) the defining equations for the combinators, for all types ρ, σ and τ,

$\Pi_{\sigma,\tau} x^\sigma y^\tau =_\sigma x, \quad \Sigma_{\rho,\sigma,\tau} x y^{\rho\to\sigma} z^\rho =_\tau xz(yz)$,

$R_\sigma xy0 =_\sigma x, \quad R_\sigma xy(Sz^{\mathbb{N}}) =_\sigma y(R_\sigma xyz)z$.

We use the system $\mathsf{N\text{-}HA}^\omega$ in Chapter 8 in connection with the interpretation of classical countable and dependent choice, using modified realizability. We shall for convenience in Chapter 8 enrich the system $\mathsf{N\text{-}HA}^\omega$ with the formation of product types and finite sequences, by saying that whenever ρ and σ are finite types then $\rho \times \sigma$ and ρ^* are also finite types. The definition of $\mathsf{level}(\rho)$ is extended as: $\mathsf{level}(\rho \times \sigma) = \max(\mathsf{level}(\rho), \mathsf{level}(\sigma))$ and $\mathsf{level}(\rho^*) = \mathsf{level}(\rho)$. In that case, we use o for an arbitrary finite type of type level 0. The classical theory $\mathsf{N\text{-}PA}^\omega$ is obtained from $\mathsf{N\text{-}HA}^\omega$ by adding the axiom scheme LEM for all formulas in the language of $\mathsf{N\text{-}HA}^\omega$.

A Different Neutral System

We shall now describe a subsystem of $\mathsf{N\text{-}HA}^\omega$, so-called HA^ω, in which equality between objects of type $\mathbb{N}$ is the only primitive relations, i.e. the *language* $\mathcal{L}^\omega$ of HA^ω is simply $\mathcal{L}^\omega_{\mathsf{h}}$ without higher type equality. The axioms of HA^ω are as follows

(i) equality axioms for type $\mathbb{N}$: $x^{\mathbb{N}} = x^{\mathbb{N}}, \;\; x^{\mathbb{N}} = y^{\mathbb{N}} \wedge z^{\mathbb{N}} = y^{\mathbb{N}} \to x^{\mathbb{N}} = z^{\mathbb{N}}$,

(ii) $x^{\mathbb{N}} = y^{\mathbb{N}} \to t[x] =_{\mathbb{N}} t[y]$, for all terms $(t[x^{\mathbb{N}}])^{\mathbb{N}}$,

(iii) successor axioms: $Sx^{\mathbb{N}} \neq 0, \;\; Sx = Sy \to x^{\mathbb{N}} = y^{\mathbb{N}}$,

(iv) the induction scheme: $A(0) \wedge \forall x(A(x) \to A(Sx)) \to \forall x A(x)$,

(v) combinators axioms, for all terms $(t[\cdot])^{\mathbb{N}}$ and types ρ, σ and τ,

$t[\Pi_{\sigma,\tau} x^\sigma y^\tau] =_{\mathbb{N}} t[x^\sigma]$,

$t[\Sigma_{\rho,\sigma,\tau} xyz] =_{\mathbb{N}} t[xz(yz)]$,

$t[R_\sigma xy0] =_{\mathbb{N}} t[x]$,

$t[R_\sigma xy(Sz)] =_{\mathbb{N}} t[y(R_\sigma xyz)z]$.

The classical theory PA^ω is obtained from HA^ω by adding the axiom scheme LEM for all formulas in the language of HA^ω. We shall use HA^ω merely as a basis for the next two systems $\mathsf{WE}\text{-}\mathsf{HA}^\omega$ and $\mathsf{E}\text{-}\mathsf{HA}^\omega$.

Extensional Arithmetic in All Finite Types

Notice that, even though in the language of $\mathsf{N}\text{-}\mathsf{HA}^\omega$ equality predicates are present for each finite type, those are not assumed in the theory to behave extensionally, more precisely objects that have the same extensional behaviour need not be equal. We shall now define an extensional version of HA^ω, so-called $\mathsf{E}\text{-}\mathsf{HA}^\omega$. The language of $\mathsf{E}\text{-}\mathsf{HA}^\omega$ is simply $\mathcal{L}^\omega$, i.e. *equality for the basic type* $\mathbb{N}$ *is the only primitive predicate*. Equality for higher types is taken to be an inductively defined notion as

$$x^{\rho\to\tau} =_{\rho\to\tau} y^{\rho\to\tau} :\equiv \forall z^\rho (xz =_\tau yz).$$

The non-logical axioms of $\mathsf{E}\text{-}\mathsf{HA}^\omega$ are:

(i) equality axioms for type $\mathbb{N}$: $x^{\mathbb{N}} = x^{\mathbb{N}}$, $x^{\mathbb{N}} = y^{\mathbb{N}} \wedge z^{\mathbb{N}} = y^{\mathbb{N}} \to x^{\mathbb{N}} = z^{\mathbb{N}}$,

(ii) the scheme of higher type extensionality[4] $(\rho = \rho_1 \to \ldots \to \rho_n \to \mathbb{N})$

$\mathsf{EXT}^\rho \;:\; \forall z^\rho, \underline{x}, \underline{y}(\bigwedge_{i=1}^n (x_i =_{\rho_i} y_i) \to z\underline{x} =_{\mathbb{N}} z\underline{y})$,

(iii) successor axioms: $Sx^{\mathbb{N}} \neq 0$, $Sx = Sy \to x^{\mathbb{N}} = y^{\mathbb{N}}$,

(iv) the induction scheme: $A(0) \wedge \forall x(A(x) \to A(Sx)) \to \forall x A(x)$,

(v) the defining equations for the combinators, for all types ρ, σ and τ,

$\Pi_{\sigma,\tau} x^\sigma y^\tau =_\sigma x$, $\quad \Sigma_{\rho,\sigma,\tau} xy^{\rho\to\sigma} z^\rho =_\tau xz(yz)$,

$R_\sigma xy0 =_\sigma x$, $\quad R_\sigma xy(Sz) =_\sigma y(R_\sigma xyz)z$.

In connection with functional interpretation (cf. Section 3.3) one uses the 'weakly extensional' theory $\mathsf{WE}\text{-}\mathsf{HA}^\omega$, which is obtained from $\mathsf{E}\text{-}\mathsf{HA}^\omega$ by replacing axiom (ii) with the *quantifier-free rule of extensionality*

$$\mathsf{EXT}\text{-}\mathsf{R}_{\mathsf{qf}} \;:\; \frac{A_0 \to s =_\rho t}{A_0 \to r[s] =_\tau r[t]}$$

[4]We denote by EXT the axiom schema $\bigcup_{\sigma\in\mathbf{T}}\{\mathsf{EXT}^\rho\}$.

where A_0 is a quantifier-free formula and s^ρ, t^ρ and $(r[x^\rho])^\tau$ are terms in the language of $\mathsf{WE\text{-}HA}^\omega$. The classical theories $\mathsf{E\text{-}PA}^\omega$ and $\mathsf{WE\text{-}PA}^\omega$ are obtained from $\mathsf{E\text{-}HA}^\omega$ and $\mathsf{WE\text{-}HA}^\omega$, respectively, by adding the axiom scheme LEM for all formulas in the language of $\mathsf{WE\text{-}HA}^\omega$.

In Chapter 3 we shall see that functional interpretation does not interpret the extensionality axiom, while it is sound for the quantifier-free extensionality rule. This will make the system $\mathsf{WE\text{-}HA}^\omega$ particularly suitable for working with functional interpretation. One should note, however, that EXT^ρ, for ρ of type level 1, is nevertheless provable in $\mathsf{WE\text{-}HA}^\omega$. Therefore, via a translation embedding $\mathsf{E\text{-}HA}^\omega$ into $\mathsf{WE\text{-}HA}^\omega$ (cf. Section 3.2), we shall see that functional interpretation can still be applied to a mathematically relevant class of theorems of $\mathsf{E\text{-}PA}^\omega$.

Remark 2.1 *In connection with monotone functional interpretation (see Section 3.4) we shall use the definable inequality predicate $\leq_\sigma$, which is defined inductively as follows*

$$x \leq_{\mathbb{N}} y :\equiv x \leq y,$$

$$x \leq_{\rho\to\tau} y :\equiv \forall z^\rho (xz \leq_\tau yz),$$

where $\leq$ is the definable inequality predicate between numbers.

The set of terms of the systems HA^ω is usually called *Gödel's* T. Subsets of Gödel's T can be obtained by restricting the type level of the higher type recursion axioms allowed. We shall denote by T_n the subclass of terms of T with recursion R_σ restricted to types σ of level $\leq n$.

The systems $\mathsf{N\text{-}HA}^\omega, \mathsf{HA}^\omega, \mathsf{WE\text{-}HA}^\omega$ and $\mathsf{E\text{-}HA}^\omega$ are combinatorially closed, in the sense that for each term $t^\tau[x^\sigma]$ in the language one can construct a new term $(\lambda x.t[x])^{\sigma\to\tau}$ such that

$$(\lambda x.t[x])s^\sigma =_\tau t[s],$$

is provable in the system.

Given a functional symbol F defined via an equation $P(F)$, we shall say that another *functional G, defined via equation $Q(G)$ is primitive recursive in F (in the sense of Gödel) w.r.t. a type structure $\mathcal{S}$ (model of* HA^ω*)*, if $\mathcal{S} \models \exists F\, P(F)$ and there exists a term $t \in \mathsf{T}$ such that $\mathcal{S} \models \forall F(P(F) \to Q(tF))$.

2.2.1 The Majorizability Relation and the Model $\mathcal{M}$

In order to show that the functional interpretation of the extensionality axioms EXT^σ cannot be realized by any functional of Gödel's T, Howard [65] made use of an interesting logical relation on the terms of T, which he called *hereditary majorizability.*

Definition 2.1 ([65]) *The relation x^* maj_σ x (x^* hereditarily majorizes x) between functionals of type σ is defined by induction on σ as follows:*

$$x^* \;\mathsf{maj}_{\mathbb{N}}\; x :\equiv x^* \geq x,$$

$$x^* \ \mathsf{maj}_{\rho\to\tau}\ x :\equiv \forall y^*, y(y^* \ \mathsf{maj}_\sigma\ y \to x^*y^* \ \mathsf{maj}_\tau\ xy),$$

Howard's definition was later extended by Bezem [18] in order to define a model for Spector's bar recursion containing discontinuous functionals.

Definition 2.2 ([18]) *The relation $x^* \ \geq^{\mathsf{m}}_\sigma \ x$ (x^* strongly majorizes x) between functionals of type σ is defined by induction on σ as follows:*

$$x^* \ \geq^{\mathsf{m}}_{\mathbb{N}}\ x :\equiv x^* \geq x,$$

$$x^* \ \geq^{\mathsf{m}}_{\rho\to\tau}\ x :\equiv \forall y^*, y(y^* \ \geq^{\mathsf{m}}_\sigma\ y \to x^*y^* \ \geq^{\mathsf{m}}_\tau\ xy \wedge x^*y^* \ \geq^{\mathsf{m}}_\tau\ x^*y),$$

The next lemma states three important properties of the strong majorizability relation.

Lemma 2.1 *The following are provable in* $\mathsf{WE\text{-}HA}^\omega$

i) $x \ \geq^{\mathsf{m}}_\rho\ y$ *implies* $x \ \geq^{\mathsf{m}}_\rho\ x$,

ii) $x \ \geq^{\mathsf{m}}_\rho\ y \wedge y \geq_\rho z \to x \geq_\rho z$, *($\geq_\rho$ as defined in Remark 2.1)*

iii) for type one objects x^1, i.e. number theoretic functions, the function

$$x^+ := \lambda n. \max_{m\leq n} x(m),$$

definable by $R_{\mathbb{N}}$, always majorizes x.

Moreover, all the closed terms of $\mathsf{WE\text{-}HA}^\omega$ and $\mathsf{E\text{-}HA}^\omega$ have strong majorants, i.e. for each closed term t^σ of e.g. $\mathsf{WE\text{-}HA}^\omega$ there exists a closed term t^* of the same type such that $\mathsf{WE\text{-}HA}^\omega \vdash t^* \ \geq^{\mathsf{m}}_\sigma\ t$.

The structure built by Bezem $\mathcal{M} :\equiv \bigcup_{\sigma\in\mathbf{T}} \mathcal{M}_\sigma$, called the *structure of strongly majorizable functionals*, is defined by simultaneous inductive definition of the sets $\mathcal{M}_\sigma$ and the relations $\geq^{\mathsf{m}}_\sigma \subseteq \mathcal{M}_\sigma \times \mathcal{M}_\sigma$ as follows:

- $x^* \ \geq^{\mathsf{m}}_{\mathbb{N}}\ x :\equiv x^* \geq x$,
- $\mathcal{M}_{\mathbb{N}} :\equiv \mathbb{N}$,
- $x^* \ \geq^{\mathsf{m}}_{\rho\to\tau}\ x :\equiv$

$$x^*, x \in \mathcal{M}_\tau^{\mathcal{M}_\sigma} \wedge \forall y^*, y \in \mathcal{M}_\sigma (y^* \ \geq^{\mathsf{m}}_\sigma\ y \to x^*y^* \ \geq^{\mathsf{m}}_\tau\ x^*y, xy),$$

- $\mathcal{M}_{\sigma\to\tau} :\equiv \{x \in \mathcal{M}_\tau^{\mathcal{M}_\sigma} : \exists x^* \in \mathcal{M}_\tau^{\mathcal{M}_\sigma}\ x^* \ \geq^{\mathsf{m}}_{\sigma\to\tau}\ x\}$.

In Chapters 4 and 8 we show that there are functionals in $\mathcal{M}$ satisfying the equations of a new form of bar recursion, so-called *modified bar recursion*, which we introduce in Chapter 8.

2.2.2 Semi-classical Theories

As we saw above, by adding the law of excluded middle LEM to an intuitionistic system such as $\mathsf{WE\text{-}HA}^\omega$ we obtain the classical system $\mathsf{WE\text{-}PA}^\omega$. There are, however, other principles which are not intuitionistically valid but are weaker than LEM.

Two of such principles are the *independence of premise for* $\exists$*-free premises*

$$\mathsf{IP}^\rho_{\mathsf{ef}} \;:\; (A_{\mathsf{ef}} \to \exists x^\rho B(x)) \to \exists x^\rho (A_{\mathsf{ef}} \to B(x)),$$

A_{ef} being $\exists$-free, and *Markov principle*

$$\mathsf{MP}^\sigma \;:\; \neg\neg\exists x^\sigma A_0(x) \to \exists x A_0(x),$$

where A_0 is quantifier free. A variation of $\mathsf{IP}_{\mathsf{ef}}$ is obtained if the premise is required to be purely universal

$$\mathsf{IP}^{\sigma,\rho}_\forall \;:\; (\forall y^\sigma A_0(y) \to \exists x^\rho B(x)) \to \exists x^\rho(\forall y A_0(y) \to B(x)),$$

where A_0 is quantifier free. In systems with decidable atomic formulas $\mathsf{IP}^{\sigma,\rho}_\forall$ is indeed a special case of $\mathsf{IP}^\rho_{\mathsf{ef}}$ (cf. Footnote 3). Those axioms give rise to the following axiom schema:

$$\mathsf{IP}_{\mathsf{ef}} :\equiv \textstyle\bigcup_{\rho\in\mathbf{T}}\{\mathsf{IP}^\rho_{\mathsf{ef}}\},$$

$$\mathsf{IP}_\forall :\equiv \textstyle\bigcup_{\sigma,\rho\in\mathbf{T}}\{\mathsf{IP}^{\sigma,\rho}_\forall\},$$

$$\mathsf{MP} :\equiv \textstyle\bigcup_{\sigma\in\mathbf{T}}\{\mathsf{MP}^\sigma\}.$$

We shall use those principles mainly in order to illustrate, in Chapter 3, how the various proof interpretation of intuitionistic theories can be extended to semi-classical theories in a simple and elegant way.

2.3 Formal Subsystems of Analysis

We obtain subsystems of analysis by extending the arithmetical systems of higher type presented above with analytical principles stating the existence of higher order objects. The two main principles we shall consider are: weak König's lemma, which enables one to create an infinite path in an infinite binary tree, and comprehension, which creates the characteristic function for a certain class of definable sets in the theory. The comprehension shall be obtained, classically, via the schema of choice.

In order to obtain a basic system for analysis, which contains RCA_0, we extend the systems $\mathsf{E\text{-}PA}^\omega$ and $\mathsf{WE\text{-}PA}^\omega$ with a special case of choice

$$\mathsf{QF\text{-}AC}^{\sigma,\tau} \;:\; \forall x^\sigma \exists y^\tau A_0(x,y) \to \exists f \forall x A_0(x, fx).$$

for *quantifier-free* formulas A_0. It is easy to show that e.g. $\mathsf{E\text{-}PA}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ already proves comprehension for Δ^0_1-definable sets. In fact, $\mathsf{E\text{-}PA}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ is much stronger than RCA_0 since it contains induction for arbitrary formulas. For a 'minimal' conservative extension of RCA_0 to all finite types one can take Feferman's [46] system $\widehat{\mathsf{E\text{-}PA}}^\omega\!\restriction$ ($\mathsf{E\text{-}PA}^\omega$ with recursion restricted to type zero and induction restricted to Σ^0_1-formulas) plus $\mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$, which in the presence of classical logic proves Δ^0_1-comprehension (see also [86]).

2.3.1 Weak König's Lemma

Weak König's lemma WKL states that every infinite binary branching tree has an infinite path. In order to fomalize this principles in the language of finite types we shall make use of primitive recursive function(al)s for coding of finite sequences of numbers (defined e.g. in [160]). Those are

- $|n|$, the length of n as a finite sequence,
- $n * m$, the concatenation of n and m,
- $\langle x \rangle$, the sequence of one element x,
- $\Phi(f, n) \equiv \overline{f}n$ which, for $f : \mathbb{N} \to \mathbb{N}$, gives the initial segment of f of length n, viewing f as an infinite sequence.

Trees shall be represented via their characteristic function, i.e. to say that a branch s belong to a tree g shall be formally stated as $g(s) = 0$. An infinite path is represented by a function $f : \mathbb{N} \to \mathbb{N}$. Therefore, weak König's lemma can be defined as (see [157])

$$\mathsf{WKL} \quad : \quad \forall g(\mathsf{Tree}(g) \wedge \mathsf{Bin}(g) \wedge \mathsf{Infinite}(g) \to \exists f \leq \lambda x.1 \forall n(g(\overline{f}n) = 0)),$$

where

$$\mathsf{Tree}(g) :\equiv \forall n, m(g(n * m) = 0 \to g(n) = 0),$$

$$\mathsf{Bin}(g) :\equiv \forall n, x(g(n * \langle x \rangle) = 0 \to x \leq 1),$$

$$\mathsf{Infinite}(g) :\equiv \forall n \exists s(|s| = n \wedge g(s) = 0).$$

The importance of the system $\mathsf{E\text{-}PA}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} + \mathsf{WKL}$ is that it includes, modulo the representation of sets via their characteristic functions, the system of reverse mathematics WKL (and hence WKL_0), and therefore proves among others: (cf. [153])

- Heine/Borel theorem (in sequential form) for $[0, 1]$,
- Every continuous function on the unit interval is uniformly continous,
- Every continuous function on the unit interval attains its infimum and supremum,
- Cauchy/Peano existence theorem,
- Gödel's compactness and completeness theorems.

As mentioned in the introduction, WKL (and consequently all the theorems listed above) are non-computational, in the sense that they are not valid in a model where all the functions/real numbers are computable. Nevertheless, it is well-known that, over various systems of analysis, the use of WKL for proving a Π^0_2-theorem can always be avoided. This was first proved ineffective, via non-constructive model theoretic arguments, by H. Friedman (and extended to Π^1_1-conservation by L. Harrington). Effective proofs were later given by Sieg [149] and Kohlenbach [89]. Kohlenbach's proof combines Gödel's functional interpretation with Howard's hereditary majorizability for functionals in all finite types and covers a whole class of analytical principles which includes WKL.

Following Kohlenbach's approach, we have analyzed Cheney's proof of Jackson's theorem, which is based on WKL. In this chapter we include the full proof given by Cheney as an example of a concrete use of WKL for proving a $\forall\exists$-theorem. The analyzes of the proof is given in Chapter 6.

Example: Jackson's Theorem

The use of WKL in Cheney's proof comes in the form of *the infimum of a strictly positive continuous functions on a closed interval is positive.* This is used in the main lemma of the proof.

Lemma 2.2 (Lemma 1, [33]) *Let f and h be elements of $C[0,1]$. If f has at most a finite number of roots and if $\int_0^1 h\,\mathsf{sgn}f \neq 0$, then for some λ,*

$$\int_0^1 |f - \lambda h| < \int_0^1 |f|.$$

Proof. Assume that $\int_0^1 h\,\mathsf{sgn}f > 0$ (in the opposite case we would take λ with different sign). If $x_1, \ldots, x_k$ are all the roots of f which lie in the open interval (0, 1), define

$$A = [r, x_1 - r] \cup [x_1 + r, x_2 - r] \cup \ldots \cup [x_k + r, 1 - r]$$

and $B = [0,1] \backslash A$, i.e.

$$B = [0, r) \cup (x_1 - r, x_1 + r) \cup \ldots \cup (x_k - r, x_k + r) \cup (1 - r, 1].$$

Take r small enough such that A consists of $k+1$ nondegenerate closed intervals and

$$(i) \quad \int_A h\,\mathsf{sgn}f > \int_B |h|.$$

Since A is closed and contains no roots of f, the number

$$\delta = \min\{|f(x)| : x \in A\}$$

is positive. Take λ such that $0 < \lambda||h||_\infty < \delta$. Then, for points in A we have $|\lambda h(x)| < \delta \leq |f(x)|$, and consequently, on A, (ii) $\mathsf{sgn}(f - \lambda h) = \mathsf{sgn}f$. Thus we have

$$\begin{array}{rcl}
\int_0^1 |f-\lambda h| & = & \int_B |f-\lambda h| + \int_A |f-\lambda h| \\
& \overset{(ii)}{=} & \int_B |f-\lambda h| + \int_A (f-\lambda h)\ \mathsf{sgn} f \\
& = & \int_B |f-\lambda h| + \int_A |f| - \lambda \int_A h\ \mathsf{sgn} f \\
& = & \int_B |f-\lambda h| - \int_B |f| + \int_0^1 |f| - \lambda \int_A h\ \mathsf{sgn} f \\
& \leq & \lambda \int_B |h| - \lambda \int_A h\ \mathsf{sgn} f + \int_0^1 |f| \\
& \overset{(i)}{<} & \int_0^1 |f|. \quad \Box
\end{array}$$

Having this lemma at hand, the proof of Jackson's theorem is mathematically quite elementary. We note that Jackson's theorem talks about the approximation of continuous functions by elements of some arbitrary Haar space[5]. In doing the analysis we have restricted our attention to the concrete Haar space of polynomials of bounded degree P_n (polynomials of degree $\leq n$). This is useful for carrying out the extraction of a concrete modulus of uniqueness, since optimizations can be performed by making use of specific properties of the space under consideration. Therefore, we also present here Cheney's proof of Jackson's theorem applied to the specific Haar space P_n (of dimension $n+1$).

Theorem 2.1 (Jackson's uniqueness theorem, [70]) *Each $f \in C[0,1]$ possesses a unique best approximation in the mean from P_n.*

Proof [33]. For the sake of contradiction, suppose that f has two best approximations from P_n: p_1 and p_2. Then by the triangle inequality for the L_1-norm, the polynomial $p = \frac{1}{2}(p_1 + p_2)$ is also a best approximation of f. Consequently

$$\int_0^1 (|f-p| - \frac{1}{2}|f-p_1| - \frac{1}{2}|f-p_2|) = 0.$$

Since the integrand is continuous and ≤ 0, it must vanish (identically) in $[0,1]$. If $f-p$ has $n+1$ roots, then $f-p_1$, $f-p_2$ and p_1-p_2 must have the same $n+1$ roots, and, $p_1 = p_2$.

Suppose, therefore, that the function $f_0 = f - p$ has at most n roots. Then there exist points $0 = x_0 < x_1 < \ldots < x_{n+1} = 1$ containing among them all the roots of f_0. By Lemma 1, the expression

$$\int_0^1 h\ \mathsf{sgn} f_0 = \Sigma_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h = \Sigma_{i=1}^{n+1} \sigma_i \phi_i(h)$$

[5] An n-dimensional subspace P in $C[0,1]$ is called a Haar subspace if 0 is the only element of P which has n (or more) roots in $[0,1]$.

must vanish for all $h \in P_n$, for otherwise we can reach the contradiction $\int_0^1 |f_0 - \lambda h| < \int_0^1 |f_0|$ by appropriately choosing $h \in P_n$ and λ. On the other hand, we can easily build a non-zero polynomial h which has the signs σ_i in the intervals $[x_{i-1}, x_i]$, which implies that $\int_0^1 h\, \mathsf{sgn} f_0$ is positive, also a contradiction. $\Box$

2.3.2 Comprehension and Choice

Stronger systems of analysis are obtained by extending $\mathsf{E\text{-}PA}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ and $\mathsf{WE\text{-}PA}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ with stronger forms of comprehension, e.g. *full comprehension over numbers*

$$\mathsf{CA} \quad : \quad \exists f \forall n^{\mathbb{N}} (f(n) = 0 \leftrightarrow A(n)),$$

where A is an arbitrary formula, or *arithmetical comprehension over numbers*

$$\mathsf{CA}_{\mathsf{ar}} \quad : \quad \exists f \forall n^{\mathbb{N}} (f(n) = 0 \leftrightarrow A_{\mathsf{ar}}(n)),$$

where A_{ar} is an arithmetical formula, which may contain parameters of higher type. Those can be obtained, in the presence of classical logic, from the schema of *countable choice for numbers*

$$\mathsf{AC}^{\mathbb{N},\mathbb{N}} \quad : \quad \forall n \exists m A(n,m) \to \exists f \forall n A(n, f(n)),$$

A being an arbitrary formula, and from *arithmetical countable choice for numbers*,

$$\mathsf{AC}_{\mathsf{ar}}^{\mathbb{N},\mathbb{N}} \quad : \quad \forall n \exists m A_{\mathsf{ar}}(n,m) \to \exists f \forall n A_{\mathsf{ar}}(n, f(n)),$$

where A_{ar} is an arithmetical formula, respectively. These are special cases of *arbitrary choice*

$$\mathsf{AC}^{\sigma,\tau} \; : \; \forall x^\sigma \exists y^\tau A(x,y) \to \exists Y^{\sigma \to \tau} \forall x A(x, Y(x)),$$

where the formulas A is arbitrary.

Whereas the addition of WKL to $\mathsf{E\text{-}PA}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ does not have any impact on the class of Π^0_2-theorems (with function parameters), the extension by $\mathsf{AC}^{\mathbb{N},\mathbb{N}}$ gives already full analysis. As we pointed out in the introduction, Spector [154] extended Gödel's functional interpretation of arithmetic to analysis by countable choice, and even dependent choice

$$\mathsf{DC}^\rho \quad : \quad \forall n, x^\rho \, \exists y^\rho \, A(n,x,y) \to \forall x \exists f \, (f(0) = x \wedge \forall n \, A(n, f(n), f(n+1))),$$

via the schema of *bar recursion.* This interpretation provides a procedure for giving programs realizing Π^0_2-theorem (with function parameter) in classical analysis by means of bar recursion. We shall now give an example of a proof of a Π^0_2-theorem (with a function parameter), due to Avigad [4], which makes use of a simple form of arithmetical comprehension. In Section 4.1.1 we analyze this proof and present a bar recursive program realizing the theorem.

Example: Avigad's Theorem

We shall deviate a bit from our notation by, in the following, taking σ and τ to denote finite partial functions from $\mathbb{N}$ to $\mathbb{N}$, i.e. partial functions which are defined on a finite domain. A partial function which is everywhere undefined is denoted by $\langle\ \rangle$, while a partial function defined only at position a (with value b) is denoted by $\langle a, b\rangle$. The finite partial functions can be viewed as finite sequences of pairs of natural numbers. For a given finite partial function σ, we define $\hat{\sigma}$ as the total function which is obtained from σ by defining the output to be 0 (zero) wherever σ is undefined. We say that τ extends σ, written as $\sigma \sqsubseteq \tau$, if τ is defined wherever σ is defined, and on those points they coincide in value. The domain of σ, written as $\mathsf{dom}(\sigma)$, is defined as usual. For a finite partial function σ and $a, b \in \mathbb{N}$ we define the finite partial function $\sigma \oplus \langle a, b\rangle$ which maps a to b and agrees with σ otherwise, i.e.

$$(\sigma \oplus \langle a, b\rangle)(x) := \begin{cases} b & \textsf{if } x = a \\ s(x) & \textsf{if } x \neq a \wedge x \in \mathsf{dom}(\sigma) \\ \uparrow & \textsf{otherwise.} \end{cases}$$

Let F be a *continuous* functional (in the sense of the Baire space) of type $(\mathbb{N} \to \mathbb{N}) \to (\mathbb{N} \times \mathbb{N})$. We say that F is a unary *update procedure* if whenever $F(\hat{\sigma}) = \langle a, b\rangle$, τ extends $\sigma \oplus \langle a, b\rangle$ and $F(\hat{\tau}) = \langle a, c\rangle$ then $b = c$.

Theorem 2.2 ([4]) *Every unary update procedure has a finite fixed point, i.e.*

$$\forall F(\mathsf{Update}(F) \to \exists\sigma(\sigma = \sigma \oplus F(\hat{\sigma}))).$$

Proof. Define the sequence of partial functions $\sigma_0, \sigma_1, \sigma_2, \ldots$ as $\sigma_0 = \langle\ \rangle$ and $\sigma_{i+1} = \sigma_i \oplus F(\hat{\sigma}_i)$. The fact that F is an update procedure implies that $\sigma_0 \sqsubseteq \sigma_1 \sqsubseteq \sigma_2 \ldots$. Let g be the partial function extending all the σ_i, that is $g := \bigcup_{i \in \mathbb{N}} \sigma_i$.The continuity of F implies that for some i we have

$$F(\hat{g}) = F(\hat{\sigma}_i) = F(\hat{\sigma}_{i+1}) = \ldots.$$

But then $\sigma_{i+1} = \sigma_i \oplus F(\hat{\sigma}_i) = \sigma_{i+1} \oplus F(\hat{\sigma}_{i+1})$. So, σ_{i+1} is the desired fixed point. $\square$

Comprehension is used in the proof above in order to obtain the function $\hat{g}$ as

$$\hat{g}(a) := \begin{cases} b & \textsf{if } \exists i(\langle a, b\rangle \in \sigma_i) \\ 0 & \textsf{otherwise.} \end{cases}$$

2.3.3 Feasible Analysis

One can also obtain an interesting fragment of analysis by restricting recursion and the amount of induction allowed to the class of so-called Σ^b_1-formulas, i.e. formulas of the form $\exists x \leq t A_0(x)$, A_0 being a quantifier-free formula (which in the system means a polynomial-time computable predicate). With an appropriate choice of the basic functions, even in the presence of Δ^0_1-comprehension, the

Π^0_2-theorems of such a system can be realized by polynomial-time computable functions. Or in other words, this gives a basic subsystem of analysis whose provably recursive functions are polynomial-time computable.

The first basic theory of feasible analysis[6] BTFA was defined by F. Ferreira [48]. The system BTFA is a second order system having as standard model the set of finite 0-1 sequences. Besides (an equivalent form of) Σ^b_1-induction and Δ^0_1-comprehension, Ferreira's system contains also *bounded collection principle* for arbitrary bounded formulas. As done by Friedman and Harrington for RCA_0, Ferreira showed non-constructively that a feasible version[7] of WKL can always be avoided in proofs of Π^0_2-theorems with function parameters (and even Π^1_1-theorems) over BTFA.

In Chapter 9, we shall define a new basic theory for feasible analysis based on the language of finite types. Our system is basically an extension of Cook and Urquhart [38] CPV^ω with quantifier-free choice $\mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$, which also enables us to prove Δ^0_1-comprehension. We use this system in order to show effectively (via functional interpretation) that WKL-proofs of Π^0_2-theorems in feasible analysis have polynomial-time computable realizers. The full description of the systems of feasible analysis which we have used, as well as the interpretation of WKL in the feasible setting, can be found in Chapter 9.

[6] BTFA can be viewed as an extension of Buss' theory S^1_2 (cf. [31]).

[7] The predicate $\mathsf{Infinite}(g)$ in the definition of WKL, i.e.

$$\forall n \exists s(|s| = n \wedge g(s) = 0),$$

states the existence of a finite branch of length n, for any given n. Since we shall measure the length of an input n as $\log n$, we see that any function producing s on input n has to run in exponential time. This implies that the premise $\mathsf{Infinite}(g)$ cannot ever be proven in a feasible theory, making WKL quite weak. Therefore, in order to strengthen WKL, in Chapter 9 (following [48]) we have weakened $\mathsf{Infinite}(g)$ to

$$\forall n \exists s(|s| = |n| \wedge g(s) = 0).$$

Chapter 3

Proof Interpretations and Translations

After having introduced concrete formal system in the last chapter, we are the position to precisely define the three proof interpretations which are going to be used in this dissertation: functional interpretation, monotone functional interpretation and modified realizability. Recall from Section 1.1.2 that a proof interpretation of a system $\mathcal{T}$ is a triple $(\mathcal{I}, \mathcal{F}, \mathcal{A})$ consisting of an interpretation $\mathcal{I}$ (in the sense of Kreisel), an enumeration of a set of functionals $\mathcal{F}$ and an algorithm $\mathcal{A}$. The interpretation $\mathcal{I}$ associates each formula A of $\mathcal{T}$ to a sequence, indexed by the enumeration of $\mathcal{F}$, of potential interpretations of A. The algorithm $\mathcal{A}$ provides the machinery for finding the right interpretation of A once a proof of A is provided.

In the case of the three interpretations mentioned above, the associated algorithm shall be provided by the *soundness* proof for the interpretation. Moreover, the class of functionals used consists of the functionals in Gödel's T. The proof interpretations apply directly to any of the intuitionistic arithmetical systems introduced in Chapter 2. Combined with the negative translation[1] (and in the case of modified realizability also the A-translation) those interpretation apply to the classical arithmetical theories as well. The interpretation of subsystems of analysis shall be obtained via an interpretation of principles such as weak König's lemma and different forms of the choice axiom.

In order to choose which interpretation to use in a given context, it is useful to look at the following five general features:

(I) the basic theories $\mathcal{T}$ to which the proof interpretation applies and how the interpretation scales to subsystems and extensions of $\mathcal{T}$,

(II) the class of formulas which, when given together with a proof, the proof interpretation gives further information,

(III) the class of formulas which, according to the interpretation, have trivial realizers,

(IV) the modularity of the algorithm $\mathcal{A}$,

(V) the complexity of the algorithm $\mathcal{A}$.

[1] The proof translations we shall use can also be viewed as proof interpretations in which $\mathcal{I}$ associates each formula A to a unique formula A', and $\mathcal{F}$ is the empty set.

Although, in the following, we shall discuss some of these points, we refrain from giving a detailed study of the relation between the different proof interpretations. The reader should pay attention, however, to the conflict between points (II) and (III), since usually the more formulas are considered as empty of information by the interpretation, the smaller the class of formulas on which the interpretation gives further information. Since all the proof translations and interpretations presented in this chapter are modular and have low complexity (polynomial of low degree on the size of the input proof, cf. [63]), the points (IV) and (V) only come into play when structural techniques such as normalization and cut-elimination are used (cf. discussion in Section 3.3.1). For a given proof interpretation $(\mathcal{I}, \mathcal{F}, \mathcal{A})$, we shall write

$$\mathcal{T} \vdash A \quad \Longrightarrow \quad \mathcal{S} \vdash \mathcal{I}(A)_F$$

to mean that from a proof of A in the system $\mathcal{T}$, one can effectively, via the algorithm $\mathcal{A}$, produce a functional $F \in \mathcal{F}$ and a proof of the formula $\mathcal{I}(A)_F$ in the system $\mathcal{S}$ (usually a subsystem of $\mathcal{T}$).

3.1 Negative Translation

The negative translation was discovered independently by Kolmogorov, Gödel and Gentzen, and provides a way of embedding classical theories into their intuitionistic counterpart. We shall use in this dissertation a version of the negative translation due to Kuroda [121]. In the following, let $\mathcal{T}_{\mathsf{i}}$ be any of the intuitionistic arithmetical theories introduced in Chapter 2 and $\mathcal{T}_{\mathsf{c}}$ its classical counterpart.

Definition 3.1 ([121, 124]) *The negative translation of $A \in \mathcal{L}(\mathcal{T}_{\mathsf{c}})$ is defined as $(A)^N :\equiv \neg\neg(A)^*$, where $(A)^*$ is defined by induction on the logical structure of A*

$(A)^* :\equiv A$, *if A is a atomic formula,*

$(\neg A)^* :\equiv \neg(A)^*$,

$(A \Box B)^* :\equiv A^* \Box B^*$, *where* $\Box \in \{\wedge, \vee, \rightarrow\}$,

$(\exists x A(x))^* :\equiv \exists x (A(x))^*$,

$(\forall x A(x))^* :\equiv \forall x \neg\neg(A(x))^*$.

Notice that the negative translation of a formula A consists of placing double negations after each universal quantifier of A and in front of the whole formula.

Theorem 3.1 (Soundness of negative translation, [124]) *For any formula A and set of formulas Δ in the language of $\mathcal{T}_{\mathsf{c}}$, the following holds*

$$\mathcal{T}_{\mathsf{c}} + \Delta \vdash A \quad \Longleftrightarrow \quad \mathcal{T}_{\mathsf{i}} + \Delta^N \vdash A^N,$$

where $\Delta^N :\equiv \{B^N : B \in \Delta\}$.

We shall make use of negative translation in Chapters 5, 6 and 9 in connection with functional interpretation and monotone functional interpretation; and in Chapter 8 together with modified realizability. For further information on the negative translation see [56, 57, 87, 108].

3.2 Elimination of Extensionality

In this section we describe a procedure for eliminating the axiom of extensionality from proofs, as developed in [124]. The procedure can be used e.g. to translate a proof of a theorem A in the system $\mathsf{E\text{-}HA}^\omega$ into a proof of a variant $(A)^e$ of A in the system $\mathsf{WE\text{-}HA}^\omega$ (and even in HA^ω). The translation is such that, if the type level of the variables in A are not greater than 1 then $(A)^e \leftrightarrow A$ is provable in $\mathsf{WE\text{-}HA}^\omega$. Let us first simultaneously define, for each finite type σ, an extensionality predicate $\mathsf{Ex}^\sigma(t)$ and a hereditarily extensional equality $t =^e_\sigma s$ on terms (assume σ has the form $\sigma_1 \to \dots \sigma_n \to \mathbb{N}$):

- $\mathsf{Ex}^{\mathbb{N}}(t^{\mathbb{N}}) :\equiv \mathsf{true}$,

  $t =^e_{\mathbb{N}} s :\equiv t =_{\mathbb{N}} s$,

- $\mathsf{Ex}^\sigma(t^\sigma) :\equiv \forall \underline{x}, \underline{y}(\bigwedge_{i=1}^n x_i =^e_{\sigma_i} y_i \to t\underline{x} =_{\mathbb{N}} t\underline{y})$,

  $t =^e_\sigma s :\equiv \mathsf{Ex}^\sigma(t) \wedge \mathsf{Ex}^\sigma(s) \wedge \forall \underline{x}(\bigwedge_{i=1}^n \mathsf{Ex}^{\sigma_i}(x_i) \to t\underline{x} =_{\mathbb{N}} s\underline{x})$.

It is easy to show that, the schema of extensionality $\cup_{\rho \in \mathbf{T}} \mathsf{EXT}^\rho$ is equivalent to $\cup_{\rho\in\mathbf{T}} \forall x^\rho \mathsf{Ex}^\rho(x)$, and that for each closed term t^ρ in the language of $\mathsf{WE\text{-}HA}^\omega$, $\mathsf{Ex}^\rho(t)$ is provable in $\mathsf{WE\text{-}HA}^\omega$. The translation $(A)^e$ below makes all the quantifiers of A to range over extensional objects.

Definition 3.2 (E-translation) *To each formula $A \in \mathcal{L}^\omega$ we associate a formula $(A)^e$ inductively in the structure of A as follows*

$(A)^e :\equiv A$, *for atomic formulas A,*

$(A \Box B)^e :\equiv (A)^e \Box (B)^e$, *where $\Box \in \{\wedge, \vee, \to\}$,*

$(\forall x^\sigma A(x))^e :\equiv \forall x(\mathsf{Ex}^\sigma(x) \to (A(x))^e)$,

$(\exists x^\sigma A(x))^e :\equiv \exists x(\mathsf{Ex}^\sigma(x) \wedge (A(x))^e)$.

The E-translation is such that, under the assumption that the free-variables of a theorem A are extensional, $(A)^e$ is also provable.

Lemma 3.1 ([124], Thm. 2.14) *If $\mathsf{PA}^\omega \vdash A(\underline{x})$ then $\mathsf{PA}^\omega \vdash \mathsf{Ex}(\underline{x}) \to (A(\underline{x}))^e$, where $\underline{x}$ are all the free-variables in A.*

The next lemma provides the main result for obtaining the elimination of extensionality theorem below.

Lemma 3.2 ([124], Thm. 2.16) *Let S_1 and S_2 be sets of closed formulas in $\mathcal{L}^\omega$. If* *(i)* $\mathsf{PA}^\omega + S_1 \vdash (A)^e$, *for all* $A \in S_1 \cup S_2$ *and*

(ii) $\mathsf{PA}^\omega + S_1 + S_2 \vdash \mathsf{EXT}^\sigma$, *for all* $\sigma \in \mathbf{T}$

then $\mathsf{PA}^\omega + S_1 + S_2 \vdash C(\underline{x}) \iff \mathsf{PA}^\omega + S_1 \vdash \mathsf{Ex}(\underline{x}) \to (C(\underline{x}))^e$.

By taking S_2 to be the axiom schema of extensionality $\cup_{\rho\in\mathbf{T}}\mathsf{EXT}^\rho$ one gets the following theorem, which allows for an effective elimination of the extensionality axiom.

Theorem 3.2 (Elimination of extensionality, [124]) *Let A be any closed formula for which* $\mathsf{WE}\text{-}\mathsf{PA}^\omega + A \vdash (A)^e$. *Then*[2]

$$\mathsf{E}\text{-}\mathsf{PA}^\omega + A \vdash C(\underline{x}) \iff \mathsf{WE}\text{-}\mathsf{PA}^\omega + A \vdash \mathsf{Ex}(\underline{x}) \to (C(\underline{x}))^e.$$

3.3 Functional Interpretation

The functional interpretation[3] was developed by Gödel with the purpose of reducing the consistency of HA (and, via negative translation, also of PA) to the consistency of the quantifier-free calculus T. It was shortly after observed [162] (cf. also the review with corrections in [156]) that the interpretation extends to appropriate systems based on finite types such as $\mathsf{WE}\text{-}\mathsf{HA}^\omega$ (and via negative translation to $\mathsf{WE}\text{-}\mathsf{PA}^\omega$). The idea of functional interpretation is to associate each formula A in the language of e.g. $\mathsf{WE}\text{-}\mathsf{HA}^\omega$ to a formula of the form $\exists\underline{x}\forall\underline{y}A_D(\underline{x},\underline{y})$, for a particular quantifier-free formula A_D. That is done inductively on the logical structure of the formula A, the most interesting passage being the treatment of implication. Suppose one has a formula A of the form $B \to C$, and suppose that B and C have already been brought into the $\exists\forall$-form

$$\exists\underline{x}\forall\underline{y}B_D(\underline{x},\underline{y}) \to \exists\underline{u}\forall\underline{v}C_D(\underline{u},\underline{v}).$$

How should one transform such a formula into an $\exists\forall$-formula? The idea of Gödel is to choose among the four possible prenexations of such a formula the *'least' non-constructive one*, which uses only $\mathsf{IP}_\forall$ and MP,

$$\forall\underline{x}\exists\underline{u}\forall\underline{v}\exists\underline{y}(B_D(\underline{x},\underline{y}) \to C_D(\underline{u},\underline{v})),$$

from which, by AC one gets

$$\exists\underline{U},\underline{Y}\forall\underline{x},\underline{v}(B_D(\underline{x},\underline{Yxv}) \to C_D(\underline{Ux},\underline{v})),$$

which is then taken to be the interpretation of A. The full interpretation is described as follows.

Definition 3.3 (Functional interpretation, [58]) *With each formula $A \in \mathcal{L}^\omega$ associate a formula $(A)^D :\equiv \exists\underline{x}\forall\underline{y}A_D(\underline{x},\underline{y})$ in $\mathcal{L}^\omega$, where A_D is quantifier free and $FV(A^D) = FV(A)$, defined inductively in the following way:*

[2] Examples of such formula A are the universal closure of $\mathsf{QF}\text{-}\mathsf{AC}^{\mathbb{N},\mathbb{N}\to\mathbb{N}}$, $\mathsf{QF}\text{-}\mathsf{AC}^{\mathbb{N}\to\mathbb{N},\mathbb{N}}$, $\mathsf{AC}^{\mathbb{N},\tau}$ or DC^ρ, which we have used in Section 2.3 to build systems of analysis, see [124].

[3] Also called *Dialectica interpretation* or *D-interpretation*.

$(A)^D :\equiv A_D :\equiv A$, *for atomic formulas* A,

and if $(A)^D :\equiv \exists \underline{x} \forall \underline{y} A_D(\underline{x}, \underline{y})$ *and* $(B)^D :\equiv \exists \underline{u} \forall \underline{v} B_D(\underline{u}, \underline{v})$ *we have,*

$$(A \wedge B)^D :\equiv \exists \underline{x}, \underline{u} \forall \underline{y}, \underline{v} (A_D(\underline{x}, \underline{y}) \wedge B_D(\underline{u}, \underline{v})),$$

$$(A \vee B)^D :\equiv \exists z^{\mathbb{N}}, \underline{x}, \underline{u} \forall \underline{y}, \underline{v} ((z = 0 \to A_D(\underline{x}, \underline{y})) \wedge (z \neq 0 \to B_D(\underline{u}, \underline{v}))),$$

$$(A \to B)^D :\equiv \exists \underline{U}, \underline{Y} \forall \underline{x}, \underline{v} (A_D(\underline{x}, \underline{Y}\, \underline{x}\, \underline{v}) \to B_D(\underline{U}\, \underline{x}, \underline{v})),$$

$$(\exists z^\rho A(z))^D :\equiv \exists z, \underline{x} \forall \underline{y} A_D(\underline{x}, \underline{y}, z),$$

$$(\forall z^\rho A(z))^D :\equiv \exists \underline{X} \forall z, \underline{y} A_D(\underline{X} z, \underline{y}, z).$$

The main feature of the functional interpretation is that, given a proof of a formula $A(\underline{z})$, having functional interpretation $\exists \underline{x} \forall \underline{y} A(\underline{x}, \underline{y}, \underline{z})$, in $\mathsf{WE\text{-}HA}^\omega$ we can effectively extract closed terms $\underline{t}$ such that $\mathsf{WE\text{-}HA}^\omega \vdash \forall \underline{y} A_D(\underline{t z}, \underline{y}, \underline{z})$. Moreover, the interpretation is modular in the sense that once new axioms are added, all we need to do is to provide the fulfilled f.i. of the axioms for the verification of the extracted term.

Theorem 3.3 (Soundness of f.i., [58, 160]) *Let* Δ *be an arbitrary set of closed formulas. Assume that for each formula* B *of* Δ *(say* $B^D \equiv \exists \underline{v} \forall \underline{w} B_D(\underline{v}, \underline{w})$*) we are given a tuple of closed terms* $\underline{r} \in \mathsf{T}$*. Let*

$$\Delta^D :\equiv \{\forall \underline{w} B_D(\underline{r}, \underline{w}) : B \in \Delta\},$$

and $A(\underline{z}) \in \mathcal{L}^\omega$*, where* $\underline{z}$ *are all the free-variables of* A*, be such that* $(A)^D \equiv \exists \underline{x} \forall \underline{y} A_D(\underline{x}, \underline{y}, \underline{z})$*. The following rule holds,*

$$\mathsf{WE\text{-}HA}^\omega + \Delta \vdash A(\underline{z}) \quad \Longrightarrow \quad \mathsf{WE\text{-}HA}^\omega + \Delta^D \vdash \forall \underline{y} A_D(\underline{t z}, \underline{y}, \underline{z}),$$

where the tuple of closed terms $\underline{t} \in \mathsf{T}$ *can extracted from the given proof of* $A(\underline{z})$*.*

Particularly interesting instances of Theorem 3.3 can be obtained as follows:

- when Δ consists only of formulas of the form $\forall x^\rho B_0(x)$, in which case $\Delta^D \equiv \Delta$, no terms $\underline{r}$ are necessary, and the verification of the extracted term t takes place in the original system,

- when Δ consists of principles B such that realizations of B^D are provable in $\mathsf{WE\text{-}HA}^\omega$, e.g. $\mathsf{MP}^\sigma, \mathsf{IP}^\rho_\forall$ and $\mathsf{AC}^{\sigma,\tau}$ (cf. [160]), in which case again the verification of the extracted term t takes place in $\mathsf{WE\text{-}HA}^\omega$,

- when A has the form $\forall x^\rho \exists y^\sigma A(x, y)$, in which case the functional interpretation of A implies $\exists Y \forall x (A(x, Y x))^D$ and the theorem guarantees the extraction of a term t satisfying $\forall x (A(x, t x))^D$. Furthermore, since over the system $\mathsf{WE\text{-}HA}^\omega + \mathsf{MP} + \mathsf{IP}_\forall + \mathsf{AC}$ one can prove $A \leftrightarrow (A)^D$, one gets that $\mathsf{WE\text{-}HA}^\omega + \mathsf{MP} + \mathsf{IP}_\forall + \mathsf{AC} \vdash \forall x A(x, t x)$ (cf. [160]).

One can extend the range of applicability of Theorem 3.3, via the negative translation, to classical systems as well.

Theorem 3.4 (Soundness of f.i. for classical systems, [46, 124]) *Let Δ be an arbitrary set of closed formulas and let ($\underline{r}$ given similarly as in Theorem 3.3)*

$$\Delta^{ND} :\equiv \{\forall \underline{w} B_{ND}(\underline{r}, \underline{w}) : B \in \Delta\},$$

with the understanding that $(B)^N$ has functional interpretation $\exists \underline{v} \forall \underline{w} B_{ND}(\underline{v}, \underline{w})$. Moreover, let $A(\underline{z}) \in \mathcal{L}^\omega$, where $\underline{z}$ are free, be such that $(A)^{ND} \equiv \exists \underline{x} \forall \underline{y} A_{ND}(\underline{x}, \underline{y}, \underline{z})$. The following rule holds,

$$\mathsf{WE\text{-}PA}^\omega + \Delta \vdash A(\underline{z}) \quad \Longrightarrow \quad \mathsf{WE\text{-}HA}^\omega + \Delta^{ND} \vdash \forall \underline{y} A_{ND}(\underline{t}\underline{z}, \underline{y}, \underline{z}),$$

where the tuple of closed terms $\underline{t} \in \mathsf{T}$ can be extracted from the given proof of $A(z)$.

Similarly as above, if Δ consists only of formulas of the form $\forall x^\rho B_0(x)$ then $\Delta^{ND} \equiv \Delta$, over $\mathsf{WE\text{-}HA}^\omega$, and if Δ consists of principles B whose functional interpretation *of its negative translation* can be realized and proven[4] in $\mathsf{WE\text{-}HA}^\omega$ then the verification of the term t takes place solely in $\mathsf{WE\text{-}HA}^\omega$. Moreover, if A has the form $\forall x^\rho \exists y^\sigma A_0(x, y)$, A_0 quantifier-free, then $\mathsf{WE\text{-}HA}^\omega \vdash A^{ND} \leftrightarrow \exists Y \forall x A_0(x, Yx)$ and the theorem guarantees the extraction of a term t satisfying $\forall x A_0(x, tx)$.

Finally, we can combine negative translation and functional interpretation with the elimination of extensionality procedure described in Section 3.2 to obtain an interpretation of classical extensional theories.

Theorem 3.5 (Soundness of f.i. for classical extensional systems, [124]) *For an arbitrary set of closed formulas Δ, let ($\underline{r}$ given similarly as in Theorem 3.3)*

$$\Delta^{eND} :\equiv \{\forall \underline{w} B_{eND}(\underline{r}, \underline{w}) : B \in \Delta\},$$

with the understanding that $(B^e)^N$ has functional interpretation $\exists \underline{v} \forall \underline{w} B_{eND}(\underline{v}, \underline{w})$. Moreover, let $A(\underline{z}) \in \mathcal{L}^\omega$, where $\underline{z}$ are all the free-variables in A, be such that $(\mathsf{Ex}(\underline{z}) \to A^e(\underline{z}))^{ND} \equiv \exists \underline{x} \forall \underline{y} A_{eND}(\underline{x}, \underline{y}, \underline{z})$. The following rule holds,

$$\mathsf{E\text{-}PA}^\omega + \Delta \vdash A(\underline{z}) \quad \Longrightarrow \quad \mathsf{WE\text{-}HA}^\omega + \Delta^{eND} \vdash \forall \underline{y} A_{eND}(\underline{t}\underline{z}, \underline{y}, \underline{z}),$$

where the closed terms $\underline{t} \in \mathsf{T}$ can be extracted from the given proof of $A(\underline{z})$.

Notice that, if B is a closed formulas of the form $\forall x^\rho B_0(x)$ then $\mathsf{WE\text{-}HA}^\omega + B \vdash (B)^e$. The same holds for the axioms $\mathsf{QF\text{-}AC}^{\mathbb{N}\to\mathbb{N},\mathbb{N}}$ and $\mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}\to\mathbb{N}}$. Therefore, by the fact that $\mathsf{WE\text{-}HA}^\omega \vdash \mathsf{Ex}^1(x)$ and Lemma 3.2, we can obtain the following special case of the theorem above.

Corollary 3.1 ([87, 124]) *Let*

$\Delta_1 :\equiv$ *arbitrary set of closed formulas of the form $\forall x^\rho B_0(x)$,*

$\Delta_2 :\equiv \{\mathsf{QF\text{-}AC}^{\mathbb{N}\to\mathbb{N},\mathbb{N}}, \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}\to\mathbb{N}}\}$.

[4] An example of such principle is $\mathsf{QF\text{-}AC}^{\rho,\sigma}$.

Moreover, let $A \equiv \forall x^1 \exists y^\tau A_0(x,y)$, $A_0(x,y)$ *being quantifier-free. The following rule holds,*

$$\mathsf{E\text{-}PA}^\omega + \Delta_1 + \Delta_2 \vdash A \qquad \Longrightarrow \qquad \mathsf{WE\text{-}HA}^\omega + \Delta_1 \vdash \forall x^1 A_0(x, tx),$$

where $t \in \mathsf{T}$ *is a closed term which can be extract from the given proof of* A.

Functional interpretation has been adapted to many other theories, e.g. full classical analysis [154] and systems of feasible arithmetic [38]. In Chapter 9, we shall use the development made in [38] in order to give a functional interpretation of weak König's lemma in the setting of feasible analysis. Two examples of the use of functional interpretation as a tool for extracting programs from classical proofs can be found in Section 4.1.1 (proof based on comprehension) and Chapter 6 (proof based on WKL). For applications of functional interpretation to fixed point theory see [100, 101, 103, 105]. For more information on functional interpretation see [5, 87] and Chapter 5.

3.3.1 Functional Interpretation versus n.c.i.

For a formula $A \in \mathcal{L}$ in prenex normal form (let us take

$$A \equiv \forall x \exists y \forall z \exists w A_0(x,y,z,w)$$

as a matter of illustration, A_0 quantifier-free) the no-counterexample interpretation is obtained by looking at the Herbrand normal form of A, which in our setting with function(al) variables can be written as

$$\forall x, f \exists y, w A_0(x, y, fy, w)$$

and asking for functionals realizing that, i.e.

$$\exists \Phi_1, \Phi_2 \forall x, f A_0(x, \Phi_1(x,f), f(\Phi_1(x,f)), \Phi_2(x,f)).$$

This provides the $\mathcal{I}$-component of the no-counterexample interpretation. Moreover, Kreisel took $\alpha(<\varepsilon_0)$-recursive functionals as the set $\mathcal{F}$, and an application of Hilbert's ε-substitution method (due to Ackermann [1]) as the algorithm $\mathcal{A}$. We note here that Kreisel could equivalently have taken $\mathcal{F}$ to be the type level 2 functionals of Gödel's T. In fact, the $\alpha(<\omega_{n+2})$-recursive functionals (via unnested recursion) correspond precisely to the type level 2 functionals of T_n, where $\omega_1 = \omega$ and $\omega_{n+1} = \omega^{\omega_n}$.

As we mentioned in the introduction, also functional interpretation combined with the negative translation can be used as the machinery $\mathcal{A}$ behind the n.c.i.. Since the no-counterexample interpretation of formulas in PA only talks about functionals of type two, the passage through higher types, necessary for using functional interpretation might seem superfluous at first sight. As pointed out in [97], however, for each number n, there are provable PA sentences A and $A \to B$, whose n.c.i. can be given in the first level of Gödel's system T, i.e. by $\alpha(<\omega^\omega)$-recursive functionals, but the n.c.i. of B is not satisfied by any functional in T_n, i.e. by any $\alpha(<\omega_{n+2})$-recursive functional. And if A and

$A \to B$ are not provable in PA one even needs Spector's bar recursion to give a n.c.i. of B. This shows that, for the n.c.i., the procedure g in condition (δ) has to be stronger than any functional in Gödel's T, since it cannot live in any fixed T_n. This means that the n.c.i. cannot be used as a proof interpretation for fragments[5] Σ^0_n-IND of PA by functionals in T_n, simply for the lack of an algorithm g satisfying condition (δ), since g should take hypothetical realizers F_1 and F_2 (in T_n) for A and $A \to B$ into a potential realizer F_3 (also in T_n) for B, such that F_3 is a realizer for B whenever F_1 and F_2 are indeed realizers for A and $A \to B$ respectively.

The problem is that the n.c.i. of A and $A \to B$ can be too weak for giving a direct realizer for B, in which case one has to go back to A by making use of the fact that, *over second order arithmetic*, the n.c.i. of A implies the original formula A. This is, however, an extremely ineffective passage, which leads to the high complexity of g in condition (δ).

Functional interpretation, on the other hand, scales very naturally to fragments of arithmetic, since the treatment of the modus ponens requires only lambda abstraction and application. The difference is that functional interpretation sets up the right induction hypothesis when producing realizers for A and $A \to B$, in such a way that the realizer for B can be obtain straightforwardly. The cost for having this nice behaviour is that higher types must be used, which requires a subsequent normalization procedure for T if one is interested in obtaining $\alpha(< \varepsilon_0)$-recursive functionals. But the advantage is that, following this approach we have clearly separated the local and global features of the proof analysis. This is explored for instance in [97] and in Chapter 9.

3.4 Monotone Functional Interpretation

As mentioned in the previous section, the modularity of functional interpretation allows one to extend the basic system $\mathsf{WE\text{-}HA}^\omega$ with principles such as $\mathsf{MP}, \mathsf{IP}_\forall$ and AC, since their f.i. have realizers in $\mathsf{WE\text{-}HA}^\omega$ (and are verifiable in $\mathsf{WE\text{-}HA}^\omega$). Moreover, since universal formulas have empty realizers they can be always taken as axioms to build more flexible systems. Formulas A of more complex logical structure, however, can only be added when $(A)^D \equiv \exists \underline{x} \forall \underline{y} A(\underline{x}, \underline{y})$ is enriched with realizers for $\underline{x}$, which makes it a universal formula again.

This limit of what can be added "for free" to the system can be expanded if one relaxes the requirement of the information to be obtained. In loose terms, if one is only looking for *bounds* instead of actual realizers, a much larger class of formulas can be taken as axioms without any cost to the extraction procedure.

We shall take as a precise notion of bound Bezem's strong majorizability relation (cf. Section 2.2.1). Let the functional interpretation of $A(\underline{z})$ be $\exists \underline{x} \forall \underline{y} A_D(\underline{x}, \underline{y}, \underline{z})$. Instead of looking for terms $\underline{t}$ which, uniformly in $\underline{z}$, realize $\underline{x}$ in $(A(z))^D$ we only look for terms $\underline{t}$ *majorizing* such $\underline{x}$, i.e.

$$\exists \underline{X} \ \leq^{\mathsf{m}} \ \underline{t} \forall \underline{z}, \underline{y} A_D(\underline{X} \underline{z}, \underline{y}, \underline{z}).$$

[5] The fragment Σ^0_n-IND of PA containing only induction for Σ^0_n formulas.

This variant of functional interpretation, called *monotone functional interpretation* (*m.f.i.* for short), was first introduced in [92]. Notice that by m.f.i. true sentences of the form $B \equiv \forall w^\rho \exists u^\sigma \leq t(x) \forall v^\tau B_0(w,u,v)$ can be freely used in proofs as if they were axioms, since any proof which uses B can also be given using $B^D \equiv \exists U \leq t \forall w^\rho, v^\tau B_0(w, Uw, v)$, whose m.f.i. can be given by any term t^* majorizing t (by Lemma 2.1). This argument works even for classical systems, where one first needs to apply the negative translation, since B^D implies intuitionistically $(B^D)^N$. The class of true sentences of the form B above constitutes a much larger class than just the true $\forall x^\rho B_0(x)$ sentences. It in fact includes ineffective principles such as weak König's lemma, as we shall discuss later.

Theorem 3.6 (Soundness for m.f.i., [87, 92]) *Let Δ be an arbitrary set of closed formulas and let ($\underline{r}$ given similarly as in Theorem 3.3)*

$$\Delta^M :\equiv \{\exists \underline{v} \leq^{\mathsf{m}} \underline{r} \forall \underline{w} B_D(\underline{v}, \underline{w}) : B \in \Delta\},$$

with the understanding that B has functional interpretation $\exists \underline{v} \forall \underline{w} B_D(\underline{v}, \underline{w})$. Moreover, let $A(\underline{z}) \in \mathcal{L}^\omega$, where $\underline{z}$ are all the free-variables of A, be such that $(A(\underline{z}))^D \equiv \exists \underline{x} \forall \underline{y} A_D(\underline{x}, \underline{y}, \underline{z})$. The following rule holds,

$$\mathsf{WE\text{-}HA}^\omega + \Delta \vdash A(\underline{z}) \implies \mathsf{WE\text{-}HA}^\omega + \Delta^M \vdash \exists \underline{X} \leq^{\mathsf{m}} \underline{t} \forall \underline{z}, \underline{y} A_D(\underline{X}\underline{z}, \underline{y}, \underline{z}),$$

where the closed terms $\underline{t}$ can be extracted from the given proof of $A(\underline{z})$.

As shown in [92], the soundness theorem for m.f.i. can be directly proved on the level of the monotone version, i.e. without the need to compute first an actual realizer t as an intermediate step. Using negative translation and elimination of extensionality, we can extend the soundness of m.f.i. to classical extensional theories as well. The main advantages of using the monotone version of functional interpretation are:

- *Wider range of applications*: Various non-effective analytical principles, such as the attainment of the infimum by continuous functions $f \in C[0,1]$, which in logical terms correspond to the non-computational principle weak König's lemma (see Section 2.3.1), have the logical form $\forall w^1 \exists u \leq_1 rw \forall v^{\mathbb{N}} B_0(w,u,v)$. As mentioned above, m.f.i. considers such principles as trivially realizable, and can therefore be added as axioms to the system.

- *Independence of parameters ranging over compact spaces*: Elements of compact spaces are represented in our formal system by bounded type one objects $y^1 \leq s^1$. Monotone functional interpretation guarantees that bounds extracted will be independent of the compact spaces involved, since any majorizing term $t^*(y)$ which depends on y can be turned into a new term $t^*(s^*)$ which does not depend on y, where s^* is a majorant for the term s.

- *Simplicity of the interpretation*: As it is well-known, when using functional interpretation the most intricate axiom schema is $A \to A \wedge A$. Those, however, are trivially handled by monotone functional interpretation.

For more information on monotone functional interpretation see Chapter 5 and [63, 87, 92].

3.5 Modified Realizability

According to the Brouwer-Heyting-Kolmogorov (BHK) constructive interpretation of the logical constants, the notion "the construction p verifies A" or "p is a proof of A" is defined as follows

- there is not proof of $\perp$,
- p is a proof of $A_1 \wedge A_2$ if $p \equiv (p_0, p_1)$ and p_i is a proof of A_i, $i \in \{0,1\}$,
- p is a proof of $A_1 \vee A_2$ if $p \equiv (i, q)$, $i \in \{0,1\}$ and q is a proof of A_i,
- p is a proof of $A \to B$ if p is a construction transforming any proof q of A into a proof $p(q)$ of B,
- p if a proof of $\forall x A(x)$ if p is a construction which, given a construction q_c of an element c, produce a proof $p(q_c)$ of $A(c)$,
- p if a proof of $\exists x A(x)$ if $p \equiv (q_c, r)$, where q_c is the construction of an element c of the domain and r proves $A(c)$,

where for atomic formulas the notion of proof is supposed to be given. Based on the BHK interpretation Kleene [77] gave an interpretation of intuitionistic arithmetic associating to each arithmetical sentence A a notion 'the number n realizes A'. That turned out to be a systematic method of making the constructive content of arithmetical sentences explicit. We present here a version of Kleene's realizability for $\mathsf{E\text{-}HA}^\omega$, called *modified realizability*[6] (*m.r.* for short), which was first formulated by Kreisel [116]. Whereas functional interpretation transforms arbitrary formulas into formulas of the form $\exists x A_\forall(x)$, $A_\forall(x)$ being purely universal, modified realizability provides a way of transforming any formula A of $\mathsf{E\text{-}HA}^\omega$ into a formula of the kind $\exists \underline{x} A_{\mathsf{ef}}(\underline{x})$, where $A_{\mathsf{ef}}(\underline{x})$ is $\exists$-free.

Definition 3.4 (Modified realizability) *With each formula $A \in \mathcal{L}^\omega$ associate a formula $(A)^{\mathsf{mr}} :\equiv \exists \underline{x} A_{\mathsf{mr}}(\underline{x})$ of $\mathsf{E\text{-}HA}^\omega$, where A_{mr} is $\exists$-free and $FV(A^{\mathsf{mr}}) = FV(A)$, defined inductively in the following way:*

$$(A)^{\mathsf{mr}} :\equiv A, \text{ for atomic formulas } A,$$

and if $(A)^{\mathsf{mr}} \equiv \exists \underline{x} A_{\mathsf{mr}}(\underline{x})$ and $(B)^{\mathsf{mr}} \equiv \exists \underline{y} A_{\mathsf{mr}}(\underline{y})$ we have,

$$(A \wedge B)^{\mathsf{mr}} :\equiv \exists \underline{x}, \underline{y}(A_{\mathsf{mr}}(\underline{x}) \wedge B_{\mathsf{mr}}(\underline{y})),$$

$$(A \vee B)^{\mathsf{mr}} :\equiv \exists z^{\mathbb{N}}, \underline{x}, \underline{y}(z = 0 \to A_{\mathsf{mr}}(\underline{x})) \wedge (z \neq 0 \to B_{\mathsf{mr}}(\underline{y})),$$

[6] Kleene's notion of realizability is based on partial type-free realizers whereas modified realizability uses total typed functionals as realizers.

$$(A \to B)^{\mathsf{mr}} :\equiv \exists \underline{Y} \forall \underline{x} (A_{\mathsf{mr}}(\underline{x}) \to B_{\mathsf{mr}}(\underline{Y}\,\underline{x})),^{7}$$

$$(\forall y^\rho A(y))^{\mathsf{mr}} :\equiv \exists \underline{X} \forall z^\rho (A_{\mathsf{mr}}(\underline{X} y, y)),$$

$$(\exists y^\rho A(y))^{\mathsf{mr}} :\equiv \exists y, \underline{x} A_{\mathsf{mr}}(\underline{x}, y).$$

Following the notation commonly found in the literature, we shall normally write $\underline{x}$ mr A (read: $\underline{x}$ modified realizes A) instead of $\exists \underline{x} A_{\mathsf{mr}}(\underline{x})$.

Note that the length and the types of the tuple $\underline{x}$ in $\exists \underline{x} A_{\mathsf{mr}}(\underline{x})$ are determined by the logical structure of A.

Theorem 3.7 (Soundness of modified realizability, [116, 160]) *Let Δ be an arbitrary set of closed formulas. Assume that for each formula B of Δ (say $B^{\mathsf{mr}} \equiv \exists \underline{x} B_{\mathsf{mr}}(\underline{x})$) we are given an appropriate tuple of closed terms $\underline{r} \in \mathsf{T}$. Let*

$$\Delta^{\mathsf{mr}} :\equiv \{\underline{r} \ \mathsf{mr} \ B : B \in \Delta\},$$

and $A \in \mathcal{L}^\omega$. Then,

$$\mathsf{E\text{-}HA}^\omega + \Delta \vdash A \qquad \Longrightarrow \qquad \mathsf{E\text{-}HA}^\omega + \Delta^{\mathsf{mr}} \vdash \underline{t} \ \mathsf{mr} \ A,$$

where the tuple of closed terms $\underline{t} \in \mathsf{T}$ can be extracted from the given proof of A. The same holds for $\mathsf{N\text{-}HA}^\omega$ instead of $\mathsf{E\text{-}HA}^\omega$.

It is easy to see that for every instance B of $\mathsf{AC} \cup \mathsf{IP}_{\mathsf{ef}}$ one can produce a tuple of terms $\underline{r}$ such that $\mathsf{E\text{-}HA}^\omega \vdash \underline{r}$ mr B, where $\underline{r}$ consists only of projection functionals. On the other hand, as opposed to f.i., the realizability interpretation of Markov principle MP is not fulfilled by terms of $\mathsf{E\text{-}HA}^\omega$. The fact that f.i. interprets MP and m.r. does not is connected to the fact that f.i. does not interpret EXT while m.r. treats EXT as empty of information.

As done for the functional interpretation (cf. Section 3.4), using the majorizability relation one can develop a monotone version of modified realizability [87, 96]. In this dissertation we make use of modified realizability in Chapter 8, when giving a bar recursive realizer for the classical countable and dependent choice. For more information on realizability see e.g. [78, 81, 87, 158–160].

3.6 Modified Realizability and Negative Translation

As mentioned above, negated formulas are considered by modified realizability as empty of information. The negative translation, however, always produces a negated formula. This means that in order to apply modified realizability to classical proofs one has to first transform the result of negative translation into a "positive" formula. In many cases, e.g. for HA, this can be done for formulas of the form $\exists y A_{\mathsf{at}}(x, y)$, A_{at} being an atomic formula[8], since intuitionistic arithmetic is closed under the Markov rule, i.e.

[7] The negation of a formula A is understood as an abbreviation for $A \to 0 = 1$, which means that $(\neg A)^{\mathsf{mr}} \equiv \forall \underline{x}(\neg A_{\mathsf{mr}}(\underline{x}))$ and $\underline{y}$ is the empty tuple. Since the information of a formula A is, according to m.r., expressed by the existential quantifier of $(A)^{\mathsf{mr}}$, it turns out the negated formulas are considered by m.r. to be empty of information.

[8] We can actually use A_0 quantifier-free whenever using systems in which quantifier-free formulas are provably equivalent to atomic formulas. Notice, however, that in the neutral systems such as $\mathsf{N\text{-}HA}^\omega$ the restriction to atomic formulas is essential.

$$\mathsf{HA} \vdash \neg\neg\exists y A_{\mathsf{at}}(\underline{x}, y) \quad \Longrightarrow \quad \mathsf{HA} \vdash \exists y A_{\mathsf{at}}(\underline{x}, y).$$

An effective version of this rule can be obtained by a nice trick due to Friedman [51], the so-called A-translation[9].

A-translation

The A-translation is basically a trick for transforming a proof of $\neg\neg\exists y A_{\mathsf{at}}(x, y)$ into a proof of $\exists y A_{\mathsf{at}}(x, y)$, so that modified realizability can be applied. Recall that $\neg\exists y A_{\mathsf{at}}(x, y)$ is actually a short hand for $\exists y A_{\mathsf{at}}(x, y) \to\perp$ or $\exists y A_{\mathsf{at}}(x, y) \to 0 = 1$. The trick basically consists of replacing, in the whole proof of $\neg\neg\exists y A_{\mathsf{at}}(x, y)$, all atomic formulas P, including $\perp$, by the disjunction of P with the desired theorem $\exists y A_{\mathsf{at}}(x, y)$. In this way, with a few manipulations one obtains a proof for $\exists y A_{\mathsf{at}}(x, y)$. The general description of the translation is as follows:

Definition 3.5 ([51]) *Let $A \in \mathcal{L}$. With each formula $F \in \mathcal{L}$ (such that the free variables of A are not bounded in F) associate a formula $(F)^A \in \mathcal{L}$, called the A-translation of F, in the following way: $(F)^A$ results when all atomic formulas P in F are replaced by $P \vee A$.*

Theorem 3.8 (Soundness of A-translation, [51]) *Let A and F be formulas in $\mathcal{L}$ such that the free variables of A are not bounded in F. The following rule holds,*

$$\mathsf{HA} \vdash F \quad \Longrightarrow \quad \mathsf{HA} \vdash (F)^A.$$

As a corollary of the Soundness of A-translation we obtain an effective closure of HA under Markov rule.

Corollary 3.2 ([51]) *The following rule holds effectively*

$$\mathsf{HA} \vdash \forall x \neg\neg\exists y A_{\mathsf{at}}(x, y) \quad \Longrightarrow \quad \mathsf{HA} \vdash \forall x \exists y A_{\mathsf{at}}(x, y),$$

for A_{at} an atomic formula.

Proof. If $\mathsf{HA} \vdash \forall x \neg\neg\exists y A_{\mathsf{at}}(x, y)$ implies,

$$\mathsf{HA} \vdash ((\exists y A_{\mathsf{at}}(x, y) \to\perp) \to\perp).$$

By the Soundness of A-translation (for $A :\equiv \exists y A_{\mathsf{at}}(x, y)$) we have,

$$\mathsf{HA} \vdash ((\exists y (A_{\mathsf{at}}(x, y) \vee \exists y A_{\mathsf{at}}(x, y)) \to \exists y A_{\mathsf{at}}(x, y)) \to \exists y A_{\mathsf{at}}(x, y)),$$

and hence, $\mathsf{HA} \vdash (\exists y A_{\mathsf{at}}(x, y) \vee \exists y A_{\mathsf{at}}(x, y) \to \exists y A_{\mathsf{at}}(x, y)) \to \exists y A_{\mathsf{at}}(x, y)$. Since $\mathsf{HA} \vdash A \vee A \to A$, for arbitrary A, we get that $\mathsf{HA} \vdash \exists y A_{\mathsf{at}}(x, y)$. □

In fact, the corollary above holds for arbitrary Π^0_2-formulas, since in HA quantifier-free formulas are provably equivalent to atomic formulas. Also, instead of HA we could have used $\mathsf{N\text{-}HA}^\omega$, but in this case the restriction to atomic formulas is essential. For more information on the A-translation see e.g. [87]. Variants of the A-translation were considered in [17, 39, 122, 159].

[9] Independently investigated by Dragalin [41]

Combining A-translation with Modified Realizability

Notice that the combination of negative translation, modified realizability and A-translation loses modularity, since lemmas in the proof are analyzed with respect to the final theorem proven. In other words, in a classical context, via modified realizability, proofs must be analyzed entirely and the analysis of its lemmas cannot a priori be used when analyzing other theorems. This also implies that, as opposed to functional interpretation, one can not give a realizer for the negative translation P^N of a new axiom P in a definitive manner, but one must give a meta-procedure which given the final theorem A, whose classical proof uses P, produces a realizer for $(P^N)^A$. This meta-procedure, however, can be made modular in the sense that one treats uniformly the theorems A to which the A-translation is applied. This idea has been developed in [16] and used in Chapter 8 in order to realize, via modified realizability, the classical (i.e. negative translated) axioms of countable and dependent choice. The main idea is that instead of replacing $\perp$ by the final theorem, we slightly change the definition of modified realizability by regarding $y \;\mathsf{mr}\; \perp$ as an (uninterpreted) atomic formula, obtaining what we shall call *classical modified realizability.* More formally we define

$$y^\tau \;\mathsf{mr}^{\mathsf{c}}_{\tau}\; \perp :\equiv P_\perp(y),$$

where $P_\perp$ is a *new unary predicate symbol* and τ is the type of the witness to be extracted. Therefore, we have a modified realizability for each type τ, according to the type of the existential quantifier in the $\forall\exists$-formula we wish to realize.

In the following theorem, Δ is an axiom system possibly containing $P_\perp$ and further constants, which has the following closure property: If $D \in \Delta$ and E is a quantifier free formula with decidable predicates, then also the universal closure of $D[\lambda y^\tau.E/P_\perp]$ is in Δ, where $D[\lambda y^\tau.E/P_\perp]$ is obtained from D by replacing any occurrence of a formula $P_\perp(L)$ in D by $E[L/y]$.

Theorem 3.9 *Let $\underline{r}$ be a tuple of closed terms (possibly in an extension of $\mathcal{L}^\omega_{\mathsf{h}}$ with new constants) such that*

$$\mathsf{N\text{-}HA}^\omega + \Delta \vdash \underline{r} \;\mathsf{mr}^{\mathsf{c}}_{\tau}\; B^N,$$

and let $\forall x^\sigma \, \exists y^\tau \, A_{\mathsf{at}}(x,y)$ be a formula in the language of $\mathsf{N\text{-}HA}^\omega$*. Then*

$$\mathsf{N\text{-}PA}^\omega + B \vdash \forall x^\sigma \, \exists y^\tau \, A_{\mathsf{at}}(x,y) \qquad \Longrightarrow \qquad \mathsf{N\text{-}HA}^\omega + \Delta \vdash \forall x \, A_{\mathsf{at}}(x, tx),$$

where $t^{\sigma\to\tau}$ is a closed term in (an extension of) $\mathcal{L}^\omega_{\mathsf{h}}$.

Proof. See Chapter 8. □

Chapter 4

Interpreting Analysis Using Bar Recursion

In this chapter we discuss how proofs of $\forall\exists$-theorems, which involve comprehension and choice principles, can be analyzed to yield sub-recursive realizing programs. We shall start by discussing Spector's bar recursion, which was used in [154] for giving a functional interpretation of (the negative translation of) full countable choice. We illustrate the use of Spector's bar recursion by analyzing Avigad's proof which we included in Section 2.3.2. In Chapter 8 we shall introduce a different form of bar recursion, so-called modified bar recursion, and we show that it can also be used to realize full countable choice via a combination of negative translation, modified realizability and A-translation. In Chapter 8 we also show that modified bar recursion combined with a version of bar recursion due to Kohlenbach [88] defines the fan functional and that any set theoretic functional which on elements of $\mathcal{M}$ satisfies the equation for modified bar recursion must also live in $\mathcal{M}$. In Section 4.2.4 we shall also show that there exists a set theoretic functional which on elements of $\mathcal{M}$ satisfies the equation for modified bar recursion. Both results together imply that $\mathcal{M}$ is a model of modified bar recursion (as it is a model of Spector's bar recursion, cf. [18]). Other results we prove about modified bar recursion include (a) modified bar recursion defines Spector's bar recursion primitive recursively, (b) modified bar recursion of the lowest type is equivalent to the functional Γ (as defined in [53]), (c) modified bar recursion is not S1-S9 computable over the total continuous functions[1] (which follows from the fact that the fan functional is not S1-S9 computable over the total continuous functionals).

Most of the results presented in this chapter were obtained in collaboration with Ulrich Berger, and can be found in [14].

4.1 Spector's Bar Recursion

Definition 4.1 *The scheme of bar recursion introduced by Spector [154] consists of a family of functional symbols* $\{\mathsf{SBR}_{\rho,\tau}\}_{\rho,\tau\in\mathbf{T}}$ *with defining equations:*

$$\mathsf{SBR}_{\rho,\tau}(Y,G,H,s) =_\tau \begin{cases} G(s) & \text{if } Y(\hat{s}) <_{\mathbb{N}} |s| \\ H(s, \lambda x^\rho.\mathsf{SBR}_{\rho,\tau}(Y,G,H,s*x)) & \text{otherwise,} \end{cases} \tag{4.1}$$

[1]For a definition of the type structure of total continuous functionals see e.g. [133].

where $=_\tau$ *denotes extensional equality of type* τ*. The functional* Y *has type* $\rho^\omega \to \mathbb{N}$ *and the functional* H *has type* $\rho^* \to (\rho \to \tau) \to \tau$*. By* SBR *we mean the family of symbols* $\{\mathsf{SBR}_{\rho,\tau}\}_{\rho,\tau\in\mathbf{T}}$ *together with their defining equations.*

We say that a model $\mathcal{S}$ satisfies Spector's bar recursion (and any of its variants, which we shall introduce) if in $\mathcal{S}$ (for any given types τ and ρ) a functional exists satisfying the defining equation for $\mathsf{SBR}_{\rho,\tau}$, $\mathsf{KBR}_{\rho,\tau}$ (or the corresponding defining equation respectively).

4.1.1 Analyzing Avigad's Proof

We shall now use the proof given in Section 2.3.2 in order to illustrate the use of Spector's bar recursion for giving a functional interpretation of proofs based on comprehension.

Formalizing the Proof

We want to prove the statement

$$\forall F(\mathsf{Update}(F) \to \exists\sigma(\sigma = \sigma \oplus F(\hat{\sigma})))$$

The sketch of the proof goes as follows: Using primitive recursion we build a family of finite sequences $(\sigma_i)_{n\in\mathbb{N}}$ as $\sigma_0 := \langle\ \rangle$ and $\sigma_{i+1} := \sigma_i \oplus F(\hat{\sigma_i})$. Then, using arithmetical comprehension we can define a function $\hat{g}$ such that

$$\hat{g}(a) := \begin{cases} b & \text{if } \exists i(\langle a, b\rangle \in \sigma_i) \\ 0 & \text{otherwise.} \end{cases}$$

In fact, $\hat{g}$ can be primitive recursively defined on F and a function h which produces for each x an index j such that $x \in \mathsf{dom}(\sigma_j)$, if this index exists, i.e. h should be such that

$$\forall x(x \in \mathsf{dom}(\sigma_{hx}) \vee \forall j(x \notin \mathsf{dom}(\sigma_j))).$$

Such a function h can be obtained by an application of arithmetical countable choice to the following instance of the law of excluded middle

$$\forall x\exists i(x \in \mathsf{dom}(\sigma_i) \vee \forall j(x \notin \mathsf{dom}(\sigma_j))).$$

Once we have the total function $\hat{g}$, we look at a point of continuity n of the function F in $\hat{g}$, which means that F only needs the sequences σ_i which are used to defined $\hat{g}$ up to the point n. This information can be recovered from the function h by taking $k := \max_{m\le n} hm$, which gives rise to the desired fixed point, i.e. $F(\hat{g}) = F(\hat{\sigma}_k) = F(\hat{\sigma}_{k+1})$ and we can take $\sigma = \sigma_{k+1}$.

Analyzing the Proof

Let us take a general look at the structure of the whole proof. The main predicates used are:

$$A(h) :\equiv \forall x(x \in \mathsf{dom}(\sigma_{hx}) \vee \forall j(x \notin \mathsf{dom}(\sigma_j)))$$

$$
\begin{aligned}
B(g) &:\equiv \forall x(\forall i(x \in \mathsf{dom}(\sigma_i) \to g(x) = \sigma_i(x)) \wedge \\
&\qquad\qquad \exists j(x \notin \mathsf{dom}(\sigma_j) \to g(x) = 0)) \\
C(s,g) &:\equiv \forall \alpha(F(s * \alpha) = Fg), \\
D(h) &:\equiv B(g_h) \wedge C(\overline{g_h} n_{g_h}, g_h) \\
E(k) &:\equiv F(\hat{\sigma}_k) = F(\hat{\sigma}_{k+1}),
\end{aligned}
$$

and those are combined in the following way:

$$
\begin{prooftree}
\AxiomC{LEM}
\AxiomC{$\Pi^0_1\text{-}\mathsf{AC}^{\mathbb{N},\mathbb{N}}$}
\BinaryInfC{$\exists h A(h)$}
\AxiomC{$[\mathsf{Update}(F)]$}
\UnaryInfC{$\forall g C(\overline{g} n_g, g)$}
\AxiomC{$[\mathsf{Update}(F)]$}
\UnaryInfC{$\forall h(A(h) \to B(g_h))$}
\BinaryInfC{$\forall h(A(h) \to D(h))$}
\AxiomC{$\forall h(A(h) \wedge D(h) \to E(k_h))$}
\BinaryInfC{$\forall h(A(h) \to E(k_h))$}
\BinaryInfC{$\exists k E(k)$}
\end{prooftree}
$$

where $n_g = \omega_F(g)$ denotes a point of continuity of F in g, g_h denotes the primitive recursive function g obtained from h by

$$
g_h(x) := \begin{cases} \sigma_{hx}(x) & \text{if } \ x \in \mathsf{dom}(\sigma_{hx}) \\ 0 & \text{otherwise,} \end{cases}
$$

and $k_h := \max_{m \leq n_{g_h}} hm$.

Realizing $(\Pi^0_1\text{-}\mathsf{AC}^{\mathbb{N},\mathbb{N}})^{ND}$

We want to find realizers for the functional interpretation of the negative translation of

$$\exists h \forall x, j(x \in \mathsf{dom}(\sigma_{hx}) \vee x \notin \mathsf{dom}(\sigma_j)).$$

Let us first look at the interpretation of the following instance of $\Pi^0_1\text{-}\mathsf{AC}^{\mathbb{N},\mathbb{N}}$

$$\forall x \exists i \forall j A_0(x,i,j) \to \exists h \forall x, j A_0(x, hx, j),$$

in which we shall take $A_0(x,i,j) :\equiv x \in \mathsf{dom}(\sigma_i) \vee x \notin \mathsf{dom}(\sigma_j)$ (the matrix of A). By Kuroda's negative translation, intuitionistic logic and the stability of A_0 we get

$$\forall x \neg\neg \exists i \forall j A_0(x,i,j) \to \neg\neg \exists h \forall x, j A_0(x, hx, j),$$

which has the functional interpretation (two steps)

$$\exists \phi \forall x, g A_0(x, \phi x g, g(\phi x g)) \to \forall \psi_1, \psi_2 \exists h A_0(\psi_1 h, h(\psi_1 h), \psi_2 h),$$

$$\forall \phi, \psi_1, \psi_2 \exists x, g, h\big(A_0(x, \phi x g, g(\phi x g)) \to A_0(\psi_1 h, h(\psi_1 h), \psi_2 h)\big).$$

So our task is to produce x, g and h, uniformly in ϕ, ψ_1, ψ_2 such that

$$A_0(x, \phi x g, g(\phi x g)) \to A_0(\psi_1 h, h(\psi_1 h), \psi_2 h),$$

which can be obtained by satisfying the following equations

$$\begin{array}{rcl} x & = & \psi_1 h \\ h(\psi_1 h) & = & \phi x g \\ \psi_2 h & = & g(\phi x g) \end{array}$$

This can be done using Spector's bar recursion, by actually solving

$$(+) \quad \exists h \forall n \leq \psi_1 h \exists g_n (\phi n g_n = hn \wedge g_n(hn) = \psi_2 h),$$

since, given an h and a sequence $(g_n)_{n\in\mathbb{N}}$ satisfying $(+)$, one can take $x := \psi_1 h$ and $g := g_x$, in order to solve the above system of equations. Let $h := \Delta'(\phi, \psi_1, \psi_2, \langle\,\rangle)$, where

$$\Delta'(\phi, \psi_1, \psi_2, s) = \begin{cases} \hat{s} & \text{if } \psi_1(\hat{s}) < |s| \\ \Delta'(\phi, \psi_1, \psi_2, s * c) & \text{otherwise,} \end{cases}$$

$c = \phi(|s|, g_{|s|})$ and $g_{|s|} = \lambda x.\psi_2(\Delta'(\phi, \psi_1, \psi_2, s * x))$. One can verify that for each s one has

(*i*) $\forall n < |s|(\Delta'(\phi, \psi_1, \psi_2, s)(n) = s(n))$, and

(*ii*) $\forall n \geq |s|(P(\Delta'(\phi, \psi_1, \psi_2, s), n))$,

where $P(h, n) :\equiv \psi_1 h \geq n \to \exists g_n(\phi n g_n = hn \wedge g_n(hn) = \psi_2 h)$. Therefore,

- $h := \Delta'(\phi, \psi_1, \psi_2, \langle\,\rangle)$ and
- $g_n := \lambda x.\psi_2(\Delta'(\phi, \psi_1, \psi_2, \overline{h}n * x))$

solves $(+)$. In this simple instance of $\Pi^0_1\text{-}\mathsf{AC}^{\mathbb{N},\mathbb{N}}$ the premise can actually be realized, i.e.

$$\exists \phi \forall x, g(x \in \mathsf{dom}(\sigma_{\phi x g}) \vee x \notin \mathsf{dom}(\sigma_{g(\phi x g)}))$$

can be realized by

$$\phi(x, g) := \begin{cases} gx & \text{if } x \in \mathsf{dom}(\sigma_{gx}) \\ x & \text{otherwise.} \end{cases}$$

By taking $\Delta(\psi_1, \psi_2, s) := \Delta'(\phi, \psi_1, \psi_2, s)$ we have

$$\Delta(\psi_1, \psi_2, s) = \begin{cases} \hat{s} & \text{if } \psi_1(\hat{s}) \leq |s| \\ \Delta(\psi_1, \psi_2, s * g_{|s|}(|s|)) & \text{if } |s| \in \mathsf{dom}(\sigma_{g_{|s|}(|s|)}) \\ \Delta(\psi_1, \psi_2, s * |s|) & \text{otherwise.} \end{cases}$$

In order to obtain the function h (the comprehension function) we need to obtain the functionals ψ_1 (which gives the tree in the bar recursion) and ψ_2. Those are obtained in the following lemma.

The Lemma $\forall h(A(h) \wedge D(h) \to E(k_h))$

This lemma says (letting n_{g_h} abbreviate $\omega_F(g_h)$)

$$\forall h(A(h) \wedge B(g_h) \wedge C(\overline{g}n_{g_h}, g_h) \to E(k_h)).$$

After presenting the quantifier in A and B we get

$$\forall h(\forall y, j A_0(h, y, j) \wedge \forall x, i \exists j B_0(g_h, x, i, j) \wedge C(\overline{g}n_{g_h}, g_h) \to E(k_h)).$$

The negative translation of this statement follows intuitionistically from the statement itself, and partial functional interpretation (disregarding the universal quantifier in C) gives, in two steps,

$$\forall h(\forall y, j A_0(h, y, j) \wedge \exists \alpha \forall x, i B_0(g_h, x, i, \alpha x i) \wedge C(\overline{g}n_{g_h}, g_h) \to E(k_h))$$

$$\forall h, \alpha \exists y, j, x, i(A_0(h, y, j) \wedge B_0(g_h, x, i, \alpha x i) \wedge C(\overline{g}n_{g_h}, g_h) \to E(k_h))$$

We now define functionals $\phi_1, \phi_2, \psi_1, \psi_2$ realizing y, j, x, and i respectively. Let $\phi_1(h, \alpha)$ and $\psi_1(h, \alpha)$ be defined as

$$\begin{cases} \min z \leq n_{g_n} \neg(A_0(h, z, \alpha(z, hz)) \wedge B_0(g_h, z, hz, \alpha(z, hz))) & \text{if such } z \text{ exists} \\ n_{g_h} & \text{otherwise,} \end{cases}$$

and (taking $z = \phi_1(h, \alpha)$) $\phi_2(h, \alpha) := \alpha(z, hz)$ and $\psi_2(h, \alpha) := hz$. We must show that for all h and α

$$A_0(h, z, \alpha(z, hz)) \wedge B_0(g_h, z, hz, \alpha(z, hz)) \wedge C(\overline{g}n_{g_h}, g_h) \to E(k_h).$$

By the way we have defined ϕ_1 and ψ_1, if $A_0(h, z, \alpha(z, hz)) \wedge B_0(g_h, z, hz, \alpha(z, hz))$ does not hold for some $z \leq n_{g_h}$ we are done. Therefore, assume

$$\forall z \leq n_{g_h}(A_0(h, z, \alpha(z, hz)) \wedge B_0(g_h, z, hz, \alpha(z, hz)))$$

i.e. for all z smaller or equal to n_{g_h}

(*i*) $z \in \mathsf{dom}(\sigma_{hz}) \vee z \notin \mathsf{dom}(\sigma_{\alpha(z,hz)})$,

(*ii*) $z \in \mathsf{dom}(\sigma_{hz}) \to g_h z = \sigma_{hz}(z)$,

(*iii*) $z \notin \mathsf{dom}(\sigma_{\alpha(z,hz)}) \to g_h z = 0$.

We show $E(k_h)$. By $C(\overline{g}n_{g_h}, g_h)$ it is enough to show that

$$\forall z \leq n_{g_h}(g_h(z) \neq 0 \to g_h(z) = \sigma_{hz}(z)).$$

Fix $z \leq n_{g_h}$ such that $g_h(z) \neq 0$. By (*iii*) $z \in \mathsf{dom}(\sigma_{\alpha(z,hz)})$. By (*i*) $z \in \mathsf{dom}(\sigma_{hz})$, and by (*ii*) $g_h(z) = \sigma_{hz}(z)$.

The Final Function

The functional α in the definition of ϕ_2 comes from the lemma $\forall h(A(h) \to B(g_h))$ and can be taken to be $\alpha_h(x, i) = hx$. Therefore, $\phi_2 = \psi_2$ (and by definition $\phi_1 = \psi_1$). The fixed point k is obtained at $\max_{m \leq n} hm$, where $n := \omega_F(g_h)$ (a point of continuity of F on g_h) and $h := \Delta(\psi_1, \psi_2, \langle\, \rangle)$.

Notice that, since the assumption of the theorem, i.e. that F is an update procedure, involves the statement that F is continuous, it is a priori clear that the final functional might depend on the computational content of the statement that F is continuous, which is given via the modulus of continuity ω_F.

4.2 Modified Bar Recursion

Let $s @ \alpha$ denotes the overwriting of the initial segment of $\alpha : \rho^\omega$ with the finite string $s : \rho^*$. We define *modified bar recursion at type* ρ as

$$\mathsf{MBR}_\rho(Y, H, s) =_{\mathbb{N}} Y(s @ H(s, \lambda x^\rho.\mathsf{MBR}_\rho(Y, H, s * x))), \tag{4.2}$$

where Y has type $\rho^\omega \to \mathbb{N}$. In Chapter 8 we show that this schema of bar recursion can be used in order to interpret countable and dependent choice via a combination of negative translation, A-translation and modified realizability. In the following we present some results which show that modified bar recursion is strictly stronger than Spector's original definition of bar recursion.

Our definition of MBR in Chapter 8 uses the more general type o instead of $\mathbb{N}$. The (inessential) restriction to $\mathbb{N}$ is convenient for our comparison of MBR and Spector's bar recursion.

4.2.1 Definability of Spector's Bar Recursion

We start by defining Spector's bar recursion primitive recursive in MBR, i.e. assuming we have a term t satisfying the equation for MBR we build a term t' (primitive recursively in t) which satisfies the equation for SBR. For that purpose we first show that MBR defines the following search operator.

Definition 4.2 $\tilde{\mu}(Y, \alpha^{\rho^\omega}, k) := \min n \geq k\, [Y(\overline{\alpha, n}) < n]$, *where*

$$(\overline{\alpha, n})(m) := \begin{cases} \alpha(m) & \text{if } m < n \\ 0^\rho & \text{otherwise.} \end{cases}$$

Kohlenbach [88] has shown that $\tilde{\mu}$ is primitive recursively definable in SBR.

Lemma 4.1 *$\tilde{\mu}$ is primitive recursively definable in* MBR*.*

Proof. Let n be the value of $\tilde{\mu}(Y, \alpha, k)$. The case when $n = k$ is simple and will be treated as a special case. Therefore, for the arguments let us assume that $n > k$. In this case we note that, by the minimality condition, $Y(\overline{\alpha, n-1}) \geq n - 1$. Hence, $Y(\overline{\alpha, n-1}) + 1$ is an upper bound for the value of n. The idea is to use MBR in order to obtain such an upper bound, via a recursion along the infinite path α, by considering longer and longer initial segments of α. Notice that, in order to obtain the upper bound on n we must give the sequence $\overline{\alpha, n-1}$ as input to Y. We show how this sequence can be computed by an appropriate functional H (in the definition of MBR) built out of Y, α and k, which, given an initial segments of α, checks for the condition $Y(\overline{\alpha, n}) < n$. The formal proof goes as follows. By MBR we can define a Φ_α satisfying $\Phi_\alpha(s) = Y(s @ (\overline{\alpha, m-1}))$, where,

$$(*)\ \ m =_{\mathbb{N}} \begin{cases} |s| + 1 & \text{if } Y(\overline{\alpha, |s|+1}) < |s| + 1 \\ \tilde{\mu}^b(Y, \alpha, k, \Phi_\alpha(s * \alpha(|s|)) + 1) & \text{otherwise,} \end{cases}$$

and $\tilde{\mu}^b$ is the bounded version of $\tilde{\mu}$ (which is primitive recursive). Notice that m will always take the desired value n, whereas $\Phi_\alpha(s) + 1$ is always a bound on n. We then define,

$$\tilde{\mu}(Y,\alpha,k) := \begin{cases} k & \text{if } Y(\overline{\alpha,k}) < k \\ \tilde{\mu}^b(Y,\alpha,k,\Phi_\alpha(\overline{\alpha}k)+1) & \text{otherwise.} \end{cases}$$

We show that this is a good definition of $\tilde{\mu}$ by showing that $\Phi_\alpha(\overline{\alpha}k)+1$ is a good upper bound on the value of $\tilde{\mu}(Y,\alpha,k)$ (assume this value is $n > k$). In fact, we show by induction on j that, for $k \leq j < n$, n is bounded by $\Phi_\alpha(\overline{\alpha}j)+1$.

i) $j = n-1$. We see that the first case of $(*)$ will be satisfied, m is equal n and $\Phi_\alpha(\overline{\alpha}j)+1 = Y(\overline{\alpha}j \;@\; (\overline{\alpha, m-1})) + 1 = Y(\overline{\alpha, n-1}) + 1 \geq n$.

ii) $j < n-1$. By induction hypothesis $\Phi_\alpha(\overline{\alpha}j * \alpha(j)) + 1$ is a bound for n. Therefore, m (see second case of $(*)$) has value n, and as above we get $\Phi_\alpha(\overline{\alpha}j)+1 \geq n$. $\Box$

Lemma 4.2 $\mathsf{SBR}_{\rho,\mathbb{N}}$ *is primitive recursively definable in* MBR_ρ.

Proof. We show how to define (primitive recursively in MBR) a Ψ satisfying the defining equation for $\mathsf{SBR}_{\rho,\mathbb{N}}$,

$$(i)\;\; \Psi(Y,G,H,s) =_{\mathbb{N}} \begin{cases} G(s) & \text{if } Y(\hat{s}) < |s| \\ H(s, \lambda x.\Psi(Y,G,H,s*x)) & \text{otherwise.} \end{cases}$$

The main idea is to use sequences of pairs, in the definition of modified bar recursion, so that given the argument α for Y we can distinguish between the part which came from s and the part which was produced by the functional H. In this way we are able to simulate Spector's bar recursion. Let Φ be a functional satisfying MBR_ρ. In the following π_0^ρ and π_1^ρ will denote the projection functionals, i.e. $\pi_i(\langle x_0^\rho, x_1^\rho \rangle) = x_i$, $i \in \{0,1\}$ (we often omit the type superscript of the projection functionals). If $s^{\langle\rho,\rho\rangle^*} = \langle s_0,\ldots,s_n\rangle$, $\pi_i(s)$ also denotes $\langle \pi_i(s_0),\ldots,\pi_i(s_n)\rangle$. In the same way we define $\pi_i(\alpha^{\langle\rho,\rho\rangle^\omega})$. We first define two tilde operations,

$$(ii)\;\; \tilde{H}(s,F) := \lambda n.\langle 1, H(\pi_1(s), \lambda x.F(\langle 0, x\rangle))\rangle,$$

which produces an infinite sequence of constant value, and [2]

$$(iii)\;\; \tilde{Y}_{G,k}(\alpha) := \begin{cases} G(\pi_1(s)) & \text{if } \bigwedge_{i=0}^{n} (\pi_0(s_i) = 0) \\ \pi_1(s_n) & \text{otherwise,} \end{cases}$$

where (in the definition of $\tilde{Y}_{G,k}$) $s = \langle s_0,\ldots,s_n\rangle = \overline{\pi_1(\alpha)}\;\tilde{\mu}(Y,\pi_1(\alpha),k)$. Note that the first operation is primitive recursive in H, s and F; and the second is primitive recursive in Y, G, k, α and MBR (since it uses $\tilde{\mu}$). Moreover,

$(+)$ for all s, if $Y(\hat{s}) \geq |s|$ then for all α extending s, $\tilde{Y}_{G,|s|}(\alpha) = \tilde{Y}_{G,|s|+1}(\alpha)$.

We abbreviate $\langle\langle 0, s_0\rangle,\ldots,\langle 0, s_{|s|-1}\rangle\rangle$ by $\langle 0, s\rangle$. Define

$$(iv)\;\; \Psi(Y,G,H,s) := \Phi(\tilde{Y}_{G,|s|}, \tilde{H}, \langle 0, s\rangle)).$$

We show that Ψ satisfies equation (i), i.e.

[2]Since, by our construction, the first element of the pair will either be 0^ρ or 1^ρ, the test $\pi_0(s_i) = 0$ in the definition (iii) is primitive recursive.

$$(v)\ \ \Phi(\tilde{Y}_{G,|s|}, \tilde{H}, \langle 0, s\rangle) = \begin{cases} G(s) & \text{if } Y(\hat{s}) < |s| \\ H(s, \lambda x.\Phi(\tilde{Y}_{G,|s|+1}, \tilde{H}, \langle 0, s*x\rangle)) & \text{otherwise.} \end{cases}$$

We first note that, by the definition of MBR (and (ii)),

$$(vi)\ \ \Phi(\tilde{Y}_{G,|s|}, \tilde{H}, \langle 0, s\rangle) = \tilde{Y}_{G,|s|}(\langle 0, s\rangle \ @\ \lambda n.\langle 1, H(s, \lambda x.\Phi(\tilde{Y}_{G,|s|}, \tilde{H}, \langle 0, s*x\rangle))\rangle).$$

We will show that (v) holds. Assume $Y(\hat{s}) < |s|$, we have,

$$\Phi(\tilde{Y}_{G,|s|}, \tilde{H}, \langle 0, s\rangle) \overset{(vi)}{=} \tilde{Y}_{G,|s|}(\langle 0, s\rangle \ @\ \ldots) \overset{(iii)}{=} G(s).$$

On the other hand, if $Y(\hat{s}) \geq |s|$ then,

$$\begin{aligned} \Phi(\tilde{Y}_{G,|s|}, \tilde{H}, \langle 0, s\rangle) &\overset{(vi)}{=} \tilde{Y}_{G,|s|}(\langle 0, s\rangle \ @\ \lambda n.\langle 1, H(s, \lambda x.\Phi(\tilde{Y}_{G,|s|}, \tilde{H}, \langle 0, s*x\rangle))\rangle) \\ &\overset{(iii)}{=} H(s, \lambda x.\Phi(\tilde{Y}_{G,|s|}, \tilde{H}, \langle 0, s*x\rangle)) \\ &\overset{(+)}{=} H(s, \lambda x.\Phi(\tilde{Y}_{G,|s|+1}, \tilde{H}, \langle 0, s*x\rangle)), \end{aligned}$$

and the proof is concluded. □

Lemma 4.3 $\mathsf{SBR}_{\rho,\tau}$ *is primitive recursively definable in* $\mathsf{SBR}_{\rho',\mathbb{N}}$*, where if* $\tau = \tau_1 \to \ldots \to \tau_n \to \mathbb{N}$ *then* $\rho' = \rho \times \tau_1 \times \ldots \times \tau_n$.

Proof. Let $\tau = \tau_1 \to \ldots \to \tau_n \to \mathbb{N}$. We will show that $\mathsf{SBR}_{\rho,\tau}$ can be defined from $\mathsf{SBR}_{\rho\times\tau_1\times\ldots\times\tau_n,\mathbb{N}}$. We have to define a functional Φ such that,

$$(i)\ \ \Phi(Y, G, H, s) =_\tau \begin{cases} G(s) & \text{if } Y(\hat{s}) <_{\mathbb{N}} |s| \\ H(s, \lambda x^\rho.\Phi(Y, G, H, s*x)) & \text{otherwise.} \end{cases}$$

We shall often omit the parameters Y, G and H, since those stay fixed during the recursion. From Y, G and H we define,

(ii) $\tilde{Y}(\alpha) := Y(\pi_0^{n+1}(\alpha))$,

(iii) $\tilde{G}(t) := G(\pi_0^{n+1}(t))(y)$,

(iv) $\tilde{H}(t, F) := H(\pi_0^{n+1}(t), \lambda x^\rho, z_1^{\tau_1}, \ldots, z_n^{\tau_n}.F(\langle x, z_1, \ldots, z_n\rangle))(y)$,

where y denotes $\pi_1^{n+1}(t_{|t|-1}), \ldots, \pi_n^{n+1}(t_{|t|-1})$ and the types are,

$$\begin{aligned} \alpha &: (\rho \times \tau_1 \times \ldots \times \tau_n)^\omega \\ y &: \tau_1 \times \ldots \times \tau_n \\ t &: (\rho \times \tau_1 \times \ldots \times \tau_n)^* \\ F &: (\rho \times \tau_1 \times \ldots \times \tau_n) \to \mathbb{N} \end{aligned}$$

and we define (using $\mathsf{SBR}_{\rho\times\tau_1\times\ldots\times\tau_n,\mathbb{N}}$),

$$(v)\ \ \Psi(\tilde{Y}, \tilde{G}, \tilde{H}, t) =_{\mathbb{N}} \begin{cases} \tilde{G}(t) & \text{if } \tilde{Y}(t \ @\ 0) <_{\mathbb{N}} |t| \\ \tilde{H}(t, \lambda x^{\rho\times\tau_1\times\ldots\times\tau_n}.\Psi(t*x)) & \text{otherwise.} \end{cases}$$

Finally we set, ($\langle s, \mathbf{y}\rangle$ abbreviates $\langle\langle s_0, \mathbf{y}\rangle, \ldots, \langle s_{|s|-1}, \mathbf{y}\rangle\rangle$)

(*vi*) $\Phi(Y, G, H, s) :=_\tau \lambda \mathbf{y}.\Psi(\tilde{Y}, \tilde{G}, \tilde{H}, \langle s, \mathbf{y}\rangle)$.

We show that equation (*i*) is satisfied by Φ. One easily verifies that

(*vii*) $\Psi(\tilde{Y}, \tilde{G}, \tilde{H}, \langle s, \mathbf{y}\rangle) = \Psi(\tilde{Y}, \tilde{G}, \tilde{H}, \langle\langle s_0, \mathbf{z}\rangle, \ldots, \langle s_{|s|-2}, \mathbf{z}\rangle, \langle s_{|s|-1}, \mathbf{y}\rangle\rangle)$,

for arbitrary $\mathbf{z}$. Let Y, G, H and s be fixed and t abbreviate $\langle s, \mathbf{y}\rangle$. By (*ii*), $Y(\hat{s}) < |s|$ iff $\tilde{Y}(\hat{t}) < |t|$. Therefore, if $Y(\hat{s}) < |s|$ then

$$\begin{aligned} \Phi(Y, G, H, s) &\overset{(vi)}{=} \lambda \mathbf{y}.\Psi(\tilde{Y}, \tilde{G}, \tilde{H}, \langle s, \mathbf{y}\rangle) \\ &\overset{(v)}{=} \lambda \mathbf{y}.\tilde{G}(\langle\langle s, \mathbf{y}\rangle) \overset{(iii)}{=} \lambda \mathbf{y}.G(s)(\mathbf{y}) = G(s). \end{aligned}$$

On the other hand, if $Y(\hat{s}) \geq |s|$ then

$$\begin{aligned} \Phi(Y, G, H, s) &\overset{(vi)}{=} \lambda \mathbf{y}.\Psi(\tilde{Y}, \tilde{G}, \tilde{H}, \langle s, \mathbf{y}\rangle) \overset{(v)}{=} \lambda \mathbf{y}.\tilde{H}(t, \lambda x.\Psi(t * x)) \\ &\overset{(iv)}{=} \lambda \mathbf{y}.H(s, \lambda x, \mathbf{z}.\Psi(t * \langle x, \mathbf{z}\rangle))(\mathbf{y}) \\ &\overset{(vii)}{=} \lambda \mathbf{y}.H(s, \lambda x, \mathbf{z}.\Psi(\langle s * x, \mathbf{z}\rangle))(\mathbf{y}) \\ &\overset{(vi)}{=} \lambda \mathbf{y}.H(s, \lambda x.\Phi(s * x))(\mathbf{y}) = H(s, \lambda x.\Phi(s * x)) \end{aligned}$$

$\square$

Theorem 4.1 SBR *is primitive recursively definable in* MBR.

Proof. Follows immediately from Lemmas 4.2 and 4.3. $\square$

4.2.2 The Equivalence of $\mathsf{MBR}_{\mathbb{N}}$ and the Functional Γ

The functional Γ (introduced in [53]) is defined as

$$\Gamma(Y, s) =_{\mathbb{N}} Y(s * 0 * \lambda n^{\mathbb{N}}.\Gamma(Y, s * (n + 1))). \tag{4.3}$$

It is easy to see that equation (4.3), in the model of all total continuous functionals, specifies a unique functional (similarly to $\mathsf{MBR}_{\mathbb{N}}$-). Gandy and Hyland's purpose for defining the functional Γ was to show that there exists a functional having a recursive associate (in the sense of Kleene's [80] schemata S1-S8, see Section 4.2.3 and also [132]) but not being S1-S9 computable in the total functionals (cf. [132]), even with the fan functional as an oracle. In the following we show that modified bar recursion of the lowest type is primitive recursively (in the sense of Kleene) equivalent to the functional Γ. Hence, one can view MBR as an extension of the functional Γ to higher types.

Theorem 4.2 *The functional* Γ *is primitive recursively equivalent to* $\mathsf{MBR}_{\mathbb{N}}$.

Proof. It is easy to see that $\mathsf{MBR}_{\mathbb{N}}$ defines the functional Γ. For the other direction the intuition is as follows. We first use Γ to compute the value of $\mathsf{MBR}_{\mathbb{N}}(Y, H, s) + 1$. The advantage of doing this is that, if the sequence s contains only positive numbers, the functional Y will get a sequence α containing only one zero, namely the one introduced by the functional Γ. Once α has the form $s*0*\beta$, it is easy to transform it (primitive recursively) into the sequence $s*H(s,\beta)$. And if β is taken to be $\lambda x.\mathsf{MBR}_{\mathbb{N}}(s*x)$ we are done. Now we give the formal proof. Define

$$(i)\ \mathsf{MBR}_{\mathbb{N}}(Y,H,s) := \Gamma(\tilde{Y},\tilde{s}) - 1,$$

where

$$\tilde{s} := \langle s_0+1,\dots,s_{|s|-1}+1\rangle$$

and (ii) $\tilde{Y}(\alpha) := Y(\hat{\alpha})+1$, where

$$\hat{\alpha}(k) := \begin{cases} \alpha(k)-1 & \text{if } \forall m \le k(\alpha(m)\neq 0) \\ H(\overline{(\alpha-1)}c_k, \lambda n.(\alpha(n+c_k+1) \dot{-} 1))(k-c_k) & \text{otherwise,} \end{cases}$$

and c_k is a shorthand for $\min m \le k(\alpha(m)=0)$. We only have to notice that if α has the form $\tilde{s}*0*\beta$ then

$$(iii)\ \hat{\alpha} = s*H(s,\lambda n.(\beta(n)-1)).$$

We then have the following:

$$\begin{array}{rcl} \mathsf{MBR}_{\mathbb{N}}(Y,H,s) & \overset{(i)}{=} & \Gamma(\tilde{Y},\tilde{s})-1 \\ & \overset{(4.3)}{=} & \tilde{Y}(\tilde{s}*0*\lambda n.\Gamma(\tilde{Y},\tilde{s}*(n+1)))-1 \\ & = & \tilde{Y}(\tilde{s}*0*\lambda n.\Gamma(\tilde{Y},\widetilde{s*n}))-1 \\ & \overset{(ii)}{=} & Y(\tilde{s}*0*\widehat{\lambda n.\Gamma(\tilde{Y},\widetilde{s*n})}) \\ & \overset{(iii)}{=} & Y(s*H(s,\lambda n.(\Gamma(\tilde{Y},\widetilde{s*n})-1))) \\ & \overset{(i)}{=} & Y(s*H(s,\lambda n.\mathsf{MBR}_{\mathbb{N}}(Y,H,s*n))) \end{array}$$

□

4.2.3 S1-S9 Computability

We show in Chapter 8 that modified bar recursion, together with a version of bar recursion due to Kohlenbach, defines primitive recursively the fan functional. In the previous section we have also shown that modified bar recursion of lowest type is already equivalent to the functional Γ, as defined in [53]. Since the fan functional is not S1-S9 computable over the total continuous functionals, and the functional Γ is not S1-S9 computable over the total continuous functionals even having the fan functional as an oracle, we obtain in this section corresponding results for modified bar recursion. For completeness we include the definition of S1-S9 computability [80].

Definition 4.3 (Axioms S1-S9) *In any applicative type structure $\mathcal{S}$ (containing $\mathbb{N}$) we define a set of relations Γ (parametrized by their arity and type of arguments) on $\mathcal{S}$ inductively as follows,* [3]

[3]We abbreviate $y_1,\dots,y_n$ (of arbitrary type) by $\underline{y}$. The variables e_1,e_2,m,n,i,k,k_1,k_2 range over natural numbers, σ and π range over codes for pure finite types and permutations, respectively, and $f,x,\underline{y}$ over functionals of appropriate types. We write $\{e\}^{\mathcal{S}}(\underline{y}) \simeq k$ instead of $\mathcal{S} \models \Gamma(e,\underline{y},k)$.

S1 $\{e\}^{\mathcal{S}}(m, \underline{y}) \simeq m + 1$, *where* $e = \langle 1, \sigma \rangle$.

S2 $\{e\}^{\mathcal{S}}(\underline{y}) \simeq k$, *where* $e = \langle 2, \sigma, k \rangle$.

S3 $\{e\}^{\mathcal{S}}(m, \underline{y}) \simeq m$, *where* $e = \langle 3, \sigma \rangle$.

S4 *If* $\{e_1\}^{\mathcal{S}}(\underline{y}) \simeq k_1$ *and* $\{e_2\}^{\mathcal{S}}(k_1, \underline{y}) \simeq k_2$ *then* $\{e\}^{\mathcal{S}}(\underline{y}) \simeq k_2$, *where* $e = \langle 4, e_1, e_2, \sigma \rangle$.

S5 *Primitive recursion,*

S6 *If* $\{e_1\}^{\mathcal{S}}(\pi(\underline{y})) \simeq k$ *then* $\{e\}^{\mathcal{S}}(\underline{y}) \simeq k$, *where* $e = \langle 6, e_1, \pi, \sigma \rangle$.

S7 $\{e\}^{\mathcal{S}}(f^1, x^{\mathbb{N}}, \underline{y}) \simeq f(x)$, *where* $e = \langle 7, \sigma \rangle$.

S8 *If* $\{e_1\}^{\mathcal{S}}(x^\rho, \underline{y}) \simeq f^{\rho \to \mathbb{N}}(x)$, *for all* x, *then* $\{e\}^{\mathcal{S}}(\underline{y}) \simeq y_1(f)$, *where* $e = \langle 8, e_1, \sigma \rangle$.

S9 *If* $\{e_1\}^{\mathcal{S}}(y_1, \ldots, y_i) \simeq k$ *then* $\{e\}^{\mathcal{S}}(e_1, \underline{y}) \simeq k$, *where* $i \leq n$ *and* $e = \langle 9, i, \sigma \rangle$.

One can prove by induction on S1-S9 *that for each* e *and* $\underline{y}$ *there exists at most one* k *such that* $\{e\}^{\mathcal{S}}(\underline{y}) \simeq k$. *Therefore, each index* e *gives rise to a partial functional (denoted by* $\{e\}^{\mathcal{S}}$*) which on input* $\underline{y}$ *takes value* k *if* $\{e\}^{\mathcal{S}}(\underline{y}) \simeq k$ *and is undefined otherwise. It is important to note that the functional* $\{e\}^{\mathcal{S}}$ *yielded by an index* e *need not belong to* $\mathcal{S}$. *The set of all indices* e *such that* $\{e\}^{\mathcal{S}} \in \mathcal{S}$ *is denoted by* $\mathcal{R}ec^{\mathcal{S}}$. *If* $\{e\}^{\mathcal{S}}$ *is a functional of the form* $\lambda\Psi, \underline{y}.\{e\}^{\mathcal{S}}(\Psi, \underline{y})$ *then* $\{e\}^{\mathcal{S}}_{\Psi}$ *denotes the functional* $\lambda\underline{y}.\{e\}(\Psi, \underline{y})$.

Definition 4.4 *A formula* P *in the language of* HA^ω *having a unique free variable is called an* specification of a functional *or just a* functional, *e.g.* SBR *having variables* Y, G, H *and* s *universally quantified is an specification for Spector's bar recursor.*

Definition 4.5 (S1-S9 computability) *Let* P, Q *be specifications and* $\mathcal{S}$ *any applicative type structure (containing* $\mathbb{N}$*). Then,*

- P *is* S1-S9 computable in $\mathcal{S}$ *if* $\mathcal{S} \models \exists e \in \mathcal{R}ec^{\mathcal{S}}.\mathsf{P}(\{e\}^{\mathcal{S}})$,
- P *is* S1-S9 + Q computable in $\mathcal{S}$ *if* $\mathcal{S} \models \exists\Psi\mathsf{Q}(\Psi)$ *and* $\mathcal{S} \models \exists e \in \mathcal{R}ec^{\mathcal{S}}\forall\Psi\,\big(\mathsf{Q}(\Psi) \to \mathsf{P}(\{e\}^{\mathcal{S}}_{\Psi})\big)$.

Moreover, we also say that P is primitive recursive (in the sense of Kleene) in Q *if* P *is S1-S8 +* Q *computable in the model of total continuous functionals.*

Note that, although we define S1-S9 computability for pure finite types, the definition can be trivially extended to all finite types.

Lemma 4.4 KBR *and* SBR *are* S1-S9 *computable in* $\mathcal{C}$.

Proof. One shows $\mathcal{C} \models \exists e \in \mathcal{R}ec^{\mathcal{C}}.\mathsf{KBR}(\{e\}^{\mathcal{C}})$ and $\mathcal{C} \models \exists e \in \mathcal{R}ec^{\mathcal{C}}.\mathsf{SBR}(\{e\}^{\mathcal{C}})$ using the recursion theorem. $\square$

Theorem 4.3 ([53, 132]) FAN *is not* S1-S9 *computable in* $\mathcal{C}$.

Lemma 4.5 FAN *is* S1-S9 + MBR *computable in* $\mathcal{C}$.

Proof. By Theorem 8.1 there exists a $\Psi \in \mathcal{C}$ such that $\mathcal{C} \models \mathsf{MBR}(\Psi)$. In Chapter 8 we have shown that $\mathcal{C} \models \exists e \in \mathcal{R}ec^{\mathcal{C}}(\forall\Psi(\mathsf{MBR}(\Psi) \to \mathsf{FAN}(\{e\}^{\mathcal{C}}_{\Psi}))$. $\square$

Corollary 4.1 MBR *is not* S1-S9 *computable in* $\mathcal{C}$.

Proof. Assume $\mathcal{C} \models \exists e \in \mathcal{R}ec^{\mathcal{C}}.\mathsf{MBR}(\{e\}^{\mathcal{C}})$. By Lemma 4.5 we have that FAN is S1-S9 computable in $\mathcal{C}$, contradicting Theorem 4.3. $\square$

Corollary 4.2 MBR *is not primitive recursively definable in* KBR *nor* SBR.

Proof. Follows from the corollary above, Lemma 4.4 and the fact that the set of functionals S1-S9 computable in $\mathcal{C}$ is closed under primitive recursion. $\square$

Gandy and Hyland also showed that the functional Γ (see Section 4.2.2) is not S1-S9 computable in $\mathcal{C}$ even in the fan functional. From Theorem 4.2 we obtain the following corollary.

Corollary 4.3 $\mathsf{MBR}_{\mathbb{N}}$ *is not* S1-S9 *computable in* $\mathcal{C}$*, even in the fan functional.*

4.2.4 Finding $\Phi \in \mathcal{M}_{\rho^\omega \to o} \times \mathcal{M}_{\rho^* \times (\rho \to o) \to \rho^\omega} \times \mathcal{M}_{\rho^*} \to \mathcal{M}_o$ Satisfying MBR

We show in Chapter 8 that any functional

$$\Phi : \mathcal{M}_{\rho^\omega \to o} \times \mathcal{M}_{\rho^* \times (\rho \to o) \to \rho^\omega} \times \mathcal{M}_{\rho^*} \to \mathcal{M}_o$$

satisfying the defining equation for MBR_ρ has a majorant, and therefore lives in $\mathcal{M}$. In order to show that $\mathcal{M}$ is model of modified bar recursion, however, we must furthermore show that such a Φ exists. That is what we shall do now.

For any type ρ, the elements s of $\mathcal{M}_{\rho^*}$ (finite sequences of elements in ρ) can be viewed as nodes of an infinite tree which we shall call T. The infinite paths of T are the elements of $\mathcal{M}_{\rho^\omega}$ (which is just $\mathcal{M}^\omega_\rho$ as shown in [18]). For fixed Y and H, the functional Φ we are looking for should assign values to the nodes of T according to MBR. For each node s the set of nodes t extending s shall be denoted by B_s.

Let $Y, H \in \mathcal{M}$ be fixed. We show that at each infinite path α there exists a point n such that a functional $\Phi_{\alpha,n} : \mathcal{M}_{\rho^*} \to \mathcal{M}_o$ can be defined satisfying MBR for all $s \in B_{\overline{\alpha}n}$. Then, by bar induction, a functional Φ can be defined for all nodes of T.

Let $\alpha \in \mathcal{M}^\omega_\rho$ be fixed, n the number whose existence is stated in Lemma 8.5, and $K :\equiv \{0, 1, \dots, n\}$. We show how to define a functional $\Phi_{\alpha,n}(s)$ such that, for $s \in B_{\overline{\alpha}n}$, equation

$$\Phi_{\alpha,n}(s) = Y(s \mathbin{@} H(s, \lambda x.\Phi_{\alpha,n}(s * x))$$

holds. Here we note that, for $s \in B_{\overline{\alpha}n}$, by Lemma 8.5, $\Phi_{\alpha,n}(s)$ must belong to K. Therefore, for those $s \in B_{\overline{\alpha}n}$, what we have is an instance of the more general equation,

$$\Psi(s) = G(s, \lambda x.\Psi(s * x)), \tag{4.4}$$

where $\mathsf{img}(G) \subseteq K$. To see that modified bar recursion becomes an instance of (4.4), let

$$G(s, F) := Y(\overline{\alpha}n * s \mathbin{@} H(\overline{\alpha}n * s, F)),$$

and, clearly, $\mathsf{img}(G) = \mathsf{img}(\lambda s, F.Y(\overline{\alpha}n * s \mathbin{@} H(\overline{\alpha}n * s, F)) \subseteq K$. Hence, it suffices to show that equations of the form (4.4) (with the mentioned restriction on G) always have a solution Ψ.

Consider the set $\mathcal{T} :\equiv T \to 2^K \backslash \{\emptyset\}$. The set $\mathcal{T}$ can be viewed as the set of labelled trees whose labels range over non-empty subsets of K. We define a partial order $\sqsubseteq$ on $\mathcal{T}$ as follows

$$f \sqsubseteq g :\equiv \forall s.(f(s) \subseteq g(s)).$$

Finally, we define an operation $\chi : \mathcal{T} \to \mathcal{T}$,

$$\chi(f)(s) := \mathsf{img}(\lambda F \in \mathsf{Cons}^f_s.G(s, F)),$$

where $\mathsf{Cons}^f_s :\equiv \{F : \forall x^\rho.F(x) \in f(s * x)\}$. We first observe the following.

Lemma 4.6 $(\mathcal{T}, \sqsubseteq)$ *is directed complete semi-lattice.*

Proof. Let S be a directed subset of $\mathcal{T}$. Since we assign non-empty finite sets to the nodes of T, it is easy to see that $\bigcap S$ belongs to $\mathcal{T}$ and it is smaller than any element in S. □

Lemma 4.7 $\chi : \mathcal{T} \to \mathcal{T}$ *is monotone.*

Proof. Let $f \sqsubseteq g$ and s be fixed. We get that $\mathsf{Cons}^f_s \subseteq \mathsf{Cons}^g_s$, which implies $\chi(f)(s) \subseteq \chi(g)(s)$. □

By the Knaster-Tarski fixed point theorem (cf. e.g. [130]) we obtain an $f \in \mathcal{T}$ such that $\chi(f) = f$, i.e. $f(s) = \mathsf{img}(\lambda F \in \mathsf{Cons}^f_s.G(s, F))$, for all s. Let F_s be a functional from $f(s)$ to Cons^f_s such that $c = G(s, F_s(c))$, for all $c \in f(s)$. Define the functional $\Phi(s)$ recursively as follows,

$$\Psi(\langle\,\rangle) := \text{arbitrary element of } f(\langle\,\rangle);$$

$$\Psi(s * x) := F_s(\Psi(s))(x).$$

Lemma 4.8 *The functional Ψ is total and satisfies equation (4.4).*

Proof. We have just shown that Φ is total. Moreover, note that, for all s, the values assigned to $\Phi(s * x)$ are such that $\Phi(s) = G(s, \lambda x.\Phi(s * x))$. □

Theorem 4.4 *There exists a functional*

$$\Phi : \mathcal{M}_{\rho^\omega \to o} \times \mathcal{M}_{\rho^* \times (\rho \to o) \to \rho^\omega} \times \mathcal{M}_{\rho^*} \to \mathcal{M}_o$$

satisfying modified bar recursion.

Part II

Papers

The following papers have been included here as they (will) appear in print, with minor changes on notation for the sake of uniformity and occasional remarks added in square brackets i.e. [...].

Chapter 5

Proof Mining: A Systematic Way of Analyzing Proofs in Mathematics

with Ulrich Kohlenbach, to appear in: *Proc. Steklov Inst. Math.*, 33 pages.

Abstract

We call *proof mining* the process of logically analyzing proofs in mathematics with the aim of obtaining new information. In this survey paper we discuss, by means of examples from mathematics, some of the main techniques used in proof mining. We show that those techniques not only apply to proofs based on classical logic, but also to proofs which involve non-effective principles such as the attainment of the infimum of $f \in C[0,1]$ and the convergence for bounded monotone sequences of reals. We also report on recent case studies in approximation theory and fixed point theory where new results were obtained.

5.1 Introduction

Many theorems in mathematics can be expressed as simple equations e.g. stating that x as an element of some Polish space (complete separable metric space) X is a root of a function $f : X \to \mathbb{R}$. Theorems of this kind have been called *complete*. Such (essentially purely universal) theorems do not ask for any effective witnessing information. On the other hand, a theorem stating that f is (strictly) positive at a point $x \in X$ is *incomplete*, for it leaves open how far from zero the value $f(x)$ actually is. As a more intricate example, consider an implication between incomplete theorems such as

$$\forall x \in X \forall y \in K\ (f(x,y) > 0 \to g(x,y) > 0), \tag{5.1}$$

where $f, g : X \times K \to \mathbb{R}$ are continuous functions from the Polish space X and the compact Polish space K to the real numbers. Theorems of the form (5.1) can also be considered incomplete, since when $f(x,y)$ is apart from zero by ε, the value $g(x,y)$ must also be apart from zero by some δ. Until the relation between ε and δ is explicitly given theorem (5.1) would be considered incomplete. An implication between complete theorems can also be viewed as incomplete. Consider a theorem of the form

$$\forall x \in X \forall y \in K \ (f(x,y) = 0 \to g(x,y) = 0). \tag{5.2}$$

Theorem (5.2) does not tell us how close to zero $f(x,y)$ must be in order to make sure that $g(x,y)$ is ε-close to zero. So, one can ask for a functional Φ satisfying: If $|f(x,y)| \le \Phi(x,y,\varepsilon)$ then $|g(x,y)| \le \varepsilon$. This, of course, is just what (5.1) would give us applied to the classically equivalent form

$$\forall x \in X \forall y \in K \ (|g(x,y)| > 0 \to |f(x,y)| > 0),$$

of (5.2).

As we shall see in the following, the compactness of the space K will in general guarantee that such a Φ can be given independently of y.

It turns out that in many cases the information missing in an incomplete theorem can be extracted by purely logical analysis out of prima-facie ineffective proofs of the theorem. That is the main goal of proof mining. The program of proof mining goes back to G. Kreisel under the name of *unwinding proofs*[1]. Already in the 50's Kreisel called for a shift of emphasis in proof theoretic research guided by the question:

"What more do we know if we have proved a theorem by restricted means than if we merely know that it is true?"

Although proof mining has been applied e.g. to number theory [125, 126], combinatorics [10, 55] and algebra [40], the area of analysis, specially numerical functional analysis, is of particular interest. In analysis ineffectivity is due not only to the use of non-constructive logical reasoning but at the core of many principles (like compactness arguments) which are used to ensure convergence and which provably rely on the existence of non-computable reals. This paper surveys the main technique of monotone functional interpretation [92] currently used in proof mining in analysis and reports on recent case studies in approximation theory and fixed point theory where new results have been obtained.

The first step in analyzing the proof of a theorem consists of *fixing the formal system needed for carrying out the proof of the theorem.* That means: restricting the mathematical language and mathematical principles to be used in the proof. Fixing a restricted language enables us to pinpoint the logical form and logical complexity of the theorem. The restriction on the principles used dictates the techniques to be applied in the extraction and at the same time provides an a priori upper bound on the computational complexity of the functional realizing the theorem. The formal system which can be used to formalize a proof is clearly not unique. By showing that the proof can be formalized in a weak system interesting *a priori* information can be already obtained in this first step of proof mining. On the other hand, stronger systems will usually

[1] For discussions on the original program of Kreisel see [47, 126].

make the formalization of the proof and the extraction of information much simpler. Therefore, the choice of the mathematical strength of formal system is a compromise between a priori information and flexibility in formalizing the proof. As is confirmed by case studies, the proof theoretic techniques we are using are faithful to the numerical content of the **actual proof** analysed and the computational complexity of the extracted functional depends only on that proof, and not on the **formal system** used for the formalization and extraction. Hence, using weak systems is only an advantage when the a priori information is the only knowledge one wants to obtain. If the extraction of an actual functional is to be carried out, it is reasonable to choose a richer formal system in which proofs can be more easily formalized. The hard part then consists in performing the extraction of the functional. Therefore, in the present paper we shall mainly use Peano arithmetic in all finite types as the underlying arithmetical framework and focus on the next two steps of proof mining (for the study of weak fragments in the context of proof mining see e.g. [93, 95]).

The second task in analysing a theorem consists of *finding out which information the theorem could provide*. We will concentrate in this paper on theorems following the patterns (5.1) and (5.2) (or rather, a generalization of those two forms to be explained in the next section) and implications between them. As we shall see, it is a task on its own to realize that a theorem has this form. We devote Section 5.2 to explaining this process.

Finally, we must *carry out the extraction*. Once we know that some information can be extracted we shall look for an appropriate proof interpretation which will guide the process of extracting the information from the proof. The main goal of the article is to present in reasonable details the method of *monotone functional interpretation* [92] (to be presented in Section 5.3) combined with negative translation. We shall furnish the different steps of the interpretation with various examples from functional analysis. Based on these examples we will argue that (the combination of negative translation with) monotone functional interpretation (but not the usual Gödel functional interpretation as considered by Bishop [21]) in many cases provides the 'right' notion of **numerical implication** in analysis.

Note that the proof interpretations used here are purely syntactical transformations. Hence, given a **completely** formalized proof the extraction of information can be in principle done automatically via a computer[2]. The difficult part of proof mining would then consist in fully formalizing a mathematical proof originally given in ordinary mathematical terms. That can be in general very tiresome and intricate. Therefore, the case studies reported here have been carried out using the approach of partially formalizing only the relevant parts of a proof to the point where one can be sure that they can be completely formalized, and then carrying out the extraction 'by hand'. This can also be viewed as an advantage since when considering a particular proof various steps of the interpretations can be simplified.

In Section 5.4, we show that statements of the form (5.1) and (5.2) are

[2]Such a tool has been developed (cf. e.g. [17]) for a different proof interpretation based on modified realizability and A-translation.

in fact very common in mathematics. We carry out the monotone functional interpretation of those statements in order to show how concepts like modulus of uniqueness, continuity, monotonicity, contractivity, asymptotic regularity etc. naturally arise. In Section 5.5.1 we exemplify how this extends to implications between such statements. In the final three sections we treat more complex classes of proofs involving ineffective principles such as the attainment of the infimum for continuous functions on compact intervals and the principle of convergence for bounded monotone sequences or reals. We also report on recent extensive case studies where proofs involving those ineffective principles have been analyzed.

5.1.1 Formal Systems

Our base formal system consists of extensional classical arithmetic in all finite types $\mathsf{E\text{-}PA}^\omega$. In places where classical logic must/can be avoided we use intuitionistic arithmetic $\mathsf{E\text{-}HA}^\omega$ (for details see [160] where $\mathsf{E\text{-}PA}^\omega$ is denoted by $\mathsf{E\text{-}HA}^\omega_c$). The finite types are inductively defined as: $\mathbb{N}$ is a finite type and if ρ and τ are finite types then $\rho \to \tau$ is a finite type. An object of type $\rho \to \tau$ denotes a mapping from objects of type ρ to objects of type τ. We often abbreviate the type $\mathbb{N} \to \mathbb{N}$ as 1.

We denote by $\mathcal{T}^\omega$ both $\mathsf{E\text{-}PA}^\omega$ as well as various subsystems of $\mathsf{E\text{-}PA}^\omega$ such as PRA^ω (cf. [5]) and $\mathsf{E\text{-}G}_n\mathsf{A}^\omega$ (cf. [93]). $\mathcal{T}_i^\omega$ is the intuitionistic counterpart of $\mathcal{T}^\omega$. We work in systems containing equality ($=$) between objects of type $\mathbb{N}$ as the only predicate symbol. Equality between higher types is defined extensionally. In the same way the (pointwise) partial order $\leq_\rho$ between objects of type ρ is defined as: $x \leq_{\rho\to\tau} y :\equiv \forall z^\rho\, (x(z) \leq_\tau y(z))$. Note that all quantifier-free formulas in our systems are decidable and can even be written as atomic formulas. We shall usually add to the base system $\mathcal{T}^\omega$ the axiom of quantifier-free choice[3]

$$\mathsf{QF\text{-}AC}^{1,\mathbb{N}} \quad : \quad \forall f^1 \exists n^{\mathbb{N}}\, A_0(f,n) \to \exists \Phi \forall f\, A_0(f, \Phi f).$$

5.2 Representation

As already mentioned, our formal systems only contain equality between natural numbers as a primitive notion. Therefore, when talking about more complex mathematical objects such as rationals, reals, continuous functions, etc. we first need to fix their representation in the system. Equality between those objects will then be defined extensionally. As a simple example we take the rational numbers which can be represented via coding of pairs into the natural numbers. Assuming the representation of the rational numbers, real numbers are represented via (representations of) Cauchy sequences $\psi : \mathbb{N} \to \mathbb{Q}$ with fixed rate of convergence say 2^{-n}, i.e. a real number x is represented by a function $\psi_x : \mathbb{N} \to \mathbb{Q}$ satisfying $\forall n \forall m, \tilde{m} \geq n\; (|\psi_x(m) -_{\mathbb{Q}} \psi_x(\tilde{m})| \leq 2^{-n})$. In a roughly similar way elements of Polish spaces X are represented as type one objects x^1 (i.e. elements in the Baire space) via the so-called standard representation (see e.g. [9]). For particular spaces, often more convenient (though essentially

[3]Here and in the following, $A_0, B_0, C_0, \ldots$ always denote quantifier-free formulas.

equivalent) representations can be used. For instance, take the Polish space $(C[0,1], \|\cdot\|_\infty)$ of all the real valued continuous functions on the interval $[0,1]$ with the uniform norm as metric. A function $f \in C[0,1]$ is represented via a pair of functions (f_r, ω_f) where f_r is the restriction of f to the rational numbers and ω_f is the modulus of uniform continuity of f (on $[0,1]$). Note that both f_r and ω_f can be further represented as type one functions. Operations on Polish spaces are then represented as type two objects and so on.

Returning to the issue of equality, given two real numbers x and y represented via ψ_x and ψ_y, the statement $x =_{\mathbb{R}} y$, on the level of representation, is defined as the Π^0_1-formula $\forall n(|\psi_x(n+1) -_{\mathbb{Q}} \psi_y(n+1)| \leq_{\mathbb{Q}} 2^{-n})$. Similarly, $x <_{\mathbb{R}} y$ is expressed by the Σ^0_1-formula $\exists n(\psi_y(n+1) -_{\mathbb{Q}} \psi_x(n+1) \geq_{\mathbb{Q}} 2^{-n})$. In order to discover the information hidden in the statement of a theorem, it is important to explicitly present all the quantifiers *hidden* in such defined equality notions for Polish spaces. In order to avoid to have to go down all the way to the intensional level of representations, it is very useful to note that $x =_{\mathbb{R}} y$ is equivalent to **both** $\forall n\ (|x-y| \leq 2^{-n})$ and $\forall n\ (|x-y| < 2^{-n})$. Although the matrices in both statements are still Π^0_1 and Σ^0_1 respectively, we can treat them as if they were quantifier-free since we can always choose the suitable form which does not increase the general logical form of the theorem is question. In this way, we have presented the hidden quantifiers of the equality without having to go into the representations of the real numbers x and y.

The representation of Polish spaces X can be arranged in such a way that **every** x^1 represents **some** element of X (see [9] and [90] for details).

For compact Polish spaces K one can achieve that the representatives ψ are always number theoretic functions which are bounded by some fixed term s (even by the constant-1 function, i.e. by elements in the Cantor space). Let X and Y be Polish spaces. Moreover, let $\{K_x\}_{x\in X}$ be a family of compact subsets of Y parametrized by elements $x \in X$ (e.g. $X = \mathbb{R}^+$, $Y = \mathbb{R}^2$ and $K_x = [-x,x]^2$). If the family $\{K_x\}_{x\in X}$ is sufficiently constructively given (see [90], Def. 3.22) the elements $z \in K_x$ can be represented as $z \leq_1 sx$, for a fixed term s. Again one can achieve that **every** function in that bounded set represents some element of the space. Details on all this can be found in [9,90] and – for very weak systems – in [95].

According to the representation, mathematical statements of the form (5.1) and (5.2) have logical counterparts

$$\forall x^1 \forall y^1 \leq s(\exists n A_0(x,y,n) \to \exists m B_0(x,y,m)), \tag{5.3}$$

$$\forall x^1 \forall y^1 \leq s(\forall n A_0(x,y,n) \to \forall m B_0(x,y,m)), \tag{5.4}$$

respectively. Note, moreover, that (5.3) and (5.4) are special cases of [4]

$$\forall x^1 \forall y^1 \leq \tilde{s}x \exists z^{\mathbb{N}}\, B_0(x,y,z), \tag{5.5}$$

which in mathematical terms corresponds to statements of the form

$$\forall x \in X \forall y \in K_x \exists z \in \mathbb{N}\, B_1(x,y,z), \tag{5.6}$$

[4] Using that '$\forall x^1, n^{\mathbb{N}}$' can be contracted to '$\forall x^1$'. Actually, we do not even need such encodings as our techniques are directly applicable to tuples $\vec{x}$ of variables of degree ≤ 1 instead of x^1.

where X is some Polish space, K_x a compact Polish space parametrized by x, and B_1 is a purely existential formula (due to the quantifiers still present in e.g. $|x-y| < 2^{-n}$ as discussed above).[5]

For fixed Polish spaces X, Y and a family of compact sets $K_x \subset Y$, (5.6) can be viewed (via the representation) as a special case of (5.5). On the other hand, (5.5) can also be considered as a special case of (5.6), taking X as the Baire space and $K_x = \{y : y \leq \tilde{s}x\}$.

5.3 Monotone Functional Interpretation

The functional ('Dialectica') interpretation introduced by Gödel [58] translates an arbitrary formula A in the language of $\mathsf{E\text{-}HA}^\omega$ into another formula A^D (in the same language) having the form $\exists x \forall y A_D(x,y)$, for some quantifier free formula A_D.[6] The translation is sound in the sense that if the formula A has been proved in $\mathsf{WE\text{-}HA}^\omega$ then from that proof one can extract a closed term t such that $A_D(t,y)$ is provable in $\mathsf{WE\text{-}HA}^\omega$.[7] The soundness theorem has been adapted to many other systems both stronger ones as well as fragments of $\mathsf{WE\text{-}HA}^\omega$. Via negative translation (and elimination of extensionality) it also applies to $\mathsf{E\text{-}PA}^\omega$ and related systems (cf. [5, 124, 160]).

Note that the formula $A_D(t,y)$ is quantifier free, but will usually contain terms of higher types, even if all the terms in the original formula A have the type $\mathbb{N}$.

Definition 5.1 (Functional Interpretation) *The interpretation associates to each formula $A \in \mathcal{L}^\omega$ (by induction on the logical structure of A) another formula $(A)^D$ of the form $\exists x \forall y A_D(x,y)$, where A_D is quantifier free, in the following manner:*

$$A^D :\equiv A, \text{ for atomic formulas } A,$$

and assuming $A^D = \exists x \forall y A_D(x,y)$ and $B^D = \exists z \forall w B_D(z,w)$ we define

$$(A \wedge B)^D :\equiv \exists x, z \forall y, w (A_D(x,y) \wedge B_D(z,w)),$$

$$(A \vee B)^D :\equiv \exists p^{\mathbb{N}} \exists x, z \forall y, w ((p = 0 \to A_D(x,y)) \wedge (p \neq 0 \to B_D(z,w))),$$

$$(A \to B)^D :\equiv \exists \Psi, \Phi \forall x, w (A_D(x, \Phi x w) \to B_D(\Psi x, w)),$$

$$(\exists z A(z))^D :\equiv \exists z, x \forall y A_D(x,y,z),$$

$$(\forall z A(z))^D :\equiv \exists \Psi \forall z, y A_D(\Psi z, y, z),$$

[5]Note that the fact that B_1 is purely existential just adds some more existential quantifiers to '$\exists z^{\mathbb{N}}$'.

[6]Actually, x, y are both tuples of variables whose length depends on the logical form of A. For simplicity we suppress the (correct) tuple notation here.

[7]Here $\mathsf{WE\text{-}HA}^\omega$ is a version of $\mathsf{E\text{-}HA}^\omega$ where the extensionality axioms in higher types are restricted to a quantifier-free rule of extensionality [160]. Such a restriction – which is necessary for the soundness theorem to hold (see [65]) – does not cause any problems for the applications treated in this paper since all the principles and theorems we consider are – because of their type restrictions – such that the 'elimination-of-extensionality'-procedure from [124] applies.

where the types of Ψ *and* Φ *can be inferred. We define* $\neg A$ *as* $A \to 0 = 1$.

The most intricate interpretation is that of the implication. Let us analyse the functional interpretation of implication when both formulas A and B have the special form $\exists x C_0(x)$ or $\forall x C_0(x)$ (with C_0 quantifier-free). Here we get (using implicitly that quantifier-free formulas $A_0(a)$ can be written as atomic ones $t_{A_0}(a) =_{\mathbb{N}} 0$ for suitable closed t_{A_0})

$$(\exists x A_0(x) \to \exists y B_0(y))^D \equiv \exists \Phi \forall x (A_0(x) \to B_0(\Phi x))$$

and

$$(\forall x A_0(x) \to \forall y B_0(y))^D \equiv \exists \Phi \forall y (A_0(\Phi y) \to B_0(y)).$$

This also holds if first negative translation has been applied, since

$$(\neg \forall x \neg C_0(x))^D \equiv (\neg\neg \exists x C_0(x))^D \equiv \exists x \neg\neg C_0(x) \leftrightarrow \exists x C_0(x),$$

modulo stability of atomic formulas under double negation.

Note that e.g. the more simple modified realizability interpretation [158] only delivers a result in the first case above (and if negative translation had been applied first, not even then). In Section 5.4 we shall see various examples of statements, commonly used in numerical analysis, having the forms $\exists x A_0(x) \to \exists y B_0(y)$ and $\forall x A_0(x) \to \forall y B_0(y)$. A detailed analysis of the treatment given to implication by functional interpretation can be found in [5].

We call *extraction procedure* the process of producing out of a proof of a sentence A a (tuple of) closed term(s) t of the underlying system and a proof of $A_D(t, y)$. The soundness proof of functional interpretation actually provides such an extraction procedure. If only a *bound* on the term t is of interest a much simpler extraction procedure can be used. This variant of the extraction procedure which looks for a hereditarily monotone bound on the realizer of $\exists x \forall y A_D(x, y)$ we call (cf. [92]) *monotone functional interpretation*, or *m.f.i.* for short. In [92] it is shown that the soundness theorem for the m.f.i. can be directly proved on the level of the monotone version, i.e. without the need to compute first a realizer t as an intermediate step.

In order to make the notion of 'bound' well behaving in higher types we use Bezem's [18] strong majorizability relation $\geq^{\mathsf{m}}_{\rho}$, which is a variant of Howard's [65] original hereditarily majorability relation. For numbers $n \geq^{\mathsf{m}}_{\mathbb{N}} m$ just means that n is greater or equal than m. For functions f and g, $f \geq^{\mathsf{m}}_{1} g$ holds when f is monotone and is pointwise bigger than g. For higher types the relation is designed to be hereditarily monotone, i.e.

$$\Phi^* \geq^{\mathsf{m}}_{\rho \to \tau} \Phi :\equiv \forall x^* \forall x \leq^{\mathsf{m}}_{\rho} x^* (\Phi^* x^* \geq^{\mathsf{m}}_{\tau} \Phi^* x \wedge \Phi^* x^* \geq^{\mathsf{m}}_{\tau} \Phi x.)$$

Three important properties of the relation $\geq^{\mathsf{m}}_{\rho}$ are:

i) $x \geq^{\mathsf{m}}_{\rho} y$ implies $x \geq^{\mathsf{m}}_{\rho} x$,

ii) $x \geq^{\mathsf{m}}_{\rho} y \wedge y \geq_{\rho} z \to x \geq_{\rho} z$, ($\geq_{\rho}$ as defined in Section 5.1.1)

iii) for type one objects x^1, i.e. number theoretic functions, the function

$$x^+ :\equiv \lambda n. \max_{m \le n} x(m)$$

always majorizes x.

Note that $\ge_\rho^{\mathsf{m}}$ is not reflexive unless $\rho = \mathbb{N}$.

Using the relation $\ge_\rho^{\mathsf{m}}$, the *monotone functional interpretation (m.f.i.)* of a formula A (having functional interpretation $\exists x^\rho \forall y^\tau A_D(x, y)$) is defined as

$$\exists x^* \exists x \le_\rho^{\mathsf{m}} x^* \forall y A_D(x, y).$$

Theorem 5.1 ([89]) *Let Δ be a set of closed axioms of the form*

$$\forall u^1 \exists v^1 \le tu \forall w^{\mathbb{N}} A_0(u, v, w), \text{ where } t \text{ is closed.}$$

Suppose that [8]

$$\mathcal{T}^\omega + \mathsf{QF\text{-}AC}^{1,\mathbb{N}} + \Delta \vdash \forall x^1 \forall y^1 \le sx \exists z^{\mathbb{N}} B_0(x, y, z).$$

From this proof one can extract a closed term Φ of $\mathcal{T}^\omega$ such that,

$$\mathcal{T}_i^\omega + \Delta_\varepsilon \vdash \forall x^1 \forall y^1 \le sx \exists z \le \Phi x\, B_0(x, y, z),$$

where Δ_ε consists of the so-called ε-weakenings of the sentences in Δ, i.e.

$$\forall u^1, w^{\mathbb{N}} \exists v^1 \le tu \forall i \le w A_0(u, v, i).$$

As shown in [89], the set of sentences Δ also includes the non-computational principle weak König lemma (WKL). Since $\mathsf{WE\text{-}HA}^\omega \vdash \mathsf{WKL}_\varepsilon$, this provides a WKL-elimination.

The result above can also be stated in more mathematical terms. Let INF denote the principle

$$\forall f \in C[0, 1] \exists x \in [0, 1](f(x) \stackrel{\mathbb{R}}{=} \inf_{y \in [0,1]} f(y)),$$

which can – using the representation of $C[0, 1]$ – be written in form Δ (see [90]). Note that INF_ε is equivalent to

$$\forall f \in C[0, 1] \forall n \exists x \in [0, 1](f(x) \le \inf_{y \in [0,1]} f(y) + 2^{-n}),$$

which, given our representation of $f \in C[0, 1]$, can be easily proved in $\mathsf{WE\text{-}HA}^\omega$. One example of a corollary of Theorem 5.1 would be the following.

Theorem 5.2 ([90]) *Let (X, d_X) be a $\mathcal{T}^\omega$-definable Polish space and $\{K_x\}_{x \in X}$ a $\mathcal{T}^\omega$-definable family of compact sets in a Polish space Y. If*

$$\mathcal{T}^\omega + \mathsf{QF\text{-}AC}^{1,\mathbb{N}} + \mathsf{INF} \vdash \forall x \in X \forall y \in K_x \exists z \in \mathbb{N} B_1(x, y, z)$$

then, from this proof one can extract a closed term Φ of $\mathcal{T}^\omega$ such that,

$$\mathcal{T}_i^\omega \vdash \forall x \in X \forall y \in K_x \exists z \le \Phi x\, B_1(x, y, z),$$

[8] $B_0(x, y, z)$ contains no other free variables than x, y, z and that s is a closed term.

where $B_1(x, y, z)$ is a Σ^0_1-formula (not containing further free variables) which is (provably in $\mathcal{T}^\omega$) extensional in x, y w.r.t. the relations $=_X$ and $=_{K_x}$.

Remark 5.1 The constructivisation of the given proof provided by the meta-theorems due to the reduction of the use of Δ to that of Δ_ε is quite independent from the construction of the bound which first uses even a stronger Skolemized version of Δ which then by subsequent manipulations can be reduced to Δ_ε. These subsequent steps can be omitted in applied proof mining. So the final proof of the result will normally again be ineffective although the meta-theorems guarantees that it can in principle be made constructive.

Note that, besides the simplicity of the extraction procedure, using m.f.i. one obtains bounds which are independent of all parameters ranging over compact spaces.

The proofs of both meta-theorems above rely on the combination of negative translation and m.f.i. These two meta-theorems are just special cases of a whole class of more general theorems proved by the first author in the papers cited and - for weak fragments - in [93]. In particular, many more analytical principles than INF can directly be seen to have the form Δ which avoids to have to analyse their proofs (say via WKL) in the proof mining process. Other WKL-related principles which do not have that form usually easily follow from a nonstandard principle of uniform boundedness (studied in [98, 102]) which is allowed to be used in the meta-theorems and can be eliminated from the proof of the conclusion. In this way large parts of given proofs can simply be skipped in the process of proof mining.

Whereas – as Theorem 5.2 shows – principles based on Heine-Borel compactness (WKL) do not contribute to the growth of extractable bounds, principles based on sequential compactness do contribute. Monotone functional interpretation (combined with a specially designed method of eliminating monotone Skolem functions) allows to calibrate the exact contribution of fixed instances of sequential compactness relative to weak fragments $\mathcal{T}^\omega$ (see [94]). We shall discuss this in more detail in Sections 5.6 and 5.7.

Another important observation is that the bound Φ above will depend on the representation of x and will therefore not be an extensional function $X \to \mathbb{N}$. In practice, however, Φ will usually be extensional in some natural enrichments of the input. The dependence on the representation is unavoidable in general. Consider the space $X = \mathbb{R}$. The only effective extensional (and therefore continuous) functions $\Phi : \mathbb{R} \to \mathbb{N}$ would be constant functions.

Notation 5.1 *For the rest of the paper all the Polish spaces are understood to be $\mathcal{T}^\omega$-definable. Examples of $\mathcal{T}^\omega$-definable Polish spaces are $(\mathbb{R}^n, d_E)$, $(\mathbb{R}^n, d_{\max})$, $(C[0,1], d_\infty)$ and (L_p, d_p) for $1 \leq p < \infty$.*

5.3.1 Monotone Functional Interpretation of Theorems Having the Form (5.5)

In Bishop [21] some arguments are given in favour of taking the functional interpretation of implication as *numerical implication*, i.e. given a theorem C of the form

$$\exists x \forall y A_0(x,y) \to \exists z \forall w B_0(z,w),$$

A_0 and B_0 quantifier free, Bishop suggests that the numerical content of the theorem C is given by the existential quantifier in

$$C^D \equiv \exists Z, Y \forall x, w (A_0(x, Yxw) \to B_0(Zx, w)).$$

In the following we argue, by considering implications between statements of the form (5.5) that if one is interested in uniform bounds (which is usually the case in analysis, see below) the m.f.i. provides exactly the right kind of numerical information. As mentioned above, statements in analysis which have the logical form (5.5) appear in the special forms (5.3) and (5.4). Let us first analyze, from a purely logical point of view, how m.f.i. treats such statements. It is important to note that for statements of this form there is no difference whether m.f.i. is applied directly or to their negative translation, since (as discussed for the usual functional interpretation above) m.f.i. treats $\neg\neg\exists y A_0(x,y)$ and $\neg\forall y \neg A_0(x,y)$ as $\exists y A_0(x,y)$.[9] This also means that m.f.i. treats the negative statement $\neg(x =_{\mathbb{R}} 0)$ as the positive $|x| >_{\mathbb{R}} 0$. Therefore, in the following we only consider the monotone functional interpretation. The m.f.i. of (5.4) gives [10]

$$\exists \Phi^* \exists \Phi \ \leq^{\mathsf{m}} \ \Phi^* \forall x \forall y \leq sx \forall m \, (A_0(x,y,\Phi xym) \to B_0(x,y,m)),$$

which is equivalent (by elementary constructive reasoning)[11] to

$$\exists \Psi^* \ \leq^{\mathsf{m}} \ \Psi^* \forall x \forall y \leq sx \forall m \exists n \leq \Psi^* x^+ (s^* x^+) m \, (A_0(x,y,n) \to B_0(x,y,m)).$$

The formula above is in turn equivalent to

$$\exists \Psi \ \leq^{\mathsf{m}} \ \Psi \forall x \forall y \leq sx \forall m \, (\forall n \leq \Psi xm \, A_0(x,y,n) \to B_0(x,y,m)).$$

In the same way, the monotone functional interpretation of (5.3) is equivalent to

$$\exists \Psi \ \leq^{\mathsf{m}} \ \Psi \forall x \forall y \leq sx \forall n \, (A_0(x,y,n) \to \exists m \leq \Psi xn \, B_0(x,y,m)).$$

In Section 5.4, we shall consider various mathematical concepts which have the logical form (5.1) and (5.2) (the mathematical counterparts of (5.3) and (5.4)) and therefore the form (5.6) where B_1 is monotone in 'z' so that any (uniform) bound in fact provides a (uniform) realizer. For each of those statements we indicate the mathematical importance of the m.f.i., by showing that

[9] In logical terms this is due to the fact that m.f.i. (as functional interpretation) satisfies the so-called Markov principle. As we are mainly interested in proofs based on full classical logic it is indeed the m.f.i. of the negative translation of a statement A which matters.

[10] Note that the universal quantifier 'hidden' in $y \leq_1 sx$ is not essential, for using extensionality one can prove that $\forall y \leq sx A(y)$ is equivalent to $\forall y A(\min_1(y, sx))$, where $\min_1(x,y) :\equiv \lambda n. \min(x(n), y(n))$.

[11] In the direction '$\to$' we can take $\Psi^* :\equiv \Phi^*$. In the other direction, suppose that Ψ^* satisfies the second formula. Then

$$\Phi^* xym :\equiv \Psi^* x^+ (s^* x^+) m \text{ and}$$

$$\Phi xym :\equiv \min i \leq \Phi^* xym \, [A_0(x,y,i) \to B_0(x,y,m)]$$

satisfy the first formula.

the modulus Ψ corresponds to an important analytical concept which has been studied extensively in the literature.

The fact that Ψ majorizes itself implies an important monotonicity behaviour. Assume we have shown that a Ψ (majorizing itself) exists such that

$$\forall x^1 \forall y \leq sx \exists n \leq \Psi x B_0(x, y, n).$$

Let t^1 be some closed term. By restricting the variable x to be bounded by t we immediately obtain the existence of a functional $\tilde{\Psi} :\equiv \Psi(t^+)$ (independent of x and y) such that

$$\forall x \leq t \forall y \leq sx \exists n \leq \tilde{\Psi} B_0(x, y, n).$$

In mathematical terms, assume that a modulus Ψ depends on an element x of some Polish space X. By restricting x to some compact subspace $K \subseteq X$ we automatically obtain a modulus Φ independent of x (but which will depend only on some information about the compact space K). An instance of this general fact can be seen in Proposition 5.3, where we restrict $f \in C[0,1]$ to functions with common modulus of uniform continuity and bounded uniform norm, therefore obtaining independence from the function f.

We shall also see in the next section that inter-relations between such moduli created by m.f.i. play an important role in numerical functional analysis. We investigate this in more detail in Section 5.5, where we explain how monotone functional interpretation naturally transforms those moduli into one another via the treatment of implications.

5.4 Applying Monotone Functional Interpretation to Mathematics

In the following we consider what m.f.i. does when applied to standard concepts used in mathematics of the logical form treated in the previous section. As we shall see, in each case the interpretation suggests the existence of a modulus which corresponds to extensively studied analytical concepts. That indicates that, via a purely logical analysis, m.f.i. will in general ask/create the 'right' effective information about a theorem. As discussed in the previous section, there is no difference between the m.f.i. of a statement (5.6) and the m.f.i of its negative translation so that we only have to consider the former.

We should keep in mind that – as mentioned already – the functionals created by m.f.i. operate on the representation of mathematical objects in the formal system, rather than on the actual objects. For instance, a functional from a Polish space X to the rational numbers will have type $\mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ and will not be extensional in general.

5.4.1 Uniqueness

Let (X, d_X) and (K, d_K) be Polish spaces, K compact. The fact that a $\mathcal{T}^\omega$-definable (and hence continuous) function $f : X \times K \to \mathbb{R}$ for each given $x \in X$ has at most one root in K can be expressed as

$$\mathsf{UNI}(f) :\equiv \forall x \in X; y_1, y_2 \in K(\bigwedge_{i=1}^{2} f(x, y_i) \stackrel{\mathbb{R}}{=} 0 \to d_K(y_1, y_2) \stackrel{\mathbb{R}}{=} 0),$$

which has the form (5.2). The *monotone functional interpretation* of a uniqueness statement of the form UNI creates a modulus $\Phi : \mathbb{N}^{\mathbb{N}} \times \mathbb{Q}_+^* \to \mathbb{Q}_+^*$ such that

$$\forall x \in X; y_1, y_2 \in K; \varepsilon \in \mathbb{Q}_+^*(\bigwedge_{i=1}^{2} |f(x, y_i)| < \Phi(x, \varepsilon) \to d_K(y_1, y_2) < \varepsilon),$$

named *modulus of uniqueness* in [90]. The notion of modulus of uniqueness shows up e.g. in approximation theory where it has been extensively studied under the name of *strong unicity* or *rate of strong uniqueness*. For the case of Chebysheff approximation this was first investigated in [131]. For L_1-approximation strong unicity was studied e.g. by Björnestål [22, 23] and Kroó [118, 120]. See [7] for a survey on the relevance of this concept.

We mention here two applications of moduli of uniqueness. First, assume that K is a compact subset of the Polish space X and that each element of $x \in X$ has a unique best approximation in K w.r.t. the metric d_X. A modulus of uniqueness Φ in this case provides necessary a priori information for computing the best approximation of x, uniformly in x, in the following way. Define $f(x, y) :\equiv d_X(x, y) - dist(x, K)$, where $dist(x, K) :\equiv \inf_{y \in K} d_X(x, y)$. If X and K are effective spaces, then one can compute approximate solutions, i.e. elements $y \in K$ such that $|f(x, y)| < \varepsilon$. Let $(y_n)_{n \in \mathbb{N}}$ be a sequence of elements of K such that $|f(x, y_n)| < \Phi(x, 2^{-n})$. Then – applying Φ to y_n and the best approximation y_b one infers that the sequence $(y_n)_{n \in \mathbb{N}}$ converges to the best approximation $y_b \in K$ of x with rate of convergence 2^{-n}, i.e. $d_X(y_b, y_n) < 2^{-n}$. Note that it is crucial for the procedure above to be useful that Φ does not depend on y_1 nor y_2, since it gets applied to context where one of the polynomials is the unknown y_b. Further details can be found in [90].

Under the assumptions above, define $\mathcal{P} : X \to K$ to be the functional which maps x to its unique best approximation in K. As shown in [90], a modulus of uniqueness Φ automatically gives a modulus of pointwise continuity for the projection $\mathcal{P}$, also called *rate of smoothness/continuity*,

$$\forall x, y \in X(d_X(x, y) < \tfrac{1}{2}\Phi(x, \varepsilon) \to d_X(\mathcal{P}(x), \mathcal{P}(y)) < \varepsilon).$$

Again, the relationship between strong uniqueness and the smoothness of the projection operator has been studied extensively in the literature (cf. [2,3,8,23]).

5.4.2 Convexity

Let $(X, \|\cdot\|)$ denote a normed linear space whose unit ball $B :\equiv \{x \in X : \|x\| \leq 1\}$ is compact (which – classically – amounts to X being finite dimensional). From the statement that X is strictly convex

$$\mathsf{CVX} :\equiv \forall x, y \in B(\|\tfrac{1}{2}(x + y)\| \stackrel{\mathbb{R}}{=} 1 \to \|x - y\| \stackrel{\mathbb{R}}{=} 0),$$

which is again of the form (5.2), *monotone functional interpretation* creates a modulus $\eta : \mathbb{Q}_+^* \to \mathbb{Q}_+^*$ satisfying

$$\forall x, y \in B; \varepsilon \in \mathbb{Q}_+^*(\|\tfrac{1}{2}(x + y)\| > 1 - \eta(\varepsilon) \to \|x - y\| < \varepsilon).$$

If a normed space has such a modulus η it is called *uniformly convex*. Moreover, η is called *modulus of uniform convexity*. The crucial feature of uniform convexity, compared to strict convexity, is that $\eta(\varepsilon)$ does not depend on x, y. It is well known that finite dimensional strictly convex normed spaces are uniformly convex. Monotone functional interpretation provides an effective version of this: From a proof of strict convexity of a compact unit ball one can extract a modulus of uniform convexity, provided the proof and the space can be represented in an appropriate formal system.

The notion of uniform convexity was introduced in 1936 by Clarkson [35] (see also [109]) and plays a crucial role in many parts of functional analysis. This is true, in particular, for the area of metric fixed point theory (see e.g. [29, 60, 61]). Here moduli of uniform convexity have been used to determine rates of convergence for Krasnoselski-Mann iterations of nonexpansive mappings which connects this concept with the concepts of *rates of monotone convergence* and *rate of asymptotic regularity* to be discussed in Sections 5.4.6 and 5.5.1 (cf. [30, 75, 100, 103]).

Moduli of uniform convexity also feature prominently in the area of best approximation theory, having a close connection with rates of strong unicity and rates of smoothness/continuity, concepts discussed in Sections 5.4.1 and 5.4.4. Among the many publications on the connection between moduli of uniform convexity and rates of strong unicity see e.g. [23, 68, 123, 140].

5.4.3 Contractivity

Let (K, d) be a compact Polish space. A function $f : K \to K$ is defined to be *contractive* if[12]

$$\mathsf{CTR}(f) :\equiv \forall x, y \in K(x \neq y \to d(f(x), f(y)) < d(x, y)),$$

which has the form (5.1). The *monotone functional interpretation* of the statement that a $\mathcal{T}^\omega$-definable f is contractive creates a modulus $\eta : \mathbb{Q}^*_+ \to \mathbb{Q}^*_+$ satisfying

$$\forall x, y \in K; \varepsilon \in \mathbb{Q}^*_+(d(x, y) > \varepsilon \to d(f(x), f(y)) + \eta(\varepsilon) < d(x, y)).$$

The concept of contractivity can be written also in the trivially equivalent form

$$\forall x, y \in K(x \neq y \to \exists n \in \mathbb{N}(d(f(x), f(y)) < (1 - 2^{-n}) \cdot d(x, y))),$$

in which case the interpretation yields a modulus $\tilde{\eta} : \mathbb{Q}^*_+ \to \mathbb{N}$ satisfying

$$\forall x, y \in K; \varepsilon \in \mathbb{Q}^*_+(d(x, y) > \varepsilon \to d(f(x), f(y)) < (1 - 2^{-\tilde{\eta}(\varepsilon)}) \cdot d(x, y)).$$

Such a modulus $\alpha(\varepsilon) :\equiv 1 - 2^{-\tilde{\eta}(\varepsilon)}$ has in fact been considered in the literature by Rakotch [141] and – in the context of Bishop style constructive analysis – in [28]. Using the boundedness of K, we can easily produce an η out of a given α and vice-versa.

[12] We may in fact consider the more general case of functions $f : X \times K \to K$, where X is a Polish space, in which case the modulus η will also depend on (a representation of) $x \in X$. Similarly in Section 5.4.4 below.

As we will show in Section 5.5.1, it is exactly such a modulus which is needed to obtain a rate of convergence in Edelstein's fixed point theorem [42, 141]. As in the case of moduli of uniqueness it is crucial here that η does not depend on x, y.

Numerous variants of the notion of 'contractive mapping' have been considered in the literature. The main purpose of those variants is to obtain generalizations of Edelstein's classical fixed point theorem to more general classes of functions. Under monotone functional interpretation, those notions again give rise to appropriate moduli, and we expect that in many of these cases explicit rates of convergence can be provided in terms of the corresponding moduli of contractivity. For a survey of 25 notions of contractivity and generalizations of Edelstein's result see [144]. This line of work is further continued in [36,129,145], to list only a few references.

5.4.4 Uniform continuity

Let (X, d_X) and (K, d_K) be Polish spaces, K compact. From the statement that a $\mathcal{T}^\omega$-definable $f : K \to X$ is a function

$$\mathsf{CTN}(f) :\equiv \forall x, y \in K(x \overset{K}{=} y \to f(x) \overset{X}{=} f(y)),$$

which has the form (5.2), *monotone functional interpretation* creates a modulus $\omega : \mathbb{Q}^*_+ \to \mathbb{Q}^*_+$ satisfying

$$\forall x, y \in K; \varepsilon \in \mathbb{Q}^*_+ (d_K(x, y) < \omega(\varepsilon) \to d_X(f(x), f(y)) < \varepsilon).$$

Such ω plays a fundamental role in constructive mathematics (see [20]) and in computable analysis (see [83], [139] and [161]) where it is called *modulus of uniform continuity.* Numerous results indicate that ω provides the right computational information on continuous functions. For example, a function $f : [0, 1] \to \mathbb{R}$ which maps computable sequences in $[0, 1]$ into computable sequences in $\mathbb{R}$ has an effective uniform approximation by polynomials iff f has a computable modulus of uniform continuity ω (see [139]). On the other hand, numerical analysts define the function

$$\Omega(\varepsilon) :\equiv \sup_{d_K(x,y) \le \varepsilon} d_X(f(x), f(y))$$

to be the *modulus of continuity of* f. The function Ω clearly satisfies

$$\forall x, y \in K; \varepsilon \in \mathbb{Q}^*_+ (d_K(x, y) \le \varepsilon \to d_X(f(x), f(y)) \le \Omega(\varepsilon))$$

and is, in contrast to ω, unique. The continuity of f is now expressed as

$$\varepsilon \searrow 0 \to \Omega(\varepsilon) \searrow 0.$$

Apparently, the notions introduced by monotone functional interpretation and numerical analysis differ. However, one can observe that in analysis (cf. [120]) the modulus Ω is often used just for building a

$$\Omega^{-1}(\varepsilon) :\equiv \inf\{\delta \in [0, 1] : \Omega(\delta) = \varepsilon\},$$

which is a roundabout and ineffective way of creating a particular modulus ω. That once again supports the thesis that monotone functional interpretation produces, by purely logical analysis, the right constructive modulus.

5.4.5 Monotonicity

Let $f : [0,1] \to \mathbb{R}$ be a $\mathcal{T}^\omega$-definable strictly increasing (decreasing) function, i.e.,

$$\mathsf{MON}(f) :\equiv \forall x, y \in [0,1](x - y > 0 \to f(x) - f(y) > 0),$$

which has the form (5.1). From this statement *monotone functional interpretation* creates a modulus $\delta : \mathbb{Q}^*_+ \to \mathbb{Q}^*_+$ such that

$$\forall x, y \in [0,1]; \varepsilon \in \mathbb{Q}^*_+ (x - y > \varepsilon \to f(x) - f(y) > \delta(\varepsilon)),$$

called *modulus of monotonicity*. Note that the modulus of monotonicity δ provides a modulus of uniform continuity for the inverse function f^{-1}.

5.4.6 Monotone Convergence

Let X and K be Polish spaces, K compact. Moreover, let $f : X \times K \times \mathbb{N} \to \mathbb{R}_+$ be a function such that for any $x \in X$ and $y \in K$ the sequence $(f(x,y,n))_{n\in\mathbb{N}}$ is non-increasing. Suppose that $(f(x,y,n))_{n\in\mathbb{N}}$ converges to zero

$$\mathsf{CVG}(f) :\equiv \forall x \in X; y \in K; \varepsilon \in \mathbb{Q}^*_+ \exists n \in \mathbb{N} \forall m \geq n\, (f(x,y,m) < \varepsilon).$$

Since the sequence is non-increasing we can omit the innermost universal quantifier and get

$$\mathsf{CVG}(f) \leftrightarrow \forall x \in X; y \in K; \varepsilon \in \mathbb{Q}^*_+ \exists n \in \mathbb{N}\, (f(x,y,n) < \varepsilon),$$

which has the form (5.6). *Monotone functional interpretation* creates a modulus $\delta : \mathbb{N}^\mathbb{N} \times \mathbb{Q}^*_+ \to \mathbb{N}$ satisfying (inserting the omitted universal quantifier back)

$$\forall x \in X; y \in K; \epsilon \in \mathbb{Q}^*_+ \forall m \geq \delta(x, \varepsilon)\, (f(x,y,m) < \varepsilon),$$

i.e. monotone functional interpretation transforms pointwise convergence into uniform convergence. The monotone functional interpretation in this case can be viewed as a form of Dini's theorem: Any non-increasing sequence $(f_n)_{n\in\mathbb{N}}$ of functions in $C[0,1]$ converging pointwise to zero converges uniformly to zero.

For a given function $f : K \to K$ and a starting point $x \in K$, let x_n denote the n-th iteration of f on x, i.e. $x_n :\equiv f^n(x)$. The convergence of the sequence $(d(x_n, x_{n+1}))_{n\in\mathbb{N}}$ to zero is normally called the *asymptotic regularity* of the function f

$$\mathsf{ASY}(f) :\equiv \forall x \in K \forall \varepsilon \in \mathbb{Q}^*_+ \exists n \forall m \geq n (d(x_m, f(x_m)) < \varepsilon).$$

In many cases the sequence $(d(x_n, x_{n+1}))_{n\in\mathbb{N}}$ is non-increasing so that, by the discussion above, the m.f.i. of $\mathsf{ASY}(f)$ (also when applied to the negative translation of $\mathsf{ASY}(f)$) creates a functional $\kappa : \mathbb{Q}^*_+ \to \mathbb{N}$ satisfying

$$\forall x \in K \forall \varepsilon \in \mathbb{Q}^*_+ \forall m \geq \kappa(\varepsilon)(d(x_m, f(x_m)) < \varepsilon).$$

The monotonicity in these convergence statements is only used to be able to write the convergence in the logical form (5.6). This is crucial for applications in a context based on classical logic in which one applies m.f.i. to the negative translation of formulas. Without monotonicity the negative translation of

$$\exists n \in \mathbb{N} \forall m \geq n \, (f(x, y, m) < \varepsilon)$$

would yield

$$\neg\neg\exists n \in \mathbb{N} \forall m \geq n \, (f(x, y, m) < \varepsilon)$$

from which m.f.i. no longer extracts a modulus of convergence (we will come back to this in Section 5.7 below). In an intuitionistic context, however, one can use m.f.i. to extract moduli of convergence even without any monotonicity assumptions. This remains true in the presence of various highly ineffective principles (see [96]).

5.5 The Monotone Functional Interpretation of Implications

As we saw in the previous section, not only the concepts created via m.f.i. but also the interconnections between these concepts have been extensively exploited in mathematics. This can again be viewed as an instance of the general logical fact that the monotone functional interpretation of an implication $A \to B$ between two statements of the form (5.5) provides a procedure to transform a modulus for the interpretation of A into one for the interpretation of B. Furthermore, if the proof of $A \to B$ is formalized in a suitable formal setting in which monotone functional interpretation applies, we are actually able to extract such a procedure from the given proof. In the following, we shall illustrate this for the so-called Edelstein fixed point theorem, where the issues involved can be explained quite easily. In Sections 5.6 and 5.7, we survey results we obtained in more substantial examples which solved open problems in the literature.

5.5.1 Example 1: Edelstein Fixed Point Theorem

In this section we illustrate with a simple example how the concepts described above interrelate via monotone functional interpretation. In this simple example the functionals required by m.f.i. can be easily provided. In more involved proofs, however, such as the ones presented in Sections 5.6.1 and 5.7.1, one also uses the interpretation to help extract from the given proof the desired functionals.

One form of the well-known Edelstein fixed point theorem can be stated as follows.

Proposition 5.1 ([42]) *Let (K, d) be a compact metric space and $f : K \to K$ be contractive (in the sense of 5.4.3). From any starting point $x \in K$, the iteration $(f^n(x))_{n\in\mathbb{N}}$ (also denoted by $(x_n)_{n\in\mathbb{N}}$) converges to the unique fixed point of f.*

We split Edelstein's proof into three lemmas. First one shows that *contractivity* implies *asymptotic regularity* of the sequence $(x_n)_{n\in\mathbb{N}}$. Note that the sequence $(d(x_n, x_{n+1}))_{n\in\mathbb{N}}$ is non-increasing. The proof of the first lemma

$\mathsf{CTN}(f) \to \mathsf{ASY}(f)$ provides a functional translating moduli of contractivity into moduli of asymptotic regularity for the function f.

Lemma 5.1 *Let D_K denote an upper bound for the diameter of the compact space K. Moreover, define $\chi_1(\eta, \varepsilon) :\equiv \frac{D_K - \varepsilon}{\eta(\varepsilon)} + 1$. For any function $f : K \to K$ having moduli of contractivity η the function $\kappa(\varepsilon) :\equiv \chi_1(\eta, \varepsilon)$ is a modulus of asymptotic regularity for f, i.e.*

$$\forall x \in K \forall \varepsilon \in \mathbb{Q}^*_+ \forall n \geq \kappa(\varepsilon)(d(x_n, f(x_n)) < \varepsilon).$$

Proof. Let $x \in K$ be arbitrary. By the definition of diameter $d(x, f(x)) = d(x_0, x_1) \leq D_K$. If $d(x_0, x_1) \leq \varepsilon$ then we are done, since $d(x_1, x_2) < \varepsilon$. Otherwise, since f is contractive we have that $d(x_1, x_2) \leq d(x_0, x_1) - \eta(\varepsilon) \leq D_K - \eta(\varepsilon)$. In general, either $d(x_m, x_{m+1}) \leq \varepsilon$ for some $m \leq n$ or $d(x_n, x_{n+1}) \leq D_K - n \cdot \eta(\varepsilon)$. Let $n \geq \frac{D_K - \varepsilon}{\eta(\varepsilon)}$. In the first case, since the sequence $(d(x_n, x_{n+1}))_{n \in \mathbb{N}}$ is non-increasing we have that $d(x_n, x_{n+1}) \leq \varepsilon$. In the second case we have $d(x_n, x_{n+1}) \leq D_K - n \cdot \eta(\varepsilon) \leq \varepsilon$. So for $n \geq \kappa(\varepsilon)$ we have $d(x_n, x_{n+1}) < d(x_{n-1}, x_n) \leq \varepsilon$. $\Box$

Remark 5.2 Note that instead of η we could have used Rakotch's notion of modulus of contractivity α. The functional $\chi_1(\alpha, \varepsilon)$ could then be defined as $\frac{\log \varepsilon - \log D_K}{\log \alpha(\varepsilon)} + 1$ in the lemma above.

In the second part we prove that *contractivity* implies *uniqueness* of the fixed point,

$$\forall x, y \in K(d(x, f(x)) = d(y, f(y)) = 0 \to d(x, y) = 0).$$

Again, the m.f.i. of the statement $\mathsf{CTN}(f) \to \mathsf{UNI}(\lambda x.d(x, f(x)))$ asks for a functional translating moduli of contractivity into moduli of uniqueness. The following lemma can be easily verified.

Lemma 5.2 *Define $\chi_2(\eta, \varepsilon) :\equiv \frac{\eta(\varepsilon)}{2}$. For any function $f : K \to K$ having moduli of contractivity η the function $\Phi(\varepsilon) :\equiv \chi_2(\eta, \varepsilon)$ is a modulus of uniqueness for the fixed point of f, i.e.*

$$\forall x, y \in K \forall \varepsilon \in \mathbb{Q}^*_+ (d(x, f(x)) < \Phi(\varepsilon) \wedge d(y, f(y)) < \Phi(\varepsilon) \to d(x, y) \leq \varepsilon).$$

Finally, the last lemma

$$\mathsf{ASY}(f) \wedge \mathsf{UNI}(\lambda x.d(x, f(x))) \to \forall x \in K((x_n)_{n \in \mathbb{N}} \text{ converges})$$

shows that *asymptotic regularity* plus *uniqueness* implies *convergence.* The statement of convergence in the conclusion has more complex logical form than (5.5). Similarly as explained in Section 5.4.1, however, one can still give a procedure for producing uniformly out of moduli of asymptotic regularity and uniqueness a modulus of convergence.

Lemma 5.3 *Define $\chi_3(\kappa, \Phi, \varepsilon) :\equiv \kappa(\Phi(\varepsilon))$. For any function $f : K \to K$ having fixed point c, modulus of asymptotic regularity κ and modulus of uniqueness of fixed point Φ, the function $\delta(\varepsilon) :\equiv \chi_3(\kappa, \Phi, \varepsilon)$ is a modulus of convergence for the fixed point of f, i.e. $\forall x \in K \forall \varepsilon \in \mathbb{Q}^*_+ \forall n \geq \delta(\varepsilon)(d(x_n, c) \leq \varepsilon)$.*

When we combine all the three lemmas we obtain the effective version of Edelstein fixed point theorem.

Proposition 5.2 *Let D_K denote the diameter of the compact space K. For any function $f : K \to K$ having modulus of contractivity η, and any starting point $x \in K$, the sequence $(x_n)_{n\in\mathbb{N}}$ converges to the fixed point c of f with rate of convergence* [13]

$$\delta(\varepsilon) :\equiv \chi_3(\lambda\varepsilon.\chi_1(\eta,\varepsilon), \lambda\varepsilon.\chi_2(\eta,\varepsilon), \varepsilon) = \frac{D_K - \eta(\varepsilon)}{\eta(\frac{\eta(\varepsilon)}{2})} + 1,$$

i.e.

$$\forall x \in K \forall \varepsilon \in \mathbb{Q}^*_+ \forall n \geq \delta(\varepsilon)\, (d(x_n, c) \leq \varepsilon).$$

Another quantitative version is given in Rakotch [141]. For a discussion of Edelstein's fixed point theorem in the context of Bishop's constructive mathematics see [28]. A recent domain theoretic approach to Edelstein's theorem can be found in [128].

5.6 Proofs Based on Heine-Borel Compactness

We have presented how the computational content of sentences of the form (5.3), (5.4) and (5.5) (in mathematical terms (5.1), (5.2) and (5.6) respectively) should be understood. Moreover, we showed how to deal with implications between statements of this from. This provides a procedure for analysing in a very simple way proofs which only involve formulas of this kind. For the rest of the paper we shall focus on more complex principles which do not fall into the general form (5.5), and how to analyze proofs involving such principles.

In this section we focus on principles related to Heine-Borel compactness such as

- **The attainment of the infimum**: Every continuous function $f : [a, b] \to \mathbb{R}$ attains its infimum.
- **Brouwer's fixed point theorem for continuous functions** $f : [0, 1]^n \to [0, 1]^n$.
- **Cauchy-Peano existence theorem**.

Each of these principles are, even when the function f is given together with the modulus of uniform continuity, equivalent to WKL (see [153]) and rely on the existence of non-computable real numbers. We analyze in details below the attainment of the infimum (for the interval $[0, 1]$) which can be written more formally as

$$\mathsf{INF} :\equiv \forall f \in C[0, 1] \exists x \in [0, 1](f(x) \stackrel{\mathbb{R}}{=} \inf_{y\in[0,1]} f(y)),$$

[13] Note that δ depends only on ε, D_K and η, but not x or f.

which, as shown in [90], has the logical form Δ. If the principle INF has been used in the proof of a theorem of the form (5.5) at some point in the proof a modus ponens over an implication

$$\forall u^1 \exists v^1 \leq tu \forall w^{\mathbb{N}} A_0(u,v,w) \to \forall x \forall y \leq sx \exists z B_0(x,y,z) \tag{5.7}$$

will take place. Negative translation of (5.7) gives

$$\forall u^1 \neg\neg \exists v^1 \leq tu \forall w^{\mathbb{N}} A_0(u,v,w) \to \forall x \forall y \leq sx \neg\neg \exists z B_0(x,y,z) \tag{5.8}$$

and hence a-fortiori

$$\forall u^1 \exists v^1 \leq tu \forall w^{\mathbb{N}} A_0(u,v,w) \to \forall x \forall y \leq sx \neg\neg \exists z B_0(x,y,z). \tag{5.9}$$

The m.f.i. of the premise of (5.9) asks for a Φ^* satisfying

$$\exists \Phi \leq t(\Phi \ \leq^{\mathsf{m}} \ \Phi^* \wedge \forall u^1 \forall w^{\mathbb{N}} A_0(u, \Phi u, w)),$$

which can be clearly taken to be $\Phi^* :\equiv t^*$, for some t^* majorizing t. The (partial) monotone functional interpretation of the implication (5.9) is realized by a functional χ^* satisfying

$$\exists \chi \ \leq^{\mathsf{m}} \ \chi^* \forall \Phi \leq t(\forall u^1 \forall w^{\mathbb{N}} A_0(u, \Phi u, w) \to \forall x \forall y \leq sx B_0(x, y, \chi(\Phi, x, y))).$$

Note that $\chi^*(t^*, x^+, s^*(x^+))$ majorizes $\chi(\Phi, x, y)$. Therefore, given the truth of the premise of (5.7) (and therefore its Skolemized version '$\exists \Phi \leq t \forall u, w A_0(u, \Phi u, w)$'), the functional $\Psi(x) :\equiv \chi^*(t^*, x^+, s^*(x^+))$ satisfies the m.f.i. of the conclusion, i.e.

$$\forall x \forall y \leq sx \exists w \leq \Psi(x)\, B_0(x, y, w).$$

The treatment of proof based on lemmas Δ presented here is due to [89], where more general forms of lemmas Δ are considered as well.

In the following section we report on a case study where a classical proof involving the principle INF has been analyzed and new results have been obtained.

5.6.1 Example 2: Jackson's Theorem

In [107] the authors have carried out the analysis of Cheney's proof [33] of the following well-known theorem in L_1-approximation theory ('approximation in the mean').

Theorem 5.3 (Jackson's theorem [70]) *Let P_n denote the space of algebraic polynomials of degree bounded by n. For any number n and continuous function $f \in C[0,1]$ there exists a unique element of P_n which best approximates f w.r.t the L_1-norm.*

This investigation yielded the first effective in all parameters modulus of uniqueness for L_1-approximation by polynomials of bounded degree. As it is clear from our Example 1, the difficulty in the analysis usually comes from the use of logically more complex principles.

Let us first outline how to bring Jackson's theorem into the form (5.2). Recall that the L_1-norm of a function $f \in C[0,1]$ is defined as

$\|f\|_1 :\equiv \int_0^1 |f(x)|\,dx,$

and $p \in P_n$ is a best L_1-approximation of f from P_n if

$\|f - p\|_1 = dist_1(f, P_n) \quad (:\equiv \inf_{p \in P_n} \|f - p\|_1).$

One easily observes that $dist_1(f, P_n) = dist_1(f, \tilde{K}_{f,n})$, where $\tilde{K}_{f,n}$ denotes the compact space $\{p \in P_n : \|p\|_1 \le 2\,\|f\|_1\}$. The existence of a best approximation of f in P_n w.r.t. the L_1-norm follows from the fact that the continuous function $G(f,p) :\equiv \|f - p\|_1$ attains its infimum in $\tilde{K}_{f,n}$. The highly non-trivial part of Theorem 5.3 is the uniqueness of the best L_1-approximation.

Define $F(f,p) :\equiv \|f - p\|_1 - dist_1(f, P_n)$. Uniqueness can be expressed as

$\forall n; f \in C[0,1]; p_1, p_2 \in P_n\, (\bigwedge_{i=1}^2 F(f, p_i) \overset{\mathbb{R}}{=} 0 \to p_1 = p_2).$

Moreover, the space P_n can be replaced by the space $\tilde{K}_{f,n}$ since any best L_1-approximation of f from P_n must belong to $\tilde{K}_{f,n}$, or the zero polynomial, which lives in $\tilde{K}_{f,n}$, would be a better approximation of f. Therefore, Theorem 5.3 can be stated as

$\forall n; f \in C[0,1]; p_1, p_2 \in K_{f,n}\, (\bigwedge_{i=1}^2 F(f, p_i) \overset{\mathbb{R}}{=} 0 \to \|p_1 - p_2\|_1 \overset{\mathbb{R}}{=} 0),$

where for technical reasons we use the larger space

$K_{f,n} :\equiv \{p \in P_n : \|p\|_1 \le \frac{5}{2}\,\|f\|_1\}.$

Note that the space $C[0,1]$ equipped with the L_1-norm is not complete, and therefore it is not a Polish space. To bring Jackson's theorem into the form (5.2) we use the Polish space $(C[0,1], \|\cdot\|_\infty)$. Since the functions f in the space $(C[0,1], \|\cdot\|_\infty)$, according to the representation, are endowed with a modulus ω_f, the functions $\|\cdot\|_1 : C[0,1] \to \mathbb{R}$ and F are PA^ω-definable. Therefore, Jackson's theorem falls into the general form described in Section 5.4.1. As we have seen, the computational content of a uniqueness statement such as the one above is given via a modulus of uniqueness Φ satisfying, for all $f \in C[0,1]$ and $n \in \mathbb{N}$,

$\forall p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}^*_+\, (\bigwedge_{i=1}^2 F(f, p_i) \le \Phi(f, n, \varepsilon) \to \|p_1 - p_2\|_1 \le \varepsilon),$

independent of the polynomials p_1 and p_2, which range over the compact space $K_{f,n}$. By the choice of the space $K_{f,n}$ the modulus Φ can be easily extended to a modulus for the whole space P_n.

Recall that Φ depends on f via its representation as an element of the Polish space $(C[0,1], \|\cdot\|_\infty)$. That is to say, Φ will (a priori) depend on the values of the function f as well as on a modulus of continuity for f. This apparent restriction of Theorem 5.1 is indeed an indication of which inputs are the right ones for the modulus of uniqueness. See, for instance, [22] and [119] where the modulus of uniform continuity is always used as an input for moduli of uniqueness.

Theorem 5.1 guarantees that from any proof of Jackson's theorem formalizable in a system like $\mathsf{E\text{-}PA}^\omega + \mathsf{QF\text{-}AC}^{1,\mathbb{N}} + \mathsf{INF}$ we are able to extract a modulus of uniqueness Φ. One such proof, as shown in [88], was presented by Cheney [33] in 1965. Therefore, by Theorem 5.1 we obtain the following *a priori* information.

Proposition 5.3 *Let $K_{\omega,M}$ be the compact subspace of $C[0,1]$ consisting of functions with modulus of continuity ω and uniform norm bounded by M. There exists a modulus of uniqueness Φ (given by a closed term of $\mathsf{E\text{-}PA}^\omega$, i.e. of Gödel's T) depending only on ω, M, n and ε for the L_1-approximation of functions $f \in K_{\omega,M}$ from the space P_n.*

In [107][14] the authors have carried out the extraction of such a modulus of uniqueness out of Cheney's proof of Jackson's theorem, providing explicitly the dependencies of Φ (*a posteriori* information).

Theorem 5.4 ([107]) *Let*

$$\Phi(\omega, n, \varepsilon) :\equiv \min\{\tfrac{c_n\varepsilon}{8(n+1)^2}, \tfrac{c_n\varepsilon}{2}\omega_n(\tfrac{c_n\varepsilon}{2})\},$$

where

$$c_n :\equiv \tfrac{\lfloor n/2\rfloor!\lceil n/2\rceil!}{2^{4n+3}(n+1)^{3n+1}} \quad \text{and} \quad \omega_n(\varepsilon) :\equiv \min\{\omega(\tfrac{\varepsilon}{4}), \tfrac{\varepsilon}{40(n+1)^4\lceil\frac{1}{\omega(1)}\rceil}\}.$$

The functional $\lambda\varepsilon.\Phi(\omega, n, \varepsilon)$ is a uniform modulus of uniqueness for the best L_1-approximation of any function f in $C[0,1]$ having modulus of uniform continuity ω from P_n, i.e. for all n and $f \in C[0,1]$

$$\forall p_1, p_2 \in P_n; \varepsilon \in \mathbb{Q}^*_+ \, (\textstyle\bigwedge_{i=1}^2 F(f, p_i) \le \Phi(\omega, n, \varepsilon) \to \|p_1 - p_2\|_1 \le \varepsilon),$$

where ω is a modulus of uniform continuity of the function f.

Note that, using Markov's inequality, from any upper bound on $\|p\|_1$ one can easily derive an upper bound on the absolute value of the actual coefficients of the polynomial p.

Although uniqueness of L_1-approximation was known since 1921, only in 1975 Björnestål [22] proved the existence of a modulus of uniqueness Φ having the form $c_{f,n}\,\varepsilon\,\omega_n(c_{f,n}\,\varepsilon)$, for some constant $c_{f,n}$ depending on f and n. Björnestål's proof is ineffective and does not supply $c_{f,n}$. In 1978, Kroó [118] improved Björnestål's results by showing that a constant $c_{\omega,n}$, depending only on the modulus of uniform continuity of f and n exists, but his proof is also ineffective and no constant is presented. Moreover, Kroó proves that the ε-dependency established by Björnestal is optimal.

By obtaining the modulus of uniqueness explicitly, as in Theorem 5.4, we get as a byproduct all those qualitative results. It should be observed that the form of the modulus Φ depends on the proof from which it was extracted. Different proofs could have given different moduli. The fact that Φ has optimal ε-dependency suggests that Cheney's proof is in some sense optimal.

The modulus of uniqueness we have obtained can be used in various ways. For instance, as already mentioned, $\Phi/2$ is a modulus for the pointwise continuity of the projection operator.

[14] In Section 3.1 of [107] [equivalently, in Section 6.3.1 of this dissertation] $\mathcal{A}^\omega$ should be replaced by $\mathcal{A}^\omega + \mathsf{QF\text{-}AC}$ (note that $\mathsf{QF\text{-}AC}$ is admissible in the metatheorem) since the σ_i-property implicitly uses the intermediate value theorem. Alternatively, one can avoid the use of the intermediate value theorem by replacing the second line in the σ_i-description by "-1 otherwise".

Theorem 5.5 ([107]) *Let $\mathcal{P}(f,n)$ denote the operator which produces the best L_1-approximation of f from P_n. Then, for all n*

$$\forall f, g \in C[0,1]; \varepsilon \in \mathbb{Q}^*_+ \left(\|f-g\|_1 \leq \tfrac{\Phi(\omega_f,\varepsilon)}{2} \to \|\mathcal{P}(f,n) - \mathcal{P}(g,n)\|_1 \leq \varepsilon\right),$$

where ω_f denotes a modulus of uniform continuity of f.

The modulus of uniqueness Φ has also been used in [134] by the second author to give the first complexity upper bound on the sequence $(p_n)_{n\in\mathbb{N}}$ of best L_1-approximations of a polynomial-time computable function.

Theorem 5.6 ([134]) *Let $f \in C[0,1]$ be polynomial-time computable, then the sequence $(p_n)_{n\in\mathbb{N}}$ is strongly* **NP** *computable in* $\mathbf{NP}[B_f]$*, where B_f is an oracle solving a left cut for integration.*

As a final remark, note that both the existence and the uniqueness proof make use of the principle INF. While the existence statement has the same logical form of INF, and therefore the use of the principle cannot be eliminated from that proof (although the constructive existence follows via the effective modulus of uniqueness), the uniqueness theorem has the simpler logical form $\forall\exists$, which indicates that INF is not really used in its full strength.

For another case study in the context of Chebycheff approximation see [90] and [91].

5.7 Proofs Based on Fixed Uses of Sequential Compactness

By proofs based on sequential compactness we mean proofs which use principles like

- PCM (**Principle of monotone convergence**) :≡ If a sequence of reals $(a_n)_{n\in\mathbb{N}}$ is non-increasing and bounded from below (say by 0) then it is convergent.
- BW (**Bolzano-Weierstraß principle**) :≡ Any sequence of reals $(a_n)_{n\in\mathbb{N}}$ belonging to the cube e.g. $[0,1]^d$ has a convergent subsequence.
- A-A (**The Arzelà-Ascoli lemma**) :≡ Any sequence $(f_n)_{n\in N} \in C[0,1]$ of equicontinuous and uniformly bounded functions has a convergent subsequence (w.r.t. $\|\cdot\|_\infty$).
- Limsup (**The existence of the limit superior**) :≡ For any sequence $(a_n)_{n\in\mathbb{N}} \subset [0,1]$ there exists a point $x \in [0,1]$ such that $x = \limsup\limits_{n\to\infty} a_n$.

By a **fixed** use of sequential compactness we mean an application of such a principle to a particular sequence of reals/functions, in general built out of the parameters of the problem. We shall denote such a fixed application of e.g. PCM to a sequence $(a_n)_{n\in\mathbb{N}}$ as PCM(a_n).

Although the principles mentioned above are equivalent to full arithmetical comprehension even over weak base systems (see [99] and [153])[15], these principles are often only used for fixed sequences in the given proof. In this case, the contribution to the growth of extractable bounds is much lower. All this has been spelled out in great detail in [94] and [99] for all of the principles mentioned above. We only discuss here briefly $\mathsf{PCM}(a_n)$ as we will need this in the application discussed in Section 5.7.1. Let $(a_n)_{n\in\mathbb{N}}$ for the rest of this section denote a sequence in $[0,N]$ for some $N\in\mathbb{N}$. $\mathsf{PCM}(a_n)$ can be written as

$$\mathsf{PCM}(a_n) :\equiv [Mon(a_n) \to \exists a \in \mathbb{R}_+(\lim_{n\to\infty} a_n = a)],$$

where $Mon(a_n) :\equiv \forall k(0 \le a_{k+1} \le a_k \le N)$. Since real numbers are represented as Cauchy sequences of rational numbers with fixed rate of convergence, $\mathsf{PCM}(a_n)$ is in fact equivalent (using $\mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$) to

$$Mon(a_n) \to \exists f \forall k \forall m \big(m \ge f(k) \to |a_{f(k)} - a_m| \le \tfrac{1}{k+1}\big).$$

It is the existence of the Cauchy modulus f which implies Π^0_1-comprehension which – by iteration – gives Π^0_∞-comprehension.

However, as mentioned already, the contribution is much weaker (under suitable conditions) when $\mathsf{PCM}(a_n)$ is applied only to a given fixed sequence (a_n) (definable in the parameters of the problem at hand) in a proof of a statement of the form (5.5) since then the iterated use of the principle is blocked. In fact over sufficiently weak fragments of classical arithmetic in all finite types (to which, though, the axioms Δ of the kind discussed above may be added) such a use of $\mathsf{PCM}(a_n)$ can be reduced to the use of the arithmetical version[16]

$$\mathsf{PCM}_{\mathsf{ar}}(a_n) :\equiv \Big[Mon(a_n) \to \forall k \exists n \forall m \big(m \ge n \to |a_n - a_m| \le \tfrac{1}{k+1}\big)\Big]$$

which in turn is equivalent to $\forall k \exists n \forall m (m \ge n \to |\tilde{a}_n - \tilde{a}_m| \le \frac{1}{k+1})$, where

$$\tilde{a}_n :\equiv \max(0, \min_{i\le n}(a_i, N)).$$

Hence, $\mathsf{PCM}_{\mathsf{ar}}(a_n)$ has the logical form $\forall k \exists n \forall m A_0(k,n,m)$, for an appropriate quantifier-free formula A_0. For simplicity we omit the parameter $(a_n)_{n\in\mathbb{N}}$ which, according to the representation of reals used, can be be encoded as a number theoretic function.

Let us now consider how monotone functional interpretation treats an implication with $\mathsf{PCM}_{\mathsf{ar}}$ as premise and a statement of the form (5.5) as conclusion:

$$\forall x^1 \forall y \le_1 sx\big(\mathsf{PCM}_{\mathsf{ar}}(t(x,y)) \to \exists z^{\mathbb{N}} B_0(x,y,z)\big), \tag{5.10}$$

[15] Nevertheless, those can also be treated by monotone functional interpretation using a weak form of monotone bar recursion (cf. Section 5.8).

[16] This reduction is very subtle and relies on a special technique of elimination of monotone Skolem functions taking into account a strong monotonicity property of the matrix of $\mathsf{PCM}_{\mathsf{ar}}$. We do not go here into this as in the application to be discussed below this passage is trivial.

where B_0 is quantifier-free and t is a term creating sequences of reals uniformly in x and y. The (partial) monotone functional interpretation of the negative translation

$$\forall x^1 \forall y \leq_1 sx \big(\forall k \neg\neg \exists n \forall m A_0(k,n,m) \to \neg\neg \exists z^{\mathbb{N}} B_0(x,y,z)\big) \tag{5.11}$$

of (5.10) is realized by a functional Ω^* satisfying

$$\begin{cases} \exists \Omega \ \leq^{\mathsf{m}} \ \Omega^* \forall \Psi, x \forall y \leq sx \\ \qquad \big(\forall k, g A_0(k, \Psi(k,g), g(\Psi(k,g))) \to B_0(x,y,\Omega(x,y,\Psi))\big). \end{cases}$$

Suppose now that we have a functional Φ^* satisfying the monotone functional interpretation of the negative translation of $\forall x^1 \forall y \leq sx \, \mathsf{PCM}_{\mathsf{ar}}(t(x,y))$, i.e.

$$\exists \Phi \ \leq^{\mathsf{m}} \ \Phi^* \forall x^1; y \leq sx, k, g \, A_0(k, \Phi(x,y,k,g), g(\Phi(x,y,k,g)))\big) \tag{5.12}$$

then $\chi(x,\Phi^*) :\equiv \Omega^*(x^+, s^* x^+, \Phi^*(x^+, s^* x^+)) \geq \Omega(x,y,\Phi(x,y))$ for all x^1 and $y \leq sx$, where $s \ \leq^{\mathsf{m}} \ s^*$ and $\Phi(x,y) :\equiv \lambda k, g.\Phi(x,y,k,g)$. Hence

$$\forall x^1 \forall y \leq_1 sx \exists z \leq \chi(x,\Phi^*) \, B_0(x,y,z).$$

So the contribution of the use of $\mathsf{PCM}_{\mathsf{ar}}(t(x,y))$ to the bound for the conclusion of (5.10) is given by a functional Φ^* satisfying (5.12). One easily verifies that we can take

$$\Phi^*(x,y,k,g) :\equiv \max_{i \leq (k+1)N} (g^i(0)), \tag{5.13}$$

i.e. Φ^* (in contrast to Φ!) basically is **independent** from the sequence $t(x,y)$ and only depends on an upper bound N on the first element of the sequence. This feature will play a crucial role in the applications to metric fixed point theory which we will discuss in the next example.

5.7.1 Example 3: Asymptotic Regularity of Iterations of Nonexpansive Mappings

One of the most active areas of nonlinear functional analysis is the fixed point theory of nonexpansive mappings (see e.g. [76]). In this section we report on the results of a recent case study of proof mining carried out by the first author (see [100, 101, 103] and – together with Laurenţiu Leuştean – [105]).

Definition 5.2 *Let $(X, \|\cdot\|)$ be a normed linear space and $C \subseteq X$ be a subset of X. A function $f : C \to C$ is called nonexpansive if*

$$\forall x, y \in C\big(\|f(x) - f(y)\| \leq \|x - y\|\big).$$

In view of Banach's result, the fixed point theory of contractions is rather simple. Even the case of contractive mappings enjoys – as we saw above – many of the features of contractions, e.g. the uniqueness of the fixed point. Things, however, change radically for nonexpansive functions. Fixed points, if existing at all, will not be unique and even if uniqueness holds the Banach iteration in general will not converge to the fixed point. Instead, other iterations play a crucial role here.

In the following, $(X, \|\cdot\|)$ will be an arbitrary normed linear space, $C \subseteq X$ a non-empty convex subset of X and $f : C \to C$ a nonexpansive mapping.

We consider the so-called Krasnoselski-Mann iteration starting from $x \in C$

$$x_0 :\equiv x, \quad x_{k+1} :\equiv (1-\lambda_k)x_k + \lambda_k f(x_k),$$

where $(\lambda_k)_{k\in\mathbb{N}}$ is a sequence of real numbers in $[0,1]$. For more information on the relevance of this kind of generalized Krasnoselski [110] iterations see e.g [25, 44, 127, 143].

Let $r_C(f) :\equiv \inf_{x\in C} \|x - f(x)\|$. For the rest of this section we assume, following [25] and [69], that $(\lambda_k)_{k\in\mathbb{N}}$ is divergent in sum, which can be expressed (since $\lambda_k \geq 0$) as[17]

$$\forall n, i \in \mathbb{N} \exists k \in \mathbb{N} \left(\sum_{j=i}^{i+k} \lambda_j \geq n \right), \tag{5.14}$$

and that

$$\forall k \in \mathbb{N}(\lambda_k \leq 1 - \frac{1}{K}) \text{ for some } K \in \mathbb{N}. \tag{5.15}$$

Theorem 5.7 ([25]) *Suppose that $(\lambda_k)_{k\in\mathbb{N}}$ satisfies the conditions* (5.14) *and* (5.15). *Then the Krasnoselski-Mann iteration $(x_n)_{n\in\mathbb{N}}$ starting from any point $x \in C$ satisfies*

$$\|x_n - f(x_n)\| \overset{n\to\infty}{\to} r_C(f).$$

Under quite general circumstances one can prove that $r_C(f) = 0$.

Theorem 5.8 ([25], [69]) *Under the assumptions of the previous theorem and the additional assumption that there exists a $x^* \in C$ such that $(x_n^*)_{n\in\mathbb{N}}$ is a bounded sequence the following holds*

$$\forall x \in C(\|x_n - f(x_n)\| \overset{n\to\infty}{\to} 0) \quad \textit{(called ‘asymptotic regularity’).}$$

Remark 5.3 The special case of Theorem 5.8 in which only the asymptotic regularity of the sequence (x_n^*) is conclude is due to Ishikawa [69].[18] The striking aspect of Ishikawa's theorem is that it does not rely on the assumption of X being uniformly convex as all results of that kind prior to [69] did. For uniformly convex spaces X, bounded C and constant $\lambda_k = \lambda$ the result was proved in [30], and for general λ_k – even more general than in Ishikawa's theorem – it follows from [62] for such spaces. If C is, moreover, compact and $\lambda = \frac{1}{2}$, asymptotic regularity was (for uniformly convex X) already proved in [110].

In oder to see that our general meta-theorem on proof mining can be applied to Theorem 5.7 we first have to find a proper formalization of the conclusion

[17] This form will be particularly suitable below.

[18] For constant $\lambda_k = \lambda$ the result was independently obtained in [43].

of the theorem. We first realize, that the ineffective existence of $r_C(f)$ is not really needed to formulate the conclusion which can be stated without $r_C(f)$ as

$$\forall \varepsilon > 0 \exists n \in \mathbb{N} \forall m \geq n \forall x^* \in C(\|x_m - f(x_m)\| < \|x^* - f(x^*)\| + \varepsilon). \tag{5.16}$$

An easy and well-known lemma shows that $(\|x_n - f(x_n)\|)_{n\in\mathbb{N}}$ is non-increasing so that the discussion from 5.4.6 applies. Therefore, the quantifier '$\forall m \geq n$' in (5.16) is in fact superfluous. Nevertheless, due to the alternation $\exists n \in \mathbb{N} \forall x^* \in C$, (5.16) still does not of the form $\forall\exists$ required (as a consequence of the use of classical logic) by our meta-theorems 5.1 and 5.2.[19] The following variant of (5.16), however, does have this form[20]

$$\forall \varepsilon > 0 \forall x^* \in C \exists n \in \mathbb{N}(\|x_n - f(x_n)\| < \|x^* - f(x^*)\| + \varepsilon). \tag{5.17}$$

Under the assumption of the existence of $r_C(f)$, formulations (5.16) and (5.17) are actually equivalent. In the following we shall study in more detail the form (5.17) of Theorem 5.7. Note that, in this case, a bound on n shall a priori depend on the additional input x^*.

Let us now consider the assumptions of Theorem 5.7 and assume for the moment that X is complete and separable and C a subset which can be explicitly represented in our underlying formal system. Observe that the assumptions of C being convex and f a nonexpansive function are purely universal[21]. Universal assumptions, however, do not change the logical form as required by our meta-theorem as they just add a couple of more existential quantifiers to the interpreted formula.

Monotone functional interpretation of the assumptions (5.14) and (5.15) on λ_k introduce new inputs, namely a bound $\alpha : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that

$$\forall n, i \in \mathbb{N}(n \leq \sum_{j=i}^{i+\alpha(i,n)} \lambda_j) \tag{5.18}$$

and a $K \in \mathbb{N}$ such that

$$\forall n \in \mathbb{N}(\lambda_n \leq 1 - \frac{1}{K}), \tag{5.19}$$

where both (5.18), (5.19) are purely universal. Given α and K as additional inputs, we can take the quantification over the sequences (λ_n) as quantification over the compact Hilbert cube $[0,1]^{\mathbb{N}}$ plus an explicit stipulation that (λ_n) satisfies (5.18) and (5.19). From this the meta-theorem provides the a priori information that the bound on the convergence in Theorem 5.7 we are about to extract might depend on α, K (and x^*) as new inputs which were not visible

[19] Indeed, an effective bound on '$\exists n$' in (5.16) would imply the computability of $r_C(f)$ (in f, x, λ_k and $\|\cdot\|$) which is unlikely to be true in the general case.

[20] One can actually consider an intermediate version where x^* is allowed to be a sequence depending on n. Bounds for this stronger form are obtained in [103].

[21] We do not even need to express explicitly that f (represented as a function on representatives of elements in $x \in C$) is extensional (i.e. respects the equivalence relation $x =_X y$ expressing that x, y represent the same X-element) since the extensionality follows from the continuity of f which in turn follows from the fact that f is nonexpansive.

in the original formulation of the theorem, but that it will be independent from any particular (λ_k) itself (cf. Section 5.3.1).

Let us now consider the lemmas used in the proof of Theorem 5.7. By far the largest part of the proof concerns a highly non-trivial inequality due to [59] (whose proof is based on [73] and also [69]): for all $n, i \in \mathbb{N}$

$$\begin{cases} \big(1 + \sum\limits_{k=i}^{n+i-1} \lambda_k\big)\|x_i - f(x_i)\| \leq \\ \|x_i - f(x_{i+n})\| + \left[\prod\limits_{k=i}^{n+i-1} \frac{1}{1-\lambda_k}\right] \big(\|x_i - f(x_i)\| - \|x_{i+n} - f(x_{i+n})\|\big). \end{cases}$$

Since this inequality is purely universal (as are two other simpler inequalities used) we can simply take it as yet another implicative assumption in the proof analysis, i.e. we do not have to consider its proof at all.

From the point of view of proof mining, the only problematic tool used in the proof is the ineffective fact that

the non-increasing sequence $(\|x_n - f(x_n)\|)$ *of reals* ≥ 0 *has a limit*,

which is just $\mathsf{PCM}(\|x_n - f(x_n)\|)$, i.e. a *fixed* instance of PCM. As we have discussed above, the use of PCM in this case can be reduced, in the poof of Theorem 5.7, to its arithmetical version $\mathsf{PCM_{ar}}(\|x_n - f(x_n)\|)$ which states that $(\|x_n - f(x_n)\|)_{n\in\mathbb{N}}$ is a Cauchy sequence. This reduction is sound provided the proof can be carried out relative to a rather weak framework like the fragment $\mathsf{E\text{-}G_3A}^\omega$ of $\mathsf{E\text{-}PA}^\omega$ which, in particular must not contain the iteration functional $\Phi_{it}(x, y, f) :\equiv f^x(y)$. In fact this is the case, though it seems at first sight impossible as the very sequence (x_n) is defined by iteration. We can, however, take

$$\forall n \big(x_{n+1} = (1 - \lambda_n)x_n + \lambda_n f(x_n)\big)$$

just as one more purely universal implicative assumption and do not need for the proof analysis to prove that such a sequence can be formed. So in total, taking A to be the conjunction of all the universal assumptions considered we get

$$\mathsf{PCM_{ar}} \to (A \to (5.17)),$$

where (5.17) (and hence $A \to (5.17)$) is a $\forall\exists$-formula. Therefore, the discussion of the modus ponens problem above applies and we can extract a bound for (5.17) in f, x, x^*, α, K which, as a consequence of the use of $\mathsf{PCM_{ar}}$, will involve a use of the iteration functional Φ_{it}. Indeed, in [101], the first author obtained the following quantitative version of Theorem 5.7 (as a matter of fact, we not even need to assume that $(X, \|\cdot\|)$ is complete or separable).

Theorem 5.9 ([101]) *Let* $(X, \|\cdot\|)$ *be a normed linear space,* $C \subseteq X$ *a non-empty convex subset and* $f : C \to C$ *a nonexpansive mapping. Let* $(\lambda_k)_{k\in\mathbb{N}}$ *be a sequence in* $[0, 1]$ *which is divergent in sum and satisfies*

$$\forall k \in \mathbb{N}\big(\lambda_k \leq 1 - \tfrac{1}{K}\big)$$

for some $K \in \mathbb{N}$. Let $\alpha : \mathbb{N} \to \mathbb{N}$ be such that

$$\forall i, n \in \mathbb{N}\big(\alpha(i,n) \le \alpha(i+1,n) \wedge n \le \sum_{s=i}^{i+\alpha(i,n)-1} \lambda_s\big).$$

Let $(x_n)_{n\in\mathbb{N}}$ be the Krasnoselski-Mann iteration

$$x_0 :\equiv x, \qquad x_{n+1} :\equiv (1-\lambda_n)x_n + \lambda_n f(x_n),$$

starting from $x \in C$. Then the following holds

$$\forall x, x^* \in C \forall \varepsilon > 0 \forall n \ge h(\varepsilon, x, x^*, f, K, \alpha)(\|x_n - f(x_n)\| < \|x^* - f(x^*)\| + \varepsilon),$$

where[22]

$$\begin{array}{l} h(\varepsilon, x, x^*, f, K, \alpha) :\equiv \widehat{\alpha}(\lceil 2\|x - f(x)\| \cdot \exp(K(M+1))\rceil \dot{-} 1, M), \\ \textit{with } M :\equiv \left\lceil \frac{1+2\|x-x^*\|}{\varepsilon} \right\rceil \textit{ and} \\ \widehat{\alpha}(0, M) :\equiv \tilde{\alpha}(0, M),\ \widehat{\alpha}(m+1, M) :\equiv \tilde{\alpha}(\widehat{\alpha}(m, M), M) \textit{ with} \\ \tilde{\alpha}(m, M) :\equiv m + \alpha(m, M)\ (m \in \mathbb{N}). \end{array}$$

Instead of M we may use any upper bound $\mathbb{N} \ni \tilde{M} \ge \frac{1+2\|x-x^*\|}{\varepsilon}$. Likewise, $\|x - f(x)\|$ may be replaced by any upper bound.

Remark 5.4 An α satisfying the conditions of the theorem can be computed from any $\beta : \mathbb{N} \to \mathbb{N}$ such that $n \le \sum_{s=0}^{\beta(n)} \lambda_s$ (for all n) by $\alpha(i,n) :\equiv \max_{j\le i}(\beta'(j,n))$, where $\beta'(i,n) :\equiv \beta(n+i) - i + 1$.

Perhaps the most useful aspect of Theorem 5.9 is that it displays the very limited dependency of the rate of uniform convergence from the input data x, f, x^*, λ_k and X, C. In fact, if C is bounded with $d \ge diam(C)$, then the dependence from x, x^* and f can be removed altogether as $\|x - f(x)\|$ and $\|x - x^*\|$ both can be replaced by d. Moreover, it follows that the bound only depends on d but not on C itself (see [101]). In fixed point theory non-trivial functional analytic embedding techniques have been used for some 20 years to obtain (partial) such uniformity results for bounded C. In this way the independence from x is proved in [43] for constant $\lambda_k :\equiv \lambda$. In [59] this is extended to uniformity also w.r.t. f (for general λ_k) but not w.r.t. C (in the sense above). In [60] it is in fact conjectured that the uniformity in C might only hold in the much simpler case of uniformly convex case (cf. [75]). For constant λ, full uniformity was finally established in [6]. Our result gives full uniformity for general λ_k and even displays that the rate of convergence is to a large extend independent from λ_k, depending only on α and K.

The next theorem, which is based on Theorem 5.9, allows to push the uniformity even further to the case where C is no longer assumed to be bounded but only to contain some point x^* whose iteration sequence (x^*_n) is bounded, i.e. the context of Theorem 5.8.

[22] $n \dot{-} 1 = \max(0, n-1)$.

Theorem 5.10 ([103]) *Under the assumptions of Theorem 5.9 the following holds. Let $d > 0, x, x^* \in C$ be such that $\forall n(\|x_n^*\| \leq d)$ and $\|x - x^*\| \leq d$. Then*

$$\forall \varepsilon > 0 \forall n \geq h(\varepsilon, d, K, \alpha)(\|x_n - f(x_n)\| \leq \varepsilon),$$

where

$$h(\varepsilon, d, K, \alpha) :\equiv \widehat{\alpha}(\lceil 12 \cdot \exp(K(M+1)) \rceil - 1, M),$$
$$\text{with } M :\equiv \left\lceil \frac{7d}{\varepsilon} \right\rceil \text{ and } \widehat{\alpha} \text{ as in Theorem 5.9.}$$

Note that the bound only depends on $\frac{d}{\varepsilon}, K$ and α!

Proof. The theorem follows from [103] (Thm. 2.5 plus Remarks 2.2 and 2.6). □

Whereas this result easily follows from the logical analysis in [101] (which resulted in Theorem 5.9) of the proof of Theorem 5.7 and does not use any functional analytic tools at all, it seems that the embedding techniques, as used e.g. in [59] and most recently in a new form in [74], are not applicable as they heavily rely on the boundedness of C. So the logical approach here not only gives new quantitative bounds but even new qualitative results which are superior to what has been achieved by more traditional functional analytic means. For more results in this direction and proofs of the results discussed see [100], [101] and [103].

Another benefit of the logical approach is that it easily generalizes to other settings for which the basic inequalities used in the proof of the Borwein-Reich-Shafrir result can be verified. Since no functional analytic embeddings are used there is no need to exploit any new analytic tool to obtain uniformity results. Very recently ([105]) the first author (together with Laurenţiu Leuştean) showed in this way that the results (as well as the basic structure of their proofs) presented above extend to hyperbolic spaces in the sense of Reich and Shafrir [142] (including the Hilbert ball with the hyperbolic metric) and – to a large extent – also to the still more general class of spaces of hyperbolic type [59] (which were first introduced in [155] under the name of 'convex metric spaces') and directionally nonexpansive mappings in the sense of [74]. In particular, strengthened versions of the main results of [74] follow as special cases.

The results just described ask for a general logical explanation for the phenomenon that here the proof analysis was possible without any assumptions on X (like being separable and representable in say $\mathsf{E\text{-}PA}^\omega$) and yielded uniformity even w.r.t. to norm bounded (i.e. not necessarily compact) convex sets. Obviously, this is related to the fact that the normed space X and its convex subset were completely general. Using a technique of "adding" structures like general normed linear spaces to finite type systems as a new ground type plus the vector space operations and the norm function as primitive constants, the first author recently obtained, generalising the technique of monotone functional interpretation, logical meta-theorems which guarantee the existence of such uniform bounds under quite general logical conditions [104]. The setting of hyperbolic spaces is particularly suitable for these meta-theorems which allow to obtain new qualitative uniformity results even without any actual proof analysis (which, however, would be necessary for the extraction of explicit bounds).

5.8 Proofs Based on Applications of Full Sequential Compactness

In the previous section we have shown how to treat proofs of theorems having the form (5.5) which make use of e.g. PCM applied to a fixed sequence $(a_n)_{n\in\mathbb{N}}$. In this section we address the problem of analyzing such proofs in which PCM (or any of the other principles) is used to obtain the convergence of an arbitrary sequence, which is not actually built in the proof.

As mentioned above, such use of PCM is in fact equivalent to arithmetical comprehension. In this case we can not expect to give a constructive treatment of the proof without making use of bar recursion (cf. [154]).

For the sake of simplicity, all the sequences $(a_n)_{n\in\mathbb{N}}$ are assumed to be contained the interval $[0, N]$. We want to produce a functional realizing the m.f.i. of the negative translation of (cf. Section 5.7)

$$\forall(a_n)_{n\in\mathbb{N}}\exists f\forall k, m(m \geq fk \rightarrow |\tilde{a}_{fk} - \tilde{a}_m| \leq \tfrac{1}{k+1}).$$

We have seen in Section 5.7 that the m.f.i. of the arithmetical version of PCM,

$$\forall(a_n)_{n\in\mathbb{N}}\forall k\exists n\forall m(m \geq n \rightarrow |\tilde{a}_n - \tilde{a}_m| \leq \frac{1}{k+1}) \tag{5.20}$$

can be easily realized using the iteration functional. Formula (5.20) has the logical form $\forall(a_n)_{n\in\mathbb{N}}\forall k\exists n\forall m A_0(k,n,m)$, for some quantifier free A_0. Note that PCM is obtained by an application of Π^0_1-AC to this formula

$$\frac{\forall(a_n)_{n\in\mathbb{N}}\forall k\exists n\forall m A_0(k,n,m) \qquad \mathsf{AC}}{\forall(a_n)_{n\in\mathbb{N}}\exists f\forall k, m A_0(k, fk, m)}$$

To make constructive sense of PCM we first apply negative translation to the proof above to get a new proof (in the following we omit $\forall(a_n)_{n\in\mathbb{N}}$)

$$\frac{\forall k\neg\neg\exists n\forall m A_0(k,n,m) \quad \mathsf{AC}^N(\forall m A_0(k,n,m))}{\neg\neg\exists f\forall k, m A_0(k, fk, m)}$$

We finally apply functional interpretation to obtain

$$\frac{\exists\Phi^2\forall k^{\mathbb{N}}, g^1\, A_0(k, \Phi kg, g(\Phi kg)) \quad (\mathsf{AC}^N(\forall m A_0(k,n,m)))^D}{\forall\Psi_1,\Psi_2\exists f A_0(\Psi_1(f), f(\Psi_1(f)), \Psi_2(f))}$$

As done in Section 5.7 (cf. (5.13)), we can define via iteration a functional Φ^*

$$\Phi^*_N(k,g) :\equiv \max_{i\leq(k+1)N}(g^i(0)).$$

which majorizes a realizer of

$$\exists\Phi^2_{(a_n)}\forall k^{\mathbb{N}}, g^1 A_0(k, \Phi kg, g(\Phi kg)),$$

i.e. $\exists\Phi \leq^{\mathsf{m}} \Phi^*\forall k^{\mathbb{N}}, g^1 A(k, \Phi kg, g(\Phi kg))$. We now set out to obtain a realizer for the monotone functional interpretation of $\mathsf{AC}^N(\forall m A_0(k,n,m))$, i.e.

$$\forall k \neg\neg\exists n \forall m A_0(k,n,m) \to \neg\neg\exists f \forall k, m A_0(k, fk, m).$$

Monotone functional interpretation provides majorants for the realizers of the existential quantifiers of

$$\forall \Phi, \Psi_1, \Psi_2 \exists f, k, g[A_0(k, \Phi kg, g(\Phi kg)) \to A_0(\Psi_1 f, f(\Psi_1 f), \Psi_2 f)]. \tag{5.21}$$

By $\mathsf{BR}_{\mathbb{N},1}$ we mean the bar recursive functional (defined by Spector [154]) satisfying the following equation:

$$\mathsf{BR}_{\mathbb{N},1}(Y,G,H,s) \stackrel{1}{=} \begin{cases} G(s) & \text{if } Y(s * \lambda n.0) \leq n \\ H(s, \lambda y^{\mathbb{N}}.\mathsf{BR}_{\mathbb{N},1}(Y,G,H,s*y)) & \text{otherwise.} \end{cases}$$

Let p be a shorthand for Φ, Ψ_1, Ψ_2. Spector showed that by taking

$$Y :\equiv \lambda p\,.\,\Psi_1$$

$$G :\equiv \lambda p, s\,.\,s * \lambda n.0$$

$$H :\equiv \lambda p, s, \gamma\,.\,\gamma(\Phi(|s|, \lambda y.\Psi_2(\gamma(y))))$$

the functionals

$$\mathcal{F} :\equiv \lambda p\,.\,\mathsf{BR}_{\mathbb{N},1}(Y(p), G(p), H(p), \langle\,\rangle)$$

$$\mathcal{K} :\equiv \lambda p\,.\,\Psi_1(\mathcal{F}(p))$$

$$\mathcal{G} :\equiv \lambda p, y\,.\,\Psi_2(\mathsf{BR}_{\mathbb{N},1}(\overline{\mathcal{F}(p)}\,\mathcal{K}(p) * y))$$

realize f, k and g in (5.21). Let $\mathsf{BR}^*_{\mathbb{N},1}$ be the majorant of $\mathsf{BR}_{\mathbb{N},1}$ presented by Bezem [18]. Since we can easily find terms Y^*, G^* and H^* which majorize Y, G and H above, we get that

$$\mathcal{F}^* :\equiv \lambda p\,.\,\mathsf{BR}^*_{\mathbb{N},1}(Y^*(p), G^*(p), H^*(p), \langle\,\rangle)$$

$$\mathcal{K}^* :\equiv \lambda p\,.\,\Psi_1(\mathcal{F}^*(p))$$

$$\mathcal{G}^* :\equiv \lambda p, y\,.\,\Psi_2(\mathsf{BR}^*_{\mathbb{N},1}(\overline{\mathcal{F}^*(p)}\mathcal{K}^*(p) * \max(\overline{\mathcal{F}^*(p)}\mathcal{K}^*(p), y))),$$

where $\max(s,x) :\equiv \max\{s_0, \ldots, s_{|s|-1}, x\}$, are terms satisfying the monotone functional interpretation of AC^N. Note that $\lambda(a_n), \Psi_1, \Psi_2.\mathcal{F}(\Phi_{(a_n)}, \Psi_1, \Psi_2)$ realizes

$$\forall (a_n) \forall \Psi_1, \Psi_2 \exists f(\Psi_2(f) \geq f(\Psi_1(f)) \to |a_{f(\Psi_1 f)} - a_{\Psi_2 f}| \leq \tfrac{1}{\Psi_1 f + 1})),$$

and $\lambda(a_n), \Psi_1, \Psi_2.\mathcal{F}^*(\Phi^*_N, \Psi_1, \Psi_2)$ is a majorant for this realizer.

Moreover, notice that this realizer is also independent of the sequence $(a_n)_{n\in\mathbb{N}}$. Therefore, in the same way as we did in Section 5.7, uniformity results can still be obtained even when the full power of PCM is used in a proof of a theorem having the form (5.5).

Remark 5.5 By the above we can treat proofs in the system $\mathcal{T}^\omega + \mathsf{QF\text{-}AC}^{1,\mathbb{N}} + \mathsf{PCM}$. Note that for $\mathcal{T}^\omega = \mathsf{PRA}^\omega$ the above system can be viewed as a finite type extension of ACA_0 known from reverse mathematics. In that case the bound extracted by m.f.i. from a proof of a theorem of form (5.6) will be a closed term of type 2 of $\mathsf{PRA}^\omega[\mathsf{BR}_{\mathbb{N},1}]$ which (by [66, 97]) denotes a functional in Gödel's primitive recursive functionals T of finite type (note that PRA^ω only contains the fragment T_0 of T with primitive recursion on type $\mathbb{N}$).

Chapter 6

Proof Mining in L_1-approximation

with Ulrich Kohlenbach, *Annals of Pure and Applied Logic*, 121:1–38, 2003.

Abstract

In this paper we present another case study in the general project of proof mining which means the logical analysis of prima facie non-effective proofs with the aim of extracting new computationally relevant data. We use techniques based on monotone functional interpretation (developed in [92]) to analyze Cheney's simplification [33] of Jackson's original proof [70] from 1921 of the uniqueness of the best L_1-approximation of continuous functions $f \in C[0,1]$ by polynomials $p \in P_n$ of degree $\leq n$. Cheney's proof is non-effective in the sense that it is based on classical logic and on the non-computational principle WKL (binary König's lemma). The result of our analysis provides the first effective (in all parameters) uniform modulus of uniqueness (a concept which generalizes 'strong uniqueness' studied extensively in approximation theory). Moreover, the extracted modulus has the optimal ε-dependency as follows from Kroó [118]. The paper also describes how the uniform modulus of uniqueness can be used to compute the best L_1-approximations of a fixed $f \in C[0,1]$ with arbitrary precision. The second author uses this result to give a complexity upper bound on the computation of the best L_1-approximation in [134].

6.1 Introduction

This paper is another case study in the general project of proof mining which means the logical analysis of prima facie non-effective proofs with the aim of extracting new computationally relevant data[1]. At the same time we obtain new results in approximation theory. More specifically, we analyze a non-effective proof of the uniqueness of best approximations of continuous functions $f \in$

[1]See [90], [91], [100], [101] and [105] for other case studies as well as more information on proof mining in general.

$C[0,1]$ by polynomials $p \in P_n$ of degree $\le n$ with respect to the L_1-norm[2]

$$\|f\|_1 := \int_0^1 |f(x)|dx.$$

In [90], the first author showed how a quite general class of (non-effective) proofs of uniqueness theorems in analysis can be analyzed such that an effective so-called modulus of uniqueness can be extracted which generalises the concept of *strong unicity*[3]. In [90] and [91] this technique has been applied to the case of best Chebycheff approximation yielding new uniform bounds on constants of strong unicity and a new quantitative version of the alternation theorem. In this paper we apply this logical approach to investigate the quantitative rate of strong unicity for the quite different case of best L_1-approximation. Like Chebycheff approximation, L_1-approximation, also called 'approximation in the mean', is a classical topic in numerical mathematics and was considered already by Chebycheff in 1859 and has been investigated ever since (see [137] for a comprehensive survey). The uniqueness of the best L_1-approximation of $f \in C[0,1]$ by polynomials of degree $\le n$ was first proved in [70]. This proof uses measure theoretic arguments. A new uniqueness proof which avoids this and only uses the Riemann integral instead was given in 1965 by Cheney (see [33], [34]). Because of this feature, Cheney called his proof 'elementary'. From a logical point of view, however, it is highly non-constructive relying both on classical logic and non-computational analytical principles which correspond – in logical terminology – to the so-called binary ('weak') König's lemma, a principle which has received considerable attention in various parts of logic in recent years (see [153]). In this paper we carry out a complete logical analysis of Cheney's proof and show how the explicit modulus mentioned above can be extracted from this (seemingly) hopelessly non-constructive proof. Consequently, our result, like Cheney's proof, does not require any measure theory.

The main result of the present paper is the following effective strong uniqueness theorem:

Main result (Theorem 6.2) *Let* $\Phi(\omega,n,\varepsilon) := \min\{\frac{c_n\varepsilon}{8(n+1)^2}, \frac{c_n\varepsilon}{2}\omega_n(\frac{c_n\varepsilon}{2})\}$, where

$$c_n := \frac{\lfloor n/2\rfloor!\lceil n/2\rceil!}{2^{4n+3}(n+1)^{3n+1}} \text{ and } \omega_n(\varepsilon) := \min\{\omega(\tfrac{\varepsilon}{4}), \frac{\varepsilon}{40(n+1)^4\lceil\frac{1}{\omega(1)}\rceil}\}.$$

The functional Φ is a uniform modulus of uniqueness for the best L_1-approximation of any function f in $C[0,1]$ having modulus of uniform continuity ω from P_n, i.e.

$$\begin{cases} \forall n \in \mathbb{N}; p_1, p_2 \in P_n; \varepsilon \in \mathbb{Q}^*_+ \\ \big(\bigwedge_{i=1}^2(\|f-p_i\|_1 - dist_1(f,P_n) < \Phi(\omega,n,\varepsilon)) \to \|p_1-p_2\|_1 \le \varepsilon\big), \end{cases}$$

[2]For $f \in L_1$ uniqueness in general fails.

[3]The term *strong unicity* was introduced by Newman and Shapiro [131] in 1963 and has been studied extensively in approximation theory. See e.g. the introduction in [7] and the references given there for a discussion of the crucial importance of estimates of strong unicity for the convergence analysis of iterative algorithms and for stability analysis.

where $dist_1(f, P_n) := \inf_{p \in P_n} \|f - p\|_1$ *and* $\omega : \mathbb{Q}^*_+ \to \mathbb{Q}^*_+$ *is a modulus of uniform continuity for* $f \in C[0,1]$ *if* [4]

$$\forall x, y \in [0,1]; \varepsilon \in \mathbb{Q}^*_+ (|x - y| < \omega(\varepsilon) \to |f(x) - f(y)| < \varepsilon).$$

Moreover, this theorem can be proved in Heyting Arithmetic HA^ω in all finite types, and consequently holds in constructive mathematics in the sense of Bishop. Such a "constructivization", however, is not necessary for the extraction of Φ which is done from the ineffective proof. In fact, our verification of Φ is also done in $\mathsf{E\text{-}PA}^\omega + \mathsf{WKL}$. The fact that Φ can be verified in HA^ω then follows from a conservation result due to the first author.

The technical details of this analysis are mainly due to the second author who is using the results in a subsequent paper [134] to determine a complexity upper bound for the sequence $(p_n)_{n \in \mathbb{N}}$ of best approximating polynomials for poly-time computable functions $f \in C[0,1]$ (in the sense of [82], [83]).

6.1.1 Logical Background

Before going into the details of the analysis we need to recall some general logical background from [90][5]. First we introduce a little amount of logical terminology:

Let $\mathcal{A}^\omega$ be a (sub-)system of classical arithmetic in all finite types (like $\mathsf{E\text{-}PA}^\omega$ from [160] or Feferman's fragment $\mathsf{E\text{-}PRA}^\omega$ with quantifier-free induction and primitive recursion on the type 0 only [46]). Let $\mathcal{A}^\omega_*$ denote the extension of $\mathcal{A}^\omega$ by the schema

$$\mathsf{QF\text{-}AC} \quad : \quad \forall f^1 \exists x^0 A_{qf}(f, x) \to \exists F^2 \forall f^1 A_{qf}(f, F(f))$$

of quantifier-free choice from functions to numbers (where A_{qf} is quantifier-free) plus certain analytical principles Γ which – described in analytical terms – correspond to applications of Heine-Borel compactness of e.g. $[0,1]^d$. In logical terms, these principles correspond to the so-called binary ('weak') König's lemma WKL which suffices to derive a substantial amount of mathematics relative to weak fragments of arithmetic (see [153])[6]. In this paper the only genuine analytical tool Γ (which goes beyond $\mathsf{E\text{-}PA}^\omega + \mathsf{QF\text{-}AC}$) is the attainment of the infimum of continuous functions on compact intervals

$$\forall f \in C[0,1] \exists x \in [0,1] \big(f(x) = \inf_{y \in [0,1]} f(y)\big). \tag{6.1}$$

(6.1) is known to fail in computable analysis and even for poly-time computable f there will be in general no computable $x \in [0,1]$ satisfying (6.1) (see [83])[7].

[4]Note that this notion – used also in constructive mathematics and computable and feasible analysis – differs from the concept of modulus of continuity used in numerical analysis which we will discuss further below.

[5]Readers only interested in the numerical results but not in the general process of proof mining might skip this passage.

[6]$\mathsf{E\text{-}PRA}^\omega + \mathsf{QF\text{-}AC} + \mathsf{WKL}$ is a finite type extension of the system WKL_0 used in reverse mathematics and is (like the latter) Π^0_2-conservative over primitive recursive arithmetic PRA (see [5], [89]).

[7]The principle (6.1) is known to be equivalent to WKL over systems like $\mathsf{E\text{-}PRA}^\omega + \mathsf{QF\text{-}AC}$ even when f is given together with a modulus of uniform continuity, see [153].

Now, let X be a Polish space, K a compact Polish space and $F : X \times K \to \mathbb{R}$ a continuous function (moreover all these objects have to be explicitly representable in $\mathcal{A}^\omega$) and assume that we can prove in $\mathcal{A}^\omega_*$ that for every $f \in X$, $F(f, \cdot)$ has at most one root in K, i.e.[8]

$$\forall f \in X \forall x_1, x_2 \in K \big(\textstyle\bigwedge_{i=1}^2 F(f, x_i) = 0 \to x_1 = x_2 \big).$$

Then by a general logical meta-theorem proved in [90] (Theorem 4.3) one can extract from such a proof an explicit bound $\Phi(f, k)$ (given by a closed term of the underlying arithmetical system $\mathcal{A}^\omega$) such that

$$\begin{cases} \forall f \in X \forall k \in \mathbb{N} \forall x_1, x_2 \in K \\ \big(\bigwedge\limits_{i=1}^{2} (|F(f, x_i)| < 2^{-\Phi(f,k)}) \to d_K(x_1, x_2) < 2^{-k} \big), \end{cases} \tag{6.2}$$

where d_K denotes the metric on K. Moreover, (6.2) can be proved without using WKL and even in the intuitionistic variant $\mathcal{A}^\omega_i$ of $\mathcal{A}^\omega$ (and hence in constructive analysis in the sense of Bishop).

The proof of this meta-theorem provides an algorithm for actually extracting Φ. This algorithm is based on the proof-theoretic technique of monotone functional interpretation [92]. It is important to note that $\Phi(f, k)$ does **not** depend on $x_1, x_2 \in K$. Because of this fact, $\Phi(f, k)$ – which we call a **modulus of uniqueness** – can be used to compute the unique root (if existent) from any algorithm $\Psi(f, k)$ computing approximate so-called $\varepsilon(= 2^{-k})$-roots of $F(f, \cdot)$:

$$\forall f \in X \forall k \in \mathbb{N} \big(\Psi(f, k) \in K \wedge |F(f, \Psi(f, k))| < 2^{-k} \big). \tag{6.3}$$

One easily verifies that (6.2) and (6.3) imply that $\Psi(f, \Phi(f, k))$ is a Cauchy sequence in K which converges with rate of convergence 2^{-k} to the unique root $x \in K$ of $F(f, \cdot)$. So $x = \lim_{k\to\infty} \Psi(f, \Phi(f, k))$ can be computed with arbitrarily prescribed precision (which can also be proved in $\mathcal{A}^\omega_i$, see [90], Theorem 4.4) and the computational complexity of x can be estimated in terms of the complexities of Φ and Ψ (cf. [134]).

Remark 6.1 (Important!) *As usual in computable analysis (see [161]), the functionals $\Phi(f, k)$ and $\Psi(f, k)$ will depend not only on $f \in X$ in the set theoretic sense but on a (computationally meaningful) representation of f. In the case of $f \in C[0, 1]$, the representation of $C[0, 1]$ as a Polish space $(C[0, 1], \|\cdot\|_\infty)$ in $\mathcal{A}^\omega$ requires that f is endowed with a modulus of uniform continuity ω_f. So when we write $\Phi(f, k)$ we tacitly understand that f is given as a pair (f, ω_f). Actually, it now suffices to use the restriction f_r of f to the rational numbers in $[0, 1]$ (which can be enumerated so that f_r can be represented as a number theoretic function), since f can be reconstructed from f_r with the help of ω_f. In this way, the representation (f_r, ω_f) of f can be viewed as an object of type 1 so that computability on f reduces to the well-known type-2 notion of computability (see again [161] for more information on this).*

[8] We may even have functions $F : X \times Y \to \mathbb{R}$, where X, Y are general Polish spaces and can allow constructively definable families $(K_f)_{f\in X}$ of compact subspaces of Y which are parametrised by $f \in X$ instead of a fixed K. See [90] for details.

6.1.2 L_1-approximation

Let us now move to the case of best L_1-approximation treated in the present paper. The uniqueness of the best approximation can be written as follows

$$\begin{cases} \forall n \in \mathbb{N} \forall f \in C[0,1] \forall p_1, p_2 \in P_n \\ \big(\bigwedge_{i=1}^{2} (\|f - p_i\|_1 = dist_1(f, P_n)) \to p_1 = p_2 \big). \end{cases} \tag{6.4}$$

Note that in (6.4) we can without loss of generality replace the non-compact subspace P_n of $C[0,1]$ with the compact one $\tilde{K}_{f,n} := \{p \in P_n : \|p\|_1 \leq 2\|f\|_1\}$ since any best approximation p has to satisfy $\|f - p\|_1 \leq \|f\|_1$ because otherwise the zero polynomial would be a better approximation. As a consequence of this, $dist_1(f, P_n) = dist_1(f, \tilde{K}_{f,n})$ can easily be seen to be computable (uniformly in f as represented above and n). We use the slightly larger space $K_{f,n} := \big\{p \in P_n : \|p\|_1 \leq \frac{5}{2}\|f\|_1\big\}$ in (6.4) since a modulus of uniqueness for $K_{f,n}$ can be extended to P_n in a particular convenient way.

In this paper we analyze the above mentioned proof of Cheney for (6.4) as given in [33], [34][9] which uses the non-computational principle (6.1) (together with classical logic) but which can be formalized in $\mathcal{A}^\omega_*$ (as was shown in [88]). So the above mentioned result on the extractability of a modulus of uniqueness is applicable, i.e. the extractability of a (primitive recursive in the sense of Gödel's T) functional Φ satisfying

$$\begin{cases} \forall n, k \in \mathbb{N} \forall f \in C[0,1] \forall p_1, p_2 \in K_{f,n} \\ \big(\bigwedge_{i=1}^{2} (\|f - p_i\|_1 - dist_1(f, P_n) < 2^{-\Phi(f,n,k)}) \to \|p_1 - p_2\|_1 < 2^{-k} \big) \end{cases} \tag{6.5}$$

is guaranteed. Moreover, a simple trick (used also in [90] in the Chebycheff case) allows to replace $K_{f,n}$ with P_n so that

$$\begin{cases} \forall n, k \in \mathbb{N} \forall f \in C[0,1] \forall p_1, p_2 \in P_n \\ \big(\bigwedge_{i=1}^{2} (\|f - p_i\|_1 - dist_1(f, P_n) < 2^{-\Phi(f,n,k)}) \to \|p_1 - p_2\|_1 < 2^{-k} \big) \end{cases}$$

Remark 6.2 *Markov inequality states that for any polynomial p of degree $\leq n$, $\|p'\|_\infty \leq 2n^2\|p\|_\infty$, where p' denotes the first derivative of p. Using this inequality one can show that for any polynomial $p \in P_n$, $\|p\|_\infty \leq 2(n+1)^2\|p\|_1$. Hence, any upper bound on $\|p_1 - p_2\|_1$ gives also an upper bound on $\|p_1 - p_2\|_\infty$ and we can use this to get a bound on the coefficients of $p_1 - p_2$. Namely, if $p_1(x) - p_2(x) = a_n x^n + \ldots + a_1 x + a_0$ and $\|p_1 - p_2\|_1 < M$ then $|a_i| \leq \frac{(2(n+1)^2)^{i+1}}{i!} M$. We present the complete proof in Section 6.3.5.*

The importance of the modulus of uniqueness $\Phi(f, k)$ can also be illustrated by the fact that $\Phi + 1$ is automatically a modulus of pointwise continuity for the operator which maps $f \in X$ to its unique best approximation $f_b \in E \subset X$ (see [90]). For the special cases of Chebycheff resp. L_1-approximation this was shown first in [34] resp. [22]. Therefore,

[9]This result was first proved in [70] and is also called Jackson's Theorem. Cheney's proof (which applies to arbitrary Chebycheff systems) is a simplification of Jackson's proof.

$$\begin{cases} \forall n, k \in \mathbb{N} \forall f, \tilde{f} \in C[0,1] \\ \big(\|f - \tilde{f}\|_1 < 2^{-\Phi(f,n,k)-1} \to \|\mathcal{P}(f,n) - \mathcal{P}(\tilde{f},n)\|_1 < 2^{-k}\big), \end{cases}$$

where $\mathcal{P}(f,n)$ is the unique best L_1-approximation of $f \in C[0,1]$ from P_n.

Since $(C[0,1], \|\cdot\|_1)$ is not a Polish space we have to represent $C[0,1]$ as the space $(C[0,1], \|\cdot\|_\infty)$ to apply the logical meta-theorem mentioned above. As we discussed already, this amounts to enriching the input f by a modulus of uniform continuity ω_f so that Φ will also depend on ω_f.

Note that if $C[0,1]$ is replaced by the (pre-)compact (w.r.t. $\|\cdot\|_\infty$) set $\mathcal{K}_{\omega,M}$ of all functions $f \in C[0,1]$ which have the common modulus of uniform continuity ω and the common bound $\|f\|_\infty \leq M$, then the same logical meta-theorem guarantees the extractability of a modulus of uniqueness Φ which only depends on $\mathcal{K}_{\omega,M}$ i.e. on ω, M (in addition to n, k). Moreover, even the M-dependency can be eliminated as the approximation problem for f can be reduced to that for $\tilde{f}(x) := f(x) - f(0)$ so that only a bound $N \geq \sup_{x\in[0,1]} |f(x) - f(0)|$ is required, which can easily be computed from ω (e.g take $N := \lceil \frac{1}{\omega(1)} \rceil$). Therefore, from the logical meta-theorem and the fact that Cheney's proof can be formalized in $\mathsf{E\text{-}PA}^\omega + \mathsf{QF\text{-}AC} + \mathsf{WKL}$ we obtain already the extractability of a primitive recursive (in the sense of Gödel's T) modulus of uniqueness Φ which only depends on ω_f, n and k: **a-priori information.** Of course, only the actual extraction of Φ by applying the algorithm provided by the logical meta-theorem gives the detailed mathematical form of Φ as presented above: **a-posteriori information.**

6.2 Analysing Proofs in Analysis

The algorithm to be used for proof mining applied in cases like Cheney's proof of Jackson's Theorem (as treated in this paper) is based on the proof theoretic technique of monotone functional interpretation combined with negative translation as developed in [92]. Whereas the meta-mathematical details of this process are of importance to establish general meta-theorems on proof mining, this is not necessary for applications to specific proofs since here all numerical data will explicitly be exhibited and verified. This is because monotone functional interpretation explicitly transforms a given proof into another numerically enriched proof (in the normal mathematical sense). It is the strategy to find that proof (and to guarantee its existence) which is provided by the logical technique.

To approach the problem of proof mining applied to a logically involved proof as Cheney's, one starts off by splitting the proof into small pieces which are analyzed separately. As a consequence of the modularity of monotone functional interpretation one can easily combine the results obtained from the analysis of the pieces into a global result (this only requires functional application and λ-abstraction). Applications of monotone functional interpretation to the lemmas in the given proof at hand consist mostly of two steps,

1) transforming a given lemma L into a variant L^* which has the form

$$\forall n \in \mathbb{N} \forall x \in X \forall y \in K \exists k A_1(n, x, y, k), \tag{6.6}$$

where X is a Polish space, K a compact Polish space and $A_1 \in \Sigma_1^0$, and

2) extracting a bound $\Phi(n, x)$ for k which is independent of y.

It turns out that all the main lemmas to be analyzed have the form of (6.6). Because of this it is worthwhile to formulate the application of monotone functional interpretation to lemmas of this form as a special meta-theorem (6.1 below) which allows us to avoid having to go into the details of the underlying mechanism of functional interpretation each time. Although in the following we perform the transformation $L \mapsto L^*$ "by hand" one should note that this transformation is also usually automatically provided by functional interpretation.

Theorem 6.1 ([90], Theorem 4.1) *Let X, K be $\mathcal{A}^\omega$-definable Polish spaces, K compact and consider a sentence which can be written (when formalized in the language of $\mathcal{A}^\omega$) in the form*

$$A :\equiv \forall n \in \mathbb{N} \forall x \in X \forall y \in K \exists k \in \mathbb{N}\, A_1(n, x, y, k),$$

where A_1 is a purely existential. Then the following rule holds:[10]

$$\begin{cases} \mathcal{A}^\omega_* \vdash \forall n \in \mathbb{N} \forall x \in X \forall y \in K \exists k \in \mathbb{N}\, A_1(n, x, y, k) \\ \textit{then one can extract an } \mathcal{A}^\omega\textit{-definable functional } \Phi \textit{ s.t.} \\ \mathcal{A}^\omega_i \vdash \forall n \in \mathbb{N} \forall x \in X \forall y \in K \exists k \leq \Phi(n, x)\, A_1(n, x, y, k). \end{cases}$$

In particular, if

$$\mathcal{A}^\omega_i \vdash (k \leq \tilde{k} \wedge A_1(n, x, y, k)) \rightarrow A_1(n, x, y, \tilde{k})$$

then

$$\mathcal{A}^\omega_i \vdash \forall n \in \mathbb{N} \forall x \in X \forall y \in K\, A_1(n, x, y, \Phi(n, x)).$$

Again it is important to note that Φ does not depend on $y \in K$[11].

It is important to observe that real numbers are represented as Cauchy sequences $(a_n)_{n \in \mathbb{N}}$ of rational number with fixed rate of convergence (say 2^{-n}) i.e. $\forall k, \tilde{k} \geq n(|a_k - a_{\tilde{k}}| \leq 2^{-n})$. In this way, equality $=_\mathbb{R}$ (similarly $\leq_\mathbb{R}$ and $\geq_\mathbb{R}$) between real numbers is a $\forall$-statement (for any point $k+1$ in the Cauchy sequence the approximants are close by 2^{-k}) and strict inequality $<_\mathbb{R}$ is a $\exists$-statement (there exists a point $k+1$ in the sequence such that the approximants are distant by 2^{-k}). We call those: 'hidden quantifiers'. For example, let $a, b \in \mathbb{R}$, then $a <_\mathbb{R} b$ is an abbreviation for $\exists k \in \mathbb{N}(a_{k+1} + 2^{-k} <_\mathbb{Q} b_{k+1})$. When observing whether a lemma has the logical form of A above also the hidden quantifiers have to be taken into consideration. We can, however, avoid going into the representation of the real numbers by observing that $a <_\mathbb{R} b$ can be written either as $\exists r \in \mathbb{Q}^*_+(a <_\mathbb{R} b + r)$ or $\exists r \in \mathbb{Q}^*_+(a \leq_\mathbb{R} b + r)$. The idea

[10] As the theorem shows the conclusion can be proved already in $\mathcal{A}^\omega_i$ instead of $\mathcal{A}^\omega_*$. This, however, is not important for the applied aspect of the present paper where only the construction of Φ matters.

[11] Recall that $\Phi(n, x)$ will depend on the representation of $x \in X$.

is that, if $a <_{\mathbb{R}} b$ occurs positively we write it as $\exists r \in \mathbb{Q}^*_+(a <_{\mathbb{R}} b + r)$ and if it occurs negatively we write it as $\exists r \in \mathbb{Q}^*_+(a \leq_{\mathbb{R}} b + r)$, in this way after prenexing these quantifiers the matrix is purely existential and (given that the prenexed quantifiers have a $\forall\exists$ form as described in Theorem 6.1) we can apply our meta-theorem 6.1.

Moreover, the extractability of a Φ such that (6.5) holds can be also justifying by an application of the meta-theorem above. We just have to write (6.4) (after presenting the hidden quantifiers) as,

$$\begin{cases} \forall n \in \mathbb{N}; f \in C[0,1]; p_1, p_2 \in K_{f,n}; k \in \mathbb{N} \exists l \in \mathbb{N} \\ \big(\bigwedge_{i=1}^2 \|f - p_i\|_1 \leq dist_1(f, P_n) + 2^{-l} \to \|p_1 - p_2\|_1 < 2^{-k}\big), \end{cases}$$

which has the form A above. In [88] it is shown that Cheney's proof can be formalized in the system $\mathsf{E\text{-}PA}^\omega + \mathsf{QF\text{-}AC} + \mathsf{WKL}$, and since (as we will show) $K_{f,n}$ can be replaced by P_n the functional Φ realizing $\exists l$ in the formula above is in fact a uniform modulus of uniqueness for L_1-approximation of functions in $C[0,1]$ by polynomials in P_n. Therefore, from the meta-theorem 6.1 and previous discussions we obtain the following corollary (see [90], Theorems 4.1 and 5.1).

Corollary 6.1 *A functional $\Phi(f, n, k)$ given by a closed term of $\mathsf{E\text{-}PA}^\omega$ (i.e. a primitive recursive functional Φ in the sense of Gödel [58]) can be extracted from Cheney's proof of Jackson's Theorem so that,*

$$\begin{cases} \mathsf{(E\text{-})HA}^\omega \vdash \forall n \in \mathbb{N}; f \in C[0,1]; p_1, p_2 \in P_n; k \in \mathbb{N} \\ \big(\bigwedge_{i=1}^{2} (\|f - p_i\|_1 - dist_1(f, P_n)) \leq 2^{-\Phi(f,n,k)} \to \|p_1 - p_2\|_1 < 2^{-k}\big). \end{cases}$$

Moreover, using the Φ above, a primitive recursive functional Ψ can be constructed such that,

$$\begin{cases} \mathsf{(E\text{-})HA}^\omega \vdash \forall n \in \mathbb{N}; f \in C[0,1] \\ \big(\Psi(f, n) \in P_n \wedge \|f - \Psi(f, n)\|_1 = dist_1(f, P_n)\big). \end{cases}$$

In this paper we carry out the extraction of a modulus of uniqueness Φ from Cheney's proof of Jackson's theorem. We shall try to keep as separate as possible the mathematical and the logical parts of the analysis. Readers interested in the mathematical results can focus upon the claims together with their proofs. Meanwhile, for readers interested in the process of proof mining we try to explain how the various steps in our concrete 'mining' correspond to steps in the monotone functional interpretation (as used in the general meta-theorems). Those explanations usually precede the treatment of each lemma. This is important to serve the twofold goal of this paper, namely not only to prove new quantitative results in L_1-approximation theory but also to get further insights into the process of proof mining in general.

6.3 Analysis of Cheney's Proof of Jackson's Theorem

6.3.1 Logical Preliminaries on Cheney's Proof

In this section we sketch how a slight modification of Cheney's proof can be seen to be formalizable in basic arithmetic like $\mathcal{A}^\omega :\equiv \mathsf{E\text{-}PA}^\omega$ plus the already mentioned analytical principle (6.1), i.e. WKL. The only part of the proof which cannot be directly formalized in $\mathcal{A}^\omega$ is the so-called 'Lemma 1' (see [34], p. 219) which reads as follows

Lemma 6.1 ([34], Lemma 1) *Let $f, h \in C[0,1]$. If f has at most finitely many roots and if $\int_0^1 h\, sgn(f) \neq 0$, then for some $\lambda \in \mathbb{R}$, $\int_0^1 |f - \lambda h| < \int_0^1 |f|$, where*

$$sgn(f)(x) \overset{\mathbb{N}}{=} \begin{cases} 1, \text{ if } f(x) >_{\mathbb{R}} 0 \\ 0, \text{ if } f(x) =_{\mathbb{R}} 0 \\ -1, \text{ if } f(x) <_{\mathbb{R}} 0. \end{cases}$$

In the context of the Cheney's proof of Jackson's theorem, h will be a polynomial in P_n. Moreover, it will be shown that if f (for the particular f at hand) has only less than $n+1$ roots one can construct an h such that $\int_0^1 h\, sgn(f) \neq 0$. So we only need the lemma with the stronger assumption that f has fewer than $n+1$ roots. The existence of $sgn(f)$ relies on the existence of the characteristic function $\chi_{=_\mathbb{R}}$ for equality between reals which in turn is equivalent to the existence of Feferman's ([46]) non-constructive μ-operator (see [86]) and hence to a strong form of arithmetical comprehension which is not available in $\mathcal{A}^\omega_* :\equiv \mathcal{A}^\omega + \mathsf{WKL}$. However, the use of sgn can be eliminated as follows: if f has less than $n+1$ roots then there exist points $x_0 < \ldots < x_{n+1}$ in $[0,1]$ (where $x_0 = 0$ and $x_{n+1} = 1$) which contain all the roots of f. By classical logic and induction one shows in $\mathcal{A}^\omega$ the existence of a vector $(\sigma_1, \ldots, \sigma_{n+1}) \in \{-1,1\}^{n+1}$ such that

$$\sigma_i =_0 \begin{cases} 1, \text{ if } f \text{ is positive on } (x_{i-1}, x_i), \\ -1, \text{ if } f \text{ is negative on } (x_{i-1}, x_i) \end{cases}$$

for $i = 1, \ldots, n+1$. Therefore, $\int_0^1 h\, sgn(f)$ can be written as $\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h$. In Section 6.3.10 we shall see that this reformulation of Lemma 1 plays a crucial role in the analysis of Cheney's proof. Monotone functional interpretation of (the negative translation of) our version of Lemma 1 will automatically introduce the main notion needed for the quantitative analysis of the proof, namely the concept of so-called 'r-clusters of δ-roots'. This concept, furthermore, is the key for the elimination of the use of (6.1) (i.e. WKL) on which Cheney's proof of Lemma 1 relies[12].

[12] It is the argument that 'δ', in the middle of page 219 in [34], is strictly positive which uses (6.1). See Section 6.3.10 and Remark 6.3.10 for more information.

6.3.2 Analysing the Structure of the Proof

The main goal of the paper is to extract from Cheney's proof [34] of Jackson's theorem [70] an effective modulus of uniqueness which can be used, as it will be shown in Section 6.5, to compute the best L_1-approximation, p_n, from P_n of a given function $f \in C[0,1]$ with arbitrary precision[13]. In order to carry out the analysis we need to formalize Cheney's proof. The first step we take in this direction is to list the main formulas used in the proof and to show how they are combined into lemmas. As mentioned before, each lemma will be analyzed separately. The functional interpretation of the lemma shows which functionals *can be* extracted from the proof of the lemma. But not all the functionals need to be presented, since some of them will disappear in the analysis of the proof (see the treatment of modus pones in the soundness of functional interpretation, e.g. in [92]). By analyzing the structure of the whole proof we can see which functionals are relevant and *need to be* extracted in order to obtain the final result. Then we construct such functionals and prove that they realize the lemma. In Section 6.4 we show how the final modulus Φ is obtained by combining these functionals.

In the propositions $A - K$ below we omitted the parameters f, n, p_1 and p_2, therefore, instead of A one should read $A(f, n, p_1, p_2)$, where n ranges over $\mathbb{N}$, $f \in C[0,1]$ and $p_1, p_2 \in P_n$, and the same holds for all the others propositions. We also use here and for the rest of this paper the defined functions $p(x) := \frac{p_1(x)+p_2(x)}{2}$ and $f_0(x) := f(x) - p(x)$ as shorthand notation. In the formulas and in the sketch of the proof presented below we use $\overline{x} := x_1, \dots, x_n$ and $\overline{\sigma} := \sigma_1, \dots, \sigma_{n+1}$. The following formulas are used in Cheney's proof:

$A :\equiv \bigwedge_{i=1}^{2}(\|f - p_i\|_1 - dist_1(f, P_n) = 0)$, i.e.
p_1 and p_2 are best L_1-*approximations* of f from P_n.

$B :\equiv \|f - p\|_1 - dist_1(f, P_n) = 0$, i.e. p is a best L_1-*approximation* of f.

$C :\equiv \|f_0\|_1 = \frac{1}{2}\|f - p_1\|_1 + \frac{1}{2}\|f - p_2\|_1$.

$C_1 :\equiv \forall \varepsilon \in \mathbb{Q}_+^* \exists \delta \in \mathbb{Q}_+^* \forall x, y \in [0,1](|x - y| < \delta \to |g(x) - g(y)| < \varepsilon)$,
where $g(x) :\equiv |f_0(x)| - \frac{1}{2}|f(x) - p_1(x)| - \frac{1}{2}|f(x) - p_2(x)|$.
The formula C_1 states that g is uniformly continuous.

$D :\equiv \forall x \in [0,1](|f_0(x)| = \frac{1}{2}(|f(x) - p_1(x)| + |f(x) - p_2(x)|))$.

$E :\equiv \exists x_0, \dots, x_n \in [0,1]\big(\bigwedge_{i=0}^{n} f_0(x_i) = 0 \wedge \bigwedge_{i=1}^{n} x_{i-1} < x_i\big)$, i.e.
f_0 has at least $n+1$ distinct roots.

$F :\equiv \exists x_0, \dots, x_n \in [0,1]\big(\bigwedge_{i=0}^{n} p_1(x_i) = p_2(x_i) \wedge \bigwedge_{i=1}^{n} x_{i-1} < x_i\big)$, i.e.
$p_1 - p_2$ has at least $n+1$ distinct roots.

$G :\equiv \forall x \in [0,1](p_1(x) = p_2(x))$, alternatively, $\|p_1 - p_2\|_1 = 0$ or $p_1 = p_2$.

[13] P_n is a Haar subspace of $C[0,1]$ of dimension $n+1$.

$H(h) :\equiv \|f_0 - h\|_1 \geq \|f_0\|_1$.

$I(\overline{x}, \overline{\sigma}, h) :\equiv \sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x)dx > 0$, where $x_0 := 0$ and $x_{n+1} := 1$.

$J(\overline{x}) :\equiv \exists y \in [0,1](f_0(y) = 0 \wedge \bigwedge_{i=0}^{n+1} x_i \neq y)$, where $x_0 := 0$ and $x_{n+1} := 1$.

$K :\equiv \forall x \in [0,1](f_0(x) = 0 \rightarrow p_1(x) = p_2(x))$.

The first part of the proof (which we call derivation $\mathcal{D}_1$) is very simple and derives K from the assumption A,

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[A] \quad \dfrac{[A] \quad A \rightarrow B}{B}}{A \wedge B} \quad A \wedge B \rightarrow C}{C} \quad C_1}{C \wedge C_1} \quad C \wedge C_1 \rightarrow D}{D} \quad D \rightarrow K}{K}$$

The most involved part of the proof (which includes the application of Lemma 1) is when we want to prove that f_0 has $n+1$ distinct roots. In the derivations below we use $\overline{\sigma}' := \sigma'_1, \ldots, \sigma'_{n+1}$, where $\sigma'_i := sgn\,(f_0)(\frac{x_{i-1}+x_i}{2})$. Moreover, $\forall \overline{x} := \forall x_1 \leq \ldots \leq x_n$, where $\forall x_1 \leq \ldots \leq x_n\ Q(\overline{x})$ is an abreviation for $\forall x_1, \ldots, x_n\ (x_1 \leq \ldots \leq x_n \rightarrow Q(\overline{x}))$. Let the following derivation

$$\dfrac{\forall \overline{x}, \overline{\sigma} \exists \tilde{h}_{\overline{x},\overline{\sigma}} I(\overline{x}, \overline{\sigma}, \tilde{h}_{\overline{x},\overline{\sigma}}) \quad \dfrac{\forall \overline{x}, h\ (\forall \lambda H(\lambda h) \wedge I(\overline{x}, \overline{\sigma}', h) \rightarrow J(\overline{x}))}{\forall \overline{x}\ (\forall \lambda H(\lambda \tilde{h}_{\overline{x},\overline{\sigma}'}) \wedge I(\overline{x}, \overline{\sigma}', \tilde{h}_{\overline{x},\overline{\sigma}'}) \rightarrow J(\overline{x}))}}{\forall \lambda H(\lambda \tilde{h}_{\overline{x},\overline{\sigma}'}) \rightarrow \forall \overline{x} J(\overline{x})}$$

be named $\mathcal{D}_2$. Using $\mathcal{D}_2$ from the assumption A we can derive that f_0 has $n+1$ distinct roots.

$$\dfrac{\dfrac{\dfrac{[A] \quad A \rightarrow B}{B} \quad B \rightarrow \forall h\ H(h)}{\forall h\ H(h)} \quad \dfrac{\mathcal{D}_2}{\forall \lambda H(\lambda \tilde{h}_{\overline{x},\overline{\sigma}'}) \rightarrow \forall \overline{x} J(\overline{x})}}{\forall \overline{x}\ J(\overline{x})}$$

We call this derivation $\mathcal{D}_3$. An outline of the whole proof in the form of an informal natural deduction derivation is presented below,

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\mathcal{D}_1}{K} \quad \dfrac{\dfrac{\mathcal{D}_3}{\forall \overline{x}\ J(\overline{x})} \quad \forall \overline{x}\ J(\overline{x}) \rightarrow E}{E}}{K \wedge E} \quad K \wedge E \rightarrow F}{F} \quad F \rightarrow G}{G}}{A \rightarrow G}\,[A]$$

Remark 6.3 *In general, we can only apply our meta-theorem 6.1 if P_n is replaced by $K_{f,n}$. As it happened, only in Section 6.3.5 this limitation really matters. Nonetheless, as we discussed already, at the end of the article we show that the final result actually holds for P_n.*

6.3.3 Lemma $A \to B$ [Triangle Inequality]

The first lemma states,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n} \\ \big(\bigwedge_{i=1}^{2} \|f - p_i\|_1 = dist_1(f, P_n) \to \|f - p\|_1 = dist_1(f, P_n)\big). \end{cases}$$

As described in the previous section, the first step is to present the hidden quantifiers,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n} \\ \big(\forall \delta \in \mathbb{Q}_+^*(\bigwedge_{i=1}^{2} \|f - p_i\|_1 - dist_1(f, P_n) \le \delta) \to \\ \forall \varepsilon \in \mathbb{Q}_+^*(\|f - p\|_1 - dist_1(f, P_n) < \varepsilon)\big). \end{cases}$$

Then we look at the functional interpretation of the lemma,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}_+^* \exists \delta \in \mathbb{Q}_+^* \\ \big(\bigwedge_{i=1}^{2} \|f - p_i\|_1 - dist_1(f, P_n) \le \delta \to \\ \|f - p\|_1 - dist_1(f, P_n) < \varepsilon\big). \end{cases} \tag{6.7}$$

We see now that (6.7) has the same structure as the formula A in Theorem 6.1. Therefore, we are sure to find a functional Φ_1, depending at most on n, f and ε, such that,[14]

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}_+^* \exists \delta \ge \Phi_1(f, n, \varepsilon) \\ \big(\bigwedge_{i=1}^{2}(\|f - p_i\|_1 - dist_1(f, P_n) < \delta) \to \\ \|f - p\|_1 - dist_1(f, P_n) < \varepsilon\big). \end{cases} \tag{6.8}$$

Since we have monotonicity in δ the functional Φ_1 actually realizes δ. The same phenomenon will happen in all the following lemmas, i.e. the lower bounds will always be realizing functionals for the variables they bound. Here, it is obvious how to construct Φ_1,

Claim 6.1 *The functional* $\boxed{\Phi_1(f, n, \varepsilon) := \Phi_1(\varepsilon) := \varepsilon}$ *does the job*[15].

Proof. Suppose (i) $\|f - p_1\|_1 - dist_1(f, P_n) < \varepsilon$ and (ii) $\|f - p_2\|_1 - dist_1(f, P_n) < \varepsilon$. Multiplying (i) and (ii) by $1/2$ and adding them together we get $1/2(\|f - p_1\|_1 + \|f - p_2\|_1) - dist_1(f, P_n) < \varepsilon$. By the triangle inequality for the L_1-*norm*, $1/2(\|2f - p_1 - p_2\|_1) - dist_1(f, P_n) < \varepsilon$, i.e. $\|f - p\|_1 - dist_1(f, P_n) < \varepsilon$. $\square$

[14] Since in Theorem 6.1 we used 2^{-k} (with $k \in \mathbb{N}$) instead of $\delta \in \mathbb{Q}_+^*$, the upper bound on k guaranteed by the meta-theorem gives a lower bound on δ.

[15] Note that in fact Φ_1 is independent of n and f. We adopt the convention that parameters not used in the definition of the functionals will be dropped.

Remark 6.4 *The reader may have noticed that from* (6.7) *to* (6.8) *we changed from* $\leq$ *to* $<$ *in the premise of the implication. The reason we wrote* $\leq$ *first was just to show that the lemma could be written in the form of* A *(from Theorem 6.1) and that a functional realizing* δ *was guaranteed by our meta-theorem. Since* $a \leq b/2$ *implies* $a < b$ *(and the reverse implication holds without the factor* $1/2$*) we normally write the relation that yields the optimal bound. When analysing the following lemmas we often claim that some sentence is an instance of our meta-theorem 6.1 without bothering to write it explicitly in the form of* A. *We hope the reader can see that through the implications mentioned above these lemmas could in fact be written in the form of* A.

6.3.4 Lemma $A \wedge B \to C$ [Basic Norm Property]

The lemma states,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n} \\ \big(\bigwedge_{i=1}^{2}(\|f - p_i\|_1 = dist_1(f, P_n)) \to \\ \|f - p\|_1 - 1/2\|f - p_1\|_1 - 1/2\|f - p_2\|_1 = 0\big). \end{cases}$$

After presenting the hidden quantifiers and performing the functional interpretation we come again to the same logical structure of the formula in Theorem 6.1, and again we know that there must exist a functional Φ_2 depending at most on n, f and ε such that,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}_+^* \\ \big(\bigwedge_{i=1}^{2}(\|f - p_i\|_1 - dist_1(f, P_n) < \Phi_2(f, n, \varepsilon)) \to \\ |\ \|f - p\|_1 - 1/2\|f - p_1\|_1 - 1/2\|f - p_2\|_1\ | < \varepsilon\big). \end{cases}$$

Again, the choice of Φ_2 is simple,

Claim 6.2 *The functional* $\boxed{\Phi_2(f, n, \varepsilon) := \Phi_2(\varepsilon) := \varepsilon}$ *does the job.*

Proof. Suppose (i) $\|f - p_1\|_1 - dist_1(f, P_n) < \varepsilon$ and (ii) $\|f - p_2\|_1 - dist_1(f, P_n) < \varepsilon$. By previous lemma we have (iii) $\|f - p\|_1 - dist_1(f, P_n) < \varepsilon$. And $\frac{(i)+(ii)}{2}$ gives (iv) $1/2(\|f - p_1\|_1 + \|f - p_2\|_1) - dist_1(f, P_n) < \varepsilon$. From (iii) and (iv), we have, $|\ \|f - p\|_1 - 1/2\|f - p_1\|_1 - 1/2\|f - p_2\|_1\ | < \varepsilon$, since if $a \in [0, m)$ and $b \in [0, m)$ then $|a - b| \in [0, m)$. $\square$

6.3.5 Lemma C_1 [Continuity of $g(x)$]

Let $g(x) := |f_0(x)| - \frac{1}{2}|f(x) - p_1(x)| - \frac{1}{2}|f(x) - p_2(x)|$. Based on the continuity of f, p_1 and p_2 we derive that g is continuous,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}_+^*; x, y \in [0,1] \exists \delta \in \mathbb{Q}_+^* \\ \big(|x - y| \leq \delta \to |g(x) - g(y)| < \varepsilon\big). \end{cases}$$

Note that here we can again apply the meta-theorem 6.1 and we are sure to find a function Δ depending only f, n and ε such that,[16]

[16] Here it is fundamental that p_1 and p_2 live in the compact space $K_{f,n}$ otherwise the modulus of continuity for g would depend also on these elements and we would be unable to get a uniform modulus of uniqueness at the end.

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}_+^*; x, y \in [0,1] \\ \big(|x-y| < \Delta(f,n,\varepsilon) \to |g(x)-g(y)| < \varepsilon\big). \end{cases}$$

We write $\Delta(f,n,\varepsilon)$ as $\omega_{f,n}(\varepsilon)$. In this section we show how the modulus of continuity $\omega_{f,n}(\varepsilon)$ can be computed using only n, the modulus of continuity of f, ω_f, and an upper bound $M_f \geq \|f\|_\infty$ (in Section 6.4 we show that we just need a bound M_f on $\sup_{x\in[0,1]} |f(x) - f(0)|$, for instance $\lceil \frac{1}{\omega_f(1)} \rceil$, so that the final result only depends on ω_f and n). It should be noted that the modulus of continuity of a function is not unique, therefore when in the following we write $\omega_f(\varepsilon) := \ldots$ we mean that $\ldots$ can be taken as the modulus of continuity of the function f.

Modulus of the Sum

Given the moduli of continuity ω_f and ω_g for the functions f and g respectively, we find the modulus of continuity for $f+g$, ω_{f+g}, in the following way. We have,

$$|x-y| < \omega_f(\varepsilon/2) \to |f(x) - f(y)| < \varepsilon/2.$$

$$|x-y| < \omega_g(\varepsilon/2) \to |g(x) - g(y)| < \varepsilon/2.$$

Therefore,

$$|x-y| < \min\{\omega_f(\varepsilon/2), \omega_g(\varepsilon/2)\} \to \\ (|f(x)-f(y)| < \varepsilon/2 \wedge |g(x)-g(y)| < \varepsilon/2).$$

$$|x-y| < \min\{\omega_f(\varepsilon/2), \omega_g(\varepsilon/2)\} \to |f(x) + g(x) - f(y) - g(y)| < \varepsilon.$$

Hence, $\omega_{f+g}(\varepsilon) := \min\{\omega_f(\varepsilon/2), \omega_g(\varepsilon/2)\}$.

Modulus of a Constant Times a Function

We show that $\omega_{af}(\varepsilon) := \omega_f(\frac{\varepsilon}{a})$. For all $a \in \mathbb{Q}_+^*$, if $|x-y| < \omega_f(\frac{\varepsilon}{a})$ then $|f(x)-f(y)| < \frac{\varepsilon}{a}$, and therefore, $|af(x) - af(y)| < \varepsilon$.

Modulus of p_1 and p_2

Let $p_i \in K_{f,n}$. Then $\|p_i\|_1 \leq \frac{5}{2}\|f\|_1 \leq \frac{5}{2}\|f\|_\infty$. If $p_i(x) = a_n x^n + \ldots + a_1 x + a_0$ and $p_i^*(x) = \frac{a_n x^{n+1}}{n+1} + \ldots + \frac{a_1 x^2}{2} + a_0 x$ then for all $x \in [0,1]$ we have,

$$|p_i^*(x)| = |\int_0^x p_i(x)dx| \leq \int_0^x |p_i(x)|dx \leq \|p_i\|_1 \leq \tfrac{5}{2}\|f\|_\infty,$$

i.e. $\|p_i^*\|_\infty \leq \|p_i\|_1 \leq \frac{5}{2}\|f\|_\infty$. By Markov inequality (see e.g. [34]),

$$\|p_i\|_\infty = \|(p_i^*)'\|_\infty \leq 2(n+1)^2\|p_i^*\|_\infty \leq 2(n+1)^2(\tfrac{5}{2}\|f\|_\infty) = 5(n+1)^2\|f\|_\infty.$$

If we apply Markov inequality once more we get,

$$\|p_i'\|_\infty \leq 2n^2 5(n+1)^2\|f\|_\infty < 10(n+1)^4\|f\|_\infty.$$

By the mean value theorem this implies that p_i has Lipschitz constant $10(n+1)^4\|f\|_\infty$ on $[0,1]$, i.e. $\frac{\varepsilon}{10(n+1)^4\|f\|_\infty}$ is a modulus of uniform continuity for p_i on $[0,1]$. Given an upper bound M_f on $\|f\|_\infty$ we have,[17]

$$\omega_{p_i}(\varepsilon) := \tfrac{\varepsilon}{10(n+1)^4 M_f}.$$

Remark 6.5 *Here we present how one gets a bound on the coefficients of* p *given* $\|p\|_1$ *(or some bound on* $\|p\|_1$*). Let* p^i *denote the* i*-th derivative of* p*. Above we have shown that* $\|p\|_\infty \le 2(n+1)^2\|p\|_1$ *which by Markov inequality yields (+)* $\|p^i\|_\infty \le (2(n+1)^2)^{i+1}\|p\|_1$*. Since* $p^i(x) = \frac{n!}{(n-i)!}a_n x^{n-i} + \ldots + i!a_i$*, from (+) we get* $|i!\, a_i| \le (2(n+1)^2)^{i+1}\|p\|_1$ *which implies* $|a_i| \le \frac{(2(n+1)^2)^{i+1}}{i!}\|p\|_1$*.*

The Modulus of Continuity $\omega_{f,n}$

Now we can present $\omega_{f,n}$ as a function of ω_f and n (note that we can take $\omega_{|f|} := \omega_f$),

$$\begin{aligned}
\omega_{f,n}(\varepsilon) &= \min\{\omega_{|f-p|}(\varepsilon/2), \omega_{1/2|f-p_1|}(\varepsilon/4), \omega_{1/2|f-p_2|}(\varepsilon/4)\} \\
&= \min\{\omega_{f-p}(\varepsilon/2), \omega_{f-p_1}(\varepsilon/2), \omega_{f-p_2}(\varepsilon/2)\} \\
&= \min\{\omega_f(\varepsilon/4), \omega_{p_1}(\varepsilon/4), \omega_{p_2}(\varepsilon/4)\} \\
&= \min\{\omega_f(\frac{\varepsilon}{4}), \frac{\varepsilon}{40(n+1)^4 M_f}\}.
\end{aligned}$$

6.3.6 Lemma $C \wedge C_1 \to D$ [Integrand is ≤ 0 and Continuous]

Let $g(x) := |f(x) - p(x)| - 1/2|f(x) - p_1(x)| - 1/2|f(x) - p_2(x)|$. The lemma says,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}\big(\int_0^1 g(x)dx = 0 \to \forall x \in [0,1](g(x) = 0)\big).$$

After presenting the hidden quantifiers and applying functional interpretation we observe that again we can apply Theorem 6.1, and we are guaranteed to find a functional $\Phi_3(f, n, \varepsilon)$ such that,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}^*_+ \\ \big(|\int_0^1 g(x)dx| \le \Phi_3(f,n,\varepsilon) \to \|g\|_\infty \le \varepsilon\big). \end{cases}$$

Let $\omega_{f,n} : \mathbb{Q}^*_+ \to \mathbb{Q}^*_+$ denote the modulus of uniform continuity of the function $g \in C[0,1]$, proved to exist in the analysis of lemma C_1 (Section 6.3.5).

Claim 6.3 *The functional* $\boxed{\Phi_3(f,n,\varepsilon) := \Phi_3(\omega_{f,n},\varepsilon) := \frac{\varepsilon}{2}\cdot\min\{\frac{1}{2},\ \omega_{f,n}(\frac{\varepsilon}{2})\}}$ *does the job.*

Proof. Assume $\|g\|_\infty > \varepsilon$, since $\forall x \in [0,1](g(x) \le 0)$ we conclude $\exists x_0 \in [0,1](g(x_0) \le -\varepsilon)$. By the continuity of g we also have,

[17] It should be clear that given f together with its modulus of continuity, ω_f, there is a simple algorithm to compute M_f, just take for instance $M_f := \max\{|f(i.\omega_f(1))| : 0 \le i \le \lfloor\frac{1}{\omega_f(1)}\rfloor\}+1$.

$$\forall x \in [0,1]\big(|x - x_0| < \omega_{f,n}(\varepsilon/2) \to g(x) < -\varepsilon/2\big).$$

If $x_0 < 1/2$ then,

$$|\int_0^1 g(x)dx| > |\int_{x_0}^{\min\{1,x_0+\omega_{f,n}(\varepsilon/2)\}} -\varepsilon/2\ dx| = \frac{\varepsilon}{2}\min\{\frac{1}{2},\omega_{f,n}(\frac{\varepsilon}{2})\},$$

otherwise ($x_0 \geq 1/2$),

$$|\int_0^1 g(x)dx| > |\int_{\max\{0,x_0-\omega_{f,n}(\varepsilon/2)\}}^{x_0} -\varepsilon/2\ dx| = \frac{\varepsilon}{2}\min\{\frac{1}{2},\omega_{f,n}(\frac{\varepsilon}{2})\}.$$

From this we conclude,

$$|\int_0^1 g(x)dx| > \frac{\varepsilon}{2}\min\{\frac{1}{2},\omega_{f,n}(\frac{\varepsilon}{2})\}. \quad \square$$

6.3.7 Lemma $D \to K$ [If $f_0(x) = 0$ then $p_1(x) = p_2(x)$]

Let $f_1(x) := 1/2(|f(x) - p_1(x)| + |f(x) - p_2(x)|)$, the lemma says,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; x \in [0,1] \\ \big(\| |f_0| - f_1\|_\infty = 0 \to (|f_0(x)| = 0 \to p_1(x) = p_2(x))\big). \end{cases}$$

Again we are sure to find functionals $\Phi_4(f,n,\varepsilon)$ and $\Phi_5(f,n,\varepsilon)$ such that,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; x \in [0,1]; \varepsilon \in \mathbb{Q}_+^* \\ \big(\| |f_0| - f_1\|_\infty \leq \Phi_4(f,n,\varepsilon) \to \\ (|f_0(x)| \leq \Phi_5(f,n,\varepsilon) \to |p_1(x) - p_2(x)| \leq \varepsilon)\big). \end{cases}$$

Claim 6.4 *The functionals* $\boxed{\Phi_4(f,n,\varepsilon) := \Phi_4(\varepsilon) := \varepsilon/8}$ *and* $\boxed{\Phi_5(f,n,\varepsilon) := \Phi_5(\varepsilon) := \varepsilon/8}$ *do the job.*

Proof. Trivial. $\square$

6.3.8 Lemma $F \to G$ [If p Has $n+1$ Roots Then $p = 0$]

The lemma states that if the polynomial $p_1(x) - p_2(x)$ has $n+1$ distinct roots in the interval $[0,1]$ then $p_1(x)$ and $p_2(x)$ are actually identical,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n} \forall x_0, \ldots, x_n \in [0,1] \\ \big(\bigwedge_{i=1}^n (x_i < x_{i+1}) \wedge \bigwedge_{i=0}^n (p_1(x_i) = p_2(x_i)) \to \|p_1 - p_2\|_\infty = 0\big). \end{cases}$$

Then we present the hidden quantifiers and apply functional interpretation,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; r, \varepsilon \in \mathbb{Q}_+^*; x_0, \ldots, x_n \in [0,1] \exists \delta \in \mathbb{Q}_+^* \\ \big(\bigwedge_{i=1}^n (x_{i-1} + r \leq x_i) \wedge \bigwedge_{i=0}^n (|p_1(x_i) - p_2(x_i)| \leq \delta) \to \|p_1 - p_2\|_\infty \leq \varepsilon\big). \end{cases}$$

By Theorem 6.1 we are sure to find a functional Φ_6 realizing δ.

Claim 6.5 *The functional* $\boxed{\Phi_6(f,n,r,\varepsilon) := \Phi_6(n,r,\varepsilon) := \frac{\lfloor n/2 \rfloor! \lceil n/2 \rceil! r^n}{(n+1)} \varepsilon}$ *does the job.*

Proof. See [90], pages 82–83. $\square$

Remark 6.6 *In fact, the functional Φ_6 does the job for $p_1, p_2 \in P_n$ (not only for $p_1, p_2 \in K_{f,n}$).*

6.3.9 Lemma $B \to \forall h\ H(h)$ [Definition of Best L_1-approximation]

This lemma is a trivial consequence of the definition of $dist_1$,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n} \\ \big(\|f_0\|_1 = dist_1(f, P_n) \to \forall h \in P_n(\|f_0 - h\|_1 \geq \|f_0\|_1)\big). \end{cases}$$

We can easily find a functional $\Phi_7(f, n, \varepsilon)$ s.t.,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}^*_+ \\ \big(\|f_0\|_1 - dist_1(f, P_n) \leq \Phi_7(f, n, \varepsilon) \to \\ \forall h \in P_n(\|f_0 - h\|_1 + \varepsilon \geq \|f_0\|_1)\big). \end{cases}$$

Claim 6.6 *The functional* $\boxed{\Phi_7(f, n, \varepsilon) := \Phi_7(\varepsilon) := \varepsilon}$ *does the job.*

Proof. Assume (i) $\|f_0\|_1 - dist_1(f, P_n) \leq \varepsilon$. By the definition of $dist_1$ we have for any $h \in P_n$ (ii) $\|f_0 - h\|_1 = \|f - (p + h)\|_1 \geq dist_1(f, P_n)$. From (i) and (ii) we have $\|f_0 - h\|_1 + \varepsilon \geq \|f_0\|_1$. $\square$

6.3.10 Lemma $\forall \overline{x}, h\ (\forall \lambda H(\lambda h) \wedge I(\overline{x}, \overline{\sigma}', h) \to J(\overline{x}))$ [Lemma 1]

This is the most intricate lemma used in the proof, hence we analyze it in greater detail. We first rewrite the lemma as it is stated in [34]. The contraposition of Lemma 1 is used in the proof.

Lemma 6.2 (Lemma 1) *Let $f \in C[0,1]$, $n \in \mathbb{N}$ and $h, p_1, p_2 \in P_n$. If f_0 has at most n roots then either $\int_0^1 (h(x)\, sgn(f_0)(x))dx = 0$ or there exists a $\lambda \in \mathbb{R}$ such that $\int_0^1 |f_0(x) - \lambda h(x)|dx < \int_0^1 |f_0(x)|dx$.*

Proof. Assume that all the roots of f_0 are among $0 = x_0 \leq x_1 \leq \ldots \leq x_{n+1} = 1$ and w.l.g. assume that $\int_0^1 (h(x)\, sgn(f_0)(x))dx > 0$. Let $B' := \bigcup_{i=0}^{n+1}(x_i - r, x_i + r)$ and $B := B' \bigcap [0,1]$. Let $A := [0,1] \backslash B$. Make r small enough so that $\int_A (h(x)\, sgn(f_0)(x))dx > \int_B |h(x)|dx$. Note that A is a finite union of closed intervals which contain no roots of f_0, therefore $\delta := \min\{|f_0(x)| : x \in A\}$ is positive. Hence we can find a λ such that $0 < \lambda \|h\|_\infty < \delta$, and for points $x \in A$, $sgn(f_0 - \lambda h)(x) = sgn(f_0)(x)$, which implies (see [34] or the proof of Claim 6.7 for details) that $\int_0^1 |f_0(x) - \lambda h(x)|dx < \int_0^1 |f_0(x)|dx$. $\square$

Logical Analysis of Lemma 1

The Lemma 1 as it is presented above does not have the logical form to which we can apply the meta-theorem 6.1. We can, however, show that a variation of the Lemma 1, which can be used in Cheney's proof does have that logical form. Let $B' := \bigcup_{i=0}^{n+1}(x_i - r, x_i + r)$, $B := B' \bigcap [0,1]$ and $A := [0,1] \backslash B$, where $x_0 := 0$ and $x_{n+1} := 1$. Note that A can be written as the union of smaller intervals[18] $A_i := [x_{i-1} + \min\{r, \frac{x_i - x_{i-1}}{2}\}, x_i - \min\{r, \frac{x_i - x_{i-1}}{2}\}]$, for $1 \leq i \leq n+1$. For the rest of Section 6.3 we use x_0, x_{n+1}, A, B and A_i as defined above and we

[18] Note that the intervals $\bigcup A_i$ and A only differ on at most a finite number of points. Clearly, however, the integrations $\sum \int_{A_i}$ and $\int_A$ coincide.

mention explicitly which r we are using when this is not clear from the context. The version of Lemma 1 we consider is: For all $f \in C[0,1]$ and $n \in \mathbb{N}$

$$\begin{cases} \forall p_1, p_2 \in P_n; x_1 \leq \ldots \leq x_n \in [0,1]; h \in P_n; r \in \mathbb{Q}^*_+ \\ \big(\forall y \in A(fy \neq 0) \wedge \int_A h\ sgn(f) > \int_B |h| \to \\ \exists \lambda \in \mathbb{R}(\|f - \lambda h\|_1 < \|f\|_1)\big) \end{cases} \tag{6.9}$$

where A, B depend on $x_1 \leq \ldots \leq x_n$ and r.

First we show how (6.9) can be used in Cheney's proof. Since f will be taken to be f_0 we can prove $\forall \lambda \in \mathbb{R}; h \in C[0,1](\|f_0 - \lambda h\|_1 \geq \|f_0\|_1)$ which leaves, for all $f \in C[0,1]$ and $n \in \mathbb{N}$

$$\begin{cases} \forall p_1, p_2 \in P_n; x_1 \leq \ldots \leq x_n \in [0,1]; h \in P_n; r \in \mathbb{Q}^*_+ \\ \big(\exists y \in A(f_0(y) = 0) \vee \int_A h\ sgn(f_0) \leq \int_B |h|\big) \end{cases}$$

but we can easily prove

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in P_n; x_1 \leq \ldots \leq x_n \in [0,1] \\ \exists h \in C[0,1]; r \in \mathbb{Q}^*_+ \big(\forall y \in A(f_0(y) \neq 0) \to \int_A h\ sgn(f_0) > \int_B |h|\big) \end{cases}$$

from which we can obtain the existence of $n+1$ roots by induction.

Now we can replace P_n with $K_{f,n}$ in (6.9) and rewrite the integral of $h\ sgn(f_0)$ over the intervals A as a sum of integrals over smaller intervals A_i (which are guaranteed by the premise to contain no root of f_0) as described in Section 6.3.1. Hence Lemma 1 can be formally written as, for all $f \in C[0,1]$ and $n \in \mathbb{N}$

$$\begin{cases} \forall p_1, p_2 \in K_{f,n}; x_1 \leq \ldots \leq x_n \in [0,1]; h \in P_n; r \in \mathbb{Q}^*_+ \\ \big(\forall y \in A(f_0(y) \neq 0) \wedge \sum_{i=1}^{n+1} \sigma_i \int_{A_i} h > \int_B |h| \to \\ \exists \lambda \in \mathbb{R}(\|f_0 - \lambda h\|_1 < \|f_0\|_1)\big) \end{cases}$$

where $\sigma_i := sgn(f_0)(\frac{x_{i-1}+x_i}{2})$, $x_0 := 0$ and $x_{n+1} := 1$. Presenting the hidden quantifiers we obtain[19], for all $f \in C[0,1]$ and $n \in \mathbb{N}$

$$\begin{cases} \forall p_1, p_2 \in K_{f,n}; x_1 \leq \ldots \leq x_n \in [0,1]; h \in P_n; \delta, r, \eta \in \mathbb{Q}^*_+\ \exists l \in \mathbb{Q}^*_+ \\ \big(\forall y \in A(|f_0(y)| \geq \delta) \wedge \sum_{i=1}^{n+1} \sigma_i \int_{A_i} h \geq \int_B |h| + \eta \to \\ \exists \lambda \in \mathbb{R}(\|f_0 - \lambda h\|_1 + l < \|f_0\|_1)\big). \end{cases}$$

This last step can be viewed as a weakening of the Lemma 1 since we replace $\forall y \in A(f_0(y) \neq 0)$ by the stronger statement $\exists \delta \in \mathbb{Q}^*_+ \forall y \in A(|f_0(y)| \geq \delta)$ in the premise. In view of WKL, however, we have that the above formula actually implies the original Lemma 1. Note that we can take $\eta = 1$ w.l.g. since $h/\eta \in P_n$. Hence, get for all $f \in C[0,1]$ and $n \in \mathbb{N}$

$$\begin{cases} \forall p_1, p_2 \in K_{f,n}; x_1 \leq \ldots \leq x_n \in [0,1]; h \in P_n; \delta, r \in \mathbb{Q}^*_+\ \exists l \in \mathbb{Q}^*_+ \\ \big(\forall y \in A(|f_0(y)| \geq \delta) \wedge \sum_{i=1}^{n+1} \sigma_i \int_{A_i} h \geq \int_B |h| + 1 \to \\ \exists \lambda \in \mathbb{R}(\|f_0 - \lambda h\|_1 + l < \|f_0\|_1)\big) \end{cases} \tag{6.10}$$

[19] Using that by WKL, $\forall y \in A(f_0(y) \neq 0) \leftrightarrow \exists \delta \in \mathbb{Q}^*_+ \forall y \in A(|f_0(y)| \geq \delta)$.

Functional Realizing Lemma 1

By observing that (6.10) has (relative to $\mathsf{E\text{-}PA}^\omega$) the same logical form as the formula A in the meta-theorem 6.1[20] we are sure to find a functional $\Phi_8(f, n, \delta, r, h)$ such that, for all $f \in C[0,1]$ and $n \in \mathbb{N}$

$$\begin{cases} \forall p_1, p_2 \in K_{f,n}; x_1 \leq \ldots \leq x_n \in [0,1]; h \in P_n; \delta, r \in \mathbb{Q}^*_+ \\ \big(\forall y \in A(|f_0(y)| > \delta) \wedge \sum_{i=1}^{n+1} \sigma_i \int_{A_i} h > \int_B |h| + 1 \rightarrow \\ \exists \lambda \in \mathbb{R}(\|f_0 - \lambda h\|_1 + \Phi_8(f, n, \delta, r, h) < \|f_0\|_1)\big). \end{cases}$$

Claim 6.7 *The functional* $\boxed{\Phi_8(f, n, \delta, r, h) := \Phi_8(n, \delta, h) := \frac{\delta}{\|h\|_\infty}}$ *does the job.*

Proof. We have to prove that, for all $f \in C[0,1]$ and $n \in \mathbb{N}$

$$\begin{cases} \forall p_1, p_2 \in K_{f,n}; x_1 \leq \ldots \leq x_n \in [0,1]; h \in P_n; \delta, r \in \mathbb{Q}^*_+ \\ \big(\forall y \in A(|f_0(y)| > \delta) \wedge \sum_{i=1}^{n+1} \sigma_i \int_{A_i} h > \int_B |h| + 1 \rightarrow \\ \exists \lambda \in \mathbb{R}(\|f_0 - \lambda h\|_1 + \frac{\delta}{\|h\|_\infty} < \|f_0\|_1)\big). \end{cases}$$

Let $f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; h \in P_n; \delta, r \in \mathbb{Q}^*_+$ be fixed. Note that now we not only require f_0 not to have roots in A but not even δ-roots (i.e. $|f_0(y)| > \delta$). As a consequence y has to be 'r-apart' from all x_i. We say that y does not belong to the (x_i, r)-clusters[21]. Now we follow the original proof. Take n points, $x_1, \ldots, x_n$, such that (i) $0 = x_0 \leq x_1 \leq \ldots \leq x_{n+1} = 1$ and suppose that (ii) all δ-roots of f_0 belong to at least one of the (x_i, r)-clusters. Moreover, suppose that (iii) $\sum_{i=1}^{n+1} \sigma_i \int_{A_i} h > \int_B |h| + 1$, where $\sigma_i = sgn(f_0)(\frac{x_{i-1}+x_i}{2})$. By assumption (ii) we have $\sigma_i = sgn(f_0)(x)$, for $x \in A_i$ and then $\sum_{i=1}^{n+1} \sigma_i \int_{A_i} h(x)\, dx = \int_A (h(x)\, sgn(f_0)(x))\, dx$. By (ii) we have $f_0(x) > \delta$ for all $x \in A$. Therefore, taking $\lambda := \frac{\delta}{\|h\|_\infty}$ we have (iv) $sgn(f_0 - \lambda h)(x) = sgn(f_0)(x)$, for $x \in A$. Hence,

$$\begin{aligned} \|f_0 - \lambda h\|_1 &= \int_A |f_0 - \lambda h| + \int_B |f_0 - \lambda h| \\ &\overset{(iv)}{=} \int_A (f_0 - \lambda h)\, sgn(f_0) + \int_B |f_0 - \lambda h| \\ &= \int_A f_0\, sgn(f_0) - \lambda \int_A h\, sgn(f_0) + \int_B |f_0 - \lambda h| \\ &\leq \int_A f_0\, sgn(f_0) - \lambda \int_A h\, sgn(f_0) + \int_B |f_0| + \lambda \int_B |h| \\ &= \int_A |f_0| + \int_B |f_0| + \lambda \int_B |h| - \lambda \int_A h\, sgn(f_0) \\ &= \int_0^1 |f_0| + \lambda \int_B |h| - \lambda \int_A h\, sgn(f_0). \end{aligned}$$

[20] Note that we can treat σ_i as $\forall \sigma_1, \ldots, \sigma_{n+1} \in \{-1, 1\}$ with the purely universal assumption $\bigwedge_{i=1}^{n+1} (\sigma_i = 1 \rightarrow sgn(f_0)(\frac{x_i - x_{i-1}}{2}) \geq 0 \wedge \sigma_i = -1 \rightarrow sgn(f_0)(\frac{x_i - x_{i-1}}{2}) \leq 0)$,
since the case where $sgn(f_0)(\frac{x_i - x_{i-1}}{2}) = 0$ does not matter.

[21] This is fundamental to the elimination of the WKL, as mentioned in Section 6.3.1. We discuss this point in more details in Section 6.3.10.

Now we can add $\frac{\delta}{\|h\|_\infty}$ on both sides of the inequality and put $\lambda = \frac{\delta}{\|h\|_\infty}$ in evidence to get,

$$\begin{aligned}\|f_0 - \lambda h\|_1 + \frac{\delta}{\|h\|_\infty} &\leq \|f_0\|_1 + \frac{\delta}{\|h\|_\infty}(1 + \int_B |h| - \int_A h\, sgn(f_0)) \\ &\overset{(iii)}{<} \|f_0\|_1. \quad \square\end{aligned}$$

Remark 6.7 *In order to be precise we should have written* $\max\{1, \|h\|_\infty\}$ *instead of* $\|h\|_\infty$ *in the definition of* Φ_8*, so that it is always defined. This can be seen to be not necessary because we only apply these functionals to an* h *with uniform norm different from zero (see Section 6.3.12). Moreover, the functional* Φ_8 *should range over* $\mathbb{Q}^*_+$*, but* $\|h\|_\infty \in \mathbb{R}_+$*. Therefore, we should have also written* $\|h\|_{\infty,\mathbb{Q}}$ *instead of* $\|h\|_\infty$ *in the definition of* Φ_8*, where* $\|h\|_{\infty,\mathbb{Q}}$ *is a rational upper bound on* $\|h\|_\infty$*.*

Remark 6.8 *As it turned out the functional* Φ_8 *can be given independently of* r*. This independency can be explained by fact that (as we will see in Section 6.3.11)* r *is taken to be a function of* $\|h\|_\infty$*, and such dependency already appears in* Φ_8*.*

Elimination of WKL

As we discussed already in the introduction, the logical method of monotone functional interpretation upon which the proof of the general logical metatheorem is based not only provides an algorithm for the extraction of the modulus of uniqueness Φ but also a constructive verification of Φ which can be formalized in intuitionistic arithmetic in all finite types HA^ω. In particular, we get from this that Jackson's theorem is provable in HA^ω despite the fact that Cheney's proof heavily relies on classical logic and the non-computational binary König's lemma WKL. We will not carry out the details of this intuitionistic **verification** since we focus in this paper on the **applied** aspect of **constructing** Φ, which is, as a special feature of monotone functional interpretation, largely independent from the "constructivization" part. However, in 6.3.10 above we can see already how the constructivisation of Cheney's proof comes out of our analysis: as said before, WKL is used in the equivalent (see [153])[22] form of

$$\forall f \in C[0,1] \forall a, b \in [0,1](a < b \to \exists x_0 \in [a,b](f(x_0) = \inf_{x_0 \in [a,b]} f(x))) \tag{6.11}$$

to conclude

$$\forall x \in [x_{i-1} + r, x_i - r](f(x) > 0) \to \inf_{x \in [x_{i-1}+r, x_i - r]} f(x) > 0.$$

After our replacement of 'roots x_i' by 'r-clusters of δ-roots' this transforms into

$$\forall x \in [x_{i-1} + r, x_i - r](f(x) > \delta) \to \inf_{x \in [x_{i-1}+r, x_i - r]} f(x) \geq \delta$$

which follows from the constructively valid 'ε-weakening'

[22] Note that $f \in C[0,1]$ is given together with a modulus of uniform continuity ω_f.

$$\begin{cases} \forall f \in C[0,1] \forall a, b \in [0,1] \\ (a < b \to \forall \varepsilon > 0 \exists x_0 \in [a,b](f(x_0) - \inf_{x_0 \in [a,b]} f(x) < \varepsilon)) \end{cases}$$

version of (6.11) which eliminates the use of WKL. Also the use of classical logic to find σ_i such that

$$\sigma_i =_0 0 \leftrightarrow f(\tfrac{x_{i-1}+x_i}{2}) \geq_{\mathbb{R}} 0$$

is no longer necessary since we now have that

$$f(\tfrac{x_{i-1}+x_i}{2}) \geq_{\mathbb{R}} \delta \vee f(\tfrac{x_{i-1}+x_i}{2}) \leq_{\mathbb{R}} -\delta$$

which can easily be decided since $\delta \in \mathbb{Q}^*_+$.

6.3.11 Lemma $\forall \overline{x}, \overline{\sigma} \exists h\ I(\overline{x}, \overline{\sigma}, h)$

In the second part of Cheney's proof he considers the case where f_0 has less than $n+1$ roots, from this assumption he arrives at a contradiction (using Lemma 1) when assuming that for any $h \in P_n$, $\int h\ sgn(f_0) = 0$. We have indicated above that a contradiction is also obtained by assuming $\exists r \in \mathbb{Q}^*_+ (\int_A h\ sgn(f) > \int_B |h|)$. Here we show that for any given n points $x_1 \leq \ldots \leq x_n$ in the interval $[0,1]$ and for any $\sigma_1, \ldots, \sigma_{n+1} \in \{-1, 1\}$ (where σ_i will denote the sign of the function f_0 in the interval A_i) it is possible to find a function $h \in P_n$ and $r \in \mathbb{Q}^*_+$ such that $\sum_{i=1}^{n+1} \sigma_i \int_{A_i} h > \int_B |h|$, where $x_0 = 0$ and $x_{n+1} = 1$. Formally,

$$\begin{cases} \forall n \in \mathbb{N}; x_1 \leq \ldots \leq x_n \in [0,1]; \sigma_1, \ldots, \sigma_{n+1} \in \{-1,1\} \exists h \in P_n; r \in \mathbb{Q}^*_+ \\ \big(\sum_{i=1}^{n+1} \sigma_i \int_{A_i} h > \int_B |h|\big). \end{cases}$$

In the same way as we did in Section 6.3.10 we present the hidden quantifier η in the inequality and since $h/\eta \in P_n$ we have,

$$\begin{cases} \forall n \in \mathbb{N}; x_1 \leq \ldots \leq x_n \in [0,1]; \sigma_1, \ldots, \sigma_{n+1} \in \{-1,1\} \exists h \in P_n; r \in \mathbb{Q}^*_+ \\ \big(\sum_{i=1}^{n+1} \sigma_i \int_{A_i} h > \int_B |h| + 1\big). \end{cases}$$

The sentence above states the existence of an $r \in \mathbb{Q}^*_+$ and a function $h \in P_n$. Therefore, there exists also a $k \in \mathbb{Q}^*_+$ such that $k \geq \|h\|_\infty$. Here we can again apply our meta-theorem 6.1 and we are sure to find functions Φ_9 and Φ_{10} depending only on n such that,[23]

$$\begin{cases} \forall n \in \mathbb{N}; x_1 \leq \ldots \leq x_n \in [0,1]; \sigma_1, \ldots, \sigma_{n+1} \in \{-1,1\} \exists h \in P_n; r \geq \Phi_9(n) \\ \big(\sum_{i=1}^{n+1} \sigma_i \int_{A_i} h > \int_B |h| + 1 \wedge \Phi_{10}(n) \geq \|h\|_\infty\big), \end{cases}$$

where A and B are defined as before.

Claim 6.8 *The functions* $\boxed{\Phi_9(n) := \frac{1}{16(n+1)^3}}$ *and* $\boxed{\Phi_{10}(n) := 8(n+1)^2}$ *do the job.*

[23]Note that Φ_9 and Φ_{10} do not depend on the points $x_1, \ldots, x_n$ nor on $\sigma_1, \ldots, \sigma_{n+1}$ since they are elements from the compact spaces $[0,1]$ and $\{-1,1\}$, respectively, and $\bigwedge_{i=1}^{n-1} x_i \leq x_{i+1}$ is purely universal.

Proof. Let $0 = x_0 \leq x_1 \leq \ldots \leq x_{n+1} = 1$ and $\sigma_1, \ldots, \sigma_{n+1} \in \{-1, 1\}$ be given. Ignore all the points x_j such that $x_i = x_j$ and $i < j$. We are left with $\tilde{n}+1$ points $0 = x_{a_0} < x_{a_1} < \ldots < x_{a_{\tilde{n}+1}} = 1$ where $a_{i-1} < a_i$, $a_i \in \{0, \ldots, n+1\}$ and $\tilde{n} \leq n$. Let $\tilde{x}_i := x_{a_i}$ and $\tilde{\sigma}_i := \sigma_{a_i}$. Since we have eliminated just empty intervals we have for any function $h \in P_n$, $\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x)\, dx = \sum_{i=1}^{\tilde{n}+1} \tilde{\sigma}_i \int_{\tilde{x}_{i-1}}^{\tilde{x}_i} h(x)\, dx$. Among the points $\tilde{x}_1, \ldots, \tilde{x}_{\tilde{n}}$ pick only the points $\tilde{x}_i$ for which $\tilde{\sigma}_i \neq \tilde{\sigma}_{i+1}$. Finally, we are left with $m+1$ points $0 = \tilde{x}_{b_0} < \tilde{x}_{b_1} < \ldots < \tilde{x}_{b_{m+1}} = 1$ where $b_{i-1} < b_i$, $b_i \in \{0, \ldots, \tilde{n}+1\}$ and $m \leq \tilde{n}$. Let $y_i := \tilde{x}_{b_i}$ and $\sigma_i^* := \tilde{\sigma}_{b_i}$. Again we have $\sum_{i=1}^{\tilde{n}+1} \tilde{\sigma}_i \int_{\tilde{x}_{i-1}}^{\tilde{x}_i} h(x)dx = \sum_{i=1}^{m+1} \sigma_i^* \int_{y_{i-1}}^{y_i} h(x)dx$, for any $h \in P_n$. Then we define $\tilde{h}(x) := (x - y_1) \ldots (x - y_m)$ and

$$h(x) := \tfrac{+/-8(n+1)^2}{\|\tilde{h}\|_\infty} \tilde{h}(x).$$

Choose $+/-$ so that $\sum_{i=1}^{m+1} \sigma_i^* \int_{y_{i-1}}^{y_i} h(x)dx = \sum_{i=1}^{m+1} \int_{y_{i-1}}^{y_i} |h(x)|dx$. Hence,

$$\textstyle\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x)\, dx = \int_0^1 |h(x)|\, dx.$$

Moreover, it is clear from the definition of h that $\|h\|_\infty = 8(n+1)^2$. Therefore, from Remark 6.2 (cf. also Section 6.3.5) we get

$$\textstyle\int_0^1 |h(x)|\, dx = \|h\|_1 \geq \frac{\|h\|_\infty}{2(n+1)^2} = 4.$$

Let $r := \Phi_9(n)$. It is clear that the intervals B as a whole (as defined above) have length at most $\frac{1}{8(n+1)^2}$. Therefore, $\int_B |h(x)|\, dx \leq 1$. Hence,

$$\textstyle\sum_{i=1}^{n+1} \sigma_i \int_{A_i} h(x)\, dx = \int_A |h(x)|\, dx > \int_B |h(x)|\, dx + 1. \quad \Box$$

Remark 6.9 *Note that (as follows from the result above) we can even allow σ_i to range over $\{-1, 0, 1\}$ as long as $\sigma_i = 0$ only when $x_i - x_{i-1} \leq 2\Phi_9(n)$. In such cases the value of σ_i has no influence on the result.*

6.3.12 Eliminating the Polynomial h in Lemma 1

We have just shown that,

$$\left\{ \begin{array}{l} \forall x_1 \leq \ldots \leq x_n \in [0,1]; \sigma_1, \ldots, \sigma_{n+1} \in \{-1,1\} \exists h \in P_n \\ \big(\sum_{i=1}^{n+1} \sigma_i \int_{A_i} h > \int_B |h| + 1 \wedge \Phi_{10}(n) \geq \|h\|_\infty \big), \end{array} \right. \tag{6.12}$$

where A_i and B are defined with r replaced by $\Phi_9(n)$. We can take $r = \Phi_9(n)$ because h is taken (cf. proof of Claim 6.8) in such way that $\sum_i \sigma_i \int_{A_i} h = \int_A |h|$ which makes the matrix of the lemma monotone on $\exists r$.

Let $f \in C[0,1], n \in \mathbb{N}, p_1, p_2 \in K_{f,n}$ and $x_1 \leq \ldots \leq x_n \in [0,1]$ be fixed, and let $\tilde{h}$ be the function from (6.12) when $\sigma_i := f_0(\frac{x_{i-1}+x_i}{2})$, where $x_0 := 0$ and $x_{n+1} := 1$. Note that here σ_i can be 0 (cf. Remark 6.9). Applying Lemma 1 to $\tilde{h}$ and $\Phi_9(n)$ (i.e. taking $h = \tilde{h}$ and $r = \Phi_9(n)$) we get,

$$\left\{ \begin{array}{l} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; x_1 \leq \ldots \leq x_n \in [0,1]; \delta \in \mathbb{Q}_+^* \\ \big(\forall \lambda \in \mathbb{R} (\|f_0 - \lambda \tilde{h}\|_1 + \Phi_8(n, \delta, \tilde{h}) \geq \|f_0\|_1) \to \exists y \in A (|f_0(y)| \leq \delta) \big). \end{array} \right.$$

Having in mind that we have $\|\tilde{h}\|_\infty \le 8(n+1)^2$ we take $\boxed{\tilde{\Phi_8}(n,\delta) := \frac{\delta}{8(n+1)^2}}$. By the monotonicity of the functional Φ_8 in $\|h\|_\infty$ we have $\tilde{\Phi_8}(n,\delta) \le \Phi_8(n,\delta,\tilde{h})$. Then,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; x_1 \le \ldots \le x_n \in [0,1]; \delta \in \mathbb{Q}_+^* \\ \big(\forall \lambda \in \mathbb{R}(\|f_0 - \lambda\tilde{h}\|_1 + \tilde{\Phi_8}(n,\delta) \ge \|f_0\|_1) \to \exists y \in A(|f_0(y)| \le \delta)\big). \end{cases}$$

We can then conclude,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \delta \in \mathbb{Q}_+^* \\ \big(\forall h \in P_n(\|f_0 - h\|_1 + \tilde{\Phi_8}(n,\delta) \ge \|f_0\|_1) \to \\ \forall x_1 \le \ldots \le x_n \in [0,1] \exists y \in A(|f_0(y)| < \delta)\big). \end{cases}$$

We can actually replace the conclusion of the implication above with the actual existence of $n+1$ roots in the following way (lemma $\forall \overline{x} J(\overline{x}) \to E$). Assume

$$\forall x_1 \le \ldots \le x_n \in [0,1] \exists y \in [0,1]\big(|f_0(y)| < \delta \wedge \bigwedge_{i=0}^{n+1} |x_i - y| \ge \Phi_9(n)\big)\big). \tag{6.13}$$

If $m < n+1$ is the biggest number of δ-roots of f_0 which are pairwise apart from each other by at least $\Phi_9(n)$ then by (6.13) we have a contradiction. Hence,

$$\exists x_0, \ldots, x_n \in [0,1](\textstyle\bigwedge_{i=0}^n |f_0(x_i)| < \delta \wedge \bigwedge_{i=1}^n (x_{i-1} + \Phi_9(n) \le x_i)).$$

Therefore, we have,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \delta \in \mathbb{Q}_+^* \\ \big(\forall h \in P_n(\|f_0 - h\|_1 + \tilde{\Phi_8}(n,\delta) \ge \|f_0\|_1) \to \\ \exists x_0, \ldots, x_n \in [0,1](\bigwedge_{i=0}^n |f_0(x_i)| < \delta \wedge \bigwedge_{i=1}^n x_{i-1} + \Phi_9(n) \le x_i)\big). \end{cases}$$

6.4 The Uniform Modulus of Uniqueness for L_1-approximation

In this section we show how the computed functionals are combined in order to obtain the uniform modulus of uniqueness. Let $f \in C[0,1]$, $n \in \mathbb{N}, p_1, p_2 \in K_{f,n}$ and $\varepsilon \in \mathbb{Q}_+^*$ be fixed. Assume (for $i \in \{1,2\}$),

$$\begin{cases} \|f - p_i\|_1 - dist_1(f, P_n) < \\ \min\{\Phi_1(\Phi_7(\tilde{\Phi_8}(n, \Phi_5(\Phi_6(n, \Phi_9(n), \varepsilon))))), \\ \Phi_1(\Phi_2(\Phi_3(\omega_{f,n}, \Phi_4(\Phi_6(n, \Phi_9(n), \varepsilon)))))\}. \end{cases} \tag{6.14}$$

By Section 6.3.3 we have, (where $f_0(x) = f(x) - \frac{p_1(x)+p_2(x)}{2}$)

$$\|f_0\|_1 - dist_1(f, P_n) < \Phi_2(\Phi_3(\omega_{f,n}, \Phi_4(\Phi_6(n, \Phi_9(n), \varepsilon)))).$$

By Section 6.3.4 (and since Φ_1 is the identity),

$$|\ \|f_0\|_1 - 1/2\|f - p_1\|_1 - 1/2\|f - p_2\|_1\ | < \Phi_3(\omega_{f,n}, \Phi_4(\Phi_6(n, \Phi_9(n), \varepsilon))).$$

By Section 6.3.6[24],

[24] Since $\int |f_0| - \frac{1}{2}|f - p_1| - \frac{1}{2}|f - p_2| = \|f_0\|_1 - \frac{1}{2}\|f - p_1\|_1 - \frac{1}{2}\|f - p_2\|_1$.

$$\| \, |f_0| - 1/2|f - p_1| - 1/2|f - p_2| \, \|_\infty \le \Phi_4(\Phi_6(n, \Phi_9(n), \varepsilon)).$$

Hence, by Section 6.3.7,

$$\begin{cases} \forall x \in [0,1] \\ (|f_0(x)| \le \Phi_5(\Phi_6(n, \Phi_9(n), \varepsilon)) \to \\ |p_1(x) - p_2(x)| \le \Phi_6(n, \Phi_9(n), \varepsilon)). \end{cases} \tag{6.15}$$

By the same assumption (6.14) and Section 6.3.3 we also have,

$$\|f_0\|_1 - dist_1(f, P_n) < \Phi_7(\tilde{\Phi_8}(n, \Phi_5(\Phi_6(n, \Phi_9(n), \varepsilon)))).$$

And by Section 6.3.9,

$$\forall h \in P_n \big(\|f_0 - h\|_1 + \tilde{\Phi_8}(n, \Phi_5(\Phi_6(n, \Phi_9(n), \varepsilon))) \ge \|f_0\|_1\big).$$

Hence, by Section 6.3.12 (taking $\delta = \Phi_5(\Phi_6(n, \Phi_9(n), \varepsilon))$),

$$\begin{cases} \exists x_0, \ldots, x_n \in [0,1] \\ (\bigwedge_{i=0}^n |f_0(x_i)| < \Phi_5(\Phi_6(n, \Phi_9(n), \varepsilon)) \wedge \bigwedge_{i=1}^n x_{i-1} + \Phi_9(n) \le x_i). \end{cases}$$

And by (6.15),

$$\begin{cases} \exists x_0, \ldots, x_n \in [0,1] \\ (\bigwedge_{i=0}^n |p_1(x_i) - p_2(x_i)| \le \Phi_6(n, \Phi_9(n), \varepsilon) \wedge \bigwedge_{i=1}^n x_{i-1} + \Phi_9(n) \le x_i). \end{cases}$$

Therefore, by Section 6.3.8 (taking $r = \Phi_9(n)$) we conclude,

$$\|p_1 - p_2\|_\infty \le \varepsilon. \tag{6.16}$$

If we substitute the linear functionals, $\Phi_1, \Phi_2, \Phi_4, \Phi_5$ and Φ_7, to make the conclusion more legible, we have (6.14) $\to$ (6.16),

$$\begin{cases} \|f - p_i\|_1 - dist_1(f, P_n) < \\ \min\{\tilde{\Phi_8}(n, \frac{\Phi_6(n,\Phi_9(n),\varepsilon)}{8}), \Phi_3(\omega_{f,n}, \frac{\Phi_6(n,\Phi_9(n),\varepsilon)}{8})\} \to \|p_1 - p_2\|_\infty \le \varepsilon. \end{cases}$$

After applying $\tilde{\Phi_8}$ and Φ_9 we get,

$$\begin{cases} \|f - p_i\|_1 - dist_1(f, P_n) < \\ \min\{\frac{\Phi_6(n, \frac{1}{16(n+1)^3}, \varepsilon)}{64(n+1)^2}, \Phi_3(\omega_{f,n}, \frac{\Phi_6(n, \frac{1}{16(n+1)^3}, \varepsilon)}{8})\} \to \|p_1 - p_2\|_\infty \le \varepsilon. \end{cases}$$

Then we apply Φ_6,

$$\begin{cases} \|f - p_i\|_1 - dist_1(f, P_n) < \\ \min\{\frac{\frac{\lfloor n/2 \rfloor ! \lceil n/2 \rceil !}{2^{4n+3}(n+1)^{3n+1}} \varepsilon}{8(n+1)^2}, \Phi_3(\omega_{f,n}, \frac{\lfloor n/2 \rfloor ! \lceil n/2 \rceil !}{2^{4n+3}(n+1)^{3n+1}} \varepsilon)\} \to \|p_1 - p_2\|_\infty \le \varepsilon. \end{cases}$$

Let $c_n := \frac{\lfloor n/2 \rfloor ! \lceil n/2 \rceil !}{2^{4n+3}(n+1)^{3n+1}}$ then we can rewrite the above formula as,

$$\begin{cases} \|f - p_i\|_1 - dist_1(f, P_n) < \\ \min\{\frac{c_n \varepsilon}{8(n+1)^2}, \Phi_3(\omega_{f,n}, c_n \varepsilon)\} \to \|p_1 - p_2\|_\infty \le \varepsilon. \end{cases}$$

And finally we apply the definition of Φ_3,

$$\begin{cases} \|f-p_i\|_1 - dist_1(f,P_n) < \\ \min\{\frac{c_n\varepsilon}{8(n+1)^2}, \frac{c_n\varepsilon}{2}\omega_{f,n}(\frac{c_n\varepsilon}{2})\} \to \|p_1-p_2\|_\infty \le \varepsilon. \end{cases}$$

Let $\tilde{\Phi}(f,n,\varepsilon) := \min\{\frac{c_n\varepsilon}{8(n+1)^2}, \frac{c_n\varepsilon}{2}\omega_{f,n}(\frac{c_n\varepsilon}{2})\}$, where

$$\omega_{f,n} := \min\{\omega_f(\tfrac{\varepsilon}{4}), \tfrac{\varepsilon}{40(n+1)^4 M_f}\}$$

and M_f is a bound on $\|f\|_\infty$. We have shown that,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n} \\ \big(\bigwedge_{i=1}^{2}(\|f-p_i\|_1 - dist_1(f,P_n) < \tilde{\Phi}(f,n,\varepsilon)) \to \|p_1-p_2\|_\infty \le \varepsilon\big). \end{cases}$$

Proposition 6.1 *The functional $\tilde{\Phi}(f,n,\varepsilon)$ is a uniform modulus of uniqueness for the best L_1-approximation of $C[0,1]$ from $K_{f,n}$, i.e.*

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}_+^* \\ \big(\bigwedge_{i=1}^{2}(\|f-p_i\|_1 - dist_1(f,P_n) < \tilde{\Phi}(f,n,\varepsilon)) \to \|p_1-p_2\|_1 \le \varepsilon\big). \end{cases}$$

Proof. Above, plus the fact that $\|p_1-p_2\|_1 \le \|p_1-p_2\|_\infty$. □

Claim 6.9 $\tilde{\Phi}(f,n,\varepsilon) \le \frac{\varepsilon}{8}$

Proof. Trivial. □

Now we show that Proposition 6.1 can be generalised to the whole space P_n (i.e. we can replace $K_{f,n}$ with P_n). Moreover, we notice that the dependency on particular values of the function f can be eliminated so that the modulus of uniqueness depends on f only through its modulus of continuity.

Theorem 6.2 *Let $\Phi(\omega,n,\varepsilon) := \min\{\frac{c_n\varepsilon}{8(n+1)^2}, \frac{c_n\varepsilon}{2}\omega_n(\frac{c_n\varepsilon}{2})\}$, where the constant $c_n := \frac{\lfloor n/2\rfloor!\lceil n/2\rceil!}{2^{4n+3}(n+1)^{3n+1}}$ and $\omega_n(\varepsilon) := \min\{\omega(\frac{\varepsilon}{4}), \frac{\varepsilon}{40(n+1)^4\lceil\frac{1}{\omega(1)}\rceil}\}$. For all $f \in C[0,1]$ with modulus of continuity ω*

$$\begin{cases} \forall n \in \mathbb{N}; p_1, p_2 \in P_n; \varepsilon \in \mathbb{Q}_+^* \\ \big(\bigwedge_{i=1}^{2}(\|f-p_i\|_1 - dist_1(f,P_n) < \Phi(\omega,n,\varepsilon)) \to \|p_1-p_2\|_1 \le \varepsilon\big). \end{cases}$$

Proof. Actually, we prove the stronger version of the theorem where instead of $\lceil\frac{1}{\omega(1)}\rceil$ in the definition of ω_n we have any upper bound on $\sup_{x\in[0,1]}|f(x)-f(0)|$. First we show that in Proposition 6.1 we can replace $K_{f,n}$ with P_n. Suppose without loss of generality that $p_1 \in P_n\backslash K_{f,n}$. Then $\|p_1\|_1 > \frac{5}{2}\|f\|_1$ and hence $\|f-p_1\|_1 > \frac{3}{2}\|f\|_1 \ge \frac{3}{2}dist_1(f,P_n)$. Assume that $\|f-p_i\|_1 < dist_1(f,P_n) + \tilde{\Phi}(f,n,\varepsilon)$. By Claim 6.9, $\|f-p_i\|_1 < dist_1(f,P_n) + \frac{\varepsilon}{8}$. Then, $\frac{\varepsilon}{8} > \frac{1}{2}dist_1(f,P_n)$, i.e. $dist_1(f,P_n) < \frac{\varepsilon}{4}$. Therefore $\|f-p_i\|_1 < dist_1(f,P_n) + \frac{\varepsilon}{8} < \frac{\varepsilon}{2}$ and we have $\|p_1-p_2\|_1 \le \varepsilon$. The second point is that some upper bound $M_f \ge \|f\|_\infty$ is used to define $\omega_{f,n}$ in Proposition 6.1. We claim that an upper bound $N_f \ge \sup_{x\in[0,1]}|f(x)-f(0)|$ is sufficient. For any function $f \in C[0,1]$ and polynomials $p_1, p_2 \in P_n$ let $\tilde{f}$, $\tilde{p_1}$ and $\tilde{p_2}$ be the functions obtained by the transposition of f, p_1 and p_2 respectively by $f(0)$ (i.e. $\tilde{f}(x) := f(x) - f(0)$ and $\tilde{p_i}(x) := p_i(x) - f(0)$). It is clear that (*i*) $\|f - p_i\|_1 = \|\tilde{f} - \tilde{p_i}\|_1$,

(*ii*) $dist_1(f, P_n) = dist_1(\tilde{f}, P_n)$ and

(*iii*) $\|p_1 - p_2\|_1 = \|\tilde{p_1} - \tilde{p_2}\|_1$.

Let ω be the modulus of continuity for f and assume $\|f - p_i\|_1 < dist(f, P_n) + \Phi(\omega, n, \varepsilon)$. By (*i*) and (*ii*) we have, $\|\tilde{f} - \tilde{p_i}\|_1 < dist(\tilde{f}, P_n) + \Phi(\omega, n, \varepsilon)$. Since ω is also a modulus of continuity for $\tilde{f}$ and $\|\tilde{f}\|_\infty = \sup_{x\in[0,1]} |f(x) - f(0)| \le N_f$ we have $\tilde{\Phi}(\tilde{f}, n, \varepsilon) = \Phi(\omega, n, \varepsilon)$, therefore,

$$\|\tilde{f} - \tilde{p_i}\|_1 < dist(\tilde{f}, P_n) + \tilde{\Phi}(\tilde{f}, n, \varepsilon),$$

which implies, by Proposition 6.1, the first part of this proof and (*iii*), $\|p_1 - p_2\|_1 \le \varepsilon$. Since $\lceil \frac{1}{\omega(1)} \rceil \ge \sup_{x\in[0,1]} |f(x) - f(0)|$ if ω is a modulus of uniform continuity for f the theorem follows. $\Box$

As mentioned in Remark 6.5, the function $\Psi(n) := \frac{n!}{2^{n+1}(n+1)^{2n+2}}$ relates the L_1-norm of a polynomial $p \in P_n$ to its actual coefficients, i.e.

$$\forall n \in \mathbb{N} \forall p \in P_n \big(\|p\|_1 \le \Psi(n) \cdot \varepsilon \to \|p\|_{\max} \le \varepsilon\big),$$

where $\|p\|_{\max}$ denotes the maximum absolute value of the coefficients of p. Therefore, we obtain the following corollary.

Corollary 6.2 *Let $\Phi(\omega, n, \varepsilon)$ be as defined above. For all $f \in C[0,1]$ with modulus of continuity ω*

$$\begin{cases} \forall n \in \mathbb{N}; p_1, p_2 \in P_n; \varepsilon \in \mathbb{Q}^*_+ \\ \big(\bigwedge_{i=1}^{2} (\|f - p_i\|_1 - dist_1(f, P_n) < \Phi(\omega, n, \Psi(n) \cdot \varepsilon)) \to \|p_1 - p_2\|_{\max} \le \varepsilon\big). \end{cases}$$

A function $f \in C[0,1]$ is said to *be Lipschitz continuous with Lipschitz constant* $\lambda \in \mathbb{R}^*_+$ if $|f(x) - f(y)| \le \lambda|x - y|$ (i.e. $\frac{\varepsilon}{\lambda}$ is a modulus of continuity for f) and is *Lipschitz-α continuous with constant* λ, $0 < \alpha \le 1$, if $|f(x) - f(y)| \le \lambda|x - y|^\alpha$ (equivalently, $(\frac{\varepsilon}{\lambda})^{1/\alpha}$ is a modulus of continuity in our sense for f)[25]. In this way, if a function f is Lipschitz continuous (or Lipschitz-α continuous) with constant λ then $\sup_{x\in[0,1]} |f(x) - f(0)| \le \lambda$ (and we can take λ instead of $\lceil \frac{1}{\omega(1)} \rceil$ in Theorem 6.2).

Corollary 6.3 *For any $f \in C[0,1]$,*

i) let $\Phi_L(\lambda, n, \varepsilon) := \min\{\frac{c_n\varepsilon}{8(n+1)^2}, \frac{c_n^2\varepsilon^2}{160(n+1)^4\lambda}\}$. If f is Lipschitz continuous with constant λ then the functional Φ_L is a modulus of uniqueness for f.

ii) let $\Phi_{L_\alpha}(\lambda, \alpha, n, \varepsilon) := \min\{\frac{c_n\varepsilon}{8(n+1)^2}, \frac{c_n\varepsilon}{2}(\frac{c_n\varepsilon}{8\lambda})^{1/\alpha}, \frac{c_n^2\varepsilon^2}{160(n+1)^4\lambda}\}$. If f is Lipschitz-α continuous with constant λ then the functional Φ_{L_α} is a modulus of uniqueness for f.

[25]In analysis the condition '$|f(x) - f(y)| \le \lambda|x - y|^\alpha$ for some λ' is called *Hölder condition with exponent α*.

And as a corollary of Proposition 5.4 from [90] and Theorem 6.2 above we get,

Theorem 6.3 *Let $\mathcal{P}(f,n)$ denote the operator which assigns to any given function $f \in C[0,1]$ and any $n \in \mathbb{N}$ the best L_1-approximation of $f \in C[0,1]$ from P_n. Then $\Phi_P(\omega_f, n, \varepsilon) := \frac{\Phi(\omega_f,n,\varepsilon)}{2}$, Φ as defined in Theorem 6.2, is a modulus of pointwise continuity for the operator $\mathcal{P}(f,n)$, i.e.,*

$$\begin{cases} \forall f, \tilde{f} \in C[0,1]; n \in \mathbb{N}; \varepsilon \in \mathbb{Q}^*_+ \\ (\|f - \tilde{f}\|_1 < \Phi_P(\omega_f, n, \varepsilon) \to \|\mathcal{P}(f,n) - \mathcal{P}(\tilde{f},n)\|_1 \le \varepsilon). \end{cases}$$

Proof. For completeness we reproduce here the proof as given in [90]. One easily verifies that $dist_1(f, P_n)$ is Lipschitz continuous in f (with respect to the L_1-norm) with $\lambda = 1$, i.e.

$$\|f - \tilde{f}\|_1 < \varepsilon \to |dist_1(f, P_n) - dist_1(\tilde{f}, P_n)| < \varepsilon. \tag{6.17}$$

Assume now that $\|f - \tilde{f}\|_1 < \Phi_P(\omega_f, n, \varepsilon) = \frac{1}{2}\Phi(\omega_f, n, \varepsilon)$. Then,

$$\begin{aligned} \|f - \mathcal{P}(\tilde{f},n)\|_1 &\le \|\tilde{f} - \mathcal{P}(\tilde{f},n)\|_1 + \|f - \tilde{f}\|_1 = dist_1(\tilde{f}, P_n) + \|f - \tilde{f}\|_1 \\ &\overset{(6.17)}{<} dist_1(f, P_n) + \frac{1}{2}\Phi(\omega_f, n, \varepsilon) + \|f - \tilde{f}\|_1 \\ &< dist_1(f, P_n) + \Phi(\omega_f, n, \varepsilon). \end{aligned}$$

Since, furthermore, $\|f - \mathcal{P}(f,n)\|_1 = dist_1(f, P_n)$, we obtain from Theorem 6.2 that $\|\mathcal{P}(f,n) - \mathcal{P}(\tilde{f},n)\|_1 \le \varepsilon$. $\square$

6.5 Computing the Sequence $(p_n)_{n\in\mathbb{N}}$

An operator $B_{f,n} : \mathbb{Q}^*_+ \to P_n$ *computes the unique best L_1-approximation, $p_n \in P_n$, of a function $f \in C[0,1]$ (given with a modulus of uniform continuity ω_f) from P_n* if for any given $\varepsilon \in \mathbb{Q}^*_+$ it generates a polynomial of degree $\le n$ with rational coefficients (i.e. a $n+1$-vector of rational coefficients) $B_{f,n}(\varepsilon)$ such that, $\|B_{f,n}(\varepsilon) - p_n\|_1 \le \varepsilon$. We indicate how this can be achieved using the uniform modulus of uniqueness, $\Phi(\omega_f, n, \varepsilon)$,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in P_n; \varepsilon \in \mathbb{Q}^*_+ \\ \big(\bigwedge_{i=1}^2 (\|f - p_i\|_1 - dist_1(f, P_n) < \Phi(\omega_f, n, \varepsilon)) \to \|p_1 - p_2\|_1 \le \varepsilon\big). \end{cases}$$

First we substitute p for p_1 and (the best L_1-approximation of f from P_n) p_n for p_2,

$$\begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p \in P_n; \varepsilon \in \mathbb{Q}^*_+ \\ \big(\|f - p\|_1 - dist_1(f, P_n) < \Phi(\omega_f, n, \varepsilon) \to \|p - p_n\|_1 \le \varepsilon\big). \end{cases}$$

Now we just need to find a $B_{f,n}(\varepsilon)$ such that, $\|f - B_{f,n}(\varepsilon)\|_1 - dist_1(f, P_n) < \Phi(\omega_f, n, \varepsilon)$. Note that now there is no explicit reference to p_n, only implicit in $dist_1(f, P_n)$.

A set $N_\varepsilon := \{p_1, p_2, \ldots, p_{n_\varepsilon}\} \subset P_n$ is said to be an *ε-net of $K_{f,n}$* if $\forall p \in K_{f,n} \exists p_i \in N_\varepsilon(\|p - p_i\|_1 \le \varepsilon)$. The algorithm for computing p_n consists in evaluating $\|f - p_i\|_1$ for each p_i in some $\Phi(\omega_f, n, \varepsilon)$-net of $K_{f,n}$ and taking the p_i which gives the minimum value. Although the elements of the net N_ε are taken to be polynomials with rational coefficients, the value of $\|f - p_i\|_1$ will in general be a real number. Therefore, we only compute $\|f - p_i\|_1$ up to some precision. By an appropriate choice of the precision the minimum value returned by the search will in fact be close the the actual minimum.

The complexity analysis of the whole algorithm has been carried out in [134] and the following result was obtained.

Theorem 6.4 ([134]) *For polynomial time computable $f \in C[0,1]$ the sequence $(p_n)_{n\in\mathbb{N}}$ is strongly* **NP** *computable in* $\mathbf{NP}[B_f]$*, where B_f is an oracle deciding left cuts for integration.*

6.6 Related Results

The first proof of the uniqueness of the best L_1-approximation of $f \in C[0,1]$ by polynomials in P_n was given in 1921 by Jackson [70]. The proof we analysed was published by Cheney [33] in 1965 and reprinted in his book [34] from 1966. Only in 1975 Björnestål [22], by analyzing the qualitative (relative to the dependencies) aspect of the continuity of the projection operator for arbitrary normed linear spaces X into a closed linear subspace of X, obtained the following result.

Theorem 6.5 (Björnestål, 75) *Let $f \in C[0,1]$ and the modulus Ω_f be defined as*

$$\Omega_f(\varepsilon) := \sup_{|x-y|<\varepsilon} |f(x) - p_n(x) - f(y) + p_n(y)|,$$

where p_n is the best L_1-approximation of f from P_n. Then, for $p \in P_n$, ε sufficiently small and for some constant c depending on f and n,

$$\|p - p_n\|_1 \geq \varepsilon \to \|f - p\|_1 - \|f - p_n\|_1 \geq 2\int_0^{\Omega_f^{-1}(c\varepsilon)} c\varepsilon - \Omega_f(x)\, dx,$$

where $\Omega_f^{-1}(\varepsilon)$ is defined as[26]

$$\Omega_f^{-1}(\varepsilon) := \inf\{\delta : \Omega_f(\delta) = \varepsilon\}.$$

We show that our Theorem 6.2 implies an effective version of Björnestål's theorem. First we can rewrite his theorem in the form we have been working with,

$$\begin{cases} \|f - p\|_1 < dist_1(f, P_n) + 2\int_0^{\Omega_f^{-1}(c\varepsilon)} c\varepsilon - \Omega_f(x)\, dx \to \\ \|p - p_n\|_1 < \varepsilon. \end{cases} \tag{6.18}$$

First we show that $\int_0^{\Omega_f^{-1}(c\varepsilon)} c\varepsilon - \Omega_f(x)\, dx$ can be written as $c'\varepsilon\Omega_f^{-1}(c'\varepsilon)$, for some constant $\frac{c}{2} \leq c' \leq c$. For that purpose note that,

[26]Note that $\Omega_f^{-1}(\varepsilon)$ (for ε small enough so that $\Omega_f^{-1}(\varepsilon)$ is defined) is a special modulus of continuity for $f - p_n$ in our sense.

$$\int_0^{\Omega_f^{-1}(c\varepsilon)} c\varepsilon - \Omega_f(x)\, dx \le \int_0^{\Omega_f^{-1}(c\varepsilon)} c\varepsilon\, dx = c\varepsilon\Omega_f^{-1}(c\varepsilon).$$

On the other hand we have,

$$\begin{aligned}\int_0^{\Omega_f^{-1}(c\varepsilon)} c\varepsilon - \Omega_f(x)\, dx &\ge \int_0^{\Omega_f^{-1}(\frac{c}{2}\varepsilon)} c\varepsilon - \Omega_f(x)\, dx \\ &\ge \int_0^{\Omega_f^{-1}(\frac{c}{2}\varepsilon)} \frac{c}{2}\varepsilon\, dx = \frac{c}{2}\varepsilon\Omega_f^{-1}(\frac{c}{2}\varepsilon).\end{aligned}$$

Therefore, for some $\frac{c}{2} \le c' \le c$, (6.18) is equivalent to,

$$\|f-p\|_1 < dist_1(f, P_n) + 2c'\,\varepsilon\Omega_f^{-1}(c'\,\varepsilon) \to \|p-p_n\|_1 < \varepsilon.$$

The constant c, however, is not presented by Björnestål and moreover the function Ω_f^{-1} is usually non-computable. We can give an effective modulus of continuity for $f - p_n$ following Section 6.3.5 (and taking $M_f = \lceil \frac{1}{\omega_f(1)} \rceil$ as suggested in the proof of Theorem 6.2),

$$\begin{aligned}\omega_{f-p_n}(\varepsilon) &\ge \min\{\omega_f(\frac{\varepsilon}{2}), \omega_{p_n}(\frac{\varepsilon}{2})\} \\ &\ge \min\{\omega_f(\frac{\varepsilon}{2}), \frac{\varepsilon}{20(n+1)^4\lceil\frac{1}{\omega_f(1)}\rceil}\}\end{aligned}$$

Therefore, let $\omega^*_{f-p_n}(\varepsilon) := \min\{\omega_f(\frac{\varepsilon}{2}), \frac{\varepsilon}{20(n+1)^4\lceil\frac{1}{\omega_f(1)}\rceil}\}$, we can restate our Theorem 6.2 and see how it relates to Björnestål's result:

Corollary 6.4 *Let $f \in C[0,1]$, ω_f be some modulus of uniform continuity of f, and $p \in P_n$. Then for $\varepsilon \le 1$,*

$$\|f-p\|_1 < dist_1(f, P_n) + \tilde{c}_n\,\varepsilon\,\omega^*_{f-p_n}(\tilde{c}_n\,\varepsilon) \to \|p-p_n\|_1 \le \varepsilon,$$

where $\tilde{c}_n := \frac{c_n}{8(n+1)^2}$ and $c_n := \frac{\lfloor n/2\rfloor!\lceil n/2\rceil!}{2^{4n+3}(n+1)^{3n+1}}$.

Proof. From Theorem 6.2 we have,

$$\begin{cases} \|f-p\|_1 < dist_1(f, P_n) + \min\{\frac{c_n\varepsilon}{8(n+1)^2}, \frac{c_n\varepsilon}{2}\omega^*_{f-p_n}(\frac{c_n\varepsilon}{4})\} \to \\ \|p-p_n\|_1 \le \varepsilon, \end{cases}$$

which implies,

$$\begin{cases} \|f-p\|_1 < dist_1(f, P_n) + \min\{\frac{c_n\varepsilon}{8(n+1)^2}, \frac{c_n\varepsilon}{8(n+1)^2}\omega^*_{f-p_n}(\frac{c_n\varepsilon}{4})\} \to \\ \|p-p_n\|_1 \le \varepsilon. \end{cases}$$

For $\varepsilon \le 1$ we have $\omega^*_{f-p_n}(\frac{c_n\varepsilon}{4}) \le 1$. Hence,

$$\|f-p\|_1 < dist_1(f, P_n) + \frac{c_n\varepsilon}{8(n+1)^2}\omega^*_{f-p_n}(\frac{c_n\varepsilon}{4}) \to \|p-p_n\|_1 \le \varepsilon.$$

Since $8(n+1)^2 > 4$ we get our result. $\Box$

Some years later, in 1978, Kroó [118] showed that the constant c in Björnestål's result needed not to depend on any particular point of the function f but only on its modulus of continuity[27]. We got an effective version of Björnestål's result where our constant c is completely independent of the function f and only depends on the dimension of the space P_n.

Remark 6.10 *In Kroó [118] the problem of L_1-approximation of continuous functions is considered for arbitrary Haar subspaces of $C[0,1]$ containing the constant functions. Kroó [120] treats uniqueness subspaces of $C[0,1]$ but in that case the constant c also depends on values of the function f and not only on its modulus of continuity. Since Cheney's proof which we analyzed works for arbitrary Haar subspaces we are also guaranteed to extract uniform moduli of uniqueness in the general setting. As done by Jackson [70] in his original proof, in the present work we focused on the specific Haar subspace P_n in order to get fully explicit results. One can observe that only Section 6.3.8 (Lagrange interpolation formula used to show that P_n is a Haar space), Section 6.3.5 (Markov inequality used to show that $K_{f,n}$ is compact by constructing a common modulus of uniform continuity) and Section 6.3.11 (Markov inequality plus the construction of a polynomial which changes sign in each x_i) made reference to the particular Haar space P_n. From results in [26](Lemma 4.3), [27](lemma) and [88](after Lemma 9.32) it follows that there exist effective and quantitative substitutes for each of these constructions for arbitrary (effectively given) Haar spaces. So it is clear that the analysis carried out in this paper can be extended to general Haar spaces H containing the constant functions*[28].

6.7 Concluding Remarks on the Extraction of Φ

We emphasize again the two important roles played by logic in the extraction of the modulus of uniqueness for best L_1-approximation presented here. First, by showing that Cheney's proof could be formalized in the system $\mathcal{A}^\omega_*$ (and by the logical meta-theorem 6.1) we were guaranteed that such a modulus Φ would exist and that it could be extracted from the mentioned proof. Moreover, the fact that Φ depends only on ω_f, n and ε (which was proved by Kroó years after Cheney's proof) is obtained immediately from the meta-theorem 6.1. The second important role is that logic not only guaranteed the existence of the modulus but it went even further and supplied a procedure (monotone functional interpretation) to extract the modulus, which enabled us to provide for the first time an explicit dependency on n and ω_f. And, as it happened, the extracted modulus of uniqueness has the optimal ε-dependency established by Kroó.

We hope it is transparent that all the mathematical tools used in our analysis

[27] As in Björnestål [22], Kroó does not present the actual constant.

[28] We only need the constant functions to belong to H if we want to get rid of the f dependency in c, i.e. obtain a constant c in the uniform modulus of uniqueness depending only on n and ω_f.

were already present in Cheney's proof,[29] which can be noticed for instance in the analysis of Lemma 1 (Section 6.3.10) where in order to prove that the functionals presented realized the lemma (see Claim 6.7) we followed line by line the original proof from [34], the only difference being that we considered the ε-version of the propositions. This visibly shows that the uniform modulus of uniqueness here extracted was really implicitly present in Cheney's proof but could only be made explicit with the help of logic. The difficulty to extract ad hoc such information can be understood because Cheney's proof (although very simple from the mathematical point of view and even called 'elementary' by the author) is logically very intricate due to the use of proof by contradiction and principles that fail in computable analysis.

[29]Except Markov inequality which was used to show that the set $K_{f,n}$ is compact (and also in Section 6.3.11) and Lagrange interpolation formula used to prove that P_n is a Haar space. These tools, however, are standard in approximation theory.

Chapter 7

On the Computational Complexity of L_1-approximation

Mathematical Logic Quarterly, 48.s1:66–77, 2002.

Abstract

It is well known that for a given continuous function $f : [0, 1] \to \mathbb{R}$ and a number n there exists a unique polynomial $p_n \in P_n$ (polynomials of degree $\leq n$) which best L_1-approximates f. We establish the first upper bound on the complexity of the sequence $(p_n)_{n\in\mathbb{N}}$, assuming f is polynomial-time computable. Our complexity analysis makes essential use of the modulus of uniqueness for L_1-approximation presented in [107].

7.1 Introduction

It is well known in approximation theory (cf. Jackson's theorem, [34] or [70]) that for a fixed continuous function f on the interval $[0, 1]$ (written $f \in C[0, 1]$) and a fixed $n \in \mathbb{N}$ there exists a unique element of P_n (polynomials of degree $\leq n$ with real coefficients) which best approximates f with respect to the L_1-norm

$$\|g\|_1 :\equiv \int_0^1 |g(x)|\, dx.$$

More precisely, given $f \in C[0, 1]$ and $n \in \mathbb{N}$ there exists a unique polynomial $p_n \in P_n$ such that

$$\|f - p_n\|_1 \leq \|f - p\|_1,$$

for any $p \in P_n$. In this paper we analyze the computational complexity of the sequence $(p_n)_{n\in\mathbb{N}}$, assuming f is a polynomial-time computable function. Since the coefficients of each p_n are potentially real numbers, in our analysis we make use of the concepts and tools developed in *computable analysis* (a brief introduction to computable analysis is presented in Section 7.2).

Our development in this paper follows the pattern used by Ko [82] in the analysis of the sequence $(p_n)_{n\in\mathbb{N}}$ of best Chebysheff approximations. The main difference in our approach is that we make a bold distinction between two steps in the analysis:

i) Finding a modulus of uniqueness Φ (see Section 7.3).

ii) Using Φ to compute (analyze the complexity of) the sequence $(p_n)_{n\in\mathbb{N}}$.

This distinction is important for understanding the difficulty in computing (or analysing the complexity of) the sequence $(p_n)_{n\in\mathbb{N}}$ of best L_1-approximations: The first modulus of uniqueness for L_1-approximation was presented very recently (cf. [107]), although uniqueness for L_1-approximation was known for over eighty years [70].

In Section 7.3 we present the notion of *modulus of uniqueness* and the modulus of uniqueness for L_1-approximation (see [107]) which is used by the algorithm described in Section 7.4. The general idea of the algorithm is taken from [90]. The algorithm is given together with the proof of correctness and complexity analysis.

Remark 7.1 *In the following we make use of well-known classical complexity classes* **P**, **NP**, **PSPACE**, **FP**, **#P** *and classes in the polynomial hierarchy. Moreover, relativized complexity classes are represented by* $\mathbf{C}[A]$, *where* **C** *is a complexity class and* A *is an oracle, e.g.* $\mathbf{P}[\mathbf{NP}]$ $(= \Delta_2^P)$ *means polynomial time with* **NP** *oracle. Readers not familiar with classical complexity theory are referred to e.g. [136].*

7.2 Computable Analysis

While classical complexity theory deals with subsets of (or functions on) countable sets (e.g. $\mathbb{N}$, Σ^* for a finite alphabet Σ, etc.) *computable* (or *effective*) *analysis* deals mainly with operations on uncountable sets (e.g. $\mathbb{R}$, $C[0,1]$, Σ^ω, etc.). In this section we give a brief introduction to Ko's approach to effective analysis as presented in [83] and [84]. Therefore, all the definitions, Theorem 7.1 and Corollary 7.1 in this subsection are taken from [83] with small changes on the notation. For other essentially equivalent approaches to computable analysis see, for instance, [139] and [161].

7.2.1 Computable real number

Real numbers are represented by converging sequence of dyadic approximations. (A rational number is dyadic if it has a finite binary representation. The set of dyadic numbers is represented as $\mathbb{D}$.) If $d \in \mathbb{D}$ has binary representation $b_m \dots b_1.e_1 \dots e_n$ $(b_i, e_j \in \{0,1\})$ then d is said to have precision n (written $prec(d) = n$). A function $\psi : \mathbb{N} \to \mathbb{D}$ is a *Cauchy name* for a real number x if $|x - \psi(n)| \le 2^{-n}$, for all $n \in \mathbb{N}$. A real number x is *computable* if it has a computable Cauchy name, i.e. if there exists a Turing machine M_x generating on input $n \in \mathbb{N}$ a $d_n \in \mathbb{D}$ such that $d_0, d_1, \dots$ is a Cauchy sequence converging to x with fixed rate 2^{-n}.

For our complexity analysis we must carefully fix how inputs are given. Natural numbers will be represented by elements of the set $S_1 = \{0\}^*$, and the dyadic numbers by elements of $S_2 \subset \{\cdot, 0, 1\}^*$ in the standard way. For the sake of simplicity we shall confuse the elements of S_1 and S_2 with the numbers they represent.

If there is a Turing machine M_x which on input $n \in S_1$ outputs a string $d_n \in S_2$ such that $\psi(n) :\equiv d_n$ is a Cauchy name for x and moreover the machine M_x works in polynomial time[1] we say that x is a *polynomial-time computable real number* (written $x \in \mathbf{P}_\mathbb{R}$). The class $\mathbf{P}_\mathbb{R}$ can be characterized via general left cuts as follows.

Definition 7.1 *Let ψ be a Cauchy name of $x \in \mathbb{R}$. The set*

$$L = \{d \in S_2 : d \leq \psi(prec(d))\}$$

is called the left cut of x associated with ψ *(or* a general left cut of x*).*

Lemma 7.1 *Let $x \in \mathbb{R}$. $x \in \mathbf{P}_\mathbb{R}$ iff x has a general left cut in* $\mathbf{P}$.[2]

Proof. If $x \in \mathbb{R}$ has a polynomial-time computable Cauchy name ψ, it is clear that the general left cut associated with this ψ will be in $\mathbf{P}$. On the other hand, suppose L is a general left cut of x in $\mathbf{P}$. Given a precision $k \in S_1$, by binary search on L, we can find a d such that $|x - d| \leq 2^{-k}$. Since $L \in \mathbf{P}$, the binary search can be performed in polynomial time. □

In this way we have reduced the problem of the complexity of a real number x to the complexity (in the sense of classical complexity theory) of a general left cut of x. The same idea can be used to define the class of *nondeterministic polynomial-time computable real number* $\mathbf{NP}_\mathbb{R}$, i.e. a real number x belongs to $\mathbf{NP}_\mathbb{R}$ if x has a general left cut in $\mathbf{NP}$.

Remark 7.2 *In Section 7.4 we make use of a general left cut L of a real number x in order to compute an approximation $d \in \mathbb{D}$ of x with precision k (i.e. $|x - d| \leq 2^{-k}$). As mentioned above, this can be done in polynomial time with oracle access to L.*

We shall now define computability and complexity for sequences of polynomials. Here we use Ko's notion of strong computability which is defined as follows. For simplicity we assume that the n-th polynomial has degree n.

Definition 7.2 *A sequence of polynomials $(p_n)_{n\in\mathbb{N}}$ is* strongly computable *if there exists a Turing machine M which, for given $n, k \in S_1$, generates an $(n+1)$-tuple $b_0, \ldots, b_n \in S_2$ such that $|a_i - b_i| \leq 2^{-k}$, for $0 \leq i \leq n$, where $p_n(x) = a_0 + \ldots + a_n x^n$.*

---

[1] The running time of a Turing machine is calculated as usual with respect to the size of the input string. It is fair to give the input n in unary since the required output $\psi(n) = d_n$ must be close to x by 2^{-n}, i.e. the string d_n ($\in S_2$) will normally have precision (and consequently length) greater n.

[2] [Notice that in order to produce a Cauchy name out of a general left cut for a real number x one needs an upper bound on the absolute value of x, so that the binary search can be performed. In the same way, a sequence of uniformly bounded real numbers given as Cauchy names can be translated into a sequence of general left cuts. This shall be used implicitly in Section 7.4.]

If the Turing machine M above works in polynomial time we say that the sequence $(p_n)_{n\in\mathbb{N}}$ is *strongly polynomial-time computable.* Strong **NP** computability is defined as follows.

Definition 7.3 *A sequence of polynomials $(p_n)_{n\in\mathbb{N}}$ is* strongly **NP** computable *if there exists a polynomial-time non-deterministic Turing machine M such that, for given $n, k \in S_1$ at least one computation path is accepting, and in each accepting path an $(n+1)$-tuple $b_0, \ldots, b_n \in S_2$ is output such that $|a_i - b_i| \leq 2^{-k}$, for $0 \leq i \leq n$, where $p_n(x) = a_0 + \ldots + a_n x^n$.*

This definition can be generalized, for instance, as follows:

i) if M is a polynomial-time deterministic oracle Turing machine with an **NP** oracle then $(p_n)_{n\in\mathbb{N}}$ is said to be strongly Δ_2^P computable;

ii) if M is a polynomial-time non-deterministic oracle Turing machine with an **NP** oracle then $(p_n)_{n\in\mathbb{N}}$ is said to be strongly Σ_2^P computable, etc.

7.2.2 Computable real valued functions

We now investigate computability of functions $f : \mathbb{R} \to \mathbb{R}$. In this case we are interested in estimating the time required to compute $f(x)$ for any given $x \in \mathbb{R}$ (even non-computable ones). Since we are only interested in the complexity of f, we abstract from the complexity of the input x. That is done by assuming that x is given via an *oracle machine* O_x which on input m returns in constant time a $d_m \in \mathbb{D}$ such that $|x - d_m| \leq 2^{-m}$.

Definition 7.4 *A function $f : \mathbb{R} \to \mathbb{R}$ is said to be* computable *if there exists an oracle Turing machines M_f which on input n (and oracle O_x) outputs $d_n \in \mathbb{D}$ such that $|f(x) - d_n| \leq 2^{-n}$.*

From the definition above, it follows that any computable function is continuous. Moreover, it can also be proved that any computable f, on a fixed compact interval $[a, b]$, has a computable modulus of uniform continuity, i.e. there exists a computable (in the sense of classical recursion theory) function $\omega_f : \mathbb{N} \to \mathbb{N}$ such that

$$\forall k \in \mathbb{N}; x, y \in [a, b](|x - y| \leq 2^{-\omega_f(k)} \to |f(x) - f(y)| \leq 2^{-k}).$$

Theorem 7.1 ([83]) *If $f : [a, b] \to \mathbb{R}$ is computable on $[a, b]$ then*

- *f is continuous on $[a, b]$ and*
- *f has a computable modulus of uniform continuity on $[a, b]$.*

As a corollary of Theorem 7.1 we get a complete characterization of the computable functions in terms of computability of two number theoretic functions.

Corollary 7.1 ([83]) *A function $f : [a, b] \to \mathbb{R}$ is computable iff there exist two computable functions $f_r : \mathbb{D} \cap [a, b] \times \mathbb{N} \to \mathbb{D}$ and $\omega_f : \mathbb{N} \to \mathbb{N}$ such that*

- $\forall d \in \mathbb{D} \cap [a,b]; n \in \mathbb{N}\ (|f(d) - f_r(d,n)| \leq 2^{-n})$,
- ω_f *is a modulus of uniform continuity for* f *on* $[a,b]$.

The restriction to a compact domain here is essential since a continuous function on $\mathbb{R}$ need not to be uniformly continuous on $\mathbb{R}$. Once we have this characterization of computable functions f on compact intervals via a pair of computable number theoretic functions (f_r, ω_f) we can easily define the complexity of real functions on $[a,b]$. A function $f : [a,b] \to \mathbb{R}$ is *polynomial-time computable* if $f_r \in \mathbf{FP}$ and ω_f is a polynomial.

7.2.3 Complexity of integration

Note that integration is an operation which takes an f (e.g. in $C[0,1]$) and returns an $x \in \mathbb{R}$. There is no well established notion of complexity classes for such operations. The best we can do is to analyze the complexity of the real number x when the complexity of f is fixed. A result which shows that integration is a difficult operation (in the sense just explained) is due to Friedman [52] and establishes that the integral of a polynomial-time computable function is always a polynomial-time computable real number iff $\mathbf{FP} = \#\mathbf{P}$. In our analysis we abstract from the complexity of integration by the use of an oracle. If we want to take into account the complexity of integration (oracle B_f of Section 7.4) the best result is given in [85]:

Theorem 7.2 *If* $f \in C[0,1]$ *is polynomial-time computable then the real number* $\int_0^1 f(x)\, dx$ *is in* $\mathbf{PSPACE}_{\mathbb{R}}$.

7.3 The modulus of uniqueness

Let U and V be Polish spaces (i.e. complete, separable metric spaces) and $G : U \times V \to \mathbb{R}$ a real-valued continuous function. The fact that $G(u,\cdot)$ has at most one root in some compact set $V_u \subseteq V$ (parametrized by u) is expressed as

$$\forall u \in U; v_1, v_2 \in V_u \big(\bigwedge_{i=1}^{2} G(u, v_i) = 0 \to v_1 = v_2 \big).$$

A *modulus of uniqueness* (notion introduced in [90]) for the function G is a functional Φ such that

$$\forall u \in U; v_1, v_2 \in V_u; k \in \mathbb{N}\ \big(\bigwedge_{i=1}^{2} |G(u, v_i)| \leq 2^{-\Phi(u,k)} \to d_V(v_1, v_2) \leq 2^{-k} \big)$$

where d_V is a metric in V. The functional Φ generally depends on the representation of u as an element of the Polish space U.

Remark 7.3 *It turns out that for a broad class of (even non-constructive) proofs of uniqueness theorems one can extract moduli of uniqueness (cf. [91]).*

Two such extractions have been carried out in the context of approximation theory (namely for the Chebyshev approximation [90, 91] and L_1-approximation [107]) as part of the project of proof mining (extraction of constructive content from prima facie ineffective proofs in mathematical analysis by means of logical analysis). For the case under study (L_1-approximation) the modulus of uniqueness was extracted from Cheney's proof of Jackson's theorem (cf. [34, 70]). Further information about the project of proof mining and other applications can be found in [92, 100, 101].

The main application of the modulus of uniqueness $\Phi(u, k)$ of a function G is its use in the computation of a root of $G(u, \cdot)$ uniformly in u, given that a root exists (see [90]). In the rest of the paper we carry out all the details of this computation for the case of L_1-approximation using the modulus of uniqueness presented in [107]. First, however, we explain how the general picture described above indeed applies to L_1-approximation. We should keep in mind though that the whole procedure is very general, and by no means confined to the area of approximation theory.

7.3.1 Best L_1-approximation

Let $(C[0,1], \|\cdot\|_1)$ denote the normed linear space of all continuous functions on the interval $[0,1]$ with metric $d(f,g) = \|f-g\|_1$. The *distance* of an element $f \in C[0,1]$ from the subspace P_n (polynomials of degree $\leq n$) with respect to the L_1-norm is defined as $dist_1(f, P_n) :\equiv \inf_{p\in P_n} \|f-p\|_1$. Therefore, an element p^* is a *best L_1-approximation of f from P_n* if $\|f-p^*\|_1 = dist_1(f, P_n)$. If we define a function $G : (C[0,1], \|\cdot\|_1) \times P_n \to \mathbb{R}$ as $G(f,p) :\equiv \|f-p\|_1 - dist_1(f, P_n)$ it is clear that $G(f, p^*) = 0$, i.e. the best L_1-approximations of f from P_n are precisely the roots of the function $G(f, \cdot)$.

We have not argued so far that any $f \in C[0,1]$ indeed has a best L_1-approximation. This can be done in the following way. Let $K_{f,n} :\equiv \{p \in P_n : \|p\|_1 \leq 2\|f\|_1\}$. Any best L_1-approximation of f from $K_{f,n}$ is also a best L_1-approximation of f from P_n (cf. Lemma 7.2). Since $K_{f,n}$ is a bounded and closed subset of the finite-dimensional subspace P_n of the normed linear space $(C[0,1], \|\cdot\|_1)$, it is compact. The existence of a best L_1-approximation then follows from the fact that G is continuous and therefore attains its infimum in the compact set $K_{f,n}$. As shown in [70], the best L_1-approximation of any $f \in C[0,1]$ from P_n is in fact unique (henceforth called p_n).

A *modulus of uniqueness for L_1-approximation* is a functional Φ such that, for all f in $C[0,1]$,

$$\forall n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; k \in \mathbb{N}\Big(\bigwedge_{i=1}^{2}(|G(f,p_i)| \leq 2^{-\Phi(f,n,k)}) \to \|p_1 - p_2\|_1 \leq 2^{-k}\Big).$$

We note that a modulus for the space $K_{f,n}$ can be easily extended to a modulus for the whole space P_n.

As pointed out in Remark 7.3, one can try to extract from a proof of the uniqueness of best L_1-approximation such a functional Φ. Such extraction is carried for Cheney's proof (cf. [34]) in [107] where a modulus of uniqueness Φ for L_1-approximation is obtained. The logical meta-theorems which guarantee the extraction of moduli of uniqueness, however, can only be applied when the function G (whose uniqueness of the root has been proved) is explicitly definable by a term of the underlying formal system and in particular continuous as a function on a Polish space. Since the space $(C[0,1], \|\cdot\|_1)$ is not complete, for the extraction of Φ we use the Polish space $U = (C[0,1], \|\cdot\|_\infty)$ instead. The functional G being continuous w.r.t. the uniform topology in $C[0,1]$ follows from the fact that $\|\cdot\|_1$ is continuous in $(C[0,1], \|\cdot\|_\infty)$ (which follows from $\|f\|_1 \leq \|f\|_\infty$).[3]

As already mentioned, it is important that the functional Φ will in general depend on f through its representation as an element of $(C[0,1], \|\cdot\|_\infty)$. Such f is represented as a pair (f_r, ω_f), where the first element is the restriction of f to the dyadic numbers and ω_f is the modulus of uniform continuity of f (cf. Corollary 7.1).

7.3.2 Modulus of uniqueness for L_1-approximation

As mentioned above, the computation of the sequence[4] $(p_n)_{n\in\mathbb{N}}$ for a given function $f \in C[0,1]$ makes essential use of the modulus of uniqueness for best L_1-approximation. We present the modulus (taken from [107], cf. Remark 7.3) in this section.

Theorem 7.3 ([107]) *Let*

$$\Phi(\omega_f, n, k) :\equiv 2k+(4n^2+10n+18)\log(n+2)+\omega_f(k+(2n^2+5n+6)\log(n+2)).$$

The functional Φ is a modulus of uniqueness for the best L_1-approximation of any $f \in C[0,1]$, having modulus of uniform continuity ω_f, from P_n, i.e. for all $n \in \mathbb{N}$, $p_1, p_2 \in P_n$,

$$\forall k \in \mathbb{N}\Big(\bigwedge_{i=1}^{2} \|f - p_i\|_1 - dist_1(\omega_f, P_n) \leq 2^{-\Phi(\omega_f,n,k)} \to \|p_1 - p_2\|_1 \leq 2^{-k}\Big).$$

Throughout the rest of the article, Φ will denote the modulus of uniqueness defined in Theorem 7.3.

It is important to notice that (besides being independent of p_1 and p_2) the modulus of uniqueness Φ depends on f only through its modulus of uniform continuity ω_f and does not depend on any particular value of the function f. Moreover, the above modulus has optimal k-dependency (as follows from [118]). In the following sections we will make use of some facts about the L_1-norm which we present here. For the rest of this section we let f and n be fixed.

[3] The continuity of G (w.r.t. the uniform topology in $C[0,1]$) also follows from the fact that G is primitive recursively definable in (f_r, ω_f) and n (cf. [90]). Actually, this is the fact which guarantees the applicability of the meta-theorems of [90] (cf. Remark 7.3) to Cheney's proof of Jackson's theorem yielding the results of [107].

[4] Throughout the rest of the paper p_n will denote the best L_1-approximation of $f \in C[0,1]$ from P_n, for a fixed f.

Lemma 7.2 *Let $K_{f,n} :\equiv \{p \in P_n : \|p\|_1 \leq 2\|f\|_1\}$. The zero polynomial (which belongs to $K_{f,n}$) L_1-approximates f better than any $p \notin K_{f,n}$.*

Proof. Let $p \notin K_{f,n}$ be fixed, i.e. $\|p\|_1 > 2\|f\|_1$. Therefore, by the triangle inequality for the L_1-norm, $\|f - p\|_1 > \|f\|_1$. □

As a consequence of Lemma 7.2 we get that $dist_1(f, P_n) = dist_1(f, K_{f,n})$. Therefore, any polynomial p^* such that $\|f - p^*\|_1 = dist_1(f, K_{f,n})$ is a best L_1-approximation of f from P_n.

Markov's inequality states that, for any given $p \in P_n$, $\|p'\|_\infty \leq 2n^2\|p\|_\infty$, where p' denotes the first derivative of p.

Lemma 7.3 *If $p \in P_n$ then $\|p\|_\infty \leq 2(n+1)^2\|p\|_1$.*

Proof. Let $p(x) = a_0 + a_1x + \ldots + a_nx^n$. Define $\tilde{p}(x) :\equiv a_0x + \frac{a_1}{2}x^2 + \ldots + \frac{a_n}{n+1}x^{n+1}$. It is clear that for any $x \in [0,1]$,

$$|\tilde{p}(x)| = |\int_0^x p(y)dy| \leq \int_0^x |p(y)|dy \leq \|p\|_1,$$

therefore, $\|\tilde{p}\|_\infty \leq \|p\|_1$. Since the derivative of $\tilde{p}$ equals p, by Markov's inequality, we have $\|p\|_\infty \leq 2(n+1)^2\|p\|_1$. □

Lemma 7.4 *Let $p(x) = a_0 + a_1x + \ldots + a_nx^n$ be an element of P_n and $M \in \mathbb{R}^*_+$. If $\|p\|_1 \leq M$ then $|a_i| \leq 4(n+1)^{2(i+1)}M$, $0 \leq i \leq n$.*

Proof. Let $p(x) = a_0 + a_1x + \ldots + a_nx^n$ and assume $\|p\|_1 \leq M$. By Lemma 7.3 we have, (1) $\|p\|_\infty \leq 2(n+1)^2M$. Let $p^{(i)}$ denote the i-th derivative of p. It is clear that $a_i = \frac{p^{(i)}(0)}{i!}$. By applying Markov's inequality i times and by (1) we have (2) $\|p^{(i)}\|_\infty \leq 2^{i+1}(n+1)^{2(i+1)}M$, and therefore,

$$|a_i| = \frac{|p^{(i)}(0)|}{i!} \overset{(2)}{\leq} \frac{2^{i+1}(n+1)^{2(i+1)}M}{i!} \leq 4(n+1)^{2(i+1)}M.$$

□

Let $p(x) = a_0 + a_1x + \ldots + a_nx^n$ be any polynomial in $K_{f,n}$ (which, by Lemma 7.2, includes p_n). By the definition of $K_{f,n}$ and Lemma 7.4 we have that $|a_i| \leq 8(n+1)^{2(i+1)}\ M$, for $0 \leq i \leq n$, where $M \in \mathbb{D}$ is an upper bound on $\|f\|_1$. Since we will use this bound on the coefficients of the elements of $K_{f,n}$ we give it a name, $C_{n,i} :\equiv 8(n+1)^{2(i+1)}\ M$. We will also need a function $\Theta(n,k)$ such that for polynomials $p(x) = a_0 + a_1x + \ldots + a_nx^n$,

$$\|p\|_1 \leq 2^{-\Theta(n,k)} \rightarrow \bigwedge_{i=0}^{n} |a_i| \leq 2^{-k}, \tag{7.1}$$

which can be easily derived from Lemma 7.4, for instance $\Theta(n,k) :\equiv 2(n+1)\log(n+1) + k + 2$.

Definition 7.5 *A set of elements $N_{n,k} \subset P_n$ is called a (n,k)-net (for $K_{f,n}$) if for any element $\tilde{p} \in K_{f,n}$ there exists an element $p \in N_{n,k}$ which is $(k+1)$-close to $\tilde{p}$, i.e. $\|p - \tilde{p}\|_1 \leq 2^{-k-1}$.*

We want to choose a net based on the representation of the dyadic numbers so that we have control over the precision of the elements.

Lemma 7.5 *Let $C_{n,k,i} :\equiv \{a \in S_2 : prec(a) \leq k + \log(\frac{n+1}{i+1})$ and $|a| \leq C_{n,i}\}$. The space of polynomials $N_{n,k} :\equiv \{a_0 + \ldots + a_n x^n : a_i \in C_{n,k,i}, 0 \leq i \leq n\}$ is a (n,k)-net.*

Proof. Take an arbitrary element of $K_{f,n}$, say $\tilde{p}(x) = b_0 + \ldots + b_n x^n$. In the way we have chosen the coefficients of the elements of $N_{n,k}$ we are able to find $p(x) = a_0 + \ldots + a_n x^n \in N_{n,k}$ such that $|a_i - b_i| \leq \frac{2^{-k-1}(i+1)}{n+1}$, $0 \leq i \leq n$, i.e.

$$
\begin{aligned}
\|p - \tilde{p}\|_1 &= \|(a_0 - b_0) + \ldots + (a_n - b_n)x^n\|_1 \\
&= \int_0^1 |(a_0 - b_0) + \ldots + (a_n - b_n)x^n| dx \\
&\leq |a_0 - b_0| + \ldots + \frac{|a_{n-1} - b_{n-1}|}{n} + \frac{|a_n - b_n|}{n+1} \\
&\leq \frac{2^{-k-1}}{n+1} + \ldots + \frac{2^{-k-1}n}{n(n+1)} + \frac{2^{-k-1}(n+1)}{(n+1)(n+1)} = 2^{-k-1}.
\end{aligned}
$$

□

7.4 The complexity of $(p_n)_{n\in\mathbb{N}}$

For the rest of the article f denotes a fixed polynomial-time computable function. As mentioned before, we will analyze the complexity of the sequence $(p_n)_{n\in\mathbb{N}}$ relative to the complexity of integration. Therefore, in the following we will make use of an oracle B_f which is supposed to answer queries about integration.

In the case of the best Chebyshev approximation the value $dist_\infty(f, P_n)$ can be computed beforehand, and that value can be used in the computation of the best Chebyshev approximation of f from P_n. For the sake of comparison between the two cases of Chebyshev and L_1-approximation, in the first part of this section we first analyze the complexity of $(p_n)_{n\in\mathbb{N}}$ relative to an oracle A_f for $dist_1(f, P_n)$ (as done in [82] for the Chebyshev case). Then, in the last section we present an algorithm which does not need the values of $dist_1(f, P_n)$ in advance. From this algorithm we obtain a complexity upper bound for the sequence $(p_n)_{n\in\mathbb{N}}$ relative solely to the oracle B_f.

7.4.1 Using oracle A_f for $dist_1(f, P_n)$

Let L_n be a general left cut of the real number $dist_1(f, P_n)$. The oracle A_f decides the set

$$\{\langle n, d\rangle : d \in L_n\}$$

where $n \in S_1$ and $d \in S_2$. The second oracle B_f answers queries about general left cuts of the real numbers $\|f - p\|_1$, uniformly in p. More precisely, let $L_{n,p}$

denote a general left cut of the real number $\|f-p\|_1$. The oracle B_f decides the set [5]

$$\{\langle n, p, e\rangle : e \in L_{n,p}\}$$

where $n \in S_1$ and $a_0, \ldots, a_n, e \in S_2$. As done in [82] for the Chebyshev case, we first show how to decide a certain set $\mathcal{G}_f$ using the oracles A_f and B_f. (The oracles are used as mentioned in Remark 7.2.)

$\langle n, k, a_0, \ldots, a_n\rangle \in \mathcal{G}_f; n, k \in S_1$ and $a_0, \ldots, a_n \in S_2$ Oracles: A_f, B_f
Let $s :\equiv \Phi(\omega_f, n, \Theta(n,k))$;   If $p \notin N_{n,s+1}$ output *no*; (cf. Lemma 7.5)   Compute $dist_1(f, P_n)$ with precision $s+3$ (let the resulting value be $d \in S_2$);   Compute $\|f-p\|_1$ with precision $s+3$ (let the resulting value be $e \in S_2$);   Output *yes* iff $\|d-e\| \leq 2^{-s-1}$.

Theorem 7.4 *Let $f \in C[0,1]$ be polynomial-time computable and ω_f a polynomial modulus of uniform continuity of f. There exists a multi-valued function α_f which on input n and k ($\in S_1$) produces a non-empty set of $(n+1)$-tuples ($\in S_2^{n+1}$) (representing elements of P_n) such that for each $\langle a_0, \ldots, a_n\rangle \in \alpha_f(n,k)$,*

(i) for $0 \leq i \leq n$, $prec(a_i) \leq \Phi(\omega_f, n, \Theta(n,k)) + \log(\frac{n+1}{i+1}) + 1$;

(ii) for $0 \leq i \leq n$, $|b_i - a_i| \leq 2^{-k}$ (where $p_n(x) = b_0 + \ldots + b_n x^n$).

Moreover,

(iii) Graph$(\alpha_f) \in \mathbf{P}[A_f, B_f]$.

Proof. Let s be a shorthand for $\Phi(\omega_f, n, \Theta(n,k))$. We define α_f to be the function that maps each $n, k \in S_1$ to all $(n+1)$-tuples $\langle a_0, \ldots, a_n\rangle \in S_2^{n+1}$ such that $\langle n, k, a_0, \ldots, a_n\rangle \in \mathcal{G}_f$, i.e. α_f is the function whose graph is $\mathcal{G}_f$. We first have to argue that α_f is total. Let n, k be fixed. By Lemma 7.5 and the fact that $p_n \in K_{f,n}$, there exists a $p \in N_{n,s+1}$ such that $\|p_n - p\|_1 \leq 2^{-s-2}$. By the triangle inequality for the L_1-norm we get $\|f-p\|_1 - dist_1(f, P_n) \leq 2^{-s-2}$. By the computation of d and e we have

$$|\, d - dist_1(f, P_n)\,| \leq 2^{-s-3} \quad \text{and} \quad |\, e - \|f-p\|_1\,| \leq 2^{-s-3},$$

which implies $|d-e| \leq 2^{-s-1}$, and the input p is accepted.

(i) Immediate consequence of the definition of a net (7.5) and the definition of α_f.

(ii) Suppose $\langle a_0, \ldots, a_n\rangle \in \alpha_f(n,k)$ (let $p(x) :\equiv \sum_{i=0}^{n} a_i x^i$). This implies $|e -$

[5] If $p(x) = a_0 + \ldots + a_n x^n \in P_n$ we also write, for convenience, $\langle \ldots, p, \ldots\rangle$ instead of $\langle \ldots, a_0, \ldots, a_n, \ldots\rangle$.

$d| \leq 2^{-s-1}$, d and e as above. We then obtain $\|f-p\|_1 - dist_1(f,P_n) \leq 2^{-s}$. By Theorem 7.3 we get $\|p_n - p\|_1 \leq 2^{-\Theta(n,k)}$. And by (7.1) of Section 7.3.2, $|b_i - a_i| \leq 2^{-k}$, for $0 \leq i \leq n$, where $p_n(x) = b_0 + \ldots + b_n x^n$.

(iii) Since ω_f is a polynomial (cf. Section 7.2.2), $\Phi(\omega_f, n, \Theta(n,k))$ is also a polynomial and the procedure $\mathcal{G}_f$ above can be performed in polynomial time (in A_f and B_f). Notice also that since f is fixed the net $N_{n,s+1}$ has size exponential on the input. □

Corollary 7.2 *Let* $f \in C[0,1]$ *be polynomial-time computable. The sequence of best* L_1*-approximation* $(p_n)_{n\in\mathbb{N}}$ *is strongly* $\mathbf{NP}[A_f,\, B_f]$ *computable.*

Proof. Let $n,k \in S_1$ be given. We define a non-deterministic oracle Turing machine M as follows. The oracles of M will be the sets A_f and B_f. Each computation path of M takes into consideration one element $p \in N_{n,s+1}$ (s as above). The machine (in each path) decides whether $\langle n,k,p\rangle$ belongs to $\mathcal{G}_f$ (i.e. Graph(α_f)) or not. If yes then the path is accepted and the machine outputs p. Note that, by Theorem 7.4 (i), the size of p is a polynomial on n and k. □

We obtain, for instance, that if A_f and B_f are in **NP** then Graph$(\alpha_f) \in \Delta_2^P$ and $(p_n)_{n\in\mathbb{N}}$ is strongly Σ_2^P computable.

Remark 7.4 *Note that the set*

$$L :\equiv \{\langle n,d\rangle \in S_1 \times S_2 : \forall p \in N_{n,k}\ (d \leq \|f-p\|_1^{k+1})\}$$

where the k *above abbreviates* $prec(d)$ *and* $\|f-p\|_1^{k+1}$ *is a* $(k+1)$*-approximation of the value* $\|f-p\|_1$*, does the job of the oracle* A_f*. In other words, the set*

$$L_n :\equiv \{d \in S_2 : \langle n,d\rangle \in L\}$$

is a general left cut of $dist_1(f,P_n)$*. An algorithm for deciding the complement of* L *can be given as follows. On input* $d \in S_2$ *(with precision* k*) and* $n \in S_1$*, non-deterministically choose a polynomial from* $N_{n,k}$ *and compute the value of* $\|f-p\|_1$ *with precision* $k+1$ *(say* e*). Then, answer yes (i.e.* $\langle n,d\rangle \notin L$*) when* $d > e$*. In this way, using the oracle* B_f *for integration, we obtain an upper bound* ***coNP***$[B_f]$ *on the complexity of the oracle* A_f*. Note also that the above procedure does not make use of the fact that the best* L_1*-approximation of* f *is unique.*

7.4.2 Absolute complexity of $(p_n)_{n\in\mathbb{N}}$

In this section we present another algorithm which only uses the oracle B_f for a general left cuts of $\|f-p\|_1$ (and does not make use of the oracle A_f). We first use B_f to define the set $\tilde{\mathcal{G}}_f$,

$\langle n, k, p, \tilde{p}\rangle \in \tilde{\mathcal{G}}_f$; $n, k \in S_1$ and $p, \tilde{p} \in S_2^{n+1}$ Oracles: B_f
Let $s :\equiv \Phi(\omega_f, n, \Theta(n, k))$; If $p \notin N_{n,s+1}$ output *no*; (cf. Lemma 7.5) Compute $\|f - p\|_1$ with precision $s + 3$ (let the resulting value be $e \in S_2$); Compute $\|f - \tilde{p}\|_1$ with precision $s + 3$ (let the resulting value be $\tilde{e} \in S_2$); Output *yes* iff $e \leq \tilde{e} + 2^{-s-1}$.

Note that deciding membership for the set $\tilde{\mathcal{G}}_f$ can be done in polynomial-time using the oracle B_f, i.e. $\tilde{\mathcal{G}}_f \in \mathbf{P}[B_f]$. Let

$$\mathcal{G}_f :\equiv \{\langle n, k, p\rangle : \forall \tilde{p} \in N_{n,\Phi(\omega_f,n,\Theta(n,k))+1}\big(\langle n, k, p, \tilde{p}\rangle \in \tilde{\mathcal{G}}_f\big)\}.$$

Theorem 7.5 *Let $f \in C[0, 1]$ be polynomial-time computable and ω_f a polynomial modulus of uniform continuity of f. There exists a multi-valued function β_f which on input n and k ($\in S_1$) produces a non-empty set of $(n+1)$-tuples ($\in S_2^{n+1}$) (representing elements of P_n) such that for each $\langle a_0, \ldots, a_n\rangle \in \beta_f(n, k)$,*

(*i*) *for $0 \leq i \leq n$, $prec(a_i) \leq \Phi(\omega_f, n, \Theta(n, k)) + \log(\frac{n+1}{i+1}) + 1$;*

(*ii*) *for $0 \leq i \leq n$, $|b_i - a_i| \leq 2^{-k}$ (where $p_n(x) = b_0 + \ldots + b_n x^n$).*

Moreover,

(*iii*) $\mathrm{Graph}(\beta_f) \in \mathbf{coNP}[B_f]$.

Proof. Let s be a shorthand for $\Phi(\omega_f, n, \Theta(n, k))$. We define β_f to be the function that maps each $n, k \in S_1$ to all $(n + 1)$-tuples $\langle a_0, \ldots, a_n\rangle \in S_2^{n+1}$ such that $\langle n, k, a_0, \ldots, a_n\rangle \in \mathcal{G}_f$, i.e. β_f is the function whose graph is $\mathcal{G}_f$. First we have to prove that β_f is total. Let p be an element of $N_{n,s+1}$ such that $\|f - p\|_1 \leq \min_{\tilde{p}\in N_{n,s+1}} \|f - \tilde{p}\|_1$. Then, clearly, $\langle n, k, p, \tilde{p}\rangle \in \tilde{\mathcal{G}}$, for all $\tilde{p} \in N_{n,s+1}$. Therefore, $\langle n, k, p\rangle \in \mathrm{Graph}(\beta_f)$.

(*i*) Immediate consequence of the definition of a net (7.5) and the definition of β_f.

(*ii*) Assume $\langle n, k, p, \tilde{p}\rangle \in \tilde{\mathcal{G}}_f$, for all $\tilde{p} \in N_{n,s+1}$. That implies

$$(*) \quad \forall \tilde{p} \in N_{n,s+1}\big(\|f - p\|_1 \leq \|f - \tilde{p}\|_1 + 3 \cdot 2^{-s-2}\big).$$

Since $p_n \in K_{f,n}$ (and by the definition of (n, k)-net) there is an element $\tilde{p} \in N_{n,s+1}$ such that $\|p_n - \tilde{p}\|_1 \leq 2^{-s-2}$ and by triangle inequality we get, $\|f - \tilde{p}\|_1 \leq dist(f, P_n) + 2^{-s-2}$. By $(*)$ we get, $\|f - p\|_1 \leq dist_1(f, P_n) + 2^{-s}$. Hence, by Theorem 7.3 we have $\|p_n - p\|_1 \leq 2^{-\Theta(n,k)}$. And by (7.1) of Section 7.3.2 $|b_i - a_i| \leq 2^{-k}$, for $0 \leq i \leq n$, where $p_n(x) = b_0 + \ldots + b_n x^n$.

(*iii*) Similar to Theorem 7.4 (*iii*). □

Corollary 7.3 *Let $f \in C[0, 1]$ be polynomial-time computable, then the sequence $(p_n)_{n\in\mathbb{N}}$ is strongly* **NP** *computable in* $\mathbf{NP}[B_f]$.

Proof. Let $n, k \in S_1$ be given. We define a non-deterministic oracle Turing machine M as follows. The oracle of M will be the set Graph(β_f) (which is in $\mathbf{coNP}[B_f]$). Each computation path of M takes into consideration one element $p \in N_{n,s+1}$ (s as above). The machine (in each path) decides whether $\langle n, k, p\rangle$ belongs to Graph(β_f) or not. If yes then the path is accepted and the machine outputs p. We also note that, as our oracle we can as well use the complement of the set Graph(β_f). $\square$

7.5 Conclusion

We have established the first complexity upper bound on the sequence $(p_n)_{n\in\mathbb{N}}$ of best L_1-approximations of a polynomial time computable $f \in C[0,1]$. For the complexity analysis we made use of two oracles A_f and B_f solving generalized left cuts of $dist_1(f, P_n)$ and $\|f-p\|_1$ respectively in two different ways:

1) *Relative to both oracles A_f and B_f.* We have shown that the sequence $(p_n)_{n\in\mathbb{N}}$ is strongly **NP** computable relative to those oracles. Since the oracle A_f has a trivial $\mathbf{coNP}[B_f]$ upper bound (cf. Remark 7.4) we obtain that $(p_n)_{n\in\mathbb{N}}$ is strongly $\mathbf{NP}[\mathbf{NP}[B_f], B_f]$ computable, i.e. strongly **NP** computable relative to an $\mathbf{NP}[B_f]$ oracle.

2) *Relative to oracle B_f.* We have also analyzed the complexity of $(p_n)_{n\in\mathbb{N}}$ without first computing the value $dist_1(f, P_n)$. In this case we concluded directly that the sequence $(p_n)_{n\in\mathbb{N}}$ is strongly **NP** computable relative to an $\mathbf{NP}[B_f]$ oracle.

One should note that our complexity analysis strongly relies on the modulus of uniqueness for L_1-approximation, first presented in [107].

In [82] a relation is established between the sequence $(p_n)_{n\in\mathbb{N}}$ (of best Chebysheff approximations of a polynomial time computable $f \in C[0,1]$) and separation of well known complexity classes. It is not known whether similar results also hold in the case under study of L_1-approximation.

Acknowledgement. I would like to thank Ulrich Kohlenbach for interesting discussions and many helpful suggestions.

Chapter 8

Modified Bar Recursion

with Ulrich Berger, to appear in: *Lecture Notes in Logic*, 19 pages.

Abstract

We introduce a variant of Spector's bar recursion in finite types (which we call "modified bar recursion") to give a realizability interpretation of the classical axiom of dependent choice allowing for the extraction of witnesses from proofs of $\forall\exists$-formulas in classical analysis. As another application, we show that the fan functional can be defined by modified bar recursion together with a version of bar recursion due to Kohlenbach. We also show that the type structure $\mathcal{M}$ of strongly majorizable functionals is a model for modified bar recursion.

8.1 Introduction

In [154], Spector extended Gödel's Dialectica Interpretation of Peano Arithmetic [58] to classical analysis using bar recursion in finite types. Although considered questionable from an intuitionistic point of view ([5], 6.6), there has been considerable interest in bar recursion, and several variants of this definition scheme and their interrelations have been studied by, e.g. Schwichtenberg [147], Bezem [19] and Kohlenbach [88]. In this paper we add another variant of bar recursion and use it to give a realizability interpretation of the negatively translated axiom of dependent choice that can be used to extract witnesses from proofs of $\forall\exists$-formulas in full classical analysis. Our interpretation is inspired by a paper by Berardi, Bezem and Coquand [11] who use a similar kind of recursion in order to interpret dependent choice. The main difference to our paper is that in [11] a rather ad-hoc infinitary term calculus and a non-standard notion of realizability are used whereas we work with a straightforward combination of negative translation, A-translation, modified realizability, and Plotkin's adequacy result for the partial continuous functional semantics of PCF [138].

As a second application of bar recursion, we show that the definition of the fan functional within PCF given in [12] and [133] can be derived from Kohlenbach's and our variant of bar recursion. Furthermore, we prove that our version of bar recursion exists in the model of majorizable functions. The relation between modified bar recursion and Spector's original definition is established in [14].

8.2 Bar recursion in finite types

We work in a suitable extension of Heyting Arithmetic in finite types, HA^ω, with equality in all types. For convenience, we enrich the type system by the formation of finite sequences. So, our *Types* are $\mathbb{N}$, function types $\rho \to \sigma$, product types $\rho \times \sigma$, and finite sequences ρ^*. We set $\rho^\omega :\equiv \mathbb{N} \to \rho$. The *level* of a type is defined by $\mathsf{level}(\mathbb{N}) = 0$, $\mathsf{level}(\rho \times \sigma) = \max(\mathsf{level}(\rho), \mathsf{level}(\sigma))$, $\mathsf{level}(\rho^*) = \mathsf{level}(\rho)$, $\mathsf{level}(\rho \to \sigma) = \max(\mathsf{level}(\rho) + 1, \mathsf{level}(\sigma))$. By o we will denote an arbitrary but fixed type of level 0, and by ρ, τ, σ arbitrary types. The terms of our version of HA^ω are a suitable extension of the terms of Gödel's system T [58] in lambda calculus notation. We use the variables $i, j, k, l, m, n \colon \mathbb{N}$ and $s, t \colon \rho^*$; $\alpha, \beta \colon \rho^\omega$, where ρ is an arbitrary type. Other letters will be used for different types in different contexts. By $=_\tau$ we denote equality of type τ for which we assume the usual equality axioms. However, equality between functions is *not* assumed to be extensional. We also do *not* assume decidability for $=_\tau$, when $\mathsf{level}(\tau) > 0$ (if $\mathsf{level}(\tau) = 0$ one can, of course, *prove* decidability). Type information will be frequently omitted when it is irrelevant or inferable from the context. We let k^ρ denote the canonical lifting of a number $k \in \mathbb{N}$ to type ρ, e.g. $k^{\rho\to\sigma} := \lambda x^\rho . k^\sigma$. By an $\exists$-*formula* respectively $\forall\exists$-*formula* we mean a formula of the form $\exists y^\tau\, B$ respectively $\forall z^\sigma\, \exists y^\tau\, B$, where B is provably equivalent to an atomic formula. We will also use the following notations:

$$
\begin{array}{rcl}
\langle x_0, \ldots, x_{n-1} \rangle & :\equiv & \text{the finite sequence with elements } x_0, \ldots, x_{n-1} \\
|s| & :\equiv & \text{the length of } s \text{, i.e. } |\langle x_0, \ldots, x_{n-1} \rangle| = n \\
s_k & :\equiv & \text{the } k\text{-th element of } s \text{ for } k < |s|, \\
 & & \qquad \text{i.e. } \langle x_0, \ldots, x_{n-1} \rangle_k = x_k \\
s * t & :\equiv & \text{the concatenation of } s \text{ and } t \\
s * x & :\equiv & s * \langle x \rangle \\
s * \alpha & :\equiv & \text{appending } \alpha \text{ to } s \text{, i.e.} \\
 & & s * \alpha :\equiv \lambda k.[\mathsf{if}\ k < |s|\ \mathsf{then}\ s_k\ \mathsf{else}\ \alpha(k - |s|)] \\
s\ @\ \alpha & :\equiv & \text{overwriting } \alpha \text{ with } s \text{, i.e.} \\
 & & s\ @\ \alpha :\equiv \lambda k.[\mathsf{if}\ k < |s|\ \mathsf{then}\ s_k\ \mathsf{else}\ \alpha(k)] \\
\overline{\alpha}k & :\equiv & \langle \alpha(0), \ldots, \alpha(k-1) \rangle \\
\beta \in \overline{\alpha}k & :\equiv & \overline{\beta}k =_{\rho^*} \overline{\alpha}k.
\end{array}
$$

Definition 8.1 *Spector's definition of bar recursion [154] reads in our notation as follows:*

$$
\Phi(Y, G, H, s) =_\tau \begin{cases} G(s) & \mathsf{if}\ Y(s\ @\ 0^{\rho^\omega}) < |s| \\ H(s, \lambda x^\rho . \Phi(Y, G, H, s * x)) & \mathsf{otherwise.} \end{cases} \tag{8.1}
$$

In his thesis [88] Kohlenbach introduced the following kind of bar recursion which differs from Spector's only in the stopping condition:

$$\Phi(Y,G,H,s) =_\tau \begin{cases} G(s) & \text{if } Y(s \mathbin{@} 0^{\rho^\omega}) =_o Y(s \mathbin{@} 1^{\rho^\omega}) \\ H(s, \lambda x^\rho.\Phi(Y,G,H,s*x)) & \text{otherwise.} \end{cases} \tag{8.2}$$

Finally, we define Modified bar recursion at type ρ*:*

$$\Phi(Y,H,s) =_o Y(s \mathbin{@} H(s, \lambda x^\rho.\Phi(Y,H,s*x))). \tag{8.3}$$

Note that each of the equations above defines a family of functionals $\Phi_{\rho,\tau}$ (Φ_ρ in the case of modified bar recursion) as ρ and τ range over arbitrary finite types. We shall often omit the parameters Y, G and H when defining a functional Φ using the equations above. We say a model $\mathcal{S}$ satisfies one of the respective variants of bar recursion if in $\mathcal{S}$ a functional exists satisfying the corresponding equation (8.1), (8.2), or (8.3) for all possible values of Y, G, H and s.

Recursive definitions similar to (8.3) occur in [11], and, in a slightly different form, in [12] and [133] in connection with the fan functional (cf. Section 8.4).

Remark 8.1 Note that replacing in equation (8.3) the operation @ by $*$ would be an inessential change. However it *is* essential that the type of $\Phi(s)$ is of level 0. If, for example, the type of $\Phi(s)$ were $\mathbb{N} \to \mathbb{N}$ we could set $Y(\alpha)(m) :=_\mathbb{N} \alpha(m) + 1$ and $H(s,F)(k) :=_\mathbb{N} F(0)(|s|+1)$, and obtain the equation

$$\Phi(s)(m) =_\mathbb{N} (s \mathbin{@} \lambda k.\Phi(s*0)(|s|+1))(m) + 1$$

implying

$$\Phi(\langle\,\rangle)(0) =_\mathbb{N} \Phi(\langle 0\rangle)(1) + 1 =_\mathbb{N} \Phi(\langle 0,0\rangle)(2) + 2 =_\mathbb{N} \ldots$$

which is unsatisfiable in $\mathbb{N}$.

The structures of primary interest to interpret bar recursion are the model $\mathcal{C}$ of *total continuous functionals* of Kleene [79] and Kreisel [115], the model $\widehat{\mathcal{C}}$ of *partial continuous functionals* of Scott [148] and Ershov [45] (see also [133]), and the model $\mathcal{M}$ of (strongly) *majorizable functionals* introduced by Howard [65] and Bezem [18].

Theorem 8.1 *The models $\mathcal{C}$ and $\widehat{\mathcal{C}}$ satisfy all three variants of bar recursion.*

Proof. In the model $\widehat{\mathcal{C}}$ all three forms of bar recursion can simply be defined as the least fixed points of suitable continuous functionals. For $\mathcal{C}$ we use Ershov's result in [45] according to which the model $\mathcal{C}$ can be identified with the total elements of $\widehat{\mathcal{C}}$. Therefore it suffices to show that all three versions of bar recursion are total in $\widehat{\mathcal{C}}$. For Spector's version this has been shown by Ershov [45], and for the other versions similar argument apply. For example, in order to see that $\Phi(s)$ defined recursively by equation (8.3) is total for given total Y, H and s one uses bar induction on the bar

$$P(s) :\Leftrightarrow Y(s \mathbin{@} \bot_\rho) \text{ is total}$$

where $\bot_\rho$ denotes the undefined element of type ρ. $P(s)$ is a bar because Y is continuous. □

Theorem 8.2 *$\mathcal{M}$ satisfies Spector's bar recursion (1), but not Kohlenbach's (2).*

Proof. See [18] and [88]. □

In Section 8.5 we will show that $\mathcal{M}$ satisfies modified bar recursion (3).

8.3 Using bar recursion to realize classical dependent choice

The aim of this section is to show how modified bar recursion can be used to extract witnesses from proofs of $\forall\exists$-formulas in classical arithmetic plus the axiom (scheme) of dependent choice [67]

$$\mathsf{DC}^\rho \quad : \quad \forall n, x^\rho \, \exists y^\rho \, A(n,x,y) \to \forall x \exists f \, (f(0) = x \wedge \forall n \, A(n, f(n), f(n+1))).$$

Actually we will need only the following *weak modified bar recursion* which is the special case of equation (8.3) where H is constant:

$$\Phi(Y, H, s) =_o Y(s \mathbin{@} \lambda k.H(s, \lambda x.\Phi(Y, H, s * x))). \tag{8.4}$$

Note that in (8.4) the returning type of H is ρ, i.e. the argument of Y consists of s followed by an infinite sequence with constant value of type ρ.

Before dealing with dependent choice we discuss our extraction method in general and then give a realizer for the (simpler) classical axiom of countable choice.

8.3.1 Witnesses from classical proofs

The method we use to extract witnesses from classical proofs is a combination of Gödel's negative translation (translation P^o in [124] page 42, see also [160]), the Dragalin/Friedman/Leivant trick, also called A-translation [159], and Kreisel's (formalized) modified realizability [158]. The method works in general for proofs in PA^ω, the classical variant of HA^ω. In order to extend it to PA^ω plus extra axioms Γ (e.g. $\Gamma \equiv \mathsf{DC}^\rho$) one has to find realizers for Γ^N, the negative translation of Γ [1], where $\bot$ is replaced by an $\exists$-formula (regarding negation, $\neg C$, is defined by $C \to \bot$). However, it is more direct and technically simpler to follow [16] and combine the Dragalin/Friedman/Leivant trick and modified realizability: instead of replacing $\bot$ by a $\exists$-formula we slightly change the definition of modified realizability by regarding $y \;\mathsf{mr}\; \bot$ as an (uninterpreted) atomic formula. More formally we define

[1] The negative translation double-negates atomic formulas, replaces $\exists x$ by $\neg\forall x \neg$ and $A \vee B$ by $\neg(\neg A \wedge \neg B)$.

$$y^\tau \; \mathsf{mr}^{\mathsf{c}}_\tau \; \bot :\equiv P_\bot(y),$$

where $P_\bot$ is a new unary predicate symbol and τ is the type of the witness to be extracted. Therefore, we have a modified realizability for each type τ, according to the type of the existential quantifier in the $\forall\exists$-formula we are realizing. The other clauses of modified realizability are as usual, e.g.

$$f \; \mathsf{mr}^{\mathsf{c}}_\tau \; (A \to B) :\equiv \forall x \, (x \; \mathsf{mr}^{\mathsf{c}}_\tau \; A \to fx \; \mathsf{mr}^{\mathsf{c}}_\tau \; B).$$

In the following proposition Δ is an axiom system possibly containing $P_\bot$ and further constants, which has the following closure property: If $D \in \Delta$ and B is a quantifier-free formula with decidable predicates, then also the universal closure of $D[\lambda y^\tau.B/P_\bot]$ is in Δ, where $D[\lambda y^\tau.B/P_\bot]$ is obtained from D by replacing any occurrence of a formula $P_\bot(L)$ in D by $B[L/y]$.

Proposition 8.1 *Assume there is a vector Φ of closed terms such that*

$$\mathsf{HA}^\omega + \Delta \vdash \Phi \; \mathsf{mr}^{\mathsf{c}}_\tau \; \Gamma^N.$$

Then from any proof

$$\mathsf{PA}^\omega + \Gamma \vdash \forall z^\sigma \, \exists y^\tau \, B(z, y),$$

where $\forall z^\sigma \, \exists y^\tau \, B(z, y)$ is a $\forall\exists$-formula in the language of HA^ω*, one can extract a closed term $M^{\sigma\to\tau}$ such that*

$$\mathsf{HA}^\omega + \Delta \vdash \forall z \, B(z, Mz).$$

Proof. The proof is folklore. The main steps are as follows. Assuming w.l.o.g. that $B(z, y)$ is atomic, we obtain from the hypothesis $\mathsf{PA}^\omega + \Gamma \vdash \forall z^\sigma \, \exists y^\tau \, B(z, y)$ via negative translation

$$\mathsf{HA}^\omega + \Gamma^N \vdash_m \forall y \, (B(z, y) \to \bot) \to \bot,$$

where $\vdash_m$ denotes derivability in minimal logic, i.e. ex-falso-quodlibet is not used. Now, soundness of modified realizability (which holds for our abstract version of modified realizability and minimal logic [16]), together with the assumption on Φ allows us to extract from this proof a closed term M such that

$$\mathsf{HA}^\omega + \Delta \vdash Mz \; \mathsf{mr}^{\mathsf{c}}_\tau \; (\forall y \, (B(z, y) \to \bot) \to \bot)$$

i.e.

$$\mathsf{HA}^\omega + \Delta \vdash \forall f^{\tau\to\tau} \, (\forall y \, (B(z, y) \to P_\bot(fy)) \to P_\bot(Mzf)).$$

Replacing $P_\bot$ by $\lambda y.B(z, y)$ respectively, and instantiating f by the identity function it follows

$$\mathsf{HA}^\omega + \Delta \vdash \forall z \, B(z, Mz(\lambda y.y)).$$

$\square$

We will apply this proposition with $\tau \equiv o$ (writing mr^{c} instead of $\mathsf{mr}^{\mathsf{c}}_o$), $\Gamma \equiv \mathsf{DC}^\rho$, or $\Gamma \equiv \mathsf{AC}^{\mathbb{N},\rho}$ (countable choice, see below), and an axiom system Δ consisting of the defining equation (8.3) for modified bar recursion, where the defined functionals Φ are new constants, together with the axiom of continuity CONT and the scheme of relativized quantifier-free bar induction $\mathsf{rBI\text{-}QF}$ which are defined as follows:

$$\mathsf{CONT} \;:\; \forall F^{\rho^\omega \to o}, \alpha\, \exists n\, \forall \beta\, (\overline{\alpha}n = \overline{\beta}n \to F(\alpha) = F(\beta)).^2$$

$\mathsf{rBI\text{-}QF}$:

$$\forall \alpha \in S\, \exists n\, P(\overline{\alpha}n) \wedge \forall s \in S\, (\forall x\, [S(s * x) \to P(s * x)] \to P(s)) \wedge S(\langle\,\rangle) \to P(\langle\,\rangle).$$

Here $S(s)$ is an arbitrary, and $P(s)$ a quantifier-free predicate in the language of $\mathsf{HA}^\omega[P_\perp]$, and $\alpha \in S$ and $s \in S$ are shorthands for $\forall n\, S(\overline{\alpha}n)$ and $S(s)$ respectively. Clearly the condition on Δ in Proposition 8.1 is satisfied.

In order to make sure that realizers can indeed be used to compute witnesses one needs to know that, 1. the axioms of $\mathsf{HA}^\omega + \Delta$ hold in a suitable model – here we can choose the model $\mathcal{C}$ of continuous functionals – and, 2. every closed term of type level 0 (e.g. of type $\mathbb{N}$) can be reduced to a numeral in an effective and provably correct way. In [11] this is solved by building the notion of reducibility to normal form into the definition of realizability. In our case we solve this problem by applying Plotkin's adequacy result [138] as follows: each term in the language of HA^ω plus the bar recursive constants can be naturally viewed as a term in the language PCF [138], by defining the bar recursors by means of the general fixed point combinator. In this way our term calculus also inherits PCF's call-by-name reduction, i.e. if M is bar recursive and M reduces to M' then M' is bar recursive. Furthermore reduction is provably correct in our system, i.e. if M reduces to M' then $M = M'$ is provable. Now let M be a closed term of type $\mathbb{N}$. By Theorem 8.1, M has a total value, which is a natural number n, in the model of partial continuous functionals. Hence, by Plotkin's adequacy theorem M reduces to the numeral denoting n.

8.3.2 Realizing $(\mathsf{AC}^{\mathbb{N},\rho})^N$

We now construct a realizer of the negatively translated axiom of countable choice

$$\mathsf{AC}^{\mathbb{N},\rho} \;:\; \forall n^{\mathbb{N}}\, \exists y^\rho\, A(n, y) \to \exists f\, \forall n\, A(n, f(n)).$$

The realizer for $(\mathsf{AC}^{\mathbb{N},\rho})^N$ is similar to the one for $(\mathsf{DC}^\rho)^N$, but technically simpler, so that the essential idea underlying the construction is more visible. Moreover we only need the following special case of relativized quantifier-free bar induction, so-called relativized quantifier-free pointwise bar induction:

$$\mathsf{pBI\text{-}QF} \;:\; \forall \alpha \in S\, \exists n\, P(\overline{\alpha}n) \wedge \forall s \in S\, (\forall x\, [S(x, |s|) \to P(s*x)] \to P(s)) \to P(\langle\,\rangle),$$

where $S(x, n)$ is arbitrary, $P(s)$ is quantifier-free, and $\alpha \in S$, $s \in S$ are shorthands for $\forall n\, S(\alpha(n), n)$ and $\forall i < |s|\, S(s_i, i)$, respectively. The principles of relativized quantifier-free bar induction, respectively pointwise bar induction, are similar to Luckhardt's general bar induction over species for quantifier-free formulas, $(\mathrm{aBI})^\rho_{\mathrm{D}}$, respectively higher bar induction over species, $(\mathrm{hBI})^\rho_{\mathrm{D}}$ ([124], page 144).

The negative translation of $\mathsf{AC}^{\mathbb{N},\rho}$ is

[2] We call any n such that $\forall \beta\, (\overline{\alpha}n = \overline{\beta}n \to F(\alpha) = F(\beta))$ a point of continuity of F at α.

$$\forall n\,(\forall y\,(A(n,y)^N \to \bot) \to \bot) \to \forall f\,(\forall n\,A(n,f(n))^N \to \bot) \to \bot.$$

Following Spector [154] we reduce $(\mathsf{AC}^{\mathbb{N},\rho})^N$ to the double negation shift

$$\mathsf{DNS} \quad : \quad \forall n\,((B(n) \to \bot) \to \bot) \to (\forall n\,B(n) \to \bot) \to \bot$$

observing that $\mathsf{AC}^{\mathbb{N},\rho} + \mathsf{DNS} \vdash_m (\mathsf{AC}^{\mathbb{N},\rho})^N$, where DNS is used with the formula $B(n) :\equiv \exists y\,A(n,y)^N$ [3]. Therefore it suffices to show that this instance of DNS is realizable. The following lemma, whose proof is trivial, is necessary to see that the weak form (8.4) of modified bar recursion suffices to realize $\mathsf{AC}^{\mathbb{N},\rho}$ and DC^{ρ}.

Lemma 8.1 *Let B be a formula such that all of its atomic subformulas occur in negated form. Then there is a closed term H such that $\forall \vec{z}\,H$ mr^{c} $(\bot \to B)$ is provable (in minimal logic), where $\vec{z}$ are the free variables of B (it is important here that H is closed, i.p. does not depend on $\vec{z}$).*

Note that the formula $B(n) :\equiv \exists y\,A(n,y)^N$ to which we apply DNS is of the form specified in Lemma 8.1.

Theorem 8.3 *The double negation shift* DNS *for a formula $B(n)$ is realizable using the weak form (8.4) of modified bar recursion provided $B(n)$ is of the form specified in Lemma 8.1.*

Proof. In order to realize the formula

$$\forall n((B(n) \to \bot) \to \bot) \to (\forall n B(n) \to \bot) \to \bot$$

we assume we are given realizers

$$Y^{\rho^\omega \to o}\ \mathsf{mr}^{\mathsf{c}}\ (\forall n B(n) \to \bot)$$

$$G^{\mathbb{N} \to (\rho \to o) \to o}\ \mathsf{mr}^{\mathsf{c}}\ \forall n((B(n) \to \bot) \to \bot)$$

and try to build a realizer for $\bot$. Using weak modified bar recursion (8.4) we define

$$\Psi(s) = Y(s\ @\ \lambda n.H(G(|s|, \lambda x^\rho.\Psi(s * x))))$$

where $H^{o \to \rho}$ is a closed term such that $\forall n\,H$ mr^{c} $(\bot \to B(n))$ is provable, according to Lemma 8.1. We set

$$S(x,n) :\equiv x\ \mathsf{mr}^{\mathsf{c}}\ B(n),$$

$$P(s) :\equiv \Psi(s)\ \mathsf{mr}^{\mathsf{c}}\ \bot,$$

and, by quantifier-free pointwise bar induction relativized to S, we show $P(\langle\,\rangle)$, i.e. $\Psi(\langle\,\rangle)$ mr^{c} $\bot$.

i) $\forall \alpha \in S\ \exists n\ P(\overline{\alpha}n)$. Let $\alpha \in S$, i.e. α mr^{c} $\forall n B(n)$. Let n be the point of continuity of Y at α, according to the continuity axiom. By assumption on Y, we get $\forall \beta\ (Y(\overline{\alpha}n\ @\ \beta)\ \mathsf{mr}^{\mathsf{c}}\ \bot)$, which implies $\Psi(\overline{\alpha}n)$ mr^{c} $\bot$.

ii) $\forall s \in S(\forall x\ [S(x,|s|) \to P(s * x)] \to P(s))$. Let $s \in S$ be fixed. Suppose $\forall x\ [S(x,|s|) \to P(s*x)]$, i.e. $\forall x\ [x\ \mathsf{mr}^{\mathsf{c}}\ B(|s|) \to \Psi(s*x)\ \mathsf{mr}^{\mathsf{c}}\ \bot]$, in other words

[3] The reduction is obvious because $(\mathsf{AC}^{\mathbb{N},\rho})^N$ is equivalent in minimal logic to $\forall n\,\neg\neg\exists y\,A(n,y)^N \to \neg\neg\exists f\,\forall n\,A(n,f(n))^N$.

$$\lambda x^\rho.\Psi(s * x) \ \mathsf{mr^c}\ (B(|s|) \to \bot).$$

Using the assumption on G we obtain

$$G(|s|, \lambda x^\rho.\Psi(s * x)) \ \mathsf{mr^c}\ \bot,$$

and from that, setting $w :=_\rho H(G(|s|, \lambda x^\rho.\Psi(s * x)))$, we obtain $w\ \mathsf{mr^c}\ B(n)$, for all n. Because $s \in S$ it follows that $s @ \lambda n.w\ \mathsf{mr^c}\ \forall n\, B(n)$ and therefore

$$Y(s @ \lambda n.w) \ \mathsf{mr^c}\ \bot.$$

Since $\Psi(s) = Y(s @ \lambda n.w)$ we have $P(s)$. □

As explained above Theorem 8.3 yields

Corollary 8.1 *The negative translation of the countable axiom of choice,* $(\mathsf{AC}^{\mathbb{N},\rho})^N$ *is realizable using the weak form (8.4) of modified bar recursion.*

8.3.3 Realizing $(\mathsf{DC}^\rho)^N$

With a similar but technically more involved construction we now prove

Theorem 8.4 *The negative translation of the axiom of dependent choice,* $(\mathsf{DC}^\rho)^N$, *is realizable using the weak form (8.4) of modified bar recursion.*

Proof. Let σ be the type of realizers of $A(n, x, y)^N$. Given x_0^ρ and realizers

$$G^{\mathbb{N}\to\rho\to(\rho\to\sigma\to o)\to o} \ \mathsf{mr^c}\ \forall n, x\, (\forall y\, (A(n, x, y)^N \to \bot) \to \bot),$$

$$Y^{\rho^\omega\to\sigma^\omega\to o} \ \mathsf{mr^c}\ \forall f\, (f(0) = x_0 \wedge \forall n\, A(n, f(n), f(n+1))^N \to \bot),$$

we have to construct a realizer of $\bot$. In the rest of this proof the variables β and t have the types $(\rho \times \sigma)^\omega$ and $(\rho \times \sigma)^*$ respectively. First we perform a trivial transformation on Y defining

$$\tilde{Y}^{(\rho\times\sigma)^\omega\to o}(\beta) := Y(x_0 * (\pi_0 \circ \beta), \pi_1 \circ \beta),$$

where π_0, π_1 are the left and right projection and $\circ$ is composition of functions. Using weak bar recursion (8.4) we now define

$$\Psi(t) = \tilde{Y}(t @ \lambda n.\pi(0^\rho, H(G(|t|, (x_0 * (\pi_0 \circ t))_{|t|}, \lambda y^\rho \lambda z^\sigma.\Psi(t * \pi(y, z)))))),$$

where $\forall n, x, y\, H\ \ \mathsf{mr^c}\ (\bot \to A(n, x, y)^N)$ according to Lemma 8.1, $\pi(.,.)$ is pairing, and $\pi_0 \circ t := \langle \pi_0(t_0), \dots, \pi_0(t_{|t|-1}) \rangle$ (hence $(\pi_0 \circ t)_i = \pi_0(t_i)$ for $i < |t|$). We define predicates

$$S(t) :\equiv \forall i < |t|\ \ (\pi_1(t_i)\ \mathsf{mr^c}\ A(i, (\langle x_0 \rangle * (\pi_0 \circ t))_i, (\pi_0 \circ t)_i)^N)$$

$$P(t) :\equiv \Psi(t)\ \mathsf{mr^c}\ \bot.$$

We show $P(\langle\ \rangle)$ by quantifier-free bar induction relativized to S. Obviously $S(\langle\,\rangle)$ holds.

i) $\forall \beta \in S\, \exists n\, P(\overline{\beta} n)$. Let $\beta \in S$. Set $f^{\rho^\omega} := \langle x_0 \rangle * (\pi_0 \circ \beta)$ and $\gamma^{\sigma^\omega} := \pi_1 \circ \beta$. Then $f(0) = x_0$ and $\forall n\, \gamma(n)\ \mathsf{mr^c}\ A(n, f(n), f(n+1))^N$. Therefore $Y(f, \gamma)\ \mathsf{mr^c}\ \bot$. Let n be a point of continuity of $\tilde{Y}$ at β. Then

$$\Psi(\overline{\beta}n) = \tilde{Y}(\beta) = Y(f, \gamma)$$

and therefore $\Psi(\overline{\beta}n)$ $\mathsf{mr^c}$ $\bot$, i.e. $P(\overline{\beta}n)$.

ii) $\forall t \in S\, (\forall q^{\rho\times\sigma}[S(t * q) \to P(t * q)] \to P(t))$. Let $t \in S$ where, say, $t = \langle \pi(x_1, z_0), \ldots, \pi(x_n, z_{n-1})\rangle$. Assume further $\forall q\, [S(t * q) \to P(t * q)]$, i.e.

$$\forall x_{n+1}, z_n\, [\forall i \leq n\, z_i\ \mathsf{mr^c}\ A(i, x_i, x_{i+1})^N \to \Psi(\langle \pi(x_1, z_0), \ldots, \pi(x_{n+1}, z_n)\rangle)\ \mathsf{mr^c}\ \bot]$$

Because $t \in S$ it follows that

$$\forall x_{n+1}, z_n\, [z_n\ \mathsf{mr^c}\ A(n, x_n, x_{n+1})^N \to \Psi(\langle \pi(x_1, z_0), \ldots, \pi(x_{n+1}, z_n)\rangle)\ \mathsf{mr^c}\ \bot]$$

i.e.

$$\lambda y \lambda z.\Psi(t * \pi(y, z))\ \mathsf{mr^c}\ \forall y\, (A(n, x_n, y)^N \to \bot).$$

By the assumption on G it follows $G(n, x_n, \lambda y\lambda z.\Psi(t * \pi(y, z)))$ $\mathsf{mr^c}$ $\bot$. Hence, for $w :=_\sigma H(G(n, x_n, \lambda y\lambda z.\Psi(t{*}\pi(y, z))))$, we have $\forall n, x, x'\, (w\ \mathsf{mr^c}\ A(n, x, x')^N)$. Now we set $f^{\rho^\omega} := \langle x_0, x_1, \ldots, x_n\rangle$ @ 0^ρ and $\gamma^{\sigma^\omega} := \langle z_0, \ldots, z_{n-1}\rangle$ @ w. Then $\forall n\, \gamma(n)$ $\mathsf{mr^c}$ $A(n, f(n), f(n+1))^N$ and therefore $Y(f, \gamma)$ $\mathsf{mr^c}$ $\bot$. But, because $x_n = (x_0 * (\pi_0 \circ t))_{|t|}$ we have

$$\Psi(t) = \tilde{Y}(t\ @\ \pi(0^\rho, a)) = Y(f, \gamma).$$

Hence $\Psi(t)$ $\mathsf{mr^c}$ $\bot$, i.e. $P(t)$. □

8.4 Bar recursion and the fan functional

A functional $\mathsf{FAN}^{(\mathbb{N}^\omega\to o)\to\mathbb{N}}$ is called *fan functional* if it computes[4] a modulus of uniform continuity for every continuous functional $Y^{\mathbb{N}^\omega\to o}$ restricted to infinite $0,1$-sequences, i.e. if FAN satisfies

$$\forall Y\, \forall \alpha, \beta\ \leq \lambda x.1(\overline{\alpha}(\mathsf{FAN}(Y)) = \overline{\beta}(\mathsf{FAN}(Y)) \to Y\alpha =_o Y\beta).$$

A recursive algorithm for $\mathsf{FAN}(Y)$ that was given in [12] and [133] uses two procedures,

$$\Phi(s^{\mathbb{N}^*}, v^o) =_{\mathbb{N}^\omega} s\ @\ [\text{if } Y(\Phi(s * 0, v)) \neq v \text{ then } \Phi(s * 0, v) \text{ else } \Phi(s * 1, v)] \quad (8.5)$$

$$\Psi(Y, s) =_{\mathbb{N}} \begin{cases} 0 & \text{if } Y(\alpha) = Y(s\ @\ \lambda k.0), \\ & \text{where } \alpha = \Phi(s, Y(s\ @\ \lambda k.0)) \\ 1 + \max\{\Psi(Y, s * 0), \Psi(Y, s * 1)\} & \text{otherwise.} \end{cases} \quad (8.6)$$

The first functional, $\Phi(s, v)$, returns an infinite path α having s as a prefix, such that $Y(s\ @\ \alpha) \neq v$, if such a path exists, and returns s extended by $\lambda x.1$, otherwise, i.e. if Y is constant v on all paths extending s. The second functional, $\Psi(Y, s)$, returns the least point of uniform continuity for Y on all extension of s. Therefore, a fan functional can be defined as $\mathsf{FAN}(Y) := \Psi(Y, \langle\ \rangle)$. A more formal proof that $\lambda Y.\Psi(Y, \langle\ \rangle)$ is indeed a fan functional can be found in [12] and [133] [5].

[4] [Here we use the term "computable" as an abreviation for computable in the partial total continous functionals $\widehat{\mathcal{C}}$. Therefore, types such as $\mathbb{N}^\omega \to o$ should actually be understood as partial continuous functionals from $\mathbb{N}_\bot \rightharpoonup \mathbb{N}_\bot$ to $o_\bot$.]

[5] The authors were informed that Robin Gandy knew a recursive definition of the fan functional in $\widehat{\mathcal{C}}$ already around 1973.

Theorem 8.5 *The functional* FAN *can be defined using bar recursions (8.3) and (8.2) together.*

Before we give the proof of the theorem we prove two lemmas.

Lemma 8.2 *Modified bar recursion (8.3) is equivalent to*

$$\Phi(s^{\rho^*}) =_o Y(s \mathbin{@} H(s, \lambda t^{\rho^*} \lambda x^\rho.\Phi(s * t * x))) \tag{8.7}$$

and also to

$$\Phi(s^{\rho^*}) =_{\rho^\omega} s \mathbin{@} H(s, \lambda t^{\rho^*} \lambda x^\rho.Y^{\rho^\omega \to o}(\Phi(s * t * x))). \tag{8.8}$$

Proof. Obviously equation (8.7) subsumes modified bar recursion. It is also easy to see that equations (8.7) and (8.8) are equivalent: Given Φ satisfying (8.7) we define $\Phi'(s) := s \mathbin{@} H(s, \lambda t \lambda x.\Phi(s*t*x))$ which satisfies (8.8), provably by relativized bar induction. Conversely, if Φ' satisfies (8.8) then Φ defined by $\Phi(s) := Y(\Phi'(s))$ satisfies (8.7). Furthermore it is clear that we can replace the operation @ in each of the equations (8.3), (8.7) and (8.8) by $*$, i.e. we prefix with s instead of overwriting (see the definitions at the beginning of Section 8.2). Hence it suffices to show that we can define a functional Φ satisfying

$$\Phi(s^{\rho^*}) =_o Y(s * H(s, \lambda t^{\rho^*} \lambda x^\rho.\Phi(s * t * x))) \tag{8.9}$$

by modified bar recursion. To this end we will use equation (8.3) (where @ is replaced by $*$) at type ρ^*. We define $\mathsf{freeze}\colon \rho^* \to \rho^{**}$ and $\mathsf{melt}\colon \rho^{**} \to \rho^*$ by $\mathsf{freeze}(\langle x_0, \ldots, x_{n-1}\rangle) = \langle\langle x_0\rangle, \ldots, \langle x_{n-1}\rangle\rangle$, $\mathsf{melt}(\langle s_0, \ldots, s_{n-1}\rangle) = s_0 * \ldots * s_{n-1}$, so that $\mathsf{melt}(\mathsf{freeze}(s)) = s$. Given $Y^{\rho^\omega \to o}$ and $H^{\rho^* \to (\rho^* \times \rho \to o) \to \rho^\omega}$ we define using modified bar recursion (8.3)

$$\Psi(q) = Y(\mathsf{melt}(q) * H(\mathsf{melt}(q), \lambda t \lambda x.\Psi(q * (t * x)))).$$

By relativized bar induction one easily proves

$$\forall q, q' \, (\mathsf{melt}(q) = \mathsf{melt}(q') \to \Psi(q) = \Psi(q')),$$

which implies, again by relativized bar induction, that Φ, defined by $\Phi(s) := \Psi(\mathsf{freeze}(s))$, satisfies (8.9). □

Lemma 8.3 *Kohlenbach's bar recursion (8.2) is equivalent to*

$$\Phi(s) =_\tau \begin{cases} G(s) & \text{if } Y(s \mathbin{@} 0^{\rho^\omega}) =_o Y(s \mathbin{@} J(s)) \\ H(s, \lambda x^\rho.\Phi(s * x)) & \text{otherwise,} \end{cases} \tag{8.10}$$

where the new parameter J is of type $\rho^* \to \rho^\omega$ and, as usual, $\Phi(s)$ is shorthand for the more accurate $\Phi(Y, G, H, J, s)$.

Proof. Our proof is based on the proof of Theorem 3.66 in [88]. The fact that (8.2) can be defined from (8.10) is trivial. To define (8.10) from (8.2) one uses the following trick. For s^{ρ^*}, $s + (\dot{-})k$ denotes pointwise addition (cut-off subtraction) of appropriate type, and $\kappa(n) := n$, $\kappa(f^{\rho \to \sigma}) := \kappa(f(0^\rho))$, $\kappa(z^{\rho \times \sigma}) := \kappa(\pi_0(z))$, so $\kappa(x^\rho + 2) > 1$ and $\kappa(n^\rho) = n$. Define

$$\eta(\beta^{\rho^\omega})(n) := \begin{cases} \beta(n) \dot{-} 2 & \text{if } \kappa(\beta(n)) > 1 \\ J(\phi(\overline{\beta} n))(n) & \text{if } \kappa(\beta(n)) = 1 \\ 0 & \text{if } \kappa(\beta(n)) = 0, \end{cases}$$

where $\phi(s) := \langle s_0, \dots, s_{k-1} \rangle$ with $k < |s|$ minimal such that $\kappa(s_k) = 1$ (if $s = \langle\ \rangle$ then k is zero). Clearly

$$\eta((s+2) @ 0^{\rho^\omega}) = s @ 0^{\rho^\omega},$$

$$\eta((s+2) @ 1^{\rho^\omega}) = s @ J(s).$$

Now we can define using Kohlenbach's bar recursion (8.2)

$$\tilde{\Phi}(s) =_\tau \begin{cases} G(s \dot{-} 2) & \text{if } Y(\eta(s @ 0^{\rho^\omega})) = Y(\eta(s @ 1^\rho)) \\ H(s \dot{-} 2, \lambda x^\rho . \tilde{\Phi}(s * (x+2))) & \text{otherwise.} \end{cases}$$

Then clearly $\Phi(s) := \tilde{\Phi}(s+2)$ satisfies (8.10). □

Proofof Theorem 8.5. We show that procedures Φ and Ψ satisfying the equations (8.5) and (8.6) respectively can be defined using equations (8.3) and (8.2).

For defining the functional $\Phi(s, v)$ we use equation (8.8) of Lemma 8.2.

$$\Phi(s, v) =_{o^\omega} s @ H(s, v, \lambda t \lambda x . Y(\Phi(s * t * x, v)))$$

where H is defined by course of value primitive recursion as

$$H(s, v, F)(n) =_o \begin{cases} s_n & \mathsf{if}\ n < |s| \\ 0 & \mathsf{if}\ n \geq |s| \wedge F(c, 0) \neq v \\ 1 & \mathsf{if}\ n \geq |s| \wedge F(c, 0) = v, \end{cases}$$

with $c := \langle H(s, v, F)(|s|), \dots, H(s, v, F)(n-1) \rangle$. Clearly Φ satisfies equation (8.5) at all $n < |s|$. For $n \geq |s|$ we first observe that

$$\Phi(s, v)(n) =_o \begin{cases} 0 & \mathsf{if}\ Y(\Phi(s * c_{s,n} * 0, v)) \neq v \\ 1 & \mathsf{if}\ Y(\Phi(s * c_{s,n} * 0, v)) = v, \end{cases}$$

where $c_{s,n} := \langle \Phi(s, v)(|s|), \dots, \Phi(s, v)(n-1) \rangle$. Now if $Y(\Phi(s * 0, v)) \neq v$ then $\Phi(s, v)(|s|) = 0$ and therefore $s * c_{s,n} = s * 0 * c_{s*0,n}$. Hence $\Phi(s, v)(n) = \Phi(s * 0, v)(n)$ as required by (8.5). The case $Y(\Phi(s * 0, v)) = v$ is similar.

One immediately sees that a functional Ψ satisfying (8.6) can be defined from an instance of equation (8.10) using the functional Φ above. □

8.5 Modified bar recursion and the model $\mathcal{M}$

The model $\mathcal{M}$ ($= \bigcup \mathcal{M}_\rho$) of strongly majorizable functionals (introduced in [18] as a variation of Howard's majorizable functionals [65]) and the strongly majorizability relation $\geq^{\mathsf{m}}_\rho \subseteq \mathcal{M}_\rho \times \mathcal{M}_\rho$ are defined simultaneously by induction on types as follows [6]

[6] For simplicity, we only consider the base type $\mathbb{N}$ and functional types. Later we extend the definition of majorizability for types ρ^*.

$$n \ \geq^{\mathsf{m}}_{\mathbb{N}} \ m :\equiv n, m \in \mathbb{N} \wedge n \geq m, \qquad \mathcal{M}_{\mathbb{N}} :\equiv \mathbb{N},$$

$$F^* \ \geq^{\mathsf{m}}_{\rho \to \tau} \ F :\equiv F^*, F \in \mathcal{M}_\rho \to \mathcal{M}_\tau \wedge \\ \forall G^*, G \in \mathcal{M}_\rho \, [G^* \ \geq^{\mathsf{m}}_{\rho} \ G \to F^*G^* \ \geq^{\mathsf{m}}_{\tau} \ F^*G, FG],$$

$$\mathcal{M}_{\rho \to \tau} :\equiv \{F \in \mathcal{M}_\rho \to \mathcal{M}_\tau : \exists F^* \in \mathcal{M}_\rho \to \mathcal{M}_\tau \ F^* \ \geq^{\mathsf{m}}_{\rho \to \tau} \ F\}.$$

In the following, by "majorizable" we always mean "strongly majorizable". We often omit the type in the relation $\geq^{\mathsf{m}}_{\rho}$. We shall sometimes write "$F \colon \rho \to \sigma$" for "$F \in \mathcal{M}_{\rho \to \sigma}$" (as opposed to "$F \colon \mathcal{M}_\rho \to \mathcal{M}_\sigma$" which just means that F is a set-theoretic function from $\mathcal{M}_\rho$ to $\mathcal{M}_\sigma$, i.e. $F \in \mathcal{M}_\rho \to \mathcal{M}_\sigma$).

In [88] it is shown that the scheme of bar recursion (8.2) is provably not primitive recursively definable from (8.1), since (8.1) yields a well defined functional in the model of (strongly) majorizable functionals $\mathcal{M}$ (cf. [18]) and (8.2) does not. Equation (8.1), however, can be primitive recursively defined from (8.2) (cf. [88]). In [14] it is shown that a functional

$$\Phi \colon \mathcal{M}_{\rho^\omega \to \mathbb{N}} \times \mathcal{M}_{\rho^* \times (\rho \to \mathbb{N}) \to \rho^\omega} \times \mathcal{M}_{\rho^*} \to \mathcal{M}_{\mathbb{N}},$$

exists satisfying equation (8.3). We now show that any such Φ indeed lives in $\mathcal{M}$, i.e. we show that there is a functional Φ^* majorizing Φ.

Recall that for continuous functionals Y of type $\rho^\omega \to \mathbb{N}$ it is the case that from some initial segment of α the value of $Y(\alpha)$ is determined. For the majorizable functionals this does not hold, but a "weak continuity" property does hold. It says that a bound on the value of $Y(\alpha)$ can be determined from an initial segment of α. We prove this result in Lemma 8.5. This turned out to be an important tool for proving the main theorem of this section. For the rest of this section all variables (unless stated otherwise) are assumed to range over the type structure $\mathcal{M}$. We first recall from [18] the following lemma:

Lemma 8.4 ([18], 1.4, 1.5) *For $F_0, \ldots, F_n \colon \rho$ we define $\max^\rho \langle F_0, \ldots, F_n \rangle \colon \rho$, also written $\max_{i \leq n}{}^{\rho} F_i \colon \rho$, as*

$$\max_{i \leq n}{}^{\mathbb{N}} \, m_i := \max\{m_0, \ldots, m_n\},$$

$$\max_{i \leq n}{}^{\tau \to \rho} F_i := \lambda x^\tau . \max_{i \leq n}{}^{\rho} F_i(x),$$

and for α^{ρ^ω}, define $\alpha^+(n) := \max_{i \leq n}{}^{\rho} \alpha(i)$. Then,

$$\forall n (\alpha(n) \ \geq^{\mathsf{m}} \ \beta(n)) \to \alpha^+ \ \geq^{\mathsf{m}} \ \beta^+, \beta.$$

We also use pointwise addition in all types ρ, denoted $x +_\rho y$.

Lemma 8.5 (Weak continuity for $\mathcal{M}$) $\forall Y^{\rho^\omega \to \mathbb{N}}, \alpha \ \exists n^{\mathbb{N}} \ \forall \beta \in \overline{\alpha} n \, (Y(\beta) \leq n)$.

Proof. Let Y and α be fixed, $\alpha^* \geq^{\mathsf{m}} \alpha$ and $Y^* \geq^{\mathsf{m}} Y$. From the assumption

$$(*)\ \forall n\, \exists \beta \in \overline{\alpha}n(Y(\beta) > n)$$

we derive a contradiction. For any n, let β_n be the functional whose existence we are assuming in $(*)$. Let

$$\beta^*_n(i) := \begin{cases} 0^\rho & i < n \\ \beta_n(i)^* & i \geq n, \end{cases}$$

where $\beta_n(i)^*$ denotes some majorant of $\beta_n(i)$. Having defined the functional β^*_n we note two of its properties,

i) $\forall i < n(\beta^*_n(i) = 0^\rho)$,

ii) $(\alpha^* +_{\rho^\omega} \beta^*_n)^+ \geq^{\mathsf{m}} \beta_n$. (by Lemma 8.4)

Consider the functional $\hat{\alpha}$ defined as $\hat{\alpha}(n) := \alpha^*(n) +_\rho \sum_{i \in \mathbb{N}} \beta^*_i(n)$. Since at each point n only finitely many β^*_i are non-zero, α^* is well defined. Let $Y^*(\hat{\alpha}^+) = l$. Note that $\hat{\alpha}^+ \geq^{\mathsf{m}} \beta_i$, for all $i \in \mathbb{N}$, and from $(*)$ we should have $l < Y(\beta_l) \leq l$, a contradiction. □

We extend, for convenience, the definition of majorizability to finite sequences, i.e., for sequences $s^*, s \in \mathcal{M}^*_\rho$ we define

$$s^* \geq^{\mathsf{m}}_{\rho^*} s :\equiv |s^*| \geq |s| \wedge \forall i \leq j < |s^*|(s^*_j \geq^{\mathsf{m}} s^*_i \wedge (i < |s| \rightarrow s^*_j \geq^{\mathsf{m}} s_i)).$$

It is clear that for any sequence $s \in \mathcal{M}^*_\rho$ we can find an $s^* \in \mathcal{M}^*_\rho$ such that $s^* \geq^{\mathsf{m}} s$. Therefore, we define $\mathcal{M}_{\rho^*}$ as $\mathcal{M}\rho^*$. Majorizability for functionals involving the type ρ^* is extended accordingly, e.g., for $F^*, F \in \mathcal{M}_{\rho^*} \rightarrow \mathcal{M}_\mathbb{N}$

$$F^* \geq^{\mathsf{m}}_{\rho^* \rightarrow \mathbb{N}} F :\equiv \forall s^*, s \in \mathcal{M}_{\rho^*}\ (s^* \geq^{\mathsf{m}}_{\rho^*} s \rightarrow F^*(s^*) \geq F^*(s), F(s)).$$

Lemma 8.6 *Let s^* and s s.t. $|s^*| = |s|$ be fixed. If $s^* \geq^{\mathsf{m}} s$ then*

$$\forall \beta \in s \exists \beta^* \in s^*\ (\beta^* \geq^{\mathsf{m}} \beta).$$

Proof. Let s^*, s and $\beta \in s$ be fixed. Moreover, assume $|s^*| = |s| = n$ and $s^* \geq^{\mathsf{m}} s$. We define β^* recursively as

$$\beta^*(i) := \begin{cases} s^*_i & \text{if } i < n \\ \max^\rho(\overline{\beta^*}(i) * \beta(i)^*) & \text{otherwise,} \end{cases}$$

where $\beta(i)^*$ is some majorant of $\beta(i)$. First note that, for all i, $\beta^*(i) \geq^{\mathsf{m}} \beta(i)$. We show that $\beta^* \geq^{\mathsf{m}} \beta$. Let $k \geq i$.

If $k < n$ then $\beta^*(k) = s^*_k \geq^{\mathsf{m}} s^*_i \geq^{\mathsf{m}} s_i = \beta(i)$.

If $k \geq n$ then $\beta^*(k) = \max^\rho\{\max_{j<k}{}^\rho \beta^*(j), \beta(k)^*\} \geq^{\mathsf{m}} \beta^*(i) \geq^{\mathsf{m}} \beta(i)$. □

In the following we shall make use of two functionals Ω and Γ defined below. The functional Ω was first introduced in [88], 3.40.

Lemma 8.7 ([88], 3.41) *Define functionals* $\min^\rho$ *(from non-empty sets $X \subseteq \mathcal{M}_\rho$ to elements of $\mathcal{M}_\rho$) and $\Omega : \mathcal{M}_\rho \rightarrow \mathcal{M}_\rho$ as*

$$\min{}^{\mathbb{N}} X :\equiv \min X, \textit{ for } \emptyset \neq X \subseteq \mathbb{N},$$

$$\min{}^{\rho\to\tau} X :\equiv \lambda y^\rho . \min{}^\tau \{Fy : F \in X\},\ \textit{for } \emptyset \neq X \subseteq \mathcal{M}_{\rho\to\tau},$$

$$\Omega(F) :\equiv \min{}^\rho \{F^* : F^* \ \geq^{\mathsf{m}}\ F\}.$$

Then,

i) *For all* F, $\Omega(F) \ \geq^{\mathsf{m}}\ F$,

ii) $\Omega \ \geq^{\mathsf{m}}\ \Omega$. *(Therefore,* $\Omega \in \mathcal{M}$.*)*

Lemma 8.8 *Define* $\Gamma : \mathcal{M}_{\rho^\omega\to\mathbb{N}} \to (\mathcal{M}_{\rho^\omega} \to \mathcal{M}_{\mathbb{N}})$

$$\Gamma(Y)(\alpha) := \min n\ [\forall \beta \in \overline{\alpha}n(\Omega(Y)(\beta) \leq n)].$$

Then,

i) $\Gamma(Y) \ \geq^{\mathsf{m}}\ Y$ *(therefore* $\Gamma(Y) \in \mathcal{M}_{\rho^\omega\to\mathbb{N}}$*),*

ii) $\Gamma(Y)$ *is continuous and* $\Gamma(Y)(\alpha)$ *is a point of continuity for* $\Gamma(Y)$ *at* α*,*

iii) $\Gamma \ \geq^{\mathsf{m}}\ \Gamma$ *(therefore,* $\Gamma \in \mathcal{M}$*).*

Proof. First of all, we note that, by Lemma 8.5, the functional Γ is well defined. By Lemma 8.7 (i), $\Omega(Y) \ \geq^{\mathsf{m}}\ Y$.

i) Let $\alpha^* \ \geq^{\mathsf{m}}\ \alpha$. We have to show $\Gamma(Y)(\alpha^*) \geq \Gamma(Y)(\alpha), Y(\alpha)$. By the definition of $\Gamma(Y)$, and Lemma 8.7 (i), we have $\Gamma(Y)(\alpha^*) \geq \Omega(Y)(\alpha^*) \geq Y(\alpha)$. It is only left to show that $\Gamma(Y)(\alpha^*) \geq \Gamma(Y)(\alpha)$. Suppose that $n = \Gamma(Y)(\alpha^*) < \Gamma(Y)(\alpha) = m$. Note that there exists a $\beta \in \overline{\alpha}(m-1)$ such that $\Omega(Y)(\beta) \geq m$ (otherwise we get a contradiction to the minimality in the definition of $\Gamma(Y)$). But since $m > n$, by Lemma 8.6, there exists a $\beta^* \in \overline{\alpha^*}n$ such that $\beta^* \ \geq^{\mathsf{m}}\ \beta$. Therefore, $\Omega(Y)(\beta^*) \leq n < m \leq \Omega(Y)(\beta)$. But by Lemma 8.7 ($i$) also $\Omega(Y)(\beta^*) \geq \Omega(Y)(\beta)$, a contradiction.

ii) Let α be fixed and take $n = \Gamma(Y)(\alpha)$. Suppose there exists a $\beta \in \overline{\alpha}n$ such that $\Gamma(Y)(\beta) \neq n$. If $\Gamma(Y)(\beta) < n$ we get, since $\alpha \in \overline{\beta}n$, that $\Gamma(Y)(\alpha) < n$, a contradiction. Suppose $\Gamma(Y)(\beta) > n$. Since $\beta \in \overline{\alpha}n$ we have, $\forall \gamma \in \overline{\beta}n(\Omega(Y)(\gamma) \leq n)$, also a contradiction.

iii) Assume $Y^* \ \geq^{\mathsf{m}}\ Y$ and $\alpha^* \ \geq^{\mathsf{m}}\ \alpha$. We show $\Gamma(Y^*)(\alpha^*) \geq \Gamma(Y)(\alpha)$. By the self majorizability of $\Gamma(Y)$ we have $\Gamma(Y)(\alpha^*) \geq \Gamma(Y)(\alpha)$. We now show $\Gamma(Y^*)(\alpha^*) \geq \Gamma(Y)(\alpha^*)$. Let $n = \Gamma(Y^*)(\alpha^*)$ and suppose $m = \Gamma(Y)(\alpha^*) > n$. By the definition of $\Gamma(Y)$, there exists a $\beta \in \overline{\alpha^*}(m-1)$ s.t. $\Omega(Y)(\beta) \geq m$. But, since $m > n$, by Lemma 8.6, there exists a $\beta^* \in \overline{\alpha^*}n$ s.t. $\beta^* \ \geq^{\mathsf{m}}\ \beta$, and by Lemma 8.7 ($ii$), $\Omega(Y^*)(\beta^*) \geq m > n$, a contradiction. $\square$

Lemma 8.9 *Let* $Y^* \ \geq^{\mathsf{m}}\ Y$ *of type* $\rho^\omega \to \mathbb{N}$ *and* α *of type* ρ^ω *be fixed. Set* $n = \Gamma(Y^*)(\alpha)$. *If* $\overline{\alpha}n \ \geq^{\mathsf{m}}\ s$ *and* $|s| = n$ *then for all sequences* β *we have*

$$\Gamma(Y^*)(s \mathbin{@} \beta), \Gamma(Y)(s \mathbin{@} \beta), Y(s \mathbin{@} \beta) \leq n.$$

Proof. We prove just that $\Gamma(Y^*)(s \ @ \ \beta) \leq n$. The other two cases follow similarly. Suppose there exists a β such that $n < \Gamma(Y^*)(s \ @ \ \beta)$. Since $\overline{\alpha}n \ \geq^{\mathsf{m}}$ s, by Lemma 8.6, there exists a β^* such that $\overline{\alpha}n * \beta^* \ \geq^{\mathsf{m}} \ s \ @ \ \beta$. Therefore, by Lemma 8.8 (*iii*), we must have $n < \Gamma(Y^*)(\overline{\alpha}n * \beta^*)$. And by the fact that n is a point of continuity for $\Gamma(Y^*)$ on α we get $\Gamma(Y^*)(\overline{\alpha}n * \beta^*) = n$, a contradiction. □

We extend the $(\cdot)^+$ operator of Lemma 8.4 to functionals $F \colon \mathcal{M}_{\rho^*} \to \mathcal{M}_{\mathbb{N}}$ by

$$F^+ := \lambda s. \max_{s' \preceq s} F(s'),$$

where $s' \preceq s :\equiv |s'| \leq |s| \ \wedge \ \forall i < |s'| \ (s'_i = s_i)$.

Lemma 8.10 *Let F and G be of type $\mathcal{M}_{\rho^*} \to \mathcal{M}_{\mathbb{N}}$. If*

$$\forall s^*, s \ [s^* \ \geq^{\mathsf{m}} \ s \wedge |s^*| = |s| \to F(s^*) \geq F(s), G(s)]$$

then $F^+ \ \geq^{\mathsf{m}} \ G^+, G$.

Proof. Let $s^* \ \geq^{\mathsf{m}} \ s$ be fixed. For all prefixes t^* (of s^*) and t (of s) of the same length, by the assumption of the lemma, we have $F(t^*) \geq F(t), G(t)$. Therefore,

$$\max_{s' \preceq s^*} F(s') \geq \max_{s' \preceq s} F(s'), \max_{s' \preceq s} G(s').$$

Therefore, $F^+ \ \geq^{\mathsf{m}} \ G^+, G$. □

Theorem 8.6 *If Φ is a functional of type*

$$\mathcal{M}_{\rho^\omega \to \mathbb{N}} \times \mathcal{M}_{\rho^* \times (\rho \to \mathbb{N}) \to \rho^\omega} \times \mathcal{M}_{\rho^*} \to \mathcal{M}_{\mathbb{N}},$$

which for any given $Y, H, s \in \mathcal{M}$ (of appropriate types) satisfies equation (8.3), then $\Phi \in \mathcal{M}$.

Proof. Our proof is based on the proof of the main result of [18]. The idea is that, if Φ satisfies equation (8.3) then the functional

$$\Phi^* := \lambda Y, H.[\lambda s.\Phi(\hat{Y}, \hat{H}, s)]^+ \ \geq^{\mathsf{m}} \ \Phi,$$

where

$$\hat{Y}(\alpha) := \Gamma(Y)(\alpha^+) \text{ and}$$

$$\hat{H}(s, F) := H(s, \lambda x.F(\{x\}_s)),$$

and $\{x\}_s$ abbreviates $\max^\rho(s * x)$. Let $Y^* \ \geq^{\mathsf{m}} \ Y$ and $H^* \ \geq^{\mathsf{m}} \ H$ be fixed. For the rest of the proof $s^* \ \geq^{\mathsf{m}} \ s$ is a shorthand for $s^* \ \geq^{\mathsf{m}} \ s \ \wedge \ |s^*| = |s|$, i.e. majorizability is only considered for sequences of equal length. The fact that $\Phi^* \ \geq^{\mathsf{m}} \ \Phi$ follows from,

$$[\lambda s.\Phi(\hat{Y^*}, \hat{H^*}, s)]^+ \ \geq^{\mathsf{m}} \ [\lambda s.\Phi(\hat{Y}, \hat{H}, s)]^+, \lambda s.\Phi(Y, H, s),$$

which follows, by Lemma 8.10, from $\forall s^* \, P(s^*)$ where

$$P(s^*) := \forall s \, [s^* \ \geq^{\mathsf{m}} \ s \to \Phi(\hat{Y^*}, \hat{H}^*, s^*) \geq \Phi(\hat{Y^*}, \hat{H}^*, s), \Phi(\hat{Y}, \hat{H}, s), \Phi(Y, H, s)].$$

We prove $\forall s^* \, P(s^*)$ by bar induction:

i) $\forall \alpha \exists n \, P(\overline{\alpha} n)$. Let α be fixed and $n := \hat{Y^*}(\alpha) = \Gamma(Y^*)(\alpha^+)$. If $\overline{\alpha} n$ does not majorize any sequence s we are done. Let s be such that $\overline{\alpha} n \ \geq^{\mathsf{m}} \ s$. Note that $\overline{\alpha^+} n = \overline{(\overline{\alpha} n \mathbin{@} \beta)^+} n$, for all β. Therefore, by Lemma 8.8 (ii) and our assumption that Φ satisfies (8.3) we get $\Phi(\hat{Y^*}, \hat{H}^*, \overline{\alpha} n) = n$. Since $\overline{\alpha^+} n \ \geq^{\mathsf{m}} \ \overline{(s \mathbin{@} \beta)^+} n$ (for all β), by Lemma 8.9, we have $n \geq \Phi(\hat{Y^*}, \hat{H}^*, s), \Phi(\hat{Y}, \hat{H}, s), \Phi(Y, H, s)$.

ii) $\forall s^* (\forall x \, P(s^* * x) \to P(s^*))$. Let s^* be fixed. Assume that $\forall x \, P(s^* * x)$, i.e.

$$\forall x, s \, [s^* * x \ \geq^{\mathsf{m}} \ s \to \Phi(\hat{Y^*}, \hat{H}^*, s^* * x) \geq \Phi(\hat{Y^*}, \hat{H}^*, s), \Phi(\hat{Y}, \hat{H}, s), \Phi(Y, H, s)].$$

We derive $P(s^*)$. Note that if s^* does not majorize any sequence we are again done. Assume s is such that $s^* \ \geq^{\mathsf{m}} \ s$. If $x^* \ \geq^{\mathsf{m}} \ x$ then (by $\forall x \, P(s^* * x)$),

$$\underbrace{\Phi(\hat{Y^*}, \hat{H}^*, s^* * \{x^*\}_{s^*})}_{\equiv:\ \Phi_1(\{x^*\}_{s^*})} \geq$$

$$\underbrace{\Phi(\hat{Y^*}, \hat{H}^*, s * \{x\}_s)}_{\equiv:\ \Phi_2(\{x\}_s)}, \underbrace{\Phi(\hat{Y}, \hat{H}, s * \{x\}_s)}_{\equiv:\ \Phi_3(\{x\}_s)}, \underbrace{\Phi(Y, H, s * x)}_{\equiv:\ \Phi_4(x)}.$$

and also $\Phi_1(\{x^*\}_{s^*}) \geq \Phi_1(\{x\}_{s^*})$, which implies

$$\lambda x.\Phi_1(\{x\}_{s^*}) \ \geq^{\mathsf{m}} \ \lambda x.\Phi_2(\{x\}_s), \lambda x.\Phi_3(\{x\}_s), \lambda x.\Phi_4(x),$$

and by the definition of majorizability

$$\underbrace{H^*(s^*, \lambda x.\Phi_1(\{x\}_{s^*}))}_{\hat{H}^*(s^*, \lambda x.\Phi_1(x))} \ \geq^{\mathsf{m}}$$

$$\underbrace{H^*(s, \lambda x.\Phi_2(\{x\}_s))}_{\hat{H}^*(s, \lambda x.\Phi_2(x))}, \underbrace{H(s, \lambda x.\Phi_3(\{x\}_s))}_{\hat{H}(s, \lambda x.\Phi_3(x))}, H(s, \lambda x.\Phi_4(x)),$$

which implies

$$\begin{aligned} (s^* \mathbin{@} \hat{H}^*(s^*, \lambda x.\Phi_1(x)))^+ \quad \geq^{\mathsf{m}} \quad & (s \mathbin{@} \hat{H}^*(s, \lambda x.\Phi_2(x)))^+, \\ & (s \mathbin{@} \hat{H}(s, \lambda x.\Phi_3(x)))^+, \\ & s \mathbin{@} H(s, \lambda x.\Phi_4(x)). \end{aligned}$$

And finally, by Lemma 8.8 (i) and (iii),

$$\begin{aligned} (\Phi(\hat{Y^*}, \hat{H}^*, s^*) =) \ \hat{Y^*}(s^* \mathbin{@} \hat{H}^*(s^*, \lambda x.\Phi_1(x))) \geq \\ & \hat{Y^*}(s \mathbin{@} \hat{H}^*(s, \lambda x.\Phi_2(x))) \ (= \Phi(\hat{Y^*}, \hat{H}^*, s)), \\ & \hat{Y}(s \mathbin{@} \hat{H}(s, \lambda x.\Phi_3(x))) \ (= \Phi(\hat{Y}, \hat{H}, s)), \\ & Y(s \mathbin{@} H(s, \lambda x.\Phi_4(x))) \ (= \Phi(Y, H, s)). \end{aligned}$$

$\square$

In [14] we show that there exists a functional

$$\Phi : \mathcal{M}_{\rho^\omega \to \mathbb{N}} \times \mathcal{M}_{\rho^* \times (\rho \to \mathbb{N}) \to \rho^\omega} \times \mathcal{M}_{\rho^*} \to \mathcal{M}_{\mathbb{N}}$$

which, for parameters Y, H, s in $\mathcal{M}$, satisfies equation (8.3). Therefore, by the theorem above, we obtain that $\mathcal{M}$ satisfies modified bar recursion.

8.6 Conclusion

In this paper, we discussed modified bar recursion a variant of Spector's bar recursion that seems to be of some significance in proof theory and the theory and higher type recursion theory. Our main result was an abstract modified realizability interpretation (where realizability for falsity is uninterpreted) of the axioms of countable and dependent choice that can be used to extract programs from non-constructive proofs using these axioms. A similar result can be found in [11], however we claim that our solution is more accessible, since it builds on the well-known model of continuous functionals and the notion of modified realizability instead of an ad-hoc model and realizability as in [11]. It can be noted here that the weak form of modified bar recursion (8.4) used for the realization of dependent choice can be implemented quite efficiently by equipping the functional with an internal memory that records the value of $H(s, \lambda x.\Phi(s * x))$ and thus avoids its repeated computation. Such an optimization does not seem to be possible for the solution given in [11]. In order to make the realizability interpretation of dependent choice useful for program synthesis, it seems necessary to combine it with optimizations of the A-translation as development e.g. in [16] and [17]. To find out whether this is possible, will be a subject of further research.

Another important result was a definition of the fan functional using modified bar recursion and a version of bar recursion due to Kohlenbach, improving [12] and [133] where a PCF definition of the fan functional was given. In [152] this definition of the fan functional has been applied to give a purely functional algorithm for exact integration of real functions.

The paper concluded with some new results on the model $\mathcal{M}$ of strongly majorizable functionals, in particular, the fact that modified bar recursion exists in $\mathcal{M}$. In [14], further results on the relation between modified bar recursion and other bar recursive definitions can be found. One important result of [14] is that modified bar recursion defines Spector bar recursion primitive recursively and that the converse does not hold.

Acknowledgements. We would like to thank Ulrich Kohlenbach for pointing out some mistakes in an early formulation of Section 8.5, and for suggesting corrections.

Chapter 9

Polynomial-time Algorithms from Ineffective Proofs

Proc. of the Eighteenth Annual IEEE *Symposium on Logic in Computer Science* LICS'03, pages 128–137, IEEE Press, 2003.

Abstract

We present the first procedure for extracting polynomial-time realizers from ineffective proofs of Π^0_2-theorems in feasible analysis. By ineffective proof we mean a proof which involves the non-computational principle weak König's lemma WKL, and by feasible analysis we mean Cook and Urquhart's system CPV^ω plus quantifier-free choice $\mathsf{QF\text{-}AC}$. We shall also discuss the relation between the system $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}$ and Ferreira's base theory for feasible analysis BTFA, for which Π^0_2-conservation of WKL has been non-constructively proven. This paper treats the case of weak König's lemma for trees defined by Π^0_1-formulas. Illustrating the applicability of $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}$ extended with this form of weak König's lemma, we indicate how to formalize the proof of the Heine/Borel covering lemma in this system. The main techniques used in the paper are Gödel's functional interpretation and a novel form of binary bar recursion.

9.1 Introduction

With the aim of capturing the notion of *feasibly constructive proof*, Stephen Cook [37] introduced in 1975 the equational system of arithmetic PV (polynomially verifiable) whose definable terms are polynomial-time computable. Later, Samuel Buss [31] developed the subsystem of classical arithmetic S^1_2 and showed that the provably recursive functions of his system are polynomial-time computable. Buss [32] also defined an intuitionistic version of S^1_2, called IS^1_2, and an intricate variant of Kleene realizability to prove that every Π^0_2-theorem of IS^1_2 has a polynomial-time computable realizer. Having as one of the motivations to simplify Buss' proof, Cook and Urquhart [38] defined systems both extending PV to higher types, obtaining PV^ω, and extending PV with intuitionistic and classical logic, obtaining IPV and CPV. A combination of those two extensions gives the systems IPV^ω and CPV^ω. Those systems have the same property of IS^1_2 that the provably recursive functions are polynomial-time computable. Cook and Urquhart then developed variants of Kreisel's modified realizability and Gödel's functional interpretation for the system IPV^ω. The latter via negative translation applies also to CPV^ω. Given a proof of a Π^0_2-theorem of IPV^ω or CPV^ω, these interpretations provide a simple procedure for extracting from this proof a polynomial-time algorithm realizing the theorem.

The main contribution of this paper is to extend Cook and Urquhart's functional interpretation, via negative translation, of CPV^ω to include quantifier-free choice $\mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ and the non-computational principle weak König's lemma (for Π^0_1-definable trees). The interpretation uses a novel form of binary bar recursion. We also show that the type one terms of the system IPV^ω extended with this new form of bar recursion are polynomial-time computable. This gives a procedure for extracting polynomial-time realizers from proofs of Π^0_2-theorems in $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} + \Pi^0_1\text{-}\mathsf{WKL}^\omega$. [1]

Weak König's lemma WKL states that every infinite finitely branching tree has an infinite path. This principle relies on the existence of non-computable functions, in the sense that it does not hold in a model where all functions are recursive. The fact that weak König's lemma is in the core of numerous ineffective proofs in analysis was first proven by Harvey Friedman. In [50], Friedman defined[2] the subsystem of second order arithmetic RCA_0, which contains the usual axioms for successor, addition and multiplication; induction restricted to Σ^0_1-formulas and comprehension for recursively defined sets. Friedman showed that, although $\mathsf{RCA}_0 + \mathsf{WKL}$ is mathematically fairly strong (in the sense that various theorems of classical analysis can be formalized in it, cf. [153]), the consistency $\mathsf{RCA}_0 + \mathsf{WKL}$ can be reduced to the consistency of primitive recursive arithmetic PRA. Actually, $\mathsf{RCA}_0 + \mathsf{WKL}$ has precisely the same Π^0_2-theorems as PRA. This implies that if

$$\mathsf{RCA}_0 + \mathsf{WKL} \vdash \forall x \exists y A(x, y),$$

where A is quantifier-free, then there exists a *primitive recursive function* h such that for all x, $A(x, hx)$ is true.

Friedman's original proof that $\mathsf{RCA}_0 + \mathsf{WKL}$ is Π^0_2-conservative over PRA is based on non-constructive model-theoretic arguments. Therefore, it does not provide a procedure for extracting primitive recursive algorithms from proofs of Π^0_2-theorems in $\mathsf{RCA}_0 + \mathsf{WKL}$. Friedman's result was later extended by Harrington, who proved (also in an unpublished paper) Π^1_1-conservation of $\mathsf{RCA}_0 + \mathsf{WKL}$ over RCA_0. The first effective version of Friedman's result was given by Sieg [151] using cut-elimination, a Herbrand analysis and a simple form of Howard's majorizability for primitive recursive terms. In [89], a combination of Gödel's functional interpretation with Howard's hereditary majorizability for functionals in all finite types is developed to extract uniform bounds for $\forall\exists$-theorems in analysis from proofs based on various analytical principles includ-

[1]Since the formulation of weak König's lemma changes from the setting of second order arithmetic to the setting of finite types, we use the superscript ω for the latter.

[2]Friedman's original definition of RCA_0 differs slightly from the one commonly used today in "reverse mathematics" (cf. [153]).

ing WKL. In particular, [89] yields effective forms of extensions of Friedman's WKL-conservation result to higher types (cf. also [5], Theorem 7.1.1).

In 1985, Sieg [150] proposed the problem of finding mathematically significant subsystems of analysis whose class of provably recursive functions consists only of *computationally feasible* ones. Fernando Ferreira took up the challenge and in [48] defined the system BTFA (Base Theory for Feasible Analysis) whose provably recursive functions are precisely the polynomial-time computable functions[3]. As done by Harrington for RCA_0, Ferreira then showed that by adding WKL (for bounded formulas Σ^b_∞) to BTFA one does not get any new Π^1_1-theorems. This shows a nice correspondence with respect to WKL between the system RCA_0, on the level of primitive recursion, and BTFA, on the level of polynomial-time. This correspondence can be expressed informally as

$$\frac{\mathsf{RCA}_0}{\mathsf{RCA}_0 + \mathsf{WKL}} \sim \frac{\mathsf{BTFA}}{\mathsf{BTFA} + \Sigma^b_\infty\text{-}\mathsf{WKL}}.$$

The congruity between the two sides of the equation goes even further. Ferreira's proof of Π^1_1-conservation, as the fore-mentioned Friedman's proof, is also based on non-constructive model-theoretic arguments and does not give a procedure for extracting, from a proof

$$\mathsf{BTFA} + \Sigma^b_\infty\text{-}\mathsf{WKL} \vdash \forall x \exists y A(x, y),$$

where A is quantifier-free, a *polynomial-time function* h such that $A(x, hx)$ holds, for all x. Therefore, to the author's knowledge, this paper presents the first effective procedure for extracting polynomial-time realizers from proofs of Π^0_2-theorems involving WKL in feasible analysis (here meaning $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$). It is important to note, however, that Ferreira proved conservation of WKL for trees defined by formulas of the kind $\forall z T(w, z)$, T being a *bounded formula*. This paper treats the case where T is a *quantifier-free formula*.

The rest of the article is organized as follows. In Sections 9.2 and 9.3 we present the systems BTFA and $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$, in order to discuss the relation between them. The reader with knowledge on CPV^ω can start reading from Section 9.4 where we introduce the new form of bar recursion, which is going to be used in the interpretation of Π^0_1-WKL^ω. In Section 9.4 we also prove that this new bar recursion does not give rise to any new functions when added to IPV^ω. The functional interpretation of the negative translation of Π^0_1-WKL^ω is given in Section 9.5. For illustrating the applicability of this Π^0_1-form of weak König's lemma, in Section 9.6 we formalize the proof of Heine/Borel covering lemma in the system $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} + \Pi^0_1\text{-}\mathsf{WKL}^\omega$.

A functional interpretation of the negative translation of weak König's lemma, using a different form of binary bar recursion, had already been given by Howard [66]. Howard's proof, however, does not carry through to the feasible setting under consideration since it is based on exponential search. We comment further on that in Section 9.7.

[3]Kohlenbach [93] also developed a subsystem of analysis (including WKL) whose Π^0_2-theorems have polynomial *bounds*, i.e. if $\forall x \exists y A(x, y)$, A quantifier-free, is a theorem of the system, then there exists effectively a polynomial $p(x)$ such that $\forall x \exists y \leq p(x) A(x, y)$.

9.2 Preliminaries

In the following we shall assume some basic knowledge on negative translation and functional interpretation. For a smooth introduction to functional interpretation see [5]. We shall use Kuroda's variant of negative translation which places double negations after universal quantifiers and in front of the whole formula. As shown in [124], the different variations of negative translation are over intuitionistic logic equivalent. The negative translation of a formula A will be denoted by A^N.

The *finite types* are defined inductively as follows: $\mathbb{N}$ is a finite type, and if ρ and σ are finite types then $\rho \to \sigma$ is also a finite type. We shall write $\ldots : \rho$ to denote that term $\ldots$ has type ρ.

The two feasible subsystems of analysis discussed here, BTFA and $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$, have two main differences. Firstly, BTFA is based on second order logic, and therefore, has variables and quantifiers for sets, whereas, the theory $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ is based on the language of functionals of all finite types, and therefore, has variables for each finite type. The second main difference is that the standard model of BTFA is based on finite 0-1 sequences $\mathbb{W}$, while $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ has standard model based on the natural number $\mathbb{N}$ (which we shall confuse with the basic finite type). We shall in this paper define both theories and discuss briefly the relation between them.

In a feasible setting, where the length of the representation matters, it is often useful to work with 0-1 sequences as basic elements. Therefore, when dealing with CPV^ω we shall view natural numbers as finite sequences of 0-1, via their binary expansion. Given a number x we shall denote the i-th bit of the binary expansion of x by $x(i)$. We often write $x0$ instead of $2x$, and $x1$ instead of $2x + 1$. In general, given a sequence of bits $b_n, \ldots, b_0 \in \{0, 1\}$ (with $b_n = 1$) we shall write $b_n \ldots b_0$ for the natural number having such binary expansion. Moreover, we write 1^n for the sequence of n bits 1 and we use $|x|$ for the length[4] of the binary expansion of x, i.e. $\lceil \log_2(x + 1) \rceil$.

In Section 9.2.1 (on BTFA) we shall talk about three relations on binary words: $x \subseteq y$ for x being a prefix of y; $x \trianglelefteq y$ for $|x|$ being less than or equal to $|y|$, and $x \subseteq^* y$ saying that x is a subword of y, i.e. if there exists a z such that $zx \subseteq y$. When treating CPV^ω we use $x \le y$ for x being a number smaller than or equal to y, and $x \preceq y$ for saying that the binary expansion of x is a prefix of the binary expansion of y. In both systems only the first relation is a primitive symbol, the others are definable relations. Based on those relations, in this paper the reader shall encounter three sorts of quantifiers:

- *unbounded quantifiers*: $Qx(\ldots)$,
- *bounded quantifiers*: $Qx \trianglelefteq t(\ldots)$ in BTFA and $Qx \le t(\ldots)$ in CPV^ω, and
- *sharply bounded quantifiers*: $Qx \subseteq t(\ldots), Qx \subseteq^* t(\ldots)$ in BTFA and $Qx \preceq t(\ldots), Qx \le |t|(\ldots)$ in CPV^ω.

[4]Although the function $|\cdot|$ is not a basic symbol in either systems BTFA or CPV^ω, it is easily definable and we shall use it freely.

Informally, bounded quantifiers correspond to an exponential search, while sharply bounded quantifiers correspond to linear or quadratic search. A formula is Π^0_1 (resp. Π^0_2) if it is of the form $\forall x A(x)$ (resp. $\forall x \exists y A(x,y)$), where A is a quantifier-free formula. While in stronger systems, such as RCA_0, a quantifier-free formula is one not containing unbounded quantifiers, in the feasible setting a *quantifier-free* formula is one containing only sharply bounded quantifiers.

Notice that, via paring, formulas of the kind $\forall x \exists y A(x,y)$, with A being quantifier-free, are as general as when A is a Σ^0_1 formula.

9.2.1 The system BTFA

Ferreira's system BTFA [48] has as basis the first order theory Σ^b_1-NIA, whose standard model is the set of finite strings over $\{0,1\}$ denoted by $\mathbb{W}$. The language of Σ^b_1-NIA contains symbols ϵ, 0 and 1, function symbols

- $x \frown y$ for the concatenation of x with y (we usually omit $\frown$ and just write xy),
- $x \times y$ for the concatenation of x with itself $|y|$ times,

and a binary relation symbol $\subseteq$ for string prefix.

The class of *subword quantification-formulas* (sw.q.-formulas for short) is the smallest class of formulas closed under boolean operations and subword quantification, i.e. quantification of the form $Qx \subseteq^* t(\ldots)$, where the variable x does not occur in the term t. The class of *bounded formulas* Σ^b_∞ is the smallest class of formulas containing the sw.q.-formulas and closed under boolean operations and bounded quantification, i.e. quantification of the form $Qx \trianglelefteq t(\ldots)$, where the variable x does not occur in the term t. The class of formulas of the form $\exists x \trianglelefteq t\, A$, A being a sw.q.-formula, is denoted by Σ^b_1.

Besides fourteen basic axioms governing the behaviour of the non-logical symbols, Σ^b_1-NIA contains the *induction scheme* Σ^b_1-IND

$$A(\epsilon) \wedge \forall x(A(x) \to A(x0) \wedge A(x1)) \to \forall x A(x),$$

for $A \in \Sigma^b_1$. The theory Σ^b_1-NIA is equivalent, in a sense that could be made precise, to Buss' theory S^1_2 (cf. [31]), and therefore, has the property that every Π^0_2-theorem has a polynomial-time realizer. The second order theory BTFA is obtained from Σ^b_1-NIA by adding the *bounded collection principle* Σ^b_∞-BC

$$\forall x \trianglelefteq t \exists y A(x,y) \to \exists z \forall x \trianglelefteq t \exists y \trianglelefteq z A(x,y),$$

for $A \in \Sigma^b_\infty$, and *comprehension* Δ^0_1-CA

$$\forall x(\exists y A(x,y) \leftrightarrow \forall z \neg B(x,z)) \to \exists \mathcal{S} \forall x(x \in \mathcal{S} \leftrightarrow \exists y A(x,y)),$$

for $A, B \in \Sigma^b_1$.

Lemma 9.1 ([48]) *Let A be a bounded formula. If*

$$\mathsf{BTFA} \vdash \forall x \exists y A(x,y)$$

then $\Sigma_1^b\text{-}\mathsf{NIA} \vdash \forall x \exists y A(x,y)$.

In the feasible setting of second order arithmetic $\mathsf{WKL}(T)$ is formulated as

$$\mathsf{Tree}_\infty(T) \to \exists \mathcal{S}(\mathsf{Path}_\infty(\mathcal{S}) \wedge \forall w(w \in \mathcal{S} \to T(w)),$$

where $\mathcal{S}$ is a set variable, $\mathsf{Tree}_\infty(T)$ is defined as

$$\forall w, v(T(w) \wedge v \subseteq w \to T(v)) \wedge \forall y \exists w(|w| = |y| \wedge T(w)),$$

and $\mathsf{Path}_\infty(\mathcal{S})$ as

$$\mathsf{Tree}_\infty(w \in \mathcal{S}) \wedge \forall x, y \in \mathcal{S}(x \subseteq y \vee y \subseteq x).$$

If Φ is a class of formulas, we shall denote by Φ-WKL the principle $\mathsf{WKL}(T)$ for T restricted to the class Φ.

Using non-constructive model-theoretic arguments, Ferreira showed that BTFA extended with Σ_∞^b-WKL has the same $\forall\exists\Sigma_\infty^b$-theorems as Σ_1^b-NIA.

Theorem 9.1 ([48]) *Let A be a bounded formula. If*

$$\mathsf{BTFA} + \Sigma_\infty^b\text{-}\mathsf{WKL} \vdash \forall x \exists y A(x,y)$$

then $\Sigma_1^b\text{-}\mathsf{NIA} \vdash \forall x \exists y A(x,y)$.

As a corollary, one obtains that the provably recursive functions of $\mathsf{BTFA} + \Sigma_\infty^b$-$\mathsf{WKL}$ are polynomial-time computable.

Corollary 9.1 *Let A be quantifier-free. If*

$$\mathsf{BTFA} + \Sigma_\infty^b\text{-}\mathsf{WKL} \vdash \forall x \exists y A(x,y)$$

then there exists a polynomial-time computable function h such that $A(x, hx)$ holds, for all x.

The main result of this paper is an effective version of Corollary 9.1 for the system $\mathsf{CPV}^\omega + \mathsf{QF}\text{-}\mathsf{AC}^{\mathbb{N},\mathbb{N}} + \Pi_1^0\text{-}\mathsf{WKL}^\omega$. In the following section we present the system $\mathsf{CPV}^\omega + \mathsf{QF}\text{-}\mathsf{AC}^{\mathbb{N},\mathbb{N}}$ and we explain how it relates to BTFA.

9.3 The system $\mathsf{CPV}^\omega + \mathsf{QF}\text{-}\mathsf{AC}^{\mathbb{N},\mathbb{N}}$

The system CPV^ω [38] builds on the equational calculus PV^ω. The language of PV^ω contains a single constant symbol 0, for the number zero. The function symbols of PV^ω, with their intended interpretation, are

- $s_0(x), s_1(x)$ extends x to the right with the bit 0 and 1, respectively;
- $\mathsf{Parity}(x)$ returns 0 if the rightmost bit of x is 0;
- $\lfloor \frac{1}{2} x \rfloor$ chops off the rightmost bit of x;
- $\mathsf{Chop}(x, y)$ chops off $|y|$ bits from the right of x;

- $\mathsf{Pad}(x,y)$ appends $|y|$ zero bits to the right of x;
- $\mathsf{Smash}(x,y)$ returns the bit '1' followed by $|x|$ times $|y|$ zeros.
- $\mathsf{Cond}(x,y,z)$ returns y if x is zero and z otherwise.

PV^ω has infinitely many variables for each finite type. Unless stated otherwise, the variables x,y,z and w shall have type $\mathbb{N}$. PV^ω has also a recursor $\mathcal{R}$ of type

$$\mathbb{N} \to (\mathbb{N} \to \mathbb{N} \to \mathbb{N}) \to (\mathbb{N} \to \mathbb{N}) \to \mathbb{N} \to \mathbb{N}.$$

The terms of PV^ω are formed out of variables and function symbols as usually done in the typed λ-calculus. PV^ω contains only the predicate symbol $=$ for the basic type $\mathbb{N}$. The formulas of PV^ω consists of all equations $s = u$, where s and u are terms of type $\mathbb{N}$. The axioms of PV^ω are the defining equations for the function symbols listed above, the axiom for *higher type limited recursion on notation* HTLRN

$$\mathcal{R}(x,h,g,y) = \begin{cases} x & \text{if } y = 0 \\ g(y) & \text{if } |t| > |g(y)| \\ t & \text{otherwise,} \end{cases}$$

where t abbreviates $h(y, \mathcal{R}(x,h,g,\lfloor \frac{1}{2}y \rfloor))$, and further axioms for normalising λ-terms. Moreover, PV^ω has four rules $R1^\omega$–$R4^\omega$ governing the behaviour of the equality predicate and a rule for induction on notation (for further details see [38]).

The system IPV^ω is defined as follows. The terms of IPV^ω are those of PV^ω. The predicate symbols of IPV^ω are $=$ and $\leq$, for type $\mathbb{N}$ only. The atomic formulas are $s = u$ and $s \leq u$, where s and u are terms of type $\mathbb{N}$. The formulas of IPV^ω are built out of atomic formulas via logical connectives and quantifiers for each finite type. The logical axioms of IPV^ω are the usual ones for many-sorted intuitionistic predicate logic. The non-logical axioms of IPV^ω consist of all the theorems of PV^ω plus[5]

- $x \leq y \leftrightarrow \mathsf{Lessequ}(x,y) = 0$,
- $x = s_0\lfloor \frac{1}{2}x \rfloor \vee x = s_1\lfloor \frac{1}{2}x \rfloor$,
- $\mathsf{Cond}(x,a,b) = c \leftrightarrow (x = 0 \wedge a = c) \vee (\neg(x = 0) \wedge b = c)$,

and the *induction axiom* $\mathsf{PIND}^\omega(A)$

$$(A(0) \wedge \forall x(A(\lfloor \frac{1}{2}x \rfloor) \to A(x))) \to \forall x A(x),$$

where A is of the form[6] $\exists y \leq t\, (s = u)$ and all the free-variables of t have type $\mathbb{N}$.

[5] $\mathsf{Lessequ}(x,y)$ is a definable function of PV^ω which represents the characteristic function of the inequality predicate.

[6] Note that in IPV^ω, for each quantifier-free formulas $A(x)$ one can build a term s such that $\mathsf{IPV}^\omega \vdash A(x) \leftrightarrow sx = 0$.

The system CPV^ω is obtained from IPV^ω by adding all instances of the law of excluded middle $A \vee \neg A$.

In the following we shall make use of two further logical principles, namely, the *scheme of quantifier-free choice*

$$\mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} \ : \ \forall x \exists y A(x,y) \rightarrow \exists h \forall x A(x, hx),$$

and *Markov's principle*

$$\mathsf{MP} \ : \ \neg\neg\exists x A(x) \rightarrow \exists x A(x),$$

where in both cases A is a quantifier-free formula, and in the case of Markov's principle the variable x can be of arbitrary type. We shall use Markov's principle in connection with the negative translation of the system $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$.

As shown in [38], the system CPV^ω contains a set of *feasible coding functions*. Therefore, one can for instance replace a sequence of quantifiers of the same kind by a singe quantifier. For simplicity, we shall state results without making it explicit that tuples of quantifiers are allowed.

The next lemma is an extension of the negative translation of CPV^ω in $\mathsf{IPV}^\omega + \mathsf{MP}$, given in [38] (Lemma 10.3), to include quantifier-free choice.

Lemma 9.2 *The theory* $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ *has a negative translation in* $\mathsf{IPV}^\omega + \mathsf{MP} + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$*.*

Proof. In [38] (Lemma 10.3) it is shown that CPV^ω has a negative translation in $\mathsf{IPV}^\omega + \mathsf{MP}$. Therefore, we just need to show that $\mathsf{IPV}^\omega + \mathsf{MP} + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ proves

$$\forall x \neg\neg\exists y A(x,y) \rightarrow \neg\neg\exists h \forall x A(x, hx),$$

A quantifier-free. By MP we have that $\neg\neg\exists y A(x,y) \leftrightarrow \exists y A(x,y)$. Therefore, it is enough to show $\mathsf{IPV}^\omega + \mathsf{MP} + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ proves

$$\forall x \exists y A(x,y) \rightarrow \neg\neg\exists h \forall x A(x, hx),$$

which follows from $\mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ and

$$\exists h \forall x A(x, hx) \rightarrow \neg\neg\exists h \forall x A(x, hx). \quad \Box$$

Since the functional interpretation of MP and $\mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ are trivial, we obtain the following extension of Theorem 10.4 of [38].

Lemma 9.3 *Let A be a quantifier-free formula. If*

$$\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} \vdash \forall x \exists y A(x,y),$$

then from this proof one can extract a closed term t of type $\mathbb{N} \rightarrow \mathbb{N}$ of IPV^ω *such that* $\mathsf{IPV}^\omega \vdash \forall x A(x, tx)$*.*

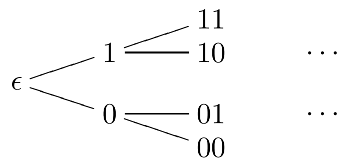
Moreover, since the terms of type $\mathbb{N} \rightarrow \mathbb{N}$ of IPV^ω denote polynomial-time computable functions, we get a procedure from extracting polynomial-time realizers from proofs of Π^0_2-theorems in $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$.

9.3.1 The system $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} + \Pi^0_1\text{-}\mathsf{WKL}^\omega$

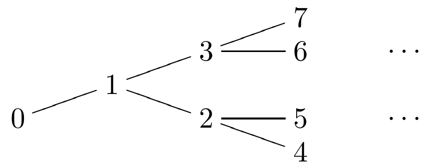
As we have mentioned, the theory BTFA has as standard model the set of finite 0-1 sequences $\mathbb{W}$. This setting is particularly convenient for working with weak König's lemma, since the prefix relation $\subseteq$ is one of the primitives of the system. The system CPV^ω, however, has the natural numbers as its standard model. Therefore, based on the bijective feasible mapping η (which assigns 0 to ϵ and positive numbers to their binary expansion) between natural number and the set of strings $1\{0,1\}^* \cup \{\epsilon\}$, we define the prefix relation $\preceq$ in CPV^ω as

$$x \preceq y :\equiv \eta(x) \subseteq \eta(y),$$

where x, y are numbers. The prefix relation $\subseteq$ in $\mathbb{W}$ is a partial order which can be depicted as



Notice that the binary words of the form $0\{0,1\}^*$ are not valid binary representation of any natural number. Therefore, under the mapping η, in $\mathbb{N}$ the prefix relation $\preceq$ gives rise to the partial order



A predicate T on numbers is said to define a *tree* if it is closed under the prefix relation $\preceq$, i.e. whenever $T(w)$ holds and $v \preceq w$ then $T(v)$ also holds. Formally

$$\mathsf{Tree}(T) :\equiv \forall w, v(T(w) \wedge v \preceq w \to T(v)).$$

A function $f : \mathbb{N} \to \mathbb{N}$ is an *infinite path* if $f(y) \in \{0,1\}$, for all y, and $f(0) = 1$, i.e.[7]

$$\mathsf{Path}(f) :\equiv \forall y(f(y) \leq 1) \wedge f(0) = 1.$$

We say that an infinite path f *belongs to a tree* T if every initial segment of f belongs to T, i.e. $\forall y T(\overline{f}y)$, where for a given path f, the function $\overline{f} : \mathbb{N} \to \mathbb{N}$ is defined as

$$\overline{f}(y) = \begin{cases} 0 & \text{if } y = 0 \\ f(0)f(1)\ldots f(1^{|y|-1}) & \text{otherwise.} \end{cases}$$

[7] We require that $f(0) = 1$ since we shall view initial segments of f as numbers, and finite 0-1 sequences of the form $0\{0,1\}^*$ do not correspond to valid natural numbers.

Therefore, in the feasible setting of finite types weak König's lemma for a predicate T is expressed as

$$\mathsf{Tree}(T) \wedge \forall y \exists w(|w| = |y| \wedge T(w)) \to \exists f(\mathsf{Path}(f) \wedge \forall y T(\overline{f}(y))).$$

The predicates Tree and Path are actually inessential (cf. [89]). Via the feasible transformation

$$T^t(w) :\equiv \forall v \preceq w T(v),$$

we can make an arbitrary predicate T into a tree T^t. The transformation is such that if $T(w)$ is already a tree, then $T^t(w)$ holds iff $T(w)$ holds. Moreover, via the transformation

$$f^p(y) = \begin{cases} 1 & \mathsf{if}\ y = 0 \\ \mathsf{Parity}(f(y)) & \mathsf{otherwise}, \end{cases}$$

we can make an arbitrary function $f : \mathbb{N} \to \mathbb{N}$ into an infinite path f^p. Again, if f is already a path then $f^p(y) = f(y)$, for all y. Using these transformations, weak König's lemma $\mathsf{WKL}^\omega(T)$ can be stated as[8]

$$\forall y \exists w(|w| = |y| \wedge T^t(w)) \to \exists f \forall y T^t(\overline{f^p}(y)). \tag{9.1}$$

Since the transformation f^p allows for quantification over infinite paths, in the following we take f as a meta-variable for infinite paths, and omit the transformation f^p.

In order to carry out the functional interpretation of $\mathsf{WKL}^\omega(T)$ it will be particularly convenient to treat it as an axiom (rather than an axiom schema)

$$\mathsf{WKL}^\omega \ : \ \forall g \, \mathsf{WKL}^\omega(gw = 0).$$

The Π^0_1-form of weak König's lemma is then stated as

$$\Pi^0_1\text{-}\mathsf{WKL}^\omega \ : \ \forall g \, \mathsf{WKL}^\omega(\forall z(gwz = 0)).$$

We shall use the superscript ω to differentiate between Ferreira's and our formulation of weak König's lemma.

9.3.2 BTFA versus $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$

In the system CPV^ω, using limited recursion on notation, sharply bounded quantifiers can be absorbed by quantifier-free matrices. Therefore, for A quantifier-free, the subword quantification of BTFA (which is definable in CPV^ω) can be also absorbed by two applications of recursion, since $Qx \subseteq^* tA(x)$ can be rewritten as

$$Qy \preceq tQx \preceq yA(\mathsf{Interv}(x, y, t)),$$

[8]Our definition of weak König's lemma (based on higher type functionals) is equivalent to Ferreira's definition (based on sets). One can define a feasible functional which given the characteristic function of a path $\mathcal{S}$ produces a path f, and vice-versa.

where the feasible function $\mathsf{Interv}(x, y, z)$ returns all the bits of z between $|x|$ and $|y|$. In this way, the sw.q.-formulas of BTFA correspond to quantifier-free formulas of CPV^ω. The predicate $\trianglelefteq$ can be easily defined using $\leq$, so that the formulas Σ^b_1 of BTFA correspond to formulas of the form $\exists x \leq t(s = u)$ in CPV^ω.

Moreover, the system $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ proves comprehension for Δ^0_1-formulas

$$\forall x(A(x) \leftrightarrow \neg B(x)) \to \exists h \forall x(hx = 0 \leftrightarrow A(x)),$$

where $A, B \in \Sigma^0_1$, which corresponds precisely to Δ^0_1-CA of BTFA; since the innermost bounded existential quantifiers of A and B, which are allowed in Δ^0_1-CA, can be actually absorbed by the unbounded existential quantifier.

The system $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$, also proves the following weaker form of bounded collection

$$\forall x \preceq t \exists y A(x, y) \to \exists z \forall x \preceq t \exists y \trianglelefteq z A(x, y),$$

for $A \in \Sigma^0_1$, but does not seem to prove the more general Σ^b_∞-BC.

One advantage of $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ over BTFA is the availability of higher order functionals. In this way one can talk about transformation between numbers (objects of type $\mathbb{N}$), real numbers and continuous functions (objects of type $\mathbb{N} \to \mathbb{N}$) in a straightforward way, as opposed to using encodings with sets.

In Section 9.6 we shall illustrate how the system $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ can be used for practical applications by sketching the proof of Heine/Borel theorem in $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} + \Pi^0_1\text{-}\mathsf{WKL}^\omega$.

9.4 A simple form of (binary) bar recursion

Howard showed in [66] that a simpler form of Spector's [154] bar recursion was sufficient for giving a functional interpretation of the negative translation of weak König's lemma. Howard's proof, however, does not seem to be suitable for weak theories such as IPV^ω, since it makes essential use of exponential search (cf. Section 9.7). For our conservation result we shall add to the language of IPV^ω the constant (of binary bar recursion) $\mathcal{B}$ having type

$$((\mathbb{N} \to \mathbb{N}) \to \mathbb{N}) \to (\mathbb{N} \to \mathbb{N}) \to \mathbb{N} \to \mathbb{N},$$

and the axiom

$$\mathcal{B}(Y, W, z) = \begin{cases} z & \text{if } |Y\hat{w}_z| \leq |w_z| \text{or } |w_z| \neq |z| \\ \mathcal{B}(Y, W, z1) & \text{otherwise,} \end{cases} \tag{9.2}$$

where w_z abbreviates Wz, and for a given $w \in \mathbb{N}$, the function $\hat{w} : \mathbb{N} \to \{0, 1\}$ is defined as

$$\hat{w}(y) := \begin{cases} w(|y|) & \text{if } |y| < |w| \\ 0 & \text{otherwise.} \end{cases}$$

The function $\hat{w}$ denotes[9] the infinite 0-1 sequence obtained by extending the binary expansion of w with 0's. In order to make sure that $\hat{w}$ always represents an infinite path (as defined in Section 9.3.1), we need to consider the particular case $w = 0$, since $\hat{0}(0) = 0$. Therefore, we change slightly the definition of $\hat{w}$ and set $\hat{0} = \hat{1}$.

The main result of this paper is based on the fact that IPV^ω is closed under the "rule version" of (9.2), i.e. if Ψ is a closed term of type $\mathbb{N} \to (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$ and Φ a closed term of type $\mathbb{N} \to \mathbb{N} \to \mathbb{N}$ then there exists a closed term $t : \mathbb{N} \to \mathbb{N}$ such that $tx = \mathcal{B}(\Psi x, \Phi x, 0)$, for all x. In fact, even 0-1 oracles are allowed (cf. Lemma 9.6).

Remark 9.1 *Note that the functional $\mathcal{B}$ only applies the first argument Y to 0-1 functions. Therefore, if Y_1 and Y_2 coincide on all 0-1 functions then $\mathcal{B}(Y_1, W, z) = \mathcal{B}(Y_2, W, z)$, for all W and z.*

Notice that the functional $\mathcal{B}(Y, W, z)$ can also be viewed as the unbounded search

$$\min y \approx z1^n \, (|Y\hat{w}_y| \leq |w_y| \vee |w_y| \neq |y|), \tag{9.3}$$

where w_y abbreviates Wy and $y \approx z1^n$ means that y has the same binary expansion as z followed by a finite number of ones. The functional $\mathcal{B}$ has a flavour of bar recursion since the sequences w_z, on the "hat transformation" of which the functional Y is applied, gets longer and longer as the recursion progresses.

For justifying this new form of binary bar recursion[10] (9.2) we can, for instance, assume *boundedness of functionals of type* $(\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$ *on 0-1 functions*

$$\mathsf{BND} \;:\; \forall Y \exists u \forall \alpha (\forall v(\alpha(v) \leq 1) \to |Y\alpha| \leq |u|),$$

which is a consequence of uniform continuity for functionals $Y : (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$ on the Cantor space. The idea is that, since $|z|$ keeps increasing in the recursion (9.2), either $|w_z| \neq |z|$ for some z, or the length of w_z also increases, and eventually reaches the length of the bound u. The condition $|Y\hat{w}_z| \leq |w_z|$ is then satisfied. We shall need BND in the verification of our interpretation of weak König's lemma (cf. Theorem 9.2).

For the rest of this section the variable x should be viewed as a sequence of variables of type $\mathbb{N}$ and α as a sequence of variables of type $\mathbb{N} \to \mathbb{N}$.

Lemma 9.4 ([71], Lemma 5.4) *For any closed term Ψ of type $\mathbb{N} \to (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$ of IPV^ω there exist constants c_1 and c_2 such that for any x and 0-1 functions α we have $|\Psi x \alpha| \leq |x|^{c_1} + c_2$.*

Using Lemma 9.4 one can show that IPV^ω is closed under the "rule version" of (9.2).

---

[9]Note that for the function $\hat{w} : \mathbb{N} \to \{0, 1\}$ only the length of the argument is considered. This shall often be used since in a feasible setting functions should be computed in polynomial-time on the length of the input. In fact, in those cases it is more convenient to use the tally part of $\mathbb{N}$ instead. We abstain from that in order to keep the basic setup of CPV^ω unchanged.

[10]Equivalently, for bounding the search (9.3).

Lemma 9.5 *Let Ψ be a closed term of type $\mathbb{N} \to (\mathbb{N} \to \mathbb{N}) \to (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$ and Φ a closed term of type $\mathbb{N} \to (\mathbb{N} \to \mathbb{N}) \to \mathbb{N} \to \mathbb{N}$ of IPV^ω. Then, there exists a closed term $t : \mathbb{N} \to (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$ such that for all x and for all 0-1 functions α, $tx\alpha = \mathcal{B}(\Psi x\alpha, \Phi x\alpha, 0)$.*

Proof. Let Ψ and Φ be fixed. We shall define t by limited recursion on notation. Let c_1 and c_2 be such that (cf. Lemma 9.4) for all 0-1 valued function β, $|\Psi x\alpha\beta| \leq |x|^{c_1} + c_2$. For a given x, let d_x denote the number $1^{|x|^{c_1}+c_2}$, then $|d_x| = |x|^{c_1} + c_2$. We then define two functions

$$h_{x,\alpha}(y, z) := \begin{cases} v & \text{if } |\Psi x\alpha\hat{w}_v| \leq |w_v| \text{ or } |w_v| \neq v \\ z & \text{otherwise,} \end{cases}$$

where v abbreviates $\mathsf{Chop}(d_x, y)$ and w_v abbreviates $\Phi x\alpha v$; and $g_x(y) := d_x$, i.e. g_x is a constant function with value d_x. Finally, we define $tx\alpha := \mathcal{R}(0, h_{x,\alpha}, g_x, d_x)$. $\Box$

The following lemma shows that arbitrary terms of type $\mathbb{N} \to (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$ (on 0-1 functions) of $\mathcal{L}(\mathsf{IPV}^\omega) + \{\mathcal{B}\}$ denote polynomial-time computable functions with boolean oracles.

Lemma 9.6 *Let $t[x, \alpha]$ be a term of $\mathcal{L}(\mathsf{IPV}^\omega) \cup \{\mathcal{B}\}$ of type $\mathbb{N}$, having as only free-variables x and α, such that (for simplicity) $\mathcal{B}$ is always applied to zero on the third argument. Then, there exists a polynomial-time computable function h (with 0-1 oracle) such that for all input x and for all 0-1 oracles α, $h(x, \alpha) = t[x, \alpha]$.*

Proof. The proof follows closely the normalisation argument given in the proof of Proposition 4.2 in [97]. In the following we say *polynomial-time computable* for *polynomial-time computable with 0-1 oracle*. We start by carrying out all possible logical reductions on the term $\lambda x, \alpha.t[x, \alpha]$. We get a term $\lambda x, \alpha.t_1[x, \alpha]$ such that $t_1[x, \alpha]$ is of the form:

- 0 or x_i (x_i in the tuple x). We are done.

- $g(t_2[x, \alpha])$, where g is either one of α or a function symbol of IPV^ω. By induction there exists a polynomial-time computable h_2 such that for all inputs x and 0-1 oracles α, $h_2(x, \alpha) = t_2[x, \alpha]$. Hence, for all inputs x and 0-1 oracles α, $h(x, \alpha) := g(h_2(x, \alpha))$ does the job.

- $\mathcal{R}(t_2[x, \alpha], t_3[x, \alpha], t_4[x, \alpha], t_5[x, \alpha])$. The terms $t_2[x, \alpha]$ and $t_5[x, \alpha]$ are again type $\mathbb{N}$, and by induction there are polynomial-time computable functions $h_2(x, \alpha)$ and $h_5(x, \alpha)$ which coincide with $t_2[x, \alpha]$ and $t_5[x, \alpha]$ on all inputs x and 0-1 oracles α. The terms $t_3[x, \alpha]$ and $t_4[x, \alpha]$ are of type $\mathbb{N} \to \mathbb{N} \to \mathbb{N}$ and $\mathbb{N} \to \mathbb{N}$ respectively. We therefore add an extra variables y and z to bring them to type $\mathbb{N}$. By induction there are polynomial-time computable functions $h_3(x, y, z, \alpha)$ and $h_4(x, y, \alpha)$ which coincide with $t_3[x, \alpha]yz$ and $t_4[x, \alpha]y$ on all inputs x, y, z and 0-1 oracles α.

Then, for all inputs x and 0-1 oracles α the polynomial-time computable function

$$h(x,\alpha) := \mathcal{R}(h_2 x\alpha, \lambda y, z.h_3 xyz\alpha, \lambda y.h_4 xy\alpha, h_5 x\alpha)$$

does the job.

- $\mathcal{B}(\Psi_2[x,\alpha], t_3[x,\alpha], 0)$. The term $\Psi_2[x,\alpha]$ is of type $(\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$. Let β be a variables of type $\mathbb{N} \to \mathbb{N}$. By induction there exists a polynomial-time computable h_2 such that for all inputs x and 0-1 oracles α, β, $h_2(x,\alpha,\beta) = \Psi_2[x,\alpha]\beta$. The term $t_3[x,\alpha]$ is of type $\mathbb{N} \to \mathbb{N}$. Adding an extra variable to bring it to type $\mathbb{N}$ we obtain, by induction hypothesis, that there exists a polynomial-time computable h_3 such that for all inputs x, y and 0-1 oracles $\alpha, h_3(x,y,\alpha) = t_3[x,\alpha]y$. By Lemma 9.5 and Remark 9.1, there exists a polynomial-time computable h such that for all inputs x and 0-1 oracles α

$$\begin{aligned} h(x,\alpha) &= \mathcal{B}(\lambda\beta.h_2(x,\alpha,\beta), \lambda y.h_3(x,y,\alpha), 0) \\ &= \mathcal{B}(\Psi_2[x,\alpha], t_3[x,\alpha], 0) \quad \Box \end{aligned}$$

9.5 Interpreting Π^0_1-WKL^ω

We shall now present the functional interpretation (via negative translation) of $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} + \Pi^0_1\text{-}\mathsf{WKL}^\omega$ in the system IPV^ω extended with a constant symbol $\mathcal{B}$, BND and the axiom (9.2).

Theorem 9.2 *The theory* $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} + \Pi^0_1\text{-}\mathsf{WKL}^\omega$ *has a functional interpretation (via negative translation) in* $\mathsf{IPV}^\omega + \mathsf{BND} + (9.2)$.

Proof. By Lemma 9.2, we just need to show that

$$\mathsf{IPV}^\omega + \mathsf{MP} + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} + (\forall g\, \mathsf{WKL}^\omega(\forall z(gwz = 0)))^N$$

has a functional interpretation in IPV^ω+BND+(9.2). The functional interpretations of MP and $\mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}}$ are trivial. Let $T(w,z)$ abbreviate $\forall v \preceq w(gvz = 0)$. The negative translation of $\forall g\, \mathsf{WKL}^\omega(\forall z(gwz = 0))$ gives

$$\neg\neg\forall g\neg\neg(\forall y\neg\neg\exists w(|w| = |y| \wedge \forall z\neg\neg T(w,z)) \to \exists f\forall y, z\neg\neg T(\overline{f}(y), z)),$$

which is equivalent to

$$\forall g(\forall y\neg\neg\exists w(|w| = |y| \wedge \forall z T(w,z)) \to \neg\neg\exists f\forall y, z T(\overline{f}(y), z)),$$

Since we shall give realizers independently of the tree g we henceforth omit the quantifier over g. Then

$$\forall y\neg\neg\exists w(|w| = |y| \wedge \forall z T(w,z)) \to \neg\neg\exists f\forall y, z T(\overline{f}(y), z),$$

has the functional interpretation (in three steps)

$$\forall y, h\exists w(|w| = |y| \wedge T(w, hw)) \to \forall Y, Z\exists f T(\overline{f}(Yf), Zf),$$

$$\exists W \forall y, h(|Wyh| = |y| \wedge T(Wyh, h(Wyh))) \to \forall Y, Z \exists f T(\overline{f}(Yf), Yf),$$

$$\forall W, Y, Z \exists y, h, f\big(|Wyh| = |y| \wedge T(Wyh, h(Wyh)) \to T(\overline{f}(Yf), Zf)\big).$$

Uniformly in W, Y, Z we produce y, h and f satisfying

$$|Wyh| = |y| \wedge T(Wyh, h(Wyh)) \to T(\overline{f}(Yf), Zf).$$

Define $\boxed{h(w) := Z(\hat{w})}$ and let w_y abbreviate Wyh. Now, we need to produce y and f satisfying

$$|w_y| = |y| \wedge T(w_y, Z(\hat{w}_y)) \to T(\overline{f}(Yf), Zf).$$

Define $\boxed{y := \mathcal{B}(Y, \lambda y.Wyh, 0)}$. By BND one can prove that

$$|Y\hat{w}_y| \leq |w_y| \vee |w_y| \neq |y|.$$

Finally, define $\boxed{f := \hat{w}_y}$. Then, assuming $|w_y| = |y|$, we have $\overline{f}(Yf) \preceq w_y$, and

$$T(w_y, Zf) \to T(\overline{f}(Yf), Zf)$$

follows from the fact that T is a tree. □

Combined with Lemma 9.6, Theorem 9.2 gives an effective procedure from extracting polynomial-time algorithms from WKL-proofs of Π^0_2-theorems in feasible analysis.

Corollary 9.2 *Let A be a quantifier-free formula. From a proof of $\forall x \exists y A(x, y)$ in the system $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} + \Pi^0_1\text{-}\mathsf{WKL}^\omega$ one can extract a polynomial-time computable function h such that for all x, $A(x, hx)$ is true.*

Proof. Via negative translation, functional interpretation and by Lemma 9.6 one can extract a term h of $\mathcal{L}(\mathsf{IPV}^\omega)$ (i.e. a polynomial-time computable function) such that $\mathsf{IPV}^\omega + \mathsf{BND} + (9.2) \vdash \forall x A(x, hx)$. Scarpellini's [146] type structure of all continuous set-theoretical functionals $\mathcal{C}$ is a model of $\mathsf{IPV}^\omega + \mathsf{BND} + (9.2)$. Therefore, since $\mathcal{C}$ coincides with the full type structure in the types zero and one, the conclusion of the corollary follows. □

By noticing that Lemma 9.6 holds even for terms t which have 0-1 oracle variables, we can strengthen Corollary 9.2.

Corollary 9.3 *Let A be a quantifier-free formula. From a proof of*

$$\forall \alpha(\forall z(\alpha(z) \leq 1) \to \forall x \exists y A(\alpha, x, y))$$

in the system $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} + \Pi^0_1\text{-}\mathsf{WKL}^\omega$ one can extract a polynomial-time computable function (with 0-1 oracle) h such that for all 0-1 oracles α and input x, $A(\alpha, x, h\alpha x)$ holds.

9.6 The Heine/Borel covering lemma

In this section we indicate how to formalize in $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} + \Pi^0_1\text{-}\mathsf{WKL}^\omega$ the proof of the Heine/Borel covering lemma. Our formalization follows closely the ones given in [49] (Theorem 1) and [153] (Lemma IV.1.1).

In the system CPV^ω we shall represent the *tally part* of $\mathbb{N}$ (to be used as unary numbers) as $\mathbb{T}$. Those are natural numbers having binary expansion in the form 1^n. The *rational numbers* $\mathbb{Q}$ shall be represented via pairs of natural number $\langle x, y\rangle$, with the convention that $\pi(x, y)$ represents the rational number $\frac{x}{2(y+1)}$, if x is even, and $\frac{-x-1}{2(y+1)}$ if x is odd. The standard functions and relations on tally and rational numbers can be easily defined in CPV^ω. In the following we use variable i, j and n to range over $\mathbb{T}$, and p, q to range over $\mathbb{Q}$. *Real numbers* $\mathbb{R}$ are represented via functions $\psi : \mathbb{T} \to \mathbb{Q}$ satisfying

$$\forall i, j(i \leq_{\mathbb{T}} j \to \mathsf{Abs}(\psi(i) -_{\mathbb{Q}} \psi(j)) \leq 2^{-i}),$$

where $\mathsf{Abs}(q)$ returns the absolute value of a rational number. A real number ψ_1 is said to be *smaller* than ψ_2, written $\psi_1 <_{\mathbb{R}} \psi_2$, if

$$\exists i(\psi_1(i +_{\mathbb{T}} 1) + 2^{-i} <_{\mathbb{Q}} \psi_2(i +_{\mathbb{T}} 1)).$$

The Heine/Borel covering lemma says that if a sequence of open sets $(\psi_i^L, \psi_i^R)_{i\in\mathbb{T}}$ covers the unit interval $[0, 1]$, then an initial segment of the sequence already covers $[0, 1]$.

Theorem 9.3 *The following is provable in* $\mathsf{CPV}^\omega + \mathsf{QF\text{-}AC}^{\mathbb{N},\mathbb{N}} + \Pi^0_1\text{-}\mathsf{WKL}^\omega$. *Given two sequences of real numbers* $(\psi_i^L)_{i\in\mathbb{T}}$ *and* $(\psi_i^R)_{i\in\mathbb{T}}$, *if*

$$\forall \psi \in [0, 1]\exists i(\psi_i^L <_{\mathbb{R}} \psi <_{\mathbb{R}} \psi_i^R)$$

then $\exists n\forall \psi \in [0, 1]\exists i \leq n(\psi_i^L <_{\mathbb{R}} \psi <_{\mathbb{R}} \psi_i^R)$.

Proof. For each positive number $w \in \mathbb{N}$ (having binary expansion $b_m b_{m-1} \dots b_0$ with $b_m = 1$) we define two rational numbers, written for simplicity in radix notation,

$$p_w := 0.b_{m-1} \dots b_0,$$

$$q_w := 0.b_{m-1} \dots b_0 + 2^{-m}.$$

For completeness we set $p_0 := p_1$ and $q_0 := q_1$. That is, for each positive number m we have partitioned the unit interval $[0, 1]$ into 2^m subintervals of length 2^{-m}. Let

$$T(w) :\equiv \neg\exists i(i \leq |w| \wedge \psi_i^L <_{\mathbb{R}} p_w <_{\mathbb{R}} q_w <_{\mathbb{R}} \psi_i^R).$$

It is easy to show that $T(w)$ defines a tree, i.e. if $T(w)$ holds and $v \preceq w$ then $T(v)$ also holds. Moreover, notice that $T(w)$ is Π^0_1. Assuming that

$$(*)\ \forall \psi \in [0, 1]\exists i(\psi_i^L <_{\mathbb{R}} \psi <_{\mathbb{R}} \psi_i^R)$$

we claim that T has no infinite path. For the sake of contradiction, assume f is an infinite path in T. Define then the real number ψ as (in radix notation)

$$\psi(n) := 0.f(1)\dots f(n -_{\mathbb{T}} 1).$$

Note that for all $j \in \mathbb{T}$, $p_{\overline{f}j} \le \psi \le q_{\overline{f}j}$, and moreover, as j increases both $p_{\overline{f}j}$ and $q_{\overline{f}j}$ converge to ψ. For such ψ, let i be as in assumption $(*)$, i.e. $\psi_i^L <_{\mathbb{R}} \psi <_{\mathbb{R}} \psi_i^R$. Let n be so large that $i \le n$ and $\psi_i^L <_{\mathbb{R}} p_{\overline{f}n} <_{\mathbb{R}} q_{\overline{f}n} <_{\mathbb{R}} \psi_i^R$. Then $\neg T(\overline{f}n)$, which proves the claim. By weak König's lemma it follows that T is finite. Let $n \in \mathbb{T}$ be such that

$$\forall w(T(w) \to |w| < n).$$

Therefore

$$\forall w(|w| = n \to \exists i \le n(\psi_i^L <_{\mathbb{R}} p_w <_{\mathbb{R}} q_w <_{\mathbb{R}} \psi_i^R)),$$

which implies

$$\exists n \forall \psi \in [0,1] \exists i \le n(\psi_i^L <_{\mathbb{R}} \psi <_{\mathbb{R}} \psi_i^R). \quad \Box$$

9.7 Related results and open problems

As mentioned above, Howard [66] gave a functional interpretation (of the negative translation of) WKL using a different form of binary bar recursion, namely

$$\mathcal{B}^H(Y, z) = \begin{cases} 0 & \text{if } Y\hat{z} \le |z| \\ t & \text{otherwise,} \end{cases}$$

where $t = 1 + \max\{\mathcal{B}^H(Y, z0), \mathcal{B}^H(Y, z1)\}$.

In the following let $T(w)$ abbreviate $\forall v \preceq w(gv = 0)$. Consider the negative translation of WKL (over intuitionistic logic and MP)[11]

$$\forall g(\forall n \neg\neg\exists w(|w| = n \wedge T(w)) \to \neg\neg\exists f \forall n T(\overline{f}n)).$$

In a system where exponential search is available, the bounded quantifier $\exists w(|w| = n \wedge \dots)$ can be absorbed by quantifier-free matrices and functional interpretation does not witness it. In such contexts the functional interpretation of WKL^N asks for n and f (uniformly in g and Y) realizing

$$\forall g, Y \exists n, f(\exists w(|w| = n \wedge T(w)) \to T(\overline{f}(Yf))).$$

The functional $\mathcal{B}^H$ is used to realize n by setting $n := \mathcal{B}^H(Y, 1)$. Then, in order to produce f satisfying

$$\forall g, Y \exists f(\exists w(|w| = n \wedge T(w)) \to T(\overline{f}(Yf))),$$

one looks for a w of length n such that $T(w)$ holds. If such w is found, meaning that the premise holds, let v be the shortest prefix of w such that $Y\hat{v} \le |v|$. We then set $f := \hat{v}$, so that (since $Yf = Y\hat{v} \le |v| \le n$) the conclusion also holds. If no such w exists we can safely take f to be an arbitrary path.

[11] In stronger settings the operation $\overline{f}$ is normally defined as

$$\overline{f}(n) = \begin{cases} 0 & \text{if } n = 0 \\ f(0)\dots f(n-1) & \text{otherwise.} \end{cases}$$

It is important to note that Howard's work concerns strong systems (such as Heyting arithmetic HA^ω), on which bounded quantifiers can be absorbed into quantifier-free matrices. In the present context of feasible arithmetic, negative translation and functional interpretation need to take those quantifiers into consideration. But, notice that by taking the quantification over w into consideration we obtain a new functional (W in the proof of Theorem 9.2) which we could use to realize y and f in a *feasible way*.

Interesting follow-ups of the present paper are:

1) Investigate whether effective proofs of WKL elimination for stronger systems (such as Sieg's and Kohlenbach's) can be translated into the feasible setting, by making a careful treatment of bounded quantifiers.

2) Find ineffective proofs of Π^0_2-theorems which can be formalized in $\mathsf{CPV}^\omega + \mathsf{QF}\text{-}\mathsf{AC}^{\mathbb{N},\mathbb{N}} + \Pi^0_1\text{-}\mathsf{WKL}^\omega$, and carry out the extraction of polynomial-time algorithms (cf. [107] where, in the context of classical analysis, a proof based on WKL has been analyzed providing the first effective realizer for the theorem).

3) Find effective proofs of WKL elimination for the setting of feasible analysis, and compare the quality of the algorithms yielded via the two different procedures.

Acknowledgement. I would like to thank Ulrich Kohlenbach for bringing the problem of extracting polynomial-time algorithms from ineffective proofs in feasible analysis to my attention, and for directing me to references such as Howard's paper on bar recursion. I am also very much indebted to him for various discussions we have had on the results here presented.

Bibliography

[1] W. Ackermann. Zur Widerspruchsfreiheit der Zahlentheorie. *Mathematische Annalen*, 117:162–194, 1940.

[2] J. R. Angelos, M. S. Henry, E. H. Kaufman, T. D. Lenker, and A. Kroó. Local and global Lipschitz constants. *Journal of Approximation Theory*, 46:137–156, 1986.

[3] J. R. Angelos and A. Kroó. The equivalence of the moduli of continuity of the best approximation operator and of strong unicity in L^1. *Journal of Approximation Theory*, 46:129–136, 1986.

[4] J. Avigad. Update procedures and the 1-consistency of arithmetic. *Mathematical Logic Quarterly*, 48:3–13, 2002.

[5] J. Avigad and S. Feferman. Gödel's functional ("Dialectica") interpretation. In S. R. Buss, editor, *Handbook of proof theory*, volume 137 of *Studies in Logic and the Foundations of Mathematics*, pages 337–405. North Holland, Amsterdam, 1998.

[6] J. Baillon and R. E. Bruck. The rate of asymptotic regularity is $O(\frac{1}{\sqrt{n}})$. In *Theory and applications of nonlinear operators of accretive and monotone type*, volume 178 of *Lecture Notes in Pure and Appl. Math.*, pages 51–81. Marcel Dekker, Inc., New York, 1996.

[7] M. W. Bartelt and W. Li. Error estimates and Lipschitz constants for best approximation in continuous function spaces. *Computers and Mathematics with Application*, 30(3-6):255–268, 1995.

[8] M. W. Bartelt and D. Schmidt. Lipschitz conditions, strong uniqueness, and almost Chebyshev subspaces of $C(x)$. *Journal of Approximation Theory*, 40:202–215, 1984.

[9] M. J. Beeson. *Foundations of Constructive Mathematics.* Metamathematical Studies. Springer, Berlin, 1985.

[10] G. Bellin. Ramsey interpreted: a parametric version of Ramsey's theorem. In *Logic and computation (Pittsburgh, PA, 1987)*, volume 106, pages 17–37. Amer. Math. Soc., Providence, RI, 1990.

[11] S. Berardi, M. Bezem, and T. Coquand. On the computational content of the axiom of choice. *Journal of Symbolic Logic*, 63(2):600–622, 1998.

[12] U. Berger. *Totale Objekte und Mengen in der Bereichstheorie*. PhD thesis, Mathematisches Institut der Universität München, 1990.

[13] U. Berger. Program extraction from normalization proofs. In M. Bezem, editor, *Typed Lambda Calculi and Applications (TLCA'93)*, volume 664 of *Lecture Notes in Computer Science*, pages 91–106. Springer, Berlin, 1993.

[14] U. Berger and P. Oliva. Modified bar recursion. BRICS Report Series RS-02-14 (23 pages), BRICS – Basic Research in Computer Science, 2002. http://www.brics.dk/RS/02/14/BRICS-RS-02-14.ps.gz.

[15] U. Berger and P. Oliva. Modified bar recursion and classical dependent choice. To appear in: Lecture Notes in Logic, 2003.

[16] U. Berger and H. Schwichtenberg. Program extraction from classical proofs. In D. Leivant, editor, *Logic and Computational Complexity Workshop (LCC'94)*, volume 960 of *Lecture Notes in Computer Science*, pages 77–97. Springer, Berlin, 1995.

[17] U. Berger, H. Schwichtenberg, and W. Buchholz. Refined program extraction from classical proofs. *Annals of Pure and Applied Logic*, 114:3–25, 2002.

[18] M. Bezem. Strongly majorizable functionals of finite type: a model for bar recursion containing discontinuous functionals. *Journal of Symbolic Logic*, 50:652–660, 1985.

[19] M. Bezem. Equivalence of bar recursors in the theory of functionals of finite type. *Archive for Mathematical Logic*, 27:149–160, 1988.

[20] E. Bishop. *Foundations of Constructive Analysis*. McGraw-Hill, New York, 1967.

[21] E. Bishop. Mathematics as a numerical language. In A. Kino, J. Myhill, and R. E. Vesley, editors, *Intuitionism and Proof Theory*, pages 53–71. North Holland, Amsterdam, 1970.

[22] B. O. Björnestål. Continuity of the metric projection operator I-III. The preprint series of Department of Mathematics 17, Royal Institute of Technology, Stockholm, TRITA-MAT, 1975.

[23] B. O. Björnestål. Local Lipschitz continuity of the metric projection operator. In *Approximation theory*, volume 4 of *Banach Center Publications*, pages 43–53. PWN-Polish Scientific Publishers, Warsaw, 1979.

[24] E. Bombieri and A. J. van der Poorten. Some quantitative results related to Roth's theorem. *Journal of the Australian Mathematical Society*, 45(2):233–248, 1988.

[25] J. Borwein, S. Reich, and I. Shafrir. Krasnoselski-Mann iterations in normed spaces. *Canadian Mathematical Bulletin*, 35:21–28, 1992.

[26] D. S. Bridges. A constructive development of Chebychev approximation theory. *Journal of Approximation Theory*, 30:99–120, 1980.

[27] D. S. Bridges. Lipschitz constants and moduli of continuity for the Chebychev projection. *Proceedings of the American Mathematical Society*, 85:557–561, 1982.

[28] D. S. Bridges, F. Richman, W. H. Julian, and R. Mines. Extensions and fixed points of contractive maps in R^n. *Journal of Mathematical Analysis and Applications*, 165(2):438–456, 1992.

[29] F. E. Browder. Nonexpansive nonlinear operators in a Banach space. In *Proceedings of the National Academy of Sciences of the U.S.A.*, volume 54, pages 1041–1044, 1965.

[30] F. E. Browder and W. V. Petryshyn. The solution by iteration of nonlinear functional equations in Banach spaces. *Bulletin of the American Mathematical Society*, 72:571–575, 1966.

[31] S. R. Buss. *Bounded arithmetic.* PhD thesis, Princeton University, Princeton, New Jersey, 1985.

[32] S. R. Buss. The polynomial hierarchy and intuitionistic bounded arithmetic. In G. Goos and J. Hartmanis, editors, *Structure in Complexity Theory*, volume 223 of *Lecture Notes in Computer Science*, pages 77–103. Springer, Berlin, 1986.

[33] E. W. Cheney. An elementary proof of Jackson's theorem on mean-approximations. *Mathematics Magazine*, 38:189–191, 1965.

[34] E. W. Cheney. *Approximation Theory.* AMS, Providence, RI, Chelsea, New York, 1966.

[35] J. A. Clarkson. Uniformly convex spaces. *Transactions of the American Mathematical Society*, 40:396–414, 1936.

[36] P. Collaço and J. Carvalho e Silva. A complete comparison of 25 contraction conditions. *Nonlinear Analysis*, 30:471–476, 1997.

[37] S. Cook. Feasibly constructive proofs and the propositional calculus. In *Proc. 7th A.C.M. Symposium on the Theory of Computation*, pages 83–97, 1975.

[38] S. Cook and A. Urquhart. Functional interpretations of feasibly constructive arithmetic. *Annals of Pure and Applied Logic*, 63:103–200, 1993.

[39] T. Coquand and M. Hofmann. A new method for establishing conservativity of classical systems over their intuitionistic version. *Mathematical Structures in Computer Science*, 9:323–333, 1999.

[40] C. Delzell. Kreisel's unwinding of Artin's proof. In P. Odifreddi, editor, *Kreiseliana: About and around Georg Kreisel*, pages 113–246. A. K. Peters, Wellesley, MA, 1996.

[41] A. Dragalin. New kinds of realizability. In *Abstracts of the 6th International Congress of Logic, Methodology and Philosophy of Sciences, Hannover, Germany*, pages 20–24, 1979.

[42] M. Edelstein. On fixed and periodic points under contractive mappings. *The Journal of the London Mathematical Society*, 37:74–79, 1962.

[43] M. Edelstein and R. C. O'Brien. Nonexpansive mappings, asymptotic regularity and successive approximations. *The Journal of the London Mathematical Society*, 17:547–554, 1978.

[44] H. W. Engl and A. Leitão. A Mann interative regularization method for elliptic Cauchy problems. *Numerical Functional Analysis and Optimization*, 22:861–884, 2001.

[45] Y. L. Ershov. Model C of partial continuous functionals. In R. Gandy and M. Hyland, editors, *Logic Colloquium 1976*, pages 455–467. North Holland, Amsterdam, Amsterdam, 1977.

[46] S. Feferman. Theories of finite type related to mathematical practice. In J. Barwise, editor, *Handbook of Mathematical Logic*, pages 913–972. North Holland, Amsterdam, 1977.

[47] S. Feferman. Kreisel's 'Unwinding Program'. In P. Odifreddi, editor, *Kreiseliana: about and around Georg Kreisel*, pages 247–273. A. K. Peters, Wellesley Massachusetts, 1996.

[48] F. Ferreira. A feasible theory for analysis. *The Journal of Symbolic Logic*, 59(3):1001–1011, 1994.

[49] F. Ferreira and A. Fernandes. Basic applications of weak König's lemma in feasible analysis. Submitted.

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

[51] H. Friedman. Classically and intuitionistically provably recursive functions. In D. Scott and G. Müller, editors, *Higher Set Theory*, volume 669 of *Lecture Notes in Mathematics*, pages 21–28. Springer, Berlin, 1978.

[52] H. Friedman. On the computational complexity of maximization and integration. *Advances in Mathematics*, 53:80–98, 1984.

[53] R. O. Gandy and M. Hyland. Computable and recursively countable functionals of higher type. In R. Gandy and M. Hyland, editors, *Logic Colloquium 1976*, pages 407–438. North Holland, Amsterdam, Amsterdam, 1977.

[54] G. Gentzen. Die Widerspruchsfreiheit der reinen Zahlentheorie. *Mathematische Annalen*, 112:493–565, 1936.

[55] J.-Y. Girard. *Proof Theory and Logical Complexity.* Studies in Proof Theory. Bibliopolis, Naples, 1987.

[56] V. I. Glivenko. Sur quelques points de la logique de M. Brouwer. *Bulletin de la Société Mathématique de Belgique*, 15, 1929.

[57] K. Gödel. Zur intuitionistischen Arithmetik und Zahlentheorie. *Ergebnisse eines Mathematischen Kolloquiums*, 4:34–38, 1933.

[58] K. Gödel. Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *Dialectica*, 12:280–287, 1958.

[59] K. Goebel and W. A. Kirk. Iteration processes for nonexpansive mappings. In *Topological Methods in Nonlinear Functional Analysis*, volume 21 of *Contemporary Mathematics*, pages 115–123. AMS, 1983.

[60] K. Goebel and W. A. Kirk. *Topics in Metric Fixed Point Theory*, volume 28 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1990.

[61] K. Goebel and S. Reich. *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*, volume 83 of *Monographs and Textbooks in Pure and Applied Mathematics.* Marcel Dekker, Inc., New York, 1984.

[62] C. W. Groetsch. A note on segmenting Mann iterates. *Journal of Mathematical Analysis and Applications*, 40:369–372, 1972.

[63] M.-D. Hernest and U. Kohlenbach. A complexity analysis of functional interpretations. BRICS Report Series RS-03-12 (70 pp.), BRICS – Basic Research in Computer Science, 2003 (Submitted). http://www.brics.dk/RS/03/12/BRICS-RS-03-12.ps.gz.

[64] D. Hilbert. Über das Unendliche. *Mathematische Annalen*, 95:161–190, 1926.

[65] W. A. Howard. Hereditarily majorizable functionals of finite type. In A. S. Troelstra, editor, *Metamathematical investigation of intuitionistic Arithmetic and Analysis*, volume 344 of *Lecture Notes in Mathematics*, pages 454–461. Springer, Berlin, 1973.

[66] W. A. Howard. Ordinal analysis of simple cases of bar recursion. *The Journal of Symbolic Logic*, 46:17–30, 1981.

[67] W. A. Howard and G. Kreisel. Transfinite induction and bar induction of types zero and one, and the role of continuity in intuitionistic analysis. *The Journal of Symbolic Logic*, 31:325–358, 1966.

[68] R. Huotari and S. Sahab. Strong unicity versus modulus of convexity. *Bulletin of the Australian Mathematical Society*, 49:305–310, 1994.

[69] S. Ishikawa. Fixed points and iterations of a nonexpansive mapping in a Banach space. *Proceedings of the American Mathematical Society*, 59:65–71, 1976.

[70] D. Jackson. Note on a class of polynomials of approximation. *Transactions of the American Mathematical Society*, 22:320–326, 1921.

[71] B. Kapron and S. Cook. A new characterization of type-2 feasibility. *SIAM Journal Computing*, 25(1):117–132, 1996.

[72] B. Kauffmann. *Application of Proof Theory to Computational Complexity Theory: Comparison of Different Methods*. PhD thesis, Carnegie Mellon University, Department of Philosophy, 2001.

[73] W. A. Kirk. Krasnoselski's iteration process in hyperbolic space. *Numerical Functional Analysis and Optimization*, 4:371–381, 1982.

[74] W. A. Kirk. Nonexpansive mappings and asymptotic regularity. *Nonlinear Analysis*, 40:323–332, 2000.

[75] W. A. Kirk and C. Martinez-Yañez. Approximate fixed points for nonexpansive mappings in uniformly convex spaces. *Annales Polonici Mathematici*, 51:189–193, 1990.

[76] W. A. Kirk and B. Sims, editors. *Handbook of Metric Fixed Point Theory*. Kluwer Academic Publishers, Dordrecht, 2001.

[77] S. C. Kleene. On the interpretation of intuitionistic number theory. *The Journal of Symbolic Logic*, 10:109–124, 1945.

[78] S. C. Kleene. *Introduction to Metamathematics*. D. Van Nostrand Co., Inc., New York, N. Y., 1952.

[79] S. C. Kleene. Countable functionals. In A. Heyting, editor, *Constructivity in Mathematics*, pages 81–100. North Holland, Amsterdam, 1959.

[80] S. C. Kleene. Recursive functionals and quantifiers of finite types I. *Trans. Amer. Math. Soc.*, 91:1–52, 1959.

[81] S. C. Kleene. *The Foundations of Intuitionistic Mathematics*. North Holland, Amsterdam, 1965.

[82] K.-I. Ko. On the computational complexity of best Chebycheff approximation. *Journal of Complexity*, 2:95–120, 1986.

[83] K.-I. Ko. *Complexity Theory of Real Functions*. Birkhäuser, Boston-Basel-Berlin, 1991.

[84] K.-I. Ko. Polynomial-time computability in Analysis. In Y. L. Ershov, S. Goncharov, A. Nerode, and J. Remmel, editors, *Handbook of recursive mathematics (vol. 2)*, volume 139 of *Studies in Logic and the Foundations of Mathematics*, pages 1271–1317. North Holland, Amsterdam, 1998.

[85] K.-I. Ko and H. Friedman. Computational complexity of real functions. *Theoretical Computer Science*, 20:323–352, 1982.

[86] U. Kohlenbach. Higher order reverse mathematics. To appear in: S. Simpson, editor, *Reverse Mathematics*, 14pp.

[87] U. Kohlenbach. Proof interpretations and the computational content of proofs. http://www.brics.dk/∼kohlenb/newcourse.ps, BRICS Lecture Notes, 2003.

[88] U. Kohlenbach. *Theory of majorizable and continuous functionals and their use for the extraction of bounds from non-constructive proofs: effective moduli of uniqueness for best approximations from ineffective proofs of uniqueness (German).* PhD thesis, Frankfurt, pp. xxii+278, 1990.

[89] U. Kohlenbach. Effective bounds from ineffective proofs in analysis: an application of functional interpretation and majorization. *The Journal of Symbolic Logic*, 57:1239–1273, 1992.

[90] U. Kohlenbach. Effective moduli from ineffective uniqueness proofs. An unwinding of de La Vallée Poussin's proof for Chebycheff approximation. *Annals of Pure and Applied Logic*, 64:27–94, 1993.

[91] U. Kohlenbach. New effective moduli of uniqueness and uniform a–priori estimates for constants of strong unicity by logical analysis of known proofs in best approximation theory. *Numerical Functional Analysis and Optimization*, 14:581–606, 1993.

[92] U. Kohlenbach. Analysing proofs in Analysis. In W. Hodges, M. Hyland, C. Steinhorn, and J. Truss, editors, *Logic: from Foundations to Applications*, pages 225–260. European Logic Colloquium (Keele, 1993), Oxford University Press, 1996.

[93] U. Kohlenbach. Mathematically strong subsystems of analysis with low rate of growth of provably recursive functionals. *Archive for Mathematical Logic*, 36:31–71, 1996.

[94] U. Kohlenbach. Arithmetizing proofs in analysis. In J. M. Larrazabal, D. Lascar, and G. Mints, editors, *Logic Colloquium'96, Lecture Notes in Logic*, volume 12, pages 115–158. Springer, Berlin, 1998.

[95] U. Kohlenbach. Proof theory and computational analysis. In A. Edalat, A. Jung, K. Keimel, and M. Kwiatkowska, editors, *Electronic Notes in Theoretical Computer Science*, volume 13. Elsevier, Amsterdam, 1998. 34 pp. (electronic).

[96] U. Kohlenbach. Relative constructivity. *The Journal of Symbolic Logic*, 63:1218–1238, 1998.

[97] U. Kohlenbach. On the no-counterexample interpretation. *The Journal of Symbolic Logic*, 64:1491–1511, 1999.

[98] U. Kohlenbach. The use of a logical principle of uniform boundedness in analysis. In A. Cantini, E. Casari, and P. Minari, editors, *Logic and*

Foundations of Mathematics, volume 280 of *Synthese Library*, pages 93–106. Kluwer Academic Publishers, Dordrecht, 1999.

[99] U. Kohlenbach. Things that can and things that can't be done in PRA. *Annals of Pure and Applied Logic*, 102:223–245, 2000.

[100] U. Kohlenbach. On the computational content of the Krasnoselski and Ishikawa fixed point theorems. In J. Blanck, V. Brattka, and P. Hertling, editors, *Computability and Complexity in Analysis, (CCA'2000)*, volume 2064 of *Lecture Notes in Computer Science*, pages 119–145. Springer, Berlin, 2001.

[101] U. Kohlenbach. A quantitative version of a theorem due to Borwein-Reich-Shafrir. *Numerical Functional Analysis and Optimization*, 22:641–656, 2001.

[102] U. Kohlenbach. Foundational and mathematical uses of higher types. In W. Sieg et al., editors, *Reflections on the Foundations of Mathematics: Essay in Honor of Solomon Feferman*, volume 15 of *Lecture Notes in Logic*, pages 92–116. A. K. Peters, Ltd., 2002.

[103] U. Kohlenbach. Uniform asymptotic regularity for Mann iterates. *Journal of Mathematical Analysis and Applications*, 279:531–544, 2003.

[104] U. Kohlenbach. Some logical metatheorems with applications in functional analysis. BRICS Research Report RS-03-21 (55pp.), BRICS – Basic Research in Computer Science, 2003 (Submitted).

[105] U. Kohlenbach and L. Leuştean. Mann iterates of directionally nonexpansive mappings in hyperbolic spaces. *Abstract and Applied Analysis*, 2003(8):449–477, 2003.

[106] U. Kohlenbach and P. Oliva. Proof mining: a systematic way of analysing proofs in mathematics. To appear in: Proc. Steklov Inst. Math., 2003.

[107] U. Kohlenbach and P. Oliva. Proof mining in L_1-approximation. *Annals of Pure and Applied Logic*, 121:1–38, 2003.

[108] A. N. Kolmogorov. On the principle of the excluded middle (russian). *Mat. Sb.*, 32:646–667, 1925.

[109] G. Köthe. *Topoplogische Lineare Räume*. Springer, Berlin, 1960.

[110] M. A. Krasnoselski. Two remarks on the method of successive approximation (russian). *Usp. Math. Nauk (N.S.)*, 10:123–127, 1955.

[111] G. Kreisel. On the interpretation of non-finitist proofs, part I. *The Journal of Symbolic Logic*, 16:241–267, 1951.

[112] G. Kreisel. On the interpretation of non-finitist proofs, part II: Interpretation of number theory. *The Journal of Symbolic Logic*, 17:43–58, 1952.

[113] G. Kreisel. Hypotheses on algebraically closed extensions. *Bulletin of the American Mathematical Society*, 63:99, 1957.

[114] G. Kreisel. Mathematical significance of consistency proof. *The Journal of Symbolic Logic*, 23(2):155–182, 1958.

[115] G. Kreisel. Interpretation of analysis by means of constructive functionals of finite types. In A. Heyting, editor, *Constructivity in Mathematics*, pages 101–128. North Holland, Amsterdam, 1959.

[116] G. Kreisel. On weak completeness of intuitionistic predicate logic. *The Journal of Symbolic Logic*, 27:139–158, 1962.

[117] G. Kreisel. Finiteness theorems in arithmetic: an application of Herbrand's theorem for Σ_2-formulas. In *Proc. of the Herbrand Symposium*, pages 39–55. North Holland, Amsterdam, 1982.

[118] A. Kroó. On the continuity of best approximations in the space of integrable functions. *Acta Mathematica Academiae Scientiarum Hungaricae*, 32:331–348, 1978.

[119] A. Kroó. On the uniform modulus of continuity of the operator of best approximation in the space of periodic functions. *Acta Mathematica Academiae Scientiarum Hungaricae*, 34:185–203, 1979.

[120] A. Kroó. On strong unicity of L_1-approximation. *Proceedings of the American Mathematical Society*, 83(4):725–729, 1981.

[121] S. Kuroda. Intuitionistische Untersuchungen der formalistischen Logik. *Nagoya Mathematical Journal*, 3:35–47, 1951.

[122] D. Leivant. Syntactic translations and provably recursive functions. *The Journal of Symbolic Logic*, 50(3):682–688, Sep., 1985.

[123] P.-K. Lin. Strongly unique best approximation in uniformly convex Banach spaces. *Journal of Approximation Theory*, 56:101–107, 1989.

[124] H. Luckhardt. *Extensional Gödel Functional Interpretation: A Consistency Proof of Classical Analysis*, volume 306 of *Lecture Notes in Mathematics*. Springer, Berlin, 1973.

[125] H. Luckhardt. Herbrand-Analysen zweier Beweise des Satzes von Roth: Polynomiale Anzahlschranken. *The Journal of Symbolic Logic*, 54:234–263, 1989.

[126] H. Luckhardt. Bounds extracted by Kreisel from ineffective proofs. In P. Odifreddi, editor, *Kreiseliana: About and around Georg Kreisel*, pages 289–300. A. K. Peters, Wellesley, MA, 1996.

[127] W. R. Mann. Mean value methods in iteration. *Proceedings of the American Mathematical Society*, 4:506–510, 1953.

[128] K. Martin. Unique fixed points in domain theory. In S. Brooks and M. Mislove, editors, *Electronic Notes in Theoretical Computer Science*, volume 45. Elsevier, Amsterdam, 2001. 11 pp. (electronic).

[129] J. Meszáros. A comparison of various definitions of contractive type mappings. *Bulletin of the Calcutta Mathematical Society*, 84:167–194, 1992.

[130] M. W. Mislove. Generalizing domain theory. http://www.math.tulane.edu/ mwm/ftp/fossacs.ps.gz.

[131] D. J. Newman and H. S. Shapiro. Some theorems on Chebycheff approximation. *Duke Mathematical Journal*, 30:673–681, 1963.

[132] D. Normann. *Recursion on the Countable Functionals*, volume 811 of *Lecture Notes in Mathematics.* Springer, Berlin, 1980.

[133] D. Normann. The continuous functionals. In E. R. Griffor, editor, *Handbook of Computability Theory*, chapter 8, pages 251–275. North Holland, Amsterdam, 1999.

[134] P. Oliva. On the computational complexity of best L_1-approximation. *Mathematical Logic Quarterly*, 48, Suppl. 1:66–77, 2002.

[135] P. Oliva. Polynomial-time algorithms from ineffective proofs. In *Proc. of the Eighteenth Annual IEEE Symposium on Logic in Computer Science LICS'03*, pages 128–137. IEEE Press, 2003.

[136] C. Papadimitriou. *Computational Complexity.* Addison-Wesley, 1994.

[137] A. Pinkus. *On L_1-Approximation*, volume 93 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge, 1989.

[138] G. D. Plotkin. LCF considered as a programming language. *Theoretical Computer Science*, 5:223–255, 1977.

[139] M. Pour-El and J. I. Richards. *Computability in Analysis and Physics.* Perspectives in Mathematical Logic. Springer, Berlin, 1989.

[140] B. Prus and R. Smarzewski. Strongly unique best approximations and centers in uniformly convex spaces. *Journal of Mathematical Analysis and Applications*, 121:10–21, 1987.

[141] E. Rakotch. A note on contractive mappings. *Proceedings of the American Mathematical Society*, 13:459–465, 1962.

[142] S. Reich and I. Shafrir. Nonexpansive iterations in hyperbolic spaces. *Nonlinear Analysis*, 15:537–558, 1990.

[143] B. E. Rhoades. Fixed point iterations using infinite matrices. *Transactions of the American Mathematical Society*, 196:161–176, 1974.

[144] B. E. Rhoades. A comparison of various definitions of contractive mappings. *Transactions of the American Mathematical Society*, 226:257–290, 1977.

[145] B. E. Rhoades. Contractive definitions. In T. M. Rassias, editor, *Nonlinear analysis*, pages 513–526. World Sci. Publishing, Singapore, 1987.

[146] B. Scarpellini. A model for bar recursion of higher types. *Compositio Mathematica*, 23:123–153, 1971.

[147] H. Schwichtenberg. On bar recursion of types 0 and 1. *The Journal of Symbolic Logic*, 44:325–329, 1979.

[148] D. S. Scott. Outline of a mathematical theory of computation. In *4th Annual Princeton Conference on Information Sciences and Systems*, pages 169–176, 1970.

[149] W. Sieg. Fragments of arithmetic. *Annals of Pure and Applied Logic*, 28:33–71, 1985.

[150] W. Sieg. Hilbert's program sixty years later. *The Journal of Symbolic Logic*, 53:338–348, 1988.

[151] W. Sieg. Herbrand analyses. *Archive for Mathematical Logic*, 30:409–441, 1991.

[152] A. Simpson. Lazy functional algorithms for exact real functionals. In L. Brim, J. Gruska, and J. Zlatuska, editors, *Mathematical Foundations of Computer Science*, volume 1450 of *Lecture Notes in Computer Science*, pages 456–464, 1998.

[153] S. G. Simpson. *Subsystems of Second Order Arithmetic.* Perspectives in Mathematical Logic. Springer, Berlin, 1999.

[154] C. Spector. Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles in current intuitionistic mathematics. In F. D. E. Dekker, editor, *Recursive Function Theory: Proc. Symposia in Pure Mathematics*, volume 5, pages 1–27. American Mathematical Society, Providence, Rhode Island, 1962.

[155] W. Takahashi. A convexity in metric space and nonexpansive mappings I. *Kodai Mathematical Seminar Reports*, 22:142–149, 1970.

[156] A. S. Troelstra. Review of: Intuitionistic analysis and Gödel's interpretation. *The Journal of Symbolic Logic*, 37:404, 1972.

[157] A. S. Troelstra. Note on the fan theorem. *The Journal of Symbolic Logic*, 39:584–596, 1974.

[158] A. S. Troelstra. Realizability. In S. Buss, editor, *Handbook of proof theory*, volume 137, pages 408–473. North Holland, Amsterdam, 1998.

[159] A. S. Troelstra and D. van Dalen. *Constructivism in Mathematics. An Introduction*, volume 121 of *Studies in Logic and the Foundations of Mathematics*. North Holland, Amsterdam, 1988.

[160] A. S. Troelstra (ed.). *Metamathematical Investigation Iintuitionistic Arithmetic and Analysis*, volume 344 of *Lecture Notes in Mathematics*. Springer, Berlin, 1973.

[161] K. Weihrauch. *Computable Analysis*. Springer, Berlin, 2000.

[162] M. Yasugi. Intuitionistic analysis and Gödel's interpretation. *Journal of the Math. Society of Japan*, 15:101–112, 1963.

Index