Applications of Proof Interpretations

Philipp Gerhardy

PhD Dissertation

Department of Computer Science
University of Aarhus
Denmark

Applications of Proof Interpretations

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

by
Philipp Gerhardy
August 16, 2006

Abstract

In this thesis, the author describes the research carried out during his PhD-studies. The results presented in this thesis have previously been published in a number of papers([38, 40, 42, 39, 41]). The subject of the author's research is the development of general, proof-theoretic methods for the extraction of effective realizers and bounds from formal (possibly ineffective) proofs in mathematics and computer science, as well as carrying out concrete case studies based on these general techniques. This branch of proof theory (or more generally, mathematical logic) has in recent years been coined "proof mining".

The results presented in this thesis roughly fall into three parts:

First, a new approach to extracting Herbrand disjunctions from proofs in first order predicate logic PL is presented(this is joint work with U.Kohlenbach[40]). Usually Herbrand disjunctions are extracted using cut elimination or ε-elimination. Carrying out a suggestion by G.Kreisel([90]) we present an algorithm for the extracting Herbrand disjunctions based on a variant of Gödel's functional interpretation due to Shoenfield[112], which we adapt to E-PL$^\omega$, extensional predicate logic in all finite types. Here, the crucial point is the interpretation of the axiom $A \vee A \vdash A$, which usually requires some arithmetic to construct characteristic terms for all quantifier-free formulas of PL. The key idea then is to explicitly add decision-by-case constants for all quantifier-free formulas, so that one can interpret this axiom for general predicate logic PL without the use of arithmetic. Using this variant of functional interpretation one may then extract higher order Herbrand terms in E-PL$^\omega$. Normalizing these terms, one can then read off the actual Herbrand terms (which are again terms in PL) from the normal form. Known upper bounds on the length of normalization sequences in the simply typed λ-calculus (see [110, 3]) furthermore provide upper bounds on the size of the Herbrand disjunction that match the best known bounds obtained via cut elimination([37, 38]).

The next part of the thesis covers the general metatheorems for the extraction of effective bounds from proofs in functional analysis developed in [42, 41]. These results are also joint work with U.Kohlenbach, and the focus here is on extending previous metatheorems due to U.Kohlenbach(see [77]). Using so-called monotone functional interpretation (i.e. by combining Gödel's functional interpretation with the Howard-Bezem strong majorization relation), one may extract effective bounds from ineffective proofs in (classical) analysis $\mathcal{A}^\omega$(see

[68]). In [77], similar metatheorems are obtained for the extension of classical analysis with abstract metric, hyperbolic and CAT(0)-spaces, respectively abstract real normed linear spaces, uniformly convex spaces, Hilbert spaces and inner product spaces. "Abstract" here refers to the fact, that the space X is added to the formal system for analysis as new type X, along with the necessary constants and axioms, expressing the defining algebraic properties of such spaces. Previous results had only covered Polish spaces, which via the so-called standard representation are explicitly representable in $\mathcal{A}^\omega$. In the case of Polish spaces, one may extract effective bounds from ineffective proofs that are uniform with regard to parameters ranging over compact Polish spaces. In [77], such uniformities were obtained for abstract bounded *not necessarily compact* metric, hyperbolic and CAT(0)-spaces and bounded convex subsets of abstract real normed linear spaces. In [41], this is further extended to *unbounded* metric spaces and *unbounded* convex subsets, where one obtains similar uniformities in the presence of relatively weak, local boundedness conditions on e.g. a parameter $x \in X$ and an accompanying function $f : X \to X$. Using monotone modified realizability, similar results may also be obtained for so-called semi-intuitionistic theories, i.e. intuitionistic base theories for analysis extended with certain nonconstructive principles such as comprehension for all negated formulas(see [42]).

Finally, this thesis also describes a case study in metric fixed point theory. In this case study, the author analyzes a fixed point theorem for so-called asymptotic contractions due to W.Kirk[59]. The original proof depends on techniques from nonstandard analysis and thus is nonelementary and highly ineffective. By a proof-theoretic analysis an elementary, effective version of Kirk's fixed point theorem was obtained([39]).

A short final chapter describes future work, an appendix contains the papers, written during the author's PhD-studies, in which the results presented in this thesis have been published.

Acknowledgements

I would like to thank BRICS, the Department of Computer Science and the University of Aarhus for hosting and financing my PhD-studies. I would also like to the thank the Department of Mathematics at the Technical University of Darmstadt, Germany for their hospitality during my almost yearlong visit. Moreover, I would like to thank a number of individual people, but I dare not form a list for fear of forgetting someone who should be included: I would like to thank all the friends, colleagues and staff – at home and wherever I have travelled – who made my PhD-studies an unforgettable and very enjoyable journey. I would also like to thank my friends and family for understanding and accepting that a budding researcher's mind sometimes operates differently. Finally and most importantly, I would like to thank my supervisor, Ulrich Kohlenbach, whose generosity with regard to time and effort and, not the least, patience as well as his competence in supervising his students is second to none. Of all the professional decisions I have ever made, signing up for that first logic class taught by my, then, future supervisor is the one I cherish the most.

Philipp Gerhardy,
Aarhus, August 16, 2006.

Contents

Part I

Overview

Chapter 1

Introduction

Proof mining is the application of proof-theoretic techniques to the analysis of proofs in mathematics and computer science with the aim of extracting additional (constructive) data from a given, sufficiently formal proof. More precisely, given a proof of a statement $\forall x \exists y A(x, y)$ one usually is interested in developing and applying general, proof-theoretic techniques to extract either an exact realizer t – given by a computable functional – such that $\forall x A(x, tx)$ holds or a computable bound t satisfying $\forall x \exists y \leq tx A(x, y)$. Extracting a bound is often much simpler than extracting an exact realizer. (Naturally, extracting a bound only makes sense, if we have a meaningful definition of the $\leq$-relation for the domain of y.)

Extracting computable bounds or realizers is not possible in all cases. Already for the formula class $\forall x \exists y \forall z A_{qf}(x, y, z)$, where A_{qf} is a decidable quantifier-free formula and the variables x, y, z range over the natural numbers, there are simple counterexamples e.g. based on the halting problem: Let the predicate $T(e, x, y)$ be the Kleene T-predicate expressing that Turing-machine e with input x terminates with computation y. Then certainly the following is logically true and hence provable in first order predicate logic:

$$\forall x \exists y \forall z (T(x, x, y) \vee \neg T(x, x, z)),$$

as each Turing machine x taking itself as input either terminates with some computation y or does not terminate for any computation z. Both Turing machines, their input and computations (with a finite number of steps) can be encoded as natural numbers. Furthermore, there is a primitive recursive algorithm (shown to be even poly-time, see e.g. [99]) to check whether Kleene's T-predicate holds for given e, x, y. Nevertheless, it is obvious that one cannot extract a realizer or even just a bound for $\exists y$ (as a computable function in the parameter x), since this would decide the special halting problem (i.e. deciding the set $\{x : \exists y T(x, x, y)\}$), which is well-known to be undecidable.

Thus proof mining aims at two things: (1) proving metatheorems that classify formal systems, theorems and proofs for which some relevant data actually can be extracted, and (2) carrying out case studies where one uses these metatheorems and the implicit proof-theoretic techniques to analyze and extract additional data from actual proofs in mathematics and computer science. Note, that this is not a sequential or mutually independent activity: On the one hand, the metatheorems suggest particular case studies that lead to new mathematical results, on the other hand, ad-hoc results in case studies that go beyond what current metatheorems predict a-priorily may inspire and lead to new proof mining techniques and, in consequence, more powerful metatheorems.

We now give a short overview of the results presented in this PhD-dissertation:

In Chapter 2, a new, special variant of Gödel's functional interpretation, jointly developed with U.Kohlenbach, is presented. The variant is specially tailored towards extracting Herbrand disjunctions from proofs in ordinary first-order predicate logic. The extraction of Herbrand disjunctions is one of the earliest considered techniques for extracting additional data from (formal) proofs in mathematics. In addition to briefly discussing potential applications, these results will be compared with results in the literature on the usual technique (i.e. cut elimination) for extraction of Herbrand disjunctions. Some of the results that will be discussed in this chapter have previously been published in [37, 38, 40]. In Chapter 3, several new metatheorems for the extraction of strongly uniform bounds from non-constructive, resp. semi-constructive, proofs in functional analysis are presented. These new metatheorems continue a line of research begun in [77], where metatheorems were developed to treat formal systems for analysis extended with abstract metric spaces and abstract (real) normed linear spaces and further variants of such spaces (rather than only treating constructively representable spaces that may be completely formalized in e.g. second order arithmetic). The results presented in Chapter 3 are also joint work with U.Kohlenbach. Finally, Chapter 4 covers a recent case study in fixed point theory carried out by the author. In [59], Kirk proves a fixed point theorem for so-called asymptotic contractions. The proof is highly ineffective relying on methods from non-standard analysis. In [39], the proof is analyzed using techniques from proof mining and an elementary proof of an effective version of Kirk's fixed point theorem is obtained. However, the focus in Chapter 4 is not on the result itself, but rather on illustrating how the techniques of proof mining are used in an actual, non-trivial example of analyzing a nonconstructive mathematical proof. It has been attempted to make each chapter reasonably self-contained, i.e. certain definitions or very simple lemmas may be repeated in several chapters, while other, more substantial results are referred to by pointing to the appropriate chapter or a suitable reference in the literature.

The rest of this chapter will be devoted to giving a general introduction to the novel area of proof mining and its historical background.

* * *

The original motivation for proof mining is best summed up by the following question formulated by G.Kreisel in the 1950s:

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

Kreisel's question was a reaction to proof theory's preoccupation with consistency proofs for the various formal systems in which one could carry out mathematical and metamathematical investigations. During the early 20th century a wide array of proof-theoretic techniques had been developed primarily to prove (or disprove) the consistency of various formal systems. Consistency of a formal system was generally expressed by asserting the unprovability of obviously absurd statements such as '0 = 1'. These efforts had been spawned by what is today known as "the foundational crisis in mathematics" and in particular Hilbert's response to the crisis: Hilbert's consistency program.

In the early 20th century Brouwer and his followers had started to question the validity of several recent results in mathematics, such as e.g. the theorem that every continuous function on a closed and bounded interval attains its maximum or the Heine-Borel covering theorem, and had shown them to lack constructive[1] meaning (see [20]), although an intuitionistic version of the Heine-Borel theorem later found justification, at least to Brouwer and his school, through the fan theorem. (We shall discuss and give an informal definition of the notion of "constructive meaning" in a moment.) Brouwer blamed this on the new axiomatic, set-theoretic (in the sense of Frege and Cantor) approach to mathematics that had grown out of the late 19th-century efforts to formalize all of mathematics and equip it with a common foundation, primarily based on Frege's work on predicate logic and Cantor's work on set theory. He claimed that this new axiomatic approach had reduced mathematics to an arbitrary game of symbols severely lacking intuitive mathematical meaning. Brouwer especially criticized the application of logical principles that had originally found their (constructive) justification in a finitary context to infinitary situations. Therefore he deemed it necessary to reconstruct mathematics and to abandon those logical principles that could not be given a satisfactory constructive interpretation.

Hilbert was not willing to give up the benefits and achievements of this new approach – sometimes termed "Cantor's paradise" – of whose usefulness to mathematics he already was thoroughly convinced. In response to both Brouwer's critique and also to the inconsistencies found in the early attempts to formalize mathematics, Hilbert and his followers focused on proving the consistency of these new formal systems. The main idea was that a formal system in which mathematics could be developed (e.g. the formal system in Russell and Whitehead's Principia Mathematica) should prove its own consistency. Moreover, if the consistency proof would only use finitary means, the proof should also be acceptable for constructive mathematicians, thus securing mathematics against Brouwer's criticism. This dispute between Brouwer, Hilbert and several other mathematicians is commonly known as "the foundational crisis in mathematics"[2].

---

[1] In this introduction we shall use "constructive" and "intuitionistic" as if they were interchangeable - in reality, there are very subtle differences between "(Bishop's) constructive mathematics" and "(Brouwer's) intuitionistic mathematics"and furthermore "Russian constructivism". See [15] for a detailed comparison.

[2] The dispute between Brouwer and Hilbert and their respective followers is to some extent

Hilbert's program suffered a major setback when Gödel published his famous incompleteness results in the early 1930s([43]). In particular, Gödel's second incompleteness theorem showed that no sufficiently expressive formal system for mathematics could prove its own consistency, and it soon became clear that extensions going beyond the original strict finitary approach proposed by Hilbert would be necessary to establish the consistency of arithmetic and analysis. Despite the fact that Gentzen soon proved the consistency of arithmetic[35], though of course not in a completely finitary manner, subsequently a revised version of Hilbert's program was pursued. Gentzen had proved the consistency of arithmetic by employing transfinite induction, but another extension of the finitary approach would prove far more important for later applications in proof mining: Instead of proving the consistency of a formal system within that system itself, one instead aimed at showing, in a finitary way, that any possible inconsistency provable in a given questionable formal system is already provable in some uncontroversial formal system, preferably acceptable to constructive mathematicians. An important tool to carry out these so-called relative consistency proofs are proof interpretations.

In general, proof interpretations are transformations between formal systems, more precisely the formulas and proofs expressible in these systems, such that certain desirable properties are preserved. Most importantly, a given proof interpretation $\phi : \Sigma_1 \to \Sigma_2$ between formal systems Σ_1 and Σ_2 is required to preserve provability such that P is a proof of A in Σ_1 implies that $\phi(P)$ is a proof of $\phi(A)$ in Σ_2. Also, one would usually ask that there is a meaningful relationship between a formula A and its interpretation $\phi(A)$. To investigate questions of (relative) consistency one would additionally ask that quantifier-free formulas remain unchanged by the interpretation ϕ, in particular that $\phi(`0 = 1') = `0 = 1'$. Finally, proof interpretations should be given by a *computable* transformation in order for themselves to be acceptable from a constructive point of view. By defining a proof interpretation (satisfying these requirements) of a formal system Σ_1 into a suitable formal system Σ_2, in which one is rather certain that'$0 = 1$' cannot be proved, one can then establish the consistency of Σ_1 relative to Σ_2. Although some such translations between formal systems had already been developed in the 1930s and 1940s, e.g. negative translation (for a survey of different negative translations see Luckhardt[93] or Troelstra[117]) and Kleene's realizability interpretation[60], the first attempt to systematically define proof interpretations is due to Kreisel in [85, 86].

A desirable target system for proof interpretations is usually either a formal system based on intuitionistic logic or a purely equational functional calculus. An example of the former is e.g. interpreting Peano arithmetic in Heyting arithmetic (i.e. interpreting classical in intuitionistic arithmetic) using negative translation. As an example of the latter, one can take the interpretation of

chronicled in [46]. In [20], Brouwer writes "[A]n incorrect theory, even if it cannot be inhibited by a contradiction that would refute it, is nonetheless incorrect." The response by Hilbert can be found in [50]: "So in recent times we come upon statements like this: even if we could introduce a notion safely (that is, without generating contradictions) and if this were demonstrated, we would still not have established that we are justified in introducing the notion. (...) [I]f justifying a procedure means anything more than proving its consistency, it can only mean determining whether the procedure is successful in fulfilling its purpose."

Heyting arithmetic into Gödel's **T** using Gödel's functional ('Dialectica') interpretation(see [44]). In that way, proof interpretations can be used to show, in a finitary way, that if there were a classical proof of e.g. '$0 = 1$', then there is already a proof of '$0 = 1$' in an intuitionistic system or even in an essentially logic-free equational calculus.

In the 1950s Kreisel suggested to shift the focus in proof theory away from proving the (relative) consistency of formal systems. Instead one should investigate how additional constructive data could be salvaged, or rather, extracted from a given apparently non-constructive proof. Ideally, a successful extraction technique would (as often as possible, but of course not in all cases, as illustrated above by the halting problem) allow one to recover exactly the kind of constructive information that Brouwer had found missing and Kreisel now claimed to be merely hidden in a number of mathematical proofs. Originally known as "Kreisel's unwinding program" such proof-theoretic investigations have recently been coined "proof mining".

The tools that Kreisel proposed for these investigations are proof interpretations, and these are still the main tool in proof mining today. As mentioned above, the target systems for proof interpretations are usually intuitionistic systems or even a purely equational, functional calculus. In that way, the target system will either directly provide or at least give easy access to the "hidden" data we want to extract. In [85, 86, 87], Kreisel first introduces his "unwinding program" and primarily discusses the use of Herbrand's theorem and his own no-counterexample interpretation (to be discussed below) as examples of how such proof interpretations can be used to "unwind" non-constructive proofs and make additional constructive data hidden in the proof explicit.

The Brouwer-Heyting-Kolmogorov(BHK) interpretation for the logical constants of intuitionistic logic gives an informal idea of the notion of "constructive meaning" of a proof and may serve as an informal guide to what additional data one hopes to be able to extract from a given non-constructive proof using proof interpretations. Implicitly, the interpretation is already present in the writings of Brouwer and Kolmogorov on intuitionistic mathematics. Later, the interpretation is made explicit by Heyting, see e.g. [49].

The BHK-interpretation defines the meaning of the logical connectives of intuitionistic logic as follows:

(i) There is no proof of $\perp$ (falsity).

(ii) A proof of $A \wedge B$ is a pair (p, q) of proofs such that p is a proof of A and q is a proof of B.

(iii) A proof of $A \vee B$ is a pair (n, p) where n is an integer and p is a proof such that p is a proof of A if $n = 0$ and p is a proof of B if $n \neq 0$.

(iv) A proof of $A \rightarrow B$ is a construction p that for every proof q of A produces a proof $p(q)$ of B.

(v) A proof of $\forall x A(x)$ is a construction which for every construction c of an element of the domain of the variable x produces a proof $p(c)$ of $A(c)$.

(vi) A proof of $\exists x A(x)$ is a pair (c, p) where c is a construction of an element of the domain of the variable x and $p(c)$ is a proof of $A(c)$.

Note, that here "proof" does not denote a derivation from certain axioms and rules, but rather describes the kind of *constructions* necessary to verify a theorem. Also note, that $\neg A$ is understood as an abbreviation of $A \to \bot$, and a proof p of $\neg A$ is a construction that transforms every hypothetical proof of A into a contradiction.

The above notion of constructive (or intuitionistic) truth implicit in the BHK-interpretation should be compared with the classical notion of truth. The crucial points are the interpretation of disjunction, implication and the existential quantifier: Classically, a disjunction is true, if not both disjuncts can be false. From a constructive point of view, the classical interpretation of a disjunction $A \vee B$ can be expressed as $\neg(\neg A \wedge \neg B)$. In a similar vein, an existential statement is true in classical logic, if it cannot simultaneously be false for all elements in the domain of the existential quantifier. From a constructive point of view the classical interpretation of $\exists x A(x)$ is $\neg \forall x \neg A(x)$. Finally, in classical logic an implication $A \to B$ is merely an abbreviation of $\neg A \vee B$ and so is expressed by $\neg(A \wedge \neg B)$, whereas the BHK-interpretation explicitly asks for a *construction* relating proofs of A to proofs of B. Some of the different negative translations are based on very similar intuitionistic interpretations of the notion of truth in classical logic.

The law of the excluded middle, i.e. the principle that $A \vee \neg A$ holds for every formula A, does *not* satisfy the Brouwer-Heyting-Kolmogorov interpretation, as this would require a decision procedure for arbitrary (possibly open) formulas A. Clearly, such a decision procedure cannot exist; again, such a procedure would e.g. decide the halting problem. The law of excluded middle is true in classical logic, but not accepted in intuitionistic logic. This lack of constructive meaning of the law of the excluded middle, at least for undecidable A, along with its indiscriminate use in a number of mathematical proofs was the main pillar of Brouwer's critique (see [20]).

One of the proof interpretations considered by Kreisel, the proof interpretation implicit in Herbrand's theorem, provides a partial computational realization of clause (*vi*) of the BHK-interpretation on the interpretation of existential statements. In its most simple form Herbrand's theorem states that given a proof in predicate logic without equality $PL_{-=}$ of an purely existential statement, i.e. $\exists \underline{x} A_{qf}(\underline{x})$ where A_{qf} is quantifier-free, we can find a finite number of candidates for a realizer for the existentially quantified variables. These so-called Herbrand terms satisfy the condition that the corresponding Herbrand disjunction over these candidates is a tautology and that the original formula may be proved from that tautology. In other words, by Herbrand's theorem we may obtain a finite list of tuples of terms $\underline{t_0}, \underline{t_1}, \ldots, \underline{t_k}$, so that $A(\underline{x})[\underline{t_i}/\underline{x}]$ cannot be simultaneously false for all candidates $\underline{t_i}$, and hence the formula follows from that disjunction. This extends to predicate logic *with* equality, where the Herbrand disjunction is then a tautological consequence of a finite number of instances of equality axioms (such a disjunction is called a quasi-tautology). Trivially, this result also extends to the class of $\forall \underline{x} \exists \underline{y} A_{qf}$-formulas by substituting new constants for the variables in leading universal quantifiers. The proof interpretation implicit in the standard effective proof of Herbrand's theorem is cut elimination, which is used to transform a given formal proof in first-order predicate logic into a purely propositional (one might also say: combinatorial) proof of it's Herbrand disjunction. Alternatively, one can also obtain an extraction algorithm for Herbrand disjunctions based on a variant of Gödel's functional interpretation. For details on the latter approach to extracting Herbrand disjunctions see Chapter 2.

Let us illustrate Herbrand's theorem by an example, here in particular focusing on the use of the law of the excluded middle in the proof. As Herbrand's theorem in the above simple form is not directly applicable to $\forall\exists\forall$-formulas, we consider the so-called Herbrand normal form formula of the special halting problem instead. The Herbrand normal form of a given formula (which has to be in prenex normal form) is slightly weaker than the original formula in the sense that even constructively $\vdash A \to A^H$, but in general $\nvdash A^H \to A$. The reverse implication only holds with regard to logical validity (and so, by the completeness theorem, w.r.t. provability in first order theories not involving the Herbrand index functions), i.e. $\models A^H \Rightarrow \models A$. We consider:

$$\forall x \exists y (T(x,x,y) \vee \neg T(x,x,h(y))),$$

where h is the Herbrand index function corresponding to the variable z in the $\forall z$-quantifier in the above formulation of the special halting problem. To prove this statement (in classical logic) one uses the law of the excluded middle. The proof starts with the following instance of the law of the excluded middle

$$T(x,x,h(y)) \vee \neg T(x,x,h(y)),$$

to argue that a given Turing machine x with itself as input either does or does not stop with computation $h(y)$. From this disjunction one cannot yet derive the original formula, as one cannot introduce the necessary quantifiers. One therefore weakens (and permutes) this statement to what in fact is the Herbrand disjunction that one may extract from this proof

$$T(x,x,y) \vee \neg T(x,x,h(y)) \vee T(x,x,h(y)) \vee \neg T(x,x,h(h(y))).$$

Now, one can introduce quantifiers $\exists y$, for the term y in the first two disjuncts and for the term $h(y)$ in the last two disjuncts. This yields

$$\exists y (T(x,x,y) \vee \neg T(x,x,h(y))) \vee \exists y (T(x,x,y) \vee \neg T(x,x,h(y))).$$

Contracting the two instances into one and introducing $\forall x$, one obtains a proof of the Herbrand normal form of the halting problem example. The Herbrand terms, which are obvious in the proof, precisely illustrate the non-constructive use of the law of the excluded middle: let some y be given, then the Turing machine x either stops with computation $h(y)$ – and then we are done – or it does not stop with computation $h(y)$ – and then we stick with the original

y. So the Herbrand terms extractable from this proof are simply y and $h(y)$. Introducing new distinct, variables for the terms involving the Herbrand index function h, we may also obtain a Herbrand disjunction for the original formula:

$$T(x, x, y_0) \vee \neg T(x, x, y_1) \vee T(x, x, y_1) \vee \neg T(x, x, y_2).$$

From this statement we get back the original formula by quantifying the variables y_0, y_1, y_2 from right to left. First we introduce a $\forall$-quantifier for y_2 and an $\exists$-quantifier for y_1. Now the variable y_1 is free for the first two disjuncts and we can introduce the remaining $\forall$- and $\exists$-quantifiers.

However, the Herbrand normal form also leads to another interesting proof interpretation: Kreisel's no-counterexample interpretation for arithmetic (in short: n.c.i.). In the above proof sketch, the index function h is quantified over only implicitly. If we consider the theorem in the context of arithmetic (extended with suitable $\alpha(< \varepsilon_0)$-recursive functionals) so that x and y are natural numbers, and explicitly make the (number-theoretic) index function h a parameter, we may define a functional realizer ϕ for y, namely

$$\phi(x, h) = \begin{cases} h(0) & \text{if } T(x, x, h(0)) \\ 0 & \text{otherwise} \end{cases}$$

As given a concrete numeral x and a computable function h the statement $T(x, x, h(0))$ is decidable, the above realizer ϕ is a well-defined, computable functional.

The Herbrand index function h can be considered as an attempt to provide a counterexample to the existence of a suitable y for all z. Such a counterexample would have to produce for each potential y a suitable $z = h(y)$ such that $T(x, x, y) \vee \neg T(x, x, h(y))$ does not hold. The above realizer $\phi(x, h)$ for y shows that we can counter each potential counterexample h and thus prove – by providing the *computable* functional ϕ – that there indeed is *no* counterexample and that therefore the formula is (classically) true. Since classically every formula is equivalent to some formula in prenex normal form and every prenex normal form is equivalent with regard to validity to its Herbrand normal form, the no-counterexample interpretation provides a kind of finitary or at least partially constructive consistency proof for classical arithmetic.

To prove the soundness of the no-counterexample interpretation one has to describe a general algorithm for extracting from a given proof of a given formula functionals that satisfy the n.c.i. of that formula. In [85], Kreisel bases his original soundness proof of the n.c.i. on Hilbert's ε-substitution method. Alternative proofs are based on cut elimination[108] or, as was also pointed out by Kreisel in [88], Gödel's functional interpretation.

Both Herbrand's theorem and the no-counterexample can also be used to analyze real mathematical proofs. In [86], Kreisel sketches several applications of Herbrand's theorem and the n.c.i. to amongst others Littlewood's theorem and Artin's solution to Hilbert's 17th problem. Later H.Luckhardt, using similar ideas, extracted the first polynomial bounds to a theorem by Roth about the number of exceptionally good rational approximations to a given algebraic irrational number (see [94]). Also, G.Bellin applied to the no-counterexample interpretation to a proof of the infinite Ramsey theorem obtain a parametric form of the Ramsey theorem (see [4]). For further details on these early applications of proof mining see e.g. [25, 32, 95].

There are, however, certain limitations with regard to the modularity of the no-counterexample interpretation. The limitations become clear when one considers the modus-ponens rule. Recall, that given a formula A^H in Herbrand normal form, i.e. $\forall\exists$-formulas where the $\forall$-quantifiers range over numbers and number-theoretic Herbrand index functions (i.e. parameters of type 0 and 1), the n.c.i. asks for realizing functionals of type 2. Assume, given a proof of A and a proof of $A \to B$, we would like to extract functionals satisfying the n.c.i. of these two formulas and combine them to obtain functionals satisfying the n.c.i. of B. In that case the functional realizing $A \to B$ must transform functionals realizing A into functionals realizing B. In [71], it is shown that solving the no-counterexample interpretation of the modus ponens rule cannot be solved uniformly in the functionals realizing the n.c.i for A and $A \to B$ by primitive recursive functionals (in the sense of Gödel), and that the solution requires so-called bar-recursion (of type 0) which was introduced by Spector in [114].

* * *

More useful for the extraction of constructive data are Kreisel's modified realizability interpretation[88] and Gödel's functional interpretation[44]. Functional interpretation is used to interpret Heyting arithmetic (in all finite types[3]) into a suitable quantifier-free functional calculus such as Gödel's **T**. Kreisel's modified realizability interpretation produces effective realizers – again in e.g. Gödel's **T** – for formulas in Heyting arithmetic, albeit without eliminating quantifiers during the interpretation. Via the additional step of applying negative translation both interpretations may also be used to interpret classical theories, although modified realizability interpretation requires the additional step of A-translation to extract computational content from classical proofs.(A-translation was discovered independently by Friedman([33]) and Dragalin([29]).) The modified realizability interpretation is an almost direct implementation of the BHK-interpretation, whereas Gödel's functional interpretation goes beyond the BHK-interpretation, as it most importantly interprets and has to interpret the Markov principle M^ω (to be defined below), which is not derivable in Heyting arithmetic and has no *effective* modified realizability interpretation. We will in the following focus on Gödel's functional interpretation.

Gödel's functional ('Dialectica') interpretation (in short: FI), introduced in [44], consists of two main parts. Let $H^\omega :\equiv \textbf{WE-HA}^\omega + \textbf{AC} + \text{IP}_\forall + M^\omega$, where $\textbf{WE-HA}^\omega$ is weakly extensional[4] Heyting arithmetic extended to all finite types,

[3]Functionals in *all* finite types – rather than type 2 functionals, which are sufficient for Kreisel's n.c.i. – are necessary to solve amongst other things the interpretation of the modus ponens rule. For a detailed exposition see e.g. [63].

[4]In [44], originally considers *intensional* Heyting arithmetic, but the results hold also for the

AC is the full axiom of choice:

$$\forall x \exists y A(x,y) \to \exists f \forall x A(x, f(x)),$$

where the types of x, y are arbitrary, $\mathrm{IP}_\forall$ is the independence of premise principle for purely universal formulas $A_\forall$:

$$IP_\forall : (A_\forall \to \exists y B(y)) \to \exists y(A_\forall \to B(y)) \quad (y \notin \mathrm{FV}(A_\forall)),$$

where the type of y is arbitrary, and M^ω is the Markov principle:

$$M^\omega : \neg\neg\exists \underline{x} A_{qf}(\underline{x}) \to \exists \underline{x} A_{qf}(\underline{x}),$$

where A_{qf} is an arbitrary quantifier-free formula and $\underline{x}$ is a tuple of variables of arbitrary types (A_{qf} may contain further free variables).

First, one assigns to each formula $A(\underline{a})$ in the language of H^ω, where $\underline{a}$ are the free variables in A, a formula $A^D :\equiv \exists \underline{x} \forall \underline{y} A_D(\underline{x}, \underline{y}, \underline{a})$, where A_D is a quantifier-free formula and hence decidable. Note, that the variables $\underline{x}, \underline{y}$ may be of arbitrary types (for convenience we will subsequently only write single variables $\exists x \forall y$ instead of tuples). Second, one proves the soundness of this translation by providing an explicit procedure to transform a given proof of A into a proof of A^D, where furthermore the leading existential quantifier is witnessed by a suitable primitive recursive functional in $\underline{a}$. In detail, to prove the soundness of the interpretation one describes explicit functional realizers for (the Gödel transformation $()^D$ of) all the axioms and rules of H^ω. In particular, this proves that given a H^ω-proof of a sentence $\forall x \exists y A(x,y)$ one can extract from that proof a primitive recursive (in the sense of Gödel) term t such that $\forall x A_D(x, tx)$ holds. By additional arguments one shows in H^ω that a formula is equivalent to its Gödel transformation $()^D$, and hence also $\forall x A(x, tx)$ holds. Using negative translation this result extends to weakly-extensional Peano arithmetic in all finite types $\mathbf{WE\text{-}PA}^\omega$ + **QF-AC** (the quantifier-free axiom of choice), but only for the more restricted class of sentences $\forall x \exists y A_{qf}(x,y)$, where again A_{qf} is a quantifier-free formula.

As shown by Spector in [114], Gödel's result can be extended to classical analysis, i.e. weakly extensional Peano arithmetic in all finite types + the axiom schema of dependent choice. The terms extracted by functional interpretation are then (Spector) bar-recursive[5] instead of primitive recursive. The bar-recursive functionals can be obtained by adding the scheme of bar-recursion (informally: recursion over well-founded trees) in all finite types to Gödel's **T**. The scheme of bar-recursion is necessary to interpret the negative translation of the axiom schema of dependent choice. To some extent bar-recursion can be considered the computation equivalent of bar induction (which is a generalization of Brouwer's bar theorem to all finite types), but the technical details of

weakly-extensional variant. The reasons for restricting the use of extensionality are discussed in Chapter 3. For proofs in numerical analysis this restriction can to some extent be overcome by the elimination-of-extensionality-technique developed by Luckhardt and Gandy (see [93]).

[5] The original definition of bar-recursion is due to Spector[114], but there are alternative definitions due to resp. Kohlenbach[64] and Berardi, Bezem and Coquand[5](the latter was termed modified bar-recursion by Berger and Oliva[8]). For more details see [9].

this equivalence are non-trivial (see [51]). Although both Gödel's and Spector's results were originally motivated by proving the relative consistency of classical arithmetic and classical analysis, they can be also used in the spirit of Kreisel to extract constructive data from non-constructive proofs.

One special variant of functional interpretation as used in the results by Gödel and Spector and further extensions of this variant have shown to be particularly useful for practical applications in functional analysis:

The first idea is to combine functional interpretation with Howard and Bezem's majorization (resp. *strong* majorization) relation. The majorization relation is a hereditary extension of the $\leq$-relation to all finite types (based on the ground type 0 representing natural numbers). Combining majorization with functional interpretation results in the so-called monotone functional interpretation, first introduced in [68] (although the interpretation is implicit already in [65]), which is used to extract bounds instead of realizers: From a given proof in Heyting arithmetic in all finite types of a formula $\forall x \forall y \leq sx \exists z A(x, y, z)$ (where the types of x, y and z are arbitrary and s is a closed term) one can extract a primitive recursive term t and a majorant t^* of t such that $\forall x \forall y \leq sx(x^* \text{ s-maj } x \wedge y^* \text{ s-maj } y \rightarrow \exists z(t^* x^* y^* \text{ s-maj } z \wedge A(x, y, z)))$ holds. If the type of z is restricted to types 0 (natural numbers) then $t^* x^* y^*$ is a bound on z in the sense of the $\leq$-relation on the natural numbers. If x is of type 1 then one can effectively construct a majorant x^M of x and likewise one can, by induction on the term structure construct a majorant s^* of s. Then $s^* x^M$ is a majorant for y and so $\phi(x) = t^* x^M (s^* x^M)$ is a majorant for z. Thus, the monotone functional interpretation allows one to extract majorants that are uniform with regard to certain parameters, e.g. the parameter y bounded by sx. Additionally the monotone functional interpretation of axioms and rules is often much simpler and allows one even to interpret certain non-constructive principles such as e.g. weak König's lemma. For Heyting arithmetic this applies to proofs of arbitrary formulas A, for Peano arithmetic we are again restricted to quantifier-free formulas A_{qf}. As before, the results can be extended to classical analysis using Spector's bar-recursion scheme - however, only at the cost of some limitations on the admissible types in the formula for which bounds are to be extracted. We will discuss the details in Chapter 3.

The aim of the next extension (which builds upon the first) is to be able to treat theorems concerning abstract metric and (real) normed linear spaces X (and further variants such as hyperbolic spaces, CAT(0)-spaces, uniformly convex spaces, Hilbert spaces, etc.). In ordinary Heyting or Peano arithmetic in all finite types one can only treat metric and normed linear spaces that are also Polish spaces, i.e. complete separable metric spaces (as such spaces are constructively representable in $\mathbf{E\text{-}HA}^\omega/\mathbf{WE\text{-}PA}^\omega$). The elements of Polish spaces are given in their so-called standard representation: the countable dense subset of the space is represented by the natural numbers, elements of the Polish space as number-theoretic functions representing a Cauchy sequence with fixed rate of convergence, etc. In that way one can treat concrete Polish spaces such as $\mathbb{R}$ or $C[0, 1]$. Simultaneously, one may also add an *abstract* metric or normed linear space to e.g. $\mathbf{WE\text{-}PA}^\omega$ as a new ground type X and further add the necessary constants and axioms, but only those that (algebraically) characterize the class of metric, resp. normed linear spaces. E.g. for an abstract metric space (X, d) one adds a constant 0_X of the new ground type X representing an element of the space (asserting the non-emptiness of the space), a constant d_X of type $X \to X \to 1$ representing the metric function $d : X \to X \to \mathbb{R}$ (where $\mathbb{R}$ is represented as a Polish space in $\mathcal{A}^\omega$), and new axioms expressing e.g. that the metric is symmetric and that it satisfies the triangle inequality. In particular, we do not allow an axiom stating the seperability of the given space, as this axiom would have no monotone functional interpretation, i.e. monotone functional interpretation would immediately strengthen this axiom into a formula expressing the compactness of the space and ask for a realizer for this transformation. As not even every separable *bounded* space is compact this cannot be. Also, we cannot allow axioms stating the extensionality of e.g. all functions $f : X \to X$ as monotone functional interpretation would turn this into a statement expressing the uniform equicontinuity of all such functions f. The restrictions in particular on extensionality will be discussed in more detail in Chapter 3. Existing axioms of $\mathbf{WE\text{-}PA}^\omega$ are extended to the set of all finite types based on ground types 0 (natural numbers) and X (for elements of the space X), resulting in a formal system $\mathbf{WE\text{-}PA}^\omega[X, d]$. As we aim for a purely universal axiomatization the axioms in fact only express that (X, d) is a pseudo-metric space and d_X is only a proper metric on the set of equivalence classes with regard to $=_X$, where $x =_X y :\equiv d_X(x, y) =_{\mathbb{R}} 0$. See [77] for a more detailed discussion of this point.

Previous results on the extraction of bounds from proofs in the formal system $\mathbf{WE\text{-}PA}^\omega$ can be extended to these new formal systems in the following way (we continue the metric space example): First, one needs to extend the majorization relation to the new type X and check that the monotone functional interpretation for the axioms and constants of $\mathbf{WE\text{-}PA}^\omega$ still is valid when extended to the types $\mathbf{T}^{\omega,X}$. Second, one needs to define "suitable" majorants for the new metric space constants and check that the new metric space axioms have a monotone functional interpretation by "suitable" closed terms in the extended language $\mathbf{WE\text{-}PA}^\omega[X, d]$. For (pseudo-)metric spaces the latter part is easily satisfied, as pseudo-metric spaces can be axiomatized by a set of purely universal sentences not containing $\vee$ and such purely universal sentences are their own (monotone) functional interpretation.

The intended meaning of "suitable" depends on how one has extended the majorization relation to the new type X and what properties one hopes to prove about the extracted bounds. With the novel combination of majorization and functional interpretation first developed in [77] and recently extended in [41] one may prove metatheorems of the following kind: Assume one proves a theorem of the form $\forall x^1 \forall y^\rho \leq sx \forall z^\tau \exists v^0 A_{qf}(x, y, z, v)$ in a suitable theory of Peano arithmetic in all finite types extended with an abstract metric space (X, d). Assume furthermore that that z only ranges over a metrically bounded *not necessarily compact* space (X, d), that the proof only uses the defining algebraic properties of the space (X, d), and that the variables x and y are of type 1, with y bounded by sx. Then one may extract a computable bound $\phi(x, b)$ for v *only* depending on x and a bound b on the diameter of the space, but not on the parameters y and z. Similar independence or uniformity results had previously only been known for parameters z ranging over *compact* metric spaces, and the extraction of computable bounds had only been proved in the setting of Polish spaces, where one proves the uniformity from parameters ranging over compact Polish spaces (see e.g. [68, 81]). While independence of extracted bounds from parameters ranging over compact Polish spaces falls into the general category of compactness results, the general independence from parameters ranging over metrically bounded spaces had previously not even be shown ineffectively, and there are no general mathematical reasons why such strong uniformities should hold.

In [77], the metatheorems only treat the extraction of bounds for proofs in the extension of weakly-extensional Peano arithmetic in all finite types with *bounded* metric, hyperbolic and CAT(0)-spaces and *norm-bounded* convex subsets of real normed linear spaces, uniformly convex spaces, Hilbert spaces, etc. (Note, that non-trivial normed linear spaces themselves always are unbounded). In these settings one can obtain bounds that are uniform on the bounded metric spaces, resp.on the bounded convex subset of the normed linear space. Recently in [41], these metatheorems were extended to also cover *unbounded* metric spaces and *unbounded* convex subsets of normed linear spaces using a novel majorization technique. Instead of having to assume that the whole space or the whole convex subset is bounded, one can with the new metatheorems obtain similar uniformities from weak *local* boundedness conditions. We will discuss these results and the novel majorization technique in detail in Chapter 3.

Monotone functional interpretation and its most recent extensions have already found a number of significant applications in mathematics, in particular in approximation theory and metric fixed point theory. A detailed treatment of these results can be found in [63], covering recent applications in Chebycheff approximation[66, 67], L_1-approximation[82] and metric fixed point theory[73, 75, 79]. Additional, recent case studies can be found in [76, 80, 92, 18].

The present author carried out an analysis of a highly ineffective proof of a fixed point theorem due to Kirk and obtained an elementary proof of an effective version of that theorem([39]); this case study is discussed in detail in Chapter 4.

Finally, one should mention that proof mining recently also has found applications in computer science, more precisely in the analysis of programming languages. A version of Kreisel's modified realizability interpretation has also been used to extract normalization algorithms from proofs of strong normalization (see e.g. [6]) and weak head normalization (see [13]) for the simply typed λ-calculus.

Chapter 2

Extracting Herbrand Disjunctions by Functional Interpretation

Some of the results presented in this chapter have previously been published in [37, 38] and (in a joint paper with U.Kohlenbach) in [40].

Among the earliest techniques for unwinding proofs are cut elimination and, related to that, ε-elimination. Both approaches immediately yield Herbrand's theorem, which states that for every proof in first-order predicate logic (without equality) of a sentence $\exists x A_{qf}(x)$ there is a collection of closed terms $t_1, \ldots t_n$ – consisting only of A_{qf}-material and possibly some arbitrary constant c, if no constant occurs in A_{qf} – so that the (Herbrand) disjunction $\bigvee_{i=1}^{n} A_{qf}(t_i)$ is a tautology (as usual, A_{qf} here denotes a quantifier-free formula). In other words, Herbrand's theorem states that for every proof of a sentence $\exists x A_{qf}(x)$ there is a list $t_1, \ldots t_n$ of terms, potentially realizing the quantifier $\exists x$, so that $A_{qf}(t_i)$ cannot simultaneously be wrong for all terms t_i. The terms $t_1, \ldots, t_n$ thus *witnessing* the proof of $\exists x A_{qf}(x)$ are called Herbrand terms.

Herbrand's theorem generalizes to tuples of existential quantifiers $\exists \underline{x} A_{qf}(\underline{x})$ and furthermore to arbitrary formulas A in prenex normal form by considering the Herbrand normal form A^H of A, where the Herbrand terms then are made of A^H-material. For open theories T, i.e. theories whose axioms are all purely universal, one obtains a Herbrand disjunction verifiable in T, that is $T \vdash \bigvee_{i=1}^{n} A^H_{qf}(\underline{t_i})$. More precisely, the disjunction is a tautological consequence of finitely many closed instances of axioms of T, where now the Herbrand terms may contain constants occurring in the axioms of T. Here, the Herbrand index functions must be new with regard to both the formula A and the theory T. Via an open axiomatization of equality one may thus also treat first order logic with equality.

The Herbrand terms are built up out of the the constants and function symbols occurring in A, resp. A^H. For open theories T the Herbrand terms may additionally contain constants and function symbols occurring in the non-logical axioms of T. The actual construction of Herbrand terms is important in the area of computational logic and has also been used to analyze actual mathematical proofs (see [91, 94]).

In addition to proof-theoretic proofs of Herbrand's theorem, using e.g. cut elimination or ε-elimination, there are also model-theoretic proofs of Herbrand's theorem, which, however, are ineffective. Only the proof-theoretic proofs provide an explicit algorithm for extracting Herbrand terms t_i from a given proof of A.

A new, alternative approach to extracting Herbrand disjunctions is via Gödel's functional ('Dialectica') interpretation (in short: FI), an approach that was suggested by G.Kreisel in his review [90] of [112]. Functional interpretation is usually applied to proofs in intuitionistic ('Heyting') arithmetic or – via negative translation – classical ('Peano') arithmetic to extract realizers and bounds (see Chapter 3). By adapting a variant of Gödel's functional interpretation due to Shoenfield[112] (which we will also denote by FI), we may also treat first-order predicate logic PL without equality. Shoenfield's variant of functional interpretation combines negative translation and functional interpretation into one step and only uses properties of arithmetic to ensure the existence of decision-by-case terms for all quantifier-free formulas[1]. By explicitly adding decision-by-case constants for all quantifier-free formulas of PL to the language $\mathcal{L}(\mathrm{PL})$, we may reuse Shoenfield's soundness proof for functional interpretation of PL in E-PL$^\omega$:= PL extended to all finite types and with extensionally defined equality plus decision functionals for all quantifier-free formulas.

Using this adapted variant of functional interpretation we can for proofs of sentences $\exists x A_{qf}$ in the language $\mathcal{L}(\mathrm{PL})$ extract realizing terms t in the extended language of E-PL$^\omega$ which are in effect higher-order Herbrand terms. From the normal form $nf(t)$ we may then read off a collection of Herbrand terms $t_1, \dots t_n$, where the t_i again are ordinary closed terms of PL not containing any higher type constructs or decision-by-case constants. From upper bounds on the length of normalization sequences in the typed λ-calculus one obtains corresponding estimates on the number of Herbrand terms extracted from a given proof. These upper bounds can be shown to match the best upper bounds obtained via cut elimination[37, 38].

The previous proof-theoretic approaches to constructing Herbrand terms (based on cut elimination or ε-substitution) have two main disadvantages: (1) the proof-theoretic approaches are not modular in an efficient way, i.e. given Herbrand terms for A and Herbrand terms for $A \to B$ one can only produce Herbrand terms for B at a non-elementary expense, and (2) the complexity of extracting the Herbrand terms from a given proof is in general non-elementary, as was shown by Statman[115]. The non-elementary complexity of the extraction of Herbrand terms is to some extent unavoidable, as Statman's proof gives a sequence S_n of sentences which have short proofs (linear in n), but for which

[1] In [83], it is shown that Shoenfield's variant of functional interpretation can be obtained by composing Gödel's functional interpretation with a negative translation due to Krivine.

every Herbrand disjunction must have at least superexponentially[2] in n many disjuncts. In the FI approach to Herbrand's theorem the extraction of the realizers is elementary (more precisely: cubic, see [48]), while only the final normalization step is of non-elementary complexity. However, already the unnormalized terms extracted by FI may be useful for certain metamathematical as well as mathematical applications. Furthermore, the extraction of the higher-order terms itself is fully modular and it is only after the final normalization step that modularity is lost.

We now briefly recall the system of first-order predicate logic PL (without equality) and its extension E-PL$^\omega$ to all finite types in which the proof of the extraction of Herbrand terms is carried out (for the complete details see [40]):

First-order predicate logic PL

The logical constants of the language $\mathcal{L}$(PL) are the connectives $\neg, \vee$ and $\forall$. The connectives $\wedge, \rightarrow$ and $\exists$ are expressed via their usual abbreviations. As we will see later, counting the number of nested negations is important for an estimate on the degree of the FI extracted terms. Therefore we translate blocks of quantifiers in the following way: $\exists \underline{x} A(\underline{x}) :\equiv \neg\forall\underline{x}\neg A(\underline{x})$. The language $\mathcal{L}$(PL) contains variables $x, y, z, \ldots$, constants $c, d, \ldots$, (possibly empty) sets of function symbols $f, g, \ldots$ for every arity n and predicate symbols $P, Q, \ldots$. We use the same symbols for free and bound variables. Formulas and terms are defined in the usual way. The axioms and rules of PL are as in [112].

Note. *We assume w.l.o.g. that there exists at least one constant symbol c in our language $\mathcal{L}$(PL), as Herbrand's theorem would fail otherwise.*

Extension of predicate logic to all finite types

The set $\mathbf{T}$ of all finite types is defined inductively:

$$(i) 0 \in \mathbf{T}, \quad (ii) \rho, \tau \in \mathbf{T} \Rightarrow (\rho \rightarrow \tau) \in \mathbf{T}$$

For convenience we write $0^n \rightarrow 0$ for $\overbrace{0 \rightarrow (0 \rightarrow (\ldots (0}^{n} \rightarrow 0) \ldots))$.

The language of E-PL$^\omega$

The language of E-PL$^\omega$ is based on a many-sorted version of PL with variables $x^\rho, y^\rho, z^\rho, \ldots$ and quantifiers $\forall^\rho, \exists^\rho$ for all types ρ. The constants $c, d, \ldots$ of PL are embedded into E-PL$^\omega$ as constants of type 0 and function symbols $f, g, \ldots$ of PL are embedded into E-PL$^\omega$ as constants of type $0^n \rightarrow 0$ for functions of arity n. In addition to the constants and functions of PL, E-PL$^\omega$ contains decision-by-case constants χ_A of type $0^n \rightarrow 0 \rightarrow 0 \rightarrow 0$ for all quantifier-free formulas A in the original language $\mathcal{L}$(PL), where n is the number of free variables in A. Moreover, E-PL$^\omega$ contains a λ-abstraction operator. The predicate symbols of E-PL$^\omega$ are the predicate of PL and equality $=_0$ of type 0. Equality for higher types $\rho = \rho_1 \rightarrow \ldots \rightarrow \rho_n \rightarrow 0$ is defined extensionally, i.e.

$$s =_\rho t :\equiv \forall x_1^{\rho_1}, \ldots, \forall x_n^{\rho_n}(s\underline{x} =_0 t\underline{x}).$$

[2]The superexponential function $s(n)$ is defined by $s(0) = 1$ and $s(n+1) = 2^{s(n)}$.

Formulas are built up in the usual way from prime formulas. The terms of E-PL$^\omega$ are built up from constants c^ρ and variables x^ρ by λ-abstraction and application: if x^ρ is a variable of type ρ and t^τ is a variable of type τ then $\lambda x.t$ is a term of type $\rho \to \tau$; if t is a term of type $\rho \to \tau$ and s is a term of type ρ then (ts) is a term of type τ. For convenience, given an n-ary function symbol f of PL and terms $t_1, \ldots, t_n$ of type 0 we usually write $f(t_1, \ldots, t_n)$ instead of $((\ldots(ft_1)\ldots)t_n)$.

Axioms and Rules of E-PL$^\omega$

- axioms and rules of PL extended to all sorts of E-PL$^\omega$,
- axioms for β-reduction in the typed λ-calculus,
- equality axioms for $=_0$,
- higher type extensionality:

$$E_\rho : \forall z^\rho, x_1^{\rho_1}, y_1^{\rho_1}, \ldots, x_k^{\rho_k}, y_k^{\rho_k} (\bigwedge_{i=1}^{k} (x_i =_{\rho_i} y_i) \to z\underline{x} =_0 z\underline{y}),$$

  where $\rho = \rho_1 \to (\rho_2 \to (\ldots \to (\rho_k \to 0)\ldots))$,
- axioms for the decision-by-case constants $\chi_{A_{qf}}$: $A_{qf}(\underline{x}) \to \chi_{A_{qf}}\underline{x}yz =_0 y$ and $\neg A_{qf}(\underline{x}) \to \chi_{A_{qf}}\underline{x}yz =_0 z$, where $\underline{x}$ are the free variables of the quantifier-free formula A_{qf} of $\mathcal{L}(\text{PL})$.

The main theorem we prove in [40] is the following:

Theorem 2.1 (**[40]**). *Assume that* PL *proves the sentence* $\exists x A_{qf}(x)$. *Then there is a collection of closed terms* $t_1, t_2, \ldots, t_n$ *in* $\mathcal{L}(\text{PL})$ *which can be obtained by normalizing a* FI *extracted realizer* t *of* $\exists x$ *s.t.* $\bigvee_{i=1}^{n} A_{qf}(t_i)$ *is a tautology. The terms* t_i *are built up out of the* A_{qf}*-material (possibly with the help of the distinguished constant* c *in case* A_{qf} *does not contain any constant). Moreover,* $n \leq 2^{\#_\chi(nf(t))}$.
The theorem also extends to tuples $\exists\underline{x}$ *of quantifiers.*

This is Theorem 11 in [40]. In the estimate $n \leq 2^{\#_\chi(nf(t))}$ on the number n of Herbrand terms, $nf(t)$ is the normal form of the realizer t for $\exists x$ extracted by FI and $\#_\chi(nf(t))$ is the number of occurrences of decision-by-case constants χ in the normal form $nf(t)$.

In order to obtain an estimate on the number of Herbrand terms from the original PL proof of $\vdash \exists x A_{qf}(x)$, we need to estimate the size and type level of the extracted realizer t. We can then use well-known estimates on the length of normalization sequences in the typed λ-calculus (see [3]) to get an estimate on the number of χs in the normal form $nf(t)$.

The crucial observation is that in Shoenfield's variant of FI only negation increases the type of the functional realizers. Since none of the derivation rules further increase the types, it suffices to count the largest number of nested negations in the cut formulas of the original proof.

In [40] the notion of ¬-depth is defined rather informally as the largest number of nested negations. The precise intended meaning is the following:

Definition 2.2. *The ¬-depth of a formula A is the largest number of nested negations in front of an innermost ∀-quantifier, i.e. a ∀-quantifier in front of a purely propositional formula.*

We do not count negations in purely propositional formulas, as such formulas need no realizer. The definition of the degree of a formula is then as follows:

Definition 2.3 ([40])**.** *Let A be a formula, then we define the degree $dg(A)$ to be the ¬-depth of A. Let ϕ be a proof, then $dg(\phi)$ is the maximum degree of cut formulas occurring in ϕ and the end-formula of ϕ. The end-formula always is purely existential, hence $dg(\phi) = \max\{1, dg(A_1), \ldots, dg(A_n)\}$ for cut formulas A_i in ϕ.*

Using the aforementioned upper bounds on the length of normalization sequences in the typed λ-calculus by Beckmann[3] (refining earlier results by Schwichtenberg[109, 110]), we prove the following:

Corollary 2.4 ([40])**.** *The number of terms extracted in Theorem 2.1 from a proof ϕ can be bounded by $2^{3\|t\|}_{dg(\phi)+1}$.*

Here, we employ the usual definition of the superexponential function 2^x_y (i.e. $2^X_0 = x$ and $2^x_{y+1} = 2^{2^x_y}$), and $\|t\|$ is the usual definition of the *size* of a term, i.e. the number of symbols in t.

The proof of Theorem 2.1 uses the following steps:

First, we define an extension of models $\mathcal{M}$ of PL to models $\mathcal{M}^\omega$ for E-PL$^\omega$ such that the models agree on the validity of sentences $A \in \mathcal{L}(\mathrm{PL})$.

Definition 2.5 ([40])**.** *Let $\mathcal{M} = \{M, \mathcal{F}\}$ be a model for $\mathcal{L}(\mathrm{PL})$. Then $\mathcal{M}^\omega = \{M^\omega, \mathcal{F}^\omega\}$ is the full set-theoretic type structure over M, i.e. $M^0 :\equiv M$, $M^{\rho\to\tau} :\equiv {}^{M^\rho}{M^\tau}$ and $M^\omega :\equiv \langle M^\rho \rangle_{\rho\in T}$. Constants, functions and predicates of $\mathcal{M}$ retain their interpretation under $\mathcal{F}$ in $\mathcal{F}^\omega$. λ-terms are interpreted in the obvious way. Furthermore, $\mathcal{F}^\omega$ defines the following interpretation of χ_A:*

For $\underline{a}, b, c \in M$ we define $[\chi_A]_{\mathcal{M}^\omega}\underline{a}bc := \begin{cases} b & \text{if } \mathcal{M} \models A_{qf}(\underline{a})^3 \\ c & \text{otherwise.} \end{cases}$

Proposition 2.6 ([40])**.** *$\mathcal{M}^\omega$ is a model of* E-PL$^\omega$*. If A is a sentence of* $\mathcal{L}(\mathrm{PL})$ *and $\mathcal{M}^\omega \models A$, then $\mathcal{M} \models A$.*

Next, we adapt Shoenfield's soundness proof of FI for Peano arithmetic PA to first-order predicate logic PL:

[3] More precisely, $\mathcal{M} \models A_{qf}(\underline{a})$ means that $A_{qf}(\underline{x})$ holds in $\mathcal{M}$ provided the free variables x_i get assigned the element $a_i \in M$.

Lemma 2.7 (**[40]**)**.** *If* PL $\vdash \exists x A_{qf}(x)$ *then* FI *extracts a closed term* t^0 *of* E-PL$^\omega$ *s.t.* E-PL$^\omega \vdash A_{qf}(t)$.
The proof of $A_{qf}(t)$ *can actually be already carried out in the quantifier-free fragment* qf-WE-PL$^\omega$ *(in the sense of* [117]*) of* WE-PL$^\omega$*, where the latter is the fragment of* E-PL$^\omega$ *which results by replacing the extensionality axioms by the quantifier-free weak rule of extensionality due to* [114] *(see also* [72]*).*

Note, that here $\exists x A_{qf}(x)$ is assumed to be a closed formula. For open formulas we introduce new constants for the free variables, carry out the extraction and then reintroduce the free variables to obtain a corresponding Herbrand disjunction for the open case.

The proof of this lemma (Lemma 8 in [40]) is essentially Shoenfield's proof. The only cases in the functional interpretation of the axioms and rules that need to be adapted are the expansion rule $B \vdash B \vee C$ and the contraction rule $A \vee A \vdash A$.

For the expansion rule, if $B \vee C$ has been inferred from B, we need closed terms of suitable type to realize C. As we assumed there exists at least one constant c of type 0, we can construct the necessary closed terms using λ-abstraction.

Adapting the interpretation of the contraction rule is somewhat more complicated. Assume we inferred A from $A \vee A$. Then we need to form one realizer for A from the two realizers for A on the left-hand side, resp. right-hand side of the $\vee$. The informal idea is this: if the realizer for the left copy of A is valid we choose this realizer for the one A in the conclusion, otherwise we choose the realizer for the right copy of A. Thus, if either copy of A in $A \vee A$ had a valid realizer, we will select that one for the A in the conclusion. Otherwise $A \vee A$ was invalid and any realizer for A will do, in particular the chosen one from the right copy of A.

To carry out this construction in Peano arithmetic PA one uses the decidability of prime formulas of PA (and thus all quantifier-free formulas of PA) which allows one to define the necessary decision-by-cases terms. In E-PL$^\omega$ we have explicitly added the necessary decision-by-case constants, allowing to track and later unwind the decision-by-case instances and form the corresponding Herbrand disjunction.

The final step in proving Theorem 2.1 is normalizing the term t extracted by functional interpretation from a given proof of $\exists x A_{qf}(x)$ (where t in effect is a higher order Herbrand term) and reading off closed terms $t_1, \ldots, t_n \in \mathcal{L}(\mathrm{PL})$ from the normal form $nf(t)$ of t. With these terms we may then form the Herbrand disjunction $\bigvee_{i=1}^{n} A_{qf}(t_i)$ and via the equivalence of models $\mathcal{M}$ for PL and extended models $\mathcal{M}^\omega$ of E-PL$^\omega$ for formulas in $\mathcal{L}(\mathrm{PL})$ show that $\bigvee_{i=1}^{n} A_{qf}(t_i)$ is a tautology.

More formally, this final step is expressed via the following lemmas (Lemmas 9 and 10 in [40]):

Lemma 2.8 (**[40]**)**.** *If* E-PL$^\omega \vdash A_{qf}(t)$ *and* $nf(t)$ *is the* β*-normal form of* t*,* *then* E-PL$^\omega \vdash A_{qf}(nf(t))$.

Lemma 2.9 (**[40]**). *If t is of type 0, closed and in β-normal form, then there exist closed terms $t_1, \ldots, t_n \in \mathcal{L}(\mathrm{PL})$, s.t. $\mathcal{M}^\omega \models t = t_1 \vee \ldots \vee t = t_n$. Moreover, $n \leq 2^{\#_\chi(nf(t))}$, where $\#_\chi(nf(t))$ is the total number of all χ-occurrences in $nf(t)$.*

In conclusion, the approach to extract Herbrand terms via functional interpretation provides several benefits:

First of all, we obtain a modular, elementary algorithm for extracting higher order Herbrand terms. Here modular refers to the fact that the terms extracted from e.g. proofs of A and $A \to B$ can be combined directly to form a realizer for B. The extraction of the higher order terms itself only is of elementary, in fact, cubic complexity (see [48]). The only non-elementary part of the extraction of Herbrand terms is the normalization step to produce the normal form from which the actual first-order Herbrand terms can be read off. However, certain structural properties, such as e.g. bounds on the computational complexity or independence from parameters may already be read off from the higher order Herbrand terms *prior to normalization.* In [91] Kreisel discusses that one may derive new results by analysing such properties of Herbrand terms. An example of this is the analysis of proofs of Roth's Theorem carried out by Luckhardt in [94].

Secondly, FI has recently been implemented in Schwichtenberg's MINLOG system by M.D.Hernest[47]. As the MINLOG system also contains an efficient normalization tool (see [10]) one can expect that the system can be adapted to yield a useful Herbrand-term extraction tool.

Finally, the FI approach to extracting Herbrand disjunctions provides – via estimates on the length of normalization sequences in the typed λ-calculus – upper bounds on the number of Herbrand terms in a Herbrand disjunction extracted from a given proof. These bounds, as we will discuss next, match the best known upper bounds obtained via cut elimination [37, 38].

The bounds obtained via cut elimination show that the size of cut-free proofs, and thereby also the size of Herbrand disjunctions, obtained from a given proof primarily depends on two things: (1) contractions on cut-formulas or subformulas of ancestors[4] to cut-formulas, and (2) the number of quantifier alternations in such contracted formulas. Note that in [37, 38], the analysis of the complexity of cut elimination is carried out for the sequent calculus LK (as formulated in [37]), rather than the calculus due to Shoenfield used above.

We call a formula A 'purely $\exists, \vee$' (resp. 'purely $\forall, \wedge$') if it only consists of atomic formulas and the connectives $\exists$ and $\vee$ (resp. $\forall$ and $\wedge$). The notion of alternating quantifier depth is defined recursively:

[4]Informally, an ancestor of a formula A is a predecessor of A in a given proof tree. E.g. in the inference $\frac{\Gamma \vdash \Delta, A \quad \Pi \vdash \Lambda, B}{\Gamma, \Pi \vdash \Delta, \Lambda, A \wedge B}$, the formulas A and B are ancestors of the formula $A \wedge B$ and all the formulas in $\Gamma, \Pi, \Delta, \Lambda$ in the premise are ancestors of the corresponding formulas in the conclusion.

Definition 2.10 (**[38]**). *We define the alternating quantifier depth $aqf(A)$ of a formula A as follows:*

- *if A is atomic, purely $\exists, \vee$ or purely $\forall, \wedge$ then $aqf(A) = 0$*
- *if A is composed of formulas $B_1, \ldots, B_n$ (each with outermost connective a quantifier) by propositional connectives only then $aqf(A) = max\{aqf(B_i)\}$*
- *if $A(:= \forall xC$ for some $C)$ is composed of connectives $\forall, \wedge$ and formulas $B_1, \ldots, B_n$ (each with outermost connective $\exists$ or $\vee$) then $aqf(A) = max\{aqf(B_i)\} + 1$*
- *if $A(:= \exists xC$ for some $C)$ is composed of connectives $\exists, \vee$ and formulas $B_1, \ldots, B_n$ (each with outermost connective $\forall$ or $\wedge$) then $aqf(A) = max\{aqf(B_i)\} + 1$*

Moreover we treat implication $B \to C$ as $\neg B \vee C$, and negation $\neg B$ simply flips the polarity of other connectives below, i.e. $\exists, \vee \mapsto \forall, \wedge$ and vice versa.

With this definition of aqf for formulas, we define aqf for proofs:

Definition 2.11 (**[38]**). $aqf(\phi) := \sup\{aqf(A) : A \text{ is a cut formula in } \phi\}$

For the role of contractions we use the following definition:

Definition 2.12 (**[38]**). *Let $B_1, \ldots B_n$ be the ancestors of a formula A that appear as main formula in a contraction inference. Then the* c*ontracted alternating quantifier depth, $caqf(A)$, of A is $\max\{aqf(B_1), \ldots, aqf(B_n)\}$.*
Moreover, $caqf(\phi) := \sup\{caqf(A) : A \text{ is a cut formula in } \phi\}$.

In [38], the following result is proved:

Theorem 2.13 (**[38]**). *Let ϕ be an* **LK***-proof of a sequent $\Gamma \vdash \Delta$. Then there is a constant c depending only on the propositional blocks of the cut formulas and a cut free proof ϕ' of the same sequent where $|\phi'| \le 2^{c \cdot |\phi|}_{caqf(\phi)+2}$.*

Here, $|\phi|$ denotes the *depth* of the proof tree representing ϕ, not counting weakenings and contractions. The overall strategy in cut elimination is to repeatedly replace cuts by one or more simpler cuts. Atomic cuts can be eliminated completely. As either the number of cuts with maximal complexity decrease or the maximal cut complexity itself decreases this process terminates and produces a cut-free proof. The proof of the above cut elimination theorem consists of two main parts. First one shows that elimination of the uncontracted parts of the cut-formulas can be done at the cost an exponentiation growth in the depth of the proof. Next one shows that decreasing the alternating quantifier depth $aqf(\phi)$ of the entire proof by one also costs at most one exponentiation. Finally, eliminating the remaining propositional or atomic cuts costs a final exponential increase in the size of the proof. The combination of these observations yields the bounds given above. Naturally, if $caqf(\phi) = aqf(\phi)$ we do need to peel off the uncontracted parts of the cut formulas first and the upper bound is then $|\phi'| \le 2^{c\cdot|\phi|}_{aqf(\phi)+1}$.

To obtain a Herbrand disjunction it is enough that the proof is almost cut-free, i.e. that all cut formulas are quantifier-free. The proof can then be rearranged, possibly changing its depth but not its size, into so-called midsequent form, from which a Herbrand disjunction then can be read off. The size of the (rearranged) proof is then an upper bound on the size of the Herbrand disjunction. The cost of obtaining an almost cut-free proof is one exponentiation less, but estimating the size of the proof in terms of the depth of the proof adds an exponentiation again. Thus the bound on the number of Herbrand terms in the Herbrand disjunction extracted from a given proof ϕ is the same, namely $\le 2^{c\cdot|\phi|}_{caqf(\phi)+2}$, resp. $\le 2^{c\cdot|\phi|}_{aqf(\phi)+1}$.

In the FI approach the number of terms in the Herbrand disjunction is determined by the $\neg$-depth of cut-formulas. As in the restricted Shoenfield calculus $\exists$-quantifiers are represented by $\neg\forall\neg$ (similarly for $\vee$ and $\wedge$), counting nested negations in the FI approach directly corresponds to counting quantifier alternations in the cut elimination approach. Similarly, the role of contractions is reflected in the FI approach, as it is the contraction inferences in Shoenfield's calculus that introduce the decision-by-case constants. Without decision-by-case constants the extracted higher order term normalizes to exactly one PL-term. Similarly, a contraction-free LK-proof yields a Herbrand-disjunction consisting of just one element. Whereas one has to carry out a subtle analysis of the cut elimination procedure to arrive at the bounds based on counting quantifier alternations, these bounds almost come for free using the FI approach combined with well-known upper bounds on the length of a normalization sequence in the typed λ-calculus.

In [40], to compare upper and lower bounds on Herbrand's theorem a variant due to Pudlak of Statman's lower bound example, i.e. a sequence S_n of sentences for which the Herbrand disjunction has superexponentially in n many elements, is discussed. Using this variant one shows that the upper and lower bounds coincide (up to a constant factor) for both the cut elimination approach and the FI approach.

Chapter 3

Applications of Monotone Proof Interpretations

In this chapter we present a number of metatheorems for the extraction of effective bounds from classical and semi-intuitionistic proofs in (nonlinear) functional analysis. These metatheorems were developed by the author jointly with U.Kohlenbach in [42, 41]. These two papers can be considered follow-up papers to U.Kohlenbach's previous paper on the subject([77]), to which we also will refer when appropriate.

In the introduction of this thesis, we sketched briefly how Gödel's functional interpretation (and similarly, Kreisel's modified realizability interpretation) can be used to systematically extract effective realizers from constructive and, via negative translation, even non-constructive proofs. In this chapter, we discuss monotone variants of these interpretations and how to adapt these monotone variants to classical and semi-intuitionistic theories extended with abstract metric and real normed linear spaces. These monotone variants of Gödel's functional interpretation and Kreisel's modified realizability interpretation form the main proof-theoretic tool underlying the metatheorems presented in [42, 41].

In [52], Howard introduces the 'majorization' relation to prove that already for functionals of type 2 the full axiom of extensionality has no functional interpretation by primitive recursive functionals. Moreover, Howard shows that for functionals of type 3 there are even models of Zermelo-Fraenkel set theory in which there do not exist functionals satisfying the functional interpretation of the corresponding full extensionality axiom. This is due to the fact that functional interpretation satisfies (and must satisfy) the Markov principle:

$$M^\omega : \neg\neg\exists \underline{x} A_{qf}(\underline{x}) \to \exists \underline{x} A_{qf}(\underline{x}),$$

where A_{qf} is an arbitrary quantifier-free formula. Instances of the extensionality axiom combined with the Markov principle allow to derive the counterexamples by Howard mentioned above. In consequence, functional interpretation of arithmetic and analysis can be carried out at most for formal systems with the Spector's weak extensionality rule([114]) instead of the full extensionality axiom.

The majorization relation is a hereditary extension of the $\leq$-relation on the natural numbers. The relation is defined as follows:

$$\begin{array}{ll} x^* \mathrm{maj}_0 x & :\equiv x^* \geq x, \\ x^* \mathrm{maj}_{\rho\to\tau} x & :\equiv \forall y^*, y(y^* \mathrm{maj}_\rho y \to x^* y^* \mathrm{maj}_\tau xy. \end{array}$$

The majorization relation was later extended by Bezem[12] to the so-called *strong* majorization relation by adding an additional clause to the inductive part of the definition:

$$x^* \text{ s-maj}_{\rho\to\tau}\ x \ :\equiv\ \forall y^*, y(y^* \text{ s-maj}_\rho\ y \to x^* y^* \text{ s-maj}_\tau\ \mathbf{x}^*\mathbf{y}, xy.$$

The additional clause ensures that a (strong) majorant also majorizes itself. In [12], Bezem uses the strong majorization relation to show that there is a model of the bar-recursive functionals, the model $\mathcal{M}$ of hereditarily strongly majorizable functionals, that contains discontinuous functions. All the previously considered models of the bar-recursive functionals had been based on continuous functionals, such as e.g. the total continuous functionals of Kleene [61] and Kreisel[88].

The strong majorization relation (for the sake of brevity we will just write *majorization* for *strong majorization* for the rest of this chapter) is also very useful for proof mining purposes. Combining functional interpretation and modified realizability interpretation with the majorization relation one obtains variants of these interpretations called *monotone* functional interpretation, resp. monotone modified realizability. These monotone variants can be used to extract effective bounds rather than exact realizers (in some cases – to be discussed later – bounds actually are realizers as well). Monotone functional interpretation was first introduced by Kohlenbach in [68] and the corresponding monotone variant of modified realizability is discussed in detail in [70].

One important motivation for using the monotone variants of functional interpretation and modified realizability instead of the usual variants is that the extraction of bounds often is much simpler than the extraction of the corresponding exact realizers. A simple, yet compelling example is the monotone functional interpretation of the logical axiom $A \to A \wedge A$: The functional interpretation of this axiom requires us to produce certain realizers for the copy of A in the premise of the implication from the two realizers for A in the conclusion. In the usual (exact) functional interpretation these realizers can only be formed using characteristic terms for (the quantifier-free matrix of) A, which in turn depends on the decidability of all quantifier-free formulas. For the theory **WE-HA**$^\omega$, weakly extensional Heyting arithmetic in all finite types, which usually is the base theory of the theories to which functional interpretation is applied, such characteristic terms indeed exist for all quantifier-free formulas. However, for the *monotone* functional interpretation, to produce the necessary realizers for the instance of A in the premise of the implication we merely take the *maximum* over the two realizers for A in the conclusion. As an additional benefit to the simpler interpretation of the theory **WE-HA**$^\omega$ itself, we may now also add certain principles, such as e.g. weak König's lemma, to the theory that may not have a computable exact functional interpretation, but do have a computable *monotone* functional interpretation, resp. *monotone* modified realizability interpretation. For a full discussion of the metatheorems that one may prove for **WE-HA**$^\omega$ and closely related theories with this combination of functional interpretation, resp. modified realizability, and majorizability see e.g. [63].

3.1 Main results

We now give a general account of the results previously published in [42, 41]. In the following we will focus mainly on the results for classical theories based on metric spaces presented in [41], while relegating the discussion of normed linear spaces, as treated in [41], and corresponding metatheorems for semi-intuitionistic theories treated in [42] to the end of this chapter. As also shown in [39], one may, instead of a single metric or normed space, simultaneously treat tuples of spaces, product spaces, functions between product spaces, etc. This extension and the various case studies in fixed point theory treated in [41] will not be discussed here.

Definition 3.1. *The set* $\mathbf{T}$ *of all finite types is defined inductively over the ground type* 0 *by the clauses*

$$(i)\ 0 \in \mathbf{T},\ (ii)\ \rho, \tau \in \mathbf{T} \Rightarrow (\rho \to \tau) \in \mathbf{T}.$$

The base systems for our metatheorems are the formal systems $\mathcal{A}^\omega$ and $\mathcal{A}^\omega_i$ for classical, resp. intuitionistic, analysis in all finite types. These formal systems are defined in detail in [77, 42], but we repeat the most important characteristics here:

Classical analysis $\mathcal{A}^\omega$ is the formal system **WE-PA**$^\omega$ + **QF-AC** + **DC**: weakly extensional Peano arithmetic in all finite types plus the quantifier-free axiom of choice plus the axiom of dependent choice. The dual formal system $\mathcal{A}^\omega_i$, which is based on intuitionistic logic, is the formal system **E-HA**$^\omega$ + **AC**: fully extensional Heyting arithmetic in all finite types plus the full axiom of choice. In both systems, only equality $=_0$ between objects of type 0 is included, whereas higher type equality is a defined notion:

$$s =_\rho t :\equiv \forall x_1^{\rho_1}, \ldots, x_k^{\rho_k} (s(x_1, \ldots, x_k) =_0 t(x_1, \ldots, x_k)),$$

where $\rho = \rho_1 \to (\rho_2 \to (\ldots(\rho_k \to 0)\ldots))$, i.e. higher type equality is defined as extensional equality. Recall, that an object F of type $\rho \to \tau$ is called extensional, if it respects the extensional equality $\forall x^\rho, y^\rho \big(x =_\rho y \to F(x) =_\tau F(y)\big)$.

A crucial difference between the formal systems for classical analysis and intuitionistic analysis is the strength of the extensionality principle included in the theory. For the classical theories we aim at providing a (monotone) functional interpretation for the axioms and rules of the theory. As mentioned above, in [52] Howard has shown that, due to the fact that functional interpretation interprets the Markov principle, already for type 3 the full extensionality axiom does not have a functional interpretation even by set-theoretic functionals. Therefore, for classical analysis, which we interpret using functional interpretation, we restrict extensionality to the weak quantifier-free extensionality rule due to Spector[114]:

$$\text{QF-ER}: \frac{A_0 \to s =_\rho t}{A_0 \to r[s] =_\tau r[t]}, \text{ where } A_0 \text{ is a quantifier-free formula.}$$

This rule still allows to derive the extensionality for type 0 objects, i.e. $x =_0 y \to t[x] =_\tau t[y]$, but not for x and y of higher types (see [52, 117]). There is yet another reason to restrict extensionality to the weak extensionality rule in the extensions of classical analysis with abstract metric and real normed linear spaces, which we will discuss below. For intuitionistic analysis (and their extensions with metric and normed linear spaces) we usually aim at a modified realizability interpretation. For modified realizability interpretation (which does not interpret the Markov principle) extensionality is unproblematic and we may therefore add the full axiom of extensionality, i.e.

$$E_\rho : \forall z^\rho, x_1^{\rho_1}, y_1^{\rho_1}, \ldots, x_k^{\rho_k}, y_k^{\rho_k} (\bigwedge_{i=1}^{k} (x_i =_{\rho_i} y_i) \to z\underline{x} =_0 z\underline{y}),$$

for all types ρ.

Recall, that for proofs in classical analysis $\mathcal{A}^\omega$ (but not for the extended theories below) one can to some extent circumvent the restriction of extensionality to Spector's extensionality rule using Luckhardt's technique for the elimination of extensionality. For details on Luckhardt's result, see [93].

Before we sketch the extensions of $\mathcal{A}^\omega$ and $\mathcal{A}_i^\omega$ with abstract metric and real normed linear spaces (where we add the space X as a kind of 'Urelement'), we briefly recall the representation of real numbers in these theories, as presented in [41]. Real numbers are represented by Cauchy sequences $(a_n)_n$ of rational numbers with fixed Cauchy modulus 2^{-n}, i.e.

$$\forall m, n(n, m \geq k \to |a_m - a_n| < 2^{-k}).$$

Rational numbers are represented by pairs of natural numbers coded into a single natural number using the Cantor pairing function j. Every natural number is the code of a unique rational number, as we take $j(n, m)$ to represent $\frac{n/2}{m+1}$ if n is even and the negative number $-\frac{(n+1)/2}{m+1}$ if n is odd. All the usual relations and operators $=_\mathbb{Q}, <_\mathbb{Q}, \leq_\mathbb{Q}, +_\mathbb{Q}, -_\mathbb{Q}, \cdot_\mathbb{Q}$ can be defined primitive recursively in the corresponding relations for natural numbers. Thus, in our theories natural and rational numbers are represented by objects of type 0 and real numbers by number-theoretic functions f, i.e. objects of type $0 \to 0 (\equiv 1)$, satisfying the following:

$$\forall n (|f(n) -_\mathbb{Q} f(n+1)| <_\mathbb{Q} 2^{-n-1}), \ (*)$$

To make sure that every function $f : \mathbb{N} \to \mathbb{N}$ represents a real number as desired, we use the following construction:

$$\widehat{f}(n) := \begin{cases} f(n) \text{ if } \forall k < n \big(|f(k) -_\mathbb{Q} f(k+1)|_\mathbb{Q} <_\mathbb{Q} 2^{-k-1}\big), \\ f(k) \text{ for } \min k < n \text{ with } |f(k) -_\mathbb{Q} f(k+1)|_\mathbb{Q} \geq_\mathbb{Q} 2^{-k-1} \text{ otherwise.} \end{cases}$$

For readability, we usually write e.g. 2^{-n} instead of its code $\langle 2^{-n}\rangle := j(2, 2^n-1)$. The construction $\widehat{f}$ can be carried out in the theories $\mathcal{A}^\omega$ and $\mathcal{A}^\omega_i$. For every f the construction $\widehat{f}$ satisfies (*) and thus every f codes a unique real number, namely the one represented by $\widehat{f}$. Furthermore, if f already satisfies (*), then $\forall n(f(n) = \widehat{f}(n))$. Equipped with the construction $\widehat{f}$ we may reduce quantifiers over the real numbers to $\forall f^1$ and $\exists f^1$. Natural numbers (and in a similar way, rational numbers) can be embedded in the real numbers using the construction $(b)_{\mathbb{R}} = \lambda n.j(2b, 0)$.

The relations $=_{\mathbb{R}}, <_{\mathbb{R}}$ and $\leq_{\mathbb{R}}$ (on representatives of real numbers) are defined notions. The relations are expressible through respectively Π^0_1-predicates, in the case of $=_{\mathbb{R}}$ and $\leq_{\mathbb{R}}$, and a Σ^0_1-predicate, for $<_{\mathbb{R}}$. The usual operators $+_{\mathbb{R}}, -_{\mathbb{R}}, \cdot_{\mathbb{R}}$, etc. can be defined by primitive recursive functionals. For more details, see [77].

Closed, bounded intervals of the real numbers can be given a special representation by number-theoretic functions that are bounded by a fixed number-theoretic function M. Of particular importance for the applications presented later in this chapter is the interval $[0,1]$ which we give the following representation:

Definition 3.2.

$$\tilde{x}(n) := j(2k_0, 2^{n+2}-1),\ \textit{where}\ k_0 = \max k \leq 2^{n+2}[\frac{k}{2^{n+2}} \leq_{\mathbb{Q}} \widehat{x}(n+2)].$$

One easily verifies the following lemma:

Lemma 3.3. *Provably in $\mathcal{A}^\omega$, for all x^1:*

1. $0_{\mathbb{R}} \leq_{\mathbb{R}} x \leq_{\mathbb{R}} 1_{\mathbb{R}} \to \tilde{x} =_{\mathbb{R}} x$,
2. $0_{\mathbb{R}} \leq_{\mathbb{R}} \tilde{x} \leq_{\mathbb{R}} 1_{\mathbb{R}}$,
3. $\tilde{x} \leq_1 M := \lambda n.j(2^{n+3}, 2^{n+2}-1)$,
4. $x >_{\mathbb{R}} 1_{\mathbb{R}} \to \tilde{x} =_{\mathbb{R}} 1_{\mathbb{R}},\ x <_{\mathbb{R}} 0_{\mathbb{R}} \to \tilde{x} =_{\mathbb{R}} 0_{\mathbb{R}}$.

The extensions of $\mathcal{A}^\omega$ and $\mathcal{A}^\omega_i$ with a *bounded* (non-empty) abstract metric space (X,d), resp. hyperbolic space or CAT(0)-space (X,d,W), as well as the extensions with a (non-trivial) abstract real normed linear space $(X, \|\cdot\|)$ with a *bounded* convex subset C are described in detail in [77, 42]. The resulting (classical) theories are denoted by $\mathcal{A}^\omega[X,d]$ for bounded metric spaces, $\mathcal{A}^\omega[X,d,W]$ for bounded hyperbolic spaces and $\mathcal{A}^\omega[X,d,W,\mathrm{CAT}(0)]$ for bounded CAT(0)-spaces. The base theories for the corresponding semi-constructive variants of these theories are denoted by $\mathcal{A}^\omega_{\mathbf{i}}[\ldots]$, which then are extended with certain non-constructive principles such as comprehension for arbitrary negated formulas. The extensions of these theories with *unbounded* abstract metric spaces and abstract normed linear spaces with *unbounded* convex subsets are treated in [41]. The corresponding unbounded theories are denoted by $\mathcal{A}^\omega[\ldots]_{-\mathbf{b}}$ and $\mathcal{A}^\omega_i[\ldots]_{-\mathbf{b}}$. Here, the '$-b$' expresses the absence of an axiom expressing the boundedness of the space (by some integer bound b)[1]. For abstract normed linear spaces $(X, \|\cdot\|)$, possibly with a bounded, resp. unbounded, convex subset C, uniformly convex spaces (with a modulus of uniform convexity η) or inner product spaces (with the inner product $\langle\cdot,\cdot\rangle$) we write $[X, \|\cdot\|]$, $[X, \|\cdot\|, C]$, $[X, \|\cdot\|, C, \eta]$, etc.

We next summarize the most important points of extending classical and intuitionistic analysis with an abstract metric or normed linear space, illustrated by the case of extending $\mathcal{A}^\omega$ with a bounded, resp. unbounded, abstract metric space (X, d). The extensions with other variants of metric and normed linear spaces are treated in a similar way.

A preliminary step is to extend the set of all finite types with the new ground type X (corresponding to adding the space X as 'Urelement'):

Definition 3.4. *The set $\mathbf{T}^X$ of all finite types over the ground types 0 and X is defined inductively by the clauses*

$$(i)\ 0, X \in \mathbf{T}^X,\ (ii)\ \rho, \tau \in \mathbf{T}^X \Rightarrow (\rho \to \tau) \in \mathbf{T}^X.$$

For the actual extension, the main ideas are to (1) add a new ground type X corresponding to the metric space (X, d) to $\mathcal{A}^\omega$, (2) extend the usual axioms of $\mathcal{A}^\omega$ to the set of types $\mathbf{T}^X$ over the ground types 0 (natural numbers) and X, and (3) add the necessary constants and axioms for an abstract (pseudo-)metric space to $\mathcal{A}^\omega$. E.g. for abstract metric spaces (X, d) one adds a constant 0_X of type X, representing an arbitrary element of the space and thus asserting its non-emptiness, and a constant d_X of type $X \to X \to 1$ (recall that real numbers are represented by type 1 objects), representing the metric function of the space. In the case bounded metric spaces treated in [77], one also needs to add an axiom stating that the space is bounded. To interpret that boundedness-axiom one then also needs to add a constant b_X representing an integer bound on the metric d_X. The reason for limiting the axiomatization to one of pseudo-metric spaces is discussed in detail in [77]. Note, that equality for type 0 remains the only primitive equality predicate. Equality $x^X =_X y^X$ for the new type X is defined as $d_X(x, y) =_{\mathbb{R}} 0$, i.e. the only primitive formulas are still $s =_0 t$.

The fact that equality for the type X merely is a defined notion yields yet another reason to restrict extensionality to the weak extensionality rule in the classical case (where bounds are extracted using functional interpretation). Already for functions of type $X \to X$ full extensionality combined with the metatheorems to be proved below would allow one to derive false statements. Suppose, we prove (from the extensionality axiom) the full extensionality of all $f^{X\to X}$, i.e.

$$\forall f^{X\to X} \forall x^X, y^X (x =_X y \to f(x) =_X f(y)).$$

This can be rewritten as

$$\forall f^{X\to X} \forall x^X, y^X \forall k \in \mathbb{N} \exists n \in \mathbb{N} (d_X(x, y) \leq_{\mathbb{R}} 2^{-n} \to d_X(f(x), f(y)) \leq 2^{-k}).$$

[1] The notation'$-b$' may seem odd, but has a historical explanation: the metatheorems for the bounded case were proved first.

Here, the Markov principle is responsible for pulling out the universal quantifier hidden in the premise $x =_X y$ as an existential quantifier, whereas modified realizability would leave that premise untouched. Extracting a bound on $\exists n$ that is independent from x, y and f from a proof of this statements – as can be done for bounded metric spaces via monotone functional interpretation – then corresponds to obtaining a common modulus of uniform continuity for *all* f. It is in general not true in e.g. an abstract bounded metric that all functions $f : X \to X$ are uniformly continuous with a common modulus of uniform continuity. Therefore, for the (monotone) functional interpretation of these theories the restriction to Spector's extensionality rule is strictly necessary. However, for many classes of functions that one considers in functional analysis (and which also will be discussed below in the context of the metatheorems), such as nonexpansive functions or Lipschitz-continuous functions, this does not cause any problems, as such functions actually *are* provably uniformly continuous and hence their extensionality follows.

In order to extend the metatheorems for $\mathcal{A}^\omega$ to the extended theories $\mathcal{A}^\omega[X, d]$ (for bounded metric spaces) and $\mathcal{A}^\omega[X, d]_{-b}$ (for unbounded metric spaces, the $-b$ expressing the absence of a bound on the metric) one again needs to carry out three main steps: (1) extend the Howard-Bezem majorizability relation to the new type X, (2) check that the axioms and constants of the theory $\mathcal{A}^\omega$ extended to the new types $\mathbf{T}^X$ still have a monotone functional interpretation, resp. suitable majorants definable by closed terms and (3) show that the new metric space axioms and constants have a monotone functional interpretation, resp. have suitable majorants.

The crucial step is the extension of the majorization relation to the new type X and, related to that, the definition of suitable majorants for the metric space constants. In [77], where only the restricted cases of bounded metric spaces and normed linear spaces with bounded convex subsets are treated two different extensions of the majorization relation are employed. For bounded metric cases the extension to the new type X is defined as

$$x^* \text{ s-maj}_X \ x :\equiv (0 = 0),$$

i.e. the majorization relation is defined to be always true for the type X. This is possible since metric of the space is bounded by some integer bound b and hence the metric d_X can be majorized by the constant b-functional of suitable type. The only other constant involving the type X, namely 0_X majorizes itself. In the case of so-called hyperbolic spaces (X, d, W), one also needs to define a majorant for the hyperbolic function W_X which is of type $X \to X \to 1 \to X$, but here a constant 0_X-functional of suitable type suffices.

For real normed linear spaces such an easy approach does not work, as non-trivial normed linear spaces always are unbounded. Here instead one defines the majorization relation for the new type X on terms of the norm $\|\cdot\|_X$:

$$x^* \text{ s-maj}_X \ x :\equiv \|x^*\|_X \geq_{\mathbb{R}} \|x\|_X.$$

The norm $\|\cdot\|_X$ (of type $X \to 1$) is then self-majorizing and most of the remaining constants of normed linear spaces have almost trivial majorants that are definable by simple closed terms. Only the constant $\cdot_X$ of type $1 \to X \to X$ for scalar multiplication (and to some extent also the norm $\|\cdot\|_X$) uses some special properties of the chosen representation for real numbers through number theoretic functions.

In [41], a generalized approach to majorization in the new type X, so called a-majorization, is presented. This new approach is derived from (and inspired by) the treatment of majorization and bound extraction for normed linear spaces in [77] via two main steps. The first (intermediate) step is to pick an arbitrary element $a \in X$ as a point of reference and define majorization in the type X in terms of an object's metric distance to the point of reference a, i.e.

$$\begin{array}{lll} x^* \text{ s-maj}_X^a\ x & :\equiv d(x^*, a) \geq_{\mathbb{R}} d(x, a) & \text{for metric spaces } (X, d) \\ x^* \text{ s-maj}_X^a\ x & :\equiv \|x^* - a\| \geq_{\mathbb{R}} \|x - a\| & \text{for normed linear spaces } (X, \|\cdot\|) \end{array}$$

Next, in [77], a relation $\widehat{\cdot}$ between types $\rho \in \mathbf{T}^X$ and $\widehat{\rho} \in \mathbf{T}$ and a relation $\sim_\rho$ between functionals of type ρ and $\widehat{\rho}$ are introduced in order to eliminate the dependency of the extracted bounds on the underlying abstract metric or normed linear spaces. As we are only interested in the numerical bounds (no longer involving the type X) resulting from transforming the extracted bounds with the $\sim_\rho$-relation anyway, this step can be built in directly into the new a-majorization relation. Thus, choosing natural numbers as the domain for majorants of elements of type X is the second important idea of the generalization of the majorization relation developed in [41]. In [77], the definition of the mapping $\widehat{\cdot}$ depends on whether we consider metric spaces or normed linear spaces, mapping the type X to the type 0 or the type 1 respectively. In [41], the type X is always mapped to type 0:

Definition 3.5. *For $\rho \in \mathbf{T}^X$ we define $\widehat{\rho} \in \mathbf{T}$ inductively as follows*

$$\widehat{0} := 0,\ \widehat{X} := 0,\ \widehat{(\rho \to \tau)} := (\widehat{\rho} \to \widehat{\tau}),$$

i.e. $\widehat{\rho}$ is the result of replacing all occurrences of the type X in ρ by the type 0.

We may then define the following so-called a-majorization relation $\gtrsim_\rho^a$ between objects of type X, $\rho \in \mathbf{T}^X$ and $\widehat{\rho} \in \mathbf{T}$:

Definition 3.6. *The ternary relation $\gtrsim_\rho^a$ between objects x, y and a of type $\widehat{\rho}, \rho$ and X respectively by induction on ρ as follows:*

- $x^0 \gtrsim_0^a y^0 :\equiv x \geq_{\mathbb{N}} y,$
- $x^0 \gtrsim_X^a y^X :\equiv (x)_{\mathbb{R}} \geq_{\mathbb{R}} d_X(y, a),$
- $x \gtrsim_{\rho\to\tau}^a y :\equiv \forall z', z(z' \gtrsim_\rho^a z \to xz' \gtrsim_\tau^a yz) \wedge \forall z', z(z' \gtrsim_{\widehat{\rho}}^a z \to xz' \gtrsim_{\widehat{\tau}}^a xz).$

For normed linear spaces we usually choose $a = 0_X$, so that $d_X(x, a) =_{\mathbb{R}} \|x\|_X$.

Note, that majorants are always of some type $\rho \in \mathbf{T}$. If we restrict the relation $\gtrsim^a$ to the types $\mathbf{T}$ it is identical to the Howard-Bezem notion of strong majorizability s-maj and hence we freely write s-maj$_\rho$ instead of $\gtrsim_\rho^a$ for $\rho \in \mathbf{T}$.

The choice of 0_X for a in the normed linear case is by no means mandatory. While one can show that it is independent of the choice of a whether a given functional is a-majorizable or not, the choice of a is crucial to obtain “nice” majorants. For normed linear spaces, if one chooses $a = 0_X$, the new majorization relation is very similar to the majorization relation for normed linear spaces defined in [77]. For any other choice of a the majorants for the constants of normed linear spaces and hence also the extracted bounds will depend heavily on (an integer bound on) $\|a - 0_X\|$. For metric spaces, all but one constant of metric (and also hyperbolic) spaces can be majorized by a suitable closed term of $\mathcal{A}^\omega$ and only the constant 0_X asserting the non-emptiness of the space needs some special care. For the a-majorant of 0_X one needs an integer bound on $d_X(0_X, a)$, but in certain cases even that requirement can be eliminated. We will discuss the details later.

A second important issue is ensuring that the new, additional axioms of the theories $\mathcal{A}^\omega[X, d]$, $\mathcal{A}^\omega[X, d, W]$, etc., as well as the usual axioms of $\mathcal{A}^\omega$ extended to all types $\mathbf{T}^X$ have a monotone functional interpretation. The latter part of this is easily checked. For the new axioms the crucial insight is that the class of abstract (pseudo-)metric, resp. real normed linear, spaces can be axiomatized by purely universal formulas not containing $\vee$. As an example, for metric spaces ((1)-(3)), hyperbolic spaces ((1)-(7)) and CAT(0)-spaces ((1)-(8)) the additional axioms given in [41] are:

(1) $\forall x^X (d_X(x, x) =_{\mathbb{R}} 0_{\mathbb{R}})$,

(2) $\forall x^X, y^X \big(d_X(x, y) =_{\mathbb{R}} d_X(y, x)\big)$,

(3) $\forall x^X, y^X, z^X \big(d_X(x, z) \le_{\mathbb{R}} d_X(x, y) +_{\mathbb{R}} d_X(y, z)\big)$,

(4) $\forall x^X, y^X, z^X \forall \lambda^1 \big(d_X(z, W_X(x, y, \lambda)) \le_{\mathbb{R}} (1_{\mathbb{R}} -_{\mathbb{R}} \tilde{\lambda}) d_X(z, x) +_{\mathbb{R}} \tilde{\lambda} d_X(z, y)\big)$,

(5) $\forall x^X, y^X \forall \lambda_1^1, \lambda_2^1 \big(d_X(W_X(x, y, \lambda_1), W_X(x, y, \lambda_2)) =_{\mathbb{R}} |\tilde{\lambda}_1 -_{\mathbb{R}} \tilde{\lambda}_2|_{\mathbb{R}} \cdot_{\mathbb{R}} d_X(x, y)\big)$,

(6) $\forall x^X, y^X \forall \lambda^1 \big(W_X(x, y, \lambda) =_X W_X(y, x, (1_{\mathbb{R}} -_{\mathbb{R}} \lambda))\big)$,

(7) $\left\{ \begin{array}{l} \forall x^X, y^X, z^X, w^X, \lambda^1 \\ \big(d_X(W_X(x, z, \lambda), W_X(y, w, \lambda)) \le_{\mathbb{R}} (1_{\mathbb{R}} -_{\mathbb{R}} \tilde{\lambda}) d_X(x, y) +_{\mathbb{R}} \tilde{\lambda} d_X(z, w)\big). \end{array} \right.$

(8) $\forall x^X, y_1^X, y_2^X \big(d_X(x, W_X(y_1, y_2, \frac{1}{2}))^2 \le_{\mathbb{R}} \frac{1}{2} d_X(x, y_1)^2 +_{\mathbb{R}} \frac{1}{2} d_X(x, y_2)^2 -_{\mathbb{R}} \frac{1}{4} d_X(y_1, y_2)^2\big)$.

Such purely universal formulas not containing $\vee$ are their own (monotone) functional interpretation. For the additional axioms for hyperbolic spaces (X, d, W) concerning the function W there is a subtle change relative to the more complicated formulation given in [77]. The (syntactical) function W_X is of type $X \to X \to 1 \to X$, i.e. it takes two elements of type X and a type 1 representative of a real number (which only potentially represents a real number in the interval $[0, 1]$) as arguments and returns a new element of type X. In contrast, the actual function W (in its intended interpretation in the model) is only well-defined for real numbers in the interval $[0, 1]$. Therefore in [41] the axiomatization of the properties of the function W_X was changed so that – using in particular axiom (5) – it implicitly ensures that

$$\forall x^X, y^X, \lambda^1 (W_X(x, y, \lambda) = W_X(x, y, \tilde{\lambda})),$$

i.e. the function W_X behaves as if every type 1 functional given as an input were a real number in the interval $[0, 1]$. (Recall, that the construction $\tilde{f}$ turns a given function f of type 1 into a corresponding function representing a real number in the interval $[0, 1]$.) In [77], this was only ensured through a suitable interpretation of W_X in the model, instead of the above, purely syntactical construction in the axioms. Moreover, the original formulation in [77] of the **CN**-inequality that characterizes CAT(0)-spaces was also simplified in [41] to the (equivalent) form given above in axiom (8).

In conclusion, using monotone functional interpretation one may, from a given formal proof in the extended theories sketched above, extract computable bounds. Depending on the underlying space and on certain additional premises, these bounds may display strong uniformities. If the underlying space is a bounded metric space (X, d), the extracted bounds will only depend on parameters ranging over the space via an integer bound b on the metric d, *even if the space is not compact*. The result holds for bounded hyperbolic and CAT(0)-spaces as well. Even in an unbounded metric space one may obtain similar uniformities with regard to input parameters ranging over the space, as soon as certain very liberal local boundedness conditions are satisfied.

To verify that these bounds actually hold in e.g. a given arbitrary metric space (X, d), one needs to verify that the full set-theoretic type structure $\mathcal{S}^{\omega,X}$ over $\mathbb{N}, X$, where X is metric space (X, d) or a normed linear space $(X, \|\cdot\|)$. As some arguments concerning the a-majorization relation only hold in the type structure $\mathcal{M}^{\omega,X}$ of strongly hereditarily a-majorizable functionals (which itself is independent of the choice of a!), one needs to place some restrictions on the types occurring in the theorem for which one wants to extract computable bounds. For suitably low types one can show that a bound that is valid in the type structure of a-majorizable functionals also is valid in the corresponding full set-theoretic type structure.

After this general discussion, we present the formal results, starting out with a number of definitions. As mentioned above, we focus on the classical metatheorems for the (unbounded) metric, hyperbolic and CAT(0)-case, corresponding to the formal theories $\mathcal{A}^\omega[X, d]_{-b}$,$\mathcal{A}^\omega[X, d, W]_{-b}$ and $\mathcal{A}^\omega[X, d, W, \mathrm{CAT}(0)]_{-b}$.

We define the following type abbreviations:

Definition 3.7 ([77, 42])**.** *We say that a type $\rho \in \mathbf{T}^X$ has degree*

- 1 *if $\rho = 0 \to \ldots \to 0$ (including $\rho = 0$),*
- $(0, X)$ *if $\rho = 0 \to \ldots \to 0 \to X$ (including $\rho = X$),*
- $(1, X)$ *if it has the form $\tau_1 \to \ldots \to \tau_k \to X$ (including $\rho = X$), where τ_i has degree* 1 *or* $(0, X)$*,*

- $(\cdot, 0)$ *if* $\rho = \tau_1 \to \ldots \to \tau_k \to 0$ *(including* $\rho = 0$*) for arbitrary types* $\tau_i \in \mathbf{T}^X$*,*
- $(\cdot, X)$ *if* $\rho = \tau_1 \to \ldots \to \tau_k \to X$ *(including* $\rho = X$*) for arbitrary types* $\tau_i \in \mathbf{T}^X$*.*

Definition 3.8 ([41])**.** *We say that a type* $\rho \in \mathbf{T}^X$ *has degree* $\widehat{1}$*, if* $\widehat{\rho}$ *has degree* 1*. Amongst others, the type degree* $\widehat{1}$ *covers types* $\mathbb{N}, X, \mathbb{N} \to \mathbb{N}, \mathbb{N} \to X, X \to \mathbb{N}$ *and* $X \to X$*.*

Definition 3.9 ([41])**.** *A formula* F *is called a* $\forall$*-formula (resp.* $\exists$*-formula) if it has the form* $F \equiv \forall \underline{a}^{\underline{\sigma}} F_{qf}(\underline{a})$ *(resp.* $F \equiv \exists \underline{a}^{\underline{\sigma}} F_{qf}(\underline{a})$*) where* F_{qf} *does not contain any quantifiers and the types in* $\underline{\sigma}$ *are of degree* $\widehat{1}$ *or* $(1, X)$*.*

The $()_\circ$-operator is defined as follows:

Definition 3.10 ([77])**.** *For* $x \in [0, \infty)$ *define* $(x)_\circ \in \mathbb{N}^{\mathbb{N}}$ *by*

$$(x)_\circ(n) := j(2k_0, 2^{n+1} - 1),$$

where

$$k_0 := \max k \Big[\frac{k}{2^{n+1}} \le x \Big].$$

The $()_\circ$-operator is used primarily as a semantic operator applied to actual real numbers rather than type 1 representatives of real numbers. This operator is used to ensure certain properties of real numbers resulting from the interpretation of e.g $d_X(x, y)$ in the model. A given real number may have many representations by number-theoretic functions. The $()_\circ$-operator selects for each actual real number a canonical representative that has certain properties beneficial for the later majorization process. The operator has a syntactical counterpart $()_\circ : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ on type 1 representatives of real numbers, which is not computable in general although it is majorizable: simply take the integer code of the first rational approximation to the given real number and add 1, then this is an integer upper bound. In our bounds this latter version of the $()_\circ$-operator will only used in the form $\lambda n^0.((n)_{\mathbb{R}})_\circ$ where it is primitive recursively computable, i.e.

$$(m)_\circ(n) := j(2m2^{n+1}, 2^{n+1} - 1), \text{ for } m \in \mathbb{N}.$$

We will use the following properties of the $()_\circ$-operator:

Lemma 3.11 ([77])**.**
1. *If* $x \in [0, \infty)$*, then* $(x)_\circ$ *is a representative of* x *in the sense of the representation of real numbers described previously in this chapter.*
2. *If* $x, y \in [0, \infty)$ *and* $x \le y$ *(in the sense of* $\mathbb{R}$*), then* $(x)_\circ \le_{\mathbb{R}} (y)_\circ$ *and also* $(x)_\circ \le_1 (y)_\circ$ *(i.e.* $\forall n \in \mathbb{N}((x)_\circ(n) \le (y)_\circ(n))$*).*
3. *If* $x \in [0, \infty)$*, then* $(x)_\circ$ *is monotone, i.e.* $\forall n \in \mathbb{N}((x)_\circ(n) \le_0 (x)_\circ(n+1))$*.*
4. *If* $x, y \in [0, \infty)$ *and* $x \le y$ *(in the sense of* $\mathbb{R}$*), then* $(y)_\circ$ *s-maj*$_1 (x)_\circ$

Definition 3.12. *Let X be a nonempty set. The full set-theoretic type structure $\mathcal{S}^{\omega,X} := \langle S_\rho \rangle_{\rho\in\mathbf{T}^X}$ over $\mathbb{N}$ and X is defined by*

$$S_0 := \mathbb{N}, \quad S_X := X, \quad S_{\rho\to\tau} := S_\tau^{S_\rho}.$$

Here $S_\tau^{S_\rho}$ is the set of all set-theoretic functions $S_\rho \to S_\tau$.

Using this and the $()_\circ$-operator we state the following definition:

Definition 3.13 ([41]). *We say that a sentence of $\mathcal{L}(\mathcal{A}^\omega[X,d,W]_{-b})$ holds in a nonempty hyperbolic space (X,d,W) if it holds in the models[2] of $\mathcal{A}^\omega[X,d,W]_{-b}$ obtained by letting the variables range over the appropriate universes of the full set-theoretic type structure $\mathcal{S}^{\omega,X}$ with the set X as the universe for the base type X, 0_X is interpreted by an arbitrary element of X, $W_X(x,y,\lambda^1)$ is interpreted as $W(x,y,r_{\tilde{\lambda}})$, where $r_{\tilde{\lambda}} \in [0,1]$ is the unique real number represented by $\tilde{\lambda}^1$ and d_X is interpreted as $d_X(x,y) :=_1 (d(x,y))_\circ$.*

Finally, we define the following functional, which is particularly useful for defining majorants for functionals of degree 1.

Definition 3.14 ([77]). *For types $0 \to \rho$ with $\rho = \rho_1 \to \ldots \to \rho_k \to 0$, we define functionals $(\cdot)^M$ of types $(0 \to \rho) \to 0 \to \rho$ by :*

$$x^M(y^0) := \lambda \underline{v}^{\underline{\rho}}.\max{}_0\{x(i,\underline{v}) \mid i = 1,\ldots,y\}.$$

We are now in a position to state the main version of the metatheorem for *unbounded* metric, hyperbolic and CAT(0)-spaces presented in [41]:

Theorem 3.15 ([41]).

1. *Let ρ be of degree $(1,X)$ or 2 and let $B_\forall(x,u)$, resp. $C_\exists(x,v)$, contain only x,u free, resp. x,v free. Assume that the constant 0_X does not occur in $B_\forall, C_\exists$ and that*

   $$\mathcal{A}^\omega[X,d]_{-b} \vdash \forall x^\rho(\forall u^0 B_\forall(x,u) \to \exists v^0 C_\exists(x,v)).$$

   Then there exists a partial functional $\Phi : S_{\widehat{\rho}} \mapsto \mathbb{N}$ whose restriction to the strongly majorizable elements of $S_{\widehat{\rho}}$ is totally computable functional of $\mathcal{M}^{\omega}$[3] and the following holds in all nonempty metric spaces (X,d): for all $x \in S_\rho$, $x^* \in S_{\widehat{\rho}}$ if there exists an $a \in X$ s.t. $x^* \gtrsim^a x$ then[4]

   $$\forall u \le \Phi(x^*) B_\forall(x,u) \to \exists v \le \Phi(x^*) C_\exists(x,v).$$

   In particular, if ρ is in addition of degree $\widehat{1}$, then $\Phi : S_{\widehat{\rho}} \times \mathbb{N} \to \mathbb{N}$ is totally computable.

   If 0_x does occur in $B_\forall$ and/or $C_\exists$, then the bound Φ depends (in addition to x^*) on an upper bound $\mathbb{N} \ni n \ge d(0_X,a)$.

[2] We use here the plural since the interpretation of 0_X is not uniquely determined.

[3] In the sense of [62] relativized to $\mathcal{M}^\omega$.

[4] Note that $x^* \gtrsim^a x$ implies that x^* s-maj$_{\widehat{\rho}}$ x^* and hence the strong majorizability of x^* so that $\Phi(x^*)$ is defined.

2. *The theorem also holds for nonempty hyperbolic spaces $\mathcal{A}^\omega[X,d,W]_{-b}$, (X,d,W) and for $\mathcal{A}^\omega[X,d,W,\mathrm{CAT}(0)]_{-b}$ where (X,d,W) is a CAT(0) space.*

Instead of single variables x,u,v and single premises $\forall u B_\forall(x,u)$ we may have tuples of variables and finite conjunctions of premises. In the case of a tuple $\underline{x}$ we then have to require that we have a tuple $\underline{x}^*$ of a-majorants for a common $a \in X$ for all the components of the tuple $\underline{x}$.

Remark ([41]). *Another way to treat parameters x^ρ, ρ of degree $(1,X)$ or 2 is to require for a majorant a computable functional t in $S_{\underline{\sigma}} \to S_{\widehat{\rho}}$,[5] where all σ_i are of degree 1. Then we may obtain a totally computable $\Phi : S_{\underline{\sigma}} \to \mathbb{N}$ such that given $\underline{c} \in S_{\underline{\sigma}}$, if there exists an $a \in X$ for which $t(\underline{c}) \gtrsim^a_\rho x$ then the bound $\Phi(\underline{c})$ holds.*

The restriction on the types of degree $(1,X)$ or 2 is made necessary by the interpretation of dependent choice using bar recursive functionals. If a given proof does not use dependent choice, we can allow arbitrary types ρ in the parameters (with majorants of type $\widehat{\rho}$).

Remark ([41]). *From the proof of Theorem 3.15 (see [41]) two further extensions follow:*

1. *The language may be extended by a-majorizable constants (in particular constants of types 0 and 1, which always are uniformly majorizable) where the extracted bounds then additionally depend on (a-majorants for) the new constants.*

2. *The theory may be extended by purely universal axioms or, alternatively, axioms which can be reformulated into purely universal axioms using new majorizable constants if the types of the quantifiers are all of degree 2 or $(1,X)$,[6] as purely universal axioms are their own functional interpretation. Again the extracted bounds depend on (a-majorants for) these new constants. Then the conclusion holds in all metric (X,d) resp. hyperbolic (X,d,W) spaces which satisfy these axioms (under a suitable interpretation of the new constants if any).*

The proof of Theorem 3.15 centers around two lemmas (for the full details of the proof see Section 9 in [41]). The lemmas are stated for hyperbolic spaces (X,d,W). From these the metric case follows by omitting the axioms and constants for the function W, the CAT(0)-case by including another purely universal axiom, which however has no impact on the proof.

Let $\mathcal{A}^\omega[X,d,W]^-_{-b} := \mathcal{A}^\omega[X,d,W]_{-b} \setminus \{\text{QF-AC}\}$.

Lemma 3.16 ([77]). *Let A be a sentence in the language of $\mathcal{A}^\omega[X,d,W]_{-b}$.*

[5]Since t is of degree 2, the computability of t implies its (strong) majorizability.

[6]This ensures that validity in $\mathcal{S}^{\omega,X}$ implies validity in $\mathcal{M}^{\omega,X}$ defined further below.

Then the following rule holds:

$$\begin{cases} \mathcal{A}^\omega[X,d,W]_{-b} \vdash A \\ \Rightarrow \text{one can construct a tuple of closed terms } \underline{t} \text{ of } \mathcal{A}^\omega[X,d,W]_{-b}\text{+(BR) s.t.} \\ \mathcal{A}^\omega[X,d,W]^-_{-b} + \text{(BR)} \vdash \forall \underline{y}\,(A')_D(\underline{t},\underline{y}). \end{cases}$$

where A' is the negative translation of A and $(A')^D \equiv \exists\underline{x}\forall\underline{y}(A')_D(\underline{x},\underline{y})$ is the Gödel functional interpretation of A'.

This is Lemma 4.4 in [77], although strictly speaking, the above lemma has been slightly modified in [41], as there is one less purely universal axiom to interpret – the axiom that the whole space (X,d) is bounded – and the axioms for the function W have been slightly reformulated. However, these minor changes do not significantly influence the proof given in [77]. It was also pointed out in [41] that by an oversight in [77] it was forgotten to state that one can QF-AC from the theory in which the extracted realizer is verified as is obvious from the fact that functional interpretation eliminates QF-AC. This fact is actually needed because QF-AC does not hold in the intended model of majorizable functionals.

With this first lemma, we may argue that from a proof of

$$\forall x^\rho(\forall u^0 B_\forall(x,u) \to \exists v^0 C_\exists(x,v)).$$

in $\mathcal{A}^\omega[X,d,W]_{-b}$ we can extract exact realizers t_U and t_V in $\mathcal{A}^\omega[X,d,W]^-_{-b}+$ (BR) which given an x produce the necessary witnesses for this theorem. However theses functional realizers may still depend e.g. on the metric d_X of the *abstract* metric space (X,d), and thus these realizers are in general *not* computable.

As the next lemma shows, we may eliminate the dependency on the space (X,d) through majorization:

Lemma 3.17 ([41])**.** *Let (X,d,W) be a nonempty hyperbolic space. Then $\mathcal{M}^{\omega,X}$ is a model of $\mathcal{A}^\omega[X,d,W]^-_{-b}$ +(BR) (for a suitable interpretation of the constants of $\mathcal{A}^\omega[X,d,W]^-_{-b}$ +(BR) in $\mathcal{M}^{\omega,X}$), where we may interpret 0_X by an arbitrary element $a \in X$.*

Moreover, for any closed term t of $\mathcal{A}^\omega[X,d,W]^-_{-b}$ +(BR) one can construct a closed term t^* of $\mathcal{A}^\omega$ +(BR) – so in particular t^* does not contain the constants $0_X, d_X$ and W_X nor any other constant involving the type X – such that

$$\mathcal{M}^{\omega,X} \models \forall a^X \forall n^0((n)_{\mathbb{R}} \geq d(0_X,a) \to t^*(n) \gtrsim^a t).$$

In particular, if we interpret 0_X by $a \in X$, then it holds in $\mathcal{M}^{\omega,X}$ that $t^*(0^0)$ is an a-majorant of t

Here, $\mathcal{M}^{\omega,X}$ is the extensional type structure of all hereditarily strongly a-majorizable set-theoretic functionals of type $\rho \in \mathbf{T}^X$ over $\mathbb{N}$ and X. (Note, that the structure of a-majorizable functionals itself is independent of the choice of a, as one can show that for any $a, b \in X$, if a functional is a-majorizable it is also

b-majorizable). This lemma is the central part of the proof. By constructing a-majorants $t^*(n)$ for each closed term t we eliminate the dependency of the extracted bounds on the underlying metric, hyperbolic or CAT(0)-space – except for the dependency on an integer bound n on $d_X(0_X, a)$.

To verify this we must provide a-majorants for the constants of $\mathcal{A}^\omega[X, d, W]^-_{-b}$ +(BR). For the constants of classical analysis $\mathcal{A}^\omega$, taken over the extended set of types $\mathbf{T}^X$, only the majorant for the bar-recursor B requires extra care (see section 9 in [41] for details on the majorization of the bar-recursor B).

The remaining constants of $\mathcal{A}^\omega[X, d, W]^-_{-b}$ can be a-majorized as follows:

- $n^0 \gtrsim^a 0_X$ for every n with $(n)_{\mathbb{R}} \geq_{\mathbb{R}} d_X(a, 0_X)$, where as just mentioned we can take $n := 0$ if we interpret 0_X by a,
- $0^0 \gtrsim^a a$, since $d_X(a, a) =_{\mathbb{R}} (0)_{\mathbb{R}}$,
- $\lambda x^0, y^0.((x + y)_{\mathbb{R}})_\circ \gtrsim^a d_X^{X \to X \to 1}$,
- $\lambda x^0, y^0, z^1. \max_0(x, y) \gtrsim^a W_X^{X \to X \to 1 \to X}$.

The verification of these majorants is straightforward. As an example of one of the more complicated cases, consider the constant d_X. The majorization of d_X additionally uses the properties of the $()_\circ$-operator as discussed above. Let $n_1 \gtrsim^a x$ and $n_2 \gtrsim^a y$ then using the triangle inequality we get

$$d(x, y) \leq d(x, a) + d(y, a) \leq n_1 + n_2.$$

In the model $\mathcal{M}^{\omega,X}$ the expression $d_X(x, y)$ is interpreted by $(d(x, y))_\circ$, so by Lemma 3.11 from $n_1 + n_2 \geq d(x, y)$ we obtain $((n_1 + n_2)_{\mathbb{R}})_\circ$ s-maj$_1 d(x, y))_\circ$ and the validity of the given majorant follows.

Thus using this lemma, from realizers t_U and t_V and an a-majorant x^* for x, we obtain a-majorants $t_{U^*}(n, x^*)$ and $t_{V^*}(n, x^*)$ for $t_U(x)$ and $t_V(x)$ respectively, where n is a bound on $d_X(0_X, a)$. Defining $\Phi(x^{\widehat{\rho}}, n) := \max(t_{U^*}(n, x), t_{V^*}(n, x))$ we obtain a common majorant for both $t_U(x)$ and $t_V(x)$.

However, with the present arguments the extracted bound Φ is only valid in the type structure $\mathcal{M}^{\omega,X}$. Thus the final step is to show that for the types ρ occurring in the formula for which we want to extract a bound, either $M_\rho = S_\rho$ or at least $M_\rho \subseteq S_\rho$, such that the extracted bound also holds in the full set-theoretic type structure $\mathcal{S}^{\omega,X}$. For the type ρ of the parameter x we explicitly stated that the bound Φ is only valid on the a-majorizable elements of S_ρ. Similarly, the types γ hidden in the quantifiers of the definition of $\forall$-formulas and $\exists$-formulas, either type $\widehat{1}$ or type $(1, X)$, at least satisfy $M_\gamma \subseteq S_\gamma$.

In conclusion, we have thus shown that

$$\mathcal{S}^{\omega,X} \models \forall u \leq \Phi(x^*, n) B_\forall(x, u) \to \exists v \leq \Phi(x^*, n) C_\exists(x, v)$$

holds for all $n \in \mathbb{N}$, $x \in S_\rho$ and $x^* \in S_{\widehat{\rho}}$ for which there exists an $a \in X$ such that $n^0 \geq d(0_X, a)$ and $x^* \gtrsim^a x$.

As stated in the theorem, the resulting functional Φ does not depend on the underlying metric, hyperbolic or CAT(0)-space. Furthermore, one can observe that if the constant 0_X does not occur in either $B_\forall$ or $C_\exists$ we may freely interpret 0_X by $a \in X$ and thus also eliminate the dependency on a bound n on $d_X(0_X, a)$. This concludes the informal presentation of the proof of Theorem 3.15.

From this very general metatheorem we may prove several interesting corollaries. For starters, Theorem 3.7 in [77], the main metatheorem for the metric case in that paper, follows (by Remark 3.1) as an easy corollary from the proof of Theorem 3.15, as we only have to treat an additional purely universal axiom which expresses that the space is bounded. The first corollary we present in detail is tailored towards concrete applications in functional analysis.

We briefly repeat the definition of the following functional, which is particularly useful for defining majorants for functionals of degree 1:

Definition 3.18 ([77]). *For types $0 \to \rho$ with $\rho = \rho_1 \to \ldots \to \rho_k \to 0$, we define functionals $(\cdot)^M$ of types $(0 \to \rho) \to 0 \to \rho$ by :*

$$x^M(y^0) := \lambda \underline{v}^{\underline{\rho}}. \max{}_0\{x(i, \underline{v}) \mid i = 1, \ldots, y\}.$$

For a given object x of type 1 the construction x^M is then a majorant for x.

Corollary 3.19. *Let P (resp. K) be a $\mathcal{A}^\omega$-definable Polish space*[7] *(resp. compact Polish space), let τ be of degree $\widehat{1}$ and let $B_\forall$, resp. $C_\exists$, contain only x, y, z, u free, resp. x, y, z, v free, where furthermore 0_X does not occur in $B_\forall, C_\exists$. If*

$$\mathcal{A}^\omega[X, d, W]_{-b} \vdash \forall x \in P \forall y \in K \forall z^\tau \big(\forall u^0 B_\forall \to \exists v^0 C_\exists\big),$$

then there exists a computable functional $\Phi : \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{(\mathbb{N} \times \ldots \times \mathbb{N})} \to \mathbb{N}$ s.t. the following holds in every nonempty hyperbolic space (X, d, W): for all representatives $r_x \in \mathbb{N}^{\mathbb{N}}$ of $x \in P$ and all $z^* \in \mathbb{N}^{(\mathbb{N} \times \ldots \times \mathbb{N})}$ if there exists an $a \in X$ for which $z^* \gtrsim^a_\tau z$ then

$$\forall y \in K \big(\forall u \leq \Phi(r_x, z^*) B_\forall \to \exists v \leq \Phi(r_x, z^*) C_\exists\big).$$

As before, instead of single variables x, y, z and a single premise $\forall u^0 B_\forall$, we may have tuples of variables and a finite conjunction of premises.
Analogously, for $\mathcal{A}^\omega[X, d]_{-b}$, where (X, d) is an arbitrary noempty metric space, and $\mathcal{A}^\omega[X, d, W, \mathrm{CAT}(0)]_{-b}$, where then (X, d, W) is an arbitrary nonempty CAT(0)-space.

This corollary is proved easily, observing that quantification over (representatives of elements in) Polish spaces, respectively compact Polish spaces, can be expressed by quantification over all x^1, resp. all $y^1 \leq s$ for some closed function term s. The bound then depends on a type 1 representative r_x of x via its majorant $(r_x)^M$, but on y only via a majorant for the closed term s, which can be constructed by induction on the structure of s. The extraction of a bound

[7] For details on this see [77] and [66].

$\Phi(r_x, z^*)$ (which has the majorant s^M for s built in) as stated in the corollary then follows from Theorem 3.15.

The next corollary illustrates the true strength and generality of the new a-majorization relation. We first present a number of general classes of functions $f : X \to X$ and later in the corollary show them to be a-majorizable by simple, elementary closed terms of $\mathcal{A}^\omega$.

Definition 3.20. *A function $f : X \to X$ on a metric space (X, d) is called*

- *nonexpansive ('f n.e.') if $d(f(x), f(y)) \leq d(x, y)$ for all $x, y \in X$,*
- *quasi-nonexpansive if $\forall p, x \in X(d(p, f(p)) = 0 \to d(f(x), f(p)) \leq d(x, p))$,*
- *weakly quasi-nonexpansive if*

$$\exists p \in X(d(p, f(p)) = 0 \wedge \forall x \in X(d(f(x), f(p)) \leq d(x, p)))$$

  or – equivalently –

$$\exists p \in X \forall x \in X(d(f(x), p) \leq d(x, p)).$$

- *Lipschitz continuous if $d(f(x), f(y)) \leq L \cdot d(x, y)$ for some $L > 0$ and for all $x, y \in X$,*
- *Hölder-Lipschitz continuous if $d(f(x), f(y)) \leq L \cdot d(x, y)^\alpha$ for some $L > 0$, $0 < \alpha \leq 1$ and for all $x, y \in X$.*

For normed linear spaces $(X, \|\cdot\|)$ those definitions are to be understood w.r.t. the induced metric $d(x, y) := \|x - y\|$.

Note, that from the fact that the relations $\leq_{\mathbb{R}}$ and $=_{\mathbb{R}}$ are expressible by Π^0_1-statements one sees that 'f is n.e.', 'f is Lipschitz continuous with constant L' and 'f is Hölder-Lipschitz continuous with constants L and α' may be written as purely universal formulas and hence are admissible as premises for our metatheorems. For Lipschitz and Hölder-Lipschitz the additional constants are assumed to be given as parameters. The statement 'f is quasi-nonexpansive' is not of a suitable form to function as a premise, as it is a $\forall \to \forall$-formula, which would prenex to a $\forall\exists$-formula. For many practical applications it has turned out that the slightly weaker premise 'f is weakly quasi-nonexpansive' actually suffices (see [78] for an example). This weaker statement can be written as purely universal formula if we take the fixed point p as an additional parameter over which we quantify and for which we then also need to find a suitable majorant.

Corollary 3.21 ([41])**.** *1. Let P (resp. K) be a $\mathcal{A}^\omega$-definable Polish space (resp. compact Polish space). Assume we can prove in $\mathcal{A}^\omega[X, d, W]_{-b}$ the following sentence:*

$$\forall x \in P \forall y \in K \forall z^X \forall f^{X \to X} \\ \big(f \text{ n.e.} \wedge \forall u^0 B_\forall(x, y, z, f, u) \to \exists v^0 C_\exists(x, y, z, f, v)\big),$$

where 0_X does not occur in $B_\forall$ and $C_\exists$. Then there exists a computable functional $\Phi : \mathbb{N}^{\mathbb{N}} \times \mathbb{N} \to \mathbb{N}$ s.t. for all representatives $r_x \in \mathbb{N}^{\mathbb{N}}$ of $x \in P$ and all $b \in \mathbb{N}$

$$\forall y \in K \forall z^X \forall f^{X\to X} \big(f \text{ n.e.} \wedge d_X(z, f(z)) \le_{\mathbb{R}} (b)_{\mathbb{R}} \\ \wedge \forall u^0 \le \Phi(r_x, b) B_\forall(x, y, z, f, u) \to \exists v^0 \le \Phi(r_x, b) C_\exists(x, y, z, f, v) \big)$$

holds in all nonempty hyperbolic spaces (X, d, W).
Analogously, for $\mathcal{A}^\omega[X, d, W, \mathrm{CAT}(0)]_{-b}$ where (X, d, W) is a CAT(0) space.

2. *The corollary also holds for an additional parameter $\forall z'^X$ if we add the additional premise $d_X(z, z') \le_{\mathbb{R}} (b)_{\mathbb{R}}$ to the conclusion.*

3. *Furthermore, the corollary holds for an additional parameter $\forall c^{0\to X}$ if one adds the premise $\forall n(d_X(z, c(n)) \le_{\mathbb{R}} (b)_{\mathbb{R}})$ or just $\forall n(d_X(z, c(n)) \le_{\mathbb{R}} (g(n))_{\mathbb{R}})$ to the conclusion, where the bound then additionally depends on $g : \mathbb{N} \to \mathbb{N}$.*

4. *Statements 1., 2. and 3. also hold if we replace 'f n.e.' with 'f Lipschitz continuous' (with constant $L \in \mathbb{Q}^*_+$), 'f Hölder-Lipschitz continuous' (with constants $L, \alpha \in \mathbb{Q}^*_+$, where $\alpha \le 1$) or 'f uniformly continuous' (with modulus $\omega : \mathbb{N} \to \mathbb{N}$). For Lipschitz and Lipschitz-Hölder continuous functions the bound additionally depends on the given constants and for uniformly continuous functions the bound additionally depends on the given modulus of uniform continuity.*

5. *Furthermore, 1., 2. and 3. hold if we replace 'f n.e.' with 'f weakly quasi-nonexpansive'. For weakly quasi-nonexpansive functions (with fixed point p) we need to state the additional premise '$d_X(z, p) \le_{\mathbb{R}} (b)_{\mathbb{R}}$' in the conclusion.*

6. *More generally, 1., 2. and 3. hold if in the conclusion f satisfies '$d_X(z, f(z)) \le_{\mathbb{R}} (b)_{\mathbb{R}}$' and if instead of 'f n.e.' we assume*

$$\forall n^0 \forall z_1^X, z_2^X (d_X(z_1, z_2) <_{\mathbb{R}} (n)_{\mathbb{R}} \to d_X(f(z_1), f(z_2)) \le_{\mathbb{R}} (\Omega_0(n))_{\mathbb{R}}), \quad (*)$$

where Ω_0 is a function $\mathbb{N} \to \mathbb{N}$. The bound then depends on Ω_0 and b.

7. *Finally, 1.,2. and 3. hold if 'f n.e.' is replaced by*

$$\forall n^0 \forall \tilde{z}^X (d_X(z, \tilde{z}) <_{\mathbb{R}} (n)_{\mathbb{R}} \to d_X(z, f(\tilde{z})) \le_{\mathbb{R}} (\Omega(n))_{\mathbb{R}}), \quad (**)$$

where Ω is a function $\mathbb{N} \to \mathbb{N}$. Then we can drop '$d_X(z, f(z)) \le_{\mathbb{R}} (b)_{\mathbb{R}}$' in the conclusion and the extracted bound only depends on Ω instead of b.

As with Theorem 3.15 we discuss the main points of the proof (for the full details see [41]). The proof primarily consists of choosing a suitable a and then defining suitable a-majorants for the various parameters z, z', c and the different kinds of functions f. The parameters ranging over Polish and compact Polish spaces are treated as before. For the remaining majorants, we choose $a = z$. The parameter z is then a-majorized by the natural number 0, the parameter z' by the given bound b on $d_X(z, z')$ and the parameter c by that same b, if the corresponding premise is $\forall n(d_X(z, c(n)) \leq_{\mathbb{R}} (b)_{\mathbb{R}})$, or by g^M, if the premise is $\forall n(d_X(z, c(n)) \leq_{\mathbb{R}} (g(n))_{\mathbb{R}})$.

For the various kinds of functions f we treat the most general case first: if f satisfies the condition $(**)$ then obviously Ω^M is an a-majorant for f. A nonexpansive function satisfies $(**)$ with $\Omega(n) = n + b$, where b is a bound on $d_X(z, f(z))$. In the same way one sees that Lipschitz and Hölder-Lipschitz functions satisfy $(**)$ with $\Omega(n) := L \cdot n + b$ and $\Omega(n) := L \cdot n^\alpha + b$ respectively. A weakly quasi-nonexpansive f with fixed point p – under the additional assumption $d_X(z, p)$ – satisfies $(**)$ with $\Omega(n) = n + 2b$. A function f satisfying $d_X(z, f(z)) \leq b$ and $(*)$ satisfies $(**)$ with $\Omega(n) := \Omega_0(n) + b$.

The trickiest case is f being uniformly continuous with modulus of uniform continuity ω, i.e. $\forall x, y \in X \forall k \in \mathbb{N}(d(x, y) < 2^{-\omega(k)} \rightarrow d(f(x), f(y)) \leq 2^{-k})$. Here, we use special properties of hyperbolic spaces (X, d, W) which allow one, using the function W, to produce a sequence of intermediate points $z_1, \ldots z_{k-1}$ between the points z and $\tilde{z}$ such that the distance between any two successive points is strictly less than $2^{-\omega(0)}$. Then, using the uniform continuity of f, the distance $d_X(f(z), f(\tilde{z}))$ is less than $n \cdot 2^{\omega(0)} + 1$, and thus f satisfies $(**)$ with $\Omega(n) := n \cdot 2^{\omega(0)} + b + 1$, as we assumed $d_X(z, f(z)) \leq b$.

For the formal details of the verification of these majorants see [41]. Note, that neither the space nor the range of f are in any way assumed to be bounded (i.e. by some constant b), and still the extracted bounds are highly uniform, depending only on e.g. a suitable Lipschitz constant L and a bound b on $d_X(z, f(z))$ for the class of Lipschitz continuous functions. Also note, that all these majorants are valid in non-hyperbolic metric spaces as well, except for the majorant for uniformly continuous functions which can only be verified using special structural properties of hyperbolic spaces.

We close the discussion of Corollary 3.19 by repeating an important remark from [41] regarding the extensionality of the functions f under consideration in Corollary 3.19:

Remark ([41]). *Note that for f nonexpansive, Lipschitz, Hölder-Lipschitz or uniformly continuous, f is provably extensional. For f weakly quasi-nonexpansive or f satisfying conditions $(*)$ or $(**)$ it does not follow that f is extensional. Thus in these cases, if an instance of the extensionality of f is used in a proof, it must either be provable via the extensionality rule (or one must explicitly require f to be (provably) extensional, e.g. by requiring that f is at least uniformly continuous).*

The next corollary to Theorem 3.15 we discuss is the following generalization of Corollary 3.11 in [77]:

Definition 3.22. *Let $f : X \rightarrow X$, then*

- *for $Fix(f) := \{x^X \mid x =_X f(x)\}$ the formula $Fix(f) \neq \emptyset$ expresses f has a fixed point,*

- *for* $Fix_\varepsilon(f,y,b) := \{x^X \mid d_X(x,f(x)) \leq_{\mathbb{R}} \varepsilon \wedge d_X(x,y) \leq_{\mathbb{R}} b\}$ *and* $\varepsilon > 0$ *the formula* $Fix_\varepsilon(f,y,b) \neq \emptyset$ *expresses* f *has an* ε*-fixed point in a* b*-neighborhood of* y.

Corollary 3.23 ([41]). *1. Let* P *(resp.* K*) be a* $\mathcal{A}^\omega$*-definable Polish space (resp. compact Polish space) and let* $B_\forall$, $C_\exists$ *be as before. If* $\mathcal{A}^\omega[X,d,W]_{-b}$ *proves that*

$$\forall x \in P\ \forall y \in K\ \forall z^X, f^{X\to X}\big(f \text{ n.e. } \wedge Fix(f) \neq \emptyset \wedge \forall u^0 B_\forall \to \exists v^0 C_\exists\big)$$

then there exists a computable functional $\Phi^{1\to 0\to 0}$ *(on representatives* $r_x : \mathbb{N} \to \mathbb{N}$ *of elements* x *of* P*) s.t. for all* $r_x \in \mathbb{N}^{\mathbb{N}}, b \in \mathbb{N}$

$$\begin{aligned}\forall y \in K\ \forall z^X, f^{X\to X}\big(&f \text{ n.e. } \wedge \forall \varepsilon > 0 Fix_\varepsilon(f,z,b) \neq \emptyset \\ &\wedge d_X(z,f(z)) \leq_{\mathbb{R}} (b)_{\mathbb{R}} \wedge \forall u^0 \leq \Phi(r_x,b) B_\forall \to \exists v^0 \leq \Phi(r_x,b) C_\exists\big).\end{aligned}$$

holds in any nonempty hyperbolic space (X,d,W).
Analogously, for $\mathcal{A}^\omega[X,d,W,\mathrm{CAT}(0)]_{-b}$ *where then* (X,d,W) *has to be a CAT(0) space.*

2. *The corollary also holds if 'f n.e.' is replaced by f Lipschitz continuous, Hölder-Lipschitz continuous or uniformly continuous, where the extracted bound then additionally will depend on the respective constants and moduli.*

3. *Considering the premise 'f weakly quasi-nonexpansive', i.e.*

$$\exists p^X(f(p) =_X p \wedge \forall w^X(d_X(f(p),f(w)) \leq_{\mathbb{R}} d_X(p,w)))$$

instead of 'f *n.e.* $\wedge Fix(f) \neq \emptyset$*' we may weaken this premise to*

$$\begin{aligned}\forall \varepsilon > 0 \exists p^X(d_X(f(p),p) \leq_{\mathbb{R}} \varepsilon \wedge d_X(z,p) \leq_{\mathbb{R}} (b)_{\mathbb{R}} \\ \wedge \forall w^X(d_X(f(p),f(w)) \leq_{\mathbb{R}} d_X(p,w))).\end{aligned}$$

4. *Let* $\Psi : (X \to X) \to X \to 1$ *be a provably extensional closed term of* $\mathcal{A}^\omega[X,d,W]_{-b}$*, then in 1. and 2. instead of '*$Fix(f) \neq \emptyset$*' we may weaken '*$\Psi(f,p) =_{\mathbb{R}} 0$*', expressing that* $\Psi(f,\cdot)$ *has a root* p*, to '*$\forall \varepsilon > 0 \exists p \in X(d(z,p) \leq b \wedge |\Psi(z,p)| \leq_{\mathbb{R}} \varepsilon)$*', expressing that* $\Psi(f,\cdot)$ *has* ε*-roots* p *which are* b*-close to* z *for every* $\varepsilon > 0$.

We discuss the idea of the proof for the first case of the corollary, the arguments for the other cases are similar. The statement $Fix(f) \neq \emptyset$ can be expressed formally as $\exists p^X \forall k^0(d_X(p,f(p)) \leq 2^{-k})$. Pulling out $\exists p^X$ as an additional parameter $\forall p^X$, the remaining part of $Fix(f) \neq \emptyset$ is a purely universal formula and hence an admissible premise. In order to extract a bound from a proof of a theorem with this premise, we must add the additional premise $d_X(z,p) \leq b$, so that b can become an a-majorant for p for the choice $a = z$. By Corollary 3.21 we then may, given the further premise $d_X(z,f(z))$, extract a bound $\Phi(r_x,b)$ on the $\exists v^0$-quantifier in the conclusion *as well as* the $\forall u^0$-quantifier (and any other $\forall$-quantifiers) in the premise. As usual, r_x is a representative for a given x. This bound is valid if f has an $2^{-\Phi(r_X,b)}$-fixed point in a b-neighborhood around z, which especially is the case, if f has ε-fixed points for every $\varepsilon > 0$.

This argument makes crucial use of the fact that the bound Φ does not depend on p, but only on the bound b on the neighborhood around z in which such a p is to be found.

The remarkable point about this corollary is not so much that we may extract a strongly uniform bound on the $\exists v^0$-quantifier in the conclusion (which already follows from Corollary 3.21), but that we may significantly weaken the premise $Fix(f) \neq \emptyset$ at the same time. In the special case of a *bounded* hyperbolic space the situation is even more favorable: A nonexpansive mapping on a bounded hyperbolic space (X, d, W) does not necessarily have exact fixed points, as required by the original premise, but always has approximate fixed points (see [45]). Thus the weaker premise is provably true and hence may be completely eliminated, while still allowing to extract the same computable bound. In general, this corollary illustrates how we, through the techniques of proof mining, may transform a given proof of an implication into a proof of a stronger conclusion from a weaker, in some cases much weaker premise.

We discuss one final corollary to Theorem 3.15. Even if one can only treat formulas that prenex to $\forall\exists$-formulas with the classical metatheorems, one may still use these metatheorems to analyze more general formulas indirectly, via their Herbrand normal forms. The extraction of bounds from Herbrand normal forms can be applied to a large class of formulas, more precisely those for which the Herbrand index functions have a suitable low type. The types of the index functions in a given Herbrand normal form A^H of a formula A depend on the $\exists\forall$-configurations in the chosen prenexation of A. A quantifier-configuration $\exists x^\rho \forall y^\tau$ leads to a Herbrand index function h_y of type $\rho \to \tau$. If we only consider formulas and prenexations where the types ρ are of degree 0 and the types of τ are of degree $(0, X)$ and 1, then the Herbrand index functions are guaranteed to have a-majorants. Although these majorants may only exist ineffectively, this still suffices to prove a Herbrand version of Corollary 3.23. In this Herbrand version of Corollary 3.23 the focus is no longer on extracting computable bounds, as they would depend on (majorants for) the Herbrand index functions and thus would be of little interest anyway, but exclusively on weakening premises.

Definition 3.24 ([41])**.** *The class $\mathcal{H}$ of formulas consists of all formulas F that have a prenexation $F' \equiv \exists x_1^{\rho_1} \forall y_1^{\tau_1} \ldots \exists x_n^{\rho_n} \forall y_n^{\tau_n} F_\exists(\underline{x}, \underline{y})$ where $F_\exists$ is an $\exists$-formula, the types ρ_i are of degree 0 and the types τ_i are of degree 1 or $(0, X)$.*

Corollary 3.25 ([41])**.** *1. Let P (resp. K) be a $\mathcal{A}^\omega$-definable Polish space (resp. compact Polish space) and let the formula A be in the class $\mathcal{H}$, where moreover A does not contain 0_X. If $\mathcal{A}^\omega[X, d, W]_{-b}$ proves a sentence*

$$\forall x \in P\ \forall y \in K\ \forall z^X, f^{X \to X} \big(f \text{ n.e.} \wedge Fix(f) \neq \emptyset \to A\big)$$

then the following holds in every nonempty hyperbolic space (X, d, W):

$$\forall x \in P\ \forall y \in K\ \forall z^X, f^{X \to X}$$
$$\big(f \text{ n.e.} \wedge \exists b^0 \forall \varepsilon > 0 (Fix_\varepsilon(f, z, b) \neq \emptyset) \to A\big).$$

Analogously, for $\mathcal{A}^\omega[X, d, W, \mathrm{CAT}(0)]_{-b}$ where (X, d, W) is a CAT(0) space.

2. *The corollary also holds if we replace 'f n.e.' with f Lipschitz continuous, Hölder-Lipschitz continuous or uniformly continuous.*

3. *Considering the premise 'f weakly quasi-nonexpansive', i.e.*

$$\exists p^X (f(p) =_X p \wedge \forall w^X (d_X(f(p), f(w)) \leq_{\mathbb{R}} d_X(p, w)))$$

instead of 'f n.e. $\wedge Fix(f) \neq \emptyset$*' we may weaken this premise to*

$$\exists b^0 \forall \varepsilon > 0 \exists p^X (d_X(f(p), p) \leq_{\mathbb{R}} \varepsilon \wedge d_X(z, p) \leq_{\mathbb{R}} (b)_{\mathbb{R}} \\ \wedge \forall w^X (d_X(f(p), f(w)) \leq_{\mathbb{R}} d_X(p, w))).$$

4. *Let* $\Psi : (X \to X) \to X \to 1$ *be a provably extensional closed term of* $\mathcal{A}^\omega[X, d, W]_{-b}$*, then in 1. and 2. instead of '*$Fix(f) \neq \emptyset$*' we may weaken '*$\exists p^X \Psi(f, p) =_{\mathbb{R}} 0$*', expressing that* $\Psi(f, \cdot)$ *has a root in p, to '*$\exists b^0 \forall \varepsilon > 0 \exists p^X (d_X(z, p) \leq_{\mathbb{R}} (b)_{\mathbb{R}} \wedge |\Psi(z, p)| \leq_{\mathbb{R}} \varepsilon)$*', expressing that* $\Psi(f, \cdot)$ *has* ε*-roots p which are b-close to z for every* $\varepsilon > 0$*.*

The general idea of the proof is the following: For a suitable prenexation and Herbrandization A^H of A the implication $A \to A^H$ is logically valid, so we may replace A by A^H in the conclusion of the implication. Next, we pull out the quantifiers ranging over the Herbrand index functions, so that the statement now has a suitable logical form to allow one, as in the Corollaries 3.21 and 3.23 the extraction of numerical bounds on the universal quantifiers in the premise and the existential quantifiers in the conclusion. These bounds depend on the parameter x through a representative r_X, on a bound b on $d_X(z, f(z))$ and on majorants for the Herbrand index functions. If a Herbrand index function h is of type 1, a majorant h^* can be obtained by the construction h^M. If the Herbrand index function h is of type $(0, X)$ it basically represents a sequence of elements of X, and then we may ineffectively choose a sequence of numbers h^* such that $h^*(n) \geq d_X(h(m), a)$ for all $n \geq m$ as a majorant for h.

Thus, reasoning as in Corollary 3.23 we may weaken the premises to their ε-version. Shifting the quantifiers ranging over the Herbrand index function back in, we get that the weakened premises imply the Herbrand normal form A^H of A. For the next step, we use that, again ineffectively(by classical logic and the axiom of choice), the Herbrand normal form A^H implies back A to let the weakened premises imply the original A, although still under the additional assumption $d_X(z, f(z)) \leq b$. Finally, as we are only interested in the truth of the weakening of premises, not in bounds, we may also eliminate that additional premise, as in any given metric space, for any given z and any given $f : X \to X$ the distance $d_X(z, f(z))$ is less than *some* b. Choosing a large enough b satisfying both $\forall \varepsilon > 0 Fix_\varepsilon(f, z, b) \neq \emptyset$ and $d_X(z, f(z)) \leq b$ we get the result stated above in Corollary 3.25.

Note, that the restrictions on the formula class $\mathcal{H}$ are essential. The simplest type of Herbrand index functions that is disallowed is the type $0 \to X$. This can be motivated by a simple counterexample. Consider the formula $Fix(f) \neq \emptyset$. This is expressed as $\exists p^X \forall n^0 (d_X(p, f(p)) \leq 2^{-n})$. Thus $Fix(f) \neq \emptyset$ gives rise to Herbrand index functions of type $0 \to X$. If that type were not disallowed we could deduce from $Fix(f) \neq \emptyset \to Fix(f) \neq \emptyset$ and the subsequent weakening of premises that every function f that has approximate fixed points has exact fixed points – a statement that cannot be true and for which one easily produces counterexamples.

This concludes the discussion of the classical metatheorems for abstract bounded and unbounded metric, hyperbolic and CAT(0)-spaces.

3.2 Normed linear spaces

For proofs of theorems about abstract real normed linear spaces with convex subsets C there are two general approaches to extracting bounds. Theorems merely concerning the convex subsets C and not using properties of the enclosing normed linear space may be treated employing an idea of Machado[96]. With two additional axioms one may characterize convex subsets of normed linear spaces in the setting of hyperbolic spaces. The axioms express additional conditions on the hyperbolic function W: (1) that the convex combinations are independent of the order in which they are carried out, and (2) that the distance is homothetic. Expressed formally, this yields

(I) $\forall x, y, z \in X \forall \lambda_1, \lambda_2, \lambda_3 \in [0,1]\big(\lambda_1 + \lambda_2 + \lambda_3 =_{\mathbb{R}} 1 \to$
$W(z, W(y, x, \frac{\lambda_1}{1-\lambda_3}), 1-\lambda_3) = W(x, W(z, y, \frac{\lambda_2}{1-\lambda_1}), 1-\lambda_1)\big)$,

(II) $\forall x, y, z \in X \forall \lambda \in [0,1]\big(d(W(z,x,\lambda), W(z,y,\lambda)) = \lambda \cdot d(x,y)\big)$.

The formal version of axiom (I) must be altered slightly, as the axiom would no longer be purely universal if we write it with the equality $\lambda_1 + \lambda_2 + \lambda_3 =_{\mathbb{R}} 1$ in the premise. Equality on the reals is a universal statement and thus the axiom in its entirety would prenex to a $\forall\exists$-statement. Instead we use a trick, requiring only λ_1 and λ_2 and then defining $\bar{\lambda_1}, \bar{\lambda_2}$ and $\bar{\lambda_3}$ such that both $\bar{\lambda_1} + \bar{\lambda_2} + \bar{\lambda_3} =_{\mathbb{R}} 1$ and if $\lambda_i \in [0,1]$ and $\lambda_1 + \lambda_2 + \lambda_3 =_{\mathbb{R}} 1$ then $\bar{\lambda_i} = \lambda_i$ for $i = 1, 2, 3$. The formal version of the Machado-axioms is:

(I) $\forall x^X, y^X, z^X \forall \lambda_1^1, \lambda_2^1$
$\big(W_X(z, W_X(y, x, \frac{\bar{\lambda_1}}{1-\bar{\lambda_3}}), 1-\bar{\lambda_3}) = W_X(x, W_X(z, y, \frac{\bar{\lambda_2}}{1-\bar{\lambda_1}}), 1-\bar{\lambda_1})\big)$,
where $\bar{\lambda_1} =_1 \tilde{\lambda}_1$, $\bar{\lambda_2} =_1 min_{\mathbb{R}}(\tilde{\lambda}_2, 1 -_{\mathbb{R}} \bar{\lambda_1})$ and $\bar{\lambda_3} =_1 1 -_{\mathbb{R}} (\bar{\lambda_1} +_{\mathbb{R}} \bar{\lambda_2})$,

(II) $\forall x^X, y^X, z^X \forall \lambda^1 \big(d_X(W_X(z,x,\lambda), W_X(z,y,\lambda)) =_{\mathbb{R}} \tilde{\lambda} \cdot_{\mathbb{R}} d_X(x,y)\big)$,

where $\tilde{\lambda}$ is the construction in Definition 3.2. In the formulation of these axioms we again use that the previous W_X-axioms(axioms (4)-(7) given above) ensure that $W_X(x,y,\lambda) =_X W_X(x,y,\tilde{\lambda})$ already syntactically and not only through the interpretation of W_X in the model.

Thus, by Remark 3.1, using the two additional Machado-axioms theorems concerning convex subsets of normed linear spaces may be treated using Theorem 3.15 and the related corollaries.

Alternatively, one may directly develop metatheorems for real normed linear spaces $(X, \|\cdot\|)$, with or without convex subsets C. The general idea is similar to developing metatheorems for metric spaces: one extends theories $\mathcal{A}^\omega$ and $\mathcal{A}_i^\omega$ with necessary additional axioms and constants for normed linear spaces and extends the monotone functional interpretation (based on functional interpretation and a-majorization) to these new theories. Compared with the approach for metric spaces, there are two important differences: (1), in the setting of normed linear spaces we fix the choice of a to $a = 0_X$. For any given choice of a virtually all the majorants for the new, additional constants of e.g $\mathcal{A}^\omega[X, \|\cdot\|, C]$ will depend heavily on a bound n on $\|0_X - a\|$ (whereas in the metric case only the majorant for 0_X depended on a bound on $d_X(0_X, a)$). Choosing $a = 0_X$ the majorants will thus depend on $\|0_X - 0_X\| = 0$, so that this is no longer a problem. (2), in the metric case one could eliminate the dependency on 0_X if the constant did not occur in the theorem to be analyzed. This was possible, because the axioms for abstract metric spaces place no requirements on the constant 0_X. For normed linear spaces this is not the case, and thus, even if 0_X does not occur in the theorem, it may influence the extracted bounds - another weighty reason for fixing the choice $a = 0_X$.

The type C for convex subsets C of a space $(X, \|\cdot\|)$ is not a ground type, but quantification over C is merely an abbreviation using the characteristic function χ_C for the subset $C \subseteq X$. We use the following abbreviations:

$$\begin{array}{ll}
\forall x^C\ A(x) & :\equiv \forall x^X(\chi_C(x^X) =_0 0 \to A(x)), \\
\forall f^{1\to C}\ A(f) & :\equiv \forall f^{1\to X}\big(\forall y^1(\chi_C(f(y)) =_0 0) \to A(f)\big), \\
\forall f^{X\to C}\ A(f) & :\equiv \forall f^{X\to X}\big(\forall y^X(\chi_C(f(y)) =_0 0) \to A(f)\big) \\
\forall f^{C\to C}\ A(f) & :\equiv \forall f^{X\to X}\big(\forall x^X(\chi_C(x) =_0 0 \to \chi_C(f(x)) =_0 0) \to A(\tilde{f})\big),
\end{array}$$

where $\tilde{f}(x) = \begin{cases} f(x), & \text{if } \chi_C(x) =_0 0 \\ c_X, & \text{otherwise.} \end{cases}$

Analogously, for the corresponding $\exists$-quantifiers with '$\wedge$' instead of '$\to$'. Up to and including types of degree $(1, X, C)$ the additional premises are $\forall$-formulas and hence admissible premises for our metatheorems. For more complicated types, e.g. $(C \to C) \to C$, the additional premises are of a too complicated logical form to appear in a premise.

Also note, that if we write 'f nonexpansive' (similarly for the other notions in Definition 3.20) for a function $f : C \to C$, this is to be understood as the $\forall$-formula

$$\forall x^X, y^X(\chi_C(x) =_0 0 =_0 \chi_C(y) \to \|f(x) - f(y)\|_X \leq_{\mathbb{R}} \|x - y\|_X).$$

Furthermore note, that we cannot assume the extensionality of the characteristic function χ_C, i.e. $x =_X y \to \chi_C(x) =_0 \chi_C(y)$, as monotone functional interpretation would extract from this statement a modulus expressing how close to an element of C an element has to be to behave *like* an element of C when e.g. passed as an argument to a function $f : C \to C$. In general, such a modulus cannot exist and thus we are again restricted to using the extensionality rule.

There is one final subtlety with parameters $f : C \to C$. When we want to find majorants for such f, we in fact must consider the extension $\tilde{f}$. In the case of bounded convex subsets C this is easily done using the bound b on the diameter of C. For unbounded convex subsets C we need to employ the given properties of the function f, such as e.g. its nonexpansivity, to majorize f. However, the extended function $\tilde{f}$ need not in general inherit these properties from f. A majorant for $\tilde{f}$ will therefor be a combination of a majorant for f on C and a majorant for c_X, as $\tilde{f}$ is defined to be constant with value c_X on $x \setminus C$. The final majorant for $\tilde{f}$ is then the maximum of these two majorants.

Next, we repeat the following definition:

Definition 3.26 ([41]). *We say that a sentence of $\mathcal{L}(\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$ holds in a non-trivial (real) normed linear space with a nonempty convex subset C, if it holds in the models[8] of $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$ obtained by letting the variables range over the appropriate universes of the full set-theoretic type structure $\mathcal{S}^{\omega,X}$ with the sets $\mathbb{N}, X$ as the universes for the base types 0 and X. Here 0_X is interpreted by the zero vector of the linear space X, 1_X by some vector $a \in X$ with $\|a\| = 1$, $+_X$ is interpreted as addition in X, $-_X$ is the inverse of x w.r.t. $+$ in X, $\cdot_X$ is interpreted as $\lambda\alpha \in \mathbb{N}^{\mathbb{N}}, x \in X.r_\alpha \cdot x$, where r_α is the unique real number represented by α and $\cdot$ refers to scalar multiplication in the $\mathbb{R}$-linear space X. Finally, $\|\cdot\|_X$ is interpreted by $\lambda x \in X.(\|x\|)_\circ$. For the nonempty convex subset $C \subseteq X$, χ_C is interpreted as the characteristic function for C and c_X by some arbitrary element of C.*

The main metatheorem for normed linear spaces in [41] is:

Theorem 3.27. *1. Let ρ be of degree $(1, X), (1, X, C)$ or 2 and let $B_\forall(x, u)$, resp. $C_\exists(x, v)$, contain only x, u free, resp. x, v free. Assume*

$$\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b} \vdash \forall x^\rho(\forall u^0 B_\forall(x, u) \to \exists v^0 C_\exists(x, v)),$$

Then there exists a partial functional $\Phi : S_{\widehat{\rho}} \mapsto \mathbb{N}$ whose restriction to the strongly majorizable elements of $S_{\widehat{\rho}}$ is totally computable functional of $\mathcal{M}^\omega$ and the following holds in all non-trivial (real) normed linear spaces $(X, \|\cdot\|, C)$ with a nonempty convex subset C: for all $x \in S_\rho$, $x^* \in S_{\widehat{\rho}}$ and $n \in \mathbb{N}$ if $x^* \gtrsim^{0_X} x$ and $(n)_{\mathbb{R}} \geq_{\mathbb{R}} \|c_X\|_X$ then

$$\forall u \leq \Phi(x^*, n) B_\forall(x, u) \to \exists v \leq \Phi(x^*, n) C_\exists(x, v)).$$

In particular, if ρ is in addition of degree $\widehat{1}$, then $\Phi : S_{\widehat{\rho}} \times \mathbb{N} \to \mathbb{N}$ is totally computable.

2. *For uniformly convex spaces with modulus of uniform convexity η statement 1. holds with $(X, \|\cdot\|, C, \eta)$, $\mathcal{A}^\omega[X, \|\cdot\|, C, \eta]_{-b}$ instead of $(X, \|\cdot\|, C)$, $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$, where the extracted bound Φ additionally depends on η.*

3. *Analogously, for real inner product spaces $(X, \langle\cdot,\cdot\rangle)$.*

As in the metric case, instead of single variables x, u, v and single premises $\forall u B_\forall(x, u)$ we may have tuples of variables and finite conjunctions of premises.

[8] Again we use the plural, as in the setting of normed linear spaces the interpretation of the constants is not uniquely determined.

The proof of this theorem follows the same pattern as the proof of Theorem 3.15. Again, for the full details of the proof see Section 9 in [41]. As before, the central part of the proof consists of defining and verifying the necessary a-majorants, or in fact 0_X-majorants, for the new constants of the theories $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$, etc. The majorants given in [41] are the following:

- $0^0 \gtrsim^{0_X} 0_X$,
- $1^0 \gtrsim^{0_X} 1_X$,
- $\lambda x^0.((x)_{\mathbb{R}})_\circ \gtrsim^{0_X} \|\cdot\|_X^{X\to 1}$,
- $\lambda x^0, y^0.x + y \gtrsim^{0_X} +_X^{X\to X\to X}$,
- $\lambda x^0.x \gtrsim^{0_X} -_X^{X\to X}$,
- $\lambda \alpha^1, x^0.(\alpha(0) + 1) \cdot x \gtrsim^{0_X} \cdot_X^{1\to X\to X}$.

For the convex subset C, we have the characteristic term χ_C for the subset C, which is majorized as follows:

$$\lambda x^0.1 \gtrsim^{0_X} \chi_C^{X\to 0}.$$

For the constant $c_X \in C$ we have, given an $n \geq \|c_X\|$, the 0_X-majorant

$$n^0 \gtrsim_X^{0_X} c_X.$$

For uniformly convex spaces we 0_X-majorize the modulus $\eta : \mathbb{N} \to \mathbb{N}$ of uniform convexity by

$$(\eta)^M \gtrsim_1^{0_X} \eta.$$

The 0_X-majorants for 0_X and 1_X are trivial. Similar to the a-majorant for the metric d_X, the 0_X-majorant of the norm employs the $()_\circ$-operator applied to a natural number. Also note, that the 0_X-majorant for the scalar product $\cdot_X$ depends on the chosen representation for real numbers. For a type 1 representative α of a real number, $\alpha(0)$ is the code of a rational 2^{-0}-approximation of the actual real number. As the coding of rational numbers is monotone, $(\alpha(0) + 1) \cdot_{\mathbb{N}} x^*$ is an integer upper bound on the norm of $\|\alpha \cdot_X x\|$, assuming that x^* is an integer upper bound on the norm of $\|x\|$.

The remaining details of the proof are largely identical to the proof of Theorem 3.15.

As corollaries of Theorem 3.27 one may prove Theorem 3.30 in [77] and the counterparts for the normed linear setting of Corollaries 3.19 and 3.23 and furthermore of Corollary 3.25. For detailed formulations, proofs and discussions of these corollaries see [41].

One important difference is that given e.g. an x and a y it is no longer sufficient to bound $\|x-y\|$, but one also needs to bound $\|x\|$. As a simple counterexample, consider $\forall x^X, y^X \exists n^0(\|x\| + \|y\| \leq n)$, which is provable for the above theories for normed linear spaces. However, one clearly cannot bound $\exists n$ just in terms of a bound on $\|x - y\|$. An actual proof of this "counterexample" would still use a certain amount of the structure of normed linear spaces, so the counterexample does not apply to the treatment of (convex subsets of) normed linear spaces via the Machado axioms, as these axioms do not provide enough structure.

This concludes the discussion of the classical metatheorems for (variants of) abstract real normed linear spaces

3.3 Semi-intuitionistic theories

We now turn towards metatheorems for theories based on intuitionistic formal systems for analysis $\mathcal{A}_i^\omega$ developed in [42], although we will restrict the discussion the most important differences between the semi-intuitionistic metatheorems and the classical metatheorems. For theories based on classical logic the (monotone) proof interpretation of choice is monotone functional interpretation. An approach using negative translation combined with modified realizability and an additional application of Friedman's A-translation also yields a technique for extracting programs from classical proofs. However, the use of the A-translation imposes some serious restrictions relative to functional interpretation, as e.g. the A-translation is not sound for Spector's extensionality rule. For the metatheorems we aim to prove it is even more critical that the interpretation of the axiom of dependent choice via A-translation and modified realizability depends on a continuity axiom that does not hold in the model $\mathcal{M}^\omega$ and uses a special variant of bar-recursion, so-called modified bar-recursion, that can only be majorized ineffectively (see [8]). For applications of (refinements of) the combination of negative translation, A-translation and modified realizability to the extraction of programs from classical proofs see e.g. [11, 7].

As shown in [70], for theories based on intuitionistic analysis the choice between functional interpretation and modified realizability (now without A-translation) is also significant. More precisely, the choice between the two different interpretations primarily translates into a choice between interpreting or not interpreting the Markov principle M^ω. As mentioned earlier, functional interpretation interprets the Markov principle and as a consequence we can then only allow the weak extensionality rule in our system. With functional interpretation, we could treat systems based on weakly extensional Heyting arithmetic + M^ω and extract bounds from proofs of formulas that prenex to $\forall x \exists y A(x, y)$ for *arbitrary* A, rather than just quantifier-free formulas A_{qf}.

On the other hand, modified realizability does not interpret the Markov principle and the (existential-free) extensionality axiom is interpreted by the empty realizer. Hence, we here can allow the full extensionality axiom in our formal system. In both cases the monotone variant of the chosen interpretation allows one to further extend theories with certain additional principles that may not have an exact interpretation, but do have a monotone interpretation. Here, it is again modified realizability that allows the more general and powerful principles to be added to the theory. As an example, for modified realizability one can add comprehension principles for *arbitrary* negated formulas (and thereby also the corresponding independence of premise principles), while for functional interpretation we can at most allow the independence of premise principle for purely universal formulas and instead of comprehension for arbitrary negated formulas at most the weak König's lemma, WKL.

Finally, for classical theories the interpretation of the axiom of dependent choice requires the use of bar-recursion and, subsequently, necessitates reasoning in the type structure of strongly hereditarily majorizable functionals in order to verify the extraction and majorization process that yields the computable bounds. In contrast, for intuitionistic theories and modified realizability the full axiom of choice has a trivial interpretation, and the extracted realizers are *primitive* recursive, so that the majorization may be carried out in the type structure of set-theoretic functionals, without a 'detour' through the majorizable functionals.

We define the following principles:

Let comprehension for negated formulas be the principle:

$$CA_{\neg}^{\underline{\rho}} : \exists \Phi \leq_{\underline{\rho} \to 0} \lambda \underline{x}^{\underline{\rho}}.1^0 \forall \underline{y}^{\underline{\rho}} (\Phi(\underline{y}) =_0 0 \leftrightarrow \neg A(\underline{y})),$$

where $\underline{y} = y_1^{\rho_1}, \ldots, y_k^{\rho_k}$ is a tuple of variables of arbitrary types and A is an arbitrary formula. The union of $CA_{\neg}^{\underline{\rho}}$ over all types ρ of the underlying formal system is denoted by $CA_{\neg}$.

In [42], the following general metatheorem is proved:

Theorem 3.28 ([42]). *1. Let σ be a type of degree 1, let ρ be a type of degree $(\cdot, 0)$ and let τ be a type of degree $(\cdot, X)$. Let $s^{\sigma \to \rho}$ be a closed term of $\mathcal{A}_i^{\omega}[X,d]$ and let A (resp. B) be an arbitrary formula with only x, y, z, n (resp. x, y, z) free. Let $\Gamma_{\neg}$ be a set of sentences of the form $\forall u^{\alpha}(C \to \exists v \leq_{\beta} tu \exists w^{\gamma} \neg D)$ with $t^{\alpha \to \beta}$ be a closed term of $\mathcal{A}_i^{\omega}[X,d]$, the type $\alpha \in \mathbf{T}^X$ arbitrary, the type β of degree $(\cdot, 0)$ and γ of degree $(\cdot, X)$. If*

$$\mathcal{A}_i^{\omega}[X,d] + CA_{\neg} + \Gamma_{\neg} \vdash \forall x^{\sigma} \forall y \leq_{\rho} s(x) \forall z^{\tau} (\neg B \to \exists n^0 A),$$

then one can extract a primitive recursive (in the sense of Gödel) functional $\Phi : \mathcal{S}_{\sigma} \times \mathbb{N} \to \mathbb{N}$ such that for all $b \in \mathbb{N}$

$$\forall x^{\sigma} \forall y \leq_{\rho} s(x) \forall z^{\tau} \exists n \leq \Phi(x,b)(\neg B \to A)$$

holds in any (non-empty) metric space (X,d) whose metric is bounded by $b \in \mathbb{N}$ and which satisfies $\Gamma_{\neg}$.[9]

2. *For bounded hyperbolic spaces (X,d,W), '1.' holds with $\mathcal{A}_i[X,d,W], (X,d,W)$ instead of $\mathcal{A}_i^{\omega}[X,d], (X,d)$.*

3. *If the premise is proved in $\mathcal{A}_i^{\omega}[X,d,W,\mathrm{CAT}(0)]$ instead of $\mathcal{A}_i^{\omega}[X,d,W]$ then the conclusion holds in all nonempty b-bounded* CAT(0) *spaces satisfying $\Gamma_{\neg}$.*

[9] Here b_X is understood to be interpreted by b.

As in the classical case, instead of single variables and single premises we may also have tuples of variables and a finite conjunction of premises.

Naturally, this theorem can also be extended to cover *unbounded* metric, hyperbolic and CAT(0)-spaces. This generalization uses the same arguments based on the new a-majorization relation sketched for the classical case earlier in this chapter. For the detailed proof of Theorem 3.28 we refer to [42]. Instead of the proof, we here sketch the main differences to the metatheorems for the classical case:

- In the classical case we were restricted to extracting bounds on conclusions $\exists v^0 C_\exists$, where $A_\exists$ is an existential formula with certain restriction on the types occurring in $A_\exists$. In the semi-intuitionistic case, we may extract bounds from $\exists n^0 A$ where A may be an arbitrary formula with no restrictions on the types.

- In the classical case the premises were restricted to formulas $\forall u^0 B_\forall$ where the definition on $\forall$-formulas again contained certain restrictions. In the semi-intuitionistic case premises may be of the form $\neg B$ where B is an arbitrary formula. However, in the classical case we could weaken the premise by extracting a bound on $\forall u^0$. In the semi-intuitionistic case the premise $\neg B$ remains unchanged.

- As mentioned earlier, in the classical case we may only use the weak extensionality rule, while we may use the full extensionality axiom in the semi-intuitionistic case (i.e. when using modified realizability).

- In addition to $CA_\neg$, we may in the semi-intuitionistic case add principles from the very general class $\Gamma_\neg$ to the theory. In general, principles from the class $\Gamma_\neg$ could not be allowed in a classical system. E.g. comprehension for arbitrary negated formulas, which falls into the class $\Gamma_\neg$, would lead to comprehension for all formulas from which one immediately could produce counterexamples to the extractability of computable bounds.

- In the classical case, depending on the extent to which the axiom of dependent choice is used in the proof, the extracted bounds are bar-recursive and the proof requires reasoning in the type structure of hereditarily majorizable functionals. In the semi-constructive case the extracted bounds are primitive recursive and the proof does not use the majorizable functionals, but only the majorizability of all functionals in Gödel's **T**. This also explains why we allow the more general types $(\cdot, 0)$ and $(\cdot, X)$ for ρ and τ instead of, as in the classical case, 1 and $(1, X)$.

We next present a corollary presented in [42] which both illustrates the extension of the semi-intuitionistic metatheorems to the unbounded setting and functions as a semi-intuitionistic counterpart to Corollary 3.23 sketched above.

Corollary 3.29. *1. Let P (resp. K) be a $\mathcal{A}_i^\omega$-definable Polish space (resp. compact Polish space) and let A and B be as before but not containing the constant 0_X. If $\mathcal{A}_i^\omega[X,d,W]_{-b} + CA_\neg$ proves that*

$$\forall x \in P \forall y \in K \forall z^X, f^{X\to X}, \Omega^1$$
$$(\forall k^0, \tilde{z}^X (d_X(z,\tilde{z}) \leq_{\mathbb{R}} (k)_{\mathbb{R}} \to d_X(z,f(\tilde{z})) \leq_{\mathbb{R}} (\Omega(k))_{\mathbb{R}}) \wedge \neg B \to \exists n^0 A)$$

then there exists a primitive recursive functional $\Phi^{1\to1\to0}$ (on representatives $x : \mathbb{N} \to \mathbb{N}$ of elements of P) such that for all $x, \Omega \in \mathbb{N}^{\mathbb{N}}$

$$\forall y \in K \forall z^X, f^{X\to X}, \Omega^1 \exists n \leq \Phi(x,\Omega)$$
$$(\forall k^0, \tilde{z}^X (d_X(z,\tilde{z}) \leq_{\mathbb{R}} (k)_{\mathbb{R}} \to d_X(z,f(\tilde{z})) \leq_{\mathbb{R}} (\Omega(k))_{\mathbb{R}}) \wedge \neg B \to A)$$

holds in any (non-empty) hyperbolic space (X,d,W).

2. The result also holds for $\mathcal{A}_i^\omega[X,d]_{-b}, (X,d)$.

Even if 'z' does not occur in B, A we need the assumption on f, Ω to hold for **some** *z in X.*

Two differences between this corollary and the classical version are particularly noteworthy:

- In the classical case, some care had to be taken with regard to extensionality. A function satisfying the above Ω-condition is not necessarily extensional and thus any extensionality used in the proof must be proven using Spector's extensionality rule. In the semi-intuitionistic case, where we have access to the full axiom of extensionality, there are no such problems.
- In the classical case, it was an important point that one could add certain premises, such as $Fix(f) \neq \emptyset$, and through functional interpretation weaken or even completely eliminate these premises. In the semi-intuitionistic case this does not apply, as the monotone modified realizability interpretation leaves (negated) premises untouched.

Naturally, semi-intuitionistic metatheorems can also be developed for the extension of $\mathcal{A}_i^\omega$ with abstract real normed linear spaces as well as uniformly convex spaces, inner product spaces and Hilbert spaces. The proofs of these semi-intuitionistic variants of the classical metatheorems for normed linear spaces are obtained as in the metric case. For the details we once more refer to [42].

Chapter 4

A Case Study in Fixed Point Theory

The results discussed in this chapter have previously been published in [39].

In this chapter we illustrate how the techniques from proof mining can be used to extract additional information from a given mathematical proof. The main example will be the quantitative version of Kirk's fixed point theorem for asymptotic contractions (see [59]) that was obtained in [39]. However, we also discuss the general considerations and ideas that are usually applied when analysing or mining mathematical proofs using the techniques and metatheorems described in Chapter 3.

The analysis of mathematical proofs is often carried out in several steps. To begin with, Gödel's Dialectica translation serves as a guiding principle for putting mathematical concepts and the theorem to be analyzed itself into a form suitable for proof mining. This often involves making the computational meaning of mathematical concepts explicit: E.g. assuming that a function is continuous means assuming the existence of a modulus of continuity or asking that a real normed linear space is strictly convex means asking for a modulus of strict convexity. In some cases, we will – with monotone functional interpretation and the extraction od uniform bounds in mind – strengthen the requirements put on a function or a space from e.g. continuity to uniform continuity or strict convexity to uniform convexity. We will discuss examples of this later.

Next, the metatheorems described in Chapter 3 characterize general classes of theorems and proofs from which additional data, i.e. computable bounds, can be extracted and moreover give a-priori guarantees for the extracted bounds on e.g. their complexity or their independence from certain parameters. Often it will not be necessary to completely formalize the theorem and proof under consideration, but rather, the techniques of proof mining serve as guiding principles for the actual extraction of the desired additional data.

If necessary one may also fully formalize the proof to be analyzed, in which case the subsequent extraction of e.g. computable bounds is a fully mechanical (and hence automatizable) procedure. Interactive theorem provers such as

Schwichtenberg's Minlog system[1] significantly lessen the burden of fully formalizing a mathematical proof and also come with a built-in program extraction tool based on functional interpretation.

We now discuss the analysis of Kirk's fixed point theorem for asymptotic contractions that was carried out in [39].

In [59], Kirk defines the notion of asymptotic contractions as follows:

Definition 4.1 (Kirk[59]). *A function $f : X \to X$ on a metric space (X, d) is called an asymptotic contraction with moduli $\phi, \phi_n : [0, \infty) \to [0, \infty)$ if ϕ, ϕ_n are continuous, $\phi(s) < s$ for all $s > 0$ and for all $x, y \in X$*

$$d(f^n(x), f^n(y)) \leq \phi_n(d(x, y))$$

and, moreover, $\phi_n \to \phi$ uniformly on the range of d for all $n \in \mathbb{N}$.

Kirk's notion of asymptotic contractions is a generalization of the notion of contractive mappings (see Edelstein[30] and Rakotch[104]), i.e. mappings $f : X \to X$ satisfying

$$x \neq y \to d(f(x), f(y)) < d(x, y),$$

In compact spaces X this may equivalently be expressed with a modulus of contractivity $\alpha : (0, \infty) \to (0, 1)$

$$d(x, y) \geq \varepsilon > 0 \to d(f(x), f(y)) \leq \alpha(\varepsilon) \cdot d(x, y).$$

The former notion is due to Edelstein, the latter due to Rakotch. In the non-compact (metric space) case Rakotch's notion of contractivity is stronger and always guarantees the existence of a fixed point, whereas there are counterexamples for Edelstein's notion already for bounded (non-compact) metric spaces.

The notion of contractive mappings is in turn a generalization of the notion of a (Banach) contraction, i.e. a mapping $f : X \to X$ satisfying (for some constant $K < 1$) $d(f(x), f(y)) \leq K \cdot d(x, y)$ for all $x, y \in X$.

In general, in fixed point theory one asks three questions: (1) does the type of mapping under consideration have a (possibly unique) fixed point at all, (2) is there an iteration converging to a/the fixed point and if so (3) is there an effective rate of convergence and what parameters does this rate of convergence depend on?

For Banach contractions (and complete metric spaces) these questions are answered quite easily and given a contraction f (with a suitable $K < 1$) and a starting point x the rate of convergence for the Picard iteration only depends on K, some bound b on $d(x, f(x))$ and the ε expressing how good an approximation to the (unique) fixed point one wants to obtain.

In [30], Edelstein extends this result (still in the context of complete metric spaces) to contractive mappings and proves that if for some starting point x_0

[1] Available at www.minlog-system.de.

the Picard iteration has a convergent subsequence, then this sequence converges to the unique fixed point of f, albeit without providing an effective rate of convergence. Using techniques from proof mining, an effective rate of convergence was extracted from Edelstein's proof in [81]. In [104], Rakotch considers contractive mappings with a modulus of contractivity α, proves that for such mappings every Picard iteration converges to the unique fixed point and provides an explicit rate of convergence. Again, this rate of convergence only depends on the modulus of contractivity α, some bound b on $d(x, f(x))$ and ε.

In [106], Rhoades surveys various generalizations of the notions of contractions and contractive mappings and the corresponding fixed point theorems, again in the context of complete metric spaces (or lack thereof illustrated by counterexamples). The study of fixed point theorems and counterexamples for the various types of contractive mappings is continued in [56, 24].

In [56], Kinces and Totik prove a fixed point theorem for the so-called generalized p-contractive mappings:

$$x \neq y \rightarrow d(f^p(x), f^p(y)) < diam\{x, y, f^p(x), f^p(y)\},$$

where the integer p is fixed. Among the variants of contractive mappings considered this is the most general notion for which one can prove a fixed point theorem. In [56], Kinces and Totik give an ineffective proof of a fixed point theorem for generalized p-contractive mappings. Recently, using proof mining to analyze the ineffective proof given by Kinces and Totik, Briseid obtained the first effective version of this fixed point theorem in his master's thesis[18].

A different generalization of contractive mappings are asymptotic contractions. In [59], Kirk proves the following theorem:

Theorem 4.2 (Kirk[59]). *Let (X, d) be a complete metric space and let $f : X \rightarrow X$ be a continuous[2] asymptotic contraction for which the mappings ϕ_n in Definition 4.1 also are continuous. Assume also that some orbit of f is bounded. Then f has a unique fixed point $x \in X$, and moreover the Picard sequence $(f^n(x))_{n=1}^{\infty}$ converges to z for each $x \in X$.*

The proof of Kirk's fixed point theorem for asymptotic contractions is highly non-elementary, as it relies on embedding the space (X, d) and the mapping f into a corresponding Banach space ultrapower (over some nontrivial ultrafilter $\mathcal{U}$) $(\tilde{X}, \tilde{d})$, where $\tilde{d}$ is the metric on $\tilde{X}$ inherited from the ultrapower norm $\|\cdot\|_{\mathcal{U}}$ in $\tilde{X}$. Using f to define a mapping $\hat{f}$ on $(\tilde{X}, \tilde{d})$ one obtains a unique fixed point $\tilde{z}$ for $\hat{f}$, from which one then in turn obtains a fixed point z for f in (X, d).

In [54, 1], elementary proofs are given of Kirk's fixed point theorem, though the proof in [54] strengthens the requirement of f being continuous to f being uniformly continuous. Neither proof provides explicit rates of convergence towards the fixed point z for the Picard iteration $(f^n(x))_{n=1}^{\infty}$. The analysis below

[2] In [54, 1], it is discussed that the requirement that f is continuous is a necessary condition for Kirk's fixed point theorem. By an oversight the requirement was left out in the original statement of Kirk's fixed point theorem in [59].

was based on Kirk's original proof and obtained before the elementary proofs in [54, 1] were known to the author. Strictly speaking, for the metatheorems from Chapter 3 to be applicable and to guarantee that effective bounds can be extracted one would need a proof that uses only elementary means (i.e. no ultrapower methods). In this particular case, it turned out that analysing the mathematical concepts involved (by proof-theoretic means) provided enough mathematical insight to produce a quantitative version of Kirk's fixed point theorem. In general, when 'mining' a nonstandard analytical proof one may consider the proof interpretations for nonstandard theories developed by Avigad in [2].

The first step of the analysis is to make the computational content of the theorem explicit. It will turn out, that a modulus of continuity is not needed although the theorem requires the asymptotic contraction to be continuous. However, continuity is only needed to prove the existence of a fixed point, while convergence towards an existing fixed point can be proven without it. Even more remarkable is the fact that in order to obtain an effective rate of convergence towards the unique fixed point an *ineffective* proof of the existence of the unique fixed point suffices.

Already Kreisel remarked that lemmas of a certain logical form do not contribute to the computational content of a theorem[3]. For purely universal formulas this result is established easily, as purely universal are their own functional interpretation and hence need no realizers. Therefore purely universal formulas may freely be added as axioms to the theory (or simply as an additional premise to the theorem) without contributing to the computational content of a given proof. In [65], the class of formulas that may freely be added as axioms is extended to formulas that may additionally include existential quantifiers that are bounded by a suitable closed term. Such formulas have a trivial monotone functional interpretation using that very same closed term. This class of "admissible" axioms is further extended in [77, 40] (also see the discussion in Chapter 3). In particular, the statement that a given function f has a fixed point can be shown to fall into this extended class of formulas.

Putting aside the computational content of the definition of asymptotic contractions for a while, we must make explicit the meaning of the premise that some orbit of f is bounded (more precisely: that some iteration sequence is bounded) and of the conclusion that for every $x \in X$ every Picard iteration $(f^n(x))_{n=1}^{\infty}$ converges to the unique fixed point of f.

One easily shows that if f is an asymptotic contraction and some iteration sequence is bounded, then every iteration sequence is bounded. Thus for the premise we require an explicit (integer) bound on the iteration sequence, i.e. we write:

$$\forall m, n \in \mathbb{N}(d(f^n(x), f^m(x)) \leq b),$$

where the starting point $x \in X$ of the iteration sequence and the bound b both

[3]For example, in [87], p.177, Kreisel writes "[I]t is hard to imagine, taking examples from the literature, what more we know about $\forall x A(x)$ [with decidable A] when it has been proved constructively than when it has been proved non-constructively".

are assumed to be given as parameters.

The conclusion then consists of two parts: uniqueness and convergence towards the fixed point, which we already know to exist by Kirk's original proof of his fixed point theorem. Proving uniqueness from asymptotic contractivity is easy, so we only express that given the unique fixed point z every iteration sequence converges towards z

$$\forall k \in \mathbb{N} \exists M \in \mathbb{N} \forall m \geq M (d(f^m(x), z) \leq 2^{-k}).$$

Even without relying on Kirk's fixed point theorem one may start out by proving convergence to an assumed fixed point and then later show, albeit only ineffectively, that every iteration indeed converges to a fixed point. At this point, extracting a bound on '$\exists M$' we may expect the bound to depend on (1) the starting point x, (2) the function f, (3) parameters expressing the asymptotic contractivity of f, (4) a bound b on the diameter of the iteration sequence $(f^n(x))_{n=1}^{\infty}$, (5) the fixed point z and (6) the parameter k expressing the desired "goodness" of approximation to the fixed point z.

Still, one problem remains with the theorem in its current form. The conclusion, a $\forall\exists\forall$-statement, is of a too complex logical form to allow the extraction of a bound. In general, one can only guarantee the extractability of bounds from theorems of the form $\forall\exists A_{qf}$, where A_{qf} is quantifier-free. Already for the class of $\forall\exists\forall A_{qf}$-formulas, with all quantifiers ranging over the natural numbers, there exist counterexamples in the form of e.g. the halting problem.

In the analysis of Edelstein's fixed point theorem for contractive mappings in [81] this problem is circumvented by the following observation: A contractive mapping f is in particular non-expansive, i.e. $\forall x, y(d(f(x), f(y)) \leq d(x, y))$. Therefore, once an iterate $f^m(x)$ is ε-close to the unique fixed point all further iterates are also ε-close (or closer), i.e. convergence towards the fixed point is monotone. Hence, it suffices to consider the following weaker conclusion:

$$\forall k \in \mathbb{N} \exists M \in \mathbb{N} (d(f^M(x), z) \leq 2^{-k}),$$

which is of a simple enough form to allow the extraction of a bound via functional interpretation. This kind of weakening of convergence statements due to a mapping being implicitly or explicitly nonexpansive, or rather, the general tactic of eliminating innermost universal quantifiers due to the monotonicity of the property under consideration also turns out to be useful in other situations.

This trick does not work for asymptotic contractions, as such mappings need not be nonexpansive (see [54]) and hence convergence is not necessarily monotone. Therefore in our case similarly weakening the conclusion will only allow one to extract a bound on an M such that for some $m \leq M$ the iterate $f^m(x)$ is ε-close to the unique fixed point z. On the other hand, if we find one iterate $f^m(x)$ sufficiently close to the unique fixed point z, deciding how much further we need to iterate to be sure that from then on iterates stay close to z depends on *how close* to z the iterate $f^m(x)$ actually is. Ironically, the closer the iterate $f^m(x)$ is to the fixed point z, the longer we may have to continue the iteration to guarantee staying close to z. If the function f additionally is nonexpansive (actually, the slightly weaker notion of 'weakly quasi-nonexpansive' suffices) the bound is a rate of convergence in the usual sense. Recently in [19], Briseid through a subtle, non-trivial argument obtained a full effective rate of convergence for the general case, the proof of which uses the effective results to be given below in an essential way.

The next step of the analysis is to put the notion of asymptotic contractions into a suitable form for the extraction of computable bounds. The notion of asymptotic contractions contains two parameters, a (continuous) function ϕ and a sequence of (continuous) functions ϕ_n together with three properties:

$$\forall s > 0(\phi(s) < s), \quad (1)$$
$$\forall x, y \in X \forall n \in \mathbb{N}(d(f^n(x), f^n(y)) \leq \phi_n(d(x, y))), \quad (2)$$
$$\phi_n \to \phi \text{ uniformly on the range of } d. \quad (3)$$

Starting with property (1), functional interpretation will ask for a witness – parametrized in s – for the inequality, i.e.

$$\exists \eta : (0, \infty) \to (0, \infty) \forall s > 0(\phi(s) + \eta(s) \leq s).$$

The requirements to this witness will undergo two further refinements: First, since we are interested in extracting bounds, we will require η to be its own majorant, i.e.

$$\exists \eta : (0, \infty) \to (0, \infty) \forall l > 0 \forall s > l(\phi(s) + \eta(l) \leq s).$$

Second, since in the main theorem Kirk requires that the iteration sequence $(f^n(x))_{n=1}^{\infty}$ is bounded, it will suffice to have an η^b witnessing the inequality for intervals $[l, b]$, where b is a bound on the diameter of $(f^n(x))_{n=1}^{\infty}$.

Thus the final version of property (1) is

$$\exists \eta^b : (0, b] \to (0, \infty) \forall l \in (0, b] \forall s \in [l, b](\phi(s) + \eta(l) \leq s), \quad (1')$$

where η^b then will be an additional parameter for the extracted bound.

Next, we consider property (3). More formally, (the functional interpretation of) this property can be expressed as follows:

$$\exists \beta : (0, \infty) \to \mathbb{N} \forall \delta > 0 \forall s \in rg(d) \forall m \geq \beta(\delta)(|\phi(s) - \phi_m(s)| \leq \delta),$$

where $rg(d)$ is the range of d. Again, it will suffice to witness the uniform convergence on intervals $[l, b]$ and instead of convergence towards a limit, we rewrite the property using Cauchy convergence. Thus we will require a modulus of convergence $\beta^b : (0, b] \times (0, \infty) \to \mathbb{N}$ s.t.

$$\forall \varepsilon > 0 \forall s \in [l, b] \forall m, n \geq \beta_l^b(\varepsilon)(|\phi_n(s) - \phi_m(s)| \leq \varepsilon), \quad (3')$$

where β_l^b is an abbreviation for $\beta^b(l, \cdot)$.

In a similar way, we will modify property (2) to be monotonous and ask only that it holds on intervals $[\varepsilon, b]$. The alternative definition of the notion of asymptotic contractions we arrive at is the following:

Definition 4.3 (**[39]**). *A function $f : X \to X$ on a metric space (X, d) is called an asymptotic contraction if for each $b > 0$ there exist moduli $\eta^b : (0, b] \to (0, 1)$ and $\beta^b : (0, b] \times (0, \infty) \to \mathbb{N}$ and additionally the following hold:*
(1) there exists a sequence of functions $\phi_n^b : (0, \infty) \to (0, \infty)$ s.t. for all $x, y \in X$, for all $\varepsilon > 0$ and for all $n \in \mathbb{N}$

$$b \geq d(x, y) \geq \varepsilon \Rightarrow d(f^n(x), f^n(y)) \leq \phi_n^b(\varepsilon) \cdot d(x, y),$$

(2) for each $0 < l \leq b$ the function $\beta_l^b := \beta^b(l, \cdot)$ is a modulus of uniform convergence for ϕ_n^b on $[l, b]$, i.e.

$$\forall \varepsilon > 0\ \forall s \in [l, b]\ \forall m, n \geq \beta_l^b(\varepsilon)\ \big(|\phi_m^b(s) - \phi_n^b(s)| \leq \varepsilon\big),$$

and (3) defining $\phi^b := \lim\limits_{n\to\infty} \phi_n^b$, then for each $0 < \varepsilon \leq b$ we have $\phi^b(s) + \eta^b(\varepsilon) \leq 1$ for each $s \in [\varepsilon, b]$.

Note, that whereas Kirk's definition mainly focuses on the functions ϕ and ϕ_n, here they are only required to exist for each $b > 0$ and the main focus is on the modulus of convergence β^b and the witness η^b for the inequality.

The alternative definition of asymptotic contractions contains Kirk's definition, as can be seen by the following argument:

Definition 4.4 (**[39]**). *Let $\phi : [0, \infty) \to [0, \infty)$, a sequence of continuous functions $\phi_n : [0, \infty) \to [0, \infty)$ and $b > 0$ be given. Define:*

$$\tilde{\phi}(s) := \tfrac{\phi(s)}{s} \text{ for } s \in (0, \infty), \qquad \tilde{\phi}_n(s) := \tfrac{\phi_n(s)}{s} \text{ for } s \in (0, \infty),$$
$$\phi^b(s) := \sup_{t\in[s,b]} \tilde{\phi}(t) \text{ for } s \in (0, b], \qquad \phi_n^b(s) := \sup_{t\in[s,b]} \tilde{\phi}_n(t) \text{ for } s \in (0, b].$$

Proposition 4.5 (**[39]**). *Let ϕ and ϕ_n be as in Definition 4.1 and let $\tilde{\phi}, \tilde{\phi}_n, \phi^b$ and ϕ_n^b be as in the above definition. Then*

- *$\tilde{\phi}$ and $\tilde{\phi}_n$ are continuous on $(0, \infty)$, $\tilde{\phi}(s) < 1$ for all $s \in (0, \infty)$ and the sequence $\tilde{\phi}_n$ converges uniformly to $\tilde{\phi}$ on $[l, \infty)$ for each $l > 0$,*
- *ϕ^b and ϕ_n^b are continuous on $(0, b]$, $\phi^b(s) < 1$ for all $s \in (0, b]$ and the sequence ϕ_n^b converges uniformly to ϕ^b on $[l, b]$ for each $0 < l \leq b < \infty$.*

Remark (**[39]**). *The moduli η^b, β^b may equivalently be given as functions $\eta^b : \mathbb{N} \to \mathbb{N}$ and $\beta^b : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, where real numbers are approximated from below by suitable rational numbers 2^{-n}. Given $b > 0$, if ϕ and a modulus β for ϕ_n (ϕ, ϕ_n as in Kirk's definition) are given as computable number-theoretic functions, then η^b and β^b are effectively computable in b.*

Proposition 4.6 (**[39]**). *If a function $f : X \to X$ on a metric space (X, d) is an asymptotic contraction (in the sense of Kirk) with moduli ϕ, ϕ_n, then the function f is an asymptotic contraction with suitable moduli η^b, β^b for every $b > 0$.*

For the extracted bound, we will be able to do completely without the functions ϕ and ϕ_n, as the following proposition shows:

Proposition 4.7 (**[39]**). *Let (X, d) be a metric space, let f be an asymptotic contraction and let $b > 0$ and η^b, β^b be given. Then for every $\varepsilon > 0$ and every $k \geq K(\eta^b, \beta^b, \varepsilon) = \beta^b_\varepsilon(\frac{\eta^b(\varepsilon)}{2})$ the following holds:*

$$b \geq d(x, y) \geq \varepsilon \Rightarrow d(f^k(x), f^k(y)) \leq (1 - \frac{\eta^b(\varepsilon)}{2}) \cdot d(x, y).$$

This proposition shows that given $d(x, y) \geq \varepsilon$ for some $\varepsilon > 0$, some iterate f^K of f almost behaves like a contractive mapping on such x, y. For convenience, we will omit the superscript b on η^b, β^b for the rest of this chapter.

The idea of the proof of the quantitative version of Kirk's fixed point theorem is now similar to the existing fully quantitative proof of a fixed point theorem for contractive mappings given in [81]. First, we produce (a variant of) a modulus of uniqueness, second, we obtain (a variant of) a modulus of uniform asymptotic regularity. Combining these modulis we obtain a variant of a rate of convergence towards the unique fixed point z for the iteration $(f^n(x))_{n=1}^\infty$. As discussed above, convergence towards the fixed point is not necessarily monotone for asymptotic contractions, so in the general case we only obtain an effective bound M on *some* $m \leq M$ s.t. x_m is close to the unique fixed point z. Such an effective bound has recently been termed a modulus of maximum proximity by Briseid[18].

A modulus of uniqueness usually is a function that for every $\varepsilon > 0$ provides a $\delta > 0$ such that $d(x, y) \leq \varepsilon$ if $d(x, f(x)), d(y, f(y)) \leq \delta$. In [39] we obtain the following variant:

Lemma 4.8 (**[39]**). *Let (X, d) be a metric space, let f be an asymptotic contraction and let $b > 0$ and η, β be given. Then for every $b \geq \varepsilon > 0$, for all $n \geq N$ and all $x, y \in X$ with $d(x, y) \leq b$*

$$d(x, f^n(x)), d(y, f^n(y)) \leq \delta \Rightarrow d(x, y) \leq \varepsilon,$$

where $\delta(\eta, \varepsilon) = \frac{\eta(\varepsilon) \cdot \varepsilon}{4}$ and $N(\eta, \beta, \varepsilon) = \beta_\varepsilon(\frac{\eta(\varepsilon)}{2})$.

For Edelstein-Rakotch-contractive mappings the next step is to find (for any given $\delta > 0$) an N s.t. $d(x_m, f(x_m)), d(x_n, f(x_n)) \leq \delta$ for all $m, n \geq N$, where x_m and x_n are the respective elements of the sequence $(f^n(x))_{n=1}^\infty$ (actually, since contractive mappings are nonexpansive finding *some* N such that $d(x_N, f(x_N)) \leq \delta$ suffices). The function producing such an N for any $\delta > 0$ is a modulus of uniform asymptotic regularity. Then using the modulus of uniqueness we deduce that for any $\varepsilon > 0$ we may find an N such that for $m, n \geq N$ the iterates x_m, x_n satisfy $d(x_m, x_n) \leq \varepsilon$, and hence the sequence $(f^n(x))_{n=1}^\infty$ is a Cauchy sequence. Since one easily proves using the continuity of f (which is implied by its contractivity) and the contractivity of f itself that the limit z of $(f^n(x))_{n=1}^\infty$ is the unique fixed point of f the two moduli combined provide a rate of convergence towards the unique fixed point.

For asymptotic contractions we only have a variant of a modulus of uniqueness and hence must produce the corresponding variant of a modulus of asymptotic regularity. As mentioned above, combining these we can only obtain a bound on an M such that for some $m \leq M$ the desired property holds. We obtain the following variant of a modulus of asymptotic regularity:

Lemma 4.9 ([**39**]). *Let (X, d) be a metric space, let f be an asymptotic contraction and let $b > 0$ and η, β be given. Then for every $\delta > 0$, for every $x_0 \in X$ s.t. $\{x_n\}$ is bounded by b and for every N there exists an $m \leq M$, s.t.*

$$d(x_m, f^N(x_m)) < \delta,$$

where $M(\eta, \beta, \delta, b) = k \cdot \lceil (\frac{lg(\delta) - lg(b)}{lg(1 - \frac{\eta(\delta)}{2})}) \rceil$ *with* $k = \beta_\delta(\frac{\eta(\delta)}{2})$.

Combining these two moduli we obtain our first main result:

Lemma 4.10 ([**39**]). *Let (X, d) be a metric space, let f be an asymptotic contraction and let $b > 0$ and η, β be given. Assume that f has a (unique) fixed point z. Then for every $\varepsilon > 0$ and every $x_0 \in X$ s.t. $\{x_n\}$ is bounded by b and $d(x_n, z) \leq b$ for all n there exists an $m \leq M$ s.t.*

$$d(x_m, z) \leq \varepsilon,$$

where $M(\eta, \beta, \varepsilon, b) = k \cdot \lceil (\frac{lg(\delta) - lg(b)}{lg(1 - \frac{\eta(\delta)}{2})}) \rceil$, $k = \beta_\delta(\frac{\eta(\delta)}{2})$, $\delta = \frac{\eta(\varepsilon) \cdot \varepsilon}{4}$.

In this lemma we obtain a bound M on some $m \leq M$ such that the iterate x_m is close to z. We mentioned earlier that the extracted bound may in the worst case depend on the starting point x, the function f, parameters expressing the asymptotic contractivity of f, a bound b on the diameter of the iteration sequence and the ε expressing the "goodness" of the the approximation to the fixed point. The above bound indeed depends on ε, the bound b and the moduli η, β, but not on x and f in any other way. This can be explained from the metatheorems in Chapter 3. Without going into the technical details, this is because in the context of the theorem the bound b is all we need to form suitable majorants for x and f.

The fact that f has a fixed point z at all (which then trivially is unique) is still an assumption. On the other hand, for the convergence towards the fixed point we need not assume that f is continuous. One may ineffectively show that the iteration of every *continuous* asymptotic contraction is a Cauchy sequence and converges to a fixed point z. The proof of this in [39]) is ineffective because it relies on a case distinction between $d(x_m, f^N(x_M)) =_{\mathbb{R}} 0$ and $d(x_m, f^N(x_M)) \neq_{\mathbb{R}} 0$. In general, if we only consider an abstract metric space (X, d) we have no access to actually computing $d(x_m, f^N(x_M))$. If we have a concretely represented metric space, we may compute the real number for $d(x_m, f^N(x_M))$, but then deciding $d(x_m, f^N(x_M)) =_{\mathbb{R}} 0$ is the problem. Equality for the real numbers – in their standard representation as Cauchy sequences of rational numbers – is undecidable. Still, for continuous asymptotic contractions the above lemma combined with Proposition 4.6 provides a quantitative version of Kirk's fixed point theorem:

Theorem 4.11 (**[39]**). *Let (X, d) be a complete metric space, let f be a continuous asymptotic contraction and let $b > 0$ and η, β be given. If for some $x_0 \in X$ the sequence $\{x_n\}$ is bounded by b then f has a unique fixed point z, $\{x_n\}$ converges to z and for every $\varepsilon > 0$ there exists an $m \leq M$ s.t.*

$$d(x_m, z) \leq \varepsilon,$$

where M is as in Lemma 4.10.

Note that in Lemma 4.10 it is only assumed that a fixed point exists. As mentioned earlier, the assumption that f has a fixed point does not contribute to the computational content of the proof. Still, we additionally need to assume that the given fixed point z is not too far away from the iteration sequence, i.e. that $\forall n(d(z, x_n) \leq b)$. In Theorem 4.11 we show that a fixed point exists and as the existence of the fixed point does not contribute to the rate of convergence, it is sufficient that this is proved ineffectively. As we actually show that the fixed point is the limit of the iteration sequence, this implies that $\forall n(d(z, x_n) \leq b)$, so that this additional premise from Lemma 4.10 may be dropped.

As mentioned earlier, if the function f additionally is weakly quasi-nonexpansive we obtain a full rate of convergence. The notion of weak quasi-nonexpansivity is introduced implicitly in [78] and is discussed further in [41]. The notion of weakly quasi-nonexpansive mapping has also been formulated independently under the name of J-type mapping in [34].

Definition 4.12. *A function $f : X \to X$ is called weakly quasi-nonexpansive if*

$$\exists p \in X(f(p) = p \wedge \forall x \in X d(f(x), p) \leq d(x, p)).$$

Corollary 4.13 (**[39]**). *Let (X, d) be a complete metric space, let f be a continuous, weakly quasi-nonexpansive asymptotic contraction and let $b > 0$ and η, β be given. If for some x_0 the sequence $\{x_n\}$ is bounded by b then f has a unique fixed point z, $\{x_n\}$ converges to z and for every $\varepsilon > 0$ and all $n \geq M$*

$$d(x_n, z) \leq \varepsilon,$$

where $M(\eta, \beta, \varepsilon, b)$ is as in Lemma 4.10 and moreover M is a rate of convergence for $\{x_n\}$.

In [18], the bound M is improved numerically by eliminating the dependency on the modulus of uniqueness:

Theorem 4.14 (**Briseid, [18]**). *Let (X, d) be a complete metric space, let f be a continuous asymptotic contraction and let $b > 0$ and η^b, β^b be given. If for some $x_0 \in X$ the sequence (x_n) is bounded by b then f has a unique fixed point z, (x_n) converges to z and for every $\varepsilon > 0$ there exists an $m \leq M_2$ such that*

$$d(x_m, z) \leq \varepsilon,$$

where

$$M_2(\eta^b, \beta^b, \varepsilon, b) = k \lceil \frac{lg(\varepsilon) - lg(b)}{lg(1 - \frac{\eta^b(\varepsilon)}{2})} \rceil$$

with $k = \beta^b_\varepsilon(\frac{\eta^b(\varepsilon)}{2})$.

The proof uses the following beautiful application of the pigeonhole principle: By Theorem 4.11 we know that (x_n) converges to the unique fixed point z and hence for every $l \in \mathbb{N}$ there is an N such that $d(x_n, z) \leq 2^{-l}$ for every $n \geq N$. Using the variant of asymptotic regularity proved in Lemma 4.9 above we know that for every $\varepsilon > 0$ there is an $m \leq M_2$ such that $d(x_m, f^N(x_m)) \leq \varepsilon$ for every N. Thus, given 2^{-l} and the corresponding N and using the triangle inequality, we get that $d(x_m, z) \leq \varepsilon + 2^{-l}$ for some $m \leq M_2$. Since there are infinitely many $l \in \mathbb{N}$ and only finitely many $m \leq M_2$, for some $m \leq M_2$ necessarily $d(x_m, z) \leq \varepsilon + 2^{-l}$ for infinitely many l and hence $d(x_m, z) \leq \varepsilon$.

As mentioned earlier, Briseid also obtained a full rate of convergence for asymptotic contractions in [19]. There is not yet a proof-theoretical, metamathematical explanation for either of these until now rather ad-hoc improvements. Finding such an explanation would be highly desirable, as this would almost certainly produce new insights into fixed point theory in particular and applications of proof mining in general.

Chapter 5

Future Work

The area of proof mining as a subdivision of mathematical logic and with its many proven applications to mathematics and computer science is a very promising field for future research. Two particular lines of future research go hand in hand: carrying out further concrete case studies in functional analysis and developing further metatheorems for analysing proofs in functional analysis. As is demonstrated by the many applications of the metatheorems presented in this thesis to the extraction of effective bounds from ineffective proofs in approximation theory and fixed point theory, the potential of this line of research is far from exhausted. At the same time, concrete case studies have regularly inspired significant generalizations of the existing metatheorems, such as the development of the a-majorization relation documented in Chapter 3. Moreover, for fixed point theory the uniformity results derived from the extraction of *uniform* bounds (using monotone functional interpretation) can also be used to prove very general properties of spaces, such as using the existence of uniform bounds to establish the approximate fixed point property for certain classes of product spaces(see [80]). This demonstrates that the potential applications of proof mining go far beyond the mere constructivization of non-constructive proofs.

However, proof mining need not be restricted to functional analysis or even analysis in general. From the beginning Kreisel demonstrated that investigations of what additional information one could extract from a given proof were just as fruitful for e.g. number theory or algebra as for analysis. In the light of previous successes in analysing proofs in analysis with techniques described in this thesis, one should at least expect for these methods to be equally fruitful when applied to other areas of mathematics that employ analytical methods, such as e.g. (analytical) number theory or certain branches of combinatorics and geometry.

As a completely different domain for proof mining consider the following example: In his Phd-dissertation([55], M.Kauers developed an algorithm for deciding a large class of combinatorial equalities, such as e.g.

$$\sum_{k=1}^{n}\sum_{i=1}^{k}\frac{1}{i} = -n + (n+1)\sum_{k=1}^{n}\frac{1}{k},$$

and similar identities involving sums, products, polynomials, Fibonacci numbers etc. Deciding such an identity (for all $n \in \mathbb{N}$) corresponds to deciding zero-equivalence for a corresponding sequence $(f_n)_n$, defined by a system of recurrences. The idea of Kauers' algorithm is to find an N such that

$$f_n = f_{n+1} = \dots f_{n+N} = 0 \to f_{n+N+1} = 0.$$

Given an N one can check the implication, by transforming the problem into an algebraic problem(testing radical membership, which can be done using Gröbner bases). If one has found such an N, one merely needs to check whether $f_i = 0$ for $i = 0, \dots N$ to either prove the zero-equivalence of the sequence $(f_n)_n$ or produce a counterexample. Kauers shows that for every sequence of a given, very general class, there indeed always exists such an N, and thus one may find such an N by successively trying $N = 0, 1, 2, \dots$. As this is essentially a $\forall\exists$-statement, it should be possible to analyze the proof and extract some additional information on N. A bound on N would not only quantitatively strengthen the result, but it would also simplify the algorithm, as one then no longer would need to successively test for radical membership, but merely would need to check whether $f_i = 0$ for $i = 0, \dots N$.

In conclusion, the area of proof mining – from the early endeavors by Kreisel to the very recent metatheorems and applications discussed in this thesis – has already shown, and continues to show, enormous promise. The author is confident that further study of proof mining, both of general techniques and concrete applications, along the lines sketched above will be most fruitful.

Part II

Papers

Chapter 6

The Role of Quantifier Alternations in Cut Elimination

The paper *The Role of Quantifier Alternations in Cut Elimination* presented in this chapter has been published in the **Notre Dame Journal of Formal Logic**, vol. 46, No.2, pp. 365-371, 2005, and can be considered a follow-up paper to [37], in which some main result from the author's master's thesis[36] were published. The paper has been slightly reformatted for inclusion in this PhD-thesis.

The Role of Quantifier Alternations in Cut Elimination

Philipp Gerhardy

Abstract

Extending previous results from [36, 37] on the complexity of cut elimination for the sequent calculus **LK**, we discuss the role of quantifier alternations and develop a measure to describe the complexity of cut elimination in terms of quantifier alternations in cut formulas and contractions on such formulas.

6.1 Introduction

In this note we will present an extension of results on the complexity of cut elimination in the sequent calculus **LK**, first developed in [36] and subsequently published in [37]. There it is shown how the complexity of cut elimination primarily depends on the nesting of quantifiers in cut formulas and contractions on ancestors of such cut formulas. A more complicated proof of the role of quantifier nestings was first given by Zhang in [120].

In this note we extend the analysis and develop a measure that describes with sufficient accuracy the role of quantifier alternations in cut elimination. The measure will be slightly more complicated than the notion of nested quantifier depth, nqf, described in [37], but will generalize with similar ease to incorporate the role of contractions. An earlier, more complicated treatment of the role of quantifier alternations by Zhang can be found in [121]. Though leading to comparable results, in particular the measure of the cut complexity described in [121] is far more complicated than the one presented in this note. For another attempt at defining a measure capturing quantifier alternations, albeit without accompanying proof of cut elimination, see Visser[119]. Neither Zhang nor Visser treat the role of contractions in the complexity of cut elimination.

6.2 Previous results

Let **LK** be the sequent calculus as defined in [37], i.e. with multiplicative rules and with no implicit contractions. Let $|\cdot|$ denote the depth and $||\cdot||$ the size of formulas and proofs, in the latter case not counting weakenings and contractions. Let $nqf(\cdot), dqf(\cdot)$ and $cnqf(\cdot)$ denote the nested quantifier depth, the deepest quantified formula (informally, the largest number of propositional connectives one has to "peel off" to get to a quantifier) and the contracted nested quantifier depth respectively, as defined in [37].

In [37] the following results are proved:

Refined Reduction Lemma. *Let ϕ be an* **LK***-proof of a sequent* $\Gamma \vdash \Delta$ *with the final inference a cut with cut formula A. Then if for all other cut formulas B*

(i) $nqf(A) \geq nqf(B)$ *and* $dqf(\phi) = dqf(A) > dqf(B)$*, then there exists a proof* ϕ' *of the same sequent with* $dqf(\phi') \leq dqf(\phi) - 1$ *and* $|\phi'| \leq |\phi| + 1$.

(ii) $nqf(\phi) = nqf(A) > nqf(B)$ *and* $dqf(A) = 0$*, then there exists a proof* ϕ' *of the same sequent with* $nqf(\phi') \leq nqf(\phi) - 1$ *and* $|\phi'| < 2 \cdot |\phi|$.

If the cut formula A is atomic and both subproofs are cut free, then there is a cut free proof ϕ' *with* $|\phi'| < 2 \cdot |\phi|$.

Lemma 6.1. *Let ϕ be an* ***LK****-proof of a sequent* $\Gamma \vdash \Delta$*. If* $dqf(\phi) = d > 0$*, then there is a proof* ϕ' *of the same sequent with* $dqf(\phi') = 0$ *and* $|\phi'| \leq 2^d \cdot |\phi|$.

Lemma 6.2. *Let ϕ be an* ***LK****-proof of a sequent* $\Gamma \vdash \Delta$*. If* $dqf(\phi) = 0$ *and* $nqf(\phi) = d > 0$*, then there is a proof* ϕ' *of the same sequent with* $nqf(\phi') \leq d-1$ *and* $|\phi'| < 2^{|\phi|}$.

First Refined Cut Elimination Theorem. *Let ϕ be an* **LK***-proof of a sequent* $\Gamma \vdash \Delta$*. If* $nqf(\phi) = d > 0$*, then there is a proof* ϕ' *of the same sequent and a constant c, depending only on the propositional nesting of the cut formulas, so that* $nqf(\phi') \leq d - 1$ *and* $|\phi'| \leq 2^{c \cdot |\phi|}$.

Corollary 6.3. *Let ϕ be an* **LK***-proof of a sequent* $\Gamma \vdash \Delta$ *and let* $nqf(\phi) = d$*. Then there is a constant c, depending only on the propositional nesting of the cut formulas, and a proof* ϕ' *of the same sequent where* ϕ' *is cut free and* $|\phi'| \leq 2^{c \cdot |\phi|}_{d+1}$.

Contraction Lemma. *Let ϕ be an* **LK***-proof of a sequent* $\Gamma \vdash \Delta$*, with* $nqf(\phi) > cnqf(\phi)$ *then there is proof* ϕ' *of the same sequent with* $nqf(\phi') = cnqf(\phi')$ *and* $||\phi'|| \leq ||\phi||$*. As a consequence also* $|\phi'| \leq 2^{|\phi|}$

Second Refined Cut Elimination Theorem. *Let ϕ be an* **LK***-proof of a sequent* $\Gamma \vdash \Delta$*. Then there is a constant c depending only on the propositional nesting of the cut formulas and a cut free proof* ϕ' *of the same sequent where* $|\phi'| \leq 2^{c \cdot |\phi|}_{cnqf(\phi)+2}$.

The main work is to prove the Refined Reduction Lemma and the Contraction Lemma from which the remaining results follow easily.

To sum up, first it is shown that the complexity of cut elimination primarily depends on the nesting of quantifiers in cut formulas, while the elimination of the propositional connectives has a negligible contribution to the complexity of cut elimination. As mentioned above this result was also shown by Zhang in [120]. Moreover, if for a cut formula none of the direct ancestors have been contracted, then the cut can be eliminated with low complexity by a mere rearrangement of the proof that does not increase the size of the proof. Thus the non-elementary complexity of cut elimination was shown to depend only on the nested quantifier depth of cut formulas whose ancestors, of sufficient quantifier depth, also have been contracted.

6.3 Quantifier alternations

In section 3.4 of [37] it is discussed that blocks of $\forall, \wedge$-connectives, respectively $\exists, \vee$, can be eliminated together, and it is shown that eliminating such a block from a cut formula at most doubles the depth of the proof. In [37] the following lemma is proved:

Lemma 6.4. *Let ϕ be a proof of a sequent $\Gamma \vdash \Delta$ with the last inference a cut. Let the cut formula be constructed from formulas $B_1, \ldots B_n$ by the connectives $\forall$ and $\wedge$ only (resp. $\exists$ and $\vee$). Then we can replace that cut by a number of smaller cuts with cut formulas B_i. For the resulting proof ϕ' we have $|\phi'| < 2 \cdot |\phi|$.*

This lemma immediately suggests a bound on cut elimination based on the number of alternations between such blocks. We propose the following cut elimination strategy: first we eliminate as many outermost propositional connectives as possible, next we eliminate all outermost $\forall, \wedge$ and $\exists, \vee$ blocks. Repeating this we eventually arrive at a cut free proof. By the Refined Reduction Lemma and subsequently Lemma 6.1, both of which can easily be adapted to some measure of the number of quantifier alternations instead of nqf, we see that the first step, eliminating propositional connectives, is not critical for the complexity of cut elimination. However, defining a new appropriate complexity measure for this cut elimination strategy is not trivial, as can also be seen by the complicated measure defined by Zhang[121] in order to prove a comparable result.

We want to define a measure aqf, the alternating quantifier depth. First consider the following very naive approach: Let us restrict the logical connectives to $\forall, \exists, \wedge$ and $\vee$ and let us count the propositional connectives $\wedge, \vee$ as the quantifiers $\forall, \exists$. Defining the aqf as the number of alternations between quantifier blocks in cut formulas would not give the desired result. Alternations between propositional connectives $\wedge, \vee$, which can be eliminated easily, would be perceived as alternations between quantifiers $\forall, \exists$, which are expensive to eliminate. Thus this definition of aqf would lead to a bound on cut elimination much worse in such situations than the bound already achieved via the nested quantifier depth nqf.

In general it turns out to be difficult to define, inductively on the formula, a measure of cut complexity that correctly captures the role of quantifier alternations. The difficulty is to decide when to increase the alternating quantifier depth.

For example, when facing a formula composed of two subformulas and one of the propositional connectives, e.g. the connective $\vee$, it is non-trivial to decide or predict whether the connective is part of a block of propositional connectives, and hence relatively harmless with respect to the alternations already present in the two subformulas, or marks the beginning of an $\exists, \vee$ block and hence leads to an increase in the number of alternations.

Consider the formula $A :\equiv B \vee C$, where the subformula C is assumed to be purely $\forall, \wedge$:

- if B is purely $\exists, \vee$ then
  - $aqf(A)$ should be 0 - eliminate the $\vee$ (simple), then two blocks each without alternations remain
  - $aqf(\exists xA)$ should be 1 - all the $\exists, \vee$ constitute one block, the $\forall, \wedge$ block below constitutes the alternation
  - $aqf(\forall xA)$ should be 1 - eliminate the $\forall$(expensive), eliminate the $\vee$(simple), eliminate the two alternation free blocks
- if B has an outermost $\exists, \vee$ block and one $\forall, \wedge$ block below
  - $aqf(A)$ should be 1 - eliminate the $\vee$(simple) then one alternation remains
  - $aqf(\exists xA)$ should be 1 - all the $\exists, \vee$ constitute one block, the $\forall, \wedge$ blocks below constitute the alternation
  - $aqf(\forall xA)$ should be 2 - eliminate the $\forall$(expensive), eliminate the $\vee$(simple) and still one alternation remains in the subformula B

The example demonstrates the problem of deciding *inductively* on the formula when to increase the alternating quantifier depth. At the point of the propositional connective we might not yet have sufficient information to decide whether to increase or not. On the other hand postponing the decision until we meet the next quantifier requires information on the exact structure of the subformulas that may no longer be available.

The solution is to let the complexity measure mirror the intended cut elimination strategy. This leads to defining the measure aqf for the cut complexity *recursively* on the cut formula instead of inductively.

Definition 6.5. *We define aqf as follows:*

- *if A is atomic, purely $\exists, \vee$ or purely $\forall, \wedge$ then $aqf(A) = 0$*
- *if A is composed of formulas $B_1, \dots, B_n$ (each with outermost connective a quantifier) by propositional connectives only then $aqf(A) = max\{aqf(B_i)\}$*
- *if A(:= $\forall xC$ for some C) is composed of connectives $\forall, \wedge$ and formulas $B_1, \dots, B_n$ (each with outermost connective $\exists, \vee$) then $aqf(A) = max\{aqf(B_i)\}+1$*
- *if A(:= $\exists xC$ for some C) is composed of connectives $\exists, \vee$ and formulas $B_1, \dots, B_n$ (each with outermost connective $\forall, \wedge$) then $aqf(A) = max\{aqf(B_i)\}+1$*

Moreover we treat implication $B \to C$ as $\neg B \vee C$, and negation $\neg B$ simply flips the polarity of other connectives below, i.e. $\exists, \vee \mapsto \forall, \wedge$ and vice versa.

With this definition of aqf for formulas, we define aqf for proofs:

Definition 6.6. $aqf(\phi) := \sup\{aqf(A) : A \text{ is a cut formula in } \phi\}$

Also the notion of deepest quantified formula dqf defined in [37] can be adapted to aqf, yielding a version of the Refined Reduction Lemma with aqf instead of nqf. Now it is easy to show the following theorem:

Theorem 6.7. *Let ϕ be an* **LK***-proof of a sequent $\Gamma \vdash \Delta$ and let $aqf(\phi) = d$. Then there is a constant c, depending only on the propositional blocks of the cut formulas, and a proof ϕ' of the same sequent where ϕ' is cut free and $|\phi'| \leq 2^{c \cdot |\phi|}_{d+1}$.*

Proof: As discussed the Refined Reduction Lemma can easily be adapted to the measure aqf instead of nqf. It then follows one can adapt Lemma 6.1 and Lemma 6.2 to the measure aqf. The theorem then follows easily from the cut elimination strategy sketched above, and the adaptions of Lemma 6.1 and Lemma 6.2 to the measure aqf. □

The definition of $caqf$, the contracted alternating quantifier depth, is defined from aqf in the same way $cnqf$ is defined from nqf (see [37]). Thus also taking the role of contractions into account we get:

Theorem 6.8. *Let ϕ be an* **LK***-proof of a sequent $\Gamma \vdash \Delta$. Then there is a constant c depending only on the propositional blocks of the cut formulas and a cut free proof ϕ' of the same sequent where $|\phi'| \leq 2^{c \cdot |\phi|}_{caqf(\phi)+2}$.*

Proof: The theorem follows easily from the above theorem and the Contraction Lemma adapted to the measure aqf. □

In conclusion both theorems follow easily from the analysis of cut elimination presented in [37], in particular Lemma 6.4, and the cut elimination strategy described above. As mentioned above a comparable result is proved in [121] but with a far more complicated complexity measure and a more complicated proof.

Furthermore, as with the upper bounds on cut elimination presented in [37], the bounds aqf and $caqf$ are optimal with regards to Statman's lower bound example, i.e. the upper and the lower bound coincide. Conversely, one can say that every proof that yields an example of non-elementary cut elimination must use cut formulas with alternating quantifiers and contractions in a way similar to Statman's lower bound example.

Finally, the exponential bound on cut elimination in the case of pure $\forall, \wedge$-cuts, respectively $\exists, \vee$-cuts, that is stated in [37], follows as a special case from these bounds.

6.4 Comparison with the literature

In this section we will briefly discuss the measures for the number of quantifier alternations proposed by resp. Zhang[121] and Visser[119].

The measure defined by Zhang uses two formula classes δ and δ':

Definition 6.9. (Zhang[121]) *A formula B is in $\delta(A)$ iff*

- $A = B$*, or*
- $A = C \wedge D$ *and* $B \in \delta(C) \cup \delta(D)$*, or*
- $A = \forall x C(x)$ *and* $B \in \delta(C(t))$ *for any term* t*.*

Definition 6.10. (Zhang[121]) *A formula B is in $\delta'(A)$ iff*

- $B \in \delta(A)$*, and*
- *B is either a disjunction of two formulas, or*
- *B is an $\exists$-formula s.t. all terms occurring in B also occur in A*

Definition 6.11. (Zhang[121]) *The cut complexity $\rho(A)$ of a formula A is defined as a polynomial in w as follows:*

- *if $\delta'(A)$ is empty, then $\rho(A) := w$,*
- *if $\delta'(A) = \{B_i | 1 \leq i \leq n\}$ and there is a formula C s.t. $\forall x C(x) \in \delta(A)$ and $nqf(C) \geq nqf(B_i)$ for $1 \leq i \leq n$, then $\rho(A) := (\rho(B_1) \oplus \ldots \oplus \rho(B_n)) \cdot w$,*
- *if $A = B \wedge C$, $\delta'(A) = \{B_i | 1 \leq i \leq n\}$ and there is no formula C s.t. $\forall x C(x) \in \delta(A)$ and $nqf(C) \geq nqf(B_i)$ for $1 \leq i \leq n$, then $\rho(A) := (\rho(B) \oplus \rho(C)) + 1$,*
- *if $\delta'(A) = \{A\}$, then $\rho(A) := \rho(\neg A)$*

where $\oplus$ is the operation of summing two polynomials by raising them to the same degree and then taking the pointwise maximum over their coefficients.

The definition is somewhat similar to the definition of the alternating quantifier depth aqf presented in this note, as the degree of the cut complexity polynomial corresponds to our notion of aqf.

In detail, the first item in the definition covers the case when the formula is atomic, purely $\forall, \wedge$ or (via the fourth item) purely $\exists, \vee$. The second item corresponds to eliminating a $\forall, \wedge$ block (or via the fourth item an $\exists, \vee$ block), and hence here the degree of the polynomial is increased. The third item corresponds to eliminating in-between propositional connectives, which only adds a constant to the polynomial.

The proof of cut elimination given the cut complexity polynomial above proceeds via several rather technical lemmas and uses an additional formula class δ^*.

Visser defines a measure "depth of quantifierchanges" via a three place function ϱ (p.281, [119]), where the first parameter is 0 when the formula under consideration occurs positively and 1 if it occurs negatively, the second parameter is

0 when we are in existential mode and 1 when we are in universal mode, while the last parameter is the formula under consideration.

The definition of ϱ is as follows:

Definition 6.12. (Visser[119]) *Let* $\varrho(A)$ *of a formula* A *be* $\varrho(0,0,A)$ *and let*

- $\varrho(i,j,A) := 0$ *of* A *is atomic,*
- $\varrho(i,j,B \wedge C) = \varrho(i,j,B \vee C) := \max\{\varrho(i,j,B), \varrho(i,j,C)\}$
- $\varrho(i,j,\neg B) := \varrho(1-i, 1-j, B)$
- $\varrho(0,j,B \to C) := \max\{\varrho(1,1-j,B), \varrho(0,j,C)\}$
- $\varrho(1,j,B \to C) := \max\{\varrho(0,j,B), \varrho(1,1-j,C)\}$
- $\varrho(i,0,\exists x B) := \varrho(i,0,B)$
- $\varrho(i,1,\exists x B) := \varrho(i,0,B) + 1$
- $\varrho(i,0,\forall x B) := \varrho(i,1,B) + 1$
- $\varrho(i,1,\forall x B) := \varrho(i,1,B)$

The merit of the measure defined by Visser is that it treats negation and implication directly. Contrary to the measure defined by Zhang and the measure aqf defined in this note, Visser's ϱ makes no distinctions for the propositional connectives as to whether they appear in existential or universal mode, i.e. below an existential or a universal quantifier. Thus the measure ϱ assigns the same "depth of quantifierchanges" to the formulas

$$\forall x(\forall y P(x,y) \vee \forall z Q(x,z))$$

and

$$\forall x(\forall y P(x,y) \wedge \forall z Q(x,z)).$$

However, one can show that the complexity of cut elimination for the two (cut) formulas is not the same, i.e. cut elimination for the second formula, which is purely $\forall, \wedge$, is simpler than cut elimination for the first formula $\forall x(\forall y P(x,y) \vee \forall z Q(x,z))$, which contains a disjunction.

Thus, although capturing the main ideas, namely that cut elimination mainly depends on quantifier alternations, the measure as it is defined in [119], is not optimal in all cases to estimate the complexity of cut elimination.

6.5 Acknowledgements

I am grateful to Lev Beklemishev, Ulrich Kohlenbach and Albert Visser for many useful discussions on the subject of this note.

Chapter 7

Extracting Herbrand Disjunctions by Functional Interpretation

The paper *Extracting Herbrand Disjunctions by Functional Interpretation* presented in this chapter has been published in the **Archive for Mathematical Logic**, vol. 44, pp. 633-644, 2005. The paper is joint work with U.Kohlenbach and has been slightly reformatted for inclusion in this PhD-thesis.

Extracting Herbrand Disjunctions by Functional Interpretation

Philipp Gerhardy, Ulrich Kohlenbach.

Abstract

Carrying out a suggestion by Kreisel, we adapt Gödel's functional interpretation to ordinary first-order predicate logic(PL) and thus devise an algorithm to extract Herbrand terms from PL-proofs. The extraction is carried out in an extension of PL to higher types. The algorithm consists of two main steps: first we extract a functional realizer, next we compute the β-normal-form of the realizer from which the Herbrand terms can be read off. Even though the extraction is carried out in the extended language, the terms are ordinary PL-terms. In contrast to approaches to Herbrand's theorem based on cut elimination or ε-elimination this extraction technique is, except for the normalization step, of low polynomial complexity, fully modular and furthermore allows an analysis of the structure of the Herbrand terms, in the spirit of Kreisel ([91]), already prior to the normalization step. It is expected that the implementation of functional interpretation in Schwichtenberg's MINLOG system can be adapted to yield an efficient Herbrand-term extraction tool.

7.1 Introduction

Herbrand's theorem states that for every proof in pure first-order logic without equality of a sentence $\exists x A_{qf}(x)$ (A_{qf} always denotes a quantifier-free formula), there is a collection of closed terms $t_1, \dots, t_n$ witnessing that proof, so that $\bigvee_{i=1}^{n} A_{qf}(t_i)$ is a tautology. Such a disjunction is called a Herbrand disjunction of A and the terms $t_1, \dots, t_n$ are called Herbrand terms. Herbrand's theorem easily generalizes to tuples of existential quantifiers $\exists \underline{x} A_{qf}(\underline{x})$, where $\underline{x} = x_1, \dots x_k$,[1] and via the Herbrand normal form A^H to arbitrary formulas A in prenex normal form. Moreover, it extends to open first order theories T (i.e. theories whose axioms are purely universal sentences), where then the disjunction is verifiable in T, i.e. $T \vdash \bigvee_{i=0}^{n} A^H(\underline{t}_i)$ (and even is a tautological consequence of a conjunction of finitely many closed instances of the non-logical axioms of T). First order logic with equality can be treated as the special case, where T is an open axiomatization of equality. For first order logic (with or without equality) the Herbrand terms are built up out of A-material (resp. A^H-material) only with possible help of some distinguished constant symbol c in case A (resp. A^H) does not contain any constant. For open first order theories T they may in addition contain some of the constants and function symbols occurring in the non-logical T-axioms used in the proof. For more details see e.g. [112, 22, 36].

[1] For notational simplicity we avoid below to write tuples.

There are both model-theoretic and proof-theoretic proofs of Herbrand's theorem. But whereas the former proofs are ineffective the latter provide a procedure for extracting Herbrand terms t_i from a given proof of A. The actual construction of Herbrand terms out of a given proof is of importance in the area of computational logic and has also been used in significant applications to mathematics (see [91, 94]).

The existing proof-theoretic approaches to Herbrand's theorem are based on cut elimination or related techniques like ε-elimination which involve global transformations of the given proof. In his review [90] of [112], G. Kreisel suggested the possibility of using Gödel's functional ('dialectica') interpretation FI ([44, 117]) to prove Herbrand's theorem. To our knowledge this suggestion has never been taken up in the literature and the present note aims at filling this lacuna: We give an extraction algorithm of Herbrand terms via functional interpretation in the variant developed in [112] which we from now on also call FI. The verifiability of the extracted disjunction as a tautology or T-provable disjunction is achieved by a simple model theoretic argument. As the case for open theories T immediately reduces (via the deduction theorem) to that of first order logic without equality PL, we only treat the latter.

From a given PL-proof of a sentence $\exists x A_{qf}(x)$, FI extracts a closed term t in an extension of typed λ-calculus by decision-by-case constants χ_A for each quantifier-free formula A of $\mathcal{L}$(PL). After computing the β-normal form $nf(t)$ of t, the Herbrand terms can be read off. The length of the resulting Herbrand disjunction is bounded by $2^{\#_\chi(nf(t))}$, where $\#_\chi(nf(t))$ is the total number of χ-occurrences in $nf(t)$.

The significance of this FI-based approach to the extraction of Herbrand terms is due to the following points:

1. FI has recently been successfully implemented by M.-D. Hernest ([47]) in H. Schwichtenberg's MINLOG system which also contains an efficient normalization tool ('normalization by evaluation', see [10]). We expect that this implementation can be adapted to yield a useful Herbrand-term extraction tool.

2. Suppose that in a PL-proof of (1) $\exists x\, A_{qf}(x)$ classical logic is only used to infer (1) from (2) $\forall x(A_{qf}(x) \to \bot) \to \bot$, where (2) is proved intuitionistically. Then already the original direct Gödel functional interpretation (i.e. without negative translation as a preprocessing step and also without Shoenfield's modification) can be used to extract a Herbrand disjunction for (1) which will in general (though not always[2]) be simpler than the detour through full classical logic. This is because the type levels will be lower resulting in a more efficient normalization and hence a shorter Herbrand disjunction.

[2]In the Statman example discussed below the original functional interpretation already creates as high types as the Shoenfield variant does. This is unavoidable here since the Statman example has the worst possible Herbrand complexity despite the fact that its form (2) is provable in intuitionistic logic.

3. When combined with known estimates ([3]) on the size of $nf(t)$ we immediately obtain bounds on Herbrand's theorem which match the most advanced estimates based on cut-elimination ([36, 37, 121]).

4. In [91] Kreisel discusses how to derive new results in mathematics by analysing the structure of Herbrand terms, e.g. growth conditions, extracted from a given proof. This has been carried out in connection with Roth's theorem by Luckhardt in [94]. Often it will be possible to read off some structural properties of the Herbrand terms already from the FI-extracted E-PL$^\omega$ term t *prior to normalization*, e.g. by analysing which constant and function symbols occur in the extracted term, thereby establishing bounds on the complexity or independence from parameters for the Herbrand terms prior to their actual construction via $nf(t)$.

7.2 An FI-based approach to Herbrand's Theorem

FI is usually applied to (appropriate formulations of) intuitionistic arithmetic (Heyting arithmetic) in all finite types. Already for the logical axioms and rules the proof of the soundness of FI relies on some minimal amount of arithmetic. Combined with negative translation FI extends to (higher type extensions of) Peano arithmetic (PA). In the following we will use Shoenfield's variant which achieves this in one step and denote this form by FI as well.

To apply FI to first-order predicate logic(PL), we will adapt the soundness proof from Shoenfield [112]. Shoenfield gives a soundness proof of FI for PA which for logical axioms and rules only uses properties of arithmetic to ensure the existence of decision-by-case terms for quantifier-free formulas. By explicitly adding decision-by-case constants χ_A for all quantifier-free formulas A in $\mathcal{L}$(PL) to the language of PL$^\omega$, we can re-use Shoenfield's proof for the soundness of FI of PL in E-PL$^\omega$:=PL extended to all finite types (based on extensionally defined equality).

We then can, for proofs of sentences $\exists x A_{qf}(x)$ in the language $\mathcal{L}$(PL), extract realizing terms t in the extended language E-PL$^\omega$. After normalizing the E-PL$^\omega$-term t one can read off from the normal form $nf(t)$ a collection of terms $t_1, \ldots, t_n$ for a Herbrand disjunction over A, where the t_i again are ordinary closed terms of PL without any higher type constructs and without the decision-by-case constants.

Remark. *At a first look one might think that the so-called Diller-Nahm version ([27, 26]) of Shoenfield's variant might be more suitable in connection with Herbrand's theorem: it avoids definitions by cases which depend on the prime formulas in favour of definition of case-functionals which do not depend on* A_{qf} *but only on cases* $x =_0 0$ *versus* $x \neq 0$*. However, our technique of eliminating all definitions by cases by explicitly writing out all cases as different terms does not distinguish between these two kinds of case-definitions. In addition to not* *being beneficial, the Diller-Nahm variant actually relies on a modest amount of arithmetic which is not available in our context of pure logic.*

We now describe the system of first-order predicate logic PL and its extension E-PL$^\omega$ to all finite types, in which our proof will be carried out.

First-order predicate logic PL

I. **The language $\mathcal{L}$(PL) of PL:**

As logical constants we use $\neg, \vee, \forall$. $\mathcal{L}$(PL) contains variables $x, y, z, \ldots$ which can be free or bound, and constants $c, d, \ldots$. Furthermore we have, for every arity n, (possibly empty) sets of function symbols $f, g, \ldots$ and predicate symbols $P, Q, \ldots$. Formulas and terms are defined in the usual way.

Abbreviations:

$A \to B :\equiv \neg A \vee B, A \wedge B :\equiv \neg(\neg A \vee \neg B), \exists x A(x) :\equiv \neg\forall x \neg A(x)$.

II. **Axioms of PL**

(i) $\neg A \vee A$

(ii) $\forall x A(x) \to A[t/x]$ (t free for x in $\forall x A(x)$)

III. **Rules of PL**

(i) $A \vdash B \vee A$ (expansion)

(ii) $A \vee A \vdash A$ (contraction)

(iii) $(A \vee B) \vee C \vdash A \vee (B \vee C)$ (associativity)

(iv) $A \vee B, \neg A \vee C \vdash B \vee C$ (cut)

(v) $A \vee B \vdash \forall x A(x) \vee B$ ($\forall$ -introduction), where x is not free in B.

Note. *As will be seen later, the degree of the terms extracted by FI depends on the* $\neg$*-depth of formulas. We treat only Shoenfield's calculus, but when translating other calculi for PL into Shoenfield's calculus, we extend Shoenfield's quantifier axioms and rules and the translation* $\exists x A(x) :\equiv \neg\forall x \neg A(x)$ *to blocks of quantifiers, i.e.*$\exists \underline{x} A(\underline{x}) :\equiv \neg\forall \underline{x} \neg A(\underline{x})$*, to avoid an artificial blow-up of the degrees when treating blocks of existential quantifiers.*

Note. *We assume w.l.o.g. that there exists at least one constant symbol, c, in our language, as Herbrand's theorem would fail otherwise.*

Extensional predicate logic in all finite types

The set $\mathbf{T}$ of all finite types is defined inductively:

$$(i)\ 0 \in \mathbf{T},\ (ii)\ \rho, \tau \in \mathbf{T} => \rho \to \tau \in \mathbf{T}$$

For convenience we write $0^n \to 0$ for $\overbrace{0 \to (0 \to (\ldots(0}^{n} \to 0)\ldots)$.

The language of E-PL$^\omega$

The language E-PL$^\omega$ is based on a many-sorted version PL$^\omega$ of PL which contains variables $x^\rho, y^\rho, z^\rho, \ldots$ and quantifiers $\forall^\rho, \exists^\rho$ for all types ρ. As constants E-PL$^\omega$ contains the constants $c, d, \ldots$ (at least one: c) of PL as constants of type 0, and the function symbols $f, g, \ldots$ of PL as constants of type $0^n \to 0$ for functions of arity n. Furthermore E-PL$^\omega$ contains decision-by-case constants χ_A of type $0^n \to 0 \to 0 \to 0$ for all quantifier-free formulas A in the original language $\mathcal{L}(\mathrm{PL})$, where n is the number of free variables in A. E-PL$^\omega$, moreover, contains a λ-abstraction operator. The predicate symbols of E-PL$^\omega$ are the predicate symbols of PL and equality of type 0 (denoted by $=_0$).

Higher type equality in E-PL$^\omega$ is defined extensionally over type 0 equality:

$$s =_\rho t :\equiv \forall x_1^{\rho_1}, \ldots, x_n^{\rho_n}(s\underline{x} =_0 t\underline{x}),$$

where $\rho = \rho_1 \to \ldots \to \rho_n \to 0$.

Formulas are defined in the usual way starting from prime formulas $s =_0 t$ and $P(t_1, \ldots, t_n)$.

Remark. *Below we often refer implicitly to the obvious embedding of* PL *into* E-PL$^\omega$*, where constants and variables of* PL *represented by their type* 0 *counterparts in* E-PL$^\omega$ *and (n-ary) function symbols of* PL *as constants of type* $0^n \to 0$*, in particular* PL *terms* $f(t_1, \ldots, t_n)$ *are represented by* $((\ldots(ft_1)\ldots)t_n)$*. Recall that the predicate symbols of* E-PL$^\omega$ *are those of* PL *plus* $=_0$*.*

Terms of E-PL$^\omega$

(i) constants c^ρ and variables x^ρ are terms of type ρ (in particular the constants $c, d, \ldots$ of PL are terms of type 0),

(ii) if x^ρ is a variable of type ρ and t^τ a term of type τ, then $\lambda x^\rho.t^\tau$ is a term of type $\rho \to \tau$,

(iii) if t is a term of type $\rho \to \tau$ and s is a term of type ρ, then (ts) is a term of type τ. In particular, if $t_1, \ldots, t_n$ are terms of type 0 and f is an n-ary function symbols of PL, then $((\ldots(ft_1)\ldots)t_n)$ is a term of type 0 which we usually will write as $f(t_1, \ldots, t_n)$.

Axioms and Rules of E-PL$^\omega$

(i) axioms and rules of PL extended to all sorts of E-PL$^\omega$,

(ii) axioms for β-normalization in the typed λ-calculus: $(\lambda x.t)s =_\rho t[s/x]$ for appropriately typed x, t and s,

(iii) equality axioms for $=_0$,

(iv) higher type extensionality:

$$E_\rho : \forall z^\rho, x_1^{\rho_1}, y_1^{\rho_1}, \ldots, x_k^{\rho_k}, y_k^{\rho_k} (\bigwedge_{i=1}^{k} (x_i =_{\rho_i} y_i) \rightarrow z\underline{x} =_0 z\underline{y}),$$

where $\rho = \rho_1 \rightarrow (\rho_2 \rightarrow (\ldots \rightarrow (\rho_k \rightarrow 0)\ldots))$,

(v) axioms for the constants $\chi_{A_{qf}}$: $A_{qf}(\underline{x}) \rightarrow \chi_{A_{qf}}\underline{x}yz =_0 y$ and $\neg A_{qf}(\underline{x}) \rightarrow \chi_{A_{qf}}\underline{x}yz =_0 z$, where $\underline{x}$ are the free variables of the quantifier-free formula A_{qf} of $\mathcal{L}(\text{PL})$.

Definition 7.1. *We define the type level $lv(t)$ of a term t inductively over the type of t as follows: $lv(0) := 0$ and $lv(\rho \rightarrow \tau) := max(lv(\tau), lv(\rho) + 1)$. The degree $dg(t)$ of a term t is then the maximum over the type levels of all subterms of t.*

Definition 7.2. *Let $\mathcal{M} = \{M, \mathcal{F}\}$ be a model for $\mathcal{L}(\text{PL})$. Then $\mathcal{M}^\omega = \{M^\omega, \mathcal{F}^\omega\}$ is the full set-theoretic type structure over M, i.e. $M^0 :\equiv M$, $M^{\rho\rightarrow\tau} :\equiv {M^\rho}_{M^\tau}$ and $M^\omega :\equiv \bigcup_{\rho \in T} M^\rho$. Constants, functions and predicates of $\mathcal{M}$ retain their interpretation under $\mathcal{F}$ in $\mathcal{F}^\omega$. λ-terms are interpreted in the obvious way. Furthermore, $\mathcal{F}^\omega$ defines the following interpretation of χ_A:*

For $\underline{a}, b, c \in M$ we define $[\chi_A]_{\mathcal{M}^\omega}\underline{a}bc := \begin{cases} b & \text{if } \mathcal{M} \models A_{qf}(\underline{a})^3 \\ c & \text{otherwise.} \end{cases}$

Proposition 7.3. *$\mathcal{M}^\omega$ is a model of* E-PL$^\omega$. *If A is a sentence of $\mathcal{L}(\text{PL})$ and $\mathcal{M}^\omega \models A$, then $\mathcal{M} \models A$.*

Proof. Obvious from the construction of $\mathcal{M}^\omega$. □

In the following $\exists x A_{qf}(x)$ will denote a closed formula. For open formulas one can replace each free variable with new distinct constants, carry out the extraction procedure and then reintroduce the variables to get a corresponding Herbrand disjunction for the open case.

Lemma 7.4. *If* PL $\vdash \exists x A_{qf}(x)$ *then* FI *extracts a closed term t^0 of* E-PL$^\omega$ *s.t.* E-PL$^\omega \vdash A_{qf}(t)$.
The proof of $A_{qf}(t)$ can actually be already carried out in the quantifier-free fragment qf-WE-PL$^\omega$ *(in the sense of* [117]*) of* WE-PL$^\omega$, *where the latter is the fragment of* E-PL$^\omega$ *which results by replacing the extensionality axioms by the quantifier-free weak rule of extensionality due to* [114] *(see also* [72]*).*

Proof. This is essentially Shoenfield's proof in [112]. The only two cases to note are the expansion rule and the contraction rule.

[3] More precisely, $\mathcal{M} \models A_{qf}(\underline{a})$ means that $A_{qf}(\underline{x})$ holds in $\mathcal{M}$ provided the free variables x_i get assigned the element $a_i \in M$.

If $B \vee C$ has been inferred from B by the expansion rule we need an arbitrary closed term of suitable type to realize C. Since we assumed there exists at least one constant c of type 0, we can, using lambda abstraction, construct closed terms $\lambda \underline{x}.c^0$ of suitable type to realize C.

For the contraction rule the argument is somewhat more involved: Let $A(\underline{a})$ be an arbitrary formula with $\underline{a}$ denoting the free variables of A. To each formula A Shoenfield assigns a formula $A^* \equiv \forall \underline{x} \exists \underline{y} A'(\underline{x}, \underline{y}, \underline{a})$, where A' is quantifier-free. The quantifier-free skeleton A_{qfs} of $A \in \mathcal{L}(\mathrm{PL})$ is the formula A with all quantifiers removed and distinct new variables substituted for the quantified variables of A, i.e. $A_{qfs}(\underline{b}, \underline{a})$, where $\underline{b}$ are the new variables and $\underline{a}$ are the original free variables of A. The formula A' is a substitution instance $A_{qfs}([\underline{x}, \underline{y}], \underline{a})$ of $A_{qfs}(\underline{b}, \underline{a})$, where $[\underline{x}, \underline{y}]$ denotes some tuple of terms which do not contain any constants but are built up exclusively out of $\underline{x}, \underline{y}$. These terms have been substituted for $\underline{b}$. For simplicity we will in the following consider only single variables x, y and a single parameter a, as the argument easily generalizes to tuples of variables.

To interpret the contraction rule $A \vee A \vdash A$ we have to produce a realizer for the conclusion

$$\forall x_3 \exists y_3 A'(x_3, y_3, a)$$

from realizers of the premise

$$\forall x_1, x_2 \exists y_1, y_2 (A'(x_1, y_1, a) \vee A'(x_2, y_2, a)),$$

where in general x_i, y_i will be of arbitrary type. However, the terms composed of x_i, y_i instantiating A_{qfs} to yield A' are of type 0, since A^* interprets the first order formula $A \in \mathcal{L}(\mathrm{PL})$. The functional interpretation of the premise yields closed terms t_1, t_2 s.t.

$$\forall x_1, x_2, a \big(A'(x_1, t_1 x_1 x_2 a, a) \vee A'(x_2, t_2 x_1 x_2 a, a)\big).$$

Substituting x_1 for x_2 gives

$$\forall x_1, a \big(A'(x_1, t_1' x_1 a, a) \vee A'(x_1, t_2' x_1 a, a)\big),$$

where $t_1' x_1 a := t_1 x_1 x_1 a$ and $t_2' x_1 a := t_2 x_1 x_1 a$.

Hence, after renaming x_3 in the conclusion into x_1, a term t_3 realizing y_3 (when applied to x_1, a) must satisfy:

$$t_3 x_1 a = \begin{cases} t_1' x_1 a & \text{if } A'(x_1, t_1' x_1 a, a) \\ t_2' x_1 a & \text{otherwise,} \end{cases}$$

i.e.

$$t_3 x_1 a = \begin{cases} t_1' x_1 a & \text{if } A_{qfs}([x_1, y](y/t_1' x_1 a), a) \\ t_2' x_1 a & \text{otherwise.} \end{cases}$$

This term t_3 can be defined via our decision-by-case constants for the quantifier-free skeleton A_{qfs} of A as follows:

$$t_3 := \lambda x_1, a, \underline{v}. \chi_{A_{qfs}}([x_1, y](y/t_1' x_1 a), a, t_1' x_1 a \underline{v}, t_2' x_1 a \underline{v}),$$

where $\underline{v}$ is a tuple of fresh variables of appropriate types such that $t_1' x_1 a \underline{v}$ is of type 0.

Hence it is sufficient to have decision-by-case constants χ_A for each quantifier-free formula A of $\mathcal{L}(\mathrm{PL})$. These have been explicitly added to the language of E-PL$^\omega$. □

Example. *As an example, consider the formula* $A \equiv \exists x \forall y (P(x) \vee \neg P(y))$. *The Shoenfield translation* A^* *of* A *is* $A^* \equiv \forall f \exists x \neg\neg(P(x) \vee \neg P(f(x)))$, *which is classically equivalent to* $\forall f \exists x (P(x) \vee \neg P(f(x)))$. *The matrix* $A' \equiv (P(x) \vee \neg P(f(x)))$ *is an instance of* $A_{qfs}(b_1, b_2) \equiv P(b_1) \vee \neg P(b_2)$, *namely* $A_{qfs}(x, f(x))$.

Functional interpretation will extract from a proof of A, *which necessarily must use the contraction rule at least once, a functional* Φ *realizing* x *in* f. *The term will also use some constant* c, *since* A *itself contains no constants. An obvious* Φ *is the following:*

$$\Phi(f) := \begin{cases} c & \text{if } P(c) \vee \neg P(f(c)) \\ f(c) & \text{otherwise.} \end{cases}$$

Lemma 7.5. *If* E-PL$^\omega \vdash A_{qf}(t)$ *and* $nf(t)$ *is the* β*-normal form of* t, *then* E-PL$^\omega \vdash A_{qf}(nf(t))$.

Proof. Since t reduces to $nf(t)$, we have E-PL$^\omega \vdash t =_\rho nf(t)$. □

Lemma 7.6. *If* t *is of type* 0, *closed and in* β*-normal form, then there exist closed terms* $t_1, \ldots, t_n \in \mathcal{L}(\mathrm{PL})$, *s.t.* $\mathcal{M}^\omega \models t = t_1 \vee \ldots \vee t = t_n$. *Moreover,* $n \leq 2^{\#_\chi(nf(t))}$, *where* $\#_\chi(nf(t))$ *is the total number of all* χ*-occurrences in* $nf(t)$.

Proof. Since t is of type 0, closed and in β-normal form and has only constants of degree ≤ 1 it contains no more λ-expressions: Assume there still is a λ-expression $\lambda x.r$ left and assume w.l.o.g. that it is not contained in any other λ-expression. Then if $\lambda x.r$ occurs with an argument $(\lambda x.r)s$ it could be further reduced, which contradicts that t is in normal form. If $\lambda x.r$ occurs without an argument it must be at least of type 1, and then since t is closed either $\lambda x.r$ must occur in another λ-expression, since the function symbols of PL only take arguments of type 0, or $t \equiv \lambda x.r$. But this contradicts that $\lambda x.r$ was not contained in any other term and that t was of type 0. Similarly, one infers that the function symbols f always occur with a full stock of arguments in t.

To read off the terms t_i by consider a tree constructed from t by "evaluating" the χ's : choose any outermost χ and build the left (resp. right) subtree by replacing the occurrence of the corresponding term $\chi(\underline{s}, t_1, t_2)$ in t with t_1 (resp. t_2). Continue recursively on the left and right subtrees until all χ's have been evaluated. Every path in the tree from the root to a leaf then represents a list of choices on the χ's and thus every leaf is a term $t_i \in \mathcal{L}(\mathrm{PL})$.

It follows trivially that $\mathcal{M}^\omega \models t = t_1 \vee \ldots \vee t = t_n$. As a simple estimate on the length n we get $n \leq 2^{\#_\chi(nf(t))}$. □

Theorem 7.7. *Assume that* PL $\vdash \exists x A_{qf}(x)$. *Then there is a collection of closed terms* $t_1, t_2, \ldots, t_n$ *in* $\mathcal{L}$(PL) *which can be obtained by normalizing a* FI *extracted realizer* t *of* $\exists x$ *s.t.* $\bigvee_{i=1}^{n} A_{qf}(t_i)$ *is a tautology. The terms* t_i *are built up out of the* A_{qf}*-material (possibly with the help of the distinguished constant* c *in case* A_{qf} *does not contain any constant). Moreover,* $n \leq 2^{\#_\chi(nf(t))}$.
The theorem also extends to tuples $\exists \underline{x}$ *of quantifiers.*

Proof. The theorem follows from the above propositions and lemmas. By the soundness of FI we can extract a closed term t in E-PL$^\omega$ realizing '$\exists x$'. We can assume that t consists exclusively of constants and function symbols for $\mathcal{L}$(PL) and some decision-by-case constants χ_B, restricted to quantifier-free formulas B built up from **predicates occurring in** A by means of propositional connectives. This restriction can be verified by a simple model-theoretic argument: give all predicates not occurring in A a trivial interpretation, e.g. interpret them as "always true", and replace decision-by-case expressions over such predicates by appropriate constants. In decision-by-case constants over combinations of predicates occurring and predicates not occurring in A, those not occurring in A can be absorbed.

We then normalize t to $nf(t)$ and read off the terms $t_1, \ldots, t_n$ from $nf(t)$ as in lemma 7.6. Let $\mathcal{M}$ be an arbitrary model of $\mathcal{L}$(PL), then $\mathcal{M}^\omega \models \bigvee_{i=1}^{n} A_{qf}(t_i)$. As the t_i are already closed terms of $\mathcal{L}$(PL), also $\mathcal{M} \models \bigvee_{i=1}^{n} A_{qf}(t_i)$. Since $\mathcal{M}$ was an arbitrary model, the completeness theorem for PL yields that also PL $\vdash \bigvee_{i=1}^{n} A_{qf}(t_i)$. Since $\bigvee_{i=1}^{n} A_{qf}(t_i)$ is quantifier-free it follows that it is a tautology (note that PL is predicate logic **without** equality).

The FI-extracted term t consists of A_{qf}-material, decision-by-case constants and λ-abstractions. The normal form $nf(t)$ contains no more λ, the extracted t_i no more decision-by-case constants, so the result follows. ☐

Corollary 7.8. *Let* $\mathcal{T}^\omega :=$ WE-PL$^\omega + \Gamma$, *where all additional axioms of the set* Γ *have a functional interpretation in by closed terms of* WE-PL$^\omega$ *(provably in* WE-PL$^\omega + \Gamma$*). If* $\mathcal{T}^\omega \vdash \exists x^0 A_{qf}(x)$, *then there is a collection of terms* $t_1, \ldots, t_n$ *in* $\mathcal{L}$(PL), *extractable via* FI, *s.t.* $\mathcal{T}^\omega \vdash \bigvee_{i=1}^{n} A_{qf}(t_i)$. *The terms* t_i *are built up out of the constant and function symbols of* $\mathcal{L}$(PL) *which occur (modulo the embedding of* PL *into* WE-PL$^\omega$*) in* A_{qf} *and* Γ.

Proof. It is sufficient to note that extending E-PL$^\omega$ with the axioms Γ adds no new constants to the language. The corollary then follows by the same arguments as in the proof of Theorem 7.7, except that $\bigvee_{i=1}^{n} A_{qf}(t_i)$ is no longer a tautology, but provable in $\mathcal{T}^\omega$. ☐

Example (continued). *For* $A \equiv \exists x \forall y(P(x) \vee \neg P(y))$ *the functional* Φ *realizing* x *in* f *can be defined in* E-PL$^\omega$ *as* $\Phi :\equiv \lambda f.\chi_{A_{qfs}}(c, f(c), c, f(c))$*. This new decision-by-case term is then applied to* f*, so that after normalization and unfolding of the* χ_A *the Herbrand disjunction will be:*

$$(P(c) \vee \neg P(f(c))) \vee (P(f(c)) \vee \neg P(f(f(c))))$$

In order to give an estimate on the number of extracted PL-terms, we need an estimate on the degree $dg(t)$ of the FI-extracted E-PL$^\omega$-term t.

Definition 7.9. *Let* A *be a formula, then we define the degree* $dg(A)$ *to be the* $\neg$*-depth of* A*. Let* ϕ *be a proof, then* $dg(\phi)$ *is the maximum degree of cut formulas occurring in* ϕ *and the end-formula of* ϕ*. The end-formula always is purely existential, hence* $dg(\phi) = \max\{1, dg(A_1), \ldots, dg(A_n)\}$ *for cut formulas* A_i *in* ϕ*.*

In Shoenfield's variant of FI only negation increases the type of the functional realizers. Since none of the derivation rules further increase the types, $dg(\phi)$ correctly estimates degree of the FI-extracted E-PL$^\omega$-term t. Refining a result by Schwichtenberg [109, 110], Beckmann [3] proves the following bound on normalization in the typed λ-calculus (which applies to our 'applied' λ-calculus by treating our constant symbols as free variables):

Theorem 7.10. (Beckmann,[3]) *Let* t *be a term in typed* λ*-calculus, then the length of any reduction sequence is bounded by* $2^{\|t\|}_{dg(t)}$

Corollary 7.11. *The number of terms extracted in Theorem 7.7 from a proof* ϕ *can be bounded by* $2^{3\|t\|}_{dg(\phi)+1}$*.*

Proof. To give a bound on $\#_\chi(nf(t))$ we use the following trick : from t construct a term t' by replacing every occurrence of χ by a term $((\lambda x^0.\chi)c^0)$. Then $\|t'\| \leq 3 \cdot \|t\|$ and t, t' have the same normal form. For t' consider a normalization sequence of the following kind : first perform all possible reduction steps except those on the terms substituted for the χ, then perform the reductions on the $((\lambda x^0.\chi)c^0)$ terms. The length of such a reduction sequence trivially is an upper bound on $\#_\chi(nf(t')) = \#_\chi(nf(t))$.

By Definition 7.9 and Theorem 7.10 we can bound the length of any reduction sequence of t' and hence $\#_\chi(nf(t))$ by $2^{3\cdot\|t\|}_{dg(\phi)}$. The result then follows from Theorem 7.7. ☐

Remark. *The dependence of the size of the Herbrand disjunction extracted by FI on the* $\neg$*-depth of cut formulas directly corresponds to the dependence of the complexity of cut elimination (and hence the length of Herbrand disjunctions extracted by cut elimination) on the quantifier alternations in the cut formulas.*

As mentioned above, the extraction of realizing terms generalizes to tuples, i.e. to formulas $\exists \underline{x} A_{qf}(\underline{x})$. For arbitrary prenex formulas we first construct the Herbrand normal form which then is a purely existential statement.

7.3 Discussion of bounds on Herbrand's Theorem

By an analysis of the E-PL$^\omega$ terms extracted by FI and using Beckmann's bounds on normalisation in the typed λ-calculus, we can extract bounds on the size of a Herbrand disjunction (i.e. the number of disjuncts), which match the best known bounds obtained via the cut elimination theorem [36, 37].

In [120, 121], Zhang gives a very technical proof that the hyperexponential complexity of cut elimination and the length of Herbrand disjunctions depend primarily on the quantifier alternations in the cut formulas, while quantifier blocks and propositional connectives do not contribute to the height of the tower of exponentials. These results on the length of the Herbrand disjunction follow easily from the extraction of Herbrand terms via FI, the bound on the degree of extracted terms and Beckmann's bounds on normalization.

In [115], Statman shows a hyperexponential lower bound on Herbrand's theorem, by describing formulas S_n for which there exist short proofs, but every Herbrand disjunction must have size at least 2_n. Later presentations of Statman's theorem are due to Orevkov and Pudlak [101, 102, 103]. The short proofs given by Pudlak are of size polynomial in n, yielding FI-extracted terms of size exponential in n (by [48]). The formulas occurring in the proof can be shown to have $\neg$-depth at most n, but by careful analysis of the extracted FI terms one can bound their degree by $n - 1$. Together with Corollary 7.11 this yields a match between an upper bound on the size of a Herbrand disjunction for S_n and Statman's lower bound as good as those obtained via cut-elimination.

Chapter 8

Strongly uniform bounds from semi-constructive proofs

The paper *Strongly uniform bounds from semi-constructive proofs* is to appear in the **Annals of Pure and Applied Logic**. The paper is joint work with U.Kohlenbach and has been slightly reformatted for inclusion in this PhD-thesis. The corrections for [77] that originally appear at the end of the paper have been omitted.

Strongly uniform bounds from semi-constructive proofs

Philipp Gerhardy, Ulrich Kohlenbach.

Abstract

In [77], the second author obtained metatheorems for the extraction of effective (uniform) bounds from classical, prima facie non-constructive proofs in functional analysis. These metatheorems for the first time cover general classes of structures like **arbitrary** metric, hyperbolic, CAT(0) and normed linear spaces and guarantee the independence of the bounds from parameters ranging over metrically bounded (not necessarily compact!) spaces. Recently (in [41]), the authors obtained generalizations of these metatheorems which allow one to prove similar uniformities even for unbounded spaces as long as certain local boundedness conditions are satisfied. The use of classical logic imposes some severe restrictions on the formulas and proofs for which the extraction can be carried out. In this paper we consider similar metatheorems for semi-intuitionistic proofs, i.e. proofs in an intuitionistic setting enriched with certain non-constructive principles. Contrary to the classical case, there are practically no restrictions on the logical complexity of theorems for which bounds can be extracted. Again, our metatheorems guarantee very general uniformities, even in cases where the existence of uniform bounds is not obtainable by (ineffective) straightforward functional analytic means. Already in the purely intuitionistic case, where the existence of effective bounds is implicit, the metatheorems allow one to derive uniformities that may not be obvious at all from a given constructive proofs. Finally, we illustrate our main metatheorem by an example from metric fixed point theory.

8.1 Introduction

Proof mining is the application of logical, or more precisely, proof theoretic methods to the analysis of formal systems and proofs with the aim of extracting additional information from (mathematical) proofs. E.g. one might want to extract from a proof that a certain iteration sequence converges an effective, computable modulus of convergence and to establish the uniformity of such a modulus or even to state general a-priori conditions for the independence of an extracted modulus from certain parameters.

In the classical case, i.e. formalizations of mathematics based on classical logic, the goal of proof mining is to extract realizers and bounds - we will focus on the extraction of bounds - from prima facie ineffective, non-constructive proofs. The technique used to prove the existence of effective bounds and, if needed, to carry out the extraction is based on an interpretation of classical proofs via some negative translation and (a suitable form of) Gödel's functional interpretation, further combined with majorization(see [65, 77]). Whereas previously only theorems involving constructively representable Polish spaces could be treated and uniformity in parameters was guaranteed only for the case of compact spaces ([65, 66]) results in [77] due to the second author allow one to treat classes of **arbitrary** metric, hyperbolic, CAT(0) and normed linear spaces X. Moreover, under very general conditions, uniformity in parameters ranging over metrically bounded spaces can be inferred a-priorily even in cases where this could not have been obtained by usual ineffective functional analytic methods. In [41], these results were recently generalized by the authors. Using a novel majorization technique developed by the authors one obtains similar uniformities even if the space as a whole is not metrically bounded but only local boundedness conditions are imposed. However, both the raw material, classical proofs, and the techniques employed for the interpretation impose certain restrictions: One can use at most weak extensionality in the proofs to be analyzed, as full extensionality can be shown to be too strong under functional interpretation. In the context of [77, 41] this is a severe restriction as it implies that not every object $f^{X\to X}$ of type $X \to X$ can be viewed as a function $f : X \to X$.[1] Also, as many classically true theorems cannot be given (a direct) computational meaning (this includes already Π^0_3-sentences), the extraction of realizers and bounds can be carried out at most for (classical) proofs of sentences of the form $\forall\exists A_{qf}$ where A_{qf} is quantifier-free with some further restrictions on the types of the quantified variables.

In this paper, we consider proof mining in the semi-intuitionistic case: intuitionistic analysis enriched with certain non-constructive principles. In the purely intuitionistic setting bounds and realizers are implicitly given. Nevertheless, even in the intuitionistic setting our results prove non-trivial consequences: as in the classical setting of [77, 41] we can now guarantee very strong uniformity results (independence from parameters ranging over metrically bounded spaces). Even in the presence of various highly ineffective principles (such as comprehension in all types for arbitrary negated or $\exists$-free formulas and many others), most of the restrictions needed in the fully classical case disappear in our semi-constructive setting: we can now use full extensionality and extract realizers and bounds from (semi-intuitionistic) proofs of arbitrary formulas, with comparatively modest restrictions on the types of the quantified variables.

The technique employed to establish these results for such semi-intuitionistic systems is a monotone variant of Kreisel's modified realizability interpretation, so-called monotone modified realizability. The metatheorem for the semi-intuitionistic case we present in this paper is to some extent based on results in [70], and the extensions presented here can be considered as the counterpart to the extensions of [65] presented in [77, 41] for the classical case. We will focus on developing the semi-intuitionistic versions of the results [77] in detail. The results in [41] can be transferred to the semi-intuitionistic setting in a similar but technically more complicated way.

As stated above, both in the classical and the semi-intuitionistic case the metatheorems allow one to derive new, strong uniformity results, by giving general, easy to check conditions under which an extracted bound will be guaranteed to be

[1] As a consequence of this, the applications given in [77, 41] mainly concern classes of functions, like nonexpansive functions, for which the extensionality can be deduced directly.

independent from certain parameters - all of this without actually having to carry out the extraction. For the independence of (effective) bounds from parameters ranging over compact spaces such results are well known and have been treated in [66, 70]. For non-compact bounded metric or hyperbolic spaces there are no general mathematical reasons why such uniformities should hold, and in metric fixed point theory similar (ineffective) uniformity results have hitherto only been obtained in special cases by non-trivial functional analytic techniques (see [77, 79] for discussions of these points). Already in the context of fully intuitionistic proofs one can derive new uniformities that may not be obvious from a given constructive proof or a bound implicit in the proof.

We illustrate the various aspects of the metatheorems by a very simple example from metric fixed point theory: First we state the original ineffective version of Edelstein's fixed point theorem [30]. The main part of Edelstein's fixed point theorem is of a too complicated logical form (namely Π^0_3) to directly allow the extraction via the classical metatheorems in [77, 41]. Therefore in [81] an effective uniform bound for Edelstein's fixed point theorem was extracted by splitting up Edelstein's proof into three lemmas, each simple enough to allow the extraction of an effective bound. We present a variant of Edelstein's fixed point theorem due to Rakotch [104], the proof of which is fully constructive. This permits us to extract a uniform bound as guaranteed by the semi-intuitionistic metatheorem. Finally, we compare the results with a treatment of Edelstein's fixed point theorem in the setting of Bishop-style constructive mathematics by Bridges, Julian, Richman and Mines [16]. Both the classical and the intuitionistic metatheorem a-priorily guarantee uniformities not stated in the constructive proof by Bridges et. al. The bound extracted from Rakotch's constructivized proof, while superior to the bound extracted in [81], is identical to the bound implicit in [16].

8.2 Formal systems

We now describe the classical and intuitionistic formal systems in which the extraction of bounds is carried out. For technical details see [77] and also [93].

Let $\mathcal{A}^\omega := \mathbf{WE\text{-}PA}^\omega + \mathbf{QF\text{-}AC} + \mathbf{DC}$ be the system of so-called weakly extensional classical analysis based upon a finite type extension $\mathbf{WE\text{-}PA}^\omega$ of first order Peano arithmetic **PA**, where **QF-AC** is the axiom schema of quantifier-free choice and **DC** is the axiom schema of dependent choice in all types. Let $\mathcal{A}^\omega_i$ be defined as $\mathbf{E\text{-}HA}^\omega + \mathbf{AC}$, where $\mathbf{E\text{-}HA}^\omega$ denotes the intuitionistic extensional counterpart of $\mathbf{WE\text{-}PA}^\omega$ and **AC** is the full axiom of choice (details are given below).

Definition 8.1. *The set* **T** *of all finite types is defined inductively by the clauses*

$$(i)\ 0 \in \mathbf{T},\ (ii)\ \rho, \tau \in \mathbf{T} \Rightarrow (\rho \to \tau) \in \mathbf{T}.$$

Objects of type 0 denote natural numbers, objects of type $\rho \to \tau$ are operations mapping objects of type ρ to objects of type τ. We only include equality $=_0$

between objects of type 0 as a primitive predicate. Equality between objects of higher types $s =_\rho t$ is a defined notion:[2]

$$s =_\rho t :\equiv \forall x_1^{\rho_1}, \ldots, x_k^{\rho_k} (s(x_1, \ldots, x_k) =_0 t(x_1, \ldots, x_k)),$$

where $\rho = \rho_1 \to \rho_2 \to \ldots \rho_k \to 0$, i.e. higher type equality is defined as extensional equality. An operation F of type $\rho \to \tau$ is called extensional if it respects this extensional equality:

$$\forall x^\rho, y^\rho \big(x =_\rho y \to F(x) =_\tau F(y)\big).$$

Ideally, we would like to have an axiom stating the extensionality for all functionals, but in the classical system $\mathcal{A}^\omega$ full extensionality would be too strong for the metatheorems we are aiming at and their applications in functional analysis to hold. Instead in $\mathcal{A}^\omega$ we include a weaker quantifier-free extensionality rule due to [114]:

$$\text{QF-ER}: \ \frac{A_0 \to s =_\rho t}{A_0 \to r[s] =_\tau r[t]}, \text{ where } A_0 \text{ is a quantifier-free formula.}$$

The rule QF-ER allows one to derive the equality axioms for type-0 objects

$$x =_0 y \to t[x] =_\tau t[y]$$

but not for objects x, y of higher types (see [117], [52]).

In the intuitionistic system $\mathcal{A}_i^\omega$ we include the much stronger extensionality axiom:

$$E_\rho : \forall z^\rho, x_1^{\rho_1}, y_1^{\rho_1}, \ldots, x_k^{\rho_k}, y_k^{\rho_k} (\bigwedge_{i=1}^{k} (x_i =_{\rho_i} y_i) \to z\underline{x} =_0 z\underline{y}),$$

for all types ρ.

The systems $\mathcal{A}^\omega$ and $\mathcal{A}_i^\omega$ are defined on top of many-sorted classical, resp. intuitionistic, logic with constants O^0 (zero), S^1 (successor), $\Pi_{\rho,\tau}^{\rho\to\tau\to\rho}$ (projectors), $\Sigma_{\delta,\rho,\tau}$ (combinators of type $(\delta \to \rho \to \tau) \to (\delta \to \rho) \to \delta \to \tau$) and constants $\underline{R}_{\underline{\rho}}$ for simultaneous primitive recursion in all types.[3] In addition to the defining equations for those constants, $\mathcal{A}^\omega$ and $\mathcal{A}_i^\omega$ contain as non-logical axioms:

1. Reflexivity, symmetry and transitivity axioms for $=_0$,
2. the axiom schema of complete induction:

   $$\text{IA} \ : \ A(0) \wedge \forall x^0 \big(A(x) \to A(S(x))\big) \to \forall x^0 A(x),$$

   where $A(x)$ is an arbitrary formula of our language,

[2] Here we write $s(x_1, \ldots, x_k)$ for $(\ldots(sx_1)\ldots x_k)$.

[3] It is well-known that simultaneous primitive recursion in all finite types (which defines primitive recursively finite tuples of functionals rather than a single functional only) can be reduced to ordinary primitive recursion in all finite types over $\mathcal{A}_i^\omega$ (see [117](1.6.16)). However, in the extensions $\mathcal{A}_{(i)}^\omega[X, \ldots]$ to be discussed below this seems to require the addition of certain product types so that we prefer to take simultaneous recursion as a primitive concept as in [77].

3. in $\mathcal{A}^\omega$:

   - the quantifier-free extensionality rule QF-ER
   - the quantifier-free axiom of choice schema in all types:

   $$\text{QF-AC} \;:\; \forall \underline{x} \exists \underline{y} A_0(\underline{x}, \underline{y}) \to \exists \underline{Y} \forall \underline{x} A_0(\underline{x}, \underline{Y}\, \underline{x}),$$

   where A_0 is quantifier-free and $\underline{x}, \underline{y}$ are tuples of variables of arbitrary types,

   - the axiom schema of dependent choice DC:= $\{\text{DC}^\rho : \rho \in \mathbf{T}\}$:

   $$\text{DC}^\rho \;:\; \forall x^0, y^\rho \exists z^\rho A(x, y, z) \to \exists f^{0 \to \rho} \forall x^0 A(x, f(x), f(S(x))),$$

   where A is an arbitrary formula and ρ an arbitrary type.

4. in $\mathcal{A}_i^\omega$:

   - the axiom schema of extensionality $E = \{E_\rho : \rho \in \mathbf{T}\}$ for all types ρ
   - the axiom schema of full choice AC:= $\{\text{AC}^{\rho,\tau} : \rho, \tau \in \mathbf{T}\}$:

   $$\text{AC}^{\rho,\tau} \;:\; \forall x^\rho \exists y^\tau A(x, y) \to \exists Y^{\rho \to \tau} \forall x A(x, Y x).$$

   where A is an arbitrary formula.

We next sketch extensions of $\mathcal{A}^\omega$ and $\mathcal{A}_i^\omega$ with an (non-empty) abstract metric space (X, d), resp. hyperbolic space or CAT(0) space (X, d, W), where for the somewhat involved details we refer to [77]:

The basic idea is to axiomatically add an abstract metric or hyperbolic space as a kind of 'Urelement' to the system. More formally, the theories $\mathcal{A}^\omega[X, d]$, $\mathcal{A}^\omega[X, d, W]$ and $\mathcal{A}^\omega[X, d, W, \text{CAT}(0)]$ result from extending $\mathcal{A}^\omega$ (and also IA, $\underline{\text{R}}$, QF-AC, DC, QF-ER, ...) to the set $\mathbf{T}^X$ of all finite types over the two ground types 0 and X, and by adding constants d_X and – in the case of $\mathcal{A}^\omega[X, d, W]$ and $\mathcal{A}^\omega[X, d, W, \text{CAT}(0)]$ – W_X representing d, W and suitable (purely universal) axioms to $\mathcal{A}^\omega$. Moreover, we add a constant b_X (of type 0) for an upper bound of d_X. Equality is defined extensionally over the base types 0 and X, where $x^X =_X y^X :\equiv (d_X(x, y) =_{\mathbb{R}} 0_{\mathbb{R}})$. Analogously, the theories $\mathcal{A}_i^\omega[X, d]$, $\mathcal{A}_i^\omega[X, d, W]$ and $\mathcal{A}_i^\omega[X, d, W, \text{CAT}(0)]$ result from an extension of $\mathcal{A}_i^\omega$.

Similarly, one defines the extensions $\mathcal{A}^\omega[X, \|\cdot\|, C]$ and $\mathcal{A}_i^\omega[X, \|\cdot\|, C]$ of $\mathcal{A}^\omega$ and $\mathcal{A}_i^\omega$ with an abstract (non-trivial) normed linear space $(X, \|\cdot\|)$ and a (non-empty) bounded convex subset $C \subset X$ (again we refer to [77] for details):

The theories $\mathcal{A}^\omega[X, \|\cdot\|, C]$ and $\mathcal{A}_i^\omega[X, \|\cdot\|, C]$ result from extending $\mathcal{A}^\omega$ and $\mathcal{A}_i^\omega$ to the set $\mathbf{T}^X$ of all finite types over the two ground types 0 and X, and by adding constants for the vector space operations and $\|\cdot\|$ as well as for the characteristic function of C and an upper bound b_X on the norm of the elements of C with appropriate (purely universal) axioms to $\mathcal{A}^\omega$ expressing the vector space and norm axioms as well as the boundedness and convexity of C. As before, equality is defined extensionally over the base types 0 and X.

Definition 8.2. *Between functionals x^ρ, y^ρ of type $\rho = \tau_1 \to \ldots \to \tau_k \to 0$ with $\tau_i \in \mathbf{T}^X$ we define a relation $\le_\rho$ as follows:*

$$x \le_\rho y :\equiv \forall \underline{z}^{\underline{\tau}}(x(\underline{z}) \le_0 y(\underline{z})).$$

For $\mathcal{A}^\omega_{(i)}[X, \|\cdot\|, C]$ we extend $\le_\rho$ to arbitrary types $\rho \in \mathbf{T}^X$ by defining for $\rho = \tau_1 \to \ldots \to \tau_k \to X$:

$$x \le_\rho y :\equiv \forall \underline{z}^{\underline{\tau}}(\|x(\underline{z})\|_X \le_{\mathbb{R}} \|y(\underline{z})\|_X).$$

Definition 8.3. *Let X be a non-empty set. The full set-theoretic type structure $\mathcal{S}^{\omega,X} := \langle S_\rho \rangle_{\rho \in \mathbf{T}^X}$ over $\mathbb{N}$ and X is defined by*

$$S_0 := \mathbb{N}, \quad S_X := X, \quad S_{\tau \to \rho} := S_\rho^{S_\tau}.$$

Here $S_\rho^{S_\tau}$ is the set of all set-theoretic functions $S_\tau \to S_\rho$.

We say that a sentence of $\mathcal{L}(\mathcal{A}^\omega[X,d])$, holds in a nonempty bounded metric space (X,d) if it holds in the model[4] of $\mathcal{A}^\omega[X,d]$ obtained by letting the variables range over the appropriate universes of the full set-theoretic type structure $\mathcal{S}^{\omega,X}$ with the set X as the universe for the base type X, and the constants of (X,d) interpreted by elements of the suitable universes as specified in [77].

Similarly for $\mathcal{L}(\mathcal{A}^\omega[X,d,W])$, $\mathcal{L}(\mathcal{A}^\omega[X,d,W,\mathrm{CAT}(0)])$ and $\mathcal{L}(\mathcal{A}^\omega[X,\|\cdot\|,C])$, and for the languages formed over the corresponding intuitionistic systems.

In the following (for $\rho \in \mathbf{T}^X$) '$\forall x^C\ A(x)$', '$\forall f^{\rho \to C}\ A(f)$', '$\forall f^{X \to C}\ A(f)$' and '$\forall f^{C \to C}\ A(f)$' abbreviate

$$\begin{array}{l} \forall x^X(\chi_C(x^X) =_0 0 \to A(x)), \\ \forall f^{\rho \to X}\big(\forall y^\rho(\chi_C(f(y)) =_0 0) \to A(f)\big), \\ \forall f^{X \to X}\big(\forall y^X(\chi_C(f(y)) =_0 0) \to A(f)\big) \text{ and} \\ \forall f^{X \to X}\big(\forall x^X(\chi_C(x) =_0 0 \to \chi_C(f(x)) =_0 0) \to A(\tilde{f})\big), \end{array}$$

where $\tilde{f}(x) = \begin{cases} f(x), & \text{if } \chi_C(x) =_0 0 \\ c_X, & \text{otherwise.} \end{cases}$

Analogously for the corresponding $\exists$-quantifiers with '$\wedge$' instead of '$\to$'. This extends to types of degree $(1, X, C)$ and (X, C) defined below.

Definition 8.4. *We say that a type $\rho \in \mathbf{T}^X$ has degree*

- 1 *if $\rho = 0 \to \ldots \to 0$ (including $\rho = 0$),*
- $(0, X)$ *if $\rho = 0 \to \ldots \to 0 \to X$ (including $\rho = X$),*
- $(1, X)$ *if it has the form $\tau_1 \to \ldots \to \tau_k \to X$ (including $\rho = X$), where τ_i has degree 1 or $(0, X)$,*

[4]Strictly speaking, we would have to use the plural here as the interpretation of constant b_X is not uniquely determined. For details see [77].

- $(\cdot, 0)$ *if* $\rho = \tau_1 \to \ldots \to \tau_k \to 0$ *(including* $\rho = 0$*) for arbitrary types* $\tau_i \in \mathbf{T}^X$,
- $(\cdot, X)$ *if* $\rho = \tau_1 \to \ldots \to \tau_k \to X$ *(including* $\rho = X$*) for arbitrary types* $\tau_i \in \mathbf{T}^X$.

Types involving C *do not belong to* $\mathbf{T}^X$ *but are only used in connection with the abbreviations mentioned above. We say that such a type has degree*

- $(1, X, C)$ *if it has the form* $\tau_1 \to \ldots \to \tau_k \to C$ *(including* $\rho = C$*), where* τ_i *has degree* 1 *or* $\tau_i = X$ *or* $\tau_i = C$,
- (X, C) *if* $\rho = \tau_1 \to \ldots \to \tau_k \to C$ *(including* $\rho = C$*) where* $\tau_i \in \mathbf{T}^X$ *or* $\tau_i = C$.

In [41], unbounded metric, hyperbolic and CAT(0) spaces, as well as normed linear spaces with an unbounded convex subset C are treated. The corresponding classical (and semi-intuitionistic) theories are defined as above, except that the axiom stating the boundedness of the metric space (X, d), resp. the convex subset C, is omitted. This is expressed by adding a '$-b$', i.e. by writing e.g. $\mathcal{A}^\omega[X, d]_{-b}$, $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$ and likewise for the unbounded variants of the other classical and semi-intuitionistic theories described in this section.

8.3 Extracting bounds from classical proofs

In this section we briefly restate material from [77] and [41].

Definition 8.5. *A formula* F *is called a* $\forall$*-formula (resp. an* $\exists$*-formula) if it has the form* $F \equiv \forall \underline{a}^{\underline{\sigma}} F_{qf}(\underline{a})$ *(resp.* $F \equiv \exists \underline{a}^{\underline{\sigma}} F_{qf}(\underline{a})$*) where* F_{qf} *does not contain any quantifier and the types in* $\underline{\sigma}$ *are of degree* 1 *or* $(1, X)$.

For metric, hyperbolic and CAT(0) spaces we have the following metatheorem:

Theorem 8.6 ([77]). *1. Let* σ, ρ *be types of degree* 1 *and* τ *be a type of degree* $(1, X)$. *Let* $s^{\sigma \to \rho}$ *be a closed term of* $\mathcal{A}^\omega[X, d]$ *and* $B_\forall(x^\sigma, y^\rho, z^\tau, u^0)$ *(resp.* $C_\exists(x^\sigma, y^\rho, z^\tau, v^0)$*) be a* $\forall$*-formula containing only* x, y, z, u *free (resp. a* $\exists$*-formula containing only* x, y, z, v *free).*
If

$$\forall x^\sigma \forall y \leq_\rho s(x) \forall z^\tau \left(\forall u^0 B_\forall(x, y, z, u) \to \exists v^0 C_\exists(x, y, z, v)\right)$$

is provable in $\mathcal{A}^\omega[X, d]$, *then one can extract a computable functional* $\Phi : \mathcal{S}_\sigma \times \mathbb{N} \to \mathbb{N}$ *such that for all* $x \in \mathcal{S}_\sigma$ *and all* $b \in \mathbb{N}$

$$\forall y \leq_\rho s(x) \forall z^\tau \left[\forall u \leq \Phi(x, b)\, B_\forall(x, y, z, u) \to \exists v \leq \Phi(x, b)\, C_\exists(x, y, z, v)\right]$$

holds in any (non-empty) metric space (X, d) *whose metric is bounded by* $b \in \mathbb{N}$.

2. *For bounded hyperbolic spaces (X, d, W) statement 1. holds with '$\mathcal{A}^\omega[X, d, W]$, (X, d, W)' instead of '$\mathcal{A}^\omega[X, d], (X, d)$'.*
3. *If the premise is proved in '$\mathcal{A}^\omega[X, d, W, \mathrm{CAT}(0)]$', instead of '$\mathcal{A}^\omega[X, d, W]$', then the conclusion holds in all b-bounded $\mathrm{CAT}(0)$-spaces.*

Instead of single variables x, y, z, u, v we may also have finite tuples of variables $\underline{x}, \underline{y}, \underline{z}, \underline{u}, \underline{v}$ as long as the elements of the respective tuples satisfy the same type restrictions as x, y, z, u, v. Moreover, instead of a single premise of the form '$\forall u^0 B_\forall(x, y, z, u)$' we may have a finite conjunction of such premises.

One of the main aspects of this theorem is that the bound $\Phi(x, b)$ does not depend on y or z.
The proof in [77] is based on an extension of Spector's[114] extension of Gödel's functional interpretation to classical analysis $\mathcal{A}^\omega$ by bar recursive functionals (i.e. recursion over well-founded trees) to $\mathcal{A}^\omega[X, d]$, resp. $\mathcal{A}^\omega[X, d, W]$ and $\mathcal{A}^\omega[X, d, W, \mathrm{CAT}(0)]$, and a subsequent interpretation of these functionals in an extension $\mathcal{M}^{\omega,X}$ of the Howard-Bezem[52, 12] strongly majorizable functionals $\mathcal{M}^\omega$ to $\mathbf{T}^X$.

These extensions rest on the following observations:

1. As is the case with $\mathcal{A}^\omega$, the prime formulas of $\mathcal{A}^\omega[X, d]$ are of the form $s =_0 t$ and hence decidable. Thus the soundness of negative translation and subsequent functional interpretation of the logical axioms and rules and the defining equations for combinators Σ, Π and the recursor R, the rule QF-ER and the axiom schema QF-AC extend to the new set of types $\mathbf{T}^X$ without any changes. Likewise the interpretation of the axiom schema of induction and the axiom schema of dependent choice extends to $\mathbf{T}^X$ using constants $\underline{\mathrm{R}}_{\underline{\rho}}$ for simultaneous primitive recursion and $\underline{\mathrm{B}}^{\underline{\rho},\underline{\tau}}$ for simultaneous bar recursion in all types $\underline{\rho}, \underline{\tau} \in \mathbf{T}^X$.
2. The functional interpretation of the negative translation of the new axioms of $\mathcal{A}^\omega[X, d], \mathcal{A}^\omega[X, d, W]$ and $\mathcal{A}^\omega[X, d, W, \mathrm{CAT}(0)]$ are equivalent to themselves as they are purely universal and don't contain $\vee$.
3. Bezem's[12] type structure of hereditarily strongly majorizable functionals $\mathcal{M}^\omega$ extends easily to all types of $\mathbf{T}^X$, taking $x^* \operatorname{maj}_X x$ always true. The realizer $\Psi \in \mathcal{M}^{\omega,X}$ for a bound on u^0, v^0 extracted by negative translation and functional interpretation depends on X via an interpretation of the constants of X. Using majorization we show that we can extract a bound which only depends on X via an interpretation of b_X by some integer bound b on the metric d.
4. Since for the restricted types γ of degree 1, $(0, X)$ or $(1, X)$ occurring in

   $$\forall x^\sigma \forall y \leq_\rho s(x) \forall z^\tau \left(\forall u^0 B_\forall(x, y, z, u) \to \exists v^0 C_\exists(x, y, z, v)\right)$$

   $M_\gamma = S_\gamma$, this bound holds in any nonempty b-bounded space (X, d), resp. (X, d, W) and $(X, d, W, \mathrm{CAT}(0))$.

For a detailed proof, see [77].

Definition 8.7. *1. Let (X, d) be a metric space. A function $f : X \to X$ is called nonexpansive (short: 'f n.e.') if*

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

2. ([58]) Let (X, d, W) be a hyperbolic space. A function $f : X \to X$ is called directionally nonexpansive (short: 'f d.n.e.') if

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

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

Definition 8.8. *Let $f : X \to X$, then $Fix(f) := \{x \in X \mid x = f(x)\}$.*

In [77], the following corollary of theorem 8.6 is derived, which is specially tailored towards applications to metric fixed point theory:

Corollary 8.9 ([77])**.** *1. Let P (resp. K) be a $\mathcal{A}^\omega$-definable Polish space (resp. compact Polish space), given in so-called standard representation, and $B_\forall(x^1, y^1, z, f, u), C_\exists(x^1, y^1, z, f, v)$ be as in the previous theorem. If $\mathcal{A}^\omega[X, d, W]$ proves that*

$$\forall x \in P \forall y \in K \forall z^X, f^{X\to X}\big(\ f \text{ n.e.} \wedge Fix(f) \neq \emptyset \wedge \forall u^0 B_\forall \to \exists v^0 C_\exists\big),$$

then there exists a computable functional $\Phi^{1\to 0\to 0}$ (on representatives $x : \mathbb{N} \to \mathbb{N}$ of elements of P) such that for all $x \in \mathbb{N}^{\mathbb{N}}, b \in \mathbb{N}$

$$\forall y \in K \forall z^X \forall f^{X\to X}\big(\ f \text{ n.e.} \wedge \forall u \le \Phi(x, b)\, B_\forall \to \exists v \le \Phi(x, b)\, C_\exists\big)$$

holds in any (non-empty) hyperbolic space (X, d, W) whose metric is bounded by b.

2. An analogous result holds if 'f n.e.' is replaced by 'f d.n.e'.

Note that in the corollary, the assumption $Fix(f) \neq \emptyset$ has disappeared in the conclusion! For a discussion of this remarkable point see [77].

For normed linear spaces, the following metatheorem is proved in [77]:

Theorem 8.10 ([77])**.** *Let σ be a type of degree 1, ρ of degree 1 or $(1, X)$ and τ of degree $(1, X, C)$. Let $s^{\sigma\to\rho}$ be a closed term of $\mathcal{A}^\omega[X, \|\cdot\|, C]$ and $B_\forall(x^\sigma, y^\rho, z^\tau, u^0)$ (resp. $C_\exists(x^\sigma, y^\rho, z^\tau, v^0)$) be a $\forall$-formula containing only x, y, z, u free (resp. an $\exists$-formula containing only x, y, z, v free).*
If

$$\forall x^\sigma \forall y \le_\rho s(x) \forall z^\tau \big(\forall u^0 B_\forall(x, y, z, u) \to \exists v^0 C_\exists(x, y, z, v)\big)$$

is provable in $\mathcal{A}^\omega[X, \|\cdot\|, C]$, then one can extract a computable functional $\Phi : \mathcal{S}_\sigma \times \mathbb{N} \to \mathbb{N}$ such that for all $x \in \mathcal{S}_\sigma$ and all $b \in \mathbb{N}$

$$\forall y \le_\rho s(x) \forall z^\tau \big[\forall u \le \Phi(x, b)\, B_\forall(x, y, z, u) \to \exists v \le \Phi(x, b)\, C_\exists(x, y, z, v)\big]$$

holds in any non-trivial normed linear space $(X, \|\cdot\|)$ *and any non-empty b-bounded convex subset* C.

Instead of single variables and a single premise we may have tuples of variables and a finite conjunction of such premises.

Remark. *In [77], there are also corresponding theorems proved for uniformly convex normed spaces* $(X, \|\cdot\|, \eta)$ *with convexity modulus* η *(then the bound* $\Phi(x, b, \eta)$ *will additionally depend on the modulus* η*) and for inner product spaces.*

The proof in [77] is based on the same fundamental ideas as the proof of Theorem 8.6, the main difference being that the majorization relation on objects of type X can no longer be treated as trivial as in the case of a bounded metric space. Instead one defines the majorization relation s-maj for elements of type X to be

$$x^* \text{ s-maj}_X \ x :\equiv \|x^*\|_X \geq_{\mathbb{R}} \|x\|_X.$$

Then one can prove, as before, the extractability of effective bounds, where the main difficulty is to define suitable majorants for the constants and constructions of $\mathcal{A}^\omega[X, \|\cdot\|, C]$.

As shown in [41], using a novel majorization technique these metatheorems can be generalized to unbounded metric spaces and normed linear spaces with unbounded convex subset C. The new majorization relation developed by the authors is technically more complicated but allows one to derive similar uniformities from far more general conditions than the boundedness of the entire metric space, resp. the convex subset C.

Discussion on extensionality: As mentioned above, one can only allow the weak extensionality rule instead of the full axiom of extensionality in the formal systems based on classical logic. In order to reverse the double negations introduced by the negative translation, it is strictly necessary that the interpretation we choose to interpret classical logic in particular interprets the Markov principle. However, together with the Markov Principle full extensionality would cause severe problems, as it allows us, when combined with functional interpretation, to obtain witnesses for potential universal quantifiers hidden in the extensionally defined equalities in the premise of implications, e.g. in the extensionality axiom itself.

The extraction of witnesses, combined with majorization, would thus transform an instance of the extensionality axiom into a statement about uniform continuity. An axiom stating the extensionality of a single function constant would allow us to prove its uniform continuity. E.g. the full extensionality axiom for type-X equality would even allow us to prove (in the context of $\mathcal{A}^\omega[X, d]$) the equicontinuity of all functions $f^{X \to X}$ which – of course – is not true in general (but does hold for the class of nonexpansive mappings $f : X \to X$, whose full extensionality follows in $\mathcal{A}^\omega[X, d]$).

A similar problem with extensionality arises from the representation of a convex subset C of a normed linear space via its characteristic function χ_C. Here we would like the characteristic function to respect the extensional equality, i.e.

$$x =_X y \to \chi_C(x) =_0 \chi_C(y).$$

In the presence of functional interpretation and majorization, this would not only yield that points $x \in X$ close to C behave similar to points in C, it would also describe a modulus for how close to C you have to be to behave 'sufficiently similar'. Unless the subset C is topologically very simple (e.g. a closed bounded ball), such statements will in general not be correct.

Therefore, we must restrict the formal system to make unwanted or simply false conclusions, drawn from extensionality statements, impossible. In turn, when it is necessary to employ an extensional equality in a proof, we cannot simply assume extensionality: every statement of extensionality that is used in a proof must itself be explicitly proved with the use of QF-ER or follow from uniform continuity. For more details, see the discussion of extensionality in section 3 of [77].

8.4 Extracting bounds from semi-constructive proofs

The metatheorems from [77] which we briefly discussed in the previous section allow one to extract bounds from proofs in fairly strong systems, namely extensions of classical analysis with an abstract metric, hyperbolic, CAT(0) resp. normed linear space. However, the fact that the formal systems were based on classical logic imposes severe restrictions on the class of formulas for which extraction of bounds is possible.

The first step in the extraction algorithm is to apply negative translation to the classical proof (of some formula F), i.e. to translate it into an essentially intutionistic proof of the negative translation F^N of F (which may, however, use the Markov principle to be discussed below). This restricts the extraction of bounds to $\forall\exists A$-formulas for which the equivalence between the formula and its negative translation can be shown to hold under the Markov Principle, namely the class of formulas $\forall\exists A_{qf}$, where A_{qf} is quantifier-free (or purely existential). In consequence, the interpretation must interpret the Markov Principle, as functional interpretation indeed does. In general, such an equivalence can be validated at most for $\forall\exists A_{qf}$-formulas, as already the formula class Π^0_3 yields counterexamples to the existence of effective bounds in the form of e.g. the halting problem.

Secondly, the interpretation of the negative translation of the axiom of dependent choice by bar recursive functionals requires arguments which hold only in the model of hereditarily strongly majorizable functionals $\mathcal{M}^{\omega,X}$ over the types $\mathbb{N}$ and X but not in the full set-theoretic model $\mathcal{S}^{\omega,X}$. In consequence, for the extracted bounds to hold in $\mathcal{S}^{\omega,X}$, we must restrict the types of the quantified variables in the theorem to be proved to types of degree 1 or $(1, X)$, as for those low types the proper inclusions between these two models hold.

We will see now that the intuitionistic counterpart of $\mathcal{A}^\omega$ and its extensions to metric, hyperbolic, CAT(0) and normed linear spaces do not suffer from such restrictions (even when strong ineffective principles are added).

In the classical case, an extension of Gödel's Dialectica interpretation combined with negative translation and majorization (monotone functional interpretation) was used to obtain the results. In the intuitionistic setting we derive these results from a monotone variant of Kreisel's modified realizability interpretation (in short: mr-interpretation), the so-called monotone modified realizability interpretation. Kreisel's mr-interpretation was introduced in [88, 89] and studied in great detail in [117, 118]. The monotone mr-interpretation was introduced in [70] and is studied in detail in [63].

This interpretation has the following nice properties:

1. As in the classical case, we can use the general metatheorem as a black box to prove (even qualitatively new) uniformity results without actually having to carry out the extraction.

2. Contrary to classical systems, we are no longer restricted to proofs of $\forall\exists A_{qf}$-statements, but can allow $\forall\exists A$-statements for arbitrary A. Furthermore, the additional restrictions on the quantifiers stated in Theorem 8.6 and Theorem 8.10 can be significantly relaxed.

3. We may add large classes of additional axioms $\Gamma_\neg$ which include highly ineffective principles such as full comprehension for arbitrary negated formulas (which is not even allowed in the classical context, where it would give full comprehension for all formulas).

The Markov Principle in all finite types is the principle

$$M^\omega : \neg\neg\exists \underline{x} A_{qf}(\underline{x}) \to \exists \underline{x} A_{qf}(\underline{x}),$$

where A_{qf} is an arbitrary quantifier-free formula and $\underline{x}$ is a tuple of variables of arbitrary types (A_{qf} may contain further free variables).

As discussed above, in the classical case it is strictly necessary that the interpretation we choose interprets the Markov principle, and this imposes certain restrictions on the formal system. In the intuitionistic setting we can choose not to include the Markov Principle. As a consequence, when extending intuitionistic analysis with non-constructive principles we have an actual choice between two main directions in which to extend the formal system: *with* or *without* the Markov Principle M^ω:

Extending the system *with* the Markov Principle would force us both to restrict extensionality to weak extensionality and to allow at most the independence of premise scheme for purely universal formulas. However, we could still – replacing the use of negative translation in the proofs of the main results in [77] by the reasoning used to prove theorem 3.18 in [70] (based on monotone functional interpretation) – extract bounds for arbitrary formulas $\forall\exists A$, instead of the restricted formula class $\forall\exists A_{qf}$.

We choose instead to extend our formal system in the direction *without* M^ω. Abandoning the Markov Principle allows us to add full extensionality and comprehension and independence of premise schemes for arbitrary negated formulas, as well as many other essentially non-constructive analytic or logical principles (see also [70]).

Let comprehension for negated formulas be the principle (also for tuples of variables $\underline{y}$):

$$CA_\neg^{\underline{\rho}} : \exists \Phi \leq_{\underline{\rho}\to 0} \lambda \underline{x}^{\underline{\rho}}.1^0 \forall \underline{y}^{\underline{\rho}}(\Phi(\underline{y}) =_0 0 \leftrightarrow \neg A(\underline{y})),$$

where $\underline{y} = y_1^{\rho_1}, \ldots, y_k^{\rho_k}$ is an arbitrary tuple of variables of arbitrary types, and let the independence-of-premise principle for negated formulas be:

$$IP_\neg^{\underline{\rho}} : (\neg A \to \exists y^\rho B(y)) \to \exists y^\rho(\neg A \to B(y)) \quad (y \notin \text{ FV}(A)),$$

where in both cases A, B are arbitrary formulas. The union of these principles over all types ρ of the underlying language are denoted by $CA_\neg$ and $IP_\neg$ where – when working over the systems $\mathcal{A}_i^\omega[X, \ldots]$ – we allow arbitrary types $\rho \in \mathbf{T}^X$.

Definition 8.11. *A formula $A \in \mathcal{A}_i^\omega$, resp. $A \in \mathcal{A}_i^\omega[\ldots]$, is called $\exists$-free (or 'negative'), if A is built up from prime formulas by means of $\wedge, \to, \neg$ and $\forall$ only, i.e. A contains neither $\exists$ nor $\vee$. We denote $\exists$-free formulas A by A_{ef}.*

The principles CA_{ef} and IP_{ef} are the principles corresponding to $CA_\neg$ and $IP_\neg$, where instead of $\neg A$ we have an $\exists$-free formula A_{ef}.

We next recall Kreisel's mr-interpretation and Bezem's[12] notion of strong majorizability, which is an extension of Howard's [52] notion of majorizability, for all types $\mathbf{T}^X$. Combining these allows us to define the monotone mr-interpretation.

For each formula $A(\underline{a})$, where $\underline{a}$ are the free variables of A, Kreisel's mr-interpretation defines, by induction on the logical structure of A, a corresponding formula '$\underline{x}$ *mr* A' (in words: $\underline{x}$ modified realizes A), where $\underline{x}$ is a (possibly empty) tuple of variables, which do not occur free in A. From a proof of A Kreisel's mr-interpretation extracts a tuple of closed terms $\underline{t}$ s.t. $\forall \underline{a}(\underline{t}\underline{a}$ *mr* $A(\underline{a}))$. For details see e.g. [117, 118].

Remark. *1. For every $\exists$-free formula A we have $(\underline{x}$ mr $A) \equiv A$ with $\underline{x}$ the empty tuple.*

2. $(\underline{x}$ mr $A)$ is always an $\exists$-free formula.

Definition 8.12 ([77], extending [52, 12])**.** *The strong majorizability relation* s-maj *is defined as follows:*

- x^* s-maj$_0$ $x :\equiv x^* \geq x$

- x^* s-maj$_X$ $x :\equiv (0 =_0 0)$ *in* $\mathcal{A}^\omega_{(i)}[X, d, \ldots]$,
- x^* s-maj$_X$ $x :\equiv \|x^*\|_X \geq_{\mathbb{R}} \|x\|_X$ *in* $\mathcal{A}^\omega_{(i)}[X, \|\cdot\|, \ldots]$,
- x^* s-maj$_{\rho\to\tau}$ $x :\equiv \forall y^*, y(y^*$ s-maj$_\rho$ $y \to x^*y^*$ s-maj$_\tau$ $x^*y, xy)$

Definition 8.13 ([70])**.** *A tuple of closed terms $\underline{t}^*$ satisfies the monotone mr-interpretation of $A(\underline{a})$ if*

$$\exists \underline{z}(\underline{t}^* \text{ s-maj } \underline{z} \wedge \forall \underline{a}(\underline{z}\underline{a} \ mr \ A(\underline{a}))$$

We briefly recall some properties of the mr-interpretation. As we have the full axiom of choice AC in $\mathcal{A}^\omega_i$, resp. $\mathcal{A}^\omega_i[\ldots]$, one shows:

Proposition 8.14 (**Troelstra[117]**)**.**

$$\mathcal{A}^\omega_i + IP_{ef} \vdash A \leftrightarrow \exists \underline{x}(\underline{x} \ mr \ A)$$

Similarly for $\mathcal{A}^\omega_i[\ldots] + IP_{ef}$.

Proof. By induction on the logical structure of A. □

Corollary 8.15. *1. For every formula $A \in \mathcal{A}^\omega_i$ we can construct an $\exists$-free formula B_{ef} s.t.*

$$\mathcal{A}^\omega_i + IP_{ef} \vdash \neg A \leftrightarrow B_{ef}.$$

Similarly for $\mathcal{A}^\omega_i[\ldots]$.

2. For every $\exists$-free formula $A_{ef} \in \mathcal{A}^\omega_i$ we have that $\mathcal{A}^\omega_i \vdash A_{ef} \leftrightarrow \neg\neg A_{ef}$. Similarly for $\mathcal{A}^\omega_i[\ldots]$.

3. Over $\mathcal{A}^\omega_i$ we have $IP_{ef} \leftrightarrow IP_\neg$ and $CA_{ef} \leftrightarrow CA_\neg$. Similarly for $\mathcal{A}^\omega_i[\ldots]$.

Proof. 1. By Proposition 8.14 we have

$$\mathcal{A}^\omega_i + IP_{ef} \vdash \neg A \leftrightarrow \forall \underline{y}((\underline{y} \ mr \ A) \to \bot),$$

where $\forall \underline{y}((\underline{y} \ mr \ A) \to \bot)$ is $\exists$-free, as $(\underline{y} \ mr \ A)$ is $\exists$-free.

2. This equivalence is provable intuitionistically in the context of decidable prime formulas.

3. $\mathcal{A}^\omega_i + IP_{ef} \vdash IP_\neg$ follows from '1.', and $\mathcal{A}^\omega_i + CA_{ef} \vdash CA_\neg$ follows from the fact that $\mathcal{A}^\omega_i + CA_{ef} \vdash IP_{ef}$ and '1.'. The converse implications follow from '2.'. □

In the following, we will omit mentioning $IP_\neg$ and IP_{ef}, as they follow from the corresponding comprehension schemes $CA_\neg$ and CA_{ef} (and the decidability of $=_0$).

Discussion of extensionality, continued: As mentioned above, in the context of functional interpretation full extensionality is much too strong, as it would allow us to derive (when combined with the generalized majorizability from [77]) statements e.g. about uniform continuity which are not true in general. In the context of (monotone) modified realizability full extensionality is harmless. Extensionally defined equalities in the premise of implications, e.g. in instances of the extensionality axiom, as indeed instances of the extensionality axiom as a whole, are ∃-free and thus realized by the empty tuple.

Informally speaking, functional interpretation is 'too eager', seeking to extract every possible and hence some unwanted bounds. In contrast, modified realizability is 'lazy enough' to only extract bounds where this is explicitly asked for, namely from positive existential statements. Where functional interpretation extracts bounds on universal premises in an implication, modified realizability leaves them alone. In practice, this allows us to remove the requirement to explicitly prove every extensional equality used in the proof and instead to simply assume it as a premise, leading to a more natural, intuitive treatment of extensionality.

We can prove the following theorem, corresponding to Theorem 8.6 in the classical setting:

Theorem 8.16. *1. Let σ be a type of degree 1, let ρ be a type of degree $(\cdot,0)$ and let τ be a type of degree $(\cdot,X)$. Let $s^{\sigma\to\rho}$ be a closed term of $\mathcal{A}^\omega_i[X,d]$ and let A (resp. B) be an arbitrary formula with only x,y,z,n (resp. x,y,z) free. Let $\Gamma_\neg$ be a set of sentences of the form $\forall u^\alpha(C \to \exists v \leq_\beta tu \exists w^\gamma \neg D)$ with $t^{\alpha\to\beta}$ be a closed term of $\mathcal{A}^\omega_i[X,d]$, the type $\alpha \in \mathbf{T}^X$ arbitrary, the type β of degree $(\cdot,0)$ and γ of degree $(\cdot,X)$. If*

$$\mathcal{A}^\omega_i[X,d] + CA_\neg + \Gamma_\neg \vdash \forall x^\sigma \forall y \leq_\rho s(x) \forall z^\tau(\neg B \to \exists n^0 A),$$

then one can extract a primitive recursive (in the sense of Gödel) functional $\Phi : \mathcal{S}_\sigma \times \mathbb{N} \to \mathbb{N}$ such that for all $b \in \mathbb{N}$

$$\forall x^\sigma \forall y \leq_\rho s(x) \forall z^\tau \exists n \leq \Phi(x,b)(\neg B \to A)$$

holds in any (non-empty) metric space (X,d) whose metric is bounded by $b \in \mathbb{N}$ and which satisfies $\Gamma_\neg$.[5]

2. *For bounded hyperbolic spaces (X,d,W), '1.' holds with $\mathcal{A}_i[X,d,W],(X,d,W)$ instead of $\mathcal{A}^\omega_i[X,d],(X,d)$.*
3. *If the premise is proved in $\mathcal{A}^\omega_i[X,d,W,\mathrm{CAT}(0)]$ instead of $\mathcal{A}^\omega_i[X,d,W]$ then the conclusion holds in all nonempty b-bounded $\mathrm{CAT}(0)$ spaces satisfying $\Gamma_\neg$.*

As in the classical case, instead of single variables and single premises we may also have tuples of variables and a finite conjunction of premises.

Proof. Since prime formulas in $\mathcal{A}^\omega_i[X,d]+CA_\neg+\Gamma_\neg$ are decidable, it follows from Corollary 8.15 that this theory is equivalent to the theory $\mathcal{A}^\omega_i[X,d]+CA_{ef}+\Gamma'_{ef}$,

[5] Here b_X is understood to be interpreted by b.

where Γ'_{ef} is the set of sentences which results from $\Gamma_\neg$ by replacing in each $S \in \Gamma_\neg$ the negated formula $\neg D$ by the $\exists$-free formula D_{ef} from Corollary 8.15 which is equivalent to $\neg D$. For the subsystem of $\mathcal{A}_i^\omega[X,d] + CA_{ef} + \Gamma'_{ef}$ not involving (X,d), i.e. restricted to the types $\mathbf{T}$, the theorem is proved in [70] by establishing that this theory has a monotone mr-interpretation in its classical counterpart (for a somewhat more restricted set Γ'_{ef} even in itself) by terms in Gödel's T ((although we use mr rather than mr-with-truth we do not have to restrict the formulas A, C to Γ_1 as in [70](thm.3.10) since in the presence of AC (and hence in $\mathcal{S}^\omega$) we can use proposition 8.14 to infer these formulas back from their mr-interpretations).

To extend the proof to the full theory $\mathcal{A}_i^\omega[X,d] + CA_{ef} + \Gamma'_{ef}$, i.e. now involving the full range of types $\mathbf{T}^X$, we observe the following:

1. By arguments similar to those used in the classical case (see [77]) the soundness of the monotone mr-interpretation of the logical axioms and rules, the defining equations for combinators Σ, Π and the recursors $\underline{R}$, axiom schemes E, AC and the axiom schema of induction extends to the types $\mathbf{T}^X$ without any changes.
2. The additional axioms of $\mathcal{A}_i^\omega[X,d]$ are purely universal and do not contain $\vee$, and hence have a trivial monotone mr-interpretation by the empty tuple.
3. The additional $\exists$-quantifiers ranging over variables of type degree $(\cdot, X)$, both in the conclusion and in sentences of the set Γ'_{ef}, can easily be majorized using appropriate constant 0_X functionals as shown in [77].
4. The monotone mr-interpretation extracts a realizer $\psi \in S^{\omega,X}$ depending only on a suitable interpretation of the constants of $\mathcal{A}_i^\omega[X,d]$: The majorization relation extends to $\mathbf{T}^X$ as defined above and given a closed term ψ of $\mathcal{A}_i^\omega[X,d]$ we can construct as in [77] a majorant ψ^*, by induction on the term structure of ψ such that
$$\mathcal{S}^{\omega,X} \models \psi^* \text{ s-maj } \psi.$$
ψ^* does not involve d_X and which depends on (X,d) only via the interpretation of the constant b_X by a bound $b \in \mathbb{N}$ on the metric d and on the interpretation of 0_X by some arbitrary element of X. Using the same techniques as in the classical case ([77]) one can eliminate the latter dependency and construct from ψ^* a functional $\Phi \in S_{0\to(\sigma\to 0)}$ which is given by a closed term of $\mathcal{A}_i^\omega$ (i.e. a primitive recursive functional in the sense of Gödel) s.t.
$$\mathcal{S}^{\omega,X} \models \forall x^\sigma \forall y \leq_\rho s(x) \forall z^\tau \exists n \leq \Phi(x,b)(\neg B \to A(x,y,z,n)).$$

Since, again by corollary 8.15, $\neg B$ is equivalent to an existential free formula it is does not in any way contribute to the extracted term. For $\mathcal{A}_i^\omega[X,d,W]$ and $\mathcal{A}_i^\omega[X,d,W,\mathrm{CAT}(0)]$ the arguments are similar. In all three cases the final extracted functional Φ is primitive recursive in the sense of Gödel, i.e. Φ is given by a closed term in Gödel's T. □

In a similar way, one can prove semi-intuitionistic counterparts to the generalized metatheorems presented in [41].

We first show the following corollary, corresponding to Corollary 8.9 in the classical case:

Corollary 8.17. *1. Let P (resp. K) be a $\mathcal{A}_i^\omega$-definable Polish space (resp. compact Polish space) and let A, B and $\Gamma_\neg$ be as in the previous theorem. If $\mathcal{A}_i^\omega[X,d,W] + CA_\neg + \Gamma_\neg$ proves that*

$$\forall x \in P \forall y \in K \forall z^X, f^{X\to X}(\neg B \to \exists n^0 A)$$

then there exists a primitive recursive functional $\Phi^{1\to 0\to 0}$ (on representatives $x : \mathbb{N} \to \mathbb{N}$ of elements of P) such that for all $x \in \mathbb{N}^{\mathbb{N}}, b \in \mathbb{N}$

$$\forall y \in K \forall z^X, f^{X\to X} \exists n \leq \Phi(x,b)(\neg B \to A)$$

holds in any (non-empty) hyperbolic space (X,d,W) whose metric is bounded by b and which satisfies $\Gamma_\neg$.

2. The result also holds for $\mathcal{A}_i^\omega[X,d], (X,d)$.

Proof. The details of the proof are similar to the classical case, i.e. by Theorem 8.16 we can extract a primitive recursive bound $\Phi(x,b)$ on n which holds in all spaces (X,d,W), resp. (X,d), whose metric is bounded by b. □

In [41] a refined version of corollary 8.9 is established which states that if the assumption is proved in $\mathcal{A}^\omega[X,d,W]_{-b}$ (i.e. without the use of the axiom stating the boundedness of d) that then the conclusion holds in arbitrary (not necessary bounded) hyperbolic spaces as long as $b \geq d(x, f(x))$. This also holds (though with '$Fix(f) \neq \emptyset$' dropped) for functions which are not nonexpansive but only have a bounding function $\Omega : \mathbb{N} \to \mathbb{N}$ such that

$$\forall k^0, \tilde{z}^X(d(z,\tilde{z}) \leq k \to d(z, f(\tilde{z})) \leq \Omega(k))$$

for some z^X, where then the bound depends on Ω. This corollary has a semi-intuitionistic counterpart analogous to the previous results:

Corollary 8.18. *1. Let P (resp. K) be a $\mathcal{A}_i^\omega$-definable Polish space (resp. compact Polish space) and let A and B be as before but not containing the constant 0_X. If $\mathcal{A}_i^\omega[X,d,W]_{-b} + CA_\neg$ proves that*

$$\forall x \in P \forall y \in K \forall z^X, f^{X\to X}, \Omega^1$$
$$(\forall k^0, \tilde{z}^X(d_X(z,\tilde{z}) \leq_{\mathbb{R}} (k)_{\mathbb{R}} \to d_X(z, f(\tilde{z})) \leq_{\mathbb{R}} (\Omega(k))_{\mathbb{R}}) \wedge \neg B \to \exists n^0 A)$$

then there exists a primitive recursive functional $\Phi^{1\to 1\to 0}$ (on representatives $x : \mathbb{N} \to \mathbb{N}$ of elements of P) such that for all $x, \Omega \in \mathbb{N}^{\mathbb{N}}$

$$\forall y \in K \forall z^X, f^{X\to X}, \Omega^1 \exists n \leq \Phi(x,\Omega)$$
$$(\forall k^0, \tilde{z}^X(d_X(z,\tilde{z}) \leq_{\mathbb{R}} (k)_{\mathbb{R}} \to d_X(z, f(\tilde{z})) \leq_{\mathbb{R}} (\Omega(k))_{\mathbb{R}}) \wedge \neg B \to A)$$

holds in any (non-empty) hyperbolic space (X,d,W).

2. The result also holds for $\mathcal{A}_i^\omega[X,d]_{-b},(X,d)$.

Even if 'z' does not occur in B, A we need the assumption on f, Ω to hold for **some** *z in X.*

Note, that the boundedness of (X, d) and the bound b as a parameter have been replaced by a far more general condition on f and the parameter Ω in the unbounded case. Still, the extracted bound Φ may display similar uniformities, i.e. independence of z, f and the underlying space (X, d). As an example, for nonexpansive functions f and the additional premise $d(z, f(z)) \leq b$ we obtain $\Omega(n) := n + b$. This yields an effective bound Φ depending only on x and b, where b is not a bound on the whole space, but *only* on $d(z, f(z))$.

Remark. *As in the classical case, we can add in corollary 8.17 additional assumptions about the function f, if of suitable logical form, to the premise. In the classical case we added the assumption 'f n.e.' and '$Fix(f) \neq \emptyset$' to the premise of the implication. Both assumptions can also be added in the semi-intuitionistic case. The condition 'f n.e.' is purely universal and hence is equivalent to its double negation. The statement '$Fix(f) \neq \emptyset$' can be written as $\exists u^X C_\forall$, where $C_\forall$ is purely universal and so again equivalent to its double negation. Thus, first pulling out the existential quantifier from the premise $\exists u^X C_\forall$ as a universal quantifier just as $\forall z^X$, we can extract a bound Φ that does not depend on u and does not depend on any of the negated premises nor $C_\forall$. Shifting the quantifier $\exists u$ back in we get the result.*

In the classical case the premise 'f n.e.' ensures that a given f indeed behaves like a function, i.e. is needed to prove the extensionality of f, as the weak extensionality rule QF-ER is not strong enough to ensure this. The weaker assumption 'f d.n.e' does not imply extensionality. This is the reason why in application 3.16 of [77] one carefully had to observe that QF-ER was in fact sufficient to formalize the proof in question. Likewise the Ω-condition in Corollary 8.18 does not imply extensionality. In the semi-intuitionistic case, where we have full extensionality included as an axiom this does not cause any difficulties.

The benefit of adding '$Fix(f) \neq \emptyset$' was that FI would weaken that assumption to 'f has approximate fixed points', which for nonexpansive and even directionally nonexpansive selfmappings of a bounded hyperbolic space is always true (see [45] and [79]) whereas, in general, '$Fix(f) \neq \emptyset$' is not. In the semi-intuitionistic case '$Fix(f) \neq \emptyset$' will not disappear from the premise, as monotone modified realizability does not weaken universal premises such as $d_X(x, f(x)) =_{\mathbb{R}} 0_{\mathbb{R}}$.

For normed linear spaces we prove the following semi-intuitionistic counterpart to Theorem 8.10:

Theorem 8.19. *1. Let σ be a type of degree 1, ρ be an arbitrary type in $\mathbf{T}^X$ and let τ be a type of degree (X, C). Let $s^{\sigma \to \rho}$ be a closed term of $\mathcal{A}_i[X, \|\cdot\|, C]$ and let A (resp. B) be an arbitrary formula with only x, y, z, n (resp. x, y, z) free. Let $\Gamma_\neg$ be a set of sentences of the form*

$\forall u^\alpha(C \to \exists v \leq_\beta tu \exists w^\gamma \neg D)$ where $t^{\alpha\to\beta}$ is a closed term of $\mathcal{A}_i^\omega[X, \|\cdot\|, C]$, the types $\alpha, \beta \in \mathbf{T}^X$ are arbitrary and γ is of degree (X, C). If

$$\mathcal{A}_i^\omega[X, \|\cdot\|, C] + CA_\neg + \Gamma_\neg \vdash \forall x^\sigma \forall y \leq_\rho s(x) \forall z^\tau(\neg B \to \exists n^0 A),$$

then one can extract a primitive recursive (in the sense of Gödel) functional $\Phi : \mathcal{S}_\sigma \times \mathbb{N} \to \mathbb{N}$ such that for all $b \in \mathbb{N}$

$$\forall x^\sigma \forall y \leq_\rho s(x) \forall z^\tau \exists n \leq \Phi(x, b)(\neg B \to A)$$

holds in any nontrivial normed linear space $(X, \|\cdot\|)$ and any b-bounded convex subset C which satisfy $\Gamma_\neg$.

Instead of single variables and single premises we may also have tuples of variables and a finite conjunction of premises.

The proof is based on arguments similar to the proof of Theorem 8.10, resp. the variations due to the change of setting from classical to semi-intuitionistic discussed in the proof of Theorem 8.16. The variables of degree (X, C) in the sentences $A \in \Gamma_\neg$ can again easily be majorized by a suitable interpretation of the constant b_X by a bound b on the norm of the elements of the convex subset C. As before, the generalized metatheorems for normed linear spaces in [41] can be transferred to the semi-intuitionistic setting in a similar way, yielding similar uniform bounds. However, for (unbounded) convex subsets C we need the additional premise $\|c_X\|, \|x\| \leq b$ and the Ω-condition is written as

$$\forall x^C(\|x\|_X \leq_\mathbb{R} (n)_\mathbb{R} \to \|f(x)\|_X \leq_\mathbb{R} (\Omega(n))_\mathbb{R}).$$

Remark. *In the classical case the construction of majorants d_X^* resp. $\|\cdot\|_X^*$ depends on the interpretation of d_X resp. $\|\cdot\|_X$ in the model $\mathcal{S}^{X,\omega}$ via an ineffective operator $()_\circ$, which from a (representative of a) real number selects a canonical representative of that real number. As an operator of type $1 \to 1$, $()_\circ$ is primitive recursive in*

$$E^2(f^1) :=_0 \begin{cases} 0, & \text{if } \forall x^0(f(x) =_0 0) \\ 1, & \text{if } \neg\forall x^0(f(x) =_0 0). \end{cases}$$

Since the functional interpretation of the defining axioms of (E^2) would require non-majorizable functionals (although E^2 itselfs is trivially majorizable) one must not include the operator $()_\circ$ to $\mathcal{A}^\omega[X, \ldots]$. This causes no problems as $()_o$ only is involved in the interpretation of the theory in the model $\mathcal{S}^{\omega,X}$. Subsequently the ineffective $()_o$ operator can be majorized effectively!

In the semi-constructive case we could actually add the $()_\circ$ operator via E^2 to the theory, as monotone modified realizability leaves the defining axioms of the E^2 untouched, and carry out part of the argument regarding the $()_\circ$ operator in the theory itself rather than in the model. The existence of E^2 actually follows from CA_{ef} and hence from $CA_\neg$.

8.5 Application to Metric Fixed Point Theory

To illustrate the various aspects of Theorem 8.16 we consider three different proofs of (variants of) Edelstein's Fixed Point Theorem: first a refinement of the original proof by Edelstein[30] developed in [81], next an alternative, constructive proof by Rakotch[104] and finally a more recent proof carried out in the framework of Bishop-style constructive mathematics by Bridges, Julian, Richman and Mines[16]. Though completely elementary, if not trivial, from a functional analytic point of view, this example serves well to demonstrate the various logical aspects of proof mining using the metatheorems presented in the previous sections. For recent *non-trivial* applications of proof mining see [39, 75, 78, 79].

In [106], Rhoades presents a survey and comparison of a large number of different notions of contractivity, compiled from the literature on metric fixed point theory, for which fixed points theorems have been proven. Many of these notions of contractivity and the accompanying proofs of fixed point theorems are far more technical than the example presented in this section. Further surveys on notions of contractivity can be found in [107, 98]. We intend to treat such more general fixed point theorems based upon the more complicated notions of contractivity discussed in these survey articles in a subsequent paper[6].

Edelstein defines contractive (self-)mappings as follows:

Definition 8.20 (Edelstein[30]). *A self-mapping f of a metric space (X, d) is contractive if for all $x, y \in X$: $x \neq y \to d(f(x), f(y)) < d(x, y)$.*

Edelstein's Fixed Point Theorem is:

Theorem 8.21 (Edelstein[30]). *Let (X, d) be a complete metric space, let f be a contractive self-mapping on X and suppose that for some $x \in X$ the sequence $\{f^n(x)\}$ has a convergent subsequence $\{f^{n_i}(x)\}$. Then $\xi = \lim_{n\to\infty} f^n(x)$ exists and is a unique fixed point of f.*

For a compact space (X, d) the sequence $\{f^n(x)\}$ always has a convergent subsequence, and thus $\{f^n(x)\}$ always converges to a unique fixed point. We are now interested in obtaining a computable (Cauchy) modulus δ for the sequence $\{f^n(x)\}$ s.t. $\forall m, n > N : d(f^m(x), f^n(x)) < \varepsilon$ for $N := \delta(\varepsilon)$. In addition to ε, we must prima facie expect the rate of convergence δ to also depend on x, the space (X, d), the function f and a modulus of contractivity for f, if such a modulus exists. In an intuitionistic setting the meaning of the implication expressing the contractivity of f is to give a procedure to transform a witness of '$d(x, y) > 0$' into a witness of '$d(f(x), f(y)) < d(x, y)$'. Proving (or assuming) contractivity of f in an intuitionistic setting yields a function that depending on x, y and an ε, by which $d(x, y)$ is larger than 0, produces an η by which

[6]Note added in proof. Meanwhile Briseid carried this out for fixed point theorem of Kincses and Totik corresponding to the most general notion of contractivity considered by Rhoades. See Ref. [18]

$d(f(x), f(y))$ is smaller than $d(x, y)$. Such a function, if uniform with regard to $x, y \in X$, is none other than a modulus of contractivity.

Remark. *On compact metric spaces or, more generally, on bounded metric spaces, monotone functional interpretation and monotone modified realizability automatically strengthen the general notion of contractivity to uniform contractivity, i.e. the existence of a modulus of contractivity. As we will see, the notion of uniform contractivity is sufficient even on unbounded metric spaces to guarantee the convergence of* $\{f^n(x)\}$ *to a unique fixed point and to state an effective rate of convergence.*

In [104] Rakotch considers functions with a multiplicative modulus of contractivity α s.t.

$$\forall x, y \in X : d(x, y) > \varepsilon \to d(f(x), f(y)) \leq \alpha(\varepsilon) \cdot d(x, y)$$

where $0 \leq \alpha(\varepsilon) < 1$ for all $\varepsilon > 0$.[7] Note that the existence of such a modulus α is a uniform version of Edelstein's notion of contractivity as α does not depend on x, y but only on ε.

Rakotch's multiplicative modulus of contractivity α is only one possible interpretation of witnessing the contractive inequality. From the point of view of logic, to witness an inequality $s < t$ one has to produce an $\varepsilon > 0$ s.t. $s + \varepsilon < t$. This leads to a additive modulus of contractivity η s.t.

$$\forall x, y \in X : d(x, y) > \varepsilon \to d(f(x), f(y)) + \eta(\varepsilon) \leq d(x, y)$$

It is easy to see that a modulus η can always be defined given a modulus α:

$$\eta(\varepsilon) := (1 - \alpha(\varepsilon)) \cdot \varepsilon$$

To define a modulus α in terms of a modulus η we have to assume that the metric d on X is bounded and define:

$$\alpha(\varepsilon) := 1 - \frac{\eta(\varepsilon)}{b}$$

As Rakotch has shown (see below) the existence of a modulus of contractivity α implies that the iteration sequence $\{f^n(x)\}$ is bounded. From this he concludes that even without assuming the boundedness of X the sequence $\{f^n(x)\}$ is Cauchy (and hence converges to a unique fixed point of f).[8] As we will see,

[7] Actually Rakotch requires α to be monotonically decreasing and to satisfy $x \neq y \to d(f(x), f(y)) \leq \alpha(d(x, y)) \cdot d(x, y)$ instead. In the proof only the above property is needed, which follows from Rakotch's requirements.

[8] With a somewhat different proof one can also show this based on an additive modulus η instead of α although to derive the existence of a global modulus α from η seems to require the boundedness of (X, d). However, as Rakotch's proof shows, the contractvity is (for given x) used only on points of the form $f^n(x)$ and on those (by the boundedness of $\{f^n(x)\}$) one can define a modulus α from η.

by 8.18 this yields the existence of a uniform Cauchy modulus which is largely independent from the starting point x and the function f but only depends on the modulus α, a bound b on $d(x, f(x))$ and the error ε.

It should be noted that it is strictly necessary for the modulus α to be uniform with regard to $x, y \in X$, as otherwise a function, although contractive, might not have a fixed point. Edelstein's non-uniform notion of contractivity $x \neq y \to d(f(x), f(y)) < d(x, y)$ is in general only sufficient to prove the existence of a fixed point in compact spaces, where that notion is equivalent to the existence of uniform moduli α and η. In most other cases the equivalence fails. As a counterexample, consider the self-mapping $f(x) := x + \frac{1}{x}$ of the interval $[1, \infty)$. It is easy to see that the function f is contractive in the sense of Edelstein. Trivially, the function f has no fixed point. One, furthermore, proves by induction that for all $n \geq 1$:

$$1 + \sum_{i=1}^{n} \frac{1}{i} \leq f^n(1) \leq n + 1$$

Since $\sum_{i=1}^{\infty} \frac{1}{i} = \infty$, the iteration sequence $\{f^n(1)\}$ is unbounded. So by the aforementioned result of Rakotch, f does not have a modulus of contractivity α (as can be also seen directly). Counterexamples even in the case of **bounded** metric spaces[9] are discussed in [113].

Using a multiplicative modulus α, Rakotch proves the following variant of Edelstein's Fixed Point Theorem:

Theorem 8.22 (**Rakotch [104]).** *Let (X, d) be a complete metric space and let f be a contractive self-mapping on X with modulus of contractivity α, then $\xi = \lim_{n\to\infty} f^n(x)$ exists and is a unique fixed point of f.*

Remark. *Whereas Edelstein's theorem requires the existence of a convergent subsequence of $\{f^n(x)\}$, which is guaranteed in general only for compact X, Rakotch's theorem avoids this by imposing a stronger uniform contractivity on f (which, however, follows from the usual one in the compact case).*

The key step in the proof is to establish the following:

Lemma 8.23. *Let (X, d) be a metric space and let f be a contractive self-mapping on X with modulus of contractivity α, then the iteration sequence $\{f^n(x)\}$ is a Cauchy sequence.*

We now expect that our metatheorems allow us to extract from a proof of Lemma 8.23 a Cauchy modulus δ; in fact it suffices to extract a bound on the modulus, as such a bound trivially also is a realizer for the modulus. Contrary to Rakotch's proof, Edelstein's original proof is a classical proof and since expressing that the sequence $\{f^n(x)\}$ is a Cauchy sequence requires a Π^0_3-statement, the metatheorem for the classical case cannot be applied directly to extract a Cauchy modulus from Edelstein's proof.

[9] In fact even in the case of the closed unit ball of the Banach space c_0.

In [81], Kohlenbach and Oliva use a trick to extract a bound from Edelstein's non-constructive proof: The proof of Edelstein's Fixed point theorem can be split up into three lemmas. Each of these lemmas is of a suitable logical form to allow extraction of a bound, and combining these bounds, the following modulus of convergence (towards the unique fixed point) for f a self-map on a compact space K is extracted[10]:

$$\delta(\alpha, b, \varepsilon) = \left\lceil \frac{\log((1-\alpha(\varepsilon))\frac{\varepsilon}{2}) - \log b}{\log \alpha((1-\alpha(\varepsilon))\frac{\varepsilon}{2})} \right\rceil + 1$$

where α is the modulus of contractivity for f, and b is a bound on the diameter of K. In accordance with Theorem 8.6, the same bound also holds if we replace the compact space K by a (more general) b-bounded metric space. Note that the Cauchy modulus δ is uniform with regard to $x \in X$ and the function f.

The treatment of (the classical proof of) Edelstein's fixed point theorem in [81] via monotone functional interpretation generalizes Edelstein's result to bounded metric spaces, where using the strengthening of contractivity to uniform contractivity a Cauchy modulus for the sequence $\{f^n(x)\}$ is extracted. Together with the observation that only the boundedness of the iteration sequence is needed and not the boundedness of the whole space, the analysis of Edelstein's classical, non-constructive proof yields essentially the same result as Rakotch's theorem. However, with regard to the numerical quality of the modulus one can do better: As mentioned Rakotch's proof is fully constructive, and one easily sees that the constructive proof can be formalized in $\mathcal{A}_i^\omega[X,d]_{-b}$. Thus, without the tedious work of splitting up Edelstein's proof, the metatheorem for the semi-intuitionistic case guarantees that we can extract an effective bound on the modulus of convergence or, without having to carry out the extraction, prove uniformities for the modulus of convergence.

In $\mathcal{A}_i^\omega[X,d]_{-b}$ we can express the fact that $f^{X\to X}$ represents a contractive function with modulus α^1 (of type degree 1), in short: 'f contr. α', as

$$\forall k^0 \forall x^X, y^X (d_X(x,y) \geq_{\mathbb{R}} 2^{-k} \to d_X(f(x), f(y)) \leq_{\mathbb{R}} (1 - 2^{-\alpha(k)}) \cdot_{\mathbb{R}} d_X(x,y))$$

Thus in the formal system $\mathcal{A}_i^\omega[X,d]_{-b}$ one can express Lemma 8.23 as:

Lemma 8.24. *$\mathcal{A}_i^\omega[X,d]_{-b}$ proves*

$$\forall f^{X\to X} \forall x^X \forall \alpha^1 \forall k^0 (f \text{ contr. } \alpha \to \exists N^0 \forall m, n \geq_0 N\ d_X(f^m(x), f^n(x)) \leq_{\mathbb{R}} 2^{-k}).$$

To see that Rakotch's proof can be formalized in $\mathcal{A}_i^\omega[X,d]_{-b}$, one notes that the proof consists of two main parts: first it is shown that for any starting point x the sequence $\{f^n(x)\}$ is bounded and that the bound depends only on α and (a bound b on) $d(x, f(x))$. Given a starting point x, the function f and an

[10] Originally in [81] an additive modulus of contractivity η is considered. The extracted modulus of convergence is then $\delta(\eta, b, \varepsilon) = \left\lceil \frac{b - \frac{\eta(\varepsilon)}{2}}{\eta(\frac{\eta(\varepsilon)}{2})} \right\rceil + 1$.

arbitrary $\rho > 0$, Rakotch shows that one can bound $d(x, f^n(x))$ for all n by[11]

$$d(x, f^n(x)) \leq b'(\alpha, b) = \max(\rho, \frac{2 \cdot b}{1 - \alpha(\rho)}),$$

where $b \geq d(x, f(x))$.
Then using this bound and the contractivity of f it is shown that $\{f^n(x)\}$ is a Cauchy sequence and hence converges to a unique fixed point.

Application 8.25. *Corollary 8.18 a-priorily guarantees that there exists a bound $\delta(\alpha, b, \varepsilon)$ on N that holds for all metric spaces (X, d), all functions f with modulus of contractivity α and all $x \in X$ s.t. $d(x, f(x)) \leq b$. Moreover, by Corollary 8.18 we can extract an effective bound $\delta(\alpha, b, \varepsilon)$ from Rakotch's constructive proof, and since a bound on N also is a realizer, this gives us the following Cauchy modulus (and hence modulus of convergence towards the unique fixed point):*

$$\delta(\alpha, b, \varepsilon) = \left\lceil \frac{\log \varepsilon - \log b'(\alpha, b)}{\log \alpha(\varepsilon)} \right\rceil \quad \textit{where}$$
$$b'(\alpha, b) = \max(\rho, \tfrac{2 \cdot b}{1 - \alpha(\rho)}) \textit{ with } b \geq d(x, f(x)) \textit{ and } \rho > 0 \textit{ arbitrary }.$$

Proof. Since the relation $\leq_{\mathbb{R}}$ can be expressed as a Π^0_1-predicate, the premise 'f contr. α' is $\exists$-free, where α is an element of the Baire space $X = \mathbb{N}^{\mathbb{N}}$. Moreover, by the comment after corollary 8.18, we can take $\Omega(n) := n + b$ since f a-fortiori is nonexpansive. The conclusion, the Cauchy property of the sequence $\{f^n(x)\}$ is of the form $\forall\exists\forall$, but contrary to the classical case there are no restrictions on the logical form, so that we can extract an effective uniform bound $\delta(\alpha, b, \varepsilon)$ on $\exists N$, i.e. an effective uniform Cauchy modulus for $(f^n(x))$.

The existence of the Cauchy modulus δ, with the described uniformities, is guaranteed by the semi-intuitionistic metatheorem, even without analyzing the proof. For the actual "extraction" of a bound $\delta(\alpha, b, \varepsilon)$, we briefly sketch the relevant, second part of Rakotch's proof:

Let $p \in \mathbb{N}$ be given, then by definition (we can assume $d(x_k, x_{k+p}) > 0$):

$$d(x_{k+1}, x_{k+p+1}) \leq \alpha(d(x_k, x_{k+p})) \cdot d(x_k, x_{k+p}).$$

Now taking the product from $k = 0$ to $n - 1$ we get

$$d(x_n, x_{n+p}) \leq d(x_0, x_p) \cdot \prod_{k=0}^{n-1} \alpha(d(x_k, x_{k+p})).$$

Since we assumed $d(x, f(x)) \leq b$ and hence $b'(\alpha, b)$ is a bound on $d(x_0, x_p)$, we get

$$d(x_n, x_{n+p}) \leq b'(\alpha, b) \cdot \prod_{k=0}^{n-1} \alpha(d(x_k, x_{k+p})).$$

[11] Here for convenience we tacitly move back to the more usual version of α as a function $\mathbb{R}^*_+ \to (0, 1)$.

If already $d(x_k, x_{k+p}) < \varepsilon$ for some $0 \leq k \leq n-1$ we would be done, so assuming $d(x_k, x_{k+p}) \geq \varepsilon$ for all $k = 0, \ldots, n-1$ and by

$$\forall x, y \in X : d(x, y) \geq \varepsilon \to d(f(x), f(y)) \leq \alpha(\varepsilon) \cdot d(x, y)$$

we get that

$$d(x_n, x_{n+p}) \leq b'(\alpha, b) \cdot (\alpha(\varepsilon))^n.$$

Then solving the inequality $b'(\alpha, b) \cdot (\alpha(\varepsilon))^n \leq \varepsilon$ with regard to n yields the following Cauchy modulus:

$$\delta(\alpha, b, \varepsilon) = \left\lceil \frac{\log \varepsilon - \log b'(\alpha, b)}{\log \alpha(\varepsilon)} \right\rceil$$

where throughout $b'(\alpha, b)$ is as described above. □

As mentioned above, extracting a bound from the classical proof of Edelstein's theorem was only possible by breaking up the proof into a couple of lemmas, each of suitable form to extract a bound, using the metatheorem for the classical case. Compared to the bound extracted from the Edelstein's proof the bound from Rakotch's constructive proof - guaranteed a-priorily by the metatheorem to exist and to be uniform on $x \in X$ and f - is both (syntactically) simpler and better. Naturally, in many cases finding a constructive proof for a classically true theorem may be far less trivial than in the case of Rakotch's variant of Edelstein's theorem and, in general, many classically true theorems may not have a constructive proof at all. However, as this example demonstrates, considering a constructive proof may yield significantly simpler and better bounds than in the classical case and may give fully uniform bounds from theorems having a logical form more complex than $\forall\exists$, where the classical metatheorem in general fails, such as for example the Cauchy property of an iteration sequence. Moreover, monotone functional interpretation or monotone modified realizability may automatically lead to the necessary strengthenings of the mathematical notions involved, as e.g. strengthening the notion of contractivity to uniform contractivity.

Finally, even for proofs that are developed in a fully constructive setting, the metatheorem for the semi-constructive case may reveal new uniformities not present in, or immediately obvious from, the theorem and proof under consideration. In [16] Bridges et al. treat Edelstein's fixed point theorem in the framework of Bishop-style constructive mathematics. A function f that is contractive in the sense of Rakotch is denoted by the concept of 'f is an almost uniform contraction'. The following theorem is proved:

Theorem 8.26 ([16]). *Let $f : X \to X$ be an almost uniform contraction on a complete metric space X. Then*

1. *f has a unique fixed point ξ in X; and*
2. *the sequence $\{f^n(x)\}$ converges to ξ uniformly on each bounded subset of X.*

This theorem largely corresponds to Rakotch's theorem discussed above, but only the uniformity with regard to $x \in X$ is stated, not the uniformity with regard to f or the bounded subset. Both uniformities follow already a-priorily from the existence of a (constructive) proof for Rakotch's theorem by means of our metatheorem. Also a modulus of convergence is not explicitly stated, though both the uniformities and the effective modulus can be seen to be implicit in the proof. An analysis of the constructive proof in [16] easily yields an explicit modulus of convergence, which is identical to the bound extracted from Rakotch's constructive proof.

Chapter 9

General logical metatheorems for functional analysis

The paper *General logical metatheorems for functional analysis* has been submitted for publication. The paper is joint work with U.Kohlenbach and has been slightly reformatted for inclusion in this PhD-thesis.

Abstract

In this paper we prove general logical metatheorems which state that for large classes of theorems and proofs in (nonlinear) functional analysis it is possible to extract from the proofs effective bounds which depend only on very sparse local bounds on certain parameters. This means that the bounds are uniform for all parameters meeting these weak local boundedness conditions. The results vastly generalize related theorems due to the second author where the global boundedness of the underlying metric space (resp. a convex subset of a normed space) was assumed. Our results treat general classes of spaces such as metric, hyperbolic, CAT(0), normed, uniformly convex and inner product spaces and classes of functions such as nonexpansive, Hölder-Lipschitz, uniformly continuous, bounded and weakly quasi-nonexpansive ones. We give several applications in the area of metric fixed point theory. In particular, we show that the uniformities observed in a number of recently found effective bounds (by proof theoretic analysis) can be seen as instances of our general logical results.

9.1 Introduction

In [77], the second author established - as part of a general project of applied proof theory - logical metatheorems which guarantee a-priorily the extractability of uniform bounds from large classes of proofs in functional analysis. 'Uniformity' here refers to the independence of the bounds from parameters ranging over compact subspaces (in the case of concrete Polish spaces) as well as abstract bounded (not necessarily compact!) metric spaces or bounded convex subsets of hyperbolic, CAT(0), normed, uniformly convex or inner product spaces. By 'abstract' spaces we mean that the proofs only use the general axioms for e.g. metric or hyperbolic spaces. If these axioms have a strong uniformity built-in (as in the classes just mentioned), then this property prevails also for theorems proved in strong theories based on these axioms. The metatheorems were derived using a monotone proof interpretation, namely an extension of Gödel's so-called functional interpretation combined with a novel form of majorizability over function spaces of arbitrary types. The theorems were applied to results in metric fixed point theory to explain the extractability of strong uniform bounds that had been observed previously in several concrete cases ([73, 75, 79]) as well as to predict new such bounds which subsequently could, indeed, be found following the extraction algorithm provided by monotone functional interpretation ([78, 76]).

In the concrete applications it usually turned out that instead of the assumption of the whole space or some convex subset being bounded only some sparse local boundedness conditions were actually needed. This observation was the starting point of the present paper which establishes far reaching extensions of the results from [77] to unbounded spaces which guarantee uniform bounds under exactly such limited local boundedness assumptions. As we will show below, in most applications our new metatheorems completely close the gap which was left between the conclusions predicted by the old metatheorems and the general form of actual bounds constructed in the case studies. In particular, we now for the first time can explain a quantitative version of a well-known theorem of Borwein, Reich and Shafrir [14] on Krasnoselski-Mann iterations of nonexpansive mappings in unbounded hyperbolic spaces, which was found in [79], as an instance of the new metatheorems. The proofs still use a combination of functional interpretation and majorization but this time in a much more subtle way: both the functional interpretation as well as the majorization relation to be applied are parametrized by a point a of the space X in question. In the applications we will be able to achieve by a suitable choice of a (which in turn depends on the parameters of the problem) that the object constructed by the a-functional interpretation can be a-majorized by a term which no longer depends on a (nor the parameters involving the space X). This applies, furthermore, to large classes of mappings between such spaces, as e.g. nonexpansive, weakly quasi-nonexpansive, Lipschitz-Hölder, uniformly continuous or bounded mappings.

The results in this paper not only allow one to strengthen known existence results in functional analysis by establishing qualitatively new forms of uniform existence as well as new quantitative bounds, but also by weakening of the assumptions needed. E.g. assumptions of the form 'f has a fixed point' can for large classes of proofs and theorems be replaced by the much weaker assumption 'f has approximate fixed points'. Finally, we will indicate how our results extend to contexts where several spaces $X_1, \ldots, X_n$ from the aforementioned classes of spaces as well as their products are simultanously present. We are confident that these results will have many more applications also outside the context of fixed point theory (see e.g. [81] for a survey of different topics to which this kind of 'proof mining' approach can be applied).

9.2 Definitions

The classes of general spaces we are considering are metric spaces, hyperbolic spaces (including CAT(0)-spaces) as well as normed spaces. Under a hyperbolic space we understand the following:

Definition 9.1. *(X, d, W) is called a hyperbolic space if (X, d) is a metric space and $W : X \times X \times [0, 1] \to X$ a function satisfying*

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

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

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

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

Remark. *As discussed in detail in [77], we obtain Takahashi's [116] 'convex metric spaces' if the axioms (ii)-(iv) are dropped, and a notion which is equivalent to the concept of 'space of hyperbolic type' from [45] if we drop only (iv). As* *observed in [105, 14] and [111], several arguments in metric fixed point theory require a bit more of linear structure which gave rise to a notion of 'hyperbolic space'*[1] *in [57, 105] which adds axiom (iv) (for $\lambda := \frac{1}{2}$) and the requirement that any two points not only are connected by a metric segment but by a metric line. As a consequence of this (just as in the case of normed spaces) nontrivial hyperbolic spaces in the sense of [57, 105] always are unbounded and convex subsets of a hyperbolic space in general are no longer hyperbolic spaces themselves. The existence of metric lines allows one to derive the general axiom (iv) from the special case of $\lambda = \frac{1}{2}$. It turns out that if we state (iv) directly for general λ as above then the proofs in metric fixed point theory we are interested in (e.g. the main results in [14] and [111]) all go through for our more liberal notion of 'hyperbolic space' which not only has a simpler logical structure but also includes all convex subsets of hyperbolic (and in particular normed) spaces as well as all CAT(0)-spaces, whereas the more restricted notion used in [57, 105, 111] only covers CAT(0)-spaces having the geodesic line extension property (see [17] for details on CAT(0)-spaces).*

As carried out in detail in [77] we formalize our classes of spaces on top of a formal system $\mathcal{A}^\omega$ of classical analysis which is based on a language of functionals of finite type.

Definition 9.2. *The set* $\mathbf{T}$ *of all finite types is defined inductively over the ground type* 0 *by the clauses*

$$(i)\ 0 \in \mathbf{T},\ (ii)\ \rho, \tau \in \mathbf{T} \Rightarrow (\rho \to \tau) \in \mathbf{T}.$$

The formal system $\mathcal{A}^\omega$ for analysis (which is based on the axioms of countable and dependent choice which, in particular, yield full comprehension for numbers) is defined as in [77]. Higher type equality is not a primitive predicate but defined extensionally. Instead of the full axiom of extensionality in all types, the system $\mathcal{A}^\omega$ only has a quantifier-free rule of extensionality.[2]

Before we can describe the extensions $\mathcal{A}^\omega[X,d], \mathcal{A}^\omega[X,d,W]$ etc. of $\mathcal{A}^\omega$ by an abstract metric space (X,d) or hyperbolic space (X,d,W) we briefly have to recall the representation of real numbers in the formal system $\mathcal{A}^\omega$:

In our formal systems based on $\mathcal{A}^\omega$, real numbers are represented by Cauchy sequences $(a_n)_n$ of rational numbers with Cauchy modulus 2^{-n}, i.e.

$$\forall m,n(m,n \geq k \to |a_m - a_n| < 2^{-k}).$$

Rational numbers are represented as pairs (n,m) of natural numbers coded into a single natural number $j(n,m)$, where j is the Cantor pairing function. If n is even $j(n,m)$ represents the rational number $\frac{n/2}{m+1}$, if n is odd $j(n,m)$ represents the negative number $-\frac{(n+1)/2}{m+1}$. Thus every natural number can be

[1] Unfortunately, Kirk calls this notion in [57] 'space of hyperbolic type' although it differs from the definition of the latter in [45].

[2] See [77] for an extensive discussion of this crucial point.

conceived as the code of a unique rational number. An equality relation $=_{\mathbb{Q}}$ on the representatives of the rational numbers, as well as the usual operators $+_{\mathbb{Q}}, -_{\mathbb{Q}}, \cdot_{\mathbb{Q}}$, etc. and the predicates $<_{\mathbb{Q}}, \leq_{\mathbb{Q}}$ are defined (primitive recursively) in the obvious way. Thus natural and rational numbers are represented by objects of type 0 and sequences of rational numbers by objects of type 1, i.e. by functions of type $0 \to 0$.

Real numbers are represented by functions $f : \mathbb{N} \to \mathbb{N}$ (i.e of type 1) s.t.

$$\forall n(|f(n) -_{\mathbb{Q}} f(n+1)| <_{\mathbb{Q}} 2^{-n-1}). \quad (*)$$

To ensure that each function $f : \mathbb{N} \to \mathbb{N}$ represents a real number we use the following construction:

$$\widehat{f}(n) := \begin{cases} f(n) \text{ if } \forall k < n\big(|f(k) -_{\mathbb{Q}} f(k+1)|_{\mathbb{Q}} <_{\mathbb{Q}} 2^{-k-1}\big), \\ f(k) \text{ for } \min k < n \text{ with } |f(k) -_{\mathbb{Q}} f(k+1)|_{\mathbb{Q}} \geq_{\mathbb{Q}} 2^{-k-1} \text{ otherwise.} \end{cases}$$

For better readability we will usually write e.g. 2^{-n} instead of its (canonical) code $\langle 2^{-n} \rangle := j(2, 2^n - 1)$.
For every $f : \mathbb{N} \to \mathbb{N}$ the construction $\widehat{f}$, which can be carried out in $\mathcal{A}^\omega$, satisfies $(*)$, and if already f satisfies $(*)$ then $\forall n(f(n) =_0 \widehat{f}(n))$. Thus every f codes a unique real number, namely the one given by the Cauchy sequence coded by $\widehat{f}$. The construction $f \mapsto \widehat{f}$ enables us to reduce quantifiers ranging over $\mathbb{R}$ to $\forall f^1$, resp. $\exists f^1$, without introducing additional quantifiers. For natural numbers $b \in \mathbb{N}$ we have the embedding $(b)_{\mathbb{R}}$ via the constructing $(b)_{\mathbb{R}} =_1 \lambda n.j(2b, 0)$.

The equivalence relation $=_{\mathbb{R}}$ and the relations $\leq_{\mathbb{R}}$ and $<_{\mathbb{R}}$ on (representatives of) real numbers are defined notions. The relations $=_{\mathbb{R}}$ and $\leq_{\mathbb{R}}$ are given by Π^0_1-predicates while $<_{\mathbb{R}}$ is given by a Σ^0_1-predicate:

$$\begin{aligned} f_1 =_{\mathbb{R}} f_2 \quad &:\equiv \forall n(|\widehat{f_1}(n+1) -_{\mathbb{Q}} \widehat{f_2}(n+1)| <_{\mathbb{Q}} 2^{-n}) \\ f_1 <_{\mathbb{R}} f_2 \quad &:\equiv \exists n(\widehat{f_2}(n+1) -_{\mathbb{Q}} \widehat{f_1}(n+1) \geq_{\mathbb{Q}} 2^{-n}) \\ f_1 \leq_{\mathbb{R}} f_2 \quad &:\equiv \neg(f_2 <_{\mathbb{R}} f_1) \end{aligned}$$

The operators $+_{\mathbb{R}}, -_{\mathbb{R}}, \cdot_{\mathbb{R}}$,etc. on representatives of real numbers can be defined by simple primitive recursive functionals. For further details see [77].

For the interval $[0, 1]$, which plays an important role in the formal treatment of hyperbolic spaces, we use a special representation by number theoretic functions $\mathbb{N} \to \mathbb{N}$(which are bounded by a fixed function M):

Definition 9.3.

$$\tilde{x}(n) := j(2k_0, 2^{n+2} - 1), \ \textit{where } k_0 = \max k \leq 2^{n+2}[\frac{k}{2^{n+2}} \leq_{\mathbb{Q}} \widehat{x}(n+2)].$$

One easily verifies the following:

Lemma 9.4. *Provably in $\mathcal{A}^\omega$, for all x^1:*

1. $0_{\mathbb{R}} \leq_{\mathbb{R}} x \leq_{\mathbb{R}} 1_{\mathbb{R}} \to \tilde{x} =_{\mathbb{R}} x$,

2. $0_{\mathbb{R}} \leq_{\mathbb{R}} \tilde{x} \leq_{\mathbb{R}} 1_{\mathbb{R}}$,

3. $\tilde{x} \leq_1 M := \lambda n.j(2^{n+3}, 2^{n+2} - 1)$,

4. $x >_{\mathbb{R}} 1_{\mathbb{R}} \to \tilde{x} =_{\mathbb{R}} 1_{\mathbb{R}},\ x <_{\mathbb{R}} 0_{\mathbb{R}} \to \tilde{x} =_{\mathbb{R}} 0_{\mathbb{R}}$.

The theories of classical analysis extended with metric or normed linear spaces and their variants are defined almost as in [77]. The crucial difference is that while in [77] only bounded metric spaces (X, d) and bounded convex subsets C of normed linear spaces $(X, \|\cdot\|)$ are considered, we now permit *unbounded* metric spaces (X, d) and *unbounded* convex subsets C. In [77], the boundedness is expressed by an axiom stating explicitly that (X, d), resp. the convex subsets C, are bounded by b. In our unbounded variants we omit this axiom. To distinguish the not-b-bounded theories from the b-bounded theories $\mathcal{A}^\omega[X, d]$ and $\mathcal{A}^\omega[X, \|\cdot\|, C]$ defined in [77], we will denote them by $\mathcal{A}^\omega[X, d]_{-b}$ and $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$. We also consider theories $\mathcal{A}^\omega[X, d, W]_{-b}$ capturing hyperbolic spaces and $\mathcal{A}^\omega[X, d, W, \mathrm{CAT}(0)]_{-b}$ capturing CAT(0)-spaces. More precisely, the theories $\mathcal{A}^\omega[X, d]_{-b}$, $\mathcal{A}^\omega[X, d, W]_{-b}$ and $\mathcal{A}^\omega[X, d, W, \mathrm{CAT}(0)]_{-b}$ result by

(i) extending $\mathcal{A}^\omega$ to the set $\mathbf{T}^X$ of all finite types over the two ground types 0 and X, i.e.

$$(i)\ 0, X \in \mathbf{T}^X,\ (ii)\ \rho, \tau \in \mathbf{T}^X \Rightarrow (\rho \to \tau) \in \mathbf{T}^X$$

(in particular, the constants $\Pi_{\rho,\tau}, \Sigma_{\delta,\rho,\tau}, \underline{R}_\rho$ for λ-abstraction and simultaneous primitive recursion (in the extended sense of Gödel [44]) and their defining axioms and the schemes IA (induction), QF-AC (quantifier-free choice in all types), DC (dependent countable choice) and the weak extensionality rule QF-ER are now taken over the extended language),

(ii) adding a constant 0_X of type X,

(iii) adding a new constant d_X of type $X \to X \to 1$ (representing the metric) together with the axioms

(1) $\forall x^X (d_X(x, x) =_{\mathbb{R}} 0_{\mathbb{R}})$,
(2) $\forall x^X, y^X \big(d_X(x, y) =_{\mathbb{R}} d_X(y, x)\big)$,
(3) $\forall x^X, y^X, z^X \big(d_X(x, z) \leq_{\mathbb{R}} d_X(x, y) +_{\mathbb{R}} d_X(y, z)\big)$.

In these axioms we refer to the representation of real numbers (including the definition of $=_{\mathbb{R}}, \leq_{\mathbb{R}}$) as sketched above.

Equality $=_0$ at type 0 is the only a primitive equality predicate. $x^X =_X y^X$ is defined as $d_X(x, y) =_{\mathbb{R}} 0_{\mathbb{R}}$. Equality for complex types is defined as before as extensional equality using $=_0$ and $=_X$ for the base cases.

$\mathcal{A}^\omega[X, d, W]_{-b}$ results from $\mathcal{A}^\omega[X, d]_{-b}$ by adding a new constant W_X of type $X \to X \to 1 \to X$ together with the axioms (where $\tilde{\lambda}$ is defined as above)

(4) $\forall x^X, y^X, z^X \forall \lambda^1 \big(d_X(z, W_X(x, y, \lambda)) \leq_{\mathbb{R}} (1_{\mathbb{R}} -_{\mathbb{R}} \tilde{\lambda}) d_X(z, x) +_{\mathbb{R}} \tilde{\lambda} d_X(z, y)\big)$,

(5) $\forall x^X, y^X \forall \lambda_1^1, \lambda_2^1 \big(d_X(W_X(x,y,\lambda_1), W_X(x,y,\lambda_2)) =_{\mathbb{R}} |\tilde{\lambda}_1 -_{\mathbb{R}} \tilde{\lambda}_2|_{\mathbb{R}} \cdot_{\mathbb{R}} d_X(x,y) \big)$,

(6) $\forall x^X, y^X \forall \lambda^1 \big(W_X(x,y,\lambda) =_X W_X(y,x,(1_{\mathbb{R}} -_{\mathbb{R}} \lambda)) \big)$,

(7) $\begin{cases} \forall x^X, y^X, z^X, w^X, \lambda^1 \\ \big(d_X(W_X(x,z,\lambda), W_X(y,w,\lambda)) \leq_{\mathbb{R}} (1_{\mathbb{R}} -_{\mathbb{R}} \tilde{\lambda}) d_X(x,y) +_{\mathbb{R}} \tilde{\lambda} d_X(z,w) \big). \end{cases}$

$\mathcal{A}^\omega[X,d,W,\mathrm{CAT}(0)]_{-b}$ results from $\mathcal{A}^\omega[X,d,W]_{-b}$ by adding as further axiom the formalized form of the Bruhat-Tits or $\mathbf{CN}^-$-inequality[21], i.e.

$$\forall x^X, y_1^X, y_2^X \big(d_X(x, W_X(y_1, y_2, \frac{1}{2}))^2 \leq_{\mathbb{R}} \frac{1}{2} d_X(x,y_1)^2 +_{\mathbb{R}} \frac{1}{2} d_X(x,y_2)^2 -_{\mathbb{R}} \frac{1}{4} d_X(y_1,y_2)^2 \big).$$

Remark. *1. The additional axioms of $\mathcal{A}^\omega[X,d]_{-b}$ express (modulo our representation of $\mathbb{R}$ sketched above) that d_X represents a pseudo-metric d (on the universe the type-X variables are ranging over).*[3] *Hence d_X represents a metric on the set of equivalence classes generated by $=_X$. We do not form these equivalence classes explicitly but talk instead only about representatives x^X, y^X. However, it is important to stress that a functional $f^{X \to X}$ represents a function $X \to X$ only if it respects this equivalence relation, i.e.*

$$\forall x^X, y^X (x =_X y \to f(x) =_X f(y)).$$

Due to our weak (quantifier-free) rule of extensionality we in general only can infer from a **proof** *of $s =_X t$ that $f(s) =_X f(t)$. The restriction on the availability of extensionality is crucial for our results to hold (see the discussion in [77]). However, the extensionality of the constants d_X, W_X as well as the constants for normed linear spaces can all be proved to be fully extensional from their defining axioms. Likewise, for most (but not all) of the classes of functions which we will consider below (notably nonexpansive functions) the full extensionlity will follow from their defining conditions.*

2. Our axiomatization of W_X given by the axioms (4)-(7) *differs slightly from the one given in [77]. Our present axiomatization is equivalent to the extension of the one given in [77] by the additional axiom*

$$W_X(x,y,\lambda) =_X W_X(x,y,\tilde{\lambda})$$

using the property

$$\widetilde{1 -_{\mathbb{R}} \lambda} =_{\mathbb{R}} 1 -_{\mathbb{R}} \tilde{\lambda}$$

of our operation $\lambda \mapsto \tilde{\lambda}$ (which follows using lemma 9.4.4). This additional axiom is trivially satisfied by the interpretation of W_X in the model $\mathcal{S}^{\omega,X}$ from [77] so that it can be added without causing problems. The benefit of this is that then the axioms on W_X can be stated in the simple form given above compared to the more complicated formulation in [77]. The intuitive interpretation of W_X in a hyperbolic spaces (X,d,W) is that for $x, y \in X$ we interpret $W_X(x,y,\lambda)$ by $W(x,y,r_{\tilde{\lambda}})$ where $r_{\tilde{\lambda}}$ is the unique real number in $[0,1]$ that is represented by $\tilde{\lambda}$.

[3]Note that (1) – (3) imply that $\forall x^X, y^X \Big(d_X(x,y) \geq_{\mathbb{R}} 0_{\mathbb{R}} \Big)$.

3. *Our additional axiom* $\mathbf{CN}^-$ *used in defining* $\mathcal{A}^\omega[X, d, W, CAT(0)]$ *differs from the one used in [77] (under the name* $\mathbf{CN}^*$*) but is in fact equivalent to the latter:* $\mathbf{CN}^-$ *is (over the* W*-axioms (i)-(iv)) equivalent to the more usual formulation* $\mathbf{CN}$ *of the Bruhat-Tits inequality[21] which states that every midpoint of* y_1, y_2 *satisfies the inequality stated for* $W(y_1, y_2, \frac{1}{2})$*. The latter point provably (in* $\mathcal{A}^\omega[X, d, W]$*) is a midpoint so that* $\mathbf{CN}$ *implies* $\mathbf{CN}^-$*. From* $\mathbf{CN}^-$ *is easily follows not only that every midpoint of* y_1, y_2 *has to coincide with* $W(y_1, y_2, \frac{1}{2})$ *(so that* $\mathbf{CN}$ *follows) but even the quantitative version* $\mathbf{CN}^*$ *of* $\mathbf{CN}$*. In [77],* $\mathbf{CN}^*$ *was used as axiom as it is (in contrast to* $\mathbf{CN}$*) purely universal which is crucially used in the proofs). However,* $\mathbf{CN}^-$ *is purely universal too and equivalent to* $\mathbf{CN}^*$*. Since it easier to state we use this formulation here.*

As for $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$, the corresponding theories for uniformly convex spaces, $\mathcal{A}^\omega[X, \|\cdot\|, C, \eta]_{-b}$, and real inner product spaces, $\mathcal{A}^\omega[X, \|\cdot\|, C, \langle\cdot,\cdot\rangle]_{-b}$, are defined as in [77] except that the axioms stating the boundedness of C is dropped.

Finally, various moduli naturally occurring in analysis, such as e.g. a modulus of uniform continuity or a modulus of uniform convexity, are also represented by number theoretic functions $\mathbb{N} \to \mathbb{N}$, i.e. objects of type 1. Thus e.g. the statement $f : X \to X$ is uniformly continuous with modulus $\omega : \mathbb{R} \to \mathbb{R}$:

$$\forall x, y \in X \forall \varepsilon > 0 (d(x, y) \leq \omega(\varepsilon) \to d(f(x), f(y)) \leq \varepsilon)$$

is translated into

$$\forall x, y \in X \forall k \in \mathbb{N} (d(x, y) < 2^{-\omega(k)} \to d(f(x), f(y)) \leq 2^{-k})$$

where $\omega : \mathbb{N} \to \mathbb{N}$ and the translated statement is purely universal.

9.3 A generalized approach to majorization

In [77] the strong majorization relation, first introduced by Bezem[12] for the finite types $\mathbf{T}$ over $\mathbb{N}$, is extended to the types $\mathbf{T}^X$ with the new ground type X for metric spaces (X, d) and normed linear spaces $(X, \|\cdot\|)$. Furthermore, a mapping $\widehat{\rho}$ between types $\rho \in \mathbf{T}^X$ and $\widehat{\rho} \in \mathbf{T}$, and a relation $\sim_\rho$ between functionals of type $\rho \in \mathbf{T}^X$ and $\widehat{\rho} \in \mathbf{T}$ are defined inductively. By relating the constants of the theories $\mathcal{A}^\omega[X, d]$ and $\mathcal{A}^\omega[X, \|\cdot\|]$ (and their variants) to suitable functionals in $\mathcal{A}^\omega$ via the relation $\sim_\rho$ one can, combined with majorization in the types $\mathbf{T}^X$, systematically eliminate the dependency on the type X in the extracted terms and obtain bounds independent of parameters ranging over bounded metric spaces, resp. bounded convex subsets of normed linear spaces.

In this section we present a generalized approach to extending the strong majorization relation to the types $\mathbf{T}^X$. The (strong) majorization relation was defined by Howard and Bezem:

Definition 9.5 (Howard-Bezem, [52, 12])**.** *The strong majorization relation* s-maj *over the finite types* $\mathbf{T}$ *is defined as follows:*

- x^* s-maj$_0$ $x :\equiv x^* \geq_{\mathbb{N}} x$, *where* $\geq_{\mathbb{N}}$ *is the usual (primitive recursive) order on* $\mathbb{N}$,
- x^* s-maj$_{\rho\to\tau}$ $x :\equiv \forall y^*, y(y^*$ s-maj$_\rho$ $y \to x^*y^*$ s-maj$_\tau$ $x^*y, xy)$.

In [77], two different approaches are employed for metric and normed linear spaces respectively to extend the majorization relation to the new type X. For metric spaces only the restricted case of b-bounded spaces is treated, where b is an integer upper bound on the metric of the space. For bounded metric spaces the relation s-maj is extended to the types $\mathbf{T}^X$ by defining:

$$x^* \text{ s-maj}_X\ x :\equiv (0 = 0), \text{ i.e. always true.}$$

Usually extending majorization to a new type X imposes a kind of order on the elements of X which the majorization of the constants $0_X, d_X$ and W_X must respect. Since here the metric d_X can be bounded independently of the elements $x, y \in X$ to which it is applied, namely by $\lambda x^X, y^X.(b)_{\mathbb{R}}$, and since the function W_X merely produces new elements of X, in [77] the majorization relation on X could be defined to be always true (corresponding to a trivial order on X).

For normed linear spaces this approach does not work as non-trivial normed linear spaces always are unbounded. Instead in [77] the extension of the majorization relation to the new type X for normed linear spaces $(X, \|\cdot\|)$ is defined via the norm:

$$x^* \text{ s-maj}_X\ x :\equiv \|x^*\|_X \geq_{\mathbb{R}} \|x\|_X.$$

The majorization of extracted terms in [77] then consists of three steps: First one majorizes the extracted terms in $\mathbf{T}^X$ – these majorants may still depend on the constants of $\mathcal{A}^\omega[X, \|\cdot\|]$. Next one eliminates the dependency on X using the relation $\sim_\rho$ and an ineffective operator $(\cdot)_\circ$ (to be defined below). Finally, the resulting terms are majorized once more in the types $\mathbf{T}$ to eliminate uses of the ineffective $(\cdot)_\circ$-operator.

As mentioned above, using these techniques it is possible to derive the independence of extracted bounds from parameters ranging over *bounded* metric spaces, resp. *norm-bounded* convex subsets C of normed linear spaces. The generalized approach to majorization we describe in this section aims to treat the more general cases of *unbounded* metric and hyperbolic spaces and normed linear spaces with *unbounded* convex subsets C, i.e. the theories $\mathcal{A}^\omega[X, d]_{-b}$, $\mathcal{A}^\omega[X, d, W]_{-b}$ and $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$, and to derive similar uniformities under certain local boundedness conditions to be discussed in detail later. This generalized approach is based upon the first two steps of the previous treatment of normed linear spaces: (strong) majorization in the types $\mathbf{T}^X$ and the relation $\sim_\rho$. In [77], in the mapping $\widehat{\cdot}$ the type X with (X, d) a metric space was mapped to the type 0, while the type X with $(X, \|\cdot\|)$ a normed linear space was mapped to the type 1. In this paper, we will map the type X to 0 in both cases:

Definition 9.6. *For* $\rho \in \mathbf{T}^X$ *we define* $\widehat{\rho} \in \mathbf{T}$ *inductively as follows*

$$\widehat{0} := 0,\ \widehat{X} := 0,\ \widehat{(\rho \to \tau)} := (\widehat{\rho} \to \widehat{\tau}),$$

i.e. $\widehat{\rho}$ *is the result of replacing all occurrences of the type* X *in* ρ *by the type* 0.

The generalized approach to majorization will again involve the $()_\circ$-operator, but restricted to cases where the circle operator is effectively computable. Hence, the second application of strong majorization, used in the previous treatment of normed linear spaces to get rid of ineffective instances of the $()_\circ$-operator, is no longer necessary.

Combining Bezem's notion of strong majorization s-maj and the idea of the relation $\sim_\rho$ we define a family of (majorization) relations $\gtrsim^a_\rho$ between objects of type $\rho \in \mathbf{T}^X$ and their majorants of type $\widehat{\rho} \in \mathbf{T}$. The relation is parametrized by an element $a \in X$, where X is the underlying metric or normed linear space and $a \in X$ serves as a reference point for comparing and majorizing elements of X. In $\mathcal{L}(\mathcal{A}[X,d])_{-b}$, resp. $\mathcal{L}(\mathcal{A}[X,\|\cdot\|])$, this is syntactically expressed as follows:

Definition 9.7. *We define a ternary relation $\gtrsim^a_\rho$ between objects x, y and a of type $\widehat{\rho}, \rho$ and X respectively by induction on ρ as follows:*

- $x^0 \gtrsim^a_0 y^0 :\equiv x \geq_{\mathbb{N}} y,$
- $x^0 \gtrsim^a_X y^X :\equiv (x)_{\mathbb{R}} \geq_{\mathbb{R}} d_X(y,a),$
- $x \gtrsim^a_{\rho\to\tau} y :\equiv \forall z', z(z' \gtrsim^a_\rho z \to xz' \gtrsim^a_\tau yz) \wedge \forall z', z(z' \gtrsim^a_{\widehat{\rho}} z \to xz' \gtrsim^a_{\widehat{\tau}} xz).$

For normed linear spaces we choose $a = 0_X$ [4]*, so that $d_X(x,a) =_{\mathbb{R}} \|x\|_X$.*

As $\gtrsim^a$ is a relation between objects of different types, the definition of $\gtrsim^a_{\rho\to\tau}$ is slightly more complicated than the corresponding definition of s-maj$_{\rho\to\tau}$. The first part of the clause ensures that x is a "majorant" for y, the second part ensures that a majorant x also majorizes itself. Since majorants are of type $\widehat{\rho} \in \mathbf{T}$ (where $\gtrsim^a_{\widehat{\rho}}$ coincides with s-maj$_{\widehat{\rho}}$), this corresponds to requiring that for all majorants x s-maj x, and so the definition of $\gtrsim^a_{\rho\to\tau}$ could equivalently be rewritten as:

$$x \gtrsim^a_{\rho\to\tau} y :\equiv \forall z', z(z' \gtrsim^a_\rho z \to xz' \gtrsim^a_\tau yz) \wedge x \text{ s-maj}_{\widehat{\rho}\to\widehat{\tau}}\, x.$$

Remark. *Restricted to the types $\mathbf{T}$ the relation $\gtrsim^a$ is identical with the Howard-Bezem notion of strong majorizability* s-maj *and hence for $\rho \in \mathbf{T}$ we may freely write s-maj$_\rho$ instead of $\gtrsim^a_\rho$, as here the parameter $a \in X$ is irrelevant. Without the requirement that "majorants" must be strongly self-majorizing, $\gtrsim^a$ restricted to $\mathbf{T}$ is identical with Howard's notion of majorizability* maj.

In the following, we call majorization in the sense of the relation $\gtrsim^a$ (strong) "a-majorization", i.e. if $t_1 \gtrsim^a t_2$ for terms t_1, t_2 we say that t_1 a-majorizes t_2 and we call t_1 an a-majorant. If neither term t_i depends on a we say that t_1 *uniformly* a-majorizes t_2. We will in general aim at uniform majorants so that we can choose a appropriately (without having an effect on the majorants of

[4] While it will turn out to be independent of the choice of a whether a given functional is a-majorizable or not, the choice of a is crucial to obtain "nice" majorants. See Section 9.9 for a detailed discussion.

the constants of our theories) to obtain bounds with the intended uniformity features.

For the normed case we also need a pointwise $\geq_\rho$ relation between functionals of type ρ :

Definition 9.8. *$\geq_\rho$ is a binary relation between functionals of of type ρ and which is defined by induction on ρ as follows:*

- $x^0 \geq_0 y^0 :\equiv x \geq_{\mathbb{N}} y$,
- $x^X \geq_X y^X :\equiv \|x\|_X \geq_{\mathbb{R}} \|y\|_X$,
- $x \geq_{\rho\to\tau} y :\equiv \forall z^\rho (xz \geq_\tau yz)$.

The only nontrivial relation between $\gtrsim^{0X}_\rho$ and $\geq_\rho$ which holds in all types is the following one:

Lemma 9.9. *For all x^*, x, y of type $\widehat{\rho}, \rho, \rho$ resp. the following holds (provably in $\mathcal{A}^\omega[X, \|\cdot\|]$):*

$$x^* \gtrsim^{0X}_\rho x \wedge x \geq_\rho y \to x^* \gtrsim^{0X}_\rho y.$$

Proof. Easy induction on ρ. □ □

9.4 Metatheorems for metric and hyperbolic spaces

Before we state the new metatheorems, we recall and add the following definitions:

Definition 9.10 ([77])**.** *We say that a type $\rho \in \mathbf{T}^X$ has degree*

- 1 *if $\rho = 0 \to \ldots \to 0$ (including $\rho = 0$),*
- $(0, X)$ *if $\rho = 0 \to \ldots \to 0 \to X$ (including $\rho = X$),*
- $(1, X)$ *if it has the form $\tau_1 \to \ldots \to \tau_k \to X$ (including $\rho = X$), where τ_i has degree 1 or $(0, X)$.*

Definition 9.11. *We say that a type $\rho \in \mathbf{T}^X$ has degree $\widehat{1}$, if $\widehat{\rho}$ has degree 1. Amongst others, the type degree $\widehat{1}$ covers types $\mathbb{N}, X, \mathbb{N} \to \mathbb{N}, \mathbb{N} \to X, X \to \mathbb{N}$ and $X \to X$.*

Definition 9.12. *A formula F is called a $\forall$-formula (resp. $\exists$-formula) if it has the form $F \equiv \forall \underline{a}^{\underline{\sigma}} F_{qf}(\underline{a})$ (resp. $F \equiv \exists \underline{a}^{\underline{\sigma}} F_{qf}(\underline{a})$) where F_{qf} does not contain any quantifiers and the types in $\underline{\sigma}$ are of degree $\widehat{1}$ or $(1, X)$.*

The $()_\circ$-operator is defined as follows:

Definition 9.13 ([77]). *For $x \in [0,\infty)$ define $(x)_\circ \in \mathbb{N}^\mathbb{N}$ by*

$$(x)_\circ(n) := j(2k_0, 2^{n+1}-1),$$

where

$$k_0 := \max k\Big[\frac{k}{2^{n+1}} \le x\Big].$$

Remark. $()_\circ$ *is a 'semantic' operator defined on the real numbers themselves (rather than representatives of real numbers). However, it has a counterpart $\circ : \mathbb{N}^\mathbb{N} \to \mathbb{N}^\mathbb{N}$ defined as $(f^1)_\circ = (|r|)_\circ$, where r is the real number represented by f, i.e. $(f^1)_\circ \in \mathbb{N}^\mathbb{N}$ is defined by*

$$(f)_\circ(n) := j(2k_0, 2^{n+1}-1),$$

where

$$k_0 := \max k\Big[\frac{k}{2^{n+1}} \le_\mathbb{R} |f|_\mathbb{R}\Big].$$

In contrast to $\tilde{f}^1$ defined before, this functional of type $1 \to 1$ is not computable, but in our bounds it will only be used in the form $\lambda n^0.((n)_\mathbb{R})_\circ$ which is (even primitive recursively) computable. It will be clear from the context whether we refer to $\circ$ defined on $[0,\infty)$ or on $\mathbb{N}^\mathbb{N}$.

We will use the following properties of the $()_\circ$-operator:

Lemma 9.14 ([77]). *1. If $x \in [0,\infty)$, then $(x)_\circ$ is a representative of x in the sense of the representation of real numbers described in Section 9.2.*

2. If $x, y \in [0,\infty)$ and $x \le y$ (in the sense of $\mathbb{R}$), then $(x)_\circ \le_\mathbb{R} (y)_\circ$ and also $(x)_\circ \le_1 (y)_\circ$ (i.e. $\forall n \in \mathbb{N}((x)_\circ(n) \le (y)_\circ(n))$).

3. If $x \in [0,\infty)$, then $(x)_\circ$ is monotone, i.e. $\forall n \in \mathbb{N}((x)_\circ(n) \le_0 (x)_\circ(n+1))$.

4. If $x, y \in [0,\infty)$ and $x \le y$ (in the sense of $\mathbb{R}$), then $(y)_\circ$ s-maj$_1 (x)_\circ$

Proof. 1.-3. are part of Lemma 2.10 in [77]. 4. follows from 2. and 3. □

Definition 9.15. *Let X be a nonempty set. The full set-theoretic type structure $\mathcal{S}^{\omega,X} := \langle S_\rho \rangle_{\rho \in \mathbf{T}^X}$ over $\mathbb{N}$ and X is defined by*

$$S_0 := \mathbb{N}, \quad S_X := X, \quad S_{\rho\to\tau} := S_\tau^{S_\rho}.$$

Here $S_\tau^{S_\rho}$ is the set of all set-theoretic functions $S_\rho \to S_\tau$.

Using this and the $()_\circ$-operator we state the following definition:

Definition 9.16. *We say that a sentence of $\mathcal{L}(\mathcal{A}^\omega[X,d,W]_{-b})$ holds in a nonempty hyperbolic space (X,d,W) if it holds in the models[5] of $\mathcal{A}^\omega[X,d,W]_{-b}$ obtained by letting the variables range over the appropriate universes of the full set-theoretic type structure $\mathcal{S}^{\omega,X}$ with the set X as the universe for the base type X, 0_X is interpreted by an arbitrary element of X, $W_X(x,y,\lambda^1)$ is interpreted as $W(x,y,r_{\tilde{\lambda}})$, where $r_{\tilde{\lambda}} \in [0,1]$ is the unique real number represented by $\tilde{\lambda}^1$ and d_X is interpreted as $d_X(x,y) :=_1 (d(x,y))_\circ$.*

[5] We use here the plural since the interpretation of 0_X is not uniquely determined.

Finally, we define the following functional, which is particularly useful for defining majorants for functionals of degree 1.

Definition 9.17 ([77]). *For types $0 \to \rho$ with $\rho = \rho_1 \to \ldots \to \rho_k \to 0$, we define functionals $(\cdot)^M$ of types $(0 \to \rho) \to 0 \to \rho$ by :*

$$x^M(y^0) := \lambda \underline{v}^{\underline{\rho}}. \max{}_0\{x(i, \underline{v}) \mid i = 1, \ldots, y\}.$$

We now state the main version of our metatheorem for *unbounded* metric, hyperbolic and CAT(0)-spaces:

Theorem 9.18. *1. Let ρ be of degree $(1, X)$ or 2 and let $B_\forall(x, u)$, resp. $C_\exists(x, v)$, contain only x, u free, resp. x, v free. Assume that the constant 0_X does not occur in $B_\forall, C_\exists$ and that*

$$\mathcal{A}^\omega[X, d]_{-b} \vdash \forall x^\rho (\forall u^0 B_\forall(x, u) \to \exists v^0 C_\exists(x, v)).$$

Then there exists a partial computable functional[6] $\Phi : S_{\widehat{\rho}} \rightharpoonup \mathbb{N}$ s.t. Φ is defined on all strongly majorizable elements of $S_{\widehat{\rho}}$ and the following holds in all nonempty metric spaces (X, d): for all $x \in S_\rho$, $x^* \in S_{\widehat{\rho}}$ if there exists an $a \in X$ s.t. $x^* \gtrsim^a x$ then[7]

$$\forall u \leq \Phi(x^*) B_\forall(x, u) \to \exists v \leq \Phi(x^*) C_\exists(x, v).$$

In particular, if ρ is in addition of degree $\widehat{1}$, then $\Phi : S_{\widehat{\rho}} \times \mathbb{N} \to \mathbb{N}$ is totally computable.

If 0_x does occur in $B_\forall$ and/or $C_\exists$, then the bound Φ depends (in addition to x^*) on an upper bound $\mathbb{N} \ni n \geq d(0_X, a)$.

2. The theorem also holds for nonempty hyperbolic spaces $\mathcal{A}^\omega[X, d, W]_{-b}$, (X, d, W) and for
$\mathcal{A}^\omega[X, d, W, \mathrm{CAT}(0)]_{-b}$ where (X, d, W) is a CAT(0) space.

Instead of single variables x, u, v and single premises $\forall u B_\forall(x, u)$ we may have tuples of variables and finite conjunctions of premises. In the case of a tuple $\underline{x}$ we then have to require that we have a tuple $\underline{x}^*$ of a-majorants for a common $a \in X$ for all the components of the tuple $\underline{x}$.

Remark. *Another way to treat parameters x^ρ, ρ of degree $(1, X)$ or 2 is to require for a majorant a computable functional t in $S_{\underline{\sigma}} \to S_{\widehat{\rho}}$,[8] where all σ_i are of degree 1. Then we may obtain a totally computable $\Phi : S_{\underline{\sigma}} \to \mathbb{N}$ such that given $\underline{c} \in S_{\underline{\sigma}}$, if there exists an $a \in X$ for which $t(\underline{c}) \gtrsim^a_\rho x$ then the bound $\Phi(\underline{c})$ holds.*

The restriction on the types of degree $(1, X)$ or 2 is made necessary by the interpretation of dependent choice using bar recursive functionals. If a given proof does not use dependent choice, we can allow arbitrary types ρ in the parameters (with majorants of type $\widehat{\rho}$).

[6] More precisely, Φ is given by a bar recursive term (in the sense of [114]) which defines a total functional in $M_{\widehat{\rho} \to 0}$ where $\mathcal{M}^\omega := \langle M_\rho \rangle$ is the type structure of all strongly majorizable functionals [12]. Note that $M_{\widehat{\rho}} \subseteq S_{\widehat{\rho}}$.

[7] Note that $x^* \gtrsim^a x$ implies that x^* s-maj$_{\widehat{\rho}}$ x^* and hence the strong majorizability of x^* so that $\Phi(x^*)$ is defined.

[8] Since t is of degree 2, the computability of t implies its (strong) majorizability.

Remark. *From the proof of Theorem 9.18 (to be given in section 9.9 below) two further extensions follow:*

1. *The language may be extended by a-majorizable constants (in particular constants of types* 0 *and* 1*, which always are uniformly majorizable) where the extracted bounds then additionally depend on (a-majorants for) the new constants.*
2. *The theory may be extended by purely universal axioms or, alternatively, axioms which can be reformulated into purely universal axioms using new majorizable constants if the types of the quantifiers are all of degree* 2 *or* $(1, X)$*,*[9] *as purely universal axioms are their own functional interpretation. Again the extracted bounds depend on (a-majorants for) these new constants. Then the conclusion holds in all metric* (X, d) *resp. hyperbolic* (X, d, W) *spaces which satisfy these axioms (under a suitable interpretation of the new constants if any).*

Remark. *The need for the restriction to* $\exists$*-formulas* $C_\exists$ *in theorem 9.18 is a consequence of the fact that our theories are based on classical logic, where one can produce counterexamples already for formulas* $\exists v^0 \forall w^0 C_{qf}(v, w)$ *with* C_{qf} *quantifier-free. If one bases the system on intuitionistic logic instead this can (even in the presence of many ineffective principles) be avoided and effective bounds for formulas* C *of* **arbitrary** *complexity can be extracted (though no longer bounds on universal premises* $\forall u^0 B_\forall$*). See [42] for this.*

As a corollary to the proof of Theorem 9.18 we obtain Theorem 3.7 in [77]:

Corollary 9.19. *1. Let* σ, ρ *be types of degree* 1 *and* τ *be a type of degree* $(1, X)$*. Let* $s^{\sigma \to \rho}$ *be a closed term of* $\mathcal{A}^\omega[X, d]$ *and let* $B_\forall(x, y, z, u)$*, resp.* $C_\exists(x, y, z, v)$*, contain only* x, y, z, u *free, resp.* x, y, z, v *free. If*

$$\forall x^\sigma \forall y \leq_\rho s(x) \forall z^\tau \left(\forall u^0 B_\forall(x, y, z, u) \to \exists v^0 C_\exists(x, y, z, v)\right)$$

is provable in $\mathcal{A}^\omega[X, d]$*, then one can extract a computable functional* $\Phi : \mathcal{S}_\sigma \times \mathbb{N} \to \mathbb{N}$ *such that for all* $x \in \mathcal{S}_\sigma$ *and all* $b \in \mathbb{N}$

$$\forall y \leq_\rho s(x) \forall z^\tau \left[\forall u \leq \Phi(x, b)\, B_\forall(x, y, z, u) \to \exists v \leq \Phi(x, b)\, C_\exists(x, y, z, v)\right]$$

holds in any nonempty metric space (X, d) *whose metric is bounded by* $b \in \mathbb{N}$*.*

2. *If the premise is proved in '*$\mathcal{A}^\omega[X, d, W]$*', instead of '*$\mathcal{A}^\omega[X, d]$*', then the conclusion holds in all b-bounded hyperbolic spaces.*
3. *If the premise is proved in '*$\mathcal{A}^\omega[X, d, W, \mathrm{CAT}(0)]$*', instead of '*$\mathcal{A}^\omega[X, d, W]$*', then the conclusion holds in all b-bounded CAT(0)-spaces.*

Instead of single variables x, y, z, u, v *we may also have finite tuples of variables*

[9] This ensures that validity in $\mathcal{S}^{\omega,X}$ implies validity in $\mathcal{M}^{\omega,X}$ defined further below.

$\underline{x}, \underline{y}, \underline{z}, \underline{u}, \underline{v}$ as long as the elements of the respective tuples satisfy the same type restrictions as x, y, z, u, v.
Moreover, instead of a single premise of the form '$\forall u^0 B_\forall(x, y, z, u)$' we may have a finite conjunction of such premises.

Proof. Take $a = 0_X$. For x, which has type σ of degree 1, we easily see (even using only strong majorization s-maj) that $x^M \gtrsim^{0_X} x$. Next, for the 0_X-majorant $s^* \gtrsim^{0_X} s$, which we can construct by induction on the structure of s as a closed term of $\mathcal{A}^\omega$ (see Lemma 9.44 in Section 9.9), we have that $s^*(x^M) \gtrsim_1^{0_X} y$ for all $y \leq_1 s(x)$. Given a bound $b \in \mathbb{N}$ on the metric, let $\tau = \tau_1 \to \ldots \to \tau_k \to X$, then also $\lambda \underline{x}^{\widehat{\tau_i}}.b \gtrsim_\tau^{0_X} z$. Likewise, independent of the choice of a we have that the distance $d(0_X, a) \leq b$, but for $a = 0_X$ even $d(0_X, a) = 0$.

Then by Theorem 9.18 we can extract a (bar recursive) functional ϕ such that $\phi((x)^M, s^*(x^M), \lambda \underline{x}^{\widehat{\tau_i}}.b, 0)$ is a bound on $\exists v$, resp. $\forall u$, for any b-bounded metric space. Since both the functional $(\cdot)^M$, the 0_X-majorant s^* for s and the 0_X-majorant $\lambda \underline{x}^{\widehat{\tau_i}}.b$ for z are given by closed terms of $\mathcal{A}^\omega$ (and hence primitive recursive in the sense of [44]), the functional

$$\Phi :\equiv \lambda x, b.\phi(x^M, s^*(x^M), \lambda \underline{x}^{\widehat{\tau_i}}.b, 0)$$

is computable and yields the desired bound.

Note, that in $\mathcal{A}^\omega[X, d]$ we have the boundedness of (X, d) as an axiom, while Theorem 9.18 only allows one to treat the boundedness of (X, d) as an implicative assumption. Since (due to the restrictions on our weak extensionality rule) our systems do not satisfy the deduction theorem[10], strictly speaking this corollary does not follow from Theorem 9.18, but rather from the *proof* of Theorem 9.18: As mentioned in Remark 9.4, we may freely add another purely universal axiom, i.e. the axiom that (X, d) is a b-bounded metric space, to the theory $\mathcal{A}^\omega[X, d]_{-b}$. $\square$

Similarly, one can derive Corollary 3.11 from [77], but we will state a generalized version of Corollary 3.11 from [77] below. For most applications to be discussed in this paper the following more concrete version of the metatheorem is sufficient:

Corollary 9.20. *Let P (resp. K) be a $\mathcal{A}^\omega$-definable Polish space[11] (resp. compact Polish space), let τ be of degree $\widehat{1}$ and let $B_\forall$, resp. $C_\exists$, contain only x, y, z, u free, resp. x, y, z, v free, where furthermore 0_X does not occur in $B_\forall, C_\exists$. If*

$$\mathcal{A}^\omega[X, d, W]_{-b} \vdash \forall x \in P \forall y \in K \forall z^\tau \big(\forall u^0 B_\forall \to \exists v^0 C_\exists\big),$$

then there exists a computable functional $\Phi : \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{(\mathbb{N} \times \ldots \times \mathbb{N})} \to \mathbb{N}$ s.t. the following holds in every nonempty hyperbolic space (X, d, W): for all representatives $r_x \in \mathbb{N}^{\mathbb{N}}$ of $x \in P$ and all $z^* \in \mathbb{N}^{(\mathbb{N} \times \ldots \times \mathbb{N})}$ if there exists an $a \in X$ for which $z^* \gtrsim_\tau^a z$ then

$$\forall y \in K \big(\forall u \leq \Phi(r_x, z^*) B_\forall \to \exists v \leq \Phi(r_x, z^*) C_\exists\big).$$

[10] See [77] for an extensive discussion of this point.
[11] For details on this see [77] and [66].

As before, instead of single variables x, y, z and a single premise $\forall u^0 B_\forall$, we may have tuples of variables and a finite conjunction of premises.
Analogously, for $\mathcal{A}^\omega[X,d]_{-b}$ or $\mathcal{A}^\omega[X,d,W,\mathrm{CAT}(0)]_{-b}$ where then (X,d) resp. (X,d,W) is an arbitrary nonempty metric resp. CAT(0)-space.

Proof. Using the representation of P and K in $\mathcal{A}^\omega$, quantification over $x \in P$ and $y \in K$ can be expressed as quantification over all x^1, resp. all $y^1 \leq s$ for some closed function term s. Then, for (type 1-)representatives r_x of elements x we have $(r_x)^M \gtrsim^a r_x$, while from s and x we obtain an a-majorant $s^*(x^M)$ for all $y \leq s(x)$. Finally, τ has degree $\widehat{1}$, so by Theorem 9.18 we obtain a totally computable bound $\Phi(r_x, z^*)$. □

Definition 9.21. *A function $f : X \to X$ on a metric space (X,d) is called*

- *nonexpansive ('f n.e.') if $d(f(x), f(y)) \leq d(x,y)$ for all $x, y \in X$,*
- *quasi-nonexpansive if $\forall p, x \in X(d(p, f(p)) = 0 \to d(f(x), f(p)) \leq d(x,p))$,*
- *weakly quasi-nonexpansive if $\exists p \in X(d(p, f(p)) = 0 \wedge \forall x \in X(d(f(x), f(p)) \leq d(x,p)))$ or – equivalently –*

$$\exists p \in X \forall x \in X(d(f(x), p) \leq d(x,p)).$$

- *Lipschitz continuous if $d(f(x), f(y)) \leq L \cdot d(x,y)$ for some $L > 0$ and for all $x, y \in X$,*
- *Hölder-Lipschitz continuous if $d(f(x), f(y)) \leq L \cdot d(x,y)^\alpha$ for some $L > 0$, $0 < \alpha \leq 1$ and for all $x, y \in X$.*

For normed linear spaces $(X, \|\cdot\|)$ those definitions are to be understood w.r.t. the induced metric $d(x,y) := \|x - y\|$.

The notion of quasi-nonexpansivity was introduced by Dotson in [28], the notion of weak quasi-nonexpansivity is (implicitly) due to B. Lambov and the second author[78] (note that in context where quasi-nonexpansive mappings are used it always is assumed that fixed points exists so that 'weakly quasi-nonexpansive' is indeed weaker than 'quasi-nonexpansive').[12] Using that $\leq_{\mathbb{R}}$ and $=_{\mathbb{R}}$ are Π^0_1-statements, we observe that the above statements, except for 'f quasi-nonexpansive' and 'f weakly quasi-nonexpansive', can – when formalized in $\mathcal{L}(\mathcal{A}^\omega[X,d,W])$ – be written as $\forall$-formulas, where in the case of Lipschitz and Hölder-Lipschitz the constants L, resp. L and α are assumed to be given as parameters. For 'f weakly quasi-nonexpansive', if we take the fixed point p as a parameter, the remaining formula can be written as a $\forall$-formula, so that to use 'f weakly quasi-nonexpansive' as a premise one needs to quantify over the additional parameter p. The statement 'f quasi-nonexpansive' is of the form

[12] The concept of weakly quasi-nonexpansive mapping has recently been formulated independently – under the name of J-type mapping – in [34] where the fixed point p is called a 'center'.

$\forall \to \forall$ and hence not of a suitable form to serve as a premise, if we want to apply our metatheorems. Most theorems involving quasi-nonexpansive functions easily extend to the 'weakly quasi-nonexpansive' functions which makes our metatheorems applicable. For examples of this see [78].

As examples of weakly quasi-nonexpansive functions (communicated by L. Leustean) consider in the setting of normed linear spaces (with a convex subset C) the class of functions satisfying $\|f(x)\| \leq \|x\|$, which are weakly quasi-nonexpansive in the fixed point 0_X. To see that such functions need not by quasi-nonexpansive consider $f : [0,1] \to [0,1]$ (on the convex subset $[0,1]$ of $\mathbb{R}$) defined by $f(x) := x^2$, which has fixed points $0, 1$, but only is weakly quasi-nonexpansive in 0.

For unbounded hyperbolic spaces (X, d, W) we now state the following corollary:

Corollary 9.22. *1. Let P (resp. K) be a $\mathcal{A}^\omega$-definable Polish space (resp. compact Polish space). Assume we can prove in $\mathcal{A}^\omega[X, d, W]_{-b}$ the following sentence:*

$$\forall x \in P \forall y \in K \forall z^X \forall f^{X \to X} \big(f \text{ n.e.} \wedge \forall u^0 B_\forall(x, y, z, f, u) \to \exists v^0 C_\exists(x, y, z, f, v)\big),$$

where 0_X does not occur in $B_\forall$ and $C_\exists$. Then there exists a computable functional $\Phi : \mathbb{N}^\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ s.t. for all representatives $r_x \in \mathbb{N}^\mathbb{N}$ of $x \in P$ and all $b \in \mathbb{N}$

$$\begin{aligned}\forall y \in K \forall z^X \forall f^{X \to X} \big(&f \text{ n.e.} \wedge d_X(z, f(z)) \leq_\mathbb{R} (b)_\mathbb{R} \\ &\wedge \forall u^0 \leq \Phi(r_x, b) B_\forall(x, y, z, f, u) \to \exists v^0 \leq \Phi(r_x, b) C_\exists(x, y, z, f, v)\big)\end{aligned}$$

holds in all nonempty hyperbolic spaces (X, d, W).
Analogously, for $\mathcal{A}^\omega[X, d, W, \mathrm{CAT}(0)]_{-b}$ where (X, d, W) is a CAT(0) space.

2. *The corollary also holds for an additional parameter $\forall z'^X$ if we add the additional premise $d_X(z, z') \leq_\mathbb{R} (b)_\mathbb{R}$ to the conclusion.*

3. *Furthermore, the corollary holds for an additional parameter $\forall c^{0 \to X}$ if one adds the premise $\forall n (d_X(z, c(n)) \leq_\mathbb{R} (b)_\mathbb{R})$ or just $\forall n (d_X(z, c(n)) \leq_\mathbb{R} (g(n))_\mathbb{R})$ to the conclusion, where the bound then additionally depends on $g : \mathbb{N} \to \mathbb{N}$.*

4. *Statements 1., 2. and 3. also hold if we replace 'f n.e.' with 'f Lipschitz continuous' (with constant $L \in \mathbb{Q}^*_+$), 'f Hölder-Lipschitz continuous' (with constants $L, \alpha \in \mathbb{Q}^*_+$, where $\alpha \leq 1$) or 'f uniformly continuous' (with modulus $\omega : \mathbb{N} \to \mathbb{N}$). For Lipschitz and Lipschitz-Hölder continuous functions the bound additionally depends on the given constants and for uniformly continuous functions the bound additionally depends on the given modulus of uniform continuity.*

5. *Furthermore, 1., 2. and 3. hold if we replace 'f n.e.' with 'f weakly quasi-nonexpansive'. For weakly quasi-nonexpansive functions (with fixed point p) we need to state the additional premise '$d_X(z, p) \leq_\mathbb{R} (b)_\mathbb{R}$' in the conclusion.*

6. *More generally, 1., 2. and 3. hold if in the conclusion f satisfies '$d_X(z, f(z)) \leq_{\mathbb{R}} (b)_{\mathbb{R}}$' and if instead of '$f$ n.e.' we assume*

$$\forall n^0, z_1^X, z_2^X (d_X(z_1, z_2) <_{\mathbb{R}} (n)_{\mathbb{R}} \to d_X(f(z_1), f(z_2)) \leq_{\mathbb{R}} (\Omega_0(n))_{\mathbb{R}}), \quad (*)$$

where Ω_0 is a function $\mathbb{N} \to \mathbb{N}$. The bound then depends on Ω_0 and b.

7. *Finally, 1.,2. and 3. hold if 'f n.e.' is replaced by*

$$\forall n^0, \tilde{z}^X (d_X(z, \tilde{z}) <_{\mathbb{R}} (n)_{\mathbb{R}} \to d_X(z, f(\tilde{z})) \leq_{\mathbb{R}} (\Omega(n))_{\mathbb{R}}), \quad (**)$$

where Ω is a function $\mathbb{N} \to \mathbb{N}$. Then we can drop '$d_X(z, f(z)) \leq_{\mathbb{R}} (b)_{\mathbb{R}}$' in the conclusion and the extracted bound only depends on Ω instead of b.

Proof. In the following we write for simplicity e.g. '$d(z, \tilde{z}) \leq n$' instead of its formal representation '$d_X(z, \tilde{z}) \leq_{\mathbb{R}} (n)_{\mathbb{R}}$' as a formula of $\mathcal{L}(\mathcal{A}^\omega[X, d])$.
For 1., by the comment after Definition 9.21 the premise 'f n.e.' is a $\forall$-formula and hence an admissible premise in Corollary 9.20. The parameters ranging over the Polish spaces P, resp, compact Polish spaces K, are treated as before. Choose $a = z$, then trivially $0 \gtrsim^z z$ and $\lambda n^0.(n+b) \gtrsim^z f$, as using $d(z, f(z)) \leq b$ and the nonexpansivity of f and assuming $d(z, \tilde{z}) \leq n$ we get[13]

$$d(z, f(\tilde{z})) \leq d(z, f(z)) + d(f(z), f(\tilde{z})) \leq b + d(z, \tilde{z}) \leq n + b.$$

For 2. and 3., note that trivially $b \gtrsim^z z, z'$ and $\lambda n^0.b \gtrsim^z c$, resp. $g^M \gtrsim^z c$.

For 4., 5. and 6., we will show that $d(z, f(z)) \leq b$ in conjunction with the requirement that f is Lipschitz continuous, Hölder-Lipschitz continuous, uniformly continuous or f satisfying $(*)$ allows one to derive an Ω such that f satisfies $(**)$, thereby reducing these cases to 7. Similarly, if f is weakly quasi-nonexpansive (with fixed point p) and the additional premise $d(z, p) \leq b$ is satisfied, f satisfies $(**)$. All these conditions on f, including $(*)$, can be written as purely universal formulas (in the case of 'f weakly quasi-nonexpansive' with a parameter p) and may hence serve as a premise according to our metatheorem.

For 7., if f satisfies $\forall \tilde{z} \in X(d(z, \tilde{z}) < n \to d(z, f(\tilde{z})) \leq \Omega(n))$, then trivially $\lambda n.\Omega^M(n+1) \gtrsim^z f$, as $d(z, \tilde{z}) \leq n$ implies $d(z, \tilde{z}) < n+1$ which by $(**)$ implies $d(z, f(\tilde{z})) \leq \Omega(n+1)$. Using the fact that $<_{\mathbb{R}}$ is a Σ_1^0-statement and $\leq_{\mathbb{R}}$ is a Π_1^0-statement we can express $(**)$ as a $\forall$-formula. Also note, that f satisfying $(**)$ implies a bound on $d(z, f(z))$: since $d(z, z) < 1$ by $(**)$ we have $d(z, f(z)) \leq \Omega(1)$.

If f is Lipschitz continuous with constant $L > 0$ then one shows, using $d(z, f(z)) \leq b$ and the triangle inequality and assuming $d(z, \tilde{z}) \leq n$

$$d(z, f(\tilde{z})) \leq d(z, f(z)) + d(f(z), f(\tilde{z})) \leq L \cdot d(z, \tilde{z}) + b \leq L \cdot n + b,$$

so f satisfies $(**)$ with $\Omega(n) := L \cdot n + b$. If f is Hölder-Lipschitz continuous, i.e. $d(f(x), f(y)) \leq L \cdot d(x, y)^\alpha$ for constants $L > 0$ and $0 < \alpha \leq 1$, then f satisfies $(**)$ with $\Omega(n) := L \cdot n^\alpha + b$.

[13]Here and in the following we write for better readability simply d and b instead of d_X and $(b)_{\mathbb{R}}$ etc.

If $f : X \to X$ with (X, d, W) a hyperbolic space is uniformly continuous with modulus[14] ω, then f satisfies $(**)$ with $\Omega(n) := n \cdot 2^{\omega(0)} + b + 1$. Given $z, \tilde{z} \in X$ with $d(z, \tilde{z}) < n$ we can (using W with $z, \tilde{z}$ and suitable λ to construct z_1, W with $z_1, \tilde{z}$ and suitable λ to construct z_2, etc.) inductively construct points $z_1, \ldots z_{k-1}$ (with $k = n \cdot 2^{\omega(0)} + 1$) such that

$$d(z, z_1), d(z_1, z_2), \ldots, d(z_{k-1}, \tilde{z}) < 2^{-\omega(0)}$$

and hence

$$d(f(z), f(z_1)), d(f(z_1), f(z_2)), \ldots, d(f(z_{k-1}), f(\tilde{z})) \leq 1 (= 2^{-0}).$$

Then by the triangle inequality $d(f(z), f(\tilde{z})) \leq k = n \cdot 2^{\omega(0)} + 1$ and another use of the triangle inequality yields $d(z, f(\tilde{z})) \leq d(z, f(z)) + d(f(z), f(\tilde{z})) \leq n \cdot 2^{\omega(0)} + b + 1$.

For weakly quasi-nonexpansive functions f – with fixed point p and with the additional premise: '$d(z, p) \leq b$' – the function f satisfies $(**)$ with $\Omega(n) := n + 2b$, as given $d(z, \tilde{z}) < n$

$$\begin{aligned} d(z, f(\tilde{z})) &\leq d(z, p) + d(f(\tilde{z}), p) \leq d(z, p) + d(\tilde{z}, p) \\ &\leq d(z, p) + d(\tilde{z}, z) + d(z, p) \leq n + 2b. \end{aligned}$$

Alternatively, choosing $a = p$ and writing $(**)$ with p instead of z (and adjusting the other majorants accordingly) f even satisfies $(**)$ with $\Omega(n) := n$, as given $d(p, \tilde{z}) < n$

$$d(p, f(\tilde{z})) \leq d(p, \tilde{z}) \leq n.$$

If f satisfies $d(z, f(z)) \leq b$ and $(*)$, then given $d(z, \tilde{z}) < n$,

$$d(z, f(\tilde{z})) \leq d(z, f(z)) + d(f(z), f(\tilde{z})) \leq \Omega_0(n) + b$$

and hence f satisfies $(**)$ with $\Omega(n) := \Omega_0(n) + b$.

The results then follow using Corollary 9.20. □

Note, that neither the space nor the range of f are in any way assumed to be bounded, but still the bound Φ is highly uniform as it depends only on b (and additional input $L, \alpha, \omega, \Omega_0$ and Ω as stated in cases 3.-7.), but not directly on the points z, z', the sequence c or the function f.

Remark. *Even if 'z' does not occur in $B_\forall, C_\exists$ so that '$\forall z$' is a 'dummy' quantifier, we still need in 1.-4. and 6. in the conclusion a number b with $b \geq d(z, f(z)$ for* **some** *z as this is used in constructing a majorant for f. In 5. we could identify z with p (and $b := 1$ say) and construct a p-majorant of f. In 7. we can construct a uniform f-majorant without reference to b.*

[14] Recall, that $f : X \to X$ uniformly continuous with modulus $\omega : \mathbb{N} \to \mathbb{N}$ is defined as $\forall x, y \in X \forall k \in \mathbb{N}(d(x, y) < 2^{-\omega(k)} \to d(f(x), f(y)) \leq 2^{-k})$.

Remark. *Note that for f nonexpansive, Lipschitz, Hölder-Lipschitz or uniformly continuous, f is provably extensional. For f weakly quasi-nonexpansive or f satisfying conditions $(*)$ or $(**)$ it does not follow that f is extensional. Thus in these cases, if an instance of the extensionality of f is used in a proof, it must either be provable via the extensionality rule (or one must explicitly require f to be (provably) extensional, e.g. by requiring that f is at least uniformly continuous).*

Remark. *Except for the case of f being uniformly continuous all results also hold for general (non-hyperbolic) metric spaces $\mathcal{A}^\omega[X,d]_{-b}$, (X,d). This also applies to corollaries 9.24 and 9.26 below. Note, that in general metric spaces uniformly continuous functions can in general not be majorized, i.e. for $(**)$ no suitable $\Omega(n)$ can be defined, because given $x, y \in X$ we cannot construct intermediate points in order to be able to make use of the uniform continuity of f.*

For a complete characterization of those metric spaces for which uniformly continuous functions f admit the definition of a suitable Ω see [100]. Otherwise, in the setting of metric spaces, we need to require explicitly that a given uniformly continuous function f with modulus ω also satisfies $(**)$ with a suitable Ω.

As a generalization of Corollary 3.11 in [77] we prove the following:

Definition 9.23. *Let $f : X \to X$, then*

- *for $Fix(f) := \{x^X \mid x =_X f(x)\}$ the formula $Fix(f) \neq \emptyset$ expresses f has a fixed point,*
- *for $Fix_\varepsilon(f,y,b) := \{x^X \mid d_X(x,f(x)) \leq_{\mathbb{R}} \varepsilon \wedge d_X(x,y) \leq_{\mathbb{R}} b\}$ and $\varepsilon > 0$ the formula $Fix_\varepsilon(f,y,b) \neq \emptyset$ expresses f has an ε-fixed point in a b-neighborhood of y.*

Corollary 9.24. 1. *Let P (resp. K) be a $\mathcal{A}^\omega$-definable Polish space (resp. compact Polish space) and let $B_\forall$ and $C_\exists$ be as before. If $\mathcal{A}^\omega[X,d,W]_{-b}$ proves that*

$$\forall x \in P\ \forall y \in K\ \forall z^X, f^{X\to X}\big(f\ n.e.\ \wedge Fix(f) \neq \emptyset \wedge \forall u^0 B_\forall \to \exists v^0 C_\exists\big)$$

then there exists a computable functional $\Phi^{1\to 0\to 0}$ (on representatives $r_x : \mathbb{N} \to \mathbb{N}$ of elements x of P) s.t. for all $r_x \in \mathbb{N}^{\mathbb{N}}, b \in \mathbb{N}$

$$\begin{aligned}\forall y \in K\ \forall z^X, f^{X\to X}\big(&f\ n.e.\ \wedge \forall \varepsilon > 0 Fix_\varepsilon(f,z,b) \neq \emptyset \\ &\wedge d_X(z,f(z)) \leq_{\mathbb{R}} (b)_{\mathbb{R}} \wedge \forall u^0 \leq \Phi(r_x,b) B_\forall \to \exists v^0 \leq \Phi(r_x,b) C_\exists\big).\end{aligned}$$

holds in any nonempty hyperbolic space (X,d,W).
Analogously, for $\mathcal{A}^\omega[X,d,W,\mathrm{CAT}(0)]_{-b}$ where then (X,d,W) has to be a CAT(0) space.

2. *The corollary also holds if 'f n.e.' is replaced by f Lipschitz continuous, Hölder-Lipschitz continuous or uniformly continuous, where the extracted bound then additionally will depend on the respective constants and moduli.*

3. *Considering the premise 'f weakly quasi-nonexpansive', i.e.*

$$\exists p^X(f(p) =_X p \wedge \forall w^X(d_X(f(p), f(w)) \leq_{\mathbb{R}} d_X(p, w)))$$

instead of 'f n.e. $\wedge Fix(f) \neq \emptyset$*' we may weaken this premise to*

$$\forall \varepsilon > 0 \exists p^X(d_X(f(p), p) \leq_{\mathbb{R}} \varepsilon \wedge d_X(z, p) \leq_{\mathbb{R}} (b)_{\mathbb{R}} \wedge \forall w^X(d_X(f(p), f(w)) \leq_{\mathbb{R}} d_X(p, w))).$$

4. *Let* $\Psi : (X \to X) \to X \to 1$ *be a provably extensional closed term of* $\mathcal{A}^\omega[X, d, W]_{-b}$*, then in 1. and 2. instead of '*$Fix(f) \neq \emptyset$*' we may weaken '*$\Psi(f, p) =_{\mathbb{R}} 0$*', expressing that* $\Psi(f, \cdot)$ *has a root p, to '*$\forall \varepsilon > 0 \exists p \in X(d(z, p) \leq b \wedge |\Psi(z, p)| \leq_{\mathbb{R}} \varepsilon)$*', expressing that* $\Psi(f, \cdot)$ *has* ε*-roots p which are b-close to z for every* $\varepsilon > 0$.

Proof. Representing P and K in $\mathcal{A}^\omega$, quantification over $x \in P$ and $y \in K$ can be expressed as quantification over all x^1, resp. all $y^1 \leq s$ for some closed function term s of $\mathcal{A}^\omega$. Thus the statement provable by assumption can be written as

$$\forall x^1\ \forall y \leq_1 s\ \forall z^X, p^X, f^{X \to X}\big(f \text{ n.e. } \wedge f(p) =_X p \wedge \forall u^0 B_\forall \to \exists v^0 C_\exists\big),$$

where $f(p) =_X p$ can be written as $\forall k^0(d_X(p, f(p)) \leq_{\mathbb{R}} 2^{-k})$ and both $d_X(p, f(p)) \leq_{\mathbb{R}} 2^{-k}$ and f n.e., resp. the other conditions on f, are $\forall$-formulas.

For x we have $x^M \gtrsim^a x$ and from s we may obtain an a-majorant s^* for s and hence for $y \leq s$. By Corollary 9.22.2, under the additional (purely universal) premises $d_X(z, p), d_X(z, f(z)) \leq_{\mathbb{R}} (b)_{\mathbb{R}}$, we extract a functional Φ s.t. for all $x \in P$ if $r_x \in \mathbb{N}^{\mathbb{N}}$ represents x then

$$\forall y \in K\ \forall z^X, p^X, f^{X \to X}\big(d_X(z, p), d_X(z, f(z)) \leq_{\mathbb{R}} (b)_{\mathbb{R}} \wedge f \text{ n.e. } \wedge \\ d_X(f(p), p) \leq_{\mathbb{R}} 2^{-\Phi(r_x, b)} \wedge \forall u^0 \leq \Phi(r_x, b) B_\forall \to \exists v^0 \leq \Phi(r_x, b) C_\exists\big)$$

holds in all nonempty hyperbolic spaces (X, d, W) (similarly for the other conditions on f, except that then the extracted bound depends on the additional constants and moduli L, α and ω).

The statement $d_X(z, p) \leq_{\mathbb{R}} (b)_{\mathbb{R}} \wedge d_X(f(p), p) \leq_{\mathbb{R}} 2^{-\Phi(r_x, b)}$ expresses that f has a $2^{-\Phi(r_x, b)}$-fixed point in a b-neighborhood of z, which, since $2^{-\Phi(r_x, b)}$ does not depend on p, is implied by $\forall \varepsilon > 0 Fix_\varepsilon(f, z, b) \neq \emptyset$, so 1. and 2. follow from Corollary 9.22. The weakening of the premise 'f weakly quasi-nonexpansive' in 3. is treated similarly.

For 4., similar to the treatment of $f(p) =_X p$ in 1., 2. and 3. we may write $\Psi(f, p) =_{\mathbb{R}} 0$ as $\forall k^0(|\Psi(f, p)|_{\mathbb{R}} \leq_{\mathbb{R}} 2^{-k})$. Then as before may weaken this statement to $\forall \varepsilon > 0 \exists p^X(d_X(z, p) \leq_{\mathbb{R}} (b)_{\mathbb{R}} \wedge |\Psi(z, p)| \leq_{\mathbb{R}} \varepsilon)$. □

Remark. *Note, that in 1. the original premise '*$Fix(f) \neq \emptyset$*' is weakened to '*$\forall \varepsilon > 0 Fix_\varepsilon(f, z, b) \neq \emptyset$*'. By Theorem 1 in [45], nonexpansive mappings on bounded hyperbolic spaces always have* ε*-fixed points for arbitrary* $\varepsilon > 0$*, while they need not have exact fixed points in general (not even in the case of bounded, closed and convex subsets of Banach spaces such as* c_0*). Hence, for bounded* *hyperbolic spaces and nonexpansive mappings the premise '$\forall \varepsilon > 0 Fix_\varepsilon(f,z,b) \neq \emptyset$' can be dropped, if b is taken as an upper bound on the metric d. For further discussion, see Remark 3.13 in [77].*

9.5 Herbrand normal forms

The metatheorems in the previous sections allow one to treat at most classical proofs of formulas that prenex to the form $\forall\exists A_{qf}$. Already for the formula class Π^0_3, i.e. $\forall x^0 \exists y^0 \forall z^0 A_{qf}(x,y,z)$, there are counterexamples where one no longer can extract effective bounds from a given classical proof. These counterexamples basically correspond to the undecidability of the halting problem for Turing machines.

However, the Herbrand normal form $B^H = \forall x^0 \forall h^1_z \exists y^0 A_{qf}(x,y,h_z(y))$ of $B = \forall x^0 \exists y^0 \forall z^0 A_{qf}(x,y,z)$ *does* have the appropriate form (and the Herbrand index function h_z has a suitable restricted type) to allow the extraction of a bound $\Phi(x,h_z)$ on $\exists y A_{qf}(x,y,h_z(y))$. Even though B and B^H are (ineffectively) equivalent, an extracted bound for $\exists y$ in the Herbrand normal form B^H does not yield a bound for $\exists y$ in the original formula B, as it may depend in addition to x on the index function h_z.

The extraction of bounds for Herbrand normal forms can be generalized to a large class of formulas, more precisely to those for which there exists a prenexation such that the Herbrand index functions are of suitable restricted type. The types of the Herbrand index functions depend on the $\exists\forall$ configurations that occur in the prenexation. A configuration $\exists y^\rho \forall z^\tau$ gives rise to $\forall h_z^{\rho\to\tau} \exists y^\rho$, i.e. Herbrand index functions of type $\rho \to \tau$. Restricting ourselves to cases where the Herbrand index functions are guaranteed to have majorants, we only allow configurations $\exists y^\rho \forall z^\tau$ where $\rho = 0$ and τ is of degree $(0,X)$ or 1. Then the types of the Herbrand index functions are of degree $(0,X)$ or 1 as well. This class of formulas covers all arithmetical formulas as well as many other interesting classes involving the extended types $\mathbf{T}^X$.

Clearly, one may extract effective bounds for the Herbrand normal form of formulas if the Herbrand index functions are of suitable restricted type, where naturally, the extracted bounds depend on (a-majorants for) the Herbrand index functions. Of even greater interest is the fact that we may, similarly to the result of Corollary 9.24, weaken or even eliminate the premises of a theorem, even though the conclusion might be of too general form to allow one to extract effective bounds on A rather than A^H, as we will show next.

Definition 9.25. *The class $\mathcal{H}$ of formulas consists of all formulas F that have a prenexation $F' \equiv \exists x_1^{\rho_1} \forall y_1^{\tau_1} \ldots \exists x_n^{\rho_n} \forall y_n^{\tau_n} F_\exists(\underline{x},\underline{y})$ where $F_\exists$ is an $\exists$-formula, the types ρ_i are of degree 0 and the types τ_i are of degree 1 or $(0,X)$.*

We state the following corollary:

Corollary 9.26. *1. Let P (resp. K) be a $\mathcal{A}^\omega$-definable Polish space (resp.*

compact Polish space) and let the formula A be in the class $\mathcal{H}$, where moreover A does not contain 0_X. If $\mathcal{A}^\omega[X,d,W]_{-b}$ proves a sentence

$$\forall x \in P\ \forall y \in K\ \forall z^X, f^{X\to X}\big(f \text{ n.e.} \wedge Fix(f) \neq \emptyset \to A\big)$$

then the following holds in every nonempty hyperbolic space (X,d,W):

$$\forall x \in P\ \forall y \in K\ \forall z^X, f^{X\to X}$$
$$\big(f \text{ n.e.} \wedge \exists b^0 \forall \varepsilon > 0(Fix_\varepsilon(f,z,b) \neq \emptyset) \to A\big).$$

Analogously, for $\mathcal{A}^\omega[X,d,W,\mathrm{CAT}(0)]_{-b}$ where (X,d,W) is a CAT(0) space.

2. *The corollary also holds if we replace 'f n.e.' with f Lipschitz continuous, Hölder-Lipschitz continuous or uniformly continuous.*
3. *Considering the premise 'f weakly quasi-nonexpansive', i.e.*

$$\exists p^X(f(p) =_X p \wedge \forall w^X(d_X(f(p),f(w)) \leq_{\mathbb{R}} d_X(p,w)))$$

instead of 'f n.e. $\wedge Fix(f) \neq \emptyset$' we may weaken this premise to

$$\exists b^0 \forall \varepsilon > 0 \exists p^X(d_X(f(p),p) \leq_{\mathbb{R}} \varepsilon \wedge d_X(z,p) \leq_{\mathbb{R}} (b)_{\mathbb{R}} \wedge \forall w^X(d_X(f(p),f(w)) \leq_{\mathbb{R}} d_X(p,w))).$$

4. *Let $\Psi : (X \to X) \to X \to 1$ be a provably extensional closed term of $\mathcal{A}^\omega[X,d,W]_{-b}$, then in 1. and 2. instead of '$Fix(f) \neq \emptyset$' we may weaken '$\exists p^X \Psi(f,p) =_{\mathbb{R}} 0$', expressing that $\Psi(f,\cdot)$ has a root in p, to '$\exists b^0 \forall \varepsilon > 0 \exists p^X(d_X(z,p) \leq_{\mathbb{R}} (b)_{\mathbb{R}} \wedge |\Psi(z,p)| \leq_{\mathbb{R}} \varepsilon)$', expressing that $\Psi(f,\cdot)$ has ε-roots p which are b-close to z for every $\varepsilon > 0$.*

Proof. Since $A \to A^H$ is logically valid, the statement

$$\forall x \in P\ \forall y \in K\ \forall z^X, f^{X\to X}\big(f \text{ n.e.} \wedge Fix(f) \neq \emptyset \to A\big)$$

trivially implies

$$\forall x \in P\ \forall y \in K\ \forall z^X, f^{X\to X}\big(f \text{ n.e.} \wedge Fix(f) \neq \emptyset \to A^H\big)$$

where A^H is the Herbrand normal form of a suitable prenexation of A as suggested by the formula class $\mathcal{H}$. Pulling outside the universal quantifiers in A^H, which range over the Herbrand index functions, the statement now has a suitable form and the index functions have a suitable type to make possible the extraction of an effective numerical bound (by Corollary 9.24) on the numerical universal quantifiers in the premise and existential quantifiers in the conclusion.

As is to be expected, the extracted bound depends on the parameter x via a representative r_x, on a bound $b \geq d(z,f(z))$ (and $b \geq d(z,0_X)$ if 0_X occurs in A) and on majorants for the Herbrand index functions. Such majorants always exist, as the Herbrand index functions h all are of type degree 1, in which case h^M is an a-majorant, or of type degree $(0,X)$, i.e. basically a sequence of elements in X, in which case we (ineffectively) choose as an a-majorant h^* any sequence of numbers such that $h^*(n) \geq d(h(m),a)$ for all $n \in \mathbb{N}$ and all $m \leq n$, e.g. we may take $h^* := \tilde{h}^M$, where $\tilde{h}(n) := \lceil d(h(n), a) \rceil$. As before, using the representation of P and K in $\mathcal{A}^\omega$, we obtain majorants for (representatives of) x and y.

Thus, by Theorem 9.18 and reasoning as in the proof of Corollary 9.24 we may weaken the universal premise '$Fix(f) \neq \emptyset$' to '$\forall \varepsilon > 0 Fix_\varepsilon(f, z, b) \neq \emptyset$'. Shifting the quantifiers ranging over the Herbrand index functions back in, we obtain:

$$\forall x \in P\ \forall y \in K\ \forall z^X, f^{X \to X} \forall b^0 \\ \big(f \text{ n.e.} \wedge \forall \varepsilon > 0 Fix_\varepsilon(f, z, b) \neq \emptyset \wedge d_X(z, f(z)) \leq_{\mathbb{R}} (b)_{\mathbb{R}} \to A^H\big).$$

But using that (ineffectively) A^H implies back A this yields that

$$\forall x \in P\ \forall y \in K\ \forall z^X, f^{X \to X} \forall b^0 \\ \big(f \text{ n.e.} \wedge \forall \varepsilon > 0 Fix_\varepsilon(f, z, b) \neq \emptyset \wedge d_X(z, f(z)) \leq_{\mathbb{R}} (b)_{\mathbb{R}} \to A\big)$$

holds in all nonempty hyperbolic spaces (X, d, W).

Finally, since here we are not interested in effective bounds but only the (classical) truth of the statement, we may furthermore omit the premise '$d_X(z, f(z)) \leq_{\mathbb{R}} (b)_{\mathbb{R}}$': if for a given $z \in X$ and $f : X \in X$ there exists a b such that '$\forall \varepsilon > 0 Fix_\varepsilon(f, z, b) \neq \emptyset$' holds, then there also exists a b' satisfying both premises, as we may simply take $b' = \lceil \max(b, d_X(z, f(z))) \rceil$.

The cases 2., 3. and 4. are treated similarly. ☐

To see that the restrictions on the types of the Herbrand index functions are necessary consider the following counterexample. In $\mathcal{A}^\omega[X, d, W]_{-b}$ one trivially proves that:

$$\forall f^{X \to X} \big(f \text{ n.e.} \wedge Fix(f) \neq \emptyset \to Fix(f) \neq \emptyset\big),$$

where '$Fix(f) \neq \emptyset$' is expressed by '$\exists z^X \forall k^0 (d_X(z, f(z)) \leq_{\mathbb{R}} 2^{-k})$'.

Without the restrictions on the types of the Herbrand index functions in A^H and hence on A, Corollary 9.26 would allow us to weaken the premise 'f has a fixed point' to 'f has ε-fixed points' and in the case of bounded hyperbolic spaces even eliminate the premise completely since nonexpansive mappings on bounded hyperbolic spaces always have approximate fixed points. Hence we could prove that for bounded hyperbolic case *every* nonexpansive mapping has exact fixed points. As we mentioned already, this is false even for bounded closed convex subsets of Banach spaces, such as e.g. c_0.

This counterexample is ruled out by the restrictions on the types of the Herbrand index functions. Since the statement 'f has a fixed point' is expressed by $\exists z^X \forall k^0 (d_X(z, f(z)) \leq_{\mathbb{R}} 2^{-k})$, the resulting Herbrand index functions have the type $X \to 0$. But already this very simple type is not allowed in the formula class $\mathcal{H}$ and hence not in our corollary.

9.6 Metatheorems for normed linear spaces

We now discuss the setting of (real) normed linear spaces with convex subsets C. As discussed in Machado[96], one may characterize convex subsets of normed spaces in the setting of hyperbolic spaces in terms of additional conditions on the function W. The additional conditions are (I) that the convex combinations do not depend on the order in which they are carried out, and (II) that the distance is homothetic. These additional conditions are:

(I) $\forall x,y,z \in X \forall \lambda_1,\lambda_2,\lambda_3 \in [0,1]\big(\lambda_1+\lambda_2+\lambda_3 =_{\mathbb{R}} 1 \to$
$W(z, W(y,x,\frac{\lambda_1}{1-\lambda_3}), 1-\lambda_3) = W(x, W(z,y,\frac{\lambda_2}{1-\lambda_1}), 1-\lambda_1)\big)$,

(II) $\forall x,y,z \in X \forall \lambda \in [0,1]\big(d(W(z,x,\lambda), W(z,y,\lambda)) = \lambda \cdot d(x,y)\big)$.

The formal version of axiom (I) will look slightly different, as expressing the axiom with $\lambda_1+\lambda_2+\lambda_3 =_{\mathbb{R}} 1$ (with the equality as a premise) is problematic for our purposes. Equality on the reals is a universal statement and hence the axiom itself would no longer be purely universal.

Instead, given λ_1, λ_2 we may explicitly define $\bar{\lambda_1}, \bar{\lambda_2}$ and $\bar{\lambda_3}$ s.t. provably (in $\mathcal{A}^\omega$) both $\bar{\lambda_1}+\bar{\lambda_2}+\bar{\lambda_3} =_{\mathbb{R}} 1$ and if $\lambda_i \in [0,1]$ and $\lambda_1+\lambda_2+\lambda_3 =_{\mathbb{R}} 1$ then $\bar{\lambda_i} = \lambda_i$ for $i = 1,2,3$. The formal versions of the axioms are then as follows:

(I) $\forall x^X, y^X, z^X \forall \lambda_1^1, \lambda_2^1$
$\big(W_X(z, W_X(y,x,\frac{\bar{\lambda_1}}{1-\bar{\lambda_3}}), 1-\bar{\lambda_3}) = W_X(x, W_X(z,y,\frac{\bar{\lambda_2}}{1-\bar{\lambda_1}}), 1-\bar{\lambda_1})\big)$,
where $\bar{\lambda_1} =_1 \tilde{\lambda_1}$, $\bar{\lambda_2} =_1 min_{\mathbb{R}}(\tilde{\lambda_2}, 1 -_{\mathbb{R}} \bar{\lambda_1})$ and $\bar{\lambda_3} =_1 1 -_{\mathbb{R}} (\bar{\lambda_1} +_{\mathbb{R}} \bar{\lambda_2})$,

(II) $\forall x^X, y^X, z^X \forall \lambda^1 \big(d_X(W_X(z,x,\lambda), W_X(z,y,\lambda)) =_{\mathbb{R}} \tilde{\lambda} \cdot_{\mathbb{R}} d_X(x,y)\big)$,

where $\tilde{\lambda}$ is the construction in Definition 9.3. As discussed for the other (X,d,W) axioms in Remark 9.2, the axiom (II) is formulated with W_X to implicitly satisfy $W_X(x,y,\lambda) =_X W_X(x,y,\tilde{\lambda})$.

Thus, theorems concerning convex subsets of normed linear spaces which can be formalized $\mathcal{A}^\omega[X,d,W]_{-b}$ + Machado's two additional axioms can already be treated using the above Theorem 9.18 (as discussed in Remark 9.4). However, as discussed in [77], metatheorems covering normed linear spaces in general rather than just convex subsets of normed linear spaces can be expected to have many more applications, than the applications in fixed point theory investigated so far.

For the new metatheorem for normed linear spaces (with convex subset C) there are, compared to the new metatheorems for (unbounded) metric spaces, two differences: (1) we fix the choice $a = 0_X$ and (2) one cannot meaningfully differentiate between 0_X occurring or not occurring in the theorem to be treated by the metatheorem since it implicitly occurs whenever the norm is used as the latter measures the distance from 0_X (in metric spaces, the only purpose of the constant 0_X was to witness the nonemptyness of the space by a closed term). It is this link between the constant 0_X and the other constants of normed linear spaces, that lets us choose $a = 0_X$ (one could also use an arbitrary a, but then the majorant of the norm would depend on a, i.e. the norm is – in contrast to the metric – not uniformly majorizable).

As in [77], the type C for the convex subset C and quantification over elements of types involving C are defined notions, i.e. an element $x \in X$ is of type C if $\chi_C(x) =_0 0$, where χ_C is a constant of type $X \to 0$ representing the characteristic function of C. Note, however, that our weakly extensional context does not allow us to prove that $x =_X y \wedge \chi_C(x) =_0 0 \to \chi_C(y) =_0 0$ but only if $s =_X t$ is provable that then $\chi_C(s) =_0 \chi_C(t)$ (see [77] for a discussion of this point).

Quantification is treated using the following abbreviations:

$$
\begin{array}{ll}
\forall x^C \ A(x) & :\equiv \forall x^X (\chi_C(x^X) =_0 0 \to A(x)), \\
\forall f^{1\to C} \ A(f) & :\equiv \forall f^{1\to X} \big(\forall y^1 (\chi_C(f(y)) =_0 0) \to A(f)\big), \\
\forall f^{X\to C} \ A(f) & :\equiv \forall f^{X\to X} \big(\forall y^X (\chi_C(f(y)) =_0 0) \to A(f)\big) \\
\forall f^{C\to C} \ A(f) & :\equiv \forall f^{X\to X} \big(\forall x^X (\chi_C(x) =_0 0 \to \chi_C(f(x)) =_0 0) \to A(\tilde{f})\big),
\end{array}
$$

where $\tilde{f}(x) = \begin{cases} f(x), & \text{if } \chi_C(x) =_0 0 \\ c_X, & \text{otherwise.} \end{cases}$

Analogously, for the corresponding $\exists$-quantifiers with '$\wedge$' instead of '$\to$'.

Note, that the additional premises to the conclusion are $\forall$-formulas if we have parameters of these defined types. This extends to types of degree $(1, X, C)$ where ρ is of degree $(1, X, C)$ if it has the form $\tau_1 \to \ldots \to \tau_k \to C$, where τ_i has degree 1 or equals X or C.
Also note, that if we write 'f nonexpansive' for a function $f : C \to C$, this is to be understood as the $\forall$-formula

$$\forall x^X, y^X (\chi_C(x) =_0 0 =_0 \chi_C(y) \to \|f(x) - f(y)\|_X \leq_{\mathbb{R}} \|x - y\|_X).$$

Analogously, for the other notions in Definition 9.21.

Remark. *When we aim to treat parameters $f : C \to C$ in our metatheorems, we need to majorize not that f, but rather the extension $\tilde{f}$ to a function $X \to C$. In [77], where only norm-bounded convex subsets C are considered, the extended function $\tilde{f}$ is easily majorized using the b-boundedness of C (as are parameters of type $(1, X, C)$ in general). In this paper, where we consider unbounded convex subsets C, majorization must employ special properties of the function f, such as e.g. f being nonexpansive. However, the extension $\tilde{f}$ does in general not inherit such properties from f, so instead a majorant for $\tilde{f}$ in general will result from deriving a majorant for f on C from special properties of f, deriving a majorant for $\tilde{f}$ on $X \setminus C$ from the definition of $\tilde{f}$ and taking the maximum over these two majorants.*

Definition 9.27. *We say that a sentence of $\mathcal{L}(\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$ holds in a non-trivial (real) normed linear space with a nonempty convex subset C, if it holds* *in the models*[15] *of* $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$ *obtained by letting the variables range over the appropriate universes of the full set-theoretic type structure* $\mathcal{S}^{\omega,X}$ *with the sets* $\mathbb{N}, X$ *as the universes for the base types* 0 *and* X. *Here* 0_X *is interpreted by the zero vector of the linear space* X, 1_X *by some vector* $a \in X$ *with* $\|a\| = 1$, $+_X$ *is interpreted as addition in* X, $-_X$ *is the inverse of* x *w.r.t.* $+$ *in* X, $\cdot_X$ *is interpreted as* $\lambda\alpha \in \mathbb{N}^{\mathbb{N}}, x \in X.r_\alpha \cdot x$, *where* r_α *is the unique real number represented by* α *and* $\cdot$ *refers to scalar multiplication in the* $\mathbb{R}$*-linear space* X. *Finally,* $\|\cdot\|_X$ *is interpreted by* $\lambda x \in X.(\|x\|)_\circ$. *For the nonempty convex subset* $C \subseteq X$, χ_C *is interpreted as the characteristic function for* C *and* c_X *by some arbitrary element of* C.

The new metatheorem for normed linear spaces is:

Theorem 9.28. *1. Let* ρ *be of degree* $(1, X), (1, X, C)$ *or* 2 *and let* $B_\forall(x, u)$, *resp.* $C_\exists(x, v)$, *contain only* x, u *free, resp.* x, v *free. Assume*

$$\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b} \vdash \forall x^\rho(\forall u^0 B_\forall(x, u) \to \exists v^0 C_\exists(x, v)),$$

Then there exists a partial computable functional $\Phi : S_{\widehat{\rho}} \times \mathbb{N} \rightharpoonup \mathbb{N}$ *s.t.* Φ *is defined on all strongly majorizable elements of* $S_{\widehat{\rho}}$ *and the following holds in all non-trivial (real) normed linear spaces* $(X, \|\cdot\|, C)$ *with a nonempty convex subset* C*: for all* $x \in S_\rho$, $x^* \in S_{\widehat{\rho}}$ *and* $n \in \mathbb{N}$ *if* $x^* \gtrsim^{0_X} x$ *and* $(n)_{\mathbb{R}} \geq_{\mathbb{R}} \|c_X\|_X$ *then*

$$\forall u \leq \Phi(x^*, n) B_\forall(x, u) \to \exists v \leq \Phi(x^*, n) C_\exists(x, v)).$$

In particular, if ρ *is in addition of degree* $\widehat{1}$, *then* $\Phi : S_{\widehat{\rho}} \times \mathbb{N} \to \mathbb{N}$ *is totally computable.*

2. *For uniformly convex spaces with modulus of uniform convexity* η *statement 1. holds with* $(X, \|\cdot\|, C, \eta)$, $\mathcal{A}^\omega[X, \|\cdot\|, C, \eta]_{-b}$ *instead of* $(X, \|\cdot\|, C)$, $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$, *where the extracted bound* Φ *additionally depends on* η.

3. *Analogously, for real inner product spaces* $(X, \langle\cdot,\cdot\rangle)$.

As in the metric case, instead of single variables x, u, v *and single premises* $\forall u B_\forall(x, u)$ *we may have tuples of variables and finite conjunctions of premises.*

Remark. *In the case of metric spaces, if* 0_X *did not occur in the formula for which we want to extract a bound, the bound did not depend on a bound on the distance between the chosen* a *and* 0_X. *This is mainly because the axioms of* $\mathcal{A}^\omega[X, d]_{-b}$ *place no requirements on* 0_X. *This is not the case for normed linear spaces, as in the theory* $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$ *the constant* c_X *is necessary for the interpretation of one of the axioms and hence in general the extracted term may depend on a bound on the norm of* c_X, *even though* c_X *does not occur in the formulas* $B_\forall$ *and* $C_\exists$. *However, if* c_X *does not occur in the formulas* $B_\forall$ *and* $C_\exists$ *and we have another parameter* $z \in C$ *for which we have a bound on the norm, we need not explicitly demand a bound on* $\|c_X\|$, *since in the model* c_X *may be interpreted by an arbitrary element of* C *and we then may interpret* c_X *by* z.

[15] Again we use the plural, as in the setting of normed linear spaces the interpretation of 1_X and c_X are not uniquely determined.

As a corollary we prove Theorem 3.30 in [77].

Corollary 9.29. *1. Let σ be of degree 1 and ρ of degree 1 or $(1, X)$ and let τ be a type of degree $(1, X, C)$. Let s be a closed term of type $\sigma \to \rho$ and $B_\forall, C_\exists$ as before. If*

$$\forall x^\sigma \forall y \leq_\rho s(x) \forall z^\tau \big(\forall u^0 B_\forall(x, y, z, u) \to \exists v^0 C_\exists(x, y, z, v)\big)$$

is provable in $\mathcal{A}^\omega[X, \|\cdot\|, C]$ then one can extract a computable functional $\Phi : S_\sigma \times \mathbb{N} \to \mathbb{N}$ s.t. for all $x \in S_\sigma$

$$\forall y \leq_\rho s(x) \forall z^\tau \big(\forall u^0 \leq \Phi(x, b) B_\forall(x, y, z, u) \to \exists v^0 \leq \Phi(x, b) C_\exists(x, y, z, v)\big)$$

holds in any non-trivial (real) normed linear space $(X, \|\cdot\|)$ and any nonempty b-bounded convex subset $C \subset X$ (with 'b_X' interpreted by 'b').

2. *For uniformly convex spaces $(X, \|\cdot\|, \eta)$ with modulus of uniform convexity η '1.' holds with $\mathcal{A}^\omega[X, \|\cdot\|, C, \eta]$ and $(X, \|\cdot\|, C, \eta)$ instead of $\mathcal{A}^\omega[X, \|\cdot\|, C]$ and $(X, \|\cdot\|, C)$. This time Φ is a computable functional in x, b and a modulus η of uniform convexity for $(X, \|\cdot\|)$ (which interprets the constant 'η').*

3. *Analogously, for real inner-product spaces $(X, \langle\cdot,\cdot\rangle)$.*

Proof. As before in the proof of Corollary 9.19, for x we have $x^M \gtrsim^{0_X} x$ and for $s(x) \geq_\rho y$ we get (using Lemma 9.9) that $s^*(x^M) \gtrsim^{0_X} y$, where s^* is some majorant of s (which exists by Lemma 9.45 as a closed term of $\mathcal{A}^\omega$). Next, given a bound $b \in \mathbb{N}$ on the diameter of C, trivially $(b)_\mathbb{R} \geq \|c_X\|$ and writing $\tau = \tau_1 \to \ldots \to \tau_k \to C$, then also $\lambda \underline{x}^{\widehat{\tau_i}}.b \gtrsim^{0_X}_\tau z$.

Then by Theorem 9.28 we can extract a bar recursive functional ϕ such that $\phi((x)^M, s^*(x^M), \lambda \underline{x}^{\widehat{\tau_i}}.b, b)$ is a bound on $\exists v$, resp. $\forall u$, for any non-trivial real normed linear space and any (*nonempty*)b-bounded convex subset C. Since both the functional $(\cdot)^M$, the a-majorant s^* for s and the a-majorant $\lambda \underline{x}^{\widehat{\tau_i}}.b$ for z are given by closed terms of $\mathcal{A}^\omega$, the functional

$$\Phi :\equiv \lambda x, b.\phi(x^M, s^*(x^M), \lambda \underline{x}^{\widehat{\tau_i}}.b, b)$$

is computable and yields the desired bound.

Note, that in $\mathcal{A}^\omega[X, \|\cdot\|, C]$ we have the boundedness of C as an axiom, while Theorem 9.28 only allows one to treat the boundedness as an implicative assumption. Therefore, as in the proof of Corollary 9.19, this corollary follows from the proof of Theorem 9.28, rather than from the theorem itself. □

We furthermore prove the analogue of Corollary 9.22, though with one important difference: it is no longer sufficient to just have a bound on $\|z - f(z)\|$, $\|z - z'\|$, etc. as in the metric case. Since the choice of a is fixed to $a = 0_X$ in the normed linear case, we also need a bound on the distance between z and 0_X, i.e. $\|z\|$. Moreover, for the functions $f : C \to C$ we consider as parameters, the majorization of f, or rather of the extension $\tilde{f} : X \to C$, requires special care (see Remark 9.6).

Corollary 9.30. *1. Let P (resp. K) be a $\mathcal{A}^\omega$-definable Polish space (resp. compact Polish space). Assume we prove in $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$ a sentence*

$$\forall x \in P \forall y \in K \forall z^C \forall f^{C\to C}(f \text{ n.e.} \wedge \forall u^0 B_\forall(x,y,z,f,u) \to \exists v^0 C_\exists(x,y,z,f,v)),$$

where c_X does not occur in $B_\forall$ and $C_\exists$. Then there exists a computable functional $\Phi : \mathbb{N}^\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ s.t. for all representatives $r_x \in \mathbb{N}^\mathbb{N}$ of $x \in P$ and all $b \in \mathbb{N}$

$$\begin{aligned}&\forall y \in K \forall z^C \forall f^{C\to C}(f \text{ n.e.} \wedge \|z\|_X, \|z - f(z)\|_X \leq_\mathbb{R} (b)_\mathbb{R}\\ &\qquad \wedge \forall u^0 \leq \Phi(r_x, b) B_\forall(x,y,z,f,u) \to \exists v^0 \leq \Phi(r_x, b) C_\exists(x,y,z,f,v))\end{aligned}$$

holds in all non-trivial normed linear spaces $(X, \|\cdot\|)$ and nonempty convex subsets C.
Analogously, for uniformly convex spaces $(X, \|\cdot\|, C, \eta)$ and inner product spaces $(X, \langle\cdot,\cdot\rangle)$, where for uniformly convex spaces the bound Φ additionally depends on the modulus of uniform convexity η.

2. *The corollary also holds for an additional parameter $\forall z'^C$, if we add the additional premise $\|z - z'\|_X \leq_\mathbb{R} (b)_\mathbb{R}$ to the conclusion.*

3. *Furthermore, the corollary holds for an additional parameter $\forall c^{0\to C}$ if we add the additional premise $\forall n(\|z - c(n)\|_X \leq_\mathbb{R} (b)_\mathbb{R})$ or just $\forall n(\|z - c(n)\|_X \leq_\mathbb{R} (g(n))_\mathbb{R})$ to the conclusion, where the bound then additionally depends on $g : \mathbb{N} \to \mathbb{N}$.*

4. *1., 2. and 3. also hold if we replace 'f n.e.' with 'f Lipschitz continuous' (with constant $L \in \mathbb{Q}^*_+$), 'f Hölder-Lipschitz continuous (with constants $L, \alpha \in \mathbb{Q}^*_+$, where $\alpha \leq 1$) or 'f uniformly continuous' (with modulus $\omega : \mathbb{N} \to \mathbb{N}$). For Lipschitz and Lipschitz-Hölder continuous functions the bound depends on the given constants, for uniformly continuous functions the bound depends on the given modulus of uniform continuity.*

5. *Furthermore, 1., 2. and 3. hold if we replace 'f n.e.' with 'f weakly quasi-nonexpansive'. For weakly quasi-nonexpansive functions (with fixed point p) we need to state the additional premise $\|p\|_X \leq_\mathbb{R} (b)_\mathbb{R}$ in the conclusion.*

6. *1.,2. and 3. also hold if we replace 'f n.e.' in the premise and the conclusion by*

$$\forall n^0, z_1^C, z_2^C(\|z_1 - z_2\|_X <_\mathbb{R} (n)_\mathbb{R} \to \|f(z_1) - f(z_2)\|_X \leq_\mathbb{R} (\Omega_0(n))_\mathbb{R}), \quad (*)$$

where Ω_0 is a function $\mathbb{N} \to \mathbb{N}$ and the bound additionally depends on Ω_0.

7. *Finally, 1.,2. and 3. hold if the previous conditions on f are replaced by*

$$\forall n^0, \tilde{z}^C(\|\tilde{z}\|_X <_\mathbb{R} (n)_\mathbb{R} \to \|f(\tilde{z})\|_X \leq_\mathbb{R} (\Omega(n))_\mathbb{R}), \quad (**)$$

where Ω is a function $\mathbb{N} \to \mathbb{N}$ and the bound additionally depends on Ω. In this case we can drop the assumption '$\|z - f(z)\|_X \leq (b)_\mathbb{R}$' in the conclusion whereas '$\|z\|_X \leq (b)_\mathbb{R}$' has to remain.

Note that $(*), (**)$ are logically equivalent to $\forall$-formulas.

Proof. This is basically the same proof as the proof of Corollary 9.22, except for two points: (1) as discussed we need to fix $a = 0_X$ and we need an additional premise, $\|z\| \leq b$ and (2) the 0_X-majorization of f (actually $\tilde{f}$) requires extra care. From the definition of $\tilde{f}$ it is obvious that $n \gtrsim_X^{0_X} f(x)$ for $x \in X \setminus C$ if $n \geq \|c_X\|$. Also note, that since we assume c_X does not occur in $B_\forall$ and $C_\exists$ we may, by Remark 9.6 interpret c_X by the parameter z in the model, so that $\|c_X\| \leq b$. Hence, given an a-majorant $\lambda n.f^*(n) \gtrsim_X^{0_X} f$ on the convex subset C, we obtain the 0_X-majorant $\lambda n.max(f^*(n), b)$ for $\tilde{f}$ and thus the extracted bound does not depend on an explicit bound on the norm of c_X. In the following we may therefore focus on 0_X-majorants for f on the convex subset C.

For 1., 2. and 3. we have that $b \gtrsim^{0_X} z$, $2b \gtrsim^{0_X} z'$ and $\lambda n^0.2b \gtrsim^{0_X} c$, resp. $\lambda n.g^M(n) + b \gtrsim^{0_X} c$. For $\lambda n^0.n + 3b \gtrsim^{0_X} f$, where f is nonexpansive, we reason as follows: assume $\|\tilde{z}\| \leq n$ then

$$\begin{aligned}
\|f(\tilde{z})\| &= \|f(\tilde{z}) - f(z) + f(z) - z + z\| \\
&\leq \|f(\tilde{z}) - f(z)\| + \|f(z) - z\| + \|z\| \\
&\leq \|\tilde{z} - z\| + b + b \\
&\leq \|\tilde{z}\| + \|z\| + 2b \\
&\leq n + 3b.
\end{aligned}$$

Similarly, for 4.,5., 6. and 7. one obtains $\lambda n^0.\Omega^M(n+1) \gtrsim^{0_X} f$ if Ω satisfies $(**)$ from 7. As in the metric case, one may obtain a bound on $\|f(z) - z\|$ using $(**)$: $\|z\| \leq b (\Rightarrow \|z\| < b+1)$ implies $\|f(z)\| \leq \Omega(b+1)$ and hence $\|f(z) - z\| \leq \Omega(b+1) + b$. For 4., 5. and 6. we derive the various Ωs, under the assumptions $\|z\|, \|f(z) - z\| \leq b$, as follows:

If f is Lipschitz continuous with constant $L > 0$ and we assume $\|\tilde{z}\| \leq n$, then using the triangle inequality and the aforementioned assumptions

$$\begin{aligned}
\|f(\tilde{z})\| &\leq \|f(\tilde{z}) - f(z)\| + \|f(z) - z\| + \|z\| \\
&\leq L \cdot \|\tilde{z} - z\| + b + b \\
&\leq L \cdot (n + b) + 2b,
\end{aligned}$$

so f satisfies $(**)$ with $\Omega(n) := L \cdot (n+b) + 2b$. Likewise, one obtains that f Hölder-Lipschitz continuous with constants L, α satisfies $(**)$ with $\Omega(n) := L \cdot (n+b)^\alpha + 2b$.

For f uniformly continuous functions with modulus ω the argument is similar to that in the hyperbolic case: assuming $\|\tilde{z}\| \leq n$ and using $\|z\| \leq b$ we obtain $\|\tilde{z} - z\| \leq n + b$. Dividing the line segment between $\tilde{z}, z$ into $(n+b) \cdot 2^{\omega(0)} + 1$ pieces of length $< 2^{-\omega(0)}$ we obtain $\|f(\tilde{z}) - f(z)\| \leq (n+b) \cdot 2^{\omega(0)} + 1$. Thus we obtain

$$\|f(\tilde{z})\| \leq \|f(\tilde{z}) - f(z)\| + \|f(z) - z\| + \|z\| \leq (n+b) \cdot 2^{\omega(0)} + 2b + 1,$$

i.e. f satisfies $(**)$ with $\Omega(n) := (n+b) \cdot 2^{\omega(0)} + 2b + 1$ for uniformly continuous functions f with modulus ω.

As in the metric case, for weakly quasi-nonexpansive functions, the fixed point p is an additional parameter and we require the additional premise $\|p\| \leq b$.

For weakly quasi-nonexpansive functions f, we then obtain $\Omega(n) := n + 2b$ as follows:

$$\|\tilde{z}\| \le n \to \|f(\tilde{z})\| \le \|f(\tilde{z})-p\|+\|p\| \le \|\tilde{z}-p\|+\|p\| \le \|\tilde{z}\|+\|p\|+\|p\| \le n+2b.$$

At last, if f satisfies $\|z - f(z)\| \le b$ and $(*)$ with an Ω_0, then f satisfies $(**)$ with $\Omega(n) := \Omega_0(n + b) + b$.

The result then follows from Theorem 9.28. □

Remark. *The previous remark 9.4 applies accordingly in the normed case.*

Defining, in the setting of normed linear spaces, the notions of $Fix(f)$ and $Fix_\varepsilon(f, z, b)$ as before, we prove the following corollary.

Corollary 9.31. *1. Let P (resp. K) be a $\mathcal{A}^\omega$-definable Polish space (resp. compact Polish space) and let $B_\forall$ and $C_\exists$ be as before. If $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$ proves a sentence*

$$\forall x \in P \forall y \in K \forall z^C, f^{C\to C}(f \text{ n.e.} \wedge Fix(f) \neq \emptyset \forall u^0 B_\forall \to \exists v^0 C_\exists),$$

then there exists a computable functional $\Phi^{1\to0\to0}$ (on representatives $r_x : \mathbb{N} \to \mathbb{N}$ of elements of P) s.t. for all $r_x \in \mathbb{N}^{\mathbb{N}}, b \in \mathbb{N}$

$$\forall y \in K \forall z^C, f^{C\to C}(f \text{ n.e. } \wedge \forall \varepsilon > 0 Fix_\varepsilon(f, z, b) \neq \emptyset$$
$$\wedge \|z\|_X, \|z - f(z)\|_X \le_{\mathbb{R}} (b)_{\mathbb{R}} \wedge \forall u^0 \le \Phi(r_x, b) B_\forall \to \exists v^0 \le \Phi(r_x, b) C_\exists)$$

holds in all non-trivial normed linear spaces $(X, \|\cdot\|)$ with nonempty convex subset C for which $\|c_X\| \le b$.
If c_X does not occur in $B_\forall, C_\exists$, we can drop the requirement that $\|c_X\| \le b$. Analogously, for uniformly convex spaces $(X, \|\cdot\|, C, \eta)$ and inner product spaces $(X, \langle\cdot,\cdot\rangle)$, where for uniformly convex spaces the bound Φ additionally depends on the modulus of uniform convexity η.

2. *The corollary also holds if 'f n.e.' is replaced by f Lipschitz continuous, Hölder-Lipschitz continuous or uniformly continuous, where the extracted bound then will depend on the respective constants and moduli.*

3. *Considering the premise 'f weakly quasi-nonexpansive', i.e.*

$$\exists p^C(f(p) =_X p \wedge \forall w^X(\|f(p) - f(w)\|_X \le_{\mathbb{R}} \|p - w\|_X))$$

instead of 'f n.e. $\wedge Fix(f) \neq \emptyset$' we may weaken this premise to

$$\forall \varepsilon > 0 \exists p^C(\|f(p)-p\|_X \le_{\mathbb{R}} \varepsilon \wedge \|z-p\|_X \le_{\mathbb{R}} (b)_{\mathbb{R}} \wedge \forall w \in X(\|f(p)-f(w)\|_X \le_{\mathbb{R}} \|p-w\|_X)).$$

4. *Let $\Psi : (X \to X) \to X \to 1$ be a provably extensional closed term of $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$, then in 1. and 2. instead of '$Fix(f) \neq \emptyset$' we may weaken '$\Psi(f,p) =_{\mathbb{R}} 0$' expressing that $\Psi(f,\cdot)$ has a root p to '$\forall \varepsilon > 0 \exists p \in X(\|z-p\|_X \le_{\mathbb{R}} (b)_{\mathbb{R}} \wedge |\Psi(f,p)| \le_{\mathbb{R}} \varepsilon)$', expressing that $\Psi(f,\cdot)$ has ε-roots p which are b-close to z for every $\varepsilon > 0$.*

Proof. This is essentially the same proof as for Corollary 9.24, except that b not only bounds the distance between z and $f(z)$ and the diameter of the subset where ε-fixed points are to be found, but also the norm of z itself and the norm of the element c_X of the convex subset C. The result then follows using Corollary 9.30. □

Similar to Corollary 9.26 we may also in the setting of normed linear spaces allow a Herbrand normal form version of the previous corollary, which allows one to weaken premises even though the conclusion is of a too general form to allow extraction of explicit bounds.

9.7 Simultaneous treatment of several spaces

The generalized approach to majorization developed in the previous section may also be extended to simultaneously cover finite collections of spaces. Instead of a single space X and a single element $a \in X$ we may have a collection of spaces $X_1, \ldots, X_n$ and corresponding elements $a_i \in X_i$ that we take as reference points for the majorization relation. We then may consider elements of products of these spaces and functions between such product-elements.

$\mathbf{T}^{X_1,\ldots,X_n}$ is the set of all finite types ρ over the ground types $0, X_1, \ldots, X_n$. For $\rho \in \mathbf{T}^{X_1,\ldots,X_n}$ the type $\widehat{\rho}$ defines the type which results from ρ by replacing all occurrences of $X_i, 1 \leq i \leq n$ by 0. The relation $\gtrsim^{\underline{a}}$ is then defined as follows:

Definition 9.32. *We define a ternary relation $\gtrsim^{\underline{a}}_{\rho}$ between objects x, y and an n-tuple $\underline{a}$ of type $\widehat{\rho}, \rho$ and $X_1, \ldots, X_n$ respectively as follows:*

- $x^0 \gtrsim^{\underline{a}}_0 y^0 :\equiv x \geq_0 y$,
- $x^0 \gtrsim^{\underline{a}}_{X_i} y^{X_i} :\equiv (x)_{\mathbb{R}} \geq_{\mathbb{R}} d_{X_i}(y, a_i)$,
- $x \gtrsim^{\underline{a}}_{\rho\to\tau} y :\equiv \forall z', z(z' \gtrsim^{\underline{a}}_{\rho} z \to xz' \gtrsim^{\underline{a}}_{\tau} yz) \wedge \forall z', z(z' \gtrsim^{\underline{a}}_{\widehat{\rho}} z \to xz' \gtrsim^{\underline{a}}_{\widehat{\tau}} xz)$.

If X_i is a normed linear spaces we require $a_i = 0_{X_i}$ s.t. $d_{X_i}(x, a_i) =_{\mathbb{R}} \|x\|_{X_i}$.

E.g. if we have two metric spaces (X_1, d_{X_1}) and (X_2, d_{X_2}) then an (a_1, a_2)-majorant for $f^{X_1\to X_2}$ is a function f^* of type 1 such that

$$\forall n^0, x^{X_1}\big(d_{X_1}(x, a_1) \leq_{\mathbb{R}} (n)_{\mathbb{R}} \to d_{X_2}(f(x), a_2) \leq_{\mathbb{R}} (f^*(n))_{\mathbb{R}}\big).$$

If f is nonexpansive and $a_2 := f(a_1)$, then f is (a_1, a_2)-majorized by the identity function $\lambda n.n^0$.

Functions involving product types are treated using "currying" in the form of the following two patterns:

- a function $f : X_1 \times \ldots \times X_n \to \rho$ is represented by $f : X_1 \to \ldots \to X_n \to \rho$,

- a function $\rho \to X_1 \times \ldots \times X_n$ is represented by an n-tuple of functions $f_i : \rho \to X_i$.

Thus e.g. a function $f : X_1 \times X_2 \to X_1 \times X_2$ will be represented by a pair $f_{1,2} : X_1 \to (X_2 \to X_{1,2})$. A function $g : (X_1 \times X_2 \to X_1 \times X_2) \to X_1 \times X_2$ by a pair $g_{1,2} : (X_1 \to (X_2 \to X_1)) \to ((X_1 \to (X_2 \to X_2)) \to X_{1,2})$ and similar for products of greater arity and functions of more complex types.

9.8 Applications

Application 9.33. *Let (X, d, W) be an arbitrary (nonempty) hyperbolic space, $k \in \mathbb{N},\ k \geq 1$ and $(\lambda_n)_{n\in\mathbb{N}}$ a sequence in $[0, 1 - \frac{1}{k}]$ with $\sum_{n=0}^{\infty} \lambda_n = \infty$ and define for $f : X \to X, x \in X$ the Krasnoselski-Mann iteration $(x_n)_n$ starting from x ([84, 97]) by*

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

In [45](Theorem 1) the following is proved[16]

$$\forall x \in X, f : X \to X\big((x_n)_n \textit{ bounded and } f \textit{ n.e.} \to \lim_{n\to\infty} d(x_n, f(x_n)) = 0\big).$$

As observed in [14], it actually suffices to assume that $(x^*_n)_n$ starting from **some** *x^* is bounded. Therefore*

$$\forall x \in X, f : X \to X\big(\exists x^* \in X((x^*_n)_n \textit{ bounded) and } f \textit{ n.e.} \to \lim_{n\to\infty} d(x_n, f(x_n)) = 0\big).$$

The proof given in [45] (and [14]) can easily be formalized in $\mathcal{A}^\omega[X, d, W]_{-b}$ (see [77] for more details on this). As an application of Corollary 9.22 we obtain (see the proof below) the following effective and uniform version:
There exists a computable bound $\Phi(k, \alpha, b, l)$ such that in any (nonempty) hyperbolic space (X, d, W), for any $l, b, k \in \mathbb{N}$ and any $\alpha : \mathbb{N} \to \mathbb{N}$ the following holds: if (λ_n) is a sequence in $[0, 1)$ such that

$$\forall n \in \mathbb{N}(\lambda_n \leq 1 - \frac{1}{k} \wedge n \leq \sum_{i=0}^{\alpha(n)} \lambda_i)$$

then

$$\forall x, x^* \in X \forall f : X \to X$$
$$(\ \forall i, j(d(x, x^*), d(x^*_i, x^*_j) \leq b) \wedge f \textit{ n.e.} \ \to \forall m \geq \Phi(k, \alpha, b, l) d(x_m, f(x_m)) < 2^{-l}).$$

Proof. As mentioned already, $\mathcal{A}^\omega[X, d, W]_{-b}$ proves the following (formalized version of Theorem 1 in [45]): if $k \geq 1, \lambda^{0\to 1}_{(\cdot)}$ represents an element of the

[16] For the case of convex subsets $C \subseteq X$ of normed linear spaces $(X, \|\cdot\|)$ this result is already due to [53]. [45] even treats spaces of hyperbolic type.

compact Polish space $[0,1]^\infty$ (with the product metric) and $\alpha : \mathbb{N} \to \mathbb{N}$ such that

$$(*)\ \forall n \in \mathbb{N}\big(\lambda_n \leq_{\mathbb{R}} 1 - \frac{1}{k} \wedge n \leq_{\mathbb{R}} \sum_{i=0}^{\alpha(n)} \lambda_i\big),$$

where $\sum_{i=0}^{\alpha(n)} \lambda_i$ represents the corresponding summation of the real numbers in $[0,1]$ represented by λ_i, then

$$\forall l, b \in \mathbb{N}, x, x^*, f^{X\to X}\big(\,(x_n^*)_n\ b\text{-bounded} \wedge f \text{ n.e.} \to \exists n \in \mathbb{N}(d_X(x_n, f(x_n)) <_{\mathbb{R}} 2^{-l})\big),$$

where '$(*)$' and '$(x_n^*)_n$ b-bounded' are a $\forall$-formulas and '$d_X(x_n, f(x_n)) <_{\mathbb{R}} 2^{-l}$' is a $\exists$-formula.

Now Corollary 9.22 yields the existence of a computable functional $\Phi(k, \alpha, b, l)$ such that for $n := \Phi(k, \alpha, b, l)$

$$\begin{array}{l}\forall(\lambda_m) \in [0,1]^\infty \forall x, x^* \in X \forall f : X \to X \\ \quad ((*) \wedge \forall i, j(d(x, f(x)), d(x, x^*), d(x_i^*, x_j^*) \leq b) \wedge f \text{ n.e.} \to \exists m \leq n(d(x_m, f(x_m)) < 2^{-l}))\end{array}$$

holds for all k, α, b, l in any (nonempty) hyperbolic space (X, d, W).

Since $(d(x_n, f(x_n)))_n$ is a non-increasing sequence ([45]) the conclusion actually implies

$$\forall m \geq \Phi(k, \alpha, b, l)\big(d(x_m, f(x_m)) < 2^{-l}\big).$$

The only thing which remains to show is that the assumption '$d(x, f(x)) \leq b$' is redundant: by Theorem 1 from [45] we know, in particular, that $d(x_n^*, f(x_n^*)) \to 0$ and so a-fortiori

$$\exists n \in \mathbb{N}\big(d(x_n^*, f(x_n^*)) \leq b\big).$$

Using $d(x, x^*), d(x_i^*, x_j^*) \leq b$ for all i, j and the nonexpansivity of f yields

$$d(x, f(x)) \leq d(x, x^*) + d(x^*, x_n^*) + d(x_n^*, f(x_n^*)) + d(f(x_n^*), f(x^*)) + d(f(x^*), f(x)) \leq 5b.$$

So replacing 'b' in the bound by '$5b$' we can drop the assumption '$d(x, f(x)) \leq b$'. □

As a corollary it follows, that for bounded hyperbolic spaces (X, d, W) the convergence $d(x_n, f(x_n)) \to 0$ is uniform in x, f and – except for a bound b on the metric – in (X, d, W). This corollary was first proved as Theorem 2 in [45][17] and was shown to follow from a general logical metatheorem in [77] where a detailed discussion of this point is given. In [79], the extraction of an actual effective uniform rate of convergence was carried out and it was noticed that the assumption on X to be bounded could be weakened to a bound b on $d(x, x^*)$ and $(x_n^*)_n$ for some $x^* \in X$. At that time, there was no explanation in terms of a general result from logic for the fact that these local bounds where sufficient. This latter fact can now for the first time be explained by our refined logical

[17] For the case of bounded convex subsets of normed spaces and constant $\lambda_n = \lambda \in (0,1)$ the uniformity in x was already shown in [31] and – for $(\lambda_n)_n$ in $[a, b] \subset (0,1)$ and non-increasing – in [23].

metatheorems as well. Note that the proof of Theorem 2 in [45] (as well as the alternative proof for constant $\lambda_n = \lambda \in (0,1)$ given in [58]) crucially uses that the whole space X is assumed to be bounded. So the uniformity result guaranteed a-priorily by the metatheorems of the present paper applied to Theorem 1 of [45] not only yields immediately Theorem 2 from [45] (called 'main result') but even a qualitatively stronger uniformity which apparently cannot be obtained by the functional analytic embedding techniques used in [45] (or in [58]).

The aforementioned explicit bound extracted in [79](see theorem 3.21 and remark 3.13) is as follows (for the case of convex subsets of normed linear spaces the result is due already to [74, 75]):

$$\begin{array}{l}\Phi(k,\alpha,b,l) := \widehat{\alpha}(\lceil 12b\cdot\exp(k(M+1))\rceil - 1, M)),\ \text{where}\\ M := (1+6b)\cdot 2^l,\ \widehat{\alpha}(0,n) := \tilde{\alpha}(0,n), \widehat{\alpha}(i+1,n) := \tilde{\alpha}(\widehat{\alpha}(i,n),n),\ \text{with}\\ \tilde{\alpha}(i,n) := i+\alpha^+(i,n),\ \text{where}\ \alpha^+(i,n) := \max\limits_{j\le i}[\alpha(n+j)-j+1].\end{array}$$

Before we come to the next application we need the following

Proposition 9.34 ([53, 45])**.** *Let (X,d,W) be a (nonempty) compact hyperbolic space and $(\lambda_n), f, (x_n)$ as in application 9.33. Then $(x_n)_n$ converges towards a fixed point of f (see [76] for details).*

Proof. By the result mentioned in application 9.33 we have that $d(x_n, f(x_n)) \to 0$ since the compactness of X implies that X – and hence $(x_n)_n$ – is bounded. Using again the compactness of X, we know that $(x_n)_n$ has a convergent subsequence $(x_{n_k})_k$ with limit $\widehat{x}$. One easily shows (using the continuity of f) that $\widehat{x}$ is a fixed point of f. The proof is concluded by verifying the easy fact that for any fixed point $\widehat{x}$ of f

$$\forall n \in \mathbb{N}(d(x_{n+1},\widehat{x}) \le d(x_n,\widehat{x}))$$

which implies that already $(x_n)_n$ converges towards $\widehat{x}$. □

In particular it follows that $(x_n)_n$ is a Cauchy sequence and for this corollary one does not need the completeness of X but only its total boundedness: suppose X is totally bounded. Then its metric completion $\widehat{X}$ (which again is a hyperbolic space) is totally bounded too and hence compact. f extends to a nonexpansive function $\widehat{f}$ on the completion so that the previous result applies. Since $\widehat{f}$ coincides with f on X, also the Krasnoselski-Mann iteration of $\widehat{f}$ coincides with that of f when starting from a point $x \in X$. Hence we conclude that $(x_n)_n$ is a Cauchy sequence.

Application 9.35. *Let us consider the proof of the Cauchy property of (x_n) from the asymptotic regularity (i.e. $d(x_n, f(x_n)) \to 0$) (taken as assumption)*[18]

[18] The proof relative to this assumption only uses that $(\lambda_n)_n$ is a sequence in $[0,1]$ but not the other assumptions on $(\lambda_n)_n$ (which are only needed to prove that $d(x_n, f(x_n)) \to 0$).

under the additional assumption of X being totally bounded, i.e. the proof of

$$(+)\ X \text{ totally bounded } \wedge \lim d(x_n, f(x_n)) = 0 \wedge f \text{ n.e. } \rightarrow \\ \forall k \in \mathbb{N} \exists n \in \mathbb{N} \forall m \geq n (d(x_n, x_m) \leq 2^{-k})$$

which can be formalized in $\mathcal{A}^\omega[X, d, W]_{-b}$.

In order to apply Corollary 9.22 we first have to modify (+) so that the logical form required in the corollary is obtained. In order to do so we first have to make the assumptions explicit:

- *due to the fact that $d(x_n, f(x_n))_n$ is non-increasing, we can write the asymptotic regularity equivalently as $\forall l \in \mathbb{N} \exists n \in \mathbb{N}(d(x_n, f(x_n)) \leq 2^{-l})$ which asks for a witnessing rate of asymptotic regularity $\delta : \mathbb{N} \rightarrow \mathbb{N}$ such that*
$$(1)\ \forall l \in \mathbb{N}(d(x_{\delta(l)}, f(x_{\delta(l)})) \leq 2^{-l}).$$
- *the total boundedness of X is expressed by the existence of a sequence $(a_n)_n$ of points in X and a function $\gamma : \mathbb{N} \rightarrow \mathbb{N}$ such that*
$$(2)\ \forall l \in \mathbb{N}, x \in X \exists n \leq \gamma(l)(d(x, a_n) \leq 2^{-l}).$$
A function γ such that a sequence $(a_n)_n$ in X satisfying (2) exists is called a modulus of total boundedness for X.

It is important to notice that both (1) and (2) are (provably equivalent to) $\forall$-formulas.

The conclusion, i.e. the Cauchy property of (x_n), is a Π^0_3-formula and so too complicated to be covered by our metatheorems. In fact, as shown in [76] there is no Cauchy rate computable in the parameters even for a very simple computable sequence of nonexpansive functions on $X = [0, 1]$ and $\lambda_n = \frac{1}{2}$. We therefore modify the conclusion to its Herbrand normal form[19]

$$(H)\ \forall l \in \mathbb{N} \forall g : \mathbb{N} \rightarrow \mathbb{N} \exists n \in \mathbb{N} \forall i, j \in [n; n + g(n)](d(x_i, x_j) < 2^{-l}),$$

where $[n; m]$ denotes the subset $\{n, n+1, \ldots, m-1, m\}$ of $\mathbb{N}$ for $m \geq n$.

Classically, (H) is a equivalent to the Cauchy property for $(x_n)_n$ but – since the proof is ineffective – a computable bound on (H) does not yield a computable Cauchy modulus for $(x_n)_n$. Note that

$$\exists n \in \mathbb{N} \forall i, j \in [n; n + g(n)](d(x_i, x_j) < 2^{-l})$$

is (equivalent to) an $\exists$-formula.

$\mathcal{A}^\omega[X, d, W]_{-b}$ proves that

$$\forall (\lambda_m) \in [0,1]^\infty \forall x^X \forall f^{X \rightarrow X}, (a_n)^{0 \rightarrow X}, l^0, \gamma^1, \delta^1, g^1 \\ ((1) \wedge (2) \wedge\ f \text{ n.e. } \rightarrow \exists m \in \mathbb{N} \forall i, j \in [m; m + g(m)](d(x_i, x_j) < 2^{-l})).$$

[19]More precisely, the Herbrand normal form of a slightly different but trivially equivalent formulation of the Cauchy property.

The total boundedness of X implies that the metric of X is bounded and a bound can be computed by $b := \max\{d(a_i, a_j) : i, j \leq \gamma(0)\} + 2$. However, in order to guarantee our result to be independent from $(a_n)_n$ we add a bound b of X as an additional input. Hence by Corollary 9.19 we obtain a computable bound $n := \Omega(l, b, \gamma, \delta, g)$ such that for all (λ_n) in $[0,1], x \in X, (a_n)$ in $X, f : X \to X, l \in \mathbb{N}$ and $\gamma, \delta, g : \mathbb{N} \to \mathbb{N}$:

$$(1) \wedge (2) \wedge \ f \ n.e. \ \to \exists m \leq n \forall i, j \in [m; m + g(m)](d(x_i, x_j) < 2^{-l})$$

holds in any (nonempty) b-bounded, totally bounded (with modulus γ) hyperbolic space (X, d, W).

A concrete bound Ω of this kind has in fact been extracted first in [76], where there extraction itself was guided by the algorithm provided by the proof of Corollary 9.19 as well as the proof-theoretic study of the Bolzano-Weierstraß principle carried out in [69]. This concrete Ω even is independent from b and is defined as follows

$$\Omega(l, g, \delta, \gamma) := \max_{i \leq \gamma(l+3)} \Psi_0(i, l, g, \delta),$$

where

$$\begin{cases} \Psi_0(0, l, g, \delta) := 0 \\ \Psi_0(n+1, l, g, \delta) := \delta\left(l + 2 + \lceil \log_2(\max_{i \leq n} g(\Psi_0(i, l, g, \delta)) + 1) \rceil\right). \end{cases}$$

For X being b-bounded and (λ_n) s.t. $\lambda_i \in [0, 1 - \frac{1}{k}], \sum_{i=0}^{\alpha(n)} \lambda_i \geq n$, we can take $\delta(l) := \Phi(k, \alpha, b, l)$ from application 9.33.

Application 9.36. *Let $(X, d, W), k, (\lambda_n), f, x$ and (x_n) be as in application 9.33. In [14], the following result is proved:*

$$\forall x \in X, f : X \to X\big(f \ n.e. \ \to \lim_{n \to \infty} d(x_n, f(x_n)) = r(f)\big),$$

where $r(f) := \inf_{y \in X} d(y, f(y))$ is the so-called minimal displacement of f. As (x_n) is no longer assumed to be bounded, $r(f)$ can very well be strictly positive: e.g. for $\mathbb{R}$ (with the natural metric) and $f : \mathbb{R} \to \mathbb{R}, f(x) := x + 1$ we have $r(f) = 1$ although f is nonexpansive.

The above theorem can be written equivalently as follows (using again that $(d(x_n, f(x_n)))$ is non-increasing):[20]

$$\forall l \in \mathbb{N} \forall x, x^* \in X, f : X \to X \exists n \in \mathbb{N}(d(x_n, f(x_n)) < d(x^*, f(x^*)) + 2^{-l}).$$

The proof given in [14] can be formalized in $\mathcal{A}^\omega[X, d, W]_{-b}$ and so Corollary 9.22 yields (like in the proof of application 9.33 above) an effective bound $n := \Psi(k, \alpha, b, l)$ such that in any (nonempty) hyperbolic space (X, d, W), for any

[20] For details see [75].

$l, b, k \in \mathbb{N}$ and any $\alpha : \mathbb{N} \to \mathbb{N}$ such that $\lambda_n \leq \frac{1}{k}$ and $n \leq \sum_{i=0}^{\alpha(n)} \lambda_i$ the following holds

$$\begin{array}{l} \forall x, x^* \in X \forall f : X \to X \\ \quad (d(x, x^*), d(x, f(x)) \leq b \wedge f \text{ n.e.} \to \exists k \leq n(d(x_k, f(x_k)) < d(x^*, f(x^*)) + 2^{-l})) \end{array}$$

and so (by the fact that $d(x_n, f(x_n))_n$ is non-increasing)

$$\begin{array}{l} \forall x, x^* \in X \forall f : X \to X \\ \quad (d(x, x^*), d(x, f(x)) \leq b \wedge f \text{ n.e.} \to \forall m \geq \Psi(k, \alpha, b, l)(d(x_m, f(x_m)) < d(x^*, f(x^*)) + 2^{-l})). \end{array}$$

An explicit such bound Ψ (which is very similar to the bound Φ mentioned in connection with application 9.33) has been extracted first in [79] (for the special case of convex subsets of normed spaces this is already due to [74] and – in a stronger form – in [75]). Our refined metatheorems for the first time allow one to explain this finding as an instance of a general result in logic.

9.9 Proofs of Theorems 9.18 and 9.28

We focus on proving Theorems 9.18 and 9.28 for the theories $\mathcal{A}^\omega[X, d, W]_{-b}$ and $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$ respectively. From these proofs, the corresponding proofs for the other variants of the theories for metric and normed linear spaces can easily be obtained by treating one of the following kinds of extensions: (1) adding another purely universal axiom to the theory (purely universal axioms are their own functional interpretation), e.g. for $\mathcal{A}^\omega[X, d, W, CAT(0)]_{-b}$, or (2) adding a new majorizable constant to the language as e.g. for $\mathcal{A}^\omega[X, \|\cdot\|, C, \eta]_{-b}$, where the modulus of uniform convexity η is given as a number theoretic function $\eta : \mathbb{N} \to \mathbb{N}$.

The proofs of Theorems 9.18 and 9.28 closely follow the general proof outline in [77], but both the interpretation of the theories $\mathcal{A}^\omega[X, d, W]_{-b}$, $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$ by bar-recursive functionals and the subsequent interpretation of these functionals in an extension of the Howard-Bezem strongly majorizable functionals to all types $\mathbf{T}^X$ are now parametrized by an element $a \in X$ for the relation $\gtrsim^a$. For the interpretation of $\mathcal{A}^\omega[X, d, W]_{-b}$ by bar-recursive functionals this in effect leads to a family of functional interpretations parametrized by $a \in X$, where the interpretation of the element $a \in X$ is fixed later during the majorization process (see Remark 9.9). For $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$ the choice $a = 0_X$ is fixed from the beginning.

Based on $\gtrsim^a$ we redefine Kohlenbach's [77] extension of Bezem's[12] type structure of hereditarily strongly majorizable set-theoretical functionals to all types $\mathbf{T}^X$ (based on $\gtrsim^a$ instead of s-maj) as follows:

Definition 9.37. *Let (X, d) be a nonempty metric space, resp. $(X, \|\cdot\|)$ a non-trivial real normed linear space, and let $a \in X$ be given. The extensional type structure $\mathcal{M}^{\omega,X}$ of all hereditarily strongly a-majorizable set-theoretic function-*

als of type $\rho \in \mathbf{T}^X$ over $\mathbb{N}$ and X is defined as

$$\begin{cases} M_0 := \mathbb{N}, \ \ n \gtrsim_0^a m :\equiv n \geq m \wedge n, m \in \mathbb{N}, \\ M_X := X, \ n \gtrsim_X^a x :\equiv n \geq d(x,a) \wedge n \in M_0, x \in M_X, \\ \qquad\qquad\qquad\qquad \textit{where } \geq \textit{ is the usual order on } \mathbb{R}. \\ x^* \gtrsim_{\rho\to\tau}^a x :\equiv x^* \in M_{\widehat{\tau}}^{M_{\widehat{\rho}}} \wedge x \in M_\tau^{M_\rho} \\ \qquad\qquad \wedge \forall y^* \in M_{\widehat{\rho}}, y \in M_\rho \, (y^* \gtrsim_\rho^a y \to x^*y^* \gtrsim_\tau^a xy) \\ \qquad\qquad \wedge \forall y^*, y \in M_{\widehat{\rho}} \, (y^* \gtrsim_{\widehat{\rho}}^a y \to x^*y^* \gtrsim_{\widehat{\tau}}^a x^*y), \\ M_{\rho\to\tau} := \left\{ x \in M_\tau^{M_\rho} \,\middle|\, \exists x^* \in M_{\widehat{\tau}}^{M_{\widehat{\rho}}} : x^* \gtrsim_{\rho\to\tau}^a x \right\} \quad (\rho, \tau \in \mathbf{T}^X) \ . \end{cases}$$

Remark. *Restricted to the types $\mathbf{T}$, this type structure is identical to Bezem's original type structure $\mathcal{M}^\omega$ of strongly hereditarily majorizable functionals, as for $\rho \in \mathbf{T}$ the relations* s-maj$_\rho$ *and $\gtrsim_\rho^a$ are the same and hence for $\rho \in \mathbf{T}$ we may freely write* s-maj$_\rho$ *instead of $\gtrsim_\rho^a$, as here the parameter $a \in X$ is irrelevant.*

Even though the a-majorization relation is parametrized by an element $a \in X$, the resulting model of all hereditarily strongly a-majorizable functionals is independent of the choice of $a \in X$, as the following lemma shows:

Lemma 9.38. *Let $a, b \in X$ be given. Then for every $\rho \in \mathbf{T}^X$ there is a mapping Φ_ρ of type $\widehat{\rho} \to 0 \to \widehat{\rho}$ s.t. for all $x^* \in M_{\widehat{\rho}}^a$, $x \in M_\rho^a$ and all $n \in \mathbb{N}$ s.t. $d(a,b) \leq n$,*

$$x^* \gtrsim_\rho^a x \to \Phi_\rho(x^*, n) \gtrsim_\rho^b x,$$

and for all $x^*, \widehat{x} \in M_{\widehat{\rho}}^a$

$$x^* \textit{ s-maj}_{\widehat{\rho}} \, \widehat{x} \to \Phi_\rho(x^*, n) \textit{ s-maj}_{\widehat{\rho}} \, \Phi_\rho(\widehat{x}, n).$$

In particular: $M_\rho^a = M_\rho^b$ and – trivially – $M_{\widehat{\rho}}^a = M_{\widehat{\rho}}^b$. Note, that this property is symmetric in $a, b \in X$.

Proof. The proof is by induction on the type $\rho \in \mathbf{T}^X$.

For $\rho = 0$ define $\Phi_0(x, n) := x$. Trivially $M_0^a = M_0^b$ by definition.

For $\rho = X$ the mapping Φ_X is the mapping $\Phi_X(x^*, n) = x^* + n$, as $x^* \gtrsim_X^a x$ is equivalent to $x^* \geq d(x,a)$ but then by the triangle inequality $x^* + n \geq d(x,b)$ and hence $x^* + n \gtrsim_X^b x$. Obviously, $M_X^a = M_X^b$.

For $\rho = \sigma \to \tau$, we need to construct the mapping $\Phi_{\sigma\to\tau}$ and show that $x \in M_{\sigma\to\tau}^a$ implies $x \in M_{\sigma\to\tau}^b$. Assume $x^* \gtrsim_{\sigma\to\tau}^a x$ for $x \in M_{\sigma\to\tau}^a$, and let $y^* \in M_{\widehat{\sigma}}$ and $y \in M_\sigma$ be given such that $y^* \gtrsim_\sigma^b y$. By the induction hypothesis for σ there is a Φ_σ such that, using the symmetry in a and b, $\Phi_\sigma(y^*, n) \gtrsim_\sigma^a y$. Next, by the definition of $\gtrsim_{\sigma\to\tau}^a$ we have that $x^*(\Phi_\sigma(y^*, n)) \gtrsim_\tau^a xy$. But then by the induction hypothesis for τ there is a mapping Φ_τ such that

$$\Phi_\tau(x^* \Phi_\sigma(y^*, n), n) \gtrsim_\sigma^b xy.$$

Also for y^* s-maj$_{\widehat{\sigma}} \, \widehat{y}$ we have by I.H. that $\Phi_\sigma(y^*, n)$ s-maj$_{\widehat{\sigma}} \, \Phi_\sigma(\widehat{y}, n)$ and so for x^* s-maj$_{\widehat{\rho}} \, \widehat{x}$ we get

$x^*(\Phi_\sigma(y^*,n))$ s-maj$_{\widehat{\tau}}\,\widehat{x}(\Phi_\sigma(\widehat{y},n))$ which in turn implies that $\Phi_\tau(x^*(\Phi_\sigma(y^*,n)))$ s-maj$_{\widehat{\tau}}\Phi_\tau(\widehat{x}(\Phi_\sigma(\widehat{y},n)))$.
The desired $\Phi_{\sigma\to\tau}$ is then obtained by λ-abstracting x^*, n and y^*. In particular, $\lambda y^*.\Phi_\tau(x^*\Phi_\sigma(y^*,n),n)$ is a b-majorant for x and hence $x \in M^b_{\sigma\to\tau}$. □

Remark. *Even though it is independent of the choice of $a \in X$ whether or not a certain functional is a-majorizable, the complexity and possible uniformities of the majorants depend crucially on the choice of $a \in X$. In particular, for normed linear spaces a-majorants for $a \neq 0_X$ will usually depend on an upper bound $n \geq \|a\|$ and hence will not have the uniformity w.r.t. a that we aim for in our applications.*

We also need the following lemmas:

Lemma 9.39. *$x^* \gtrsim^a_\rho x \to x^* \gtrsim^a_{\widehat{\rho}} x^*$ for all $\rho \in \mathbf{T}^X$.*

Proof. By induction on ρ using that by definition of $\gtrsim^a$ if $x^* \gtrsim^a_{\rho\to\tau} x$ then $\forall z^*, z(z^* \gtrsim^a_{\widehat{\rho}} z \to x^*z^* \gtrsim^a_{\widehat{\tau}} x^*z)$. □

Lemma 9.40. *Let $\rho = \rho_1 \to \ldots \to \rho_k \to \tau$. Then $x^* \gtrsim^a_\rho x$ iff*
(I) $\forall y_1^*, y_1, \ldots, y_k^*, y_k\big(\bigwedge_{i=1}^k (y_i^* \gtrsim^a_{\rho_i} y_i) \to x^*y_1^*\ldots y_k^* \gtrsim^a_\tau xy_1\ldots y_k\big)$ and
(II) $\forall y_1^*, y_1, \ldots, y_k^*, y_k\big(\bigwedge_{i=1}^k (y_i^* \gtrsim^a_{\widehat{\rho}_i} y_i) \to x^*y_1^*\ldots y_k^* \gtrsim^a_{\widehat{\tau}} x^*y_1\ldots y_k\big)$.

Proof. By induction on k. The case $k=1$ follows from the definition of $\gtrsim^a$.

$k = n+1$: Let $\tau_0 = \rho_{n+1} \to \tau$. For '⇒', we have by induction hypothesis

$$\forall y_1^*, y_1, \ldots y_n^*, y_n\big(\bigwedge_{i=1}^n (y_i^* \gtrsim^a_{\rho_i} y_i) \to x^*y_1^*\ldots y_n^* \gtrsim^a_{\tau_0} xy_1\ldots y_n\big).$$

Now assume $y_{n+1}^* \gtrsim^a_{\rho_{n+1}} y_{n+1}$. Then by definition of $\gtrsim^a$

$$x^*y_n^*\ldots y_n^*y_{n+1}^* \gtrsim^a_\tau xy_1\ldots y_ny_{n+1},$$

so (I) follows. (II) can be treated analogously.

For '⇐', assume

$$\forall y_1^*, y_1, \ldots y_{n+1}^*, y_{n+1}\big(\bigwedge_{i=1}^{n+1} (y_i^* \gtrsim^a_{\rho_i} y_i) \to x^*y_1^*\ldots y_{n+1}^* \gtrsim^a_\tau xy_1\ldots y_{n+1}\big)$$

and

$$\forall y_1^*, y_1, \ldots y_{n+1}^*, y_{n+1}\big(\bigwedge_{i=1}^{n+1} (y_i^* \gtrsim^a_{\widehat{\rho}_i} y_i) \to x^*y_1^*\ldots y_{n+1}^* \gtrsim^a_{\widehat{\tau}} x^*y_1\ldots y_{n+1}\big).$$

We need to show that under these assumptions (1) $x^*y_1^*\ldots y_n^* \gtrsim^a_{\tau_0} xy_1\ldots y_n$ and (2) $x^*y_1^*\ldots y_n^* \gtrsim^a_{\widehat{\tau}_0} x^*y_1\ldots y_n$ hold. Then using the induction hypothesis we are done.

There are three cases to check:

(1a) $\bigwedge_{i=1}^{n+1} y_i^* \gtrsim_{\rho_i}^a y_i \to (x^* y_1^* \dots y_n^*) y_{n+1}^* \gtrsim_{\tau}^a (x y_1 \dots y_n) y_{n+1}$,

(1b) $\bigwedge_{i=1}^{n+1} y_i^* \gtrsim_{\widehat{\rho}_i}^a y_i \to (x^* y_1^* \dots y_n^*) y_{n+1}^* \gtrsim_{\widehat{\tau}}^a (x^* y_1^* \dots y_n^*) y_{n+1}$,

(2) $\bigwedge_{i=1}^{n+1} y_i^* \gtrsim_{\widehat{\rho}_i}^a y_i \to (x^* y_1^* \dots y_n^*) y_{n+1}^* \gtrsim_{\widehat{\tau}}^a (x^* y_1 \dots y_n) y_{n+1}$,

(1a) and (2) hold by assumption, (1b) follows from (2) using Lemma 9.39. $\square$

We need following (primitive recursive) functionals for the types $\rho \in \mathbf{T}$ whose definitions we recall here:

Definition 9.41. *For $\rho = \rho_1 \to \dots \to \rho_k \to 0 \in \mathbf{T}$ we define $\max_\rho$ by*

$$\max{}_{\rho}(x, y) := \lambda v_1^{\rho}, \dots, v_k^{\rho_k}. \max{}_{\mathbb{N}}(x\underline{v}, y\underline{v})$$

For types $0 \to \rho$ with $\rho = \rho_1 \to \dots \to \rho_k \to 0$, we define functionals $(\cdot)^M$ of types $(0 \to \rho) \to 0 \to \rho$ by :

$$x^M(y^0) := \lambda \underline{v}^{\rho}. \max{}_{\mathbb{N}}\{x(i, \underline{v}) \mid i = 1, \dots, y\}.$$

The next lemma is easy:

Lemma 9.42. *If $\forall n(x^*(n) \gtrsim_{\rho}^a x(n))$ then $(x^*)^M \gtrsim_{0\to\rho}^a x$.*

We now prove Theorem 9.18. We focus on hyperbolic spaces (X, d, W) and the corresponding theory $\mathcal{A}^\omega[X, d, W]_{-b}$. The case of ordinary metric spaces (X, d) follows by simply omitting the axioms concerning W, while for CAT(0) spaces we merely need to consider the additional purely universal axiom $\mathbf{CN}^-$.
The next lemma is an adaptation of the corresponding result from [77] and states that $\mathcal{A}^\omega[X, d, W]_{-b}$ has (via its so-called negative translation) a Gödel functional interpretation in $\mathcal{A}^\omega[X, d, W]_{-b} \setminus \{\text{QF-AC}\}$ (actually even without DC and in a quantifier-free fragment of this theory) augmented by the schema (BR) of simultaneous bar recursion in all types of $\mathbf{T}^X$ (for $\mathcal{A}^\omega$ this fundamental result is due to [114] (see [93] for a comprehensive treatment) which extends [44], where functional interpretation was introduced).
Let $\mathcal{A}^\omega[X, d, W]_{-b}^- := \mathcal{A}^\omega[X, d, W]_{-b} \setminus \{\text{QF-AC}\}$. Analogously, we define $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}^- := \mathcal{A}^\omega[X, \|\cdot\|, C]_{-b} \setminus \{\text{QF-AC}\}$ etc.

Lemma 9.43 ([77])**.** *Let A be a sentence in the language of $\mathcal{A}^\omega[X, d, W]_{-b}$. Then the following rule holds:*

$$\begin{cases} \mathcal{A}^\omega[X, d, W]_{-b} \vdash A \\ \Rightarrow \textit{one can construct a tuple of closed terms } \underline{t} \textit{ of } \mathcal{A}^\omega[X, d, W]_{-b} + (\text{BR}) \textit{ s.t.} \\ \mathcal{A}^\omega[X, d, W]_{-b}^- + (\text{BR}) \vdash \forall \underline{y}\, (A')_D(\underline{t}, \underline{y}). \end{cases}$$

where A' is the negative translation of A and $(A')^D \equiv \exists \underline{x} \forall \underline{y} (A')_D(\underline{x}, \underline{y})$ is the Gödel functional interpretation of A'.

Proof. This is Lemma 4.4 in [77],[21] except that we have one less purely universal axiom to interpret: the axiom that the metric of (X, d, W) is bounded by b. Also, as discussed in Section 9.2 the axioms concerning the hyperbolic function have been reformulated to implicitly satisfy $W_X(x, y, \lambda) =_X W_X(x, y, \tilde{\lambda})$. Recall that, in general, purely universal axioms not containing $\vee$, such as the additional axioms for metric, hyperbolic and CAT(0) spaces, are interpreted by themselves. □

Lemma 9.44. *Let (X, d, W) be a nonempty hyperbolic space. Then $\mathcal{M}^{\omega,X}$ is a model of $\mathcal{A}^\omega[X, d, W]^-_{-b}$ +(BR) (for a suitable interpretation of the constants of $\mathcal{A}^\omega[X, d, W]^-_{-b}$ +(BR) in $\mathcal{M}^{\omega,X}$), where we may interpret 0_X by an arbitrary element $a \in X$.*

Moreover, for any closed term t of $\mathcal{A}^\omega[X, d, W]^-_{-b}$ +(BR) one can construct a closed term t^* of $\mathcal{A}^\omega$ +(BR) – so, in particular, t^* does not contain the constants $0_X, d_X$ and W_X – such that

$$\mathcal{M}^{\omega,X} \models \forall a^X \forall n^0 ((n)_{\mathbb{R}} \geq d(0_X, a) \to t^*(n) \gtrsim^a t).$$

In particular, if we interpret 0_X by $a \in X$, then it holds in $\mathcal{M}^{\omega,X}$ that $t^*(0^0)$ is an a-majorant of t (note that $t^*(0^0)$ does not depend on a).

Proof. The constants of $\mathcal{A}^\omega[X, d, W]^-_{-b}$ +(BR) are interpreted as in [77], so next we need to show that all these functionals are in $\mathcal{M}^{\omega,X}$ by constructing a-majorants. To show that we can construct a suitable a-majorant t^* for any closed term t of $\mathcal{A}^\omega[X, d, W]^-_{-b}$ +(BR) it suffices to describe a-majorants for the constants of $\mathcal{A}^\omega[X, d, W]^-_{-b}$ +(BR).

We first describe the (trivially uniform) a-majorants for the constants of classical analysis $\mathcal{A}^\omega$, which now are taken over the extended set of types $\mathbf{T}^X$. Using Lemma 9.40 one easily verifies that:

- $0 \gtrsim^a_0 0$,
- $S \gtrsim^a_1 S$,
- $\Pi_{\widehat{\rho},\widehat{\tau}} \gtrsim^a \Pi_{\rho,\tau}$,
- $\Sigma_{\widehat{\sigma},\widehat{\rho},\widehat{\tau}} \gtrsim^a \Sigma_{\sigma,\rho,\tau}$.

To produce a-majorants for the (simultaneous) recursor(s) $\underline{R}$ and the (simultaneous) bar-recursor(s) $\underline{B}$[22], we only need the functional $\max_\rho$ defined for all types $\rho \in \mathbf{T}$. As a-majorants for R and B only operate on the types $\mathbf{T}$, we do not need to extend $\max_\rho$ to the types $\mathbf{T}^X$ as it was done in [77].

[21] Some errata to [77] are listed at the end of [42].

[22] As in [77], our formal systems are formulated with simultaneous recursion $\underline{R}$ and simultaneous bar-recursion $\underline{B}$, both of which could be defined primitive recursively in ordinary recursion R and ordinary bar-recursion B on the presence of appropriate product types. For convenience we supress the tuple notation in the following. See [77] for a detailed discussion.

By induction on n and using Lemma 9.40 one easily proves $\forall n(R_{\widehat{\rho}}n \gtrsim^a_\rho R_\rho n)$ and hence by Lemma 9.42 $R^M_{\widehat{\rho}} \gtrsim^a_\rho R_\rho$.

The majorant for the bar-recursor B is defined as $B^*_{\rho,\tau} := \lambda x, z, u, n, y.(B_{\rho,\tau}x^M z^M u_z)^M ny$, where $x^M(y^{0\to\rho}) := x(y^M)$, $z^M ny := zny^M$ and $u_z := \lambda v, n, y. \max(zny^M, uvny^M)$. As the defining axioms of B involve 0_X, we here assume $a = 0_X$, though by Lemma 9.38 given a bound on $d(a, 0_X)$ we may transform this majorant into a majorant for any choice of a. If 0_X does not occur, such that we may interpret 0_X by $a \in X$, the dependency on 0_X, resp. a bound on $d(a, 0_X)$ disappears.

The crucial step in proving $B^*_{\widehat{\rho},\widehat{\tau}} \gtrsim^{0_X} B_{\rho,\tau}$ is to establish the following: let $x^*, z^*, u^*, x, z, u, \widehat{x}, \widehat{z}, \widehat{u}$ be given s.t. $x^* \gtrsim_\alpha x$, $x^* \gtrsim_{\widehat{\alpha}} \widehat{x}$, $z^* \gtrsim_\beta z$, $z^* \gtrsim_{\widehat{\beta}} \widehat{z}$, $u^* \gtrsim_\gamma u$ and $u^* \gtrsim_{\widehat{\gamma}} \widehat{u}$, where α, β and γ are determined by ρ and τ. Then

$$(+)\ \forall y \in M_{0\to\widehat{\rho}} \forall n^0 Q(\overline{y, n}; n),$$

where $Q(\overline{y, n}; n) :\equiv$

$$\begin{cases} \forall y^*, \tilde{y}, \widehat{\tilde{y}} \in \mathcal{M}^{\omega,X}\big(\forall k(y^*k \gtrsim_\rho \tilde{y}k \wedge y^*k \gtrsim_{\widehat{\rho}} \widehat{\tilde{y}}k) \ \wedge\ \overline{y^*, n} =_{0\to\widehat{\rho}} \overline{y, n} \Rightarrow \\ \quad B^+x^*z^*u^*ny^* \gtrsim_\tau Bxzun\tilde{y} \wedge \\ \quad B^+x^*z^*u^*ny^* \gtrsim_{\widehat{\tau}} B^+x^*z^*u^*n\widehat{\tilde{y}}, B^+\widehat{x}\widehat{z}\widehat{u}n\widehat{\tilde{y}}\big), \end{cases}$$

where $B^+_{\rho,\tau} := \lambda x, z, u, n, y.B_{\rho,\tau}x^M zu_z ny$. By Lemma 9.40 and 9.42 it then follows that $B^* \gtrsim B$ and $B^*, B \in \mathcal{M}^{\omega,X}$.

The proof of (+) uses the following form of dependent choice, also called bar induction (which holds in $\mathcal{M}^{\omega,X}$ since by lemma 9.42 we have $M_{0\to\rho} = (M_\rho)^{\mathbb{N}}$):

$$\begin{cases} \forall y \in M_{0\to\widehat{\rho}} \exists n_0 \in \mathbb{N} \forall n \geq n_0\, Q(\overline{y, n}; n) \ \wedge \\ \forall y \in M_{0\to\widehat{\rho}}, n \in \mathbb{N}(\forall D \in M_{\widehat{\rho}}\, Q(\overline{y, n} * D; n+1) \to Q(\overline{y, n}; n)) \\ \to \forall y \in M_{0\to\widehat{\rho}}, n_0 \in \mathbb{N} Q(\overline{y, n}; n). \end{cases}$$

For the additional constants of $\mathcal{A}^\omega[X, d, W]^-_{-b}$ we define the following a-majorants:

- $n^0 \gtrsim^a 0_X$ for every n with $(n)_{\mathbb{R}} \geq_{\mathbb{R}} d_X(a, 0_X)$, where as just mentioned we can take $n := 0$ if we interpret 0_X by a,
- $0^0 \gtrsim^a a$, since $d_X(a, a) =_{\mathbb{R}} (0)_{\mathbb{R}}$,
- $\lambda x^0, y^0.((x + y)_{\mathbb{R}})_\circ \gtrsim^a d_X^{X\to X\to 1}$,[23]
- $\lambda x^0, y^0, z^1. \max_0(x, y) \gtrsim^a W_X^{X\to X\to 1\to X}$.

The a-majorants for 0_X and a are obvious. The a-majorant for d_X follows from the triangle inequality: assume $n_1 \gtrsim^a x$ and $n_2 \gtrsim^a y$ then

$$d(x, y) \leq d(x, a) + d(y, a) \leq n_1 + n_2.$$

[23] Here we refer to remark 9.4.

Hence, as in the model $\mathcal{M}^{\omega,X}$ the expression $d_X(x,y)$ is interpreted by $(d(x,y))_\circ$ (see [77]) and by Lemma 9.14 $n_1 + n_2 \geq d(x,y)$ implies $((n_1+n_2)_{\mathbb{R}})_\circ$ s-maj$_1 d(x,y))_\circ$, the validity of the a-majorant for d_X follows.

Finally, the a-majorant for W_X can be justified by the first axiom for hyperbolic spaces:

$$\forall x^X, y^X, z^X \forall \lambda^1 (d_X(z, W_X(x,y,\lambda)) \leq_{\mathbb{R}} (1 -_{\mathbb{R}} \tilde{\lambda}) d_X(z,x) +_{\mathbb{R}} \tilde{\lambda} d_X(z,y)),$$

The construction $\tilde{\lambda}$ turns representatives λ of arbitrary real numbers into representatives $\tilde{\lambda}$ of real numbers in the interval $[0,1]$. Hence we may reason that $d_X(z, W_X(x,y,\lambda))$ is less than the maximum of $d_X(z,x)$ and $d_X(z,y)$ and hence less than the maximum of respective upper bounds on $d_X(z,x)$ and $d_X(z,y)$. Note that without the reformulation of axioms (4)-(7) for hyperbolic spaces discussed in Section 9.2 (see Remark 9.2), this reasoning only holds in the model $\mathcal{M}^{\omega,X}$ (in which $W_X(x,y,\lambda)$ is interpreted by $W(x,y,r_{\tilde{\lambda}})$, where W is the function of the hyperbolic space (X,d,W) and $r_{\tilde{\lambda}}$ is the real number in $[0,1]$ represented by $\tilde{\lambda}$) whereas now it is even provable in $\mathcal{A}^\omega[X,d,W]_{-b}$.

Note, that the a-majorants for d_X, W_X are uniform, i.e. they do not depend on a. The a-majorant for a also does not depend on a, other than the requirement that the variable a and the element a in $\gtrsim^a$ denote the same element. Only the a-majorant for 0_X depends on a. Also note, that the $(\cdot)_\circ$-operator, which is ineffective in general, only is applied to natural numbers, where it is effectively (even primitive recursively) computable.

Thus given a closed term t of $\mathcal{A}^\omega[X,d,W]^-_{-b}$ +(BR) we construct an a-majorant t^* by induction on the term structure from the a-majorants given above, where we furthermore λ-abstract the majorant n for 0_X. Then one easily shows that

$$\mathcal{M}^{\omega,X} \models \forall a^X \forall n^0 ((n)_{\mathbb{R}} \geq_{\mathbb{R}} d_X(0_X, a) \to t^*(n) \gtrsim^a t).$$

where t^* does not contain $0_X, d_X$ and W_X and we may take $n := 0$ if we interpret 0_X by a. ☐

Lemma 9.44 also covers $\mathcal{A}^\omega[X,d]_{-b}$, simply by omitting the parts concerning the W-function, and $\mathcal{A}^\omega[X,d,W,\mathrm{CAT}(0)]_{-b}$, as this theory contains no additional constants that need to be majorized but only another purely universal axiom which is interpreted by itself.

Proof of Theorem 9.18.

Assume

$$\mathcal{A}^\omega[X,d,W]_{-b} \vdash \forall x^\rho \big(\forall u^0 B_\forall(x,u) \to \exists v^0 C_\exists(x,v)\big).$$

As in [77], this yields (using that negative translation and (partial) functional interpretation of the formula in question results in $\exists U, V \forall x^\rho \big(B_\forall(x, U(x)) \to C_\exists(x, V(x))\big)$) by Lemmas 9.43 and 9.44 the extractability of closed terms t_U, t_V of $\mathcal{A}^\omega[X,d,W]$+(BR) and closed terms t_{U^*}, t_{V^*} of $\mathcal{A}^\omega$+(BR) (in particular,

t_{U^*}, t_{V^*} do not contain $0_X, d_X, W_X$ and the other constants only for types in **T**) such that for all $n^0 \geq d(0_X, a)$

$$\mathcal{M}^{\omega,X} \models \begin{cases} t_{U^*}(n) \gtrsim^a t_U \wedge t_{V^*}(n) \gtrsim^a t_V \wedge \\ \quad \forall x^\rho \big(B_\forall(x, t_U(x)) \to C_\exists(x, t_V(x))\big). \end{cases}$$

Next, define the functional $\Phi(x^{\widehat{\rho}}, n) := \max(t_{U^*}(n,x), t_{V^*}(n,x))$, then

$$(+)\ \mathcal{M}^{\omega,X} \models \forall u \leq \Phi(x^*, n) B_\forall(x, u) \to \exists v \leq \Phi(x^*, n) C_\exists(x, v)$$

holds for all $n \in \mathbb{N}$, $x \in M_\rho$ and $x^* \in M_{\widehat{\rho}}$ for which there exists an $a \in X$ such that $n \geq d(0_X, a)$ and $x^* \gtrsim^a x$.

For the types γ of degree $\widehat{1}$ or $(1, X)$ of the quantifiers hidden in the definition of $\forall/\exists$-formulas we have at least $M_\gamma \subseteq S_\gamma$, which is sufficient for our purposes. This is because types of that kind have arguments for whose types δ one has – using lemma 9.42 – that $M_\delta = S_\delta$. For parameters x^ρ with ρ of degree 2 or $(1, X)$, we restricted ourselves to those $x \in S_\rho$ which have a-majorants $x^* \in S_{\widehat{\rho}}$. Since functionals of such types ρ only have arguments of types τ for which $M_\tau = S_\tau$ we get from $x^* \gtrsim^a_\rho x$ (which implies that $x^* \gtrsim^a_{\widehat{\rho}} x^*$) that $x^* \in M_{\widehat{\rho}}, x \in M_\rho$. Hence $\Phi(x^*, n)$ is defined and (+) yields

$$(++)\ \mathcal{S}^{\omega,X} \models \forall u \leq \Phi(x^*, n) B_\forall(x, u) \to \exists v \leq \Phi(x^*, n) C_\exists(x, v)$$

holds for all $n \in \mathbb{N}$, $x \in S_\rho$ and $x^* \in S_{\widehat{\rho}}$ for which there exists an $a \in X$ such that $n^0 \geq d(0_X, a)$ and $x^* \gtrsim^a x$ (here $\Phi(x^*, n)$ is interpreted as $[\Phi(x^*, n)]_{\mathcal{M}^\omega}$).

This finishes the proof, as Φ is a partial functional (which is always defined on the majorizable elements of $S_{\widehat{\rho}}$) in $S_{\widehat{\rho} \to 0 \to 0}$ which does not depend on (X, d, W).

Finally, if 0_X does not occur in either $B_\forall$ or $C_\exists$ we may freely interpret 0_X by $a \in X$. We then get majorants t_{U^*}, t_{V^*} and a resulting term Φ which no longer depend on a bound n on $d(0_X, a)$ (as we can take $n := 0$). □

Remark. *The proof of the soundness theorem of Gödel's functional interpretation (by closed terms) requires that we have closed terms for each type which we can use e.g. in order to construct the functional interpretation of axioms such as $\perp \to A$. That closed term can be arbitrarily chosen and usually is taken as the constant-0-functional of suitable type, where for the type X one takes by default 0_X. However, one could have also chosen an open term such as the constant-a-functional for types which map arguments to the type X. So rather than having just one term extracted by functional interpretation we have a whole family of such terms parametrized by a^X. In the last step of the previous proof we make use of this by picking the a-th term according to our choice of the reference point a for $\gtrsim^a$. By letting both $\gtrsim^a$ and the functional interpretation depend on a in a simultaneous way we achieve that the extracted majorant does not depend on the distance between a and any arbitrarily fixed constant such as 0_X.*

For normed linear spaces we focus on the theory $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$. The further cases $\mathcal{A}^\omega[X, \|\cdot\|, C, \eta]_{-b}$ and $\mathcal{A}^\omega[X, \|\cdot\|, C, \langle\cdot,\cdot\rangle]_{-b}$ follow by extending $\mathcal{A}^\omega[X, \|\cdot$

$\|, C]_{-b}$ respectively with an additional (majorizable) constant η for the modulus of uniform convexity or an additional purely universal axiom expressing the properties of the inner product in terms of the norm.

Lemma 9.45. *Let $(X, \|\cdot\|)$ be a non-trivial real normed linear space with a nonempty convex subset C. Then $\mathcal{M}^{\omega,X}$ is a model of $\mathcal{A}^\omega[X, \|\cdot\|, C]^-_{-b}$+(BR) (for a suitable interpretation of the constants of $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$ +(BR) in $\mathcal{M}^{\omega,X}$ where we have to interpret 0_X by the zero vector 0 in X and use $\gtrsim^{0_X}$).*

Moreover, for any closed term t of $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$ +(BR) one can construct a closed term t^* of $\mathcal{A}^\omega$ +(BR) such that

$$\mathcal{M}^{\omega,X} \models \forall n^0((n)_{\mathbb{R}} \geq_{\mathbb{R}} \|c_X\|_X \to t^*(n) \gtrsim^{0_X} t).$$

Similarly for $\mathcal{A}^\omega[X, \|\cdot\|, C, \eta]^-_{-b}$ +(BR) and $\mathcal{A}^\omega[X, \|\cdot\|, C, \langle\cdot,\cdot\rangle]^-_{-b}$ +(BR).

Proof. The proof is almost the same as the above proof of Lemma 9.44 and as before the interpretation of the constants of $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$+(BR) is as in [77].

The main difference to the proof of Lemma 9.44 is that we fix $a = 0_X$ (where 0_X now has to be interpreted by the zero vector of X), as otherwise we cannot define suitable (i.e. suitably uniform) majorants for the new constants of $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$. So now it suffices to state 0_X-majorants for the new constants:

- $0^0 \gtrsim^{0_X} 0_X$,
- $1^0 \gtrsim^{0_X} 1_X$,
- $\lambda x^0.((x)_{\mathbb{R}})_\circ \gtrsim^{0_X} \|\cdot\|_X^{X\to 1}$,
- $\lambda x^0, y^0.x + y \gtrsim^{0_X} +_X^{X\to X\to X}$,
- $\lambda x^0.x \gtrsim^{0_X} -_X^{X\to X}$,
- $\lambda \alpha^1, x^0.(\alpha(0) + 1)\cdot x \gtrsim^{0_X} \cdot_X^{1\to X\to X}$.

For the convex subset C, we have the characteristic term χ_C for the subset C, which is majorized as follows:

$$\lambda x^0.1 \gtrsim^{0_X} \chi_C^{X\to 0}.$$

For the constant $c_X \in C$ we have, given an $n \geq \|c_X\|$, the 0_X-majorant

$$n^0 \gtrsim^{0_X}_X c_X.$$

For uniformly convex spaces we 0_X-majorize the modulus $\eta : \mathbb{N} \to \mathbb{N}$ of uniform convexity by

$$(\eta)^M \gtrsim^{0_X}_1 \eta.$$

In [77], inner product spaces are defined by adding the so-called parallelogram law as another axiom. A norm satisfying the parallelogram law, allows one to define an inner product in terms of the norm and hence the inner product is immediately 0_X-majorizable. Inversely, if only an inner product is given the norm $\|\cdot\|$ can be recovered by defining $\|x\| := \sqrt{\langle x, x\rangle}$.

The majorants for the constants of normed linear spaces, with the exception of the modulus η of uniform convexity, are only seemingly uniform, since they depend on properties of the norm and hence on the choice of $a = 0_X$. The 0_X-majorants for $0_X, 1_X, \chi_C$ and c_X are obvious. For $\|\cdot\|_X$ we need to consider the interpretation of $\|\cdot\|_X$ in the model $\mathcal{M}^{\omega,X}$: the norm $\|x\|_X$ of an element $x \in X$ is interpreted by the actual norm using the $()_\circ$-operator, i.e. by $(\|x\|)_\circ$. In order to show that (in the model) $\lambda x.((x)_{\mathbb{R}})_\circ \gtrsim^a \|\cdot\|_X$ we need to show two things: (1) if $n \gtrsim^{0_X} x$ then $((n)_{\mathbb{R}})_\circ$ s-maj$_1(\|x\|)_\circ$ and (2) if $n \geq m$ then $((n)_{\mathbb{R}})_\circ$ s-maj$_1((m)_{\mathbb{R}})_\circ$ (recall that for $\rho \in \mathbf{T}$ s-maj$_1$ and $\gtrsim_1^a$ are equivalent). For (1), if $n \gtrsim^{0_X} x$ then by definition $(n)_{\mathbb{R}} \geq_{\mathbb{R}} \|x\|_X$ and the result then follows by Lemma 9.14. For (2) the result follows directly from Lemma 9.14. For $-_X$ the 0_X-majorant is derived straightforwardly from basic properties of the norm $\|\cdot\|_X$. For $+_X$ we additionally use the triangle inequality to verify the majorant, i.e. $\|x+y\| \leq \|x\| + \|y\|$ and then if $n_1 \gtrsim^{0_X} x$ and $n_2 \gtrsim^{0_X} y$ we have that $n_1 + n_2 \geq \|x+y\|$, and the validity of the majorant follows.

Finally, for scalar multiplication $\cdot_X$ we use that α codes a real number via a Cauchy sequence of rational numbers with fixed rate of convergence. The rational numbers in turn are represented by natural numbers using a monotone coding function such that $(\alpha(n))_{\mathbb{Q}} \geq_{\mathbb{Q}} |\alpha(n)|_{\mathbb{Q}}$ for all n. Since $|\lambda n^0.\alpha(0) -_{\mathbb{R}} \alpha| \leq 1$ the natural number $\alpha(0)+1$ is an upper bound for the real number represented by $|\alpha|_{\mathbb{R}}$. Now let α^* s-maj α. Then $\alpha^*(0)+1 \geq \alpha(0)+1$. Since $\|\alpha \cdot x\|_X =_{\mathbb{R}} |\alpha|_{\mathbb{R}} \cdot_{\mathbb{R}} \|x\|_X$ we, therefore, have that $\alpha^*(0)+1$ taken as a natural number multiplied with an $n \gtrsim^{0_X} x$ is a 0_X-majorant for $\alpha \cdot_X x$. □

Proof of Theorem 9.28. As in the proof of Theorem 9.18

$$\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b} \vdash \forall x^\rho \big(\forall u^0 B_\forall(x,u) \to \exists v^0 C_\exists(x,v)\big)$$

yields (using an easy adaptation of lemma 9.43 and Lemma 9.45) the extractability of closed terms t_U, t_V of $\mathcal{A}^\omega[X, \|\cdot\|, C]_{-b}$+(BR) and closed terms t_{U^*}, t_{V^*} of $\mathcal{A}^\omega$+(BR) (so, in particular, t_{U^*}, t_{V^*} do not contain $0_X, 1_X, c_X, +_X, -_X, \cdot_X$ or $\|\cdot\|_X$) such that for all $n^0 \geq \|c_X\|$

$$\mathcal{M}^{\omega,X} \models \begin{cases} t_{U^*}(n) \gtrsim^{0_X} t_U \wedge t_{V^*}(n) \gtrsim^{0_X} t_V \wedge \\ \quad \forall x^\rho \big(B_\forall(x, t_U(x)) \to C_\exists(x, t_V(x))\big). \end{cases}$$

As before, defining $\Phi(x^{\widehat{\rho}}, n) := \max(t_{U^*}(n,x), t_{V^*}(n,x))$, we then have that

$$\mathcal{M}^{\omega,X} \models \forall u \leq \Phi(x^*, n) B_\forall(x,u) \to \exists v \leq \Phi(x^*, n) C_\exists(x,v)$$

holds for all $n \in \mathbb{N}$, $x \in M_\rho$ and $x^* \in M_{\widehat{\rho}}$ for which $n \geq \|c_X\|$ and $x^* \gtrsim^{0_X} x$.

Also for normed linear spaces one verifies that for the types γ of degree $\widehat{1}$ and $(1,X)$ hidden in the definition of $\forall/\exists$-formulas, we have the necessary inclusion $M_\gamma \subseteq S_\gamma$. For parameters x^ρ with types ρ of degree $(1,X), (1,X,C)$ or 2, we again restricted ourselves to functionals which have 0_X-majorants x^* and hence the necessary inclusions hold.

Thus, also

$$\mathcal{S}^{\omega,X} \models \forall u \leq \Phi(x^*, n) B_\forall(x, u) \to \exists v \leq \Phi(x^*, n) C_\exists(x, v)$$

holds for all $n \in \mathbb{N}$, $x \in S_\rho$ and $x^* \in S_{\widehat{\rho}}$ for which $n^0 \geq \|c_X\|$ and $x^* \gtrsim^{0_X} x$. In uniformly convex spaces the bound additionally depends on a modulus η of uniform convexity via its majorant η^M.

This finishes the proof, as Φ is a partial (resp. total, if ρ is in addition of degree $\widehat{1}$) computable functional in $S_{\widehat{\rho} \to 0 \to 0}$, defined on the majorizable elements of $S_{\widehat{\rho}}$, which does not depend on $(X, \|\cdot\|, C)$. $\square$

Chapter 10

A quantitative version of Kirk's fixed point theorem for asymptotic contractions

The paper *A quantitative version of Kirk's fixed point theorem for asymptotic contractions* presented in this chapter has been published in the **Journal of Mathematical Analysis and Applications**, vol. 316, No.1, pp. 339-345, 2006. The paper has been slightly reformatted for inclusion in this PhD-thesis.

A quantitative version of Kirk's fixed point theorem for asymptotic contractions

Philipp Gerhardy

Abstract

In [J.Math.Anal.App.277(2003) 645-650], W.A.Kirk introduced the notion of asymptotic contractions and proved a fixed point theorem for such mappings. Using techniques from proof mining, we develop a variant of the notion of asymptotic contractions and prove a quantitative version of the corresponding fixed point theorem.

10.1 Introduction

In [59], W.A. Kirk proved a fixed-point theorem for so-called asymptotic contractions on complete metric spaces, showing that given a *continuous*[1] asymptotic contraction f for every starting point x the iteration sequence $\{f^n(x)\}$ converges to the unique fixed point of f. The proof is non-elementary, as it uses an ultrapower construction to establish the fixed point theorem. Recent alternative proofs by Jachymski and Jóźwik[54], additionally assuming that f is *uniformly* continuous, and by Arandelović [1], under the same assumptions as Kirk, are elementary and avoid ultrapowers, but neither of the three proofs provides explicit rates of convergence.

Using techniques from proof mining as developed e.g. in [81, 77], we first derive a suitable variant of the notion of asymptotic contractivity and subsequently give an elementary proof of Kirk's fixed point theorem, providing an explicit "rate of convergence"[2] (to the unique fixed point) for sequences $\{f^n(x)\}$.

In detail, we show that:

- the rate of convergence only depends on the starting point x via a bound on the iteration sequence $\{f^n(x)\}$,
- the rate of convergence only depends on the function f via suitable moduli expressing its asymptotic contractivity,
- the continuity of f is only necessary to prove the existence of a unique fixed point, while the convergence to such a fixed point can be proved without the continuity of f.

[1]In [54, 1], it is discussed that the requirement that f is continuous is a necessary condition for Kirk's fixed point theorem. By an oversight the requirement was left out in the original statement of Kirk's fixed point theorem in [59]

[2]Since an asymptotic contraction need not be non-expansive (cf. Example 2 in [54]), convergence need not be monotone, and hence in the general case can at most produce a bound M s.t. $f^m(x)$ is close to the unique fixed point for some $m \leq M$. We will discuss the details later.

10.2 Preliminaries

In [59], Kirk defines asymptotic contractions as follows:

Definition 10.1 (Kirk[59]). *A function $f : X \to X$ on a metric space (X, d) is called an asymptotic contraction with moduli $\phi, \phi_n : [0, \infty) \to [0, \infty)$ if ϕ, ϕ_n are continuous, $\phi(s) < s$ for all $s > 0$ and for all $x, y \in X$*

$$d(f^n(x), f^n(y)) \leq \phi_n(d(x, y))$$

and moreover $\phi_n \to \phi$ uniformly on the range of d.

What is needed to prove the fixed point theorem are not so much the moduli ϕ, ϕ_n, but instead a function η producing a witness of the inequality $\phi(s) < s$ and a modulus of convergence β for ϕ_n yielding a K s.t. for all $k \geq K$ ϕ_k is close enough to ϕ and hence f^k is a contraction. For η it is sufficient to provide a witness for every interval $[l, b]$, for β it suffices to have uniform convergence on every interval $[l, b]$, in both cases with $0 < l \leq b < \infty$.

Thus, to give an elementary and effective proof of the fixed point theorem proved by Kirk, we derive the following alternative definition of asymptotic contractions:

Definition 10.2. *A function $f : X \to X$ on a metric space (X, d) is called an asymptotic contraction if for each $b > 0$ there exist moduli $\eta^b : (0, b] \to (0, 1)$ and $\beta^b : (0, b] \times (0, \infty) \to \mathbb{N}$ and the following hold:*
(1) there exists a sequence of functions $\phi_n^b : (0, \infty) \to (0, \infty)$ s.t. for all $x, y \in X$, for all $\varepsilon > 0$ and for all $n \in \mathbb{N}$

$$b \geq d(x, y) \geq \varepsilon \Rightarrow d(f^n(x), f^n(y)) \leq \phi_n^b(\varepsilon) \cdot d(x, y),$$

(2) for each $0 < l \leq b$ the function $\beta_l^b := \beta^b(l, \cdot)$ is a modulus of uniform convergence for ϕ_n^b on $[l, b]$, i.e.

$$\forall \varepsilon > 0\ \forall s \in [l, b]\ \forall m, n \geq \beta_l^b(\varepsilon)\ \left(|\phi_m^b(s) - \phi_n^b(s)| \leq \varepsilon\right),$$

and (3) defining $\phi^b := \lim_{n\to\infty} \phi_n^b$, then for each $0 < \varepsilon \leq b$ we have $\phi^b(s) + \eta^b(\varepsilon) \leq 1$ for each $s \in [\varepsilon, b]$.

Remark. *The moduli η^b, β^b are necessary to derive explicit bounds later on. Conditions (2) and (3) may equivalently be defined without moduli η^b, β^b:*
(2') there is a $\phi^b : (0, b] \to (0, 1)$ s.t. for each $0 < l \leq b$ the sequence $\phi_n^b|_{[l,b]}$ converges uniformly to $\phi^b|_{[l,b]}$,
(3') for each $t \in (0, b]$, $\phi^b(t) < 1$ and $\limsup_{s\to t} \phi^b(s) < 1$ (here: if $t := b$, then we consider the leftside limit).

All the relevant information is contained in the moduli η^b and β^b and we do not need to refer to ϕ^b, ϕ_n^b at all, as the following proposition shows:

Proposition 10.3. *Let (X, d) be a metric space, let f be an asymptotic contraction and let $b > 0$ and η^b, β^b be given. Then for every $\varepsilon > 0$ there is a $K(\eta^b, \beta^b, \varepsilon)$ s.t. for all $k \geq K$, where $K = \beta^b_\varepsilon(\frac{\eta^b(\varepsilon)}{2})$,*

$$b \geq d(x, y) \geq \varepsilon \Rightarrow d(f^k(x), f^k(y)) \leq (1 - \frac{\eta^b(\varepsilon)}{2}) \cdot d(x, y).$$

Proof: Let $K = \beta^b_\varepsilon(\frac{\eta^b(\varepsilon)}{2})$, let a suitable sequence ϕ^b_n be given and let $\phi^b := \lim_{n\to\infty} \phi^b_n$. By the definition of η^b we have that $\phi^b(s) + \eta^b(\varepsilon) \leq 1$ for $s \in [\varepsilon, b]$. By the definition of β^b the function ϕ^b_k is at least $\frac{\eta^b(\varepsilon)}{2}$-close to ϕ^b for all $k \geq K$ and for all $s \in [\varepsilon, b]$ and hence also $\phi^b_k(s) \leq 1 - \frac{\eta^b(\varepsilon)}{2}$. □

Remark. *Requiring moduli η^b and β^b parametrized by b where $b > 0$ instead of one pair of moduli η, β for all $b > 0$ is no restriction. In the proof given in [59], it is assumed that some iteration sequence of the asymptotic contraction f is bounded, which allows to prove that every iteration sequence is bounded. Given $b > 0$, we say that a subset of X is b-bounded if its diameter is not greater than b. As we will see, to prove the fixed point theorem it suffices to have moduli η^b and β^b for the corresponding b-bounded subsets of (X, d).*

Next we show that Definition 10.2 covers Kirk's notion of asymptotic contractivity.

Definition 10.4. *Let $\phi : [0, \infty) \to [0, \infty)$, a sequence of functions $\phi_n : [0, \infty) \to [0, \infty)$ and $b > 0$ be given. Define:*

$\tilde{\phi}(s) := \frac{\phi(s)}{s}$ *for* $s \in (0, \infty)$, $\quad \tilde{\phi}_n(s) := \frac{\phi_n(s)}{s}$ *for* $s \in (0, \infty)$,

$\phi^b(s) := \sup_{t\in[s,b]} \tilde{\phi}(t)$ *for* $s \in (0, b]$, $\quad \phi^b_n(s) := \sup_{t\in[s,b]} \tilde{\phi}_n(t)$ *for* $s \in (0, b]$.

Proposition 10.5. *Let ϕ and ϕ_n be as in Definition 10.1 and let $\tilde{\phi}, \tilde{\phi}_n, \phi^b$ and ϕ^b_n be as in the above definition. Then*

- *$\tilde{\phi}$ and $\tilde{\phi}_n$ are continuous on $(0, \infty)$, $\tilde{\phi}(s) < 1$ for all $s \in (0, \infty)$ and the sequence $\tilde{\phi}_n$ converges uniformly to $\tilde{\phi}$ on $[l, \infty)$ for each $l > 0$,*
- *ϕ^b and ϕ^b_n are continuous on $(0, b]$, $\phi^b(s) < 1$ for all $s \in (0, b]$ and the sequence ϕ^b_n converges uniformly to ϕ^b on $[l, b]$ for each $0 < l \leq b < \infty$.*

Proof: Obvious. □

Remark. *The moduli η^b, β^b may equivalently be given as functions $\eta^b : \mathbb{N} \to \mathbb{N}$ and $\beta^b : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, where real numbers are approximated from below by suitable rational numbers 2^{-n}. Given $b > 0$, if ϕ and a modulus β for ϕ_n (ϕ, ϕ_n as in Kirk's definition) are given as computable number-theoretic functions, then η^b and β^b are effectively computable in b.*

Proposition 10.6. *If a function $f : X \to X$ on a metric space (X,d) is an asymptotic contraction (in the sense of Kirk) with moduli ϕ, ϕ_n, then the function f is an asymptotic contraction with suitable moduli η^b, β^b for every $b > 0$.*

Proof: Follows from the above remarks and constructions. □

10.3 Main results

We are now in position to give an elementary proof of Kirk's fixed point theorem. The general idea of the proof is similar to the constructivization of Edelstein's fixed point theorem in [81]. We first derive (variants of) a modulus of uniqueness and of a modulus of asymptotic regularity. Combining these two moduli one proves the convergence of the iteration sequence and thereby the convergence to a unique fixed point (additionally providing effective bounds).

Throughout this section we assume that $f : X \to X$ is a self-mapping on a metric space (X,d). Given $x_0 \in X$ we write x_n for $f^n(x_0)$ and $\{x_n\}$ for the corresponding iteration sequence. When there is no ambiguity we will omit the superscript b from the moduli η^b, β^b.

Lemma 10.7. *Let (X,d) be a metric space, let f be an asymptotic contraction and let $b > 0$ and η, β be given. Then for every $b \geq \varepsilon > 0$, for all $n \geq N$ and all $x, y \in X$ with $d(x,y) \leq b$*

$$d(x, f^n(x)), d(y, f^n(y)) \leq \delta \Rightarrow d(x,y) \leq \varepsilon,$$

where $\delta(\eta, \varepsilon) = \frac{\eta(\varepsilon)\cdot\varepsilon}{4}$ and $N(\eta, \beta, \varepsilon) = \beta_\varepsilon(\frac{\eta(\varepsilon)}{2})$.

Proof: Let $b \geq \varepsilon > 0$ be given and assume $d(x,y) \leq b$. Let $n \geq N$, then by Proposition 10.3

$$b \geq d(x,y) \geq \varepsilon \Rightarrow d(f^n(x), f^n(y)) \leq (1 - \frac{\eta(\varepsilon)}{2}) \cdot d(x,y).$$

Let $d(x, f^n(x)), d(y, f^n(y)) \leq \delta$, with $\delta = \frac{\eta(\varepsilon)\cdot\varepsilon}{4}$ and assume $d(x,y) > \varepsilon$. Then by the triangle inequality

$$\begin{aligned} d(x,y) &\leq d(x, f^n(x)) + d(f^n(x), f^n(y)) + d(y, f^n(y)) \\ &\leq \tfrac{\eta(\varepsilon)\cdot\varepsilon}{2} + (1 - \tfrac{\eta(\varepsilon)}{2}) \cdot d(x,y) \end{aligned}$$

and hence $\frac{\eta(\varepsilon)}{2} \cdot d(x,y) \leq \frac{\eta(\varepsilon)}{2} \cdot \varepsilon$ which implies $d(x,y) \leq \varepsilon$. But this contradicts the assumption $d(x,y) > \varepsilon$ and therefore $d(x,y) \leq \varepsilon$. □

Lemma 10.8. *Let (X,d) be a metric space, let f be an asymptotic contraction and let $b > 0$ and η, β be given. Then for every $\delta > 0$, for every $x_0 \in X$ s.t. $\{x_n\}$ is bounded by b and for every N there exists an $m \leq M$, s.t.*

$$d(x_m, f^N(x_m)) < \delta,$$

where $M(\eta, \beta, \delta, b) = k \cdot \lceil (\frac{lg(\delta) - lg(b)}{lg(1 - \frac{\eta(\delta)}{2})}) \rceil$ with $k = \beta_\delta(\frac{\eta(\delta)}{2})$.

Proof: Let $k = \beta_\delta(\frac{\eta(\delta)}{2})$. Assume for some M_0 and all $m < M_0$ we have $d(x_{mk}, f^N(x_{mk})) \geq \delta$, then repeatedly using Proposition 10.3

$$d(x_{M_0 k}, f^N(x_{M_0 k})) \leq (1 - \frac{\eta(\delta)}{2})^{M_0} d(x_0, f^N(x_0)) \leq (1 - \frac{\eta(\delta)}{2})^{M_0} \cdot b$$

since by assumption $d(x_0, f^N(x_0)) \leq b$.

Solving the inequality $(1 - \frac{\eta(\delta)}{2})^{M_0} \cdot b < \delta$ w.r.t. M_0 yields the described upper bound $M = k \cdot M_0$ on an m s.t. $d(x_m, f^N(x_m)) < \delta$. □

Remark. *Bounding m by M is in this context optimal. Since f^k only behaves like a (Banach) contraction mapping with constant $(1 - \frac{\eta(\delta)}{2})$ for x, y s.t. $d(x, y) \geq \delta$, we cannot be certain that $d(x_M, f^N(x_M)) < \delta$. An asymptotic contraction need not be nonexpansive (see [54]); hence the existence of an $m \leq M$ such that $d(x_m, f^N(x_m)) < \delta$ does not imply the distances between further f^k-iterates of $x_m, f^N(x_m)$ are less than δ. In particular, we do not know if $d(x_M, f^N(x_M)) < \delta$.*
If the function f and the space (X, d) have a computable representation one can of course check $x_0, \ldots, x_M$ to find which one satisfies $d(x_m, f^N(x_m)) \leq \delta$.

Lemma 10.9. *Let (X, d) be a metric space, let f be an asymptotic contraction and let $b > 0$ and η, β be given. Assume that f has a (unique) fixed point z. Then for every $\varepsilon > 0$ and every $x_0 \in X$ s.t. $\{x_n\}$ is bounded by b and $d(x_n, z) \leq b$ for all n there exists an $m \leq M$ s.t.*

$$d(x_m, z) \leq \varepsilon,$$

where $M(\eta, \beta, \varepsilon, b) = k \cdot \lceil (\frac{lg(\delta) - lg(b)}{lg(1 - \frac{\eta(\delta)}{2})}) \rceil$, $k = \beta_\delta(\frac{\eta(\delta)}{2})$, $\delta = \frac{\eta(\varepsilon)\cdot\varepsilon}{4}$.

Proof: By Lemma 10.7 for every $\varepsilon > 0$ there exist δ, N as described above s.t. if $d(x, f^N(x)), d(y, f^N(y)) \leq \delta$ then $d(x, y) \leq \varepsilon$. Any (trivially unique) fixed point z of f satisfies $d(z, f^N(z)) = 0 \leq \delta$, so if $d(x, f^N(x)) \leq \delta$ then $d(x, z) \leq \varepsilon$.

Now, by Lemma 10.8 for every δ and every N we can find an $m \leq M$ as described above s.t. $d(x_m, f^N(x_m)) < \delta$ and hence x_m is ε-close to the fixed point z. □

Note, that the functional M does not depend on the starting point x_0, but only on a bound b on $\{x_n\}$. Also, M only depends on f via the moduli η, β. Finally, the continuity of f was not necessary to prove this theorem.

Lemma 10.10. *Let (X, d) be a metric space, let f be an asymptotic contraction and let $b > 0$ and η, β be given. Then for every $\delta > 0$, for every $x_0 \in X$ s.t. $\{x_n\}$ is bounded by b and for every N there exists an M s.t. for all $m \geq M$*

$$d(x_m, f^N(x_m)) < \delta.$$

Proof: By Lemma 10.8 there exists an m s.t. $d(x_m, f^N(x_m)) < \delta$. Either $d(x_m, f^N(x_m)) = 0$ – then we are done – or $d(x_m, f^N(x_m)) > \varepsilon_0$ for some $\varepsilon_0 > 0$.

Let $K = \beta_{\varepsilon_0}(\frac{\eta(\varepsilon_0)}{2})$, then it follows by Proposition 10.3 that for all $k \geq K$

$$d(x_{m+k}), f^N(x_{m+k})) \leq (1 - \frac{\eta(\varepsilon_0)}{2})d(x_m, f^N(x_m)) < \delta.$$

Let $M = m + K$ and the result follows. □

Lemma 10.11. *Let (X, d) be a metric space, let f be an asymptotic contraction and let $b > 0$ and η, β be given. If $\{x_n\}$ is bounded by b then $\{x_n\}$ is a Cauchy sequence.*

Proof: By Lemma 10.7 for every $\varepsilon > 0$ there exists a $\delta > 0$ and an N s.t. $d(x, y) \leq \varepsilon$ for all $x, y \in X$ with $d(x, f^N(x)), d(y, f^N(y) \leq \delta$. By Lemma 10.10 for every $\delta > 0$ and every N there exists an M s.t. $d(x_m, f^N(x_m)) < \delta$ for all $m \geq M$. Then $d(x_m, x_n) \leq \varepsilon$ for all $m, n \geq M$. □

Theorem 10.12. *Let (X, d) be a complete metric space, let f be a continuous asymptotic contraction and let $b > 0$ and η, β be given. If for some $x_0 \in X$ the sequence $\{x_n\}$ is bounded by b then f has a unique fixed point z, $\{x_n\}$ converges to z and for every $\varepsilon > 0$ there exists an $m \leq M$ s.t.*

$$d(x_m, z) \leq \varepsilon,$$

where M is as in Lemma 10.9.

Proof: By Lemma 10.11 every iteration sequence $\{x_n\}$ which is bounded is a Cauchy sequence. Since (X, d) is complete the limit z of $\{x_n\}$ exists and using the continuity of f one then easily shows that $f(z) = z$, i.e. z is a fixed point of f. Every fixed point of f is trivially unique.

The bound M follows by Lemma 10.9. □

Remark. *By Remark 10.2 and Proposition 10.6, Theorem 10.12 implies Kirk's fixed point theorem for asymptotic mappings in [59].*

As mentioned in Remark 10.3, we do not know which x_m of $x_0, \ldots, x_M$ is ε-close to the fixed point z, and hence M is merely a bound on m. A bound M on m only is a rate of convergence under additional requirements that ensure that the convergence of $\{x_n\}$ towards z is monotone, such as e.g. weak quasi-nonexpansivity:

Definition 10.13. *A function $f : X \to X$ is called weakly quasi-nonexpansive if*

$$\exists p \in X (f(p) = p \land \forall x \in X d(f(x), p) \leq d(x, p)).$$

Corollary 10.14. *Let (X, d) be a complete metric space, let f be a continuous, weakly quasi-nonexpansive asymptotic contraction and let $b > 0$ and η, β be given. If for some x_0 the sequence $\{x_n\}$ is bounded by b then f has a unique fixed point z, $\{x_n\}$ converges to z and for every $\varepsilon > 0$ and all $n \geq M$*

$$d(x_n, z) \leq \varepsilon,$$

where $M(\eta, \beta, \varepsilon, b)$ is as in Lemma 10.9 and moreover M is a rate of convergence for $\{x_n\}$.

Proof: By Theorem 10.12 there exists $m \leq M$ s.t. $d(x_m, z) \leq \varepsilon$ where z is the unique fixed point of f and M is given as in Lemma 10.9. If the function f is weakly quasi-nonexpansive, convergence to the fixed point is monotone, so for all $n \geq M \geq m$ we have that $d(x_n, z) \leq d(x_m, z)$ and hence also $d(x_n, z) \leq \varepsilon$. □

Acknowledgements: I am grateful to Ulrich Kohlenbach for many useful discussions on the subject and helpful suggestions for improving the presentation of the material in this article. I would also like to thank the reviewer for making many helpful suggestions to improve the presentation.

Bibliography

[1] I. D. Arandelović. On a fixed point theorem by Kirk. *J. Math. Anal. Appl.*, 301:384–385, 2005.

[2] J. Avigad. Weak theories of nonstandard arithmetic and analysis. In S. Simpson, editor, *Reverse Mathematics 2001*, pages 19–46. A K Peters, Wellesley, MA, 2005.

[3] A. Beckmann. Exact bounds for lengths of reductions in typed λ-calculus. *J. of Symbolic Logic*, 66:1277–1285, 2001.

[4] G. Bellin. Ramsey interpreted: a parametric version of Ramsey's theorem. In *Logic and computation*, volume 106 of *Contemporaray Mathematics*, pages 17–37. AMS, Providence, R.I., 1990.

[5] S. Berardi, M. Bezem, and T. Coquand. On the computational content of the axiom of choice. *J. of Symbolic Logic*, 63:600–622, 1998.

[6] U. Berger. Program extraction from normalization proofs. In M. Bezem and J. Groote, editors, *Typed Lambda Calculi and Applications (TLCA'93)*, volume 664 of *Lecture Notes in Computer Science*, pages 91–106. Springer Verlag, Berlin, 1993.

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

[8] U. Berger and P. Oliva. Modified bar recursion and classical dependent choice. In M. Baaz, S.-D. Friedman, and J. Krajcek, editors, *Proceedings of the Annual European Summer Meeting of the Association for Symbolic Logic, Vienna, Austria*, volume 20 of *Lecture Notes in Logic*, pages 89–107. ASL and A K Peters, 2001.

[9] U. Berger and P. Oliva. Modified bar recursion. To appear in: Mathematical Structures in Computer Science, 2005.

[10] U. Berger and H. Schwichtenberg. An inverse of the evaluation functional for typed lambda-calculus. In R. Vemuri, editor, *Proceedings of the 6. Annual IEEE Symposium on Logic in Computer Sciences(LICS)*, pages 203–211. IEEE Press, Los Alamitos, 1991.

[11] U. Berger and H. Schwichtenberg. Program extraction from classical proofs. In D. Leivant, editor, *Logic and Computation Workshop, LCC '94*, volume 960 of *Lecture Notes in Computer Science*, pages 77–97. Springer Verlag, Berlin, 1995.

[12] M. Bezem. Strongly majorizable functionals of finite type: a model of bar recursion containing discontinous functionals. *J. of Symbolic Logic*, 50:652–660, 1985.

[13] M. Biernacka, O. Danvy, and K. Støvring. Program extraction from proofs of weak head normalization. In M. Escardò, editor, *Proceedings of the 21st Conference on the Mathematical Foundations of Programming Semantics*, Electronic Notes in Theoretical Computer Science, pages 105–124. Elsevier Science B.V., 2005.

[14] J. M. Borwein, S. Reich, and I. Shafrir. Krasnoselski-Mann iterations in normed spaces. *Canad. Math. Bull.*, 35:21–28, 1992.

[15] D. F. Bridges and F. Richman. *Varieties of constructive mathematics*, volume 97 of *London Mathematical Society Lecture Note Series.* Cambridge University Press, Cambridge, 1988.

[16] D. F. Bridges, F. Richman, W. H. Julian, and R. Mines. Extensions and Fixed Points of Contractive Maps in $\mathbb{R}^n$. *J. Math. Anal. Appl.*, 165:438–456, 1992.

[17] M. Bridson and A. Haefliger. *Metric spaces of non-positive curvature.* Springer Verlag, Berlin, 1999.

[18] E. M. Briseid. Proof Mining Applied to Fixed Point Theorems for Mappings of Contractive Type. Master's thesis, Oslo, 2005.

[19] E. M. Briseid. A rate of convergence for asymptotic continuous contractions. Preprint, 2006.

[20] L. E. Brouwer. On the siginificance of the principle of the excluded middle in mathematics, especially in function theory. In J. v. Heijenort, editor, *From Frege to Gödel: A sourcebook in mathematical logic, 1879-1931*, pages 334–345. Harvard University Press, Cambridge, MA, 1923.

[21] F. Bruhat and J. Tits. Groupes réductifs sur un corps local. I. Données radicielles valuées. *Inst. Hautes Études Sci. Publ. Math.*, 41:5–251, 1972.

[22] S. R. Buss. An Introduction to Proof Theory. In S. R. Buss, editor, *Handbook of Proof Theory*, pages 2–78. Elsevier Science B.V., 1998.

[23] C. E. Chidume. On the approximation of fixed points of nonexpansive mappings. *Houston Journal of Mathematics*, 7:345–355, 1981.

[24] P. Collaco and J. C. E. Silva. A complete comparison of 25 contractive conditions. *Nonlinear Analysis*, 30:471–476, 1997.

[25] C. A. Delzell. Kreisel's unwinding of Artin's proof - part I. In P. Odifreddi, editor, *Kreiseliana*, pages 113–246. A K Peters, Wellesley, MA, 1996.

[26] J. Diller. Logical problems of functional interpretations. *Annals of Pure and Applied Logic*, 114:27–42, 2002.

[27] J. Diller and W. Nahm. Eine Variante zur Dialectica-Interpretation der Heyting-Arithmetik Endlicher Typen. *Arch. Math. Logik*, 16:49–66, 1974.

[28] W. Dotson. On the Mann iterative process. *Trans. Amer. Math. Soc.*, 149:65–73, 1970.

[29] A. G. Dragalin. New kinds of realizability and the Markov rule. *Dokl. Akad. Nauk. SSSR*, 251:534–537, 1980. English translation in: *Soviet Math. Dokl.*, 21: 461-464, 1980.

[30] M. Edelstein. On Fixed and Periodic Points under Contractive Mappings. *J London Math. Soc.*, 37:74–79, 1962.

[31] M. Edelstein and R. O'Brien. Nonexpansive mappings, asymptotic regularity and successive approximations. *J. London Math. Soc.*, 17:547–554, 1978.

[32] S. Feferman. Kreisel's "unwinding" program. In P. Odifreddi, editor, *Kreiseliana*, pages 247–237. A K Peters, Wellesley, MA, 1996.

[33] H. Friedman. Classical and intuitionistically provably recursive functions. In G. Müller and D. Scott, editors, *Higher Set Theory*, volume 669 of *Springer LNM*, pages 21–27. Springer Verlag, Berlin, 1978.

[34] J. Garcia-Falset, E. Llorens-Fuster, and S. Prus. The fixed point property for mappings admitting a center. Talk given at '7th International Conference on Fixed Point Theory and its Applications', Juky 17-23, 2005, Guanajuato, Mexico.

[35] G. Gentzen. Die Widerspruchsfreiheit der reinen Zahlentheorie. *Mathematische Annalen*, 112:493–565, 1936.

[36] P. Gerhardy. Improved Complexity Analysis of Cut Elimination and Herbrand's Theorem. Master's thesis, Aarhus, 2003.

[37] P. Gerhardy. Refined Complexity Analysis of Cut Elimination. In M. Baaz and J. Makovsky, editors, *Proceedings of the 17th International Workshop CSL 2003*, volume 2803 of *LNCS*, pages 212–225. Springer-Verlag, Berlin, 2003.

[38] P. Gerhardy. The Role of Quantifier Alternations in Cut Elimination. *Notre Dame Journal of Formal Logic*, 46, No.2:165–171, 2005.

[39] P. Gerhardy. A quantitative version of Kirk's fixed point theorem for asymptotic contractions. *J. Math. Anal. Appl.*, 316, No.1:339–345, 2006.

[40] P. Gerhardy and U. Kohlenbach. Extracting Herbrand Disjunctions by Functional Interpretation. *Arch. Math. Logic*, 44:633–644, 2005.

[41] P. Gerhardy and U. Kohlenbach. General logical metatheorems for functional analysis, 2005. Submitted, 42pp.

[42] P. Gerhardy and U. Kohlenbach. Strongly uniform bounds from semi-constructive proofs. *Annals of Pure and Applied Logic*, 141:89–107, 2006.

[43] K. Gödel. Über formal unentscheidbare Sätze der Principia mathematica und verwandter Systeme i. *Monatshefte für Mathematik und Physik*, 38:173–198, 1931.

[44] K. Gödel. Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *Dialectica*, 12:280–287, 1958.

[45] K. Goebel and W. A. Kirk. Iteration processes for nonexpansive mappings. In S. Singh, S. Thomeier, and B. Watson, editors, *Topological Methods in Nonloinear Functional Analysis*, volume 21 of *Contemporaray Mathematics*, pages 115–123. AMS, Providence, R.I., 1983.

[46] J. v. Heijenort, editor. *From Frege to Gödel: A sourcebook in mathematical logic, 1879-1931.* Harvard University Press, Cambridge, MA, 1967.

[47] M.-D. Hernest. A comparison between two techniques of program extraction from classical proofs. In M. Baaz, J. Makovsky, and A. Voronkov, editors, *LPAR 2002: Short Contributions and CSL 2003: Extended Posters*, volume VIII of *Kurt Gödel Society's Collegium Logicum*, pages 99–102. Springer-Verlag, Berlin, 2004.

[48] M.-D. Hernest and U. Kohlenbach. A Complexity Analysis of Functional Interpretations. To appear in: Theoretical Computer Science.

[49] A. Heyting. *Mathematische Grundlagenforschung. Intuitionismus. Beweistheorie.* Springer-Verlag, Berlin, 1934.

[50] D. Hilbert. On the infinite. In J. v. Heijenort, editor, *From Frege to Gödel: A sourcebook in mathematical logic, 1879-1931*, pages 367–392. Harvard University Press, Cambridge, MA, 1925.

[51] W. A. Howard. Functional interpretation of bar induction by bar recursion. *Compositio Mathematica*, 20:107–124, 1968.

[52] W. A. Howard. Hereditarily majorizable functionals of finite type. In A. Troelstra, editor, *Metamathematical investigation of intuitionistic arithmetic and analysis*, volume 344 of *Springer LNM*, pages 454–461. Springer-Verlag, Berlin, 1973.

[53] S. Ishikawa. Fixed points and iterations of a nonexpansive mapping in a Banach space. *Proc. Amer. Math. Soc.*, 59:65–71, 1976.

[54] J. Jachymski and I. Jóźwik. On Kirk's asymptotic contractions. *J. Math. Anal. Appl.*, 300:147–159, 2004.

[55] M. Kauers. Algorithms for nonlinear higher order difference equations. Ph.D. Thesis, Linz, 2005.

[56] J. Kinces and V. Totik. Theorems and Counterexamples on Contractive Mappings. *Mathematica Balkanica*, 4:69–90, 1990.

[57] W. A. Kirk. Krasnosel'skii iteration process in hyperbolic spaces. *Numer. Funct. Anal. Optimiz.*, 4:371–381, 1982.

[58] W. A. Kirk. Nonexpansive mappings and asymptotic regularity. *Nonlinear Analysis*, 40:323–332, 2000.

[59] W. A. Kirk. Fixed points of asymptotic contractions. *J. Math. Anal. Appl.*, 277:645–650, 2003.

[60] S. C. Kleene. On the interpretation of intuitionistic number theory. *J. of Symbolic Logic*, 10:109–124, 1945.

[61] S. C. Kleene. Countable functionals. In A. Heyting, editor, *Constructivity in Mathematics*, pages 81–100. North-Holland Publishing Company, Amsterdam, 1959.

[62] S. C. Kleene. Recursive functionals and quantifiers of finite types, i. *Trans. Am. Math. Soc.*, 91:1–52, 1959.

[63] U. Kohlenbach. Proof interpretations and the computational content of proofs. Draft, January 2006, ii+412pp, http://www.mathematik.tu-darmstadt.de/~kohlenbach/newcourse.ps.gz.

[64] U. Kohlenbach. Theorie der stetigen und majorisierbaren funktionale und ihre anwendung bei der extraktion von schranken aus inkonstruktiven beweisen: Effektive eindeutigkeitsmodule bei besten approximationen aus ineffektiven eindeutigkeitsbeweisen. PhD thesis, Frankfurt, pp. xxii+278, 1990.

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

[66] U. Kohlenbach. Effective moduli from ineffective uniqueness proofs. An unwinding of de La Vallée Poussin's proof for Chebycheff approximation. *Annals of Pure and Applied Logic*, 64, No.1:27–94, 1993.

[67] U. Kohlenbach. New effective moduli of uniqueness and uniform a-priori estimates for constants of strong unicity by logical analysis of known proofs in best approximation theory. *Numer. Funct. Anal. Optimiz.*, 14:581–606, 1993.

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

[69] U. Kohlenbach. Arithmetizing proofs in analysis. In *Logic Colloquium '96*, volume 12 of *Springer Lecture Notes in Logic*, pages 115–158. Springer-Verlag, Berlin, 1998.

[70] U. Kohlenbach. Relative Constructivity. *J. of Symbolic Logic*, 63:1218–1238, 1998.

[71] U. Kohlenbach. On the no-counterexample interpretation. *J. of Symbolic Logic*, 64:1491–1511, 1999.

[72] U. Kohlenbach. A note on Spector's quantifier-free rule of extensionality. *Arch. Math. Logic*, 40:89–92, 2001.

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

[74] U. Kohlenbach. On the computational content of the Krasnoselski and Ishikawa fixed point theorems. In *Proc. of the Fourth Workshop on Computability and Complexity in Analysis*, pages 119–145. Spinger Lecture Notes in Computer Science LNCS 2064, 2001.

[75] U. Kohlenbach. Uniform asymptotic regularity for Mann iterates. *J. Math. Anal. Appl.*, 279:531–544, 2003.

[76] U. Kohlenbach. Some computational aspects of metric fixed point theory. *Nonlinear Analysis*, 61, no.5:823–837, 2005.

[77] U. Kohlenbach. Some logical metatheorems with applications in functional analysis. *Trans. Amer. Math. Soc.*, 357:89–128, 2005.

[78] U. Kohlenbach and B. Lambov. Bounds on iterations of asymptotically quasi-non-expansive mappings. In *Proc. International Conference on Fixed Point Theory and Applications, Valencia 2003*, pages 143–172. Yokohama Publishers, 2004.

[79] U. Kohlenbach and L. Leuştean. Mann iterates of directionally nonexpansive mappings in hyperbolic spaces. *Abstract and Applied Analysis*, 8:449–477, 2003.

[80] U. Kohlenbach and L. Leuştean. The approximate fixed point property in product spaces. To appear in: Nonlinear Analysis: Theory, Methods & Applications, 2005.

[81] U. Kohlenbach and P. Oliva. Proof mining: a systematic way of analyzing proofs in mathematics. *Proc. Steklov Inst. Math*, 242:136–164, 2003.

[82] U. Kohlenbach and P. Oliva. Proof mining in L_1-approximation. *Annals of Pure and Applied Logic*, 121:1–38, 2003.

[83] U. Kohlenbach and T. Streicher. Shoenfield = Gödel ∘ Krivine. Unpublished note.

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

[85] G. Kreisel. On the interpretation of non-finitist proofs I. *J. of Symbolic Logic*, 16:241–267, 1951.

[86] G. Kreisel. On the interpretation of non-finitist proofs II. *J. of Symbolic Logic*, 17:312–340, 1951.

[87] G. Kreisel. Mathematical significance of consistency proofs. *J. of Symbolic Logic*, 23:155–182, 1958.

[88] G. Kreisel. Interpretation of analysis by means of constructive functionals of finite types. In A. Heyting, editor, *Constructivity in Mathematics*, pages 101–128. North-Holland Publishing Company, Amsterdam, 1959.

[89] G. Kreisel. On weak completeness of intuitionistic predicate logic. *J. of Symbolic Logic*, 27:139–158, 1962.

[90] G. Kreisel. Review of [112]. in Mathematical Reviews, 1967. 37 #1224.

[91] G. Kreisel. Finiteness Theorems in Arithmetic : An Application of Herbrand's Theorem for Σ_2-formulas. In J. Stern, editor, *Logic Colloquium '81*, pages 39–55. North-Holland Publishing Company, 1982.

[92] L. Leuştean. A quadratic rate of asymptotic regularity for cat(0)-spaces. To appear in: Journal of Mathematical Analysis and Applications, 2005.

[93] H. Luckhardt. *Extensional Gödel Functional Interpretation.*, volume 306 of *Springer LNM.* Springer-Verlag, Berlin, 1973.

[94] H. Luckhardt. Herbrand-Analysen zweier Beweise des Satzes von Roth: Polynomiale Anzahlsschranken. *J. of Symbolic Logic*, 54:234–263, 1989.

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

[96] H. V. Machado. A characterization of convex subsets of normed spaces. *Kodai Math. Sem. Rep.*, 25:307–320, 1973.

[97] W. R. Mann. Mean value methods in iteration. *Proc. Amer. Math. Soc.*, 4:506–510, 1953.

[98] J. Meszáros. A Comparison of Various Definitions of Contractive Type Mappings. *Bull. Calcutta Math. Soc.*, 84:167–194, 1992.

[99] P. Odifreddi. *Classical recursion theory*, volume 125 of *Studies in Logic and the Foundations of Mathematics.* North-Holland Publishing, Amsterdam, 1989.

[100] A. G. O'Farrell. When uniformly-continuous implies bounded. *Irish Math. Soc. Bulletin*, 53:53–56, 2004.

[101] V. P. Orevkov. Lower bounds on the increase in complexity of deductions in cut elimination. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov*, 88:137–162, 1979. English transl., J. Soviet Math. **20**(1982), no.4.

[102] V. P. Orevkov. *Complexity of Proofs and Their Transformations in Axiomatic Theories*, volume 128 of *Translations of Mathematical Monographs.* AMS, Providence, R.I., 1993.

[103] P. Pudlak. The Length of Proofs. In S. R. Buss, editor, *Handbook of Proof Theory*, pages 548–637. Elsevier Science B.V., 1998.

[104] E. Rakotch. A Note on Contractive Mappings. *Proc. Am. Math. Soc.*, 13:459–465, 1962.

[105] S. Reich and I. Shafrir. Nonexpansive iterations in hyperbolic spaces. *Nonlinear Analysis, Theory, Methods and Applications*, 15:537–558, 1990.

[106] B. E. Rhoades. A Comparison of Various Definitions of Contractive Mappings. *Trans. Am. Math. Soc.*, 226:257–290, 1977.

[107] B. E. Rhoades. Contractive Definitions. In T. M. Rassias, editor, *Nolinear Analysis*, pages 513–526. World Sci. Publishing, Singapore, 1987.

[108] H. Schwichtenberg. Proof Theory : Some Applications of Cut-Elimination. In J. Barwise, editor, *Handbook of Mathematical Logic*, pages 868–894. North-Holland Publishing Company, 1977.

[109] H. Schwichtenberg. Complexity of normalization in the pure typed lambda-calculus. In A. Troelstra and D. van Dalen, editors, *The L.E.J. Brouwer Centenary Symposium*, pages 453–457. North-Holland Publishing Company, 1982.

[110] H. Schwichtenberg. An upper bound for reduction sequences in the typed λ-calculus. *Arch. Math. Logic*, 30:405–408, 1991.

[111] I. Shafrir. The approximate fixed point property in Banach and hyperbolic spaces. *Israel J. Math.*, 71:211–223, 1990.

[112] J. R. Shoenfield. *Mathematical Logic.* Addison-Wesley, Reading, MA, 1967.

[113] B. Sims. Examples of fixed point free mappings. In W. Kirk and B. Sims, editors, *Handbook of Metric Fixed Point Theory*, pages 35–48. Kluwer Academic Publishers, 2001.

[114] C. Spector. Provably recursive functionals of analysis : a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics. In J. Dekker, editor, *Proceedings of Symposia in Pure Mathematics*, volume 5, pages 1–27. AMS, Providence, R.I., 1962.

[115] R. Statman. Lower Bounds on Herbrand's Theorem. *Proc. Am. Math. Soc.*, 75:104–107, 1979.

[116] W. Takahashi. A convexity in metric and nonexpansive mappings. *I. Kodai Math. Sem. Rep.*, 22:142–149, 1970.

[117] A. S. Troelstra, editor. *Metamathematical investigation of intuitionistic arithmetic and analysis*, volume 344 of *Springer LNM.* Springer-Verlag, Berlin, 1973.

[118] A. S. Troelstra. Realizability. In S. Buss, editor, *Handbook of Proof Theory*, pages 407–473. Elsevier Science B.V., 1998.

[119] A. Visser. The unprovability of small inconsistency. *Arch. Math. Logic*, 32:275–298, 1993.

[120] W. Zhang. Cut elimination and automatic proof procedures. *Theoretical Computer Science*, 91:265–284, 1991.

[121] W. Zhang. Depth of proofs, depth of cut-formulas and complexity of cut formulas. *Theoretical Computer Science*, 129:193–206, 1994.