PROOF MINING: A SYSTEMATIC WAY OF ANALYSING PROOFS IN MATHEMATICS

ULRICH KOHLENBACH AND PAULO OLIVA

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.

§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

$$\forall x \in X \forall y \in K \; (f(x,y) > 0 \to g(x,y) > 0), \tag{1}$$

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 (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 (1) would be considered incomplete. An implication between complete theorems can also be viewed as incomplete. Consider a theorem of the form

$$\forall x \in X \forall y \in K \; (f(x,y) = 0 \to g(x,y) = 0). \tag{2}$$

Theorem (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 (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 (2).

Basic Research in Computer Science, funded by the Danish National Research Foundation.

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 [68, 69], combinatorics [8, 27] and algebra [22], 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 [47] 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 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

[1]For discussions on the original program of Kreisel see [26, 69].

two steps of proof mining (for the study of weak fragments in the context of proof mining see e.g. [48, 50]).

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 (1) and (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 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* [47] (to be presented in Section 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 [12]) 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 4, we show that statements of the form (1) and (2) are 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 5.1 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.

1.1. Formal systems. Our base formal system consists of extensional classical arithmetic in all finite types $\mathbf{E\text{-}PA}^\omega$. In places where classical logic must/can be avoided we use intuitionistic arithmetic $\mathbf{E\text{-}HA}^\omega$ (for details see [85] where $\mathbf{E\text{-}PA}^\omega$ is denoted by $\mathbf{E\text{-}HA}^\omega_c$). The finite types are inductively defined as: 0 is

[2]Such a tool has been developed (cf. e.g. [9]) for a different proof interpretation based on modified realizability and A-translation.

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 $0 \to 0$ as 1.

We denote by $\mathcal{T}^\omega$ both $\mathbf{E}\text{-}\mathbf{PA}^\omega$ as well as various subsystems of $\mathbf{E}\text{-}\mathbf{PA}^\omega$ such as $\mathbf{PRA}^\omega$ (cf. [3]) and $\mathbf{E}\text{-}\mathbf{G}_n\mathbf{A}^\omega$ (cf. [48]). $\mathcal{T}_i^\omega$ is the intuitionistic counterpart of $\mathcal{T}^\omega$. We work in systems containing equality ($=$) between objects of type 0 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]

$$\mathbf{QF}\text{-}\mathbf{AC}^{1,0} \quad : \quad \forall f^1 \exists n^0\, A_0(f,n) \to \exists \Phi \forall f\, A_0(f, \Phi f).$$

§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. [7]). For particular spaces, often more convenient (though essentially 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

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

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 [7] and [45] 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 [45], 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 [7, 45] and – for very weak systems – in [50].

According to the representation, mathematical statements of the form (1) and (2) have logical counterparts

$$\forall x^1 \forall y^1 \leq s(\exists n A_0(x, y, n) \to \exists m B_0(x, y, m)), \tag{3}$$

$$\forall x^1 \forall y^1 \leq s(\forall n A_0(x, y, n) \to \forall m B_0(x, y, m)), \tag{4}$$

respectively. Note, moreover, that (3) and (4) are special cases of [4]

$$\forall x^1 \forall y^1 \leq \tilde{s}x \exists z^0 \, B_0(x, y, z), \tag{5}$$

which in mathematical terms corresponds to statements of the form

$$\forall x \in X \forall y \in K_x \exists z \in \mathbb{N} \, B_1(x, y, z), \tag{6}$$

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 $K_x \subset Y$, (6) can be viewed (via the representation) as a special case of (5). On the other hand, (5) can also be considered as a special case of (6), taking X as the Baire space and $K_x = \{y : y \leq \tilde{s}x\}$.

§3. Monotone functional interpretation. The functional ('Dialectica') interpretation introduced by Gödel [28] translates an arbitrary formula A in the language of $\mathbf{E\text{-}HA}^\omega$ 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 $\mathbf{HA}^\omega$ then from that

[4]Using that '$\forall x^1, n^0$' 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.

[5]Note that the fact that B_1 is purely existential just adds some more existential quantifiers to '$\exists z^0$'.

[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.

proof one can extract a closed term t such that $A_D(t,y)$ is provable in $\mathbf{HA}^\omega$.[7] The soundness theorem has been adapted to many other systems both stronger ones as well as fragments of $\mathbf{HA}^\omega$. Via negative translation (and elimination of extensionality) it also applies to $\mathbf{E\text{-}PA}^\omega$ and related systems (cf. [3, 67, 85]).

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 0.

Definition 3.1 (Functional Interpretation). *The interpretation associates to each formula $A \in \mathcal{L}(\mathbf{HA}^\omega)$ (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:*

$A^D :\equiv A$, *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

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

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 $t_{A_0}(a) =_0 0$ 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 [86] only delivers a result in the first case above (and if negative translation had been applied first, not even then). In Section 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 [3].

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

[7]Here $\mathbf{HA}^\omega$ is a version of $\mathbf{E\text{-}HA}^\omega$ where the extensionality axioms in higher types are restricted to a quantifier-free rule of extensionality ([85]). Such a restriction – which is necessary for the soundness theorem to hold (see [33]) – 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 [67] applies.

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. [47]) *monotone functional interpretation*, or *m.f.i.* for short. In [47] 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 [10] strong majorizability relation $\geq^*_\rho$, which is a variant of Howard's [33] original hereditarily majorability relation. For numbers $n \geq^*_0 m$ just means that n is greater or equal than m. For functions f and g, $f \geq^*_1 g$ holds when f is monotone and is pointwise bigger than g. For higher types the relation is designed to be hereditarily monotone, i.e.

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

Three important properties of the relation $\geq^*_\rho$ are:

i) $x \geq^*_\rho y$ implies $x \geq^*_\rho x$,

ii) $x \geq^*_\rho y \wedge y \geq_\rho z \to x \geq_\rho z$, ($\geq_\rho$ as defined in Section 1.1)

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

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

always majorizes x.

Note that $\geq^*_\rho$ is not reflexive unless $\rho = 0$.

Using the relation $\geq^*$, 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

$$\exists x^* \exists x \leq^*_\rho x^* \forall y A_D(x,y).$$

THEOREM 3.2 ([44]). *Let Δ be a set of closed axioms of the form*

$\forall u^1 \exists v^1 \leq tu \forall w^0 A_0(u,v,w)$, *where t is closed.*

Suppose that [8]

$\mathcal{T}^\omega + \textbf{QF-AC}^{1,0} + \Delta \vdash \forall x^1 \forall y^1 \leq sx \exists z^0 B_0(x,y,z).$

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

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

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

$\forall u^1, w^0 \exists v^1 \leq tu \forall i \leq w A_0(u,v,i).$

As shown in [44], the set of sentences Δ also includes the non-computational principle weak König lemma (**WKL**). Since $\textbf{HA}^\omega \vdash \textbf{WKL}_\varepsilon$, this provides a **WKL**-elimination.

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

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

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

which can – using the representation of $C[0,1]$ – be written in form Δ (see [45]). Note that INF_ε is equivalent to

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

which, given our representation of $f \in C[0,1]$, can be easily proved in $\mathbf{HA}^\omega$. One example of a corollary of Theorem 3.2 would be the following.

Theorem 3.3 ([45]). *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*

$$\mathcal{T}^\omega + \mathbf{QF\text{-}AC}^{1,0} + \mathsf{INF} \vdash \forall x \in X \forall y \in K_x \exists z \in \mathbb{N} B_1(x,y,z)$$

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

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

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 3.4. 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 [48]. 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 [53, 57]) 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 3.3 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 [49]). We shall discuss this in more detail in Sections 6 and 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 3.5. *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* $1 \leq p < \infty$.

3.1. Monotone functional interpretation of theorems having the form (5). In Bishop [12] 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) 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) appear in the special forms (3) and (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 (4) gives [10]

$$\exists \Phi^* \exists \Phi \leq^* \Phi^* \forall x \forall y \leq sx \forall m \, (A_0(x, y, \Phi xym) \to B_0(x, y, m)),$$

which is equivalent (by elementary constructive reasoning)[11] to

$$\exists \Psi^* \leq^* \Psi^* \forall x \forall y \leq sx \forall m \exists n \leq \Psi^* x^+(s^* x^+) m \, (A_0(x, y, n) \to B_0(x, y, m)).$$

The formula above is in turn equivalent to

$$\exists \Psi \leq^* \Psi \forall x \forall y \leq sx \forall m \, (\forall n \leq \Psi xm \, A_0(x, y, n) \to B_0(x, y, m)).$$

[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 $\forall y \leq sxA(y)$ is equivalent to $\forall y A(\min_1(y, sx))$, where $\min_1(x, y) := \lambda n. \min(x(n), y(n))$.

[11] In the direction '$\to$' we can take $\Psi^* := \Phi^*$. In the other direction, suppose that Ψ^* satisfies the second formula. Then

$\Phi^* xym := \Psi^* x^+(s^* x^+) m$ and
$\Phi xym := \min i \leq \Phi^* xym \, [A_0(x, y, i) \to B_0(x, y, m)]$

satisfy the first formula.

In the same way, the monotone functional interpretation of (3) is equivalent to

$$\exists \Psi \leq^* \Psi \forall x \forall y \leq sx \forall n\, (A_0(x,y,n) \to \exists m \leq \Psi xn\, B_0(x,y,m)).$$

In Section 4, we shall consider various mathematical concepts which have the logical form (1) and (2) (the mathematical counterparts of (3) and (4)) and therefore the form (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 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 $\tilde{\Psi} := \Psi(t^+)$ (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 6.2, 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, where we explain how monotone functional interpretation naturally transforms those moduli into one another via the treatment of implications.

§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 (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.

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 (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 [45]. 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 [73]. For L_1-approximation strong unicity was studied e.g. by Björnestål [13, 14] and Kroó [63, 65]. See [5] 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 $f(x, y) := d_X(x, y) - dist(x, K)$, where $dist(x, K) := \inf_{y \in K} d_X(x, y)$. 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 [45].

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 [45], a modulus of uniqueness Φ automatically gives a modulus of pointwise continuity for the projection $\mathcal{P}$, also called *rate of smoothness/continuity*,

$$\forall x, y \in X(d_X(x, y) < \frac{1}{2}\Phi(x, \varepsilon) \to d_X(\mathcal{P}(x), \mathcal{P}(y)) < \varepsilon).$$

Again, the relationship between strong uniqueness and the smoothness of the projection operator has been studied extensively in the literature (cf. [1, 2, 6, 14]).

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

$$\mathsf{CVX} :\equiv \forall x, y \in B(\|\frac{1}{2}(x + y)\| \stackrel{\mathbb{R}}{=} 1 \to \|x - y\| \stackrel{\mathbb{R}}{=} 0),$$

which is again of the form (2), *monotone functional interpretation* creates a modulus $\eta : \mathbb{Q}^*_+ \to \mathbb{Q}^*_+$ satisfying

$$\forall x, y \in B; \varepsilon \in \mathbb{Q}^*_+ (\|\frac{1}{2}(x+y)\| > 1 - \eta(\varepsilon) \to \|x - y\| < \varepsilon).$$

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 [20] (see also [61]) 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. [17, 30, 31]). 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 4.6 and 5.1 (cf. [18, 40, 55, 58]).

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 4.1 and 4.4. Among the many publications on the connection between moduli of uniform convexity and rates of strong unicity see e.g. [14, 35, 66, 76].

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 (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 $\alpha(\varepsilon) := 1 - 2^{-\tilde{\eta}(\varepsilon)}$ has in fact been considered in the literature by Rakotch [77] and – in the context of Bishop style constructive analysis – in [16]. 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 4.4 below.

As we will show in Section 5.1, it is exactly such a modulus which is needed to obtain a rate of convergence in Edelstein's fixed point theorem [23, 77]. 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 [80]. This line of work is further continued in [21, 72, 81], to list only a few references.

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 (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 [11]) and in computable analysis (see [42], [75] and [87]) 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 [75]). On the other hand, numerical analysts define the function

$$\Omega(\varepsilon) := \sup_{d_K(x,y) \le \varepsilon} d_X(f(x), f(y))$$

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. [65]) the modulus Ω is often used just for building a

$$\Omega^{-1}(\varepsilon) := \inf\{\delta \in [0, 1] : \Omega(\delta) = \varepsilon\},$$

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.

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 (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}.

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 (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. $x_n := f^n(x)$. 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 (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 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 [51]).

§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) 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 6 and 7, we survey results we obtained in more substantial examples which solved open problems in the literature.

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 6.1 and 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 ([23]). *Let (K,d) be a compact metric space and $f : K \to K$ be contractive (in the sense of 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 5.2. *Let D_K denote an upper bound for the diameter of the compact space K. Moreover, define $\chi_1(\eta, \varepsilon) := \frac{D_K - \varepsilon}{\eta(\varepsilon)} + 1$. For any function $f : K \to K$ having moduli of contractivity η the function $\kappa(\varepsilon) := \chi_1(\eta, \varepsilon)$ 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 $d(x_n, x_{n+1}) < d(x_{n-1}, x_n) \leq \varepsilon$. $\dashv$

Remark 5.3. 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 5.4. *Define* $\chi_2(\eta, \varepsilon) := \frac{\eta(\varepsilon)}{2}$. *For any function* $f : K \to K$ *having moduli of contractivity* η *the function* $\Phi(\varepsilon) := \chi_2(\eta, \varepsilon)$ *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). Similarly as explained in Section 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 5.5. *Define* $\chi_3(\kappa, \Phi, \varepsilon) := \kappa(\Phi(\varepsilon))$. *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* $\delta(\varepsilon) := \chi_3(\kappa, \Phi, \varepsilon)$ *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 5.6. *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]

$$\delta(\varepsilon) := \chi_3(\lambda\varepsilon.\chi_1(\eta, \varepsilon), \lambda\varepsilon.\chi_2(\eta, \varepsilon), \varepsilon) = \frac{D_K - \eta(\varepsilon)}{\eta(\frac{\eta(\varepsilon)}{2})} + 1,$$

i.e.

$$\forall x \in K \forall\varepsilon \in \mathbb{Q}^*_+ \forall n \geq \delta(\varepsilon)\,(d(x_n, c) \leq \varepsilon).$$

[13]Note that δ depends only on ε, D_K and η, but not x or f.

Another quantitative version is given in Rakotch [77]. For a discussion of Edelstein's fixed point theorem in the context of Bishop's constructive mathematics see [16]. A recent domain theoretic approach to Edelstein's theorem can be found in [71].

§6. Proofs based on Heine-Borel compactness. We have presented how the computational content of sentences of the form (3), (4) and (5) (in mathematical terms (1), (2) and (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), 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 [82]) 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

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

which, as shown in [45], has the logical form Δ. If the principle INF has been used in the proof of a theorem of the form (5) at some point in the proof a modus ponens over an implication

$$\forall u^1 \exists v^1 \leq tu \forall w^0 A_0(u, v, w) \to \forall x \forall y \leq sx \exists z B_0(x, y, z) \tag{7}$$

will take place. Negative translation of (7) gives

$$\forall u^1 \neg\neg \exists v^1 \leq tu \forall w^0 A_0(u, v, w) \to \forall x \forall y \leq sx \neg\neg \exists z B_0(x, y, z) \tag{8}$$

and hence a-fortiori

$$\forall u^1 \exists v^1 \leq tu \forall w^0 A_0(u, v, w) \to \forall x \forall y \leq sx \neg\neg \exists z B_0(x, y, z). \tag{9}$$

The m.f.i. of the premise of (9) asks for a Φ^* satisfying

$$\exists \Phi \leq t(\Phi \leq^* \Phi^* \wedge \forall u^1 \forall w^0 A_0(u, \Phi u, w)),$$

which can be clearly taken to be $\Phi^* := t^*$, for some t^* majorizing t. The (partial) monotone functional interpretation of the implication (9) is realized by a functional χ^* satisfying

$$\exists \chi \leq^* \chi^* \forall \Phi \leq t(\forall u^1 \forall w^0 A_0(u, \Phi u, w) \to \forall x \forall y \leq sx B_0(x, y, \chi(\Phi, x, y))).$$

Note that $\chi^*(t^*, x^+, s^*(x^+))$ majorizes $\chi(\Phi, x, y)$. Therefore, given the truth of the premise of (7) (and therefore its Skolemized version '$\exists \Phi \leq t \forall u, w A_0(u, \Phi u, w)$'), the functional $\Psi(x) := \chi^*(t^*, x^+, s^*(x^+))$ 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 [44], 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.

6.1. Example 2: Jackson's theorem. In [60] the authors have carried out the analysis of Cheney's proof [19] of the following well-known theorem in L_1-approximation theory ('approximation in the mean').

Theorem 6.1 (Jackson's theorem [37]). *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 (2). Recall that the L_1-norm of a function $f \in C[0,1]$ is defined as

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

and $p \in P_n$ is a best L_1-approximation of f from P_n if

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

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 $\{p \in P_n : \|p\|_1 \leq 2\,\|f\|_1\}$. 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 $G(f,p) := \|f - p\|_1$ attains its infimum in $\tilde{K}_{f,n}$. The highly non-trivial part of Theorem 6.1 is the uniqueness of the best L_1-approximation.

Define $F(f,p) := \|f - p\|_1 - dist_1(f, P_n)$. Uniqueness can be expressed as

$$\forall n; f \in C[0,1]; p_1, p_2 \in P_n \,(\bigwedge_{i=1}^{2} F(f, p_i) \stackrel{\mathbb{R}}{=} 0 \to p_1 = p_2).$$

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 6.1 can be stated as

$$\forall n; f \in C[0,1]; p_1, p_2 \in K_{f,n} \,(\bigwedge_{i=1}^{2} F(f, p_i) \stackrel{\mathbb{R}}{=} 0 \to \|p_1 - p_2\|_1 \stackrel{\mathbb{R}}{=} 0),$$

where for technical reasons we use the larger space

$$K_{f,n} :\equiv \{p \in P_n : \|p\|_1 \leq \frac{5}{2}\|f\|_1\}.$$

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 (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 $\mathbf{PA}^\omega$-definable. Therefore, Jackson's theorem falls into the general form described in Section 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 $n \in \mathbb{N}$,

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

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 3.2 is indeed an indication of which inputs are the right ones for the modulus of uniqueness. See, for instance, [13] and [64] where the modulus of uniform continuity is always used as an input for moduli of uniqueness.

Theorem 3.2 guarantees that from any proof of Jackson's theorem formalizable in a system like $\mathbf{E\text{-}PA}^\omega + \mathbf{QF\text{-}AC}^{1,0} + \mathsf{INF}$ we are able to extract a modulus of uniqueness Φ. One such proof, as shown in [43], was presented by Cheney [19] in 1965. Therefore, by Theorem 3.2 we obtain the following *a priori* information.

PROPOSITION 6.2. *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 $\mathbf{E\text{-}PA}^\omega$, 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 [60] 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 6.3 ([60]). *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

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

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 $f \in C[0,1]$

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

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 [13] 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 [63] 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 6.3, 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.

Theorem 6.4 ([60]). *Let $\mathcal{P}(f,n)$ denote the operator which produces the best L_1-approximation of f from P_n. Then, for all n*

$$\forall f, g \in C[0,1]; \varepsilon \in \mathbb{Q}_+^* \, (\|f - g\|_1 \leq \frac{\Phi(\omega_f, \varepsilon)}{2} \to \|\mathcal{P}(f,n) - \mathcal{P}(g,n)\|_1 \leq \varepsilon),$$

where ω_f denotes a modulus of uniform continuity of f.

The modulus of uniqueness Φ has also been used in [74] 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 6.5 ([74]). *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 [45] and [46].

§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**) $:\equiv$ 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**) $:\equiv$ 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**) $:\equiv$ 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**) $:\equiv$ 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 $\mathsf{PCM}(a_n)$.

Although the principles mentioned above are equivalent to full arithmetical comprehension even over weak base systems (see [54] and [82])[14], 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 [49] and [54] 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 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 $\mathbf{QF\text{-}AC}^{0,0}$) to

$$Mon(a_n) \to \exists f \forall k \forall m \big(m \ge f(k) \to |a_{f(k)} - a_m| \le \frac{1}{k+1}\big).$$

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) 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

[14]Nevertheless, those can also be treated by monotone functional interpretation using a weak form of monotone bar recursion (cf. Section 8).

$\mathsf{PCM}(a_n)$ can be reduced to the use of the arithmetical version[15]

$$\mathsf{PCM_{ar}}(a_n) :\equiv \Big[Mon(a_n) \to \forall k \exists n \forall m \big(m \geq n \to |a_n - a_m| \leq \frac{1}{k+1}\big)\Big]$$

which in turn is equivalent to $\forall k \exists n \forall m (m \geq n \to |\tilde{a}_n - \tilde{a}_m| \leq \frac{1}{k+1})$, where

$$\tilde{a}_n := \max(0, \min_{i \leq n}(a_i, N)).$$

Hence, $\mathsf{PCM_{ar}}(a_n)$ 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 $\mathsf{PCM_{ar}}$ as premise and a statement of the form (5) as conclusion:

$$\forall x^1 \forall y \leq_1 sx \big(\mathsf{PCM_{ar}}(t(x,y)) \to \exists z^0 B_0(x,y,z)\big), \tag{10}$$

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

$$\forall x^1 \forall y \leq_1 sx \big(\forall k \neg\neg \exists n \forall m A_0(k,n,m) \to \neg\neg \exists z^0 B_0(x,y,z)\big) \tag{11}$$

of (10) is realized by a functional Ω^* satisfying

$$\begin{cases} \exists \Omega \leq^* \Omega^* \forall \Psi, x \forall y \leq sx \\ \qquad \big(\forall k, g A_0(k, \Psi(k,g), g(\Psi(k,g))) \to B_0(x,y,\Omega(x,y,\Psi))\big). \end{cases}$$

Suppose now that we have a functional Φ^* satisfying the monotone functional interpretation of the negative translation of $\forall x^1 \forall y \leq sx\, \mathsf{PCM_{ar}}(t(x,y))$, i.e.

$$\exists \Phi \leq^* \Phi^* \forall x^1; y \leq sx, k, g\, A_0(k, \Phi(x,y,k,g), g(\Phi(x,y,k,g)))\big) \tag{12}$$

then $\chi(x,\Phi^*) := \Omega^*(x^+, s^*x^+, \Phi^*(x^+, s^*x^+)) \geq \Omega(x,y,\Phi(x,y))$ for all x^1 and $y \leq sx$, where $s \leq^* s^*$ and $\Phi(x,y) := \lambda k, g.\Phi(x,y,k,g)$. 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 $\mathsf{PCM_{ar}}(t(x,y))$ to the bound for the conclusion of (10) is given by a functional Φ^* satisfying (12). One easily verifies that we can take

$$\Phi^*(x,y,k,g) := \max_{i \leq (k+1)N}(g^i(0)), \tag{13}$$

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.

[15]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 $\mathsf{PCM_{ar}}$. We do not go here into this as in the application to be discussed below this passage is trivial.

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. [41]). In this section we report on the results of a recent case study of proof mining carried out by the first author (see [55, 56, 58] and – together with Laurenţiu Leuştean – [59]).

Definition 7.1. *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 $x \in C$

$$x_0 := x, \quad x_{k+1} := (1 - \lambda_k)x_k + \lambda_k f(x_k),$$

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 [62] iterations see e.g [15, 25, 70, 79].

Let $r_C(f) := \inf_{x\in C} \|x - f(x)\|$. For the rest of this section we assume, following [15] and [36], that $(\lambda_k)_{k\in\mathbb{N}}$ is divergent in sum, which can be expressed (since $\lambda_k \geq 0$) as[16]

$$\forall n, i \in \mathbb{N} \exists k \in \mathbb{N} \left(\sum_{j=i}^{i+k} \lambda_j \geq n \right), \tag{14}$$

and that

$$\forall k \in \mathbb{N}(\lambda_k \leq 1 - \frac{1}{K}) \text{ for some } K \in \mathbb{N}. \tag{15}$$

Theorem 7.2 ([15]). *Suppose that $(\lambda_k)_{k\in\mathbb{N}}$ satisfies the conditions* (14) *and* (15). *Then the Krasnoselski-Mann iteration $(x_n)_{n\in\mathbb{N}}$ starting from any point $x \in C$ satisfies*

$$\|x_n - f(x_n)\| \stackrel{n\to\infty}{\to} r_C(f).$$

Under quite general circumstances one can prove that $r_C(f) = 0$.

Theorem 7.3 ([15], [36]). *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*

$$\forall x \in C(\|x_n - f(x_n)\| \stackrel{n\to\infty}{\to} 0) \quad \textit{(called ‘asymptotic regularity’).}$$

[16] This form will be particularly suitable below.

Remark 7.4. The special case of Theorem 7.3 in which only the asymptotic regularity of the sequence (x_n^*) is conclude is due to Ishikawa [36].[17] 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 [36] did. For uniformly convex spaces X, bounded C and constant $\lambda_k = \lambda$ the result was proved in [18], and for general λ_k – even more general than in Ishikawa's theorem – it follows from [32] for such spaces. If C is, moreover, compact and $\lambda = \frac{1}{2}$, asymptotic regularity was (for uniformly convex X) already proved in [62].

In oder to see that our general meta-theorem on proof mining can be applied to Theorem 7.2 we first have to find a proper formalization of the conclusion 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

$$\forall \varepsilon > 0 \exists n \in \mathbb{N} \forall m \geq n \forall x^* \in C(\|x_m - f(x_m)\| < \|x^* - f(x^*)\| + \varepsilon). \tag{16}$$

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 4.6 applies. Therefore, the quantifier '$\forall m \geq n$' in (16) is in fact superfluous. Nevertheless, due to the alternation $\exists n \in \mathbb{N} \forall x^* \in C$, (16) still does not of the form $\forall\exists$ required (as a consequence of the use of classical logic) by our meta-theorems 3.2 and 3.3.[18] The following variant of (16), however, does have this form[19]

$$\forall \varepsilon > 0 \forall x^* \in C \exists n \in \mathbb{N}(\|x_n - f(x_n)\| < \|x^* - f(x^*)\| + \varepsilon). \tag{17}$$

Under the assumption of the existence of $r_C(f)$, formulations (16) and (17) are actually equivalent. In the following we shall study in more detail the form (17) of Theorem 7.2. 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 7.2 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[20]. 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 (14) and (15) on λ_k introduce new inputs, namely a bound $\alpha : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that

$$\forall n, i \in \mathbb{N}(n \leq \sum_{j=i}^{i+\alpha(i,n)} \lambda_j) \tag{18}$$

[17] For constant $\lambda_k = \lambda$ the result was independently obtained in [24].

[18] Indeed, an effective bound on '$\exists n$' in (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.

[19] 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 [58].

[20] 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.

and a $K \in \mathbb{N}$ such that

$$\forall n \in \mathbb{N}(\lambda_n \leq 1 - \frac{1}{K}), \tag{19}$$

where both (18), (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 (18) and (19). From this the meta-theorem provides the a priori information that the bound on the convergence in Theorem 7.2 we are about to extract might depend on α, K (and x^*) as new inputs which were not visible in the original formulation of the theorem, but that it will be independent from any particular (λ_k) itself (cf. Section 3.1).

Let us now consider the lemmas used in the proof of Theorem 7.2. By far the largest part of the proof concerns a highly non-trivial inequality due to [29] (whose proof is based on [38] and also [36]): 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 7.2, to its arithmetical version $\mathsf{PCM}_{\mathsf{ar}}(\|x_n - f(x_n)\|)$ 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 $\textbf{E-G}_3\textbf{A}^\omega$ of $\textbf{E-PA}^\omega$ which, in particular must not contain the iteration functional $\Phi_{it}(x, y, f) := f^x(y)$. 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

$$\forall n \left(x_{n+1} = (1 - \lambda_n)x_n + \lambda_n f(x_n)\right)$$

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

$$\mathsf{PCM}_{\mathsf{ar}} \to (A \to (17)),$$

where (17) (and hence $A \to (17)$) is a $\forall\exists$-formula. Therefore, the discussion of the modus ponens problem above applies and we can extract a bound for (17) in f, x, x^*, α, K which, as a consequence of the use of $\mathsf{PCM}_{\mathsf{ar}}$, will involve a use of the iteration functional Φ_{it}. Indeed, in [56], the first author obtained the following quantitative version of Theorem 7.2 (as a matter of fact, we not even need to assume that $(X, \|\cdot\|)$ is complete or separable).

THEOREM 7.5 ([56]). *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*

$$\forall k \in \mathbb{N}\big(\lambda_k \leq 1 - \frac{1}{K}\big)$$

for some $K \in \mathbb{N}$. Let $\alpha : \mathbb{N} \to \mathbb{N}$ be such that

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

Let $(x_n)_{n\in\mathbb{N}}$ be the Krasnoselski-Mann iteration

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

starting from $x \in C$. Then the following holds

$$\forall x, x^* \in C \forall \varepsilon > 0 \forall n \geq h(\varepsilon, x, x^*, f, K, \alpha)(\|x_n - f(x_n)\| < \|x^* - f(x^*)\| + \varepsilon),$$

where[21]

$$\begin{aligned}
&h(\varepsilon, x, x^*, f, K, \alpha) := \widehat{\alpha}(\lceil 2\|x - f(x)\| \cdot \exp(K(M+1))\rceil \dot{-} 1, M),\\
&\textit{with } M := \left\lceil \tfrac{1+2\|x-x^*\|}{\varepsilon} \right\rceil \textit{ and}\\
&\widehat{\alpha}(0, M) := \tilde{\alpha}(0, M),\ \widehat{\alpha}(m+1, M) := \tilde{\alpha}(\widehat{\alpha}(m, M), M) \textit{ with}\\
&\tilde{\alpha}(m, M) := m + \alpha(m, M)\ (m \in \mathbb{N}).
\end{aligned}$$

Instead of M we may use any upper bound $\mathbb{N} \ni \tilde{M} \geq \frac{1+2\|x-x^*\|}{\varepsilon}$. Likewise, $\|x - f(x)\|$ may be replaced by any upper bound.

REMARK 7.6. An α satisfying the conditions of the theorem can be computed from any $\beta : \mathbb{N} \to \mathbb{N}$ such that $n \leq \sum\limits_{s=0}^{\beta(n)} \lambda_s$ (for all n) by $\alpha(i,n) := \max_{j\leq i}(\beta'(j,n))$, where $\beta'(i,n) := \beta(n+i) - i + 1$.

Perhaps the most useful aspect of Theorem 7.5 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 $d \geq diam(C)$, 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 [56]). 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 [24] for constant $\lambda_k := \lambda$. In [29] 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 [30] it is in fact conjectured that the uniformity in C might only hold in the much simpler case of uniformly convex case (cf. [40]). For constant λ, full uniformity was finally established in [4]. 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 7.5, allows to push the uniformity even further to the case where C is no longer assumed to be bounded but

[21] $n \dot{-} 1 = \max(0, n-1)$.

only to contain some point x^* whose iteration sequence (x_n^*) is bounded, i.e. the context of Theorem 7.3.

THEOREM 7.7 ([58]). *Under the assumptions of Theorem 7.5 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

$$h(\varepsilon, d, K, \alpha) := \widehat{\alpha}(\lceil 12 \cdot \exp(K(M+1)) \rceil - 1, M),$$
$$\text{with } M := \left\lceil \frac{7d}{\varepsilon} \right\rceil \text{ and } \widehat{\alpha} \text{ as in Theorem 7.5.}$$

Note that the bound only depends on $\frac{d}{\varepsilon}, K$ *and* α*!*

PROOF. The theorem follows from [58] (Thm. 2.5 plus Remarks 2.2 and 2.6). $\dashv$

Whereas this result easily follows from the logical analysis in [56] (which resulted in Theorem 7.5) of the proof of Theorem 7.2 and does not use any functional analytic tools at all, it seems that the embedding techniques, as used e.g. in [29] and most recently in a new form in [39], 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 [55], [56] and [58].

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 ([59]) 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 [78] (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 [29] (which were first introduced in [84] under the name of 'convex metric spaces') and directionally nonexpansive mappings in the sense of [39]. In particular, strengthened versions of the main results of [39] 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 $\mathbf{E\text{-}PA}^\omega$) 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 (U. Kohlenbach, Some logical metatheorems with applications in functional analysis, Preprint). 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).

§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) 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. [83]).

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 7)

$$\forall (a_n)_{n\in\mathbb{N}} \exists f \forall k, m(m \geq fk \to |\tilde{a}_{fk} - \tilde{a}_m| \leq \frac{1}{k+1}).$$

We have seen in Section 7 that the m.f.i. of the arithmetical version of PCM,

$$\forall (a_n)_{n\in\mathbb{N}} \forall k \exists n \forall m(m \geq n \to |\tilde{a}_n - \tilde{a}_m| \leq \frac{1}{k+1}) \tag{20}$$

can be easily realized using the iteration functional. Formula (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

$$\frac{\forall (a_n)_{n\in\mathbb{N}} \forall k \exists n \forall m A_0(k, n, m) \qquad \mathbf{AC}}{\forall (a_n)_{n\in\mathbb{N}} \exists f \forall k, m A_0(k, fk, m)}$$

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 $\forall (a_n)_{n\in\mathbb{N}}$)

$$\frac{\forall k \neg\neg \exists n \forall m A_0(k, n, m) \qquad \mathbf{AC}^N(\forall m A_0(k, n, m))}{\neg\neg \exists f \forall k, m A_0(k, fk, m)}$$

We finally apply functional interpretation to obtain

$$\frac{\exists \Phi^2 \forall k^0, g^1\, A_0(k, \Phi kg, g(\Phi kg)) \qquad (\mathbf{AC}^N(\forall m A_0(k, n, m)))^D}{\forall \Psi_1, \Psi_2 \exists f A_0(\Psi_1(f), f(\Psi_1(f)), \Psi_2(f))}$$

As done in Section 7 (cf. (13)), we can define via iteration a functional Φ^*

$$\Phi^*(k, g) := \max_{i \leq (k+1)N} (g^i(0)).$$

which majorizes a realizer of

$$\exists \Phi^2 \forall k^0, g^1 A_0(k, \Phi kg, g(\Phi kg)),$$

i.e. $\exists \Phi \leq^* \Phi^* \forall k^0, g^1 A(k, \Phi kg, g(\Phi kg))$. We now set out to obtain a realizer for the monotone functional interpretation of $\mathbf{AC}^N(\forall m A_0(k, n, m))$, 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

$$\forall \Phi, \Psi_1, \Psi_2 \exists f, k, g[A_0(k, \Phi k g, g(\Phi k g)) \to A_0(\Psi_1 f, f(\Psi_1 f), \Psi_2 f)]. \tag{21}$$

By $\mathsf{BR}_{0,1}$ we mean the bar recursive functional (defined by Spector [83]) satisfying the following equation:

$$\mathsf{BR}_{0,1}(Y,G,H,s) \stackrel{1}{=} \begin{cases} G(s) & \textsf{if } Y(s * \lambda n.0) \leq n \\ H(s, \lambda y^0 . \mathsf{BR}_{0,1}(Y,G,H,s*y)) & \textsf{otherwise.} \end{cases}$$

Let p be a shorthand for Φ, Ψ_1, Ψ_2. Spector showed that by taking

$$\begin{aligned} &Y := \lambda p\,.\,\Psi_1 \\ &G := \lambda p, s\,.\,s * \lambda n.0 \\ &H := \lambda p, s, \gamma\,.\,\gamma(\Phi(|s|, \lambda y.\Psi_2(\gamma(y)))) \end{aligned}$$

the functionals

$$\begin{aligned} &\mathcal{F} := \lambda p\,.\,\mathsf{BR}_{0,1}(Y(p), G(p), H(p), \langle\,\rangle) \\ &\mathcal{K} := \lambda p\,.\,\Psi_1(\mathcal{F}(p)) \\ &\mathcal{G} := \lambda p, y\,.\,\Psi_2(\mathsf{BR}_{0,1}(\overline{\mathcal{F}(p)}\,\mathcal{K}(p) * y)) \end{aligned}$$

realize f, k and g in (21). Let $\mathsf{BR}^*_{0,1}$ be the majorant of $\mathsf{BR}_{0,1}$ presented by Bezem [10]. Since we can easily find terms Y^*, G^* and H^* which majorize Y, G and H above, we get that

$$\begin{aligned} &\mathcal{F}^* := \lambda p\,.\,\mathsf{BR}^*_{0,1}(Y^*(p), G^*(p), H^*(p), \langle\,\rangle) \\ &\mathcal{K}^* := \lambda p\,.\,\Psi_1(\mathcal{F}^*(p)) \\ &\mathcal{G}^* := \lambda p, y\,.\,\Psi_2(\mathsf{BR}^*_{0,1}(\overline{\mathcal{F}^*(p)}\mathcal{K}^*(p) * \max(\overline{\mathcal{F}^*(p)}\mathcal{K}^*(p), y))), \end{aligned}$$

where $\max(s,x) := \max\{s_0, \ldots, s_{|s|-1}, x\}$, are terms satisfying the monotone functional interpretation of $\mathbf{AC}^N$. Note that $\lambda(a_n), \Psi_1, \Psi_2.\mathcal{F}((a_n), \Phi, \Psi_1, \Psi_2)$ realizes

$$\forall (a_n) \forall \Psi_1, \Psi_2 \exists f(\Psi_2(f) \geq f(\Psi_1(f)) \to |a_{f(\Psi_1 f)} - a_{\Psi_2 f}| \leq \frac{1}{\Psi_1 f + 1})),$$

and $\lambda(a_n), \Psi_1, \Psi_2.\mathcal{F}^*(N, \Phi^*, \Psi_1, \Psi_2)$ 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 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).

Remark 8.1. By the above we can treat proofs in the system $\mathcal{T}^\omega + \mathbf{QF\text{-}AC}^{1,0} + \mathsf{PCM}$. Note that for $\mathcal{T}^\omega = \mathbf{PRA}^\omega$ the above system can be viewed as a finite type extension of $\mathbf{ACA}_0$ known from reverse mathematics. In that case the bound extracted by m.f.i. from a proof of a theorem of form (6) will be a closed term of type 2 of $\mathbf{PRA}^\omega[\mathsf{BR}_{0,1}]$ which (by [34, 52]) denotes a functional in Gödel's primitive recursive functionals T of finite type (note that $\mathbf{PRA}^\omega$ only contains the fragment T_0 of T with primitive recursion on type 0).

REFERENCES

[1] J. R. Angelos, M. S. Henry, E. H. Kaufman, T. D. Lenker, and A. Kroó, *Local and global Lipschitz constants*, ***Journal of Approximation Theory***, vol. 46 (1986), pp. 137–156.

[2] 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***, vol. 46 (1986), pp. 129–136.

[3] J. Avigad and S. Feferman, *Gödel's functional ("Dialectica") interpretation*, ***Handbook of proof theory*** (S. R. Buss, editor), Studies in Logic and the Foundations of Mathematics, vol. 137, North Holland, Amsterdam, 1998, pp. 337–405.

[4] J. Baillon and R. E. Bruck, *The rate of asymptotic regularity is* $0(\frac{1}{\sqrt{n}})$, ***Theory and applications of nonlinear operators of accretive and monotone type***, Lecture Notes in Pure and Appl. Math., vol. 178, Marcel Dekker, Inc., New York, 1996, pp. 51–81.

[5] M. W. Bartelt and W. Li, *Error estimates and Lipschitz constants for best approximation in continuous function spaces*, ***Computers and Mathematics with Application***, vol. 30 (1995), no. 3-6, pp. 255–268.

[6] M. W. Bartelt and D. Schmidt, *Lipschitz conditions, strong uniqueness, and almost Chebyshev subspaces of* $C(x)$, ***J. Approx. Theory***, vol. 40 (1984), pp. 202–215.

[7] M. J. Beeson, ***Foundations of constructive mathematics***, Metamathematical Studies, Springer, Berlin, 1985.

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

[9] U. Berger, H. Schwichtenberg, and W. Buchholz, *Refined program extraction from classical proofs*, ***Annals of Pure and Applied Logic***, vol. 114 (2002), pp. 3–25.

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

[11] E. Bishop, ***Foundations of constructive analysis***, McGraw-Hill, New York, 1967.

[12] ———, *Mathematics as a numerical language*, ***Intuitionism and proof theory*** (A. Kino, J. Myhill, and R. E. Vesley, editors), North Holland, Amsterdam, 1970, pp. 53–71.

[13] 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.

[14] ———, *Local Lipschitz continuity of the metric projection operator*, ***Approximation theory. Banach Center Publications***, vol. 4 (1979), pp. 43–53.

[15] J. Borwein, S. Reich, and I. Shafrir, *Krasnoselski-Mann iterations in normed spaces*, ***Canadian Mathematical Bulletin***, vol. 35 (1992), pp. 21–28.

[16] 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***, vol. 165 (1992), no. 2, pp. 438–456.

[17] F. E. Browder, *Nonexpansive nonlinear operators in a Banach space*, ***Proceedings of the National Academy of Sciences of the U.S.A.***, vol. 54 (1965), pp. 1041–1044.

[18] F. E. Browder and W. V. Petryshyn, *The solution by iteration of nonlinear functional equations in Banach spaces*, ***Bulletin of the American Mathematical Society***, vol. 72 (1966), pp. 571–575.

[19] E. W. Cheney, *An elementary proof of Jackson's theorem on mean-approximations*, ***Mathematics Magazine***, vol. 38 (1965), pp. 189–191.

[20] J. A. Clarkson, *Uniformly convex spaces*, ***Transactions of the American Mathematical Society***, vol. 40 (1936), pp. 396–414.

[21] P. Collaço and J. Carvalho e Silva, *A complete comparison of 25 contraction conditions*, ***Nonlinear Analysis***, vol. 30 (1997), pp. 471–476.

[22] C. Delzell, *Kreisel's unwinding of Artin's proof*, ***Kreiseliana: About and around Georg Kreisel*** (P. Odifreddi, editor), A. K. Peters, Wellesley, MA, 1996, pp. 113–246.

[23] M. Edelstein, *On fixed and periodic points under contractive mappings*, ***The Journal of the London Mathematical Society***, vol. 37 (1962), pp. 74–79.

[24] M. Edelstein and R. C. O'Brien, *Nonexpansive mappings, asymptotic regularity and successive approximations*, ***The Journal of the London Mathematical Society***, vol. 17 (1978), pp. 547–554.

[25] H. W. Engl and A. Leitão, *A Mann interative regularization method for elliptic Cauchy problems*, ***Numerical Functional Analysis and Optimization***, vol. 22 (2001), pp. 861–884.

[26] S. Feferman, *Kreisel's 'Unwinding Program'*, ***Kreiseliana: about and around Georg Kreisel*** (P. Odifreddi, editor), A. K. Peters, Wellesley Massachusetts, 1996, pp. 247–273.

[27] J.-Y. Girard, ***Proof theory and logical complexity***, Studies in Proof Theory, Bibliopolis, Naples, 1987.

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

[29] K. Goebel and W. A. Kirk, *Iteration processes for nonexpansive mappings*, ***Topological methods in nonlinear functional analysis***, Contemporary Mathematics, vol. 21, AMS, 1983, pp. 115–123.

[30] ———, ***Topics in metric fixed point theory***, Cambridge studies in advanced mathematics, vol. 28, Cambridge University Press, Cambridge, 1990.

[31] K. Goebel and S. Reich, ***Uniform convexity, hyperbolic geometry, and nonexpansive mappings***, Monographs and Textbooks in Pure and Applied Mathematics, vol. 83, Marcel Dekker, Inc., New York, 1984.

[32] C. W. Groetsch, *A note on segmenting Mann iterates*, ***Journal of Mathematical Analysis and Applications***, vol. 40 (1972), pp. 369–372.

[33] W. A. Howard, *Hereditarily majorizable functionals of finite type*, ***Metamathematical investigation of intuitionistic Arithmetic and Analysis*** (A. S. Troelstra, editor), Lecture Notes in Mathematics, vol. 344, Springer, Berlin, 1973, pp. 454–461.

[34] ———, *Ordinal analysis of simple cases of bar recursion*, ***The Journal of Symbolic Logic***, vol. 46 (1981), pp. 17–30.

[35] R. Huotari and S. Sahab, *Strong unicity versus modulus of convexity*, ***Bulletin of the Australian Mathematical Society***, vol. 49 (1994), pp. 305–310.

[36] S. Ishikawa, *Fixed points and iterations of a nonexpansive mapping in a Banach space*, ***Proceedings of the American Mathematical Society***, vol. 59 (1976), pp. 65–71.

[37] D. Jackson, *Note on a class of polynomials of approximation*, ***Transactions of the American Mathematical Society***, vol. 22 (1921), pp. 320–326.

[38] W. A. Kirk, *Krasnoselski's iteration process in hyperbolic space*, ***Numererical Functional Analysis and Optimization***, vol. 4 (1982), pp. 371–381.

[39] ———, *Nonexpansive mappings and asymptotic regularity*, ***Nonlinear Analysis***, vol. 40 (2000), pp. 323–332.

[40] W. A. Kirk and C. Martinez-Yañez, *Approximate fixed points for nonexpansive mappings in uniformly convex spaces*, ***Annales Polonici Mathematici***, vol. 51 (1990), pp. 189–193.

[41] W. A. Kirk and B. Sims (editors), ***Handbook of metric fixed point theory***, Kluwer Academic Publishers, Dordrecht, 2001.

[42] K.-I. Ko, ***Complexity theory of real functions***, Birkhäuser, Boston-Basel-Berlin, 1991.

[43] 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)*, ***Ph.D. thesis***, Frankfurt, pp. xxii+278, 1990.

[44] ———, *Effective bounds from ineffective proofs in analysis: an application of functional interpretation and majorization*, ***The Journal of Symbolic Logic***, vol. 57 (1992), pp. 1239–1273.

[45] ———, *Effective moduli from ineffective uniqueness proofs. An unwinding of de La Vallée Poussin's proof for Chebycheff approximation*, ***Annals of Pure and Applied Logic***, vol. 64 (1993), pp. 27–94.

[46] ———, *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***, vol. 14 (1993), pp. 581–606.

[47] ———, *Analysing proofs in Analysis*, ***Logic: from foundations to applications*** (W. Hodges, M. Hyland, C. Steinhorn, and J. Truss, editors), European Logic Colloquium (Keele, 1993), Oxford University Press, 1996, pp. 225–260.

[48] ———, *Mathematically strong subsystems of analysis with low rate of growth of provably recursive functionals*, ***Archive for Mathematical Logic***, vol. 36 (1996), pp. 31–71.

[49] ———, *Arithmetizing proofs in analysis*, ***Lecture Notes in Logic***, vol. 12 (1998), pp. 115–158.

[50] ———, *Proof theory and computational analysis*, ***Electronic notes in theoretical computer science*** (A. Edalat, A. Jung, K. Keimel, and M. Kwiatkowska, editors), vol. 13, Elsevier, Amsterdam, 1998, 34 pp. (electronic).

[51] ———, *Relative constructivity*, ***The Journal of Symbolic Logic***, vol. 63 (1998), pp. 1218–1238.

[52] ———, *On the no-counterexample interpretation*, ***The Journal of Symbolic Logic***, vol. 64 (1999), pp. 1491–1511.

[53] ———, *The use of a logical principle of uniform boundedness in analysis*, ***Logic and foundations of mathematics*** (A. Cantini, E. Casari, and P. Minari, editors), Synthese Library, vol. 280, Kluwer Academic Publishers, Dordrecht, 1999, pp. 93–106.

[54] ———, *Things that can and things that can't be done in PRA*, ***Annals of Pure and Applied Logic***, vol. 102 (2000), pp. 223–245.

[55] ———, *On the computational content of the Krasnoselski and Ishikawa fixed point theorems*, ***Computability and complexity in analysis, (CCA'2000)*** (J. Blanck, V. Brattka, and P. Hertling, editors), Lecture Notes in Computer Science, vol. 2064, Springer, Berlin, 2001, pp. 119–145.

[56] ———, *A quantitative version of a theorem due to Borwein-Reich-Shafrir*, ***Numerical Functional Analysis and Optimization***, vol. 22 (2001), pp. 641–656.

[57] ———, *Foundational and mathematical uses of higher types*, ***Reflections on the Foundations of Mathematics: Essay in Honor of Solomon Feferman***, Lecture Notes in Logic, vol. 15, A. K. Peters, Ltd., 2002, pp. 92–116.

[58] ———, *Uniform asymptotic regularity for Mann iterates*, ***Journal of Mathematical Analysis and Applications***, vol. 279 (2003), pp. 531–544.

[59] U. Kohlenbach and L. Leuştean, *Mann iterates of directionally nonexpansive mappings in hyperbolic spaces*, To appear in: Abstract and Applied Analysis.

[60] U. Kohlenbach and P. Oliva, *Proof mining in L_1-approximation*, ***Annals of Pure and Applied Logic***, vol. 121 (2003), pp. 1–38.

[61] G. Köthe, ***Topoplogische lineare Räume***, Springer, Berlin, 1960.

[62] M. A. Krasnoselski, *Two remarks on the method of successive approximation (russian)*, ***Usp. Math. Nauk (N.S.)***, vol. 10 (1955), pp. 123–127.

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

[64] ———, *On the uniform modulus of continuity of the operator of best approximation in the space of periodic functions*, ***Acta Mathematica Academiae Scientiarum Hungaricae***, vol. 34 (1979), pp. 185–203.

[65] ———, *On strong unicity of L_1-approximation*, ***Proceedings of the American Mathematical Society***, vol. 83 (1981), no. 4, pp. 725–729.

[66] P.-K. Lin, *Strongly unique best approximation in uniformly convex Banach spaces.*, ***Journal of Approximation Theory***, vol. 56 (1989), pp. 101–107.

[67] H. Luckhardt, ***Extensional Gödel functional interpretation: A consistency proof of classical analysis***, Lecture Notes in Mathematics, vol. 306, Springer, Berlin, 1973.

[68] ———, *Herbrand-Analysen zweier Beweise des Satzes von Roth: Polynomiale Anzahlschranken*, ***The Journal of Symbolic Logic***, vol. 54 (1989), pp. 234–263.

[69] ———, *Bounds extracted by Kreisel from ineffective proofs*, ***Kreiseliana: About and around Georg Kreisel*** (P. Odifreddi, editor), A. K. Peters, Wellesley, MA, 1996, pp. 289–300.

[70] W. R. Mann, *Mean value methods in iteration*, ***Proceedings of the American Mathematical Society***, vol. 4 (1953), pp. 506–510.

[71] K. Martin, *Unique fixed points in domain theory*, ***Electronic notes in theoretical computer science*** (Stephen Brooks and Michael Mislove, editors), vol. 45, Elsevier, Amsterdam, 2001, 11 pp. (electronic).

[72] J. Meszáros, *A comparison of various definitions of contractive type mappings*, ***Bulletin of the Calcutta Mathematical Society***, vol. 84 (1992), pp. 167–194.

[73] D. J. Newman and H. S. Shapiro, *Some theorems on Chebycheff approximation*, ***Duke Mathematical Journal***, vol. 30 (1963), pp. 673–681.

[74] P. Oliva, *On the computational complexity of best L_1-approximation*, ***Mathematical Logic Quarterly***, vol. 48, Suppl. 1 (2002), pp. 66–77.

[75] M. Pour-El and J. Ian Richards, ***Computability in analysis and physics***, Perspectives in Mathematical Logic, Springer, Berlin, 1989.

[76] B. Prus and R. Smarzewski, *Strongly unique best approximations and centers in uniformly convex spaces*, ***Journal of Mathematical Analysis and Applications***, vol. 121 (1987), pp. 10–21.

[77] E. Rakotch, *A note on contractive mappings*, ***Proceedings of the American Mathematical Society***, vol. 13 (1962), pp. 459–465.

[78] S. Reich and I. Shafrir, *Nonexpansive iterations in hyperbolic spaces*, ***Nonlinear Analysis***, vol. 15 (1990), pp. 537–558.

[79] B. E. Rhoades, *Fixed point iterations using infinite matrices*, ***Transactions of the American Mathematical Society***, vol. 196 (1974), pp. 161–176.

[80] ———, *A comparison of various definitions of contractive mappings*, ***Transactions of the American Mathematical Society***, vol. 226 (1977), pp. 257–290.

[81] ———, *Contractive definitions*, ***Nonlinear analysis*** (Th M. Rassias, editor), World Sci. Publishing, Singapore, 1987, pp. 513–526.

[82] S. G. Simpson, ***Subsystems of second order arithmetic***, Perspectives in Mathematical Logic, Springer, Berlin, 1999.

[83] C. Spector, *Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles in current intuitionistic mathematics*, ***Recursive function theory: Proc. symposia in pure mathematics*** (F. D. E. Dekker, editor), vol. 5, American Mathematical Society, Providence, Rhode Island, 1962, pp. 1–27.

[84] W. Takahashi, *A convexity in metric space and nonexpansive mappings i*, ***Kodai Mathematical Seminar Reports***, vol. 22 (1970), pp. 142–149.

[85] A. S. Troelstra, ***Metamathematical investigation of intuitionistic Arithmetic and Analysis***, Lecture Notes in Mathematics, vol. 344, Springer, Berlin, 1973.

[86] ———, *Realizability*, ***Handbook of proof theory*** (S.R. Buss, editor), vol. 137, North Holland, Amsterdam, 1998, pp. 408–473.

[87] K. Weihrauch, ***Computable analysis***, Springer, Berlin, 2000.

BRICS
DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF AARHUS
NY MUNKEGADE
DK-8000 AARHUS C, DENMARK
E-mail: kohlenb@brics.dk

E-mail: pbo@brics.dk