PROOF INTERPRETATIONS AND THE COMPUTATIONAL CONTENT OF PROOFS IN MATHEMATICS

Ulrich Kohlenbach*

1 Introduction

This survey reports on some recent developments in the project of applying proof theory to proofs in core mathematics. The historical roots, however, go back to Hilbert's central theme in the foundations of mathematics which can be paraphrased by the following question

"How is it that abstract methods ('ideal elements') can be used to prove 'real' statements e.g. about the natural numbers and is this use necessary in principle?"

Hilbert's original aim, to show the consistency of the use of such ideal elements by finitistic means, which would suffice to eliminate these ideal elements from proofs of purely universal ('real') theorems ('Hilbert's program'), turned out to be impossible for theories containing a sufficient amount of number theory by K. Gödel's 2nd incompleteness theorem. Nevetheless, various partial realizations of this program and many relative consistency proofs could be achieved. One class of tools used to obtain this are so-called proof interpretations I which transform proofs p in theories $\mathcal{T}$ of theorems A into new proofs p^I in theories $\mathcal{T}^I$ of the interpretation A^I of A. If $(0 = 1)^I \equiv (0 = 1)$, then this yields a consistency proof for $\mathcal{T}$ relative to the assumed consistency of $\mathcal{T}^I$. Whereas, Hilbert's program focusses exclusively on (ideal proofs of) purely universal theorems, a natural 'shift of emphasis' (G. Kreisel) is to try to apply proof interpretations to interesting proofs of existential theorems with the goal to extract new information from the proof which was not visible beforehand.[1] This new information often consists of effective data such as algorithms or effective bounds (extracted from prima facie ineffective proofs) but also continuous dependence or even full independence of

*Fachbereich Mathematik, Technische Universität Darmstadt, Schlossgartenstraße 7, D-64289 Darmstadt, Germany, kohlenbach@mathematik.tu-darmstadt.de

[1] As stressed by Kreisel, for this applied purpose, proofs of universal lemmas do not matter at all (but only their truth) so that such lemmas may be taken simply as axioms. So Kreisel's emphasis is sort of opposite to that in Hilbert's program.

solutions from certain parameters (uniformity). Another aspect is the generalization of proofs by a weakening of the premises.
Already in the 50's G. Kreisel had asked

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

Kreisel proposed to apply proof theoretic techniques – originally developed for foundational purposes – to concrete proofs in mathematics which mathematicians could not 'unwind' themselves (see e.g. [78, 85, 27] and - more recently - [86]).

Kreisel's idea of **unwinding proofs**, i.e. the logical analysis of proofs using techniques from proof theory, has been applied e.g. to number theory ([77, 84]), combinatorics ([32, 3, 47, 100]) (though the latter two papers also apply combinatorics to proof theory) and algebra ([19, 20, 21, 26]). During the last 10-15 years, however, the most systematic development of such an applied proof theory (also called 'Proof Mining') took place in connection with applications to numerical analysis (approximation theory) and functional analysis (nonlinear analysis, fixed point theory and ergodic theory), see e.g. [1, 14, 15, 16, 30, 55, 61, 62, 64, 70, 71, 72, 69, 73, 74, 83]. The area of analysis and, in particular, functional analysis seems specially suited for this approach as here issues of representations of analytical objects (usually not made explicit in mathematics) play a crucial role and are systematically addressed by techniques such as proof interpretations.
Moreover, in the context of applications to functional analysis, new logical metatheorems have been developed which not only a priori guarantee the extractability of effective bounds but also qualitatively new uniformity results (see [63, 31, 82]). These metatheorems and the resulting applications to concrete proofs in analysis are all based on suitable extensions and refinements of so-called functional interpretations which have the root in Gödel's 'Dialectica Interpretation' ([36, 68]). In particular, a monotone functional interpretation due to the author [56] is crucially used.

Recently, Terence Tao ([98, 99]) arrived, prompted by some of his famous work[2] on the use of ergodic theory in combinatorics and number theory, at a proposal of so-called 'hard analysis' (as opposed to 'soft analysis') which roughly can be understood as carrying out analysis on the level of uniform bounds in the sense of monotone functional interpretation which in many cases allows one to 'finitize' analytic assumptions and to arrive at qualitatively stronger results. Indeed, one of the main benefits of the metatheorems proved in [63, 31] is that they allow one to do exactly this (e.g. it is shown how to remove assumptions like the existence of

[2]Tao was awarded a fields medal for this work in 2006. One particularly celebrated result is his ergodic theoretic proof (together with B. Green) that there are arithmetic progressions of arbitrary length in the prime numbers.

fixed points etc. in proofs). Tao illustrates his ideas using two examples: a finite convergence principle and a ‘finitary’ infinite pigeonhole principle. We indicate below how the former and a variant of the latter directly result from monotone functional interpretation.

In this survey we sketch some of the central techniques for applied proof theory and list a few keynote applications in various areas of mathematics. As to the proof theoretic methods we discuss we focus (rather than aiming at a complete list of the various techniques that have been developed over the years) on those which have been instrumental in finding new mathematical results (when applied to specific proofs).
Since the number of applications in functional analysis is rather large and we just published a comprehensive survey on the results obtained up to 2006 in [67] we will mainly restrict our selection in this area to some recent results from 2007 (not covered in [67]).

Obviously, a short paper like this can only touch in a very superficial way most of the issues it addresses. For a more serious and comprehensive treatment of both the logical aspects of proof interpretations as well as their uses in mathematics we have to refer the reader to the forthcoming book [66].

For a proper understanding of the rest of this article we presuppose some basic knowledge of first order logic and type systems as well as of basic notions from abstract analysis such as metric, normed and Hilbert spaces while all other concepts used will be defined. Some of the results are formulated in the general context of so-called hyperbolic spaces (which comprises both normed spaces as well as important structures frequently used in geometric group theory such as CAT(0)-spaces in the sense of Gromov). However, one can get a proper understanding of the main results already by just replacing ‘hyperbolic space’ everywhere by ‘convex subset of a normed space’ and ‘$(1-\lambda)x \oplus \lambda y$’ by ‘$(1-\lambda)x + \lambda y$’.

Notation: Throughout this paper $\mathbb{N} := \{0, 1, 2, \ldots\}$.

2 Some tools from proof theory

One of the oldest tools from proof theory, which has been effectively used by H. Luckhardt in the unwinding of proofs of the famous finiteness theorem of Roth in diophantine approximation (see section 3 below), is Herbrand’s theorem which we formulate here only for the case of Π^0_3-sentences (which covers finiteness statements such as the one proved by Roth).

Let

$$A \equiv \forall x \exists y \forall z \, A_{qf}(x, y, z)$$

be sentence where A_{qf} is quantifier-free.
The so-called Herbrand normal form A^H of A is defined as

$$A^H :\equiv \forall x \exists y\, A_{qf}(x, y, f(y)),$$

where f is a new function symbol (also called index function). Herbrand's theorem states two things:

1) From a given proof of A in predicate logic without equality one can extract finitely many terms $t_1, \dots, t_n$ which are built up out of x, f and the constants occurring in A such that

$$A^{H,D} :\equiv \bigvee_{i=1}^{n} A_{qf}(x, t_i, f(t_i))$$

is a tautology.

2) There is a direct derivation (using only appropriate quantifier introduction rules and contractions) from any disjunction of the form $A^{H,D}$ to A.

Remark 2.1. *Over 2nd order logic where we can quantify over functions and can write A^H as $\forall f, x \exists y\, A_{qf}(x, y, f(y))$, the Herbrand normal form A^H and A can be shown to be equivalent (using the axiom of choice).*

Let us illustrate things using the following logically valid sentence of the required form

$$A :\equiv \forall x \exists y \forall z\, (P(x, y) \vee \neg P(x, z)),$$

where

$$A^H \equiv \forall x \exists y\, (P(x, y) \vee \neg P(x, f(y))).$$

Whereas no single term instantiated for '$\exists y$' produces a tautology, two terms $t_1 := x, t_2 := f(x)$ will do since

$$A^{H,D} :\equiv (P(x, x) \vee \neg P(x, f(x))) \vee (P(x, f(x)) \vee \neg P(x, f(f(x))))$$

is a tautology.

For the 2nd claim in Herbrand's theorem one argues as follows: $A^{H,D}$ remains being a tautology if we replace the f-terms starting from the term of greatest depth successively by new variables resulting in

$$A^D :\equiv (P(x, x) \vee \neg P(x, y)) \vee (P(x, y) \vee \neg P(x, z))$$

from which we arrive back to A by an obvious direct proof (first introducing the quantifiers in the 2nd disjunct and then the ones in the first disjunct and finally applying a contraction).

It is an easy exercise to show that in general for sentences $A \equiv \forall x \exists y \forall z\, A_{qf}(x, y, z)$, A^D can always be written in the linearly ordered form

$$(L) \;\; (A_{qf}(x, t_1, b_1) \vee A_{qf}(x, t_2, b_2) \vee \ldots \vee A_{qf}(x, t_k, b_k),$$

where the b_i are new variables and t_i does not contain any b_j with $i \leq j$ (see [77]). It is this form of Herbrand's theorem, which – appropriately reformulated – also extends theories having only purely universal axioms, that is used by H. Luckhardt.

For theories which logically complex axioms (e.g. general induction axioms) such as Peano Arithmetic PA, Herbrand's theorem does not extend. However, the formulation $A^{H,D}$ suggests the following generalization: suppose that we work in a theory with decidable prime formulas (and hence decidable quantifier-free formulas) such as PA. Then the above Herbrand disjunction $A^{H,D}$ can be written as (using function quantifiers)

$$\forall f\, A_{qf}(x, \Phi(x, f), f(\Phi(x, f))),$$

with

$$\Phi(x, f) := t_i[x, f],$$

where $1 \leq i \leq n$ is least such that

$$A_{qf}(x, t_i[x, f], f(t_i[x, f]))$$

holds. If one is working in a system containing some arithmetic and interested only in a **bound** on a number quantifier '$\exists y$' in A^H one can take simply

$$\Phi^*(x, f) := \max\{t_1[x, f], \ldots, t_n[x, f]\}$$

which is even independent from the prime formulas in A.

Whereas for theorems A proved in theories $\mathcal{T}$ having only universal axioms (so-called 'open' theories) it suffices to use terms t_i that are built of the A^H- and $\mathcal{T}$-material, so that for Φ, Φ^* in addition to this only definition-by-case functions resp. a maximum function are required, it is clear that for more complicated theories $\mathcal{T}$ more complicated classes of functionals Φ are required. This leads to the following definition

Definition 2.2 (Kreisel [75, 76]). *A computable functional $\Phi : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ satisfies the no-counterexample interpretation (short: n.c.i.) of a sentence*

$$A \equiv \forall x \exists y \forall z\, A_{qf}(x, y, z) \in \mathcal{L}(\mathrm{PA})$$

if

$$\forall x \in \mathbb{N} \forall f \in \mathbb{N}^{\mathbb{N}} A_{qf}(x, \Phi(x, f), f(\Phi(x, f)))$$

is true.

We illustrate the no-counterexample interpretation by the following example used already by G. Kreisel in [75] and which recently received new attention by T. Tao's discussion in [98]: Let (a_n) be a nonincreasing sequence in $[0, 1]$. Then, clearly, (a_n) is convergent and so a Cauchy sequence. For convenience, we express this fact in the following form

$$(1)\ \forall k \in \mathbb{N} \exists n \in \mathbb{N} \forall m \in \mathbb{N} \forall i, j \in [n; n+m]\ (|a_i - a_j| \leq 2^{-k}),$$

where $[n; n+m] := \{n, n+1, \ldots, n+m\}$.
Then (treating for the moment $\leq$ between real numbers as a primitive predicate and disregarding the bounded quantifiers '$\forall i, j \in [n; n+m]$') the Herbrand normal form of this statement is

$$(2)\ \forall k \in \mathbb{N} \forall g \in \mathbb{N}^{\mathbb{N}} \exists n \in \mathbb{N} \forall i, j \in [n; n+g(n)]\ (|a_i - a_j| \leq 2^{-k}).$$

By the well-known counterexamples due to E. Specker ('Specker sequences') there exist easily computable such sequences (a_n) even of rational numbers for which there is no computable bound on '$\exists n$' in (1). By contrast, there is a simple primitive recursive (in the sense of [51]) functional $\Phi^*(g, k)$ which provides a bound on (2) (also referred to as 'metastability' in Tao [98]):

Proposition 2.3 (see e.g. [66]). *Let (a_n) be any nonincreasing sequence in $[0, 1]$ then*

$$\forall k \in \mathbb{N} \forall g \in \mathbb{N}^{\mathbb{N}} \exists n \leq \Phi^*(g, k) \forall i, j \in [n; n+g(n)]\ (|a_i - a_j| \leq 2^{-k}),$$

where

$$\Phi^*(g, k) := \tilde{g}^{(2^k)}(0) \text{ with } \tilde{g}(n) := n + g(n).$$

Moreover, there exists an $i < 2^k$ such that n can be taken as $\tilde{g}^{(i)}(0)$.[3]

It is of interest here that the bound $\Phi^*(g, k)$ in proposition 2.3 does not depend on (a_n) at all (see also the discussion of this point further below).

Remark 2.4. *If (a_n) is a sequence of rational numbers so that $\leq$ becomes (primitive recursively) decidable, then using the bound Φ^* and primitive recursive bounded search one can obtain a functional Φ (this time depending on (a_n)) that satisfies the no-counterexample interpretation of this principle. Such a solution*

[3] For $g : \mathbb{N} \to \mathbb{N}$ and $n \in \mathbb{N}$ we define $g^{(0)}(0) := 0,\ g^{(n+1)}(0) := g(g^{(n)}(0))$.

is also possible for sequences (a_n) of real numbers a_n using rational approximations as the proper n.c.i.-treatment of $\leq$ defined in terms of representatives of real numbers given as Cauchy sequences of rational numbers with some fixed rate of convergence would insist on doing. We bypassed this issue since we are only interested in the bound Φ^*.

As an immediate consequence of proposition 2.3 we obtain the following (explicit version of the) 'finite convergence principle' which recently has been discussed by T. Tao ([98, 99]):

Corollary 2.5. *For all $k \in \mathbb{N}, g \in \mathbb{N}^{\mathbb{N}}$ there exists an $M \in \mathbb{N}$ such that for all nonincreasing finite sequences $0 \leq a_M \leq \ldots \leq a_0 \leq 1$ of length $M+1$ in $[0,1]$ there exists an $n \in \mathbb{N}$ with*

$$n + g(n) \leq M \wedge \forall i, j \in [n; n+g(n)](|a_i - a_j| \leq 2^{-k}).$$

Moreover, we can compute M as $M := \Phi^*(g,k)$, where Φ^* is as in proposition 2.3.

Let us now consider general prenex normal sentences

$$A \equiv \exists x_1 \forall y_1 \ldots \exists x_n \forall y_n \, A_{qf}(x_1, y_1, \ldots, x_n, y_n)$$

and their Herbrand normal form (written again with function quantifiers)

$$A^H \equiv \forall f_1, \ldots, f_n \exists x_1, \ldots, x_n \, A_{qf}(x_1, f_1(x_1), \ldots, x_n, f_n(x_1, \ldots, x_n)).$$

Then we say that $\Phi_1, \ldots, \Phi_n$ satisfy the n.c.i. of A if (writing $\underline{f}$ for $f_1, \ldots, f_n$)

$$\forall \underline{f} \, A_{qf}(\Phi_1(\underline{f}), f_1(\Phi_1(\underline{f})), \ldots, \Phi_n(\underline{f}), f_n(\Phi_1(\underline{f}), \ldots, \Phi_n(\underline{f})))$$

holds.
The problem with the n.c.i. is that for sentences that no longer are Π^0_3 the Herbrand normal form in general is too much of a weakening of the original sentence such that a witness (or bound) as required in its no-counterexample interpretation would reflect the correct computational contribution the use of such a sentence in a proof of some Π^0_2-theorem might have. In fact, even for Π^0_3-sentences

$$A \equiv \forall x \exists y \forall z \, A_{qf}(x, y, z)$$

the n.c.i. of any prenex normal form $(A \to B)^{pr}$ of an implication $A \to B$, where

$$B \equiv \forall u \exists v \, B_{qf}(u, v)$$

in Π^0_2 is too weak to allow for a solution of the modus ponens by constructing a realizing function for B (out of functionals satisfying the n.c.i. of A and $(A \to$

$B)^{pr}$) without an explosion in the computational complexity (see [59] for a detailed discussion of the 'modus ponens' problem for n.c.i.).
If

$$A \equiv \exists x \forall y \exists z\, A_{qf}(x, y, z)$$

is in Σ^0_3 (or of higher complexity) then already the n.c.i. of A might fail to give the correct result. This has to do with the fact that in order to infer A back from A^H for such A one needs AC (though – in the case of $A \in \mathcal{L}(\mathrm{PA})$ – only from $\mathbb{N}$ to $\mathbb{N}$ and so rather than a proper choice axiom just a form of arithmetical comprehension) applied to Π^0_1-formulas (or of higher complexity) whereas in the case of Π^0_3-sentences it was needed only for quantifier-free (and hence decidable) formulas. This already matters even for bounded quantifiers '$\exists x \le a$' as can be illustrated by the infinitary pigeonhole principle

$$(\mathrm{IPP}): \forall n \in \mathbb{N} \forall f : \mathbb{N} \to C_n \exists i \le n \forall k \in \mathbb{N} \exists m \ge k\, (f(m) = i),$$

where $C_n := \{0, 1, \ldots, n\}$.
The Herbrand normal form of (IPP) is

$$(\mathrm{IPP})^H \equiv \forall n \in \mathbb{N} \forall f : \mathbb{N} \to C_n \forall F : C_n \to \mathbb{N} \exists i \le n \exists m \ge F(i)\, (f(m) = i)$$

which gives rise to the following computationally almost trivial solution for the n.c.i. of IPP:

$$M(n, f, F) := \max\{F(i) : i \le n\} \text{ and } I(n, f, F) := f(M(n, f, F))$$

are realizers for '$\exists i$' and '$\exists m$' in $(\mathrm{IPP})^H$, despite of the fact that the proof of IPP requires some substantial amount of induction (more precisely the so-called bounded collection principle for Π^0_1-formulas whose strength is in between Σ^0_2- and Σ^0_1-induction).

We use this principle to motivate the Gödel functional ('Dialectica') interpretation D (and its monotone variant), where we refer here always to its combination ND with some negative translation N (e.g. one may use the so-called Shoenfield variant from [95], see [97]). This interpretation produces a $A^{ND} \equiv \forall X \exists Y\, A_*(X, Y)$ normal form for arbitrary formulas A, where A_* is quantifier-free, such that the equivalence between A and A^{ND} follows using only (classical logic and) **quantifier-free** choice

$$\text{QF-AC} : \forall x \exists y\, F_{qf}(x, y) \to \exists Y \forall x\, F_{qf}(x, Y(x)) \quad (F_{qf} \text{ quantifier-free}).$$

The price to be paid for this is that we need QF-AC for objects of arbitrary finite type (over some base type such as $\mathbb{N}$) and also X, Y in A^{ND} will be functionals of higher type (where the types depend on the logical complexity of A).

The ND-interpretation of (IPP) is arrived at in the following way (strictly speaking A^{ND} is defined as the $\exists\forall$-form resulting from a final QF-AC application to our $\forall\exists$-form which we omit here for better readability)

$$\begin{array}{l}
(\text{IPP}) \overset{\text{QF-AC}}{\Leftrightarrow} \\
\forall n \in \mathbb{N} \forall f : \mathbb{N} \to C_n \exists i \leq n \exists g : \mathbb{N} \to \mathbb{N} \forall k \in \mathbb{N} (g(k) \geq k \wedge f(g(k)) = i) \overset{\text{QF-AC}}{\Leftrightarrow} \\
\forall n \in \mathbb{N} \forall f : \mathbb{N} \to C_n \forall K : C_n \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N} \exists i \leq n \exists g : \mathbb{N} \to \mathbb{N} \\
\qquad\qquad (g(K(i,g)) \geq K(i,g) \wedge f(g(K(i,g))) = i) \equiv: (\text{IPP})^{ND}.
\end{array}$$

The functional interpretation of (IPP) requires functionals $I(n, f, K)$ and $G(n, f, K)$ realizing '$\exists i$' and '$\exists g$'.
Note that the implication '$\Leftarrow$' in the second equivalence above only needs computable (in f) and hence continuous functionals K.

These functionals reflect the correct computational contribution of the use of (IPP) in a proof as follows from the general soundness theorem of functional interpretation. I, K can be defined using primitive recursion of type level 1. However, to get a usable notation for these functionals it is advisable to use a finite form of bar recursion as is done in [90]. We refer to that paper as well as [66] for a detailed discussion and urge the reader to try to come up directly with a solution for I, G to appreciate the highly nontrivial task performed by functional interpretation even restricted to a principle as simple as (IPP).

In the applications to analysis discussed in sections 5, 6 and 7 below, one actually is interested in the extraction of bounds (rather than realizers) which, however, need to be **uniform**, i.e. independent from various parameters, to be useful. This can be achieved by modifying Gödel's functional interpretation in such a way that instead of realizers for the functional interpretation so-called majorants (a sort of hereditarily monotone bounds) are extracted. This variant has been introduced in [56] under the name of **monotone functional interpretation**.[4] Not only does this interpretation directly extract uniform bounds, it also nicely extends to important subsystems of analysis (based on the binary 'weak' König's lemma WKL) and simplifies the extraction algorithm (see e.g. [41] for a detailed complexity analysis). In its simplest form the notion of majorizability (due to W.A. Howard [44]) is defined as follows for functionals of finite type over $\mathbb{N}$:

Definition 2.6. *Between functionals of type ρ the following binary relation is defined by induction on ρ :*

$$\begin{cases} x^* \gtrsim_{\mathbb{N}} x :\equiv x^* \geq x, \\ x^* \gtrsim_{\rho\to\tau} x :\equiv \forall y^*, y\, (y^* \gtrsim_\rho y \to x^* y^* \gtrsim_\tau xy). \end{cases}$$

[4]A related so-called bounded functional interpretation was recently designed in Ferreira-Oliva [28] and has interesting applications to systems of feasible analysis (see [29]).

$x \leq_\rho y$ *is defined as the pointwise inequality relation.*

Remark 2.7. *Sometimes, a variant of* $\gtrsim$ *due to Bezem [8] with a clause '*$x^*y^* \gtrsim_\tau x^*y$*' added is useful too.*

Let us recall that a metric space (X, d) is called *complete* if every Cauchy sequence in X converges and it is called *separable* if there exists a countable subset $X_{count} \subseteq X$ such that each $x \in X$ can be arbitrarily good approximated by elements in X_{count}. A complete separable metric space is also called a *Polish metric space*. One way of defining the *compactness* of a complete metric space is by its *total boundedness* which means that for each $\varepsilon > 0$ there are finitely many points $x_1, \ldots x_n \in X$ such that any $x \in X$ is ε-close to one of these points. E.g. $[0,1]^n$ is compact (w.r.t. the Euclidean metric) whereas $\mathbb{R}^n$ is not. Obviously, any compact metric space is separable and so a Polish space. Via an appropriate so-called standard representation of Polish (i.e. complete separable) metric spaces P and compact (i.e. complete and totally bounded) metric spaces K via the Baire space $\mathbb{N}^{\mathbb{N}}$ respectively an appropriate compact subspace $\{f \in \mathbb{N}^{\mathbb{N}} : \forall n \in \mathbb{N}(f(n) \leq M(n)\}$ (for some effective M) it follows that the monotone functional interpretation extracts bounds Φ^* that are independent from parameters ranging over compact metric spaces but depend on representatives $f_x \in \mathbb{N}^{\mathbb{N}}$ for parameters $x \in P$ in Polish spaces P (see theorem 2.10 below).

The uniformity of the bound in proposition 2.3 above (and hence Tao's finite convergence principle) can be seen as an instance of monotone functional interpretation using the representation of the space of sequences in $[0, 1]$ as the compact metric space $[0,1]^{\mathbb{N}}$ (with the product metric).

In the case of (IPP) one obtains majorants I^* and G^* for I, G which – using that the constant-n function majorizes $f : \mathbb{N} \to C_n$ – no longer depend on f but require a majorant K^* for the argument K in the case of G^* (as I^* we simply can take the constant-n functional). Not every functional K does possess a majorant K^* but among others e.g. all K's that are continuous in g (w.r.t. to the product topology on $\mathbb{N}^{\mathbb{N}}$) do. For continuous K one can even replace g by some initial segment encoded in a number m (we use $[m]$ to denote the function which continues this initial segment by zeroes). Then as a consequence of the monotone interpretation and the uniform continuity of K (in g) on $\{g : g \leq_1 G^*(n, K^*)\}$ we get the following semi-finite form

$$\begin{aligned}&\forall n \in \mathbb{N} \forall K : C_n \times \mathbb{N}^{\mathbb{N}} \overset{cont.}{\to} \mathbb{N} \exists M \in \mathbb{N} \forall f : C_M \to C_n \exists i \leq n \exists m \leq M \\ &\qquad (Image([m]) \subseteq C_M \wedge [m](K(i,[m])) \geq K(i,[m]) \wedge f([m](K(i,[m]))) = i)\end{aligned}$$

of (IPP) which is a kind of reformulation of Tao's 'finitary' infinite pigeonhole principle discussed in [98, 99].

Both for carrying out (monotone) functional interpretation as well as for formalizing proofs in analysis, formal systems based on (fragments of) arithmetic PA^ω in the language of functionals in all finite types over $\mathbb{N}$ augmented by suitable analytical principles such as WKL are most convenient. For many such systems $\mathcal{T}^\omega$, e.g. PA^ω+QF-AC+WKL, one can prove results of the following type (using '1' to denote the type $\mathbb{N} \to \mathbb{N}$):

Theorem 2.8 (Kohlenbach [53, 56, 57]). *Let $A_\exists(x^1, y^1, z^{\mathbb{N}}) \equiv \exists \underline{\tilde{z}} A_{qf}(x, y, z, \tilde{z})$, where A_{qf} a quantifier-free formula and $A_\exists$ contains only x, y, z free. Here $\underline{z}$ is any tuple of variables up to type 2. Let s be a closed term.*

$$\begin{cases} \text{From a } \mathcal{T}^\omega\text{-proof of } \forall x^1 \forall y \leq_1 sx \exists z^{\mathbb{N}}\, A_\exists(x, y, z) \\ \text{monotone functional interpretation extracts a closed term } \Phi \text{ of } \mathcal{T}^\omega \text{ such that} \\ \mathcal{T}^\omega \vdash \forall x^1 \forall y \leq_1 sx \exists z \leq \Phi(x)\, A_\exists(x, y, z). \end{cases}$$

Note that the bound $\Phi(x)$ does not depend on y.

Remark 2.9. *In fact, even the existential quantifiers $\exists \underline{z}$ can be bounded (in the sense of $\leq_\rho$ where $\rho (\leq 2)$ is the type of the respective variable).*
If $\mathcal{T}^\omega$ is based on a weak form of extensionality given by a quantifier-free rule (see below), then y might have any type.
Instead of single variables x, y, z we may have tuples as well.

In the case of PA^ω+QF-AC+WKL the bound Φ will be a primitive recursive functional in the sense of Hilbert [43] and Gödel [36]. If PA^ω is restricted to Σ^0_1-induction and primitive recursion in the sense of Kleene only, then Φ will be primitive recursive in the sense of Kleene. If one uses systems of bounded arithmetic such as $\mathrm{G}_2\mathrm{A}^\omega$ (from [57]) instead of PA^ω, then $\Phi(x)$ will be given by a term built up out of $0, S, +, \cdot$ and $x^M(k) := \max\{x(i) : i \leq k\}$ only, i.e. a polynomial in x^M. So in the presence of WKL (corresponding to the Heine-Borel compactness of spaces such as $[0, 1]^n$) as the only genuine analytical principle the complexity of Φ is determined solely by the strength of the underlying arithmetical system. In the presence of (instances of) sequential compactness (corresponding to instances of arithmetical comprehension) this is no longer the case but the contribution of this form of compactness still is rather limited in many cases (see [58]). If higher comprehension over numbers is used or even dependent choice in all types – we denote the resulting system by $\mathcal{A}^\omega$ – then Φ will be a so-called bar recursive functional in the sense of C. Spector ([96], see [66] for a modern treatment).

In order to apply theorem 2.8 to actual proofs in analysis one first has to verify that the theorem to be proved (and to a certain extent its proof)[5] are can be formalized

[5] Proofs of universal lemmas and even of lemmas of the form $\forall u \exists v \leq tu \forall w\, F_{qf}$ for functionals u, v, w of moderate types can be disregarded as such lemmas can be treated simply as axioms, see the remarks in the introduction and [53, 56].

in a suitable formal system which requires a representation of the various analytic objects involved. After that formalization one has to check whether the statement has or can be transformed into (using, if necessary, an appropriate enrichment of data) the required logical form. The process is much simplified using the following kind of 'macro': using the standard (or 'Cauchy'-) representation of Polish spaces P and compact metric spaces K mentioned already above one can reduce quantification over P (resp. over K) to quantification $\forall x^1$ resp. $(\forall y \leq_1 M)$ without introducing new quantifiers. So it suffices to formalize the proof up to quantification over such spaces (provided they are representable in the formal system at hand). This yields the following applied form of theorem 2.8:

Theorem 2.10 (Kohlenbach [54, 56]).

$$\begin{cases} \textit{From a } \mathcal{T}^\omega\textit{-proof of } \forall x \in P \forall y \in K \exists z^{\mathbb{N}} A_\exists(x,y,z) \\ \textit{monotone functional interpretation extracts a closed term } \Phi \textit{ of } \mathcal{T}^\omega \textit{ such that} \\ \mathcal{T}^\omega \vdash \forall x \in P \forall y \in K \exists z \leq \Phi(f_x)\, A_\exists(x,y,z), \end{cases}$$

where Φ depends on a given representative $f_x \in \mathbb{N}^{\mathbb{N}}$ of $x \in P$ but is independent of $y \in K$. Here $A_\exists$ is a purely existential formula as above (when expressed in terms of the representation of P, K) that is (provably in $\mathcal{T}^\omega$) extensional in x, y with respect to $=_X, =_Y$.[6]

For the special Polish spaces $\mathbb{N}$ and $\mathbb{N}^{\mathbb{N}}$ no representation is necessary since $\mathcal{T}^\omega$ contains quantification over these spaces as a primitive concept. Let us illustrate the use of theorem 2.10 by two extremely simple examples:

$$(3)\ \mathrm{PA}^\omega \vdash \forall x \in \mathbb{R} \exists n \in \mathbb{N}\, (n > x)$$

and

$$(4)\ \mathrm{PA}^\omega \vdash \forall x \in (0,1] \exists n \in \mathbb{N}\, (1 < n \cdot x).$$

Clearly, $\mathbb{R}$ has a standard representation as Polish space via the Cauchy completion of $\mathbb{Q}$, i.e. real numbers are represented as Cauchy sequences of rational numbers with fixed rate of convergence (say 2^{-n}) and $>$ (expressed on such representatives) is in Σ^0_1. Hence theorem 2.10 applies and there exists a primitive recursive functional which computes given a representative (r_n) of x a number $n \in \mathbb{N}$ satisfying (3). E.g. we may take $\Phi(r_n) := \lceil r_0 \rceil + 1$.
In (4) there does not exist any subrecursive such Ψ which – given a representative (r_n) of a strictly positive real number x – would produce an n satisfying (4). This corresponds to the fact that one cannot treat $\forall x \in (0,1]$ as a primitive notion (since $(0,1]$ is not complete) but only

$$\forall x \in \mathbb{R}(\exists m \in \mathbb{N}(x \geq 1/(m+1)) \to \exists n \in \mathbb{N}(1 < n \cdot x)),$$

[6]This extensionality is not needed for the extraction of Φ but only to justify that $A_\exists$ indeed speaks about $x \in P$ and $y \in K$ rather than $x, y \in \mathbb{N}^{\mathbb{N}}$.

i.e.

$$\forall x \in \mathbb{R} \forall m \in \mathbb{N} \exists n \in \mathbb{N} \, (x \geq 1/(m+1) \rightarrow 1 < n \cdot x),$$

where now '$x \geq 1/(m+1) \rightarrow 1 > n \cdot x$)' is (equivalent to) a Σ^0_1-formula and '$\forall x \in \mathbb{R} \forall n \in \mathbb{N}$' **can** be treated as primitive quantifiers. Indeed, in (r_n) **plus** m as input the solution is trivial: $\Psi((r_n), m) := m + 2$. Alternatively, one can rewrite (4) using the Polish space $[1, \infty)$ as

$$\mathrm{PA}^{\omega} \vdash \forall x \in [1, \infty) \exists n \in \mathbb{N} (1 < n \cdot (1/x)).$$

We then can take $\Psi(r_n) := \Phi(r_n)$ with Φ from the first example (note that $1/x$ is primitive recursively definable on $[1, \infty)$ but not on $(0, \infty)$ in the Cauchy representation (r_n)). This example a-fortiori shows that if the compactness of K is weakened to total boundedness than one does not have in general uniform bounds anymore (independent from $y \in K$). This also is the case if K is only Polish and bounded (but not totally bounded): Consider the closed unit ball B in the space $C[0, 1]$ of all continuous functions $f : [0, 1] \rightarrow \mathbb{R}$ w.r.t. the uniform norm $\|f\|_\infty := \sup\{|f(x)| : x \in [0, 1]\}$ (note that w.r.t. the metric induced by $\|\cdot\|_\infty$, $C[0, 1]$ is a Polish space and B a bounded Polish space). Then

$$\mathrm{PA}^{\omega} \vdash \forall f \in B \exists n \in \mathbb{N} (n \text{ is code of some } p \in \mathbb{Q}[x] \text{ with } \|f - p\|_\infty < 1/2).$$

Clearly, there is no uniform bound on '$\exists n$' (i.e. a bound independent of $f \in B$) as the sequence $f_n(x) := \sin(nx) \in B$ shows. However, given a representative of f (in the sense of the standard representation of the Polish space $C[0, 1]$) one can easily compute a suitable n.

At first sight, these results essentially seem to show that the requirements on K being complete and totally bounded are necessary to be guaranteed the extractability of bounds Φ that do not depend on parameters y from K. However, these counterexamples only show that this can be the case for concrete spaces K. Looking at the example concerning $B \subset C[0, 1]$ above one realizes that it is the (provable) separability of $C[0, 1]$ that was crucially used (likewise it was the provable incompleteness of $(0, 1]$ that was used in the first counterexample). In fact, to have a B-uniform bound on the very property of separability is nothing else but requiring B to be totally bounded. This opens up the possibility to extract uniform bounds from proofs that do not use the separability of K but rather treat K as an abstract metric space (X, d). Since the only direct way to talk about metric spaces in systems like $\mathcal{T}^{\omega}$ is via their standard representation which is based on separability, one then has to extend $\mathcal{T}^{\omega}$ by an abstract space (X, d) as a kind of atom resulting in a new system $\mathcal{T}^{\omega}[X, d]$. This is done by introducing a new ground type X (for objects in X) and all finite types over $\mathbb{N}$ and X together with a constant d_X plus axioms expressing that d_X is a pseudo-metric. Equality $x =_X y$ is defined as $d_X(x,y) =_{\mathbb{R}} 0_{\mathbb{R}}$, where the real numbers still are represented as type-1 objects and $=_{\mathbb{R}}$ is the equivalence relation on those corresponding to equality in $\mathbb{R}$. Higher type equality is defined extensionally. Whereas in the previous cases the issue of extensionality was not important due to the availability of an elimination-of-extensionality procedure (note that functional interpretation is not sound for full extensionality in higher types) it now becomes crucial that we only include a weak quantifier-free **rule** of extensionality which allows one to infer $r[s] =_\tau r[t]$ only once $s =_\rho t$ has been established. Fortunately, for most applications in fixed point theory and ergodic theory this does not cause any problems since the extensionality of functions $f : X \to X$ usually follows from the f-properties assumed such as f being nonexpansive. In [63] such systems were introduced and very general metatheorems on the extractability of bounds that are independent from parameters $x \in X$ and even $f : X \to X$ were obtained under the only assumption that (X, d) was bounded. In [31] this is further refined and it is shown that the global boundedness assumption on (X, d) can be replaced by local bounds. This was achieved by a new majorizability relation $x^* \gtrsim^a_\rho x$ that is parametrized by some reference point $a \in X$. For a finite type ρ over $\mathbb{N}, X$, the majorant x^* always has a finite type over just $\mathbb{N}$ resulting from replacing in ρ the type X by $\mathbb{N}$. For the ground type X, the relation is defined as follows

$$x^* \gtrsim^a_X x :\equiv x^* \geq_{\mathbb{R}} d_X(a, x),$$

where x has type X while x^* has the type $\mathbb{N}$. For $\rho := X \to X$ the relation is defined between objects f^* of type $1 = (\mathbb{N} \to \mathbb{N})$ and f of type $X \to X$ as follows

$$f^* \gtrsim^a_{X\to X} f :\equiv \begin{cases} \forall n \in \mathbb{N}(f^*(n+1) \geq f^*(n)) \wedge \\ \forall x^* \in \mathbb{N} \forall x \in X(x^* \geq_{\mathbb{R}} d_X(a, x) \to f^*(x^*) \geq_{\mathbb{R}} d_X(a, f(x))). \end{cases}$$

Our approach extracts effective bounds in terms of such majorants $x^* \in \mathbb{N}$, $f^* : \mathbb{N} \to \mathbb{N}$ etc. rather than $x \in X$, $f : X \to X$ and does not presuppose at all that (X, d) comes together with some notion of effectivity on X.

In [63, 31] not only metric spaces (X, d) but also other classes of structures such as hyperbolic spaces, CAT(0)-spaces, normed spaces, uniformly convex spaces and inner product spaces are treated. Further examples ($\mathbb{R}$-trees, δ-hyperbolic spaces and uniformly convex hyperbolic spaces) are discussed in [82]. Since hyperbolic spaces play an important role in section 6 below we give the definition

Definition 2.11 ([63, 34, 92]). *(X, d, W) is called a hyperbolic space if (X, d) is a metric space and $W : X \times X \times [0, 1] \to X$ a function satisfying*

(i) $\forall x, y, z \in X \forall \lambda \in [0, 1](d(z, W(x, y, \lambda)) \leq (1 - \lambda)d(z, x) + \lambda d(z, y))$,

(ii) $\forall x, y \in X \forall \lambda_1, \lambda_2 \in [0, 1](d(W(x, y, \lambda_1), W(x, y, \lambda_2)) = |\lambda_1 - \lambda_2| \cdot d(x, y))$,

(iii) $\forall x, y \in X \forall \lambda \in [0,1](W(x,y,\lambda) = W(y,x,1-\lambda))$,

(iv) $\begin{cases} \forall x,y,z,w \in X, \lambda \in [0,1] \\ (d(W(x,z,\lambda), W(y,w,\lambda)) \le (1-\lambda)d(x,y) + \lambda d(z,w)). \end{cases}$

Definition 2.12. *Let (X,d,W) be a hyperbolic space. The set*

$$seg(x,y) := \{ W(x,y,\lambda) : \lambda \in [0,1] \}$$

is called the metric segment with endpoints x, y.
We usually write $(1-\lambda)x \oplus \lambda y$ for $W(x,y,\lambda)$.

Remark 2.13. *Every convex subset C of a normed linear space is a hyperbolic space with $W(x,y,\lambda) := (1-\lambda)x + \lambda y$.*

In the case of theorems 2.8 and 2.10 it is mainly the concrete numerical information obtained from specific extracted bounds Φ of (in most cases) relatively low complexity which is of interest (see section 5 below) as the existence of a computable uniform bound follows (in the case where all $\exists$-quantifiers in A_{qf} have type $\mathbb{N}$) by unbounded search and the fact that computable functionals $\mathbb{N}^{\mathbb{N}} \times 2^{\mathbb{N}} \to \mathbb{N}$ have computable moduli of uniform continuity $\omega_\Phi : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ for $\Phi(x,\cdot)$ restricted to $y \in 2^{\mathbb{N}}$) so that also $\Phi_M(x) := \max\{\Phi(x,y) : y \in 2^{\mathbb{N}}\}$ (and even $\Phi_M(x) := \max\{\Phi(x,y) : y \le_1 sx\}$) is computable as well. In the case of noncompact spaces, however, even the existence of a uniform bound at all is of interest. We, therefore, use in the following the strongest system $\mathcal{A}^\omega$ extended by an abstract hyperbolic space resulting in $\mathcal{A}^\omega[X,d,W]_{-b}$, where '$-b$' indicates that we do not assume (X,d) to be bounded.

Rather than formulating here one of the general metatheorems from [31] we confine ourselves to some special case which, however, is typical for the kind of results used in obtaining the bounds in sections 6 and 7 below.

Definition 2.14. *Let (X,d) be a metric space. A mapping $f : X \to X$ is called nonexpansive (short: n.e.) if*

$$\forall x, y \in X (d(f(x), f(y)) \le d(x,y)).$$

Theorem 2.15 (Gerhardy - Kohlenbach [31]). *Let $A_\exists$ be an $\exists$-formula and P, K Polish resp. compact metric spaces in standard representation by $\mathcal{A}^\omega$-definable terms.*
If $\mathcal{A}^\omega[X,d,W]_{-b}$ proves a sentence

$$\forall x \in P \forall y \in K \forall z^X, \tilde{z}^X, f^{X \to X}(f \text{ nonexpansive} \to \exists v^{\mathbb{N}} A_\exists)$$

then one can extract a (bar recursively) computable functional $\Phi(g_x, b)$ *s.t. for all* $x \in P, g_x \in \mathbb{N}^{\mathbb{N}}$ *representative of* x*,* $b \in \mathbb{N}$

$$\forall y \in K \forall z, \tilde{z} \in X \forall f : X \to X (f \text{ n.e. } \wedge d(z, f(z)), d(z, \tilde{z}) \le b \to \exists v \le \Phi(g_x, b)\, A_{\exists})$$

holds in **any** *nonempty hyperbolic space* (X, d, W)*.*

The most crucial thing about theorem 2.15 is that the bound Φ depends on (X, d, W), f, z and $\tilde{z}$ **only** via an upper bound $b \ge d(z, \tilde{z}), d(z, f(z))$. In particular, if (X, d) is b-bounded one obtains a bound that is fully independent from $f, z, \tilde{z}$ (this was already proved in Kohlenbach [63]).

Theorem 2.15 also holds for normed spaces (as well as subclasses of those such as Hilbert spaces) if we additionally require that $b \ge \|z\|$ (due to the fact that we now have to take the reference point a in our majorization relation as the zero vector O_X in this case). This is unavoidable as e.g. the following trivial example

$$\forall z^X, \tilde{z}^X \exists y^{\mathbb{N}} (y > \|z\| + \|\tilde{z}\|)$$

shows.

Remark 2.16. *For systems based on* PA^{ω}+QF-AC+WKL *etc. instead of* $\mathcal{A}^{\omega}$ *we obtain bounds of the respective limited complexity classes discussed above.*

3 Applications of proof mining in number theory

A famous theorem of K.F. Roth says

Theorem 3.1 (Roth [93]). An algebraic irrational number α has only finitely many exceptionally good rational approximations, i.e. for $\varepsilon > 0$ there are only finitely many $q \in \mathbb{N}$ such that

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

Weaker results (with smaller exponents) had been obtained before by Dirichlet 1842, Liouville 1844, Thue 1909, Siegel 1921, Schneider 1936 and Dyson and Gelfond 1947 (see [84, 85]).

In 1983, Esnault and Viehweg found a new proof of Roth's theorem. Both proofs are ineffective and prima facie give no bounds neither on the size nor on the number of q's. Nevertheless, Davenport and Roth obtained an exponential bound on the number of q's analyzing Roth's proof.

In 1985 Luckhardt (see [84]) applied a systematic logical analysis (based on Herbrand's theorem in the form (L) discussed in section 2 and using prior work of

Kreisel [77]) to both proofs of Roth's theorems, obtaining from Roth's proof a bound which roughly is the fourth root of the bound found by Davenport and Roth and from the proof due to Esnault and Viehweg the first polynomial bound on the number of q's, more precisely:

Theorem 3.2 (Luckhardt [84]). The following upper bound on $\#\{q : R(q)\}$ holds:

$$\#\{q : R(q)\} < \frac{7}{3}\varepsilon^{-1} \log N_\alpha + 6 \cdot 10^3 \varepsilon^{-5} \log^2 d \cdot \log(50\varepsilon^{-2} \log d),$$

where $N_\alpha < \max(21 \log 2h(\alpha), 2\log(1 + |\alpha|))$ and h is the logarithmic absolute homogeneous height.

Independently, Bombieri and van der Poorten obtained in 1988 [10] a roughly similar bound using a more ad hoc strategy of proof.

4 Applications of proof mining in algebra

Artin's solution of Hilbert's 17th problem can be formulated as follows (see e.g. [85]):
Let k be an ordered field and let R be a real-closed order extension of k. If $f \in k[x_1, \ldots, x_n]$ of degree d is positive semi-definite over R, then f can be represented as a non-negative weighted sum of squares of the form

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{\lambda} p_i \cdot g_i(x_1, \ldots, x_n)^2,$$

where $p_i \in k$, $p_i \geq 0$ and $g_i \in k(x_1, \ldots, x_n)$.
Artin's proof can be formalized in weak formal systems based on the weak König's lemma WKL (which are conservative over Primitive Recursive Arithmetic PRA as can be shown e.g. by monotone functional interpretation, see [53, 57]). This was already observed by Kreisel in the late 50's who concluded from this the existence of primitive recursive bounds (in n, d) for λ and the degrees of the rational functions involved. If one carries this out one obtains an exponential tower whose height is given by the number of variables. Later L. Henkin showed that the p_i and the coefficients of the g_i can be piecewise-rationally constructed from the coefficients of f. Kreisel asked whether such a case distinction is necessary and whether there would exist a continuous solution. These questions where settled completely by Delzell in a series of papers ([22, 23, 24, 25], see [26] for a thorough discussion of the role of proof theory played in these developments). The details are too technical to be stated here.

Another type of applications of proof theory is presented by recent work of T.

Coquand and H. Lombardi (see e.g. [19, 20, 21]). Here the use of ideal (so-called strict Π^1_1-) statements (on objects such as prime ideals or maximal ideals) in abstract algebra is replaced by elementary (Σ^0_1) syntactical ones which can effectively be witnessed.[7] This effective reduction of strict Π^1_1-formulas to Σ^0_1-formulas again can be viewed as a WKL-elimination mentioned in the introduction (WKL in this work shows up indirectly via the completeness theorem for propositional logic).

Among other things this has led to a new non-Noetherian version of Serre's splitting-off theorem (1958) and the Forster-Swan theorem (1964-67) improving results of Heitmann from 1984.

5 Application of proof mining in approximation theory

An important classical theorem in best approximation theory is the following:

Theorem 5.1 (Jackson [46]). Let $f \in C[0,1]$ and $n \in \mathbb{N}$. There exists a unique polynomial $p_b \in P_n$ of degree $\leq n$ that approximates f best in the L_1-norm, i.e.

$$\|f - p_b\|_1 = \inf_{p \in P_n} \|f - p\|_1 =: dist_1(f, P_n).$$

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

Both the (easy) existence as well as the (difficult) uniqueness part are proved ineffectively involving noncomputable real numbers in the form of – logically speaking – WKL. Applying the extraction algorithm provided by the proof of theorem 2.10, the following result was extracted from another ineffective uniqueness proof due to E.W. Cheney [18]:

Theorem 5.2 (Kohlenbach-Oliva [73]). 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

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

---

[7] A is strict Π^1_1 if is has the form $\forall X \exists y A_{qf}$, where X and y are set resp. number variables and A_{qf} is quantifier-free. Alternatively, one can uses quantitication over 0/1-functions instead of X.

Then $\Phi(\omega, n, \varepsilon)$ is an effective rate of strong unicity 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 $f \in C[0,1]$

$$\forall p_1, p_2 \in P_n; \varepsilon \in \mathbb{Q}^*_+ \, (\bigwedge_{i=1}^{2} (\|f - p_i\|_1 \leq dist_1(f, P_n) + \Phi(\omega, n, \varepsilon)) \rightarrow \|p_1 - p_2\|_1 \leq \varepsilon),$$

where ω is a modulus of uniform continuity of the function f, i.e.

$$\forall \varepsilon \in \mathbb{Q}^*_+ \forall x, y \in [0,1](|x - y| < \omega(\varepsilon) \rightarrow |f(x) - f(y)| < \varepsilon).$$

Note that Φ only depends on f only via the modulus ω.

What makes Φ a rate of 'strong unicity' in the sense of numerical analysis is that it does not depend at all on $p_1, p_2 \in P_n$. This can be explained in terms of theorem 2.10 as it is easy to see that it suffices to construct such a rate on the compact subset $K_{f,n} := \{p \in P_n : \|p\|_1 \leq \frac{5}{2}\|f\|_1\}$ since such a rate can be extended to whole P_n (see [54, 73, 66]). Strong unicity plays a vast role in approximation theory. In particular, any rate of strong unicity provides a modulus of pointwise continuity ('stability') of the corresponding projection operator that maps a function to its unique best approximation. The fact that Φ does not depend on f as such but only on ω is also guaranteed a priori based on theorem 2.10 (see [73]).

Remark 5.3. *The fact that the bound depends on ω comes from the representation of $C[0,1]$ as Polish space w.r.t. $\|\cdot\|_\infty$. $(C[0,1], \|\cdot\|_1)$ is easily seen to be incomplete and hence not a Polish space!*

Although the uniqueness of the best L_1-approximation was known since 1921, only in 1975 Björnestål [9] proved the existence of a rate of strong unicity 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 describe $c_{f,n}$ explicitly. In 1978, Kroó [80] 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. Again, the constant is not presented. Kroó also showed that the ε-dependency established by Björnestal is optimal. Since the above rate also has this optimal dependency, theorem 5.2 is an explicit effective version of Kroó's result (see the detailed discussion in [66]). This effective rate of strong unicity allows one for the first time to give a subrecursive algorithm for the computation of the best approximation for general $f \in C[0,1]$. The complexity of that procedure is analyzed in [89].

Effective bounds on strong unicity for best Chebycheff approximation were extracted in [54, 55] improving earlier results of D. Bridges and K.-I. Ko.

6 Application of proof mining in metric fixed point theory

A substantial part of metric fixed point theory studies the fixed point (and approximate fixed point) property of nonexpansive selfmappings $f : C \to C$ of convex subsets of normed spaces $(X, \|\cdot\|)$ or – more generally – hyperbolic spaces (which includes the class of CAT(0)-spaces), see e.g. [50] and – for fixed point theory in the context of CAT(0)-spaces – [48, 49]. Whereas the fixed point theory of contractions, i.e. of mappings $f : X \to X$ satisfying

$$\forall x, y \in X \, (d(f(x), f(y)) \leq c \cdot d(x, y))$$

for some $c \in (0, 1)$, essentially is trivial and fully effective due to the Banach fixed point theory, this is radically different for the more general class of nonexpansive mappings. Here in general neither do fixed points exist nor are they necessarily unique in cases where they do exist. Moreover, even in the case of a unique fixed point, the trivial iteration $f^n(x)$ might not converge to the fixed point as the example $f : [0, 1] \to [0, 1]$, $f(x) := 1 - x$ shows: $f^n(x)$ converges to the unique fixed point $1/2$ iff $x = 1/2$. As a result of these difficulties, the fixed point theory of nonexpansive mappings is one of the most active research areas in nonlinear analysis. However, the fixed point theory of such mappings still has a certain computational flavor as one can define other effective iteration schemata which under general conditions converge towards a fixed point or - at least - provide approximate fixed point sequences.

The most common schema is the so-called Krasnoselski-Mann iteration which for a given sequence (λ_n) in $[0, 1]$ and starting point $x_0 \in C$ is defined as follows (for the general case of hyperbolic spaces)

$$x_{n+1} := (1 - \lambda_n)x_n \oplus \lambda_n f(x_n).$$

In the following theorems we assume (following [45]) that $(\lambda_k)_{k \in \mathbb{N}}$ satisfies the following conditions:

- (λ_k) is divergent in sum,
- $\exists K \in \mathbb{N} \forall k \in \mathbb{N} (\lambda_k \in [0, 1 - \frac{1}{K}])$.

Theorem 6.1 (Ishikawa [45]).
Under the assumptions above the following holds:

$$(x_n)_{n \in \mathbb{N}} \text{ bounded } \to \|x_n - f(x_n)\| \overset{n \to \infty}{\to} 0.$$

The theorem has been extended to hyperbolic spaces in [34]. Both proofs are prima facie ineffective and do not provide any rate of convergence.

Another important result is the Borwein-Reich-Shafrir theorem

Theorem 6.2 (Borwein-Reich-Shafrir [11]). *Let* (X, d, W) *be a hyperbolic space,* $f : X \to X$ *nonexpansive and* (λ_n) *as above. Then*

$$d(x_n, f(x_n)) \overset{n\to\infty}{\to} r(f) := \inf\{d(y, f(y)) : y \in X\}.$$

Note that in the Borwein-Reich-Shafrir theorem, (x_n) is not assumed to be bounded. The Ishikawa-Gobel-Kirk theorem combined with the Borwein-Reich-Shafrir theorem implies that

$$d(x_n, f(x_n)) \overset{n\to\infty}{\to} 0$$

if there exists an x^* such that the Krasnoselski-Mann iteration starting from x^* is bounded. In [70] (and – for the normed case – in [61]) a logical analysis of the ineffective proof of the Borwein-Reich-Shafrir theorem based on the extraction algorithm from the proof of theorem 2.15 has been carried out resulting in a quantitative version of the latter guaranteed a priori by theorem 2.15. Combined with the (mere truth of the) Ishikawa-Goebel-Kirk theorem the following bound on this convergence is extracted (since $(d(x_n, f(x_n)))_n$ is nonincreasing the convergence towards 0 can be written as a Π^0_2-statement so that theorem 2.15 applies):

Theorem 6.3 (Kohlenbach-Leuştean [70]). Let (X, d, W) be a hyperbolic space and $f : X \to X$ be a nonexpansive mapping, $(\lambda_n)_{n\in\mathbb{N}}, \alpha$ and K be such that $\alpha : \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ be such that $\lambda_n \in [0, 1 - \frac{1}{K}]$ and

$$\forall i, n \in \mathbb{N}((\alpha(i, n) \leq \alpha(i + 1, n)) \wedge (n \leq \sum_{s=i}^{i+\alpha(i,n)-1} \lambda_s)).$$

Let $b > 0, x, x^* \in X$ be such that

$$d(x, x^*) \leq b \wedge \forall n, m \in \mathbb{N}(d(x^*_n, x^*_m) \leq b).$$

Then the following holds

$$\forall \varepsilon > 0 \forall n \geq h(\varepsilon, b, K, \alpha)(d(x_n, f(x_n)) \leq \varepsilon),$$

where

$h(\varepsilon, b, K, \alpha) := \widehat{\alpha}(\lceil 10b \cdot \exp(K(M + 1))\rceil - 1, M)),$ with
$M \in \mathbb{N}$ such that $M \geq \frac{1+4b}{\varepsilon}$ and $\widehat{\alpha}(0, n) := \tilde{\alpha}(0, n),\ \widehat{\alpha}(i + 1, n) := \tilde{\alpha}(\widehat{\alpha}(i, n), n)$ with
$\tilde{\alpha}(i, n) := i + \alpha(i, n)\ \ (i, n \in \mathbb{N}).$

Note that the rate of convergence depends on x, f, X only via b and on (λ_n) only via α, K. In [70], many more results of this type are proved.

Since our notion of hyperbolic space, in particular, contains all CAT(0)-spaces, the result applies to these spaces as well.

For the special case of (convex subsets of) normed spaces the result was proved already in [61] and [62]. For the general logical background see [63, 74]. In [63, 31] it is shown that the quantitative version of such a kind is guaranteed by a general logical metatheorem whose proof provides an algorithm for the extraction of the bound in this and many other contexts.

The result implies, in particular, that for bounded X the convergence $d(x_n, f(x_n)) \to 0$ is uniform in x and f (ineffectively this is due to [34]).

Since (x^*_n) is assumed to be bounded, a natural question to be addressed by proof mining is to analyze how much of this boundedness actually is needed. As a consequence of the fact that the above result is based on the logical analysis of the Borwein-Reich-Shafrir theorem (which does not involved any boundedness assumption) and uses only the truth of the Ishikawa-Goebel-Kirk theorem (rather than its proof), this question is not answered by these results but requires a direct logical analysis of the Ishikawa-Goebel-Kirk theorem which (for $x^* := x$) gives the following answer:

Theorem 6.4 (Kohlenbach [66]). *Let (X, d, W) be a nonempty hyperbolic space, $f : X \to X$ a nonexpansive mapping, $(\lambda_n), K, \alpha$ as in theorem 6.3, $x \in X$ and (x_n) the Krasnoselski-Mann iteration of f starting from x and $\tilde{b} > 0$ so that $d(x, f(x)) \le \tilde{b}$. Then for every $\varepsilon, b > 0$ the following holds (abbreviating $h^*(\varepsilon, b, \tilde{b}, K, \alpha)$ by h^*):*

$$\forall i \le h^* \forall j \le \alpha(h^*, M)\ (d(x_i, x_{i+j}) \le b) \to \forall n \ge h^*\ (d(x_n, f(x_n)) < \varepsilon),$$

where

$$h^*(\varepsilon, b, \tilde{b}, K, \alpha) := \widehat{\alpha}\left(\left\lceil \frac{\tilde{b} \cdot \exp\left(K \cdot \left(\frac{3\tilde{b}+b}{\varepsilon} + 1\right)\right)}{\varepsilon} \right\rceil \dot{-} 1, M\right)$$

with $\widehat{\alpha}$ as before.

For the case of sequences (λ_n) in $[a, b]$ with $0 < a < b < 1$ we obtain from theorem 6.4 the following qualitative improvement of the Ishikawa-Goebel-Kirk theorem concerning the requirement of (x_n) being bounded (which for the case of constant $\lambda_n := \lambda \in (0, 1)$ and convex subsets of normed spaces was first observed in [2] (theorem 2.1)):

Corollary 6.5 (Kohlenbach [66]). *Let (X, d, W) be a hyperbolic space and $f : X \to X$ nonexpansive. For $x \in X$ and (λ_n) in $[a, b]$, where $0 < a < b < 1$, let (x_n) be the corresponding Krasnoselski-Mann iteration of f starting from x. Let*

$$c(n) := \max\{d(x, x_j) : j \le n\}.$$

Then

$$\lim_{n\to\infty} \frac{c(n)}{n} \to 0$$

implies that

$$\lim_{n\to\infty} d(x_n, f(x_n)) = 0.$$

The proof of this corollary is based on the fact that for $K \in \mathbb{N}$, $K \geq 2$ satisfying $\lambda_n \in [\frac{1}{K}, 1 - \frac{1}{K}]$ for all $n \in \mathbb{N}$ one can take $\alpha(i, M) := K \cdot M$.

Remark 6.6. *The previous result shows that* $d(x_n, f(x_n)) \to 0$ *provided* (x_n) *grows with a lower than linear (in n) rate. This is optimal in the sense that linear growth does not suffice as follows from the following simple example:* $X := \mathbb{R}$, $f(x) := x+1$ *and* $\lambda := \frac{1}{2}$. *For the starting point* $x_0 := 0$ *we have for the Krasnoselski-Mann iteration* (x_n) *that* $x_n = \frac{n}{2}$, *but* $d(x_n, f(x_n)) = 1$ *for all* $n \in \mathbb{N}$.

Since the fundamental paper Goebel-Kirk [33], an extension of the class of nonexpansive functions has been studied extensively in fixed point theory:

Definition 6.7. *Let* (X, d) *be a metric space. A function* $f : X \to X$ *is called asymptotically nonexpansive if for some sequence* (k_n) *in* $[0, \infty)$ *with* $\lim_{n\to\infty} k_n = 0$ *one has (with* f^n *denoting the n-th iteration of f)*

$$d(f^n x, f^n y) \leq (1 + k_n) d(x, y), \quad \forall n \in \mathbb{N}, \forall x, y \in X.$$

In the case of asymptotically nonexpansive mappings one considers the following version of the Krasnoselski-Mann iteration:

$$x_0 := x, \;\; x_{n+1} := (1 - \lambda_n)x_n + \lambda_n f^n(x_n).$$

Definition 6.8 ([35, 82]). *A hyperbolic space* (X, d, W) *is uniformly convex if for any* $r > 0$ *and any* $\varepsilon \in (0, 2]$ *there exists* $\delta \in (0, 1]$ *such that for all* $a, x, y \in X$,

$$\left.\begin{array}{l} d(x,a) \leq r \\ d(y,a) \leq r \\ d(x,y) \geq \varepsilon r \end{array}\right\} \quad \Rightarrow \quad d\left(\frac{1}{2}x \oplus \frac{1}{2}y, a\right) \leq (1-\delta)r. \tag{1}$$

A mapping $\eta : (0, \infty) \times (0, 2] \to (0, 1]$ *providing such a* $\delta := \eta(r, \varepsilon)$ *for given* $r > 0$ *and* $\varepsilon \in (0, 2]$ *is called a* modulus of uniform convexity.
We say that η *is* **monotone** *if it decreases with r (for any fixed* ε*).*

Examples of uniformly convex hyperbolic spaces (with monotone modulus of uniform convexity) are Hilbert spaces as well as CAT(0)-spaces which are the generalization of Hilbert spaces satisfying instead of the parallelogram equality only a parallelogram inequality (also called Bruhat-Tits inequality; see e.g. [12]).

For Krasnoselski-Mann iterations (x_n) of asymptotically nonexpansive mappings f we in general no longer have that $(d(x_n, f(x_n)))_n$ is noninceasing. Hence we this time have to formulate our quantitative bound in terms of 'metastability' rather than the Π^0_3-form of convergence. Using an appropriate form of theorem 2.15 the following result has been obtained (see also Kohlenbach-Lambov [69] which treats the uniformly convex **normed** case and the logical discussion provided there):

Theorem 6.9 (Kohlenbach-Leuştean [72]). *Let (X, d, W) be a uniformly convex hyperbolic space with a monotone modulus of uniform convexity η and $f : X \to X$ be asymptotically nonexpansive with sequence (k_n).*
Assume that $K \geq 0$ is such that $\sum\limits_{n=0}^{\infty} k_n \leq K$ and that $L \in \mathbb{N}, L \geq 2$ is such that $\frac{1}{L} \leq \lambda_n \leq 1 - \frac{1}{L}$ for all $n \in \mathbb{N}$.
Let $x \in X$ and $b > 0$ be such that for any $\delta > 0$ there is $p \in X$ with

$$d(x, p) \leq b \wedge d(f(p), p) \leq \delta. \tag{2}$$

Then for all $\varepsilon \in (0, 1]$ and for all $g : \mathbb{N} \to \mathbb{N}$,

$$\exists N \leq \Phi(K, L, b, \eta, \varepsilon, g) \forall m \in [N, N + g(N)] \left(d(x_m, f(x_m)) < \varepsilon\right), \tag{3}$$

where

$$\begin{aligned}
&\Phi(K, L, b, \eta, \varepsilon, g) := h^{(M)}(0), \quad h(n) := g(n+1) + n + 2,\\
&M := \left\lceil \frac{3\left(5KD + D + \frac{11}{2}\right)}{\delta} \right\rceil, \quad D := e^K(b+2),\\
&\delta := \frac{\varepsilon}{L^2 F(K)} \cdot \eta\left((1+K)D + 1, \frac{\varepsilon}{F(K)((1+K)D+1)}\right),\\
&F(K) := 2(1 + (1+K)^2(2+K)).
\end{aligned}$$

Moreover, $N = h^{(i)}(0) + 1$ for some $i < M$.

The special case $g(n) := 0$ yields that for M as defined above

$$(*)\ \exists N \leq 2M\ (d(x_N, f(x_N)) < \varepsilon).$$

In the case of CAT(0)-spaces one obtains a quadratic bound on N in $(*)$ (see [72]) which even for nonexpansive mappings and the special case of Hilbert space X is expected to be optimal.
For further results obtained by proof mining in the context of fixed point theory see [13, 14, 15, 16, 30, 60, 62, 64, 66, 69, 70, 71, 72, 83, 81]. Most of these results (except for [72] partly summarized above) are included in the survey [67].

7 Applications of proof mining in ergodic theory

Ergodic theory has close connections to metric fixed point theory and nonexpansive mappings f again play an important role. In particular, one studies the asymptotic behavior of the averaging operator defined by

$$A_n(x) := \frac{1}{n} S_n(x), \text{ where } S_n(x) := \sum_{i=0}^{n-1} f^i(x).$$

The context, typically, is that of Hilbert spaces (or, more generally, uniformly convex Banach spaces).
A classical result is the following

Theorem 7.1 (von Neumann mean ergodic theorem). *Let X be a Hilbert space and $f : X \to X$ a nonexpansive linear operator. Then for any point $x \in X$ the sequence $(A_n(x))_n$ defined above converges (in the Hilbert space norm).*

In [1], a detailed unwinding of a standard (ineffective) proof of this theorem is given. Since, as the authors show, there is (in general) no computable bound on the convergence itself, one again has (in order to obtain computational information) to move to the no-counterexample version (i.e. metastability in the sense of Tao) of this result:

$$\forall g : \mathbb{N} \to \mathbb{N}\, \forall \varepsilon > 0\, \forall x \in X\, \exists n \forall i, j \in [n; n + g(n)]\, (\|A_i(x) - A_j(x)\| < \varepsilon).$$

From the general logical metatheorem proved in [31] and discussed in section 2 above (see theorem 2.15) it can be inferred (after some preprocessing of the proof) that one can extract a computable bound $\Phi(d, \varepsilon, g)$ on $\exists n$ that only depends on a norm upper bound $d \geq \|x\|, \varepsilon$ and g (note that since f is linear and nonexpansive, one has $\|f(x)\| \leq \|x\|$ so that $\|f(x) - x\| \leq 2\|x\|$).
In [1] the following bound Φ is extracted:

Theorem 7.2 (Avigad-Gerhardy-Towsner [1]). *Let X and f be as in theorem 7.1. Then*

$$\forall g : \mathbb{N} \to \mathbb{N}\, \forall \varepsilon > 0\, \forall x \in X\, \exists n \leq \Phi\, \forall i, j \in [n; n + g(n)]\, (\|A_i(x) - A_j(x)\| \leq \varepsilon),$$

where $\Phi = h^{(k)}(0)$ *with* $\rho := \lceil \frac{\|x\|}{\varepsilon} \rceil,\ k := 2^9 \rho^2,\ h(n) := n + 2^{13} \rho^4 \tilde{g}((n+1)\tilde{g}(2n\rho)\rho^2)$ *and* $\tilde{g}(n) := \max\{i + g(i) : i \leq n\}$.

An area closely related to ergodic theory is ‘topological dynamics’. For an early use of proof mining in connection with Furstenberg and Weiss’ proof (based on topological dynamics) of van der Waerden’s theorem see [32].

8 Applications of proof mining in computer science

In the previous section we have shown show interesting new computational information can be extracted using proof theoretic tools even from prima facie highly ineffective proofs in different areas of mathematics. This extraction of computational information in itself is an application of proof mining that it relevant to computer science as it opens up a general logic-based approach toward computational mathematics.

In this section we mention briefly some more specific uses of proof mining to proofs **inside** of theoretical computer science and logic as well experiments with machine-extractions of algorithms from (simple) proofs.

A normalization algorithm for typed λ-terms (and various extensions thereof) that has received quite some attention is the so-called 'normalization by evaluation' algorithm first described in Berger-Schwichtenberg [6] (though it has some roots in early work of Martin-Löf as well). While this algorithm was first found without the use of proof mining it was shown by U. Berger in [4] that it in fact coincides with the algorithm extractable from the standard Tait-Troelstra strong normalization proof using Kreisel's so-called modified realizability (which can be viewed as a simplified form of functional interpretation that suffices here since the normalization proof can be carried out with constructive logic; see [66] for a detailed discussion of modified realizability and its relation to functional interpretation). Whereas the extraction in [4] was done by hand, more recently, a machine extraction of this algorithm has been carried out in [5] thereby comparing different tools: MinLog, Coq and Isabelle/HOL.

Normalization by evaluation can be viewed as a kind of reduction-free algorithm (based on a logical relation used to define the computability predicate).

In [52] we gave a reduction-free proof (based on monotone functional interpretation) for the fact that primitive recursive functionals in the sense of Gödel of type $\mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ are uniformly continuous on $\{f : f \leq_1 g\}$ with a primitive recursive (in g) modulus of uniform continuity. Recently, M.-D. Hernest used his implementation ([38]) of (an optimized so-called 'light' version [37] of) monotone functional interpretation to machine-extract such moduli for concrete functionals with essentially optimal results (see [40]).

Various experiments with machine extractions of algorithms from (though rather simple) proofs in elementary arithmetic, combinatorics and algebra have been carried out e.g. in Berger-Schwichtenberg [7] (using a combination of modified realizability and a so-called refined Friedman/Dragalin-translation) which, in particular, treats Dickson's lemma and in Hernest [39] (using 'light' functional interpretation). An approach based on functional interpretation of Dickson's lemma and the Hilbert Basis Theorem, carried out (on paper) by A. Hertz [42], provides particularly good results in terms of complexity theory. From a practical point of view the algorithm extracted in Raffalli [91] (based on methods related to Krivine [79]) seems to be quite efficient in test cases. In [94], Schwichtenberg uses functional interpretation to extract an algorithm close to Euclid's from an almost trivial proof.

An interesting use of the Shoenfield variant of functional interpretation for the extraction of a new cut-elimination algorithm from the ineffective semantical proof of cut-elimination was recently given by G. Mints [87]. Conservation results of appropriate forms of weak König's lemma over systems of feasible analysis have been proved in Ferreira-Oliva [29] using their bounded functional interpretation ([28]).

References

[1] Avigad, J., Gerhardy, P., Towsner, H., Local stability of ergodic averages. arXiv:0706.1512v1 [math.DS] (2007).

[2] Baillon,J.B., Bruck, R.E., Reich, S., On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces. Houston J. Math. **4**, pp. 1-9 (1978).

[3] Bellin, G., Ramsey interpreted: a parametric version of Ramsey's theorem. In: Logic and computation (Pittsburgh, PA, 1987), pp. 17-37, Contemp. Math., 106, Amer. Math. Soc., Providence, RI, (1990).

[4] Berger, U., Program extraction from normalization proofs. In: M. Bezem et al.(eds.), TLCA'93, pp. 91-106. Springer LNCS **664** (1993).

[5] Berger, U., Berghofer, S., Letouzey, P., Schwichtenberg, H., Program extraction from normalization proofs. Studia Logica **82**, pp. 25-49 (2006).

[6] Berger, U., Schwichtenberg, H., An inverse of the evaluation functional for typed λ-calculus. In: Vemuri, R. (ed.), Proc. of the Sixth Annual IEEE Symposium on Logic in Computer Science, pp. 203-211. IEEE Computer Society, Los Alamitos, 1991.

[7] Berger, U., Schwichtenberg, H., Seisenberger, M., The Warshall Algorithm and Dickson's Lemma: Two examples of realistic program extraction. Journal of Automated Reasoning **26**, pp. 205-221 (2001).

[8] Bezem, M., Strongly majorizable functionals of finite type: a model for bar recursion containing discontinuous functionals. J. Symbolic Logic **50** pp. 652–660 (1985).

[9] Björnestal, B.O., Continuity of the metric projection operator I-III. The preprint series of Department of Mathematics. Royal Institute of Technology. Stockholm, TRITA-MAT **17** (1974), **20** (1974), **12** (1975).

[10] Bombieri, E., van der Poorten, A.J., Some quantitative results related to Roth's theorem. J. Austral. Math. Soc. (Series A) **45**, pp. 233-248 (1988).

[11] Borwein, J., Reich, S., Shafrir, I., Krasnoselski-Mann iterations in normed spaces. Canad. Math. Bull. **35**, pp. 21-28 (1992).

[12] Bridson, M.R., Haefliger, A., Metric spaces of non-positive curvature. Springer Verlag, Berlin (1999).

[13] Briseid, E., Proof Mining Applied to Fixed Point Theorems for Mappings of Contractive Type. Master Thesis, 70pp., Oslo 2005.

[14] Briseid, E.M., Fixed points of generalized contractive mappings. To appear in: J. Nonlinear and Convex Analysis.

[15] Briseid, E.M., A rate of convergence for asymptotic contractions. J. Math. Anal. Appl. **330**, pp. 364-376 (2007).

[16] Briseid, E.M., Some results on Kirk's asymptotic contractions. Fixed Point Theory **8**, No.1, pp. 17-27 (2007).

[17] Briseid, E.M., Logical aspects of rates of convergence in metric spaces. In preparation.

[18] Cheney, E.W., Approximation Theory. AMS Chelsea Publishing, 1966.

[19] Coquand, Th., Sur un théorème de Kronecker concernant les variétés algébriques. C.R. Acad. Sci. Paris, Ser. I **338**, pp. 291-294 (2004).

[20] Coquand, Th., Lombardi, H., Quitte, C., Generating non-Noetherian modules constructively. Manuscripta mathematica **115**, pp. 513-520 (2004).

[21] Coste, M., Lombardi, H., Roy, M.F., Dynamical methods in algebra: effective Nullstellensätze. Ann. Pure Appl. Logic **111**, pp. 203-256 (2001).

[22] Delzell, C., Continuous sums of squares of forms. In: Proc. L.E.J. Brouwer Centenary Symposium. Noordwijkerhout, pp. 65-75 (1981).

[23] Delzell, C., Case distinctions are necessary for representing polynomials as sums of squares. Proc. Herbrand Symposium, pp. 87-103 (1981).

[24] Delzell, C., A finiteness theorem for open semi-algebraic sets, with applications to Hilbert's 17th problem. Contemporary Math. **8**, pp. 79-97 (1982).

[25] Delzell, C., A continuous, constructive solution to Hilbert's 17th problem. Invent. math. **76**, pp. 365-384 (1984).

[26] Delzell, C., Kreisel's unwinding of Artin's proof-Part I. In: Odifreddi, P., Kreiseliana, 113-246, A K Peters, Wellesley, MA (1996).

[27] Feferman, S., Kreisel's 'Unwinding Program'. In: P. Odifreddi (ed.), Kreiseliana: about and around Georg Kreisel, A.K. Peters, Wellesley Massachusetts, pp. 247-273 (1996).

[28] Ferreira, F., Oliva, P., Bounded functional interpretation. Ann. Pure Appl. Logic **135**, pp. 73-112 (2005).

[29] Ferreira, F., Oliva, P., Bounded functional interpretation and feasible analysis. Ann. Pure Appl. Logic **145**, pp. 115-129 (2007).

[30] Gerhardy, P., A quantitative version of Kirk's fixed point theorem for asymptotic contraction. J. Math. Anal. Appl. **316**, pp. 339-345 (2006).

[31] Gerhardy, P., Kohlenbach, U., General logical metatheorems for functional analysis. To appear in: Trans. Amer. Math. Soc.

[32] Girard, J.-Y., Proof Theory and Logical Complexity Vol.I. Studies in Proof Theory. Bibliopolis (Napoli) and Elsevier Science Publishers (Amsterdam) 1987.

[33] Goebel, K., Kirk, W.A., A fixed point theorem for asymptotically nonexpansive mappings, Proc. Amer. Math. Soc. 35 (1972), 171-174.

[34] Goebel, K., Kirk, W.A., Iteration processes for nonexpansive mappings. In: Singh, S.P., Thomeier, S., Watson, B., eds., Topological Methods in Nonlinear Functional Analysis. Contemporary Mathematics **21**, AMS, pp. 115-123 (1983).

[35] Goebel, K., Reich, S., Uniform convexity, hyperbolic geometry, and nonexpansive mappings, Marcel Dekker, Inc., New York and Basel, 1984.

[36] Gödel, K., Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. Dialectica **12**, pp. 280–287 (1958).

[37] Hernest, M.-D., Light functional interpretation. An optimization of Gödel's technique twoards the extraction of (more) efficient programs from (classical) proofs. In: Ong, L. (ed.), CSL 2005, Springer LNCS **3634**, pp. 477-492 (2005).

[38] Hernest, M.-D., The MinLog proof-system for Dialectica program-extraction. Free software available at http://www.brics.dk/ danher/MinLogForDialectica.

[39] Hernest, M.-D., Light Dialectica program extraction for a classical Fibonacci proof. Electonic Notes in Theoretical Computer Science **171**, pp. 43-53 (2007).

[40] Hernest, M.-D., Synthesis of moduli of uniform continuity by the monotone Dialectica interpretation in the proof-system MinLog. Electronic Notes in Theoretical Computer Science **174**, pp. 141-149 (2007).

[41] Hernest, M.-D., Kohlenbach, U., A complexity analysis of functional interpretations. Theoretical Computer Science **338**, pp. 200-246 (2005).

[42] Hertz, A., A constructive version of the Hilbert basis theorem. Master Thesis, Carnegie Mellon University (2004).

[43] Hilbert, D., Über das Unendliche. Math. Ann. **95**, pp. 161-190 (1926).

[44] Howard, W.A., Hereditarily majorizable functionals of finite type. In: Troelstra (ed.), Metamathematical investigation of intuitionistic arithmetic and analysis, pp. 454-461. Springer LNM 344 (1973).

[45] Ishikawa, S., Fixed points and iterations of a nonexpansive mapping in a Banach space. Proc. Amer. Math. Soc. **59**, pp. 65-71 (1976).

[46] Jackson, D., Note on a class of polynomials of approximation. Trans. Amer. Math. Soc. **22**, pp. 320-326 (1921).

[47] Ketonen, J., Solovay, R., Rapidly growing Ramsey functions. Ann. Math. **113**, pp. 267-314 (1981).

[48] Kirk, W.A., Geodesic geometry and fixed point theory. Seminar of Mathematical Analysis (Malaga/Seville, 2002/2003), pp. 195-225, Colecc. Abierta, 64, Univ. Seville Secr. Publ., Seville (2003).

[49] Kirk, W.A., Geodesic geometry and fixed point theory II. In: G-Falset, J., L-Fuster, E., Sims, B. (eds.), Proc. International Conference on Fixed Point Theory, Valencia 2003, pp. 113-142, Yokohama Press 2004.

[50] Kirk, W.A., Sims, B. (eds.), Handbook of Metric Fixed Point Theory. Kluwer Academic Publishers, Dordrecht, Boston, London, xi+703pp. (2001).

[51] Kleene, S.C., Introduction to Metamathematics. North-Holland (Amsterdam), Noordhoff (Groningen), Van Nostrand (New-York) 1952.

[52] Kohlenbach, U., Pointwise hereditary majorization and some applications. Arch. Math. Logic **31**, pp. 227-241 (1992).

[53] Kohlenbach, U., Effective bounds from ineffective proofs in analysis: an application of functional interpretation and majorization. J. Symbolic Logic **57**, pp. 1239-1273 (1992).

[54] Kohlenbach, U., Effective moduli from ineffective uniqueness proofs. An unwinding of de La Vallée Poussin's proof for Chebycheff approximation. Ann. Pure Appl. Logic **64**, pp. 27–94 (1993).

[55] Kohlenbach, U., 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. Numer. Funct. Anal. and Optimiz. **14**, pp. 581–606 (1993).

[56] Kohlenbach, U., Analysing proofs in analysis. In: W. Hodges, M. Hyland, C. Steinhorn, J. Truss, editors, *Logic: from Foundations to Applications. European Logic Colloquium* (Keele, 1993), pp. 225–260, Oxford University Press (1996).

[57] Kohlenbach, U., Mathematically strong subsystems of analysis with low rate of growth of provably recursive functionals. Arch. Math. Logic **36**, pp. 31–71 (1996).

[58] Kohlenbach, U., Arithmetizing proofs in analysis. In: Larrazabal, J.M., Lascar, D., Mints, G. (eds.), Logic Colloquium '96, Springer Lecture Notes in Logic **12**, pp. 115-158 (1998).

[59] Kohlenbach, U., On the no-counterexample interpretation. J. Symbolic Logic **64**, pp. 1491-1511 (1999).

[60] Kohlenbach, U., On the computational content of the Krasnoselski and Ishikawa fixed point theorems. In: Proceedings of the Fourth Workshop on Computability and Complexity in Analysis, J. Blanck, V. Brattka, P. Hertling (eds.), Springer LNCS **2064**, pp. 119-145 (2001).

[61] Kohlenbach, U., A quantitative version of a theorem due to Borwein-Reich-Shafrir. Numer. Funct. Anal. and Optimiz. **22**, pp. 641-656 (2001).

[62] Kohlenbach, U., Uniform asymptotic regularity for Mann iterates. J. Math. Anal. Appl. **279**, pp. 531-544 (2003).

[63] Kohlenbach, U., Some logical metatheorems with applications in functional analysis. Trans. Amer. Math. Soc. vol. 357, no. 1, pp. 89-128 (2005).

[64] Kohlenbach, U., Some computational aspects of metric fixed point theory. Nonlinear Analysis **61**, pp. 823-837 (2005).

[65] Kohlenbach, U., A logical uniform boundedness principle for abstract metric and hyperbolic spaces. Electronic Notes in Theoretical Computer Science **165** (Proc. WoLLIC 2006), pp. 81-93 (2006).

[66] Kohlenbach, U., Applied Proof Theory: Proof Interpretations and their Use in Mathematics. Book in prepration for 'Springer Monographs in Mathematics'. Approx. 500pp. Expected to appear: Spring 2008.

[67] Kohlenbach, U., Effective uniform bounds from proofs in abstract functional analysis. To appear in: Cooper, B., Loewe, B., Sorbi, A. (eds.), 'CiE 2005 New Computational Paradigms: Changing Conceptions of What is Computable'. Springer Publisher.

[68] Kohlenbach, U., Gödel's functional interpretation and its use in current mathematics. To appear in: Horizons of Truth, Gödel Centenary. Cambridge University Press.

[69] Kohlenbach, U., Lambov, B., Bounds on iterations of asymptotically quasi-nonexpansive mappings. In: Falset, J.G., Fuster, E.L., Sims, B. (eds.), Proc. International Conference on Fixed Point Theory and Applications, Valencia 2003, pp. 143-172, Yokohama Publishers (2004)

[70] Kohlenbach, U., Leuştean, L., Mann iterates of directionally nonexpansive mappings in hyperbolic spaces. Abstr. Appl. Anal. vol. 2003, no.8, pp. 449-477 (2003).

[71] Kohlenbach, U., Leuştean, L., The approximate fixed point property in product spaces. Nonlinear Analysis **66** , pp. 806-818 (2007).

[72] Kohlenbach, U., Leuştean, L., Asymptotically nonexpansive mappings in uniformly convex hyperbolic spaces. arXiv:0707.1626 [math.LO] (2007).

[73] Kohlenbach, U., Oliva, P., Proof mining in L_1-approximation. Ann. Pure Appl. Logic **121**, pp. 1-38 (2003).

[74] Kohlenbach, U., Oliva, P., Proof mining: a systematic way of analysing proofs in mathematics. Proc. Steklov Inst. Math. **242**, pp. 1-29 (2003).

[75] Kreisel, G., On the interpretation of non-finitist proofs, part I. J. Symbolic Logic **16**, pp.241-267 (1951).

[76] Kreisel, G., On the interpretation of non-finitist proofs, part II: Interpretation of number theory, applications. J. Symbolic Logic **17**, pp. 43-58 (1952).

[77] Kreisel, G., Finiteness theorems in arithmetic: an application of Herbrand's theorem for Σ_2-formulas. Proc. of the Herbrand symposium (Marseille, 1981), North-Holland (Amsterdam), pp. 39-55 (1982).

[78] Kreisel, G., Macintyre, A., Constructive logic versus algebraization I. Proc. L.E.J. Brouwer Centenary Symposium (Noordwijkerhout 1981), North-Holland (Amsterdam), pp. 217-260 (1982).

[79] Krivine, J.-L., Dependent choice, 'quote' and the clock. Theoretical Computer Science **308**, pp. 259-276 (2003).

[80] Kroó, A., On the continuity of best approximations in the space of integrable functions. Acta Mathematica Academiae Scientiarum Hungaricae **32**, pp. 331-348 (1978).

[81] Lambov, B., Rates of convergence of recursively defined sequences. In: Brattka, V., Staiger, L., Weihrauch, E., Proc. of the 6th Workshop on Computability and Complexity in Analysis, vol. 120 of Electronic Notes in Theoretical Computer Science, pp. 125-133 (2005).

[82] Leuştean, L., Proof mining in $\mathbb{R}$-trees and hyperbolic spaces. Electronic Notes in Theoretical Computer Science **165** (Proc. WoLLIC 2006), pp. 95-106 (2006).

[83] Leuştean, L., A quadratic rate of asymptotic regularity for CAT(0)-spaces.J. Math. Anal. Appl. **325**, pp. 386-399 (2007).

[84] Luckhardt, H., Herbrand-Analysen zweier Beweise des Satzes von Roth: Polynomiale Anzahlschranken. J. Symbolic Logic **54**, pp. 234-263 (1989).

[85] Luckhardt, H., Bounds extracted by Kreisel from ineffective proofs. In: Odifreddi, P., Kreiseliana, 289–300, A K Peters, Wellesley, MA, 1996.

[86] Macintyre, A., The mathematical significance of proof theory. Phil. Trans. R. Soc. A **363**, pp. 2419-2435 (2005).

[87] Mints, G.E., Unwinding a non-effective cut elimination proof. In: Grigoriev, D., Harrison, J., Hirsch, E.A. (Eds.): Computer Science - Theory and Applications, First International Computer Science Symposium in Russia, CSR 2006, St. Petersburg, Russia, June 8-12, 2006, Proceedings. Springer LNCS **3967**, pp. 259-269 (2006).

[88] Odifreddi, P. (ed.), Kreiseliana. About and around Greorg Kreisel. A K Peters, Wellesley, Massachusetts, xiii+495 pp. (1996).

[89] Oliva, P., On the computational complexity of best L_1-Approximation. Math. Logic. Quart. **48, suppl. I**, pp. 66-77 (2002).

[90] Oliva, P., Understanding and using Spector's bar recursive interpretation of classical analysis. In: Proceedings of CiE 2006, Springer LNCS **3988**, pp. 423-434 (2006).

[91] Raffalli, C., Getting results from programs extracted from classical proofs. Theoretical Computer Science **323**, pp. 49-70 (2004).

[92] Reich, S., Shafrir, I., Nonexpansive iterations in hyperbolic spaces. Nonlinear Analysis, Theory, Methods and Applications **15**, pp. 537-558 (1990).

[93] Roth, K.F., Rational approximations to algebraic numbers. Mathematika **2**, pp. 1-20 (1955).

[94] Schwichtenberg, H., Dialectica interpretation of well-founded induction. To appear in: Math. Log. Quart.

[95] Shoenfield, J.S., Mathematical Logic. Addison-Wesley Publishing Company (Reading, Massachusetts) 1967.

[96] Spector, C., Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics. In: Recursive function theory, Proceedings of Symposia in Pure Mathematics, vol. 5 (J.C.E. Dekker (ed.)), AMS, Providence, R.I., pp. 1–27 (1962).

[97] Streicher, T., Kohlenbach, U., Shoenfield is Gödel after Krivine. Math. Log. Quart. **53**, pp. 176-179 (2007).

[98] Tao, T., Soft analysis, hard analysis, and the finite convergence principle. Essay posted May 23, 2007. Available at: http://terrytao.wordpress.com/2007/05/23/soft-analysis-hard-analysis-and-the-finite-convergence-principle/.

[99] Tao, T., Norm convergence of multiple ergodic averages for commuting transformations. arXiv:0707.1117v1 [math.DS] (2007).

[100] Weiermann, A., A classification of rapidly growing Ramsey functions. Proc. Amer. Math. Soc. 132, no. 2, pp. 553–561 (2004).

[101] Weiermann, A., Phasenübergänge in Logik und Kombinatorik. DMV-Mitteilungen **13**,no.3, pp. 152-156 (2005).