PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 2 (79–100)

PROOF-THEORETIC METHODS IN NONLINEAR ANALYSIS

ULRICH KOHLENBACH

Abstract

We discuss applications of methods from proof theory, so-called proof interpretations, for the extraction of explicit bounds in convex optimization, fixed point theory, ergodic theory and nonlinear semigroup theory.

1 Introduction: Proof Theory, Hilbert's Program and Kreisel's 'Unwinding of Proofs'

Proof theory has its origin in what has been called 'Hilbert's program': Since the 19th century noneffective and nonfinitary (set-theoretic) principles became increasingly important which raised the issue of their legitimacy. Hilbert's approach was to establish the consistency of a suitable formalization T of mathematics (first number theory and then analysis and set theory) within some finitary reasoning T_{fin}. In the language of number theory and with a minimal amount of number-theoretic tools one can express the consistency of T (axiomatized by an effective list of axioms) as a purely universal number-theoretic sentence (a so-called Π^0_1-sentence)

$$Con_T := \forall n \in \mathbb{N} \neg Prov_T(n, \ulcorner 0 = 1 \urcorner)$$

which states that no $n \in \mathbb{N} := \{0, 1, 2, \ldots\}$ is the code of a T-proof of $0 = 1$.

Consider now an arbitrary Π^0_1-sentence (called a 'real statement' by Hilbert) $S :\equiv \forall n \in \mathbb{N}\ (t(n) = 0)$, where t is some primitive recursive function term. If S is provable in T (using any nonfinitary 'ideal elements' of T), then also $T_{fin} + Con_T$ proves S (see Smorynski [1977][5.2.1]). So if Con_T could be proved in T_{fin}, one could convert the 'ideal' proof of S in T into a finitistic proof of S in T_{fin}.

Obviously, Gödel's second incompleteness theorem rules out that the consistency of any

MSC2010: primary 03F10; secondary 03F35, 47H10, 47H20, 47H25, 90C30.
Keywords: Proof mining, effective bounds, convex optimization, fixed points, ergodic theory, nonlinear semigroups.

$T \supseteq T_{fin}$ can be established inside of T_{fin}. Nevertheless, Hilbert's program gave rise to many 'relative consistency proofs' where the consistency of T is reduced to that of an in **some** sense more elementary theory T'. 'More elementary' often is related to being 'quantifier-free': e.g. Gentzen's proof-theoretic analysis of first-order number theory PA (Gentzen [1936]) reduces logically complex instances of ordinary induction to quantifier-free instances of transfinite induction (along ordinals $\alpha < \varepsilon_0$) and Gödel's consistency proof for PA (Gödel [1958]) reduces PA to a quantifier-free calculus of so-called primitive recursive functionals of finite type (considered already by Hilbert himself in Hilbert [1926]).
In the early 50's, Georg Kreisel suggested to re-orient proof theory by applying proof-theoretic methods - which in some way eliminate quantifiers in terms of quantifier-free constructions - to proofs of theorems which are not purely universal (as consistency statements) but e.g. of the form

$$(*)\ \forall n \in \mathbb{N}\, \exists m \in \mathbb{N}\, A_{qf}(n,m) \quad (A_{qf} \text{ quantifier-free}).$$

Kreisel noted that the respective consistency proofs for PA due to Gentzen (see Kreisel [1951, 1952]) and Gödel resp. (see Kreisel [1959] (3.4)) actually characterize the class of subrecursive functions f needed to realize $(*)$ in the form

$$\forall n \in \mathbb{N}\, A_{qf}(n, f(n))$$

for theorems $(*)$ which are provable in PA, namely as the class of $\alpha < \varepsilon_0$-recursive functions (in the case of Gentzen's proof) and - equivalently - as the class of functions definable in the aforementioned calculus of primitive recursive functionals (in the case of Gödel's proof Gödel [1941, 1958]), see also Parsons [1972].
While such results concern (the provability of $\forall\exists$-sentences in) formal systems such as PA rather than individual proofs, Kreisel already in Kreisel [1952] also launched the program of analyzing specific prima facie nonconstructive proofs with the aim of extracting new (e.g. effective) information on the theorem proven:

Input: A (prima facie) noneffective proof P of a conclusion C.

Goal: Additional information on C such as:

- effective bounds,
- algorithms,
- continuous dependency or full independence from certain parameters,
- generalizations of proofs: weakening of premises.

Kreisel's examples and suggestions for applications mainly concerned proofs in number theory. E.g. in Kreisel [1982], Kreisel suggested to analyze finiteness statements such as Roth's theorem in diophantine approximation with the aim of extracting bounds on the number of solutions. In Luckhardt [1989], Luckhardt extracted the first polynomial such bound for Roth's theorem from a proof due to Esnault and Viehweg (independently, this result was also obtained in Bombieri and van der Poorten [1988]). Since the 90's, the program has been developed most systematically and with specially designed so-called logical metatheorems (see the next section) in the context of nonlinear analysis ('proof mining'). Also while Kreisel's unwindings were based on techniques related to cut-elimination (Herbrand theory, ε-substitution etc.) the applications to analysis are all based on functional interpretations which have their origin in Gödel's 'Dialectica' interpretation on which Gödel's aforementioned consistency proof is based.

2 Logical metatheorems for bound extractions

In order to establish general theorems on the extractability of effective uniform bounds from given proofs one has to set up an appropriate formal deductive context. As the bound extraction methods are based on modern ('monotone') extensions and variants (see Kohlenbach [2008a]) of Gödel's functional interpretation (Gödel [1941, 1958]) one uses formal systems formulated in the language of functionals in all finite types such as appropriate forms of Peano arithmetic in all finite types PA^{ω}. In such systems one already can represent complete separable metric ('Polish') spaces (X, d) as continuous images of the Baire space $\mathbb{N}^{\mathbb{N}}$. However, this requires the separability of the space X and for separable spaces one can show that the independence of the extracted bounds from parameters in subspaces of X in general can only be expected if these subspaces are compact (see Kohlenbach [2008a] for discussions of this point). Many theorems in nonlinear analysis, however, involve - in addition to **concrete** Polish spaces such as $\mathbb{R}$ - general classes of **abstract** spaces X (e.g. general Hilbert spaces) which are not required to be separable and one can extract bounds that are independent from parameters in X (and even functions $T : X \to X$) if general metric bounds ('majorants') are given.

Many abstract types of metric structures can be added as atoms to our formal systems. E.g. this applies to metric, W-hyperbolic (see below), CAT(0), CAT(1), δ-hyperbolic, normed, uniformly convex, Hilbert, abstract L^p, abstract $C(K)$ spaces and $\mathbb{R}$-trees, and, in fact, all normed structures that are axiomatizable in so-called positive bounded logic (see Günzel and Kohlenbach [2016]). In order to be able to speak about such spaces one adds a **new base type** X to the formal system and forms all finite types over $\mathbb{N}, X$ (see Kohlenbach [2005b]; one may also have several such types: see Kohlenbach [2008a], section 17.6). One also adds constants for the metric d_X or normed space operators with appropriate axioms that characterize the class of structures in question.

Condition: the defining axioms must have a monotone functional interpretation (possibly with the addition of appropriate moduli, see Kohlenbach [2008a]).

Counterexamples (to the extractability of uniform bounds) exist for the classes of strictly convex or separable spaces which get upgraded by the monotone functional interpretation to uniformly convex resp. boundedly compact spaces.

Formal systems for analysis with abstract spaces X

Types: (i) $\mathbb{N}, X$ are types, (ii) with ρ, τ also $\rho \to \tau$ is a type.

Functionals of type $\rho \to \tau$ map type-ρ objects to type-τ objects.

$\mathrm{PA}^{\omega,X}$ is the extension of PA to all types, $\mathcal{A}^{\omega,X}$:=$\mathrm{PA}^{\omega,X}$+DC, where

DC: axiom schema of dependent choice for all types,

which implies the axiom schema of countable choice and so, applied to the law-of-excluded middle, full comprehension for numbers

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

where $A(n)$ may contain quantifiers (and parameters) of arbitrary types.

$\mathcal{A}^{\omega}[X, d, \ldots]$ results by adding constants $d_X, \ldots$ with axioms expressing that $(X, d, \ldots)$ is a nonempty metric, hyperbolic …space (we deviate here from the notation used in Kohlenbach [ibid.] where this theory is denoted by $\mathcal{A}^{\omega}[X, d, \ldots]_{-b}$, and $\mathcal{A}^{\omega}[X, d, \ldots]$ denotes the theory with an axiom stating the boundedness of (X, d) by some constant b being added).

A warning concerning equality: our formal theories only have a **quantifier-free rule of extensionality** (with A_{qf} being a quantifier-free formula)

$$\frac{A_{qf} \to s =_{\rho} t}{A_{qf} \to r[s/x] =_{\tau} r[t/x]},$$

where only $x =_{\mathbb{N}} y$ is a primitive predicate but for X and $\rho \to \tau$ one defines

$$x^X =_X y^X :\equiv d_X(x, y) =_{\mathbb{R}} 0_{\mathbb{R}},\ x =_{\rho\to\tau} y :\equiv \forall v^{\rho}(x(v) =_{\tau} y(v)).$$

This is crucial as the uniform quantitative rendering of the extensionality axiom $x =_X y \to Tx =_X Ty$ for T of type $X \to X$ implies the uniform continuity of T (on bounded subsets) and we want (in contrast to the setting of current continuous model theory; see, however, the recent Cho [2016]) also to be able to treat discontinuous situations (see Kohlenbach [2008a] for extensive discussions of this point).

Extension of majorizability to the new types: A crucial notion used is an extension of Howard's (Howard [1973]) concept of majorizability to the new types, where we 'bound' an element in a metric space by the distance it has from a fixed reference point $a \in X$ (where $a = 0_X$ in the normed case): let y, x be functionals be of types $\rho, \widehat{\rho} := \rho[\mathbb{N}/X]$ and a^X of type X:

$$x^{\mathbb{N}} \gtrsim^a_{\mathbb{N}} y^{\mathbb{N}} :\equiv x \geq y,\ x^{\mathbb{N}} \gtrsim^a_X y^X :\equiv x \geq d(y, a).$$

For complex types $\rho \to \tau$ this is extended in a hereditary fashion.

Example: for monotone T^* one defines

$$T^* \gtrsim^a_{X \to X} T \equiv \forall n \in \mathbb{N}, x \in X[n \geq d(a, x) \to T^*(n) \geq d(a, T(x))]$$

(see Gerhardy and Kohlenbach [2008] and Kohlenbach [2008a]).
$T : X \to X$ is nonexpansive (n.e.) if $d(T(x), T(y)) \leq d(x, y)$.
Then $\lambda n.n + b \gtrsim^a_{X \to X} T$, if $d(a, T(a)) \leq b$.
Proof mining exhibits the finitary combinatorial kernel of a proof and as a consequence of this it often is easy to generalize things from a normed linear setting to some geodesic setting. In fact, the approach has been particularly useful in the context of hyperbolic spaces which is a variant of notions considered by Takahashi [1970], Goebel and Kirk [1983] and Kirk [1981/82] and Reich and Shafrir [1990] (see Kohlenbach [2005b] for the precise relationship):

Definition 2.1 (Kohlenbach [ibid.]). *A $(W$-$)$hyperbolic space is a triple (X, d, W) where (X, d) is a metric space and $W : X \times X \times [0, 1] \to X$ s.t. for all $x, y, z \in W$ and $\lambda, \tilde{\lambda} \in [0, 1]$*

(i) $d(z, W(x, y, \lambda)) \leq (1 - \lambda)d(z, x) + \lambda d(z, y)$,

(ii) $d(W(x, y, \lambda), W(x, y, \tilde{\lambda})) = |\lambda - \tilde{\lambda}| \cdot d(x, y)$,

(iii) $W(x, y, \lambda) = W(y, x, 1 - \lambda)$,

(iv) $d(W(x, z, \lambda), W(y, w, \lambda)) \leq (1 - \lambda)d(x, y) + \lambda d(z, w)$.

CAT(0)-spaces (Gromov) are hyperbolic spaces (X, d, W) which satisfy the CN-inequality of Bruhat-Tits (determining W uniquely): for all $x, y_0, y_1, y_2 \in X$

$$\begin{cases} d(y_0, y_1) = \frac{1}{2}d(y_1, y_2) = d(y_0, y_2) \to \\ d(x, y_0)^2 \leq \frac{1}{2}d(x, y_1)^2 + \frac{1}{2}d(x, y_2)^2 - \frac{1}{4}d(y_1, y_2)^2. \end{cases}$$

Small types (over $\mathbb{N}, X$) include: $\mathbb{N},\ \mathbb{N} \to \mathbb{N},\ X,\ \mathbb{N} \to X,\ X \to X$.

Theorem 2.2 (Gerhardy and Kohlenbach [2008] and Kohlenbach [2008a]). *Let P, K be Polish resp. compact metric spaces (definable in $\mathcal{A}^\omega$), $A_\exists$*[1] *be an $\exists$-formula and $\underline{\tau}$ be a tuple of small types.*
If $\mathcal{A}^\omega[X, d, W]$ proves

$$\forall x \in P\, \forall y \in K\, \forall \underline{z}^{\underline{\tau}}\, \exists v^{\mathbb{N}}\, A_\exists(x, y, \underline{z}, v),$$

then one can extract a computable $\Phi : \mathbb{N}^{\mathbb{N}} \times \underline{\mathbb{N}}^{(\mathbb{N})} \to \mathbb{N}$ s.t. the following holds in every nonempty hyperbolic space: for all representatives $r_x \in \mathbb{N}^{\mathbb{N}}$ of $x \in P$ and all $\underline{z}^{\underline{\tau}}$ and $\underline{z}^* \in \mathbb{N}^{(\mathbb{N})}$ s.t. $\exists a \in X(\underline{z}^* \gtrsim^a_{\underline{\tau}} \underline{z})$:

$$\forall y \in K\, \exists v \leq \Phi(r_x, \underline{z}^*)\, A_\exists(x, y, \underline{z}, v).$$

For the case of bounded hyperbolic spaces, see Kohlenbach [2005b].

As a special case of the above metatheorem one has:

Corollary 2.3 (Gerhardy and Kohlenbach [2008] and Kohlenbach [2008a]). *If $\mathcal{A}^\omega[X, d, W]$ proves*

$$\forall x \in P\, \forall y \in K\, \forall z \in X\, \forall T : X \to X\, \big(T \text{ n.e.} \to \exists v \in \mathbb{N}\, A_\exists\big),$$

then one can extract a computable $\Phi : \mathbb{N}^{\mathbb{N}} \times \mathbb{N} \to \mathbb{N}$ s.t. for all $x \in P, b \in \mathbb{N}$

$$\forall y \in K\, \forall z \in X\, \forall T : X \to X\, \big(T \text{ n.e.} \wedge d_X(z, T(z)) \leq b \to \exists v \leq \Phi(r_x, b) A_\exists\big)$$

holds in all nonempty hyperbolic spaces (X, d, W).

Similar results hold for the other classes of metric and normed structures listed above. In the normed case, one additionally needs $\|z\| \leq b$ as an assumption in the conclusion of the corollary.

Remark 2.4. *Usually, proofs in ordinary mathematics only require a small fragment of $\mathcal{A}^\omega[X, d, \ldots]$ with e.g. the binary ('weak') König's lemma* WKL *instead of* DC *and Σ^0_1-induction only, which guarantees the extractability of primitive recursive (in the sense of Kleene) bounds.* WKL *is equivalent to a sentence of the form $\forall f^{\mathbb{N}\to\mathbb{N}}\, \exists b \leq_{\mathbb{N}\to\mathbb{N}} 1 \forall x^{\mathbb{N}}\, A_{qf}(f, b, x)$ and can be added to the system via a Skolem constant B with the purely universal axiom $\forall f, x(Bf \leq 1 \wedge A_{qf}(f, Bf, x))$ which is satisfiable in the full set-theoretic model and B is trivially majorized by the constant-1 functional in the extracted bound (see* Kohlenbach [2008a]*).*

[1] There are some mild restrictions on the types of the quantified variables in $A_\exists$.

3 General types of applications

3.1 Asymptotic regularity theorems. Consider a metric space (X, d) and a continuous function $F : X \to \mathbb{R}$. Many problems can be stated in the form of finding a zero $z \in X$ of F. Such problems are often algorithmically approached by setting up some iterative procedure resulting in a sequence (x_n) in X which converges to a zero z of F :

$$(*)\ F(\lim_{n\to\infty} x_n) = F(z) = 0.$$

In this case one, in particular, has that

$$(**)\ F(x_n) \to 0.$$

Quite often, $(**)$ holds under much more general conditions than those needed to ensure the convergence of (x_n) itself. In the case of fixed point problems for mappings $T : X \to X$, i.e. the case where $F(x) := d(x, Tx)$, results of the form $(**)$ are usually referred to as asymptotic regularity statements where this term was originally introduced by Browder and Petryshyn [1966] to refer to the property of T that the sequence $x_n := T^n x$ of Picard iterates satisfies $d(x_n, Tx_n) \to 0$. In many cases (see below) $(d(x_n, Tx_n))_{n\in\mathbb{N}}$ for some iterative process not only converges to 0 but does so in a nonincreasing way. In this situation the asymptotic regularity statement can be equivalently written in the form

$$\forall k \in \mathbb{N}\ \exists n \in \mathbb{N}\ (d(x_n, Tx_n) < 2^{-k}) \in \forall\exists$$

and any upper bound $\Phi(k)$ on '$\exists n$' provides a rate of convergence. This means that one can apply the logical metatheorems mentioned in the previous section to extract effective and highly uniform rates of asymptotic regularity even from prima facie noneffective proofs of asymptotic regularity. In fact, this has been achieved in many instances in the context of nonlinear analysis (see some of the applications below and Kohlenbach [2008b, 2017] for general surveys).

3.2 Strong convergence theorems. Suppose that the theorem to be studied is not about an asymptotic regularity result but about the convergence of the sequence (x_n) itself, e.g. towards a zero of F or a fixed point of T. Already the Cauchy property of (x_n)

$$(+)\ \forall k \in \mathbb{N}\ \exists n \in \mathbb{N}\ \forall i, j \geq n\ (d(x_i, x_j) \leq 2^{-k}) \in \forall\exists\forall$$

has too complicated a logical form to directly apply the logical metatheorems on uniform bound extractions and, in fact, there are already simple cases of computable monotone sequences of rational numbers in $[0, 1]$ which do not have a computable rate of convergence (Specker [1949]).

Roughly speaking, one can distinguish the following situations:
1) The proof of the Cauchy property of (x_n) (or of the convergence of (x_n) to some known element $x \in X$) uses - on top of constructive ('intuitionistic') reasoning - at most the law-of-excluded-middle schema LEM for negated formulas

$$\mathrm{LEM}_{\neg} : \neg A \vee \neg\neg A$$

which, in particular covers the case where A is $\exists$-free (e.g. $A \in \Pi^0_1$) as such formulas are equivalent to their double negation (using the stability of the prime formulas in our formal systems).
Alternatively (but not combined), one may use the so-called Markov principle

$$\mathrm{M} : \neg\neg\exists n \in \mathbb{N}\, A_{qf}(n) \rightarrow \exists n \in \mathbb{N}\, A_{qf}(n) \;\; (A_{qf} \text{ quantifier-free with parameters})$$

together with the following weak form of LEM (weaker than LEM for Π^0_1-formulas):

$$\mathrm{LLPO} : \neg(\exists n \in \mathbb{N}\, A_{qf}(n) \wedge \exists n \in \mathbb{N}\, B_{qf}(n)) \rightarrow \forall n \in \mathbb{N}\, \neg A_{qf}(n) \vee \forall n \in \mathbb{N}\, \neg B_{qf}(n),$$

where A_{qf}, B_{qf} are quantifier-free formulas. In both scenarios one can set up logical bound extraction metatheorems, where instead of the purely existential formula $A_{\exists}$ one may now have an arbitrary formula (see Kohlenbach [2008a]). Since $(+)$ is monotone w.r.t. '$\exists n \in \mathbb{N}$' any upper bound on $n \leq \Phi(k)$ in fact is a Cauchy rate for (x_n) and so one can in these cases extract effective rates of convergence.
2) If the proof of the Cauchy property of (x_n) uses LEM for Σ^0_1-formulas (purely existential formulas for natural numbers) as in the case of the Specker sequences from Specker [1949], then one often has the following dichotomy: either one can show that (x_n) converges to the **unique** zero of F or fixed point of T, or one can use the non-uniqueness of the solution to construct an instance of the Cauchy statement in question which provably does not allow for an effective Cauchy rate.

(i) **Unique existence:** in many cases one can obtain effective rates of convergence (and in fact also with a constructive verification of this fact) for (x_n) if (x_n) converges towards a **unique** zero of F resp. fixed point of T: consider a function $F : X \rightarrow \mathbb{R}$ on some metric space (X, d) which has exactly one zero z. The uniqueness part

$$(a)\; \forall x, y \in X\, (F(x) = 0 = F(y) \rightarrow x = y)$$

can be written equivalently as

$$(b)\; \forall x, y \in X\, \forall k \in \mathbb{N}\, \exists n \in \mathbb{N}\, (|F(x)|, |F(y)| \leq 2^{-n} \rightarrow d(x, y) < 2^{-k}) \in \forall\exists.$$

Then logical metatheorems can be applied to extract from a proof of (a) an effective uniform bound $\Phi(k)$ on '$\exists n \in \mathbb{N}$' in (b), which we called in Kohlenbach [1993] a 'modulus of uniqueness', where $\Phi(k)$ depends on x, y only via general majorizing data and, in particular, is independent of x, y if X is bounded (in the case where X can be treated as an abstract space and, otherwise, if X is compact). Suppose now that we can construct some (bounded) sequence (x_n) of approximate zeros, i.e.

$$(c)\ \forall k \in \mathbb{N}\, \exists n \in \mathbb{N} (|F(x_n)| < 2^{-k}) \in \forall\exists$$

from which we then can extract (using again a logical metatheorem) an effective bound $\Psi(k)$ on '$\exists n \in \mathbb{N}$' in (c), then for $\chi(k) := \Psi(\Phi(k))$ we have

$$\forall k \in \mathbb{N}\, \exists n \leq \chi(k)\, (d(x_n, z) < 2^{-k})$$

and, if we even have that $(|F(x_n)|)_n$ is nonincreasing, it follows that χ is a rate of convergence for $\lim x_n = z$. In Briseid [2009], it is shown that for Picard iterations $x_n = T^n x$ for suitable classes of mappings T the aforementioned logical metatheorems can be used to obtain such rates of convergence even when $(|F(x_n)|)_n$ (for $F(x) := d(x, Tx)$) is not nonincreasing which explains the explicit construction of effective rates of convergence for the classes of asymptotic contractions in the sense of Kirk and of uniformly generalized p-contractive mappings given by Briseid (see Briseid [ibid.] and the literature cited there).

(ii) **Non-unique existence:** when F or T possess many zeros resp. fixed points, one usually can construct computable instances of iterative procedures (x_n) (converging to some zero or fixed point) that do not have a computable rate of convergence. In fact, Neumann [2015] shows that this is the case for the usual iterative schemes used in metric fixed point theory, ergodic theory and convex optimization which even for (firmly) nonexpansive selfmappings $T : [0, 1] \to [0, 1]$ fail to have a computable rate of convergence for simple computable such mappings T. One then has to weaken the goal to what has been called an effective rate of metastability: Noneffectively, (+) is equivalent to

$$(++)\ \forall k \in \mathbb{N}\, \forall g \in \mathbb{N}^{\mathbb{N}}\, \exists n \in \mathbb{N}\, \forall i, j \in [n, n + g(n)]\, (d(x_i, x_j) < 2^{-k}) \in \forall\exists,$$

the so-called Herbrand normal of $(+)$, and a bound $\Phi(k, g)$ on '$\exists n$' is a bound for the Kreisel 'no-counterexample interpretation' (Kreisel [1951, 1952]) of the Cauchy property. Since Tao [2008b] calls an interval $[n, n + g(n)]$ with the property in $(++)$ an interval of 'metastability', we call bounds $\Phi(k, g)$ on '$\exists n$' in $(++)$ rates of metastability. If one additionally knows that (x_n) is converging to a zero of F or a fixed point of T with some rate of metastability then one can actually combine both rates into a common one (formulated here for the case of fixed points), i.e. a bound $\Phi(k, g)$ such that for all $k \in \mathbb{N}$ and $g : \mathbb{N} \to \mathbb{N}$

$$(+++)\ \exists n \leq \Phi(k, g)\, \forall i, j \in [n, n + g(n)]\, (d(x_i, x_j), d(x_i, Tx_i) \leq 2^{-k}).$$

If one has a rate of convergence for $d(x_n, Tx_n) \to 0$, one can even achieve that

$$\exists n \leq \Phi(k,g)\, \forall i,j \in [n, n+g(n)]\, \forall l \geq n\ (d(x_i, x_j), d(x_l, Tx_l) \leq 2^{-k})$$

(see e.g. Kohlenbach, Leuştean, and Nicolae [2018] and Kohlenbach [2016], Rem.2.11). The extraction of explicit bounds Φ on the metastable form of Cauchy or convergence statements is of interest for the following reasons:

a) Disregarding bounded quantifiers, the statement $(+++)$ is purely universal ('real') and captures all the mathematical content of the theorem $\lim x_n = x = Tx$: by a fixed piece of proof it implies back the original convergence theorem: forgetting the bound Φ gives the Herbrand normal form which by recursive comprehension (more precisely QF-AC^{0,0} in the terminology of Kohlenbach [2008a]) and LEM implies the Cauchy property and so by arithmetical comprehension (more precisely Π^0_1-AC^{0,0} in our formal context, see Kohlenbach [ibid.]) the convergence of (x_n). Applying $(+++)$ to the constant function $g(n) := K \in \mathbb{N}$ shows the existence of $i \geq K$ with $d(x_i, Tx_i) < 2^{-k}$ which - together with the continuity of T - gives $Tx = x$ for $x := \lim x_n$.

b) The proof-theoretic extraction of a rate of metastability from a convergence proof exhibits the finitary combinatorial content of that proof which may lead to generalizations of the resulting metastable statement and so - when unpacked into the full convergence statement (see above) - to generalized convergence theorems.

c) The concrete bounds extracted are of numerically interest: often they provide explicit information on the **algorithmic learnability** of a rate of convergence which - if a gap condition is satisfied - yields **oscillation bounds** (Avigad and Rute [2015] and Kohlenbach and Safarik [2014] and Section 5 below).

d) In many cases, asymptotic regularity is just the special case of metastability where $g(n) := 1$, e.g. for Picard iterates of nonexpansive functions T.
Some history:

- 2004, first rate of metastability (for the asymptotic regularity of asymptotically nonexpansive mappings) extracted (Kohlenbach and Lambov [2004]).
- 2005, rate of metastability for Krasnoselski-Mann iterations of nonexpansive self-mappings $T : X \to X$ of compact hyperbolic spaces X (Kohlenbach [2005a]).
- 2007, Tao [2008b] introduced the term 'metastability' in connection with the von Neumann Mean Ergodic Theorem (MET).
- 2007, independently from Tao, the first rate of asymptotic regularity for MET was extracted in Avigad, Gerhardy, and Towsner [2010].

- 2008, Kohlenbach and Leuştean [2009] generalized this with a better bound to uniformly convex Banach spaces which, subsequently, led to oscillation bounds by Avigad and Rute [2015] (see below).
- Since then, many papers extracting explicit rates of metastability have appeared, including Avigad and Rute [2015], Kohlenbach [2011, 2012, 2016], Kohlenbach and Koutsoukou-Argyraki [2015], Kohlenbach and Leuştean [2009, 2012, 2014], Kohlenbach, Leuştean, and Nicolae [2018], Kohlenbach, López-Acedo, and Nicolae [2017b], Körnlein [2016], Leuştean and Nicolae [2016], Safarik [2012], and Sipoş [2017a].

We like to emphasize that sometimes in analyzing convergence proofs one uses a combination of the approach used in the semi-constructive context discussed further above (applied to those parts of the proof that do not require Σ^0_1-LEM) and the approach to proofs based on full classical logic (applied to the more noneffective parts of the proof). E.g. Leuştean [2014] and Sipoş [2017b] provide interesting instances of such a hybrid approach.

In very special, but important, cases for applications one can extract rates of convergence for iterative procedures towards some non-unique zero of F or fixed point of T, namely when one has an effective so-called modulus of regularity which is closely related to the concepts of weak sharp minima and metric regularity used in convex optimization (see Kohlenbach, López-Acedo, and Nicolae [2017a]).

3.3 Inclusions between sets of solutions. Consider functions $F, G : X \to \mathbb{R}$ on a metric space (X, d) such that every zero of F is also one of G :

$$\forall x \in X \ (F(x) = 0 \to G(x) = 0)$$

which can be re-written in $\forall\exists$-form as

$$\forall x \in X \ \forall k \in \mathbb{N} \ \exists n \in \mathbb{N} \ \overbrace{(|F(x)| \le 2^{-n} \to |G(x)| < 2^{-k})}^{\in \Sigma^0_1}$$

so that logical metatheorems can be applied to extract effective uniform bounds (which due to monotonicity are in fact realizers) for '$\exists n$', i.e.

$$\forall k \in \mathbb{N} \ (|F(x)| \le 2^{-\Phi(x^*,k)} \to |G(x)| < 2^{-k}),$$

where x^* are appropriate majorizing data for x.
For concrete instances of such applications see sections 4 and 6 below.

3.4 Extraction of effective moduli. The first applications of the proof mining methodology in analysis concerned the extraction of explicit moduli of uniqueness in the aforementioned sense (as well as so-called constants of strong unicity) in Chebycheff approximation by us in 1990-1993 which in 2003 - together with Paulo Oliva - was also carried out for best L^1-approximation (see Kohlenbach [2008a] for an extensive coverage of this and the references given there). However, many more concepts of quantitative 'moduli' exist in mathematics or have been introduced as quantitative proof-theoretic versions of qualitative concepts in analysis. Proof mining has been used to explicitly transform moduli for one situation into moduli for another one. This e.g. is used essentially in Bačák and Kohlenbach [2018] and Kohlenbach, López-Acedo, and Nicolae [2017a].

In the rest of the paper we give a few typical examples of explicit bounds which have been obtained by the proof-theoretic machinery discussed so far. For more comprehensive surveys, see Kohlenbach [2008b] for results up to 2008 and Kohlenbach [2017] for applications since 2008.

4 Proof Mining in Convex Analysis

A polynomial rate of asymptotic regularity in Bauschke's solution of the 'zero displacement conjecture'

Consider a real Hilbert space H and nonempty closed and convex subsets $C_1, \ldots, C_N \subseteq H$ with metric projections P_{C_i}, define $T := P_{C_N} \circ \ldots \circ P_{C_1}$. In 2003, Bauschke proved the 'zero displacement conjecture' (Bauschke [2003]) which was first stated in Bauschke, Borwein, and Lewis [1997]:

$$\|T^{n+1}x - T^n x\| \to 0 \quad (x \in H).$$

Previously, this was only known for $N = 2$ or $Fix(T) \neq \emptyset$ (or even $\bigcap_{i=1}^N C_i \neq \emptyset$) or C_i half spaces etc.

The proof uses the Bruck and Reich [1977] theory of firmly and strongly nonexpansive mappings and the abstract theory of maximal monotone operators: Minty's theorem, Brézis-Haraux theorem, Rockafellar's maximal monotonicity and sum theorems, conjugate functions, normal cone operator.

The sequence $(\|T^{n+1}x - T^n x\|)_{n\in\mathbb{N}}$ is nonincreasing and hence the conclusion in Bauschke's theorem is of the form $\forall\exists$. Logical metatheorems as discussed above, therefore, guarantee (modulo the formalizability of the proof in the resp. formal system which, however, does not need to be checked if one explicitly has carried out the extraction) the extractability of an effective uniform rate of asymptotic regularity which only depends on the error $\varepsilon > 0$, $N \in \mathbb{N}$ and majorants for $x \in H$ and $P_{C_1}, \ldots, P_{C_N}$, i.e. $b \geq \|x\|$ and

$K \geq \|c_1\|, \ldots, \|c_N\|$ for some points $c_1 \in C_1, \ldots, c_N \in C_N$ since

$$n \geq \|y\| \to n + K \geq \|P_{C_i} y - P_{C_i} 0\| + \|P_{C_i} 0\| \geq \|P_{C_i} y\|.$$

So one gets a computable $\Phi(\varepsilon, N, b, K)$ s.t. for $b \geq \|x\|$

$$\forall \varepsilon > 0\, \forall n \geq \Phi(\varepsilon, N, b, K)\, (\|T^{n+1}x - T^n x\| < \varepsilon).$$

Strongly nonexpansive mappings

Definition 4.1 (Kohlenbach [2016]). *Let $S \subseteq X$ be a nonempty subset of a normed space X. $T : S \to X$ is strongly nonexpansive with SNE-modulus $\omega : \mathbb{R}^*_+ \times \mathbb{R}^*_+ \to \mathbb{R}^*_+$ if*

$$\forall d, \varepsilon > 0\, \forall x, y \in S\, (\|x - y\| \leq d \ \wedge \|x - y\| - \|Tx - Ty\| < \omega(d, \varepsilon) \\ \to \|(x - y) - (Tx - Ty)\| < \varepsilon).$$

Remark: T is strongly nonexpansive in the sense of Bruck and Reich [1977] iff it possesses an SNE-modulus.

Recall that in Hilbert spaces $H = X$, a function $T : S \to H$ is called firmly nonexpansive if

$$\forall x, y \in S\, (\|Tx - Ty\|^2 \leq \langle x - y, Tx - Ty \rangle)$$

and metric projections onto closed convex subsets of H are firmly nonexpansive.

The next two results have been obtained by a proof-theoretic analysis of Bruck and Reich [ibid.]:

Lemma 4.2 (Kohlenbach [2016]). *Let H be a real Hilbert space and $T = T_N \circ \ldots \circ T_1$ with firmly nonexpansive $T_1, \ldots, T_N : H \to H$. Then T is SNE with modulus*

$$\omega_T(d, \varepsilon) := \frac{1}{16d} \left(\frac{\varepsilon}{N}\right)^2.$$

A rate of asymptotic regularity for SNE-mappings

Theorem 4.3 (Kohlenbach [2018]). *Let $T : S \to S$ be SNE with modulus ω s.t.* $\inf\{\|x - Tx\| : x \in S\} = 0$ *and let $\alpha : \mathbb{R}^*_+ \to \mathbb{R}^*_+$ be such that*

$$\forall \varepsilon > 0\, \exists y \in S\ \ (\|y\| \leq \alpha(\varepsilon) \wedge \|y - Ty\| \leq \varepsilon)\,.$$

Then for $x \in S$, $x_n := T^n x$ and $D > 0$ such that $\|x - Tx\| \leq D$ one has

$$\forall \varepsilon > 0\, \forall n \geq \psi(\varepsilon, b, D, \alpha, \omega)\ \ (\|x_{n+1} - x_n\| < \varepsilon)\,, \text{ where}$$

$$\psi(\varepsilon, b, D, \alpha, \omega) := \left\lceil \frac{18b + 12\alpha(\varepsilon/6)}{\varepsilon} - 1 \right\rceil \left\lceil \left(\frac{D}{\omega(D, \tilde{\varepsilon})} \right) \right\rceil,\ \tilde{\varepsilon} := \frac{\varepsilon^2}{27b + 18\alpha(\varepsilon/6)}.$$

The proof-theoretic analysis of the **operator-theoretic part of Bauschke's proof** gives:

Theorem 4.4 (Kohlenbach [2018]). *Let H be real Hilbert space, $C_1, \ldots, C_N \subseteq H$ nonempty closed and convex subsets, P_{C_i} metric projections onto C_i for $i = 1, \ldots, N$. Let $c = (c_1, \ldots, c_N) \in C_1 \times \ldots \times C_N$ be arbitrary and $K \geq \|c\| = \sqrt{\sum_{i=1}^N \|c_i\|^2}$. Let $T := P_{C_N} \circ \ldots \circ P_{C_1}$. Then for every $\varepsilon \in (0,1)$ there exists a point $y \in C_N$ with*

$$\|y\| \leq \alpha(\varepsilon) \text{ and } \|Ty - y\| \leq \varepsilon, \text{ where}$$

$$\alpha(\varepsilon) := \frac{(K^2 + N^3(N-1)^2K^2)N^2}{\varepsilon}.$$

Corollary 4.5 (Kohlenbach [ibid.]).

$$\Phi(\varepsilon, N, b, K) := \left\lceil \frac{18b + 12\alpha(\varepsilon/6))}{\varepsilon} - 1 \right\rceil \left\lceil \left(\frac{D}{\omega(D, \tilde{\varepsilon})} \right) \right\rceil$$

is a rate of asymptotic regularity in Bauschke's result, where

$$\tilde{\varepsilon} := \frac{\varepsilon^2}{27b + 18\alpha(\varepsilon/6)}, \quad D := 2b + NK, \quad \omega(D, \tilde{\varepsilon}) := \frac{1}{16D}(\tilde{\varepsilon}/N)^2.$$

$$\alpha(\varepsilon) := \frac{(K^2 + N^3(N-1)^2K^2)N^2}{\varepsilon}.$$

The case where $Fix(T) \neq \varnothing$ is much simpler:

Theorem 4.6 (Kohlenbach [2016]). *Let $C \subseteq H$ be any nonempty subset of a real Hilbert space H, $T_1, \ldots, T_N : C \to C$ be firmly nonexpansive. Let $T := T_N \circ \ldots \circ T_1$ possess a fixed point $p \in C$ and, for $x \in C$, let $b \geq \|x - p\|, b > 0$. Then for $x_n := T^n x$:*

$$\forall \varepsilon > 0 \, \forall n \geq \lceil b/\omega_T(b, \varepsilon) \rceil \, (\|x_{n+1} - x_n\| < \varepsilon), \text{ where}$$

$$\omega_T(b, \varepsilon) := \frac{1}{16b}(\varepsilon/N)^2.$$

Convex feasibility problems

If in Theorem 4.6 the fixed point sets $Fix(T_1), \ldots, Fix(T_N)$ have a nonempty intersection, then any fixed point of T in fact is a common fixed point of $T_1, \ldots, T_N$. This even holds for arbitrary strongly nonexpansive mappings $T_1, \ldots, T_N$ in arbitrary Banach spaces X. In Kohlenbach [ibid.], an explicit bound $\rho(b, \varepsilon)$ (in terms of SNE-moduli for $T_1, \ldots, T_N$) is extracted from the classical proof of this fact such that for $x, p \in C$, p a common fixed point of $T_1, \ldots, T_N$ and $b \geq \|x - p\|$

$$\forall \varepsilon > 0 \, (\|T_N T_{N-1} \ldots T_1 x - x\| < \rho(b, \varepsilon) \to \bigwedge_{i=1}^{N} (\|T_i x - x\| < \varepsilon)).$$

Combined with a rate of asymptotic regularity for $T = T_N \circ \ldots \circ T_1$ (which even in this generality is provided in Kohlenbach [ibid.]) this quantitatively solves the problem of constructing a common approximate fixed point of $T_1, \ldots, T_N$.

All this largely holds even in general metric spaces and for strongly quasi-nonexpansive mappings in the sense of Bruck [1982]. Metric projections in so-called CAT(κ)-spaces X (in the sense of Gromov) with $\kappa > 0$ are strongly quasi-nonexpansive and one can construct an explicit modulus for this property which then makes it possible to quantitatively solve the problem to construct a point in the intersection of (ε-neighbourhoods of) finitely many overlapping closed convex subsets of X (i.e. the so-called convex feasibility problem for CAT(κ)-spaces). In the case where X is compact one obtains a rate of metastability for the strong convergence of the iterative use of the composition of the corresponding projections towards a point in the intersection of these sets (see Kohlenbach [2016]).

Other quantitative results in convex optimization have been obtained in

- Ariza-Ruiz, López-Acedo, and Nicolae [2015] and Kohlenbach, López-Acedo, and Nicolae [2017b]: rates of asymptotic regularity and - for compact X - metastability for iterations of compositions of two resolvents in CAT(0)-spaces.
- Kohlenbach, Leuştean, and Nicolae [2018], Kohlenbach, López-Acedo, and Nicolae [2017a], and Sipoş [2017a] rates of asymptotic regularity, strong convergence (in special cases) resp. metastability for the proximal point algorithm.
- Körnlein [2016] explicit such rates for Yamada's hybrid steepest descent method.

5 Proof Mining in Ergodic Theory

Let H be a real Hilbert space, $T : H \to H$ be linear and $\|T(x)\| \leq \|x\|$ for all $x \in H$. Consider the Cesàro mean of the iterates of T :

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

The von Neumann Mean Ergodic Theorem in the formulation of Riesz states:

Theorem 5.1 (von Neumann Mean Ergodic Theorem)**.** *For every $x \in H$, the sequence $(A_n(x))_n$ strongly converges.*

In Avigad, Gerhardy, and Towsner [2010], it is shown that in general there is no computable rate of convergence, but a primitive recursive rate of metastability is extracted using the proof-theoretic methods discussed above. Tao [2008a] also established (without bound) a uniform metastable version of the Mean Ergodic Theorem in Hilbert space and used that uniformity as a base step for a generalization to commuting families of operators. On the connection to the proof-theoretic approach he comments:

'We shall establish Theorem 1.6 by "finitary ergodic theory" techniques, reminiscent of those used in [Green-Tao]...' 'The main advantage of working in the finitary setting ... is that the underlying dynamical system becomes extremely explicit'...'In proof theory, this finitisation is known as Gödel functional interpretation...which is also closely related to the Kreisel no-counterexample interpretation' (T. Tao [2008a]).
In 1939, Garrett Birkhoff proved:

Theorem 5.2 (Birkhoff [1939]). *The Mean Ergodic Theorem holds for arbitrary uniformly convex Banach spaces.*

Remark 5.3. *In the same year as Birkhoff [ibid.], Lorch [1939] showed that the mean ergodic theorem even holds in all reflexive spaces. However, the class of reflexive spaces does not have enough uniformity to allow for a logical metatheorem on uniform bound extractions and, in fact, in Avigad and Rute [2015] it is shown that a uniform rate of metastability has to depend on the modulus of uniform convexity.*

Since Birkhoff's proof formalizes in the deductive framework of uniformly convex normed spaces (with modulus η) $\mathfrak{A}^\omega[X,\|\cdot\|,\eta]$ (see Kohlenbach [2008a] for the definition of this system) the following is guaranteed a-priorily:

Let X be a uniformly convex Banach space with modulus η and $T: X\to X$ nonexpansive linear operator. Let $b>0$. Then there is an effective functional Φ in ε, g, b, η s.t. for all $x\in X$ with $\|x\|\le b$, all $\varepsilon>0$, all $g:\mathbb{N}\to\mathbb{N}$:

$$\exists n\le \Phi(\varepsilon,g,b,\eta)\,\forall i,j\in[n,n+g(n)]\,\big(\|A_i(x)-A_j(x)\|<\varepsilon\big).$$

Note that $T^* := id$ majorizes T.

Based on the logical metatheorem above (for uniformly convex normed spaces) the following rate of metastability was extracted from Birkhoff's proof:

Theorem 5.4 (Kohlenbach and Leuştean [2009]). *Let X be a uniformly convex Banach space, η a modulus of uniform convexity, $T: X\to X$ be as above and $b>0$. Then for all $x\in X$ with $\|x\|\le b$, all $\varepsilon>0$ and all $g:\mathbb{N}\to\mathbb{N}$:*

$$\exists n\le \Phi(\varepsilon,g,b,\eta)\,\forall i,j\in[n,n+g(n)]\,\big(\|A_i(x)-A_j(x)\|<\varepsilon\big),\ \textit{where}$$

$$\Phi(\varepsilon,g,b,\eta):=M\cdot\tilde{h}^{(K)}(1),\ \textit{with}\ M:=\left\lceil\frac{16b}{\varepsilon}\right\rceil, \gamma:=\frac{\varepsilon}{16}\eta\left(\frac{\varepsilon}{8b}\right),\quad K:=\left\lceil\frac{b}{\gamma}\right\rceil,$$
$$h,\tilde{h}:\mathbb{N}\to\mathbb{N},\ h(n):=2(Mn+g(Mn)),\quad \tilde{h}(n):=\max\nolimits_{i\le n}h(i).$$

If $\eta(\varepsilon) = \varepsilon \cdot \tilde{\eta}(\varepsilon)$ with increasing $\tilde{\eta}$, then we can replace 'η' by '$\tilde{\eta}$' and '16' by '8'. In particular, for $X = L^p$ with $1 < p < \infty$, we may take $\tilde{\eta}(\varepsilon) = \varepsilon^{p-1}/(p2^p)$.

Bounding the number of fluctuations: We say that (x_n) admits k ε-fluctuations if there are $i_1 \leq j_1 \leq \dots i_k \leq j_k$ s.t. $\|x_{j_n} - x_{i_n}\| \geq \varepsilon$ for $n = 1, \dots, k$.

Using the analysis of Birkhoff's proof in Kohlenbach and Leuştean [ibid.], Avigad and Rute subsequently improved the rate of metastability to a bound on the number of ε-fluctuations:

Theorem 5.5 (Avigad and Rute [2015]). *$(A_n(x))$ admits at most*

$$\left\lfloor 4\log(M) \cdot \frac{b}{\varepsilon} \right\rfloor + \left\lfloor \frac{b}{\gamma} \right\rfloor \cdot \left\lfloor (4\log(2M) \cdot \frac{b}{\varepsilon} \right\rfloor + \left\lfloor \frac{b}{\gamma} \right\rfloor$$

many ε-fluctuations with b, M, γ as in Theorem 5.4.

In the Hilbert space case, fluctuation bounds had already been obtained in Jones, Ostrovskii, and Rosenblatt [1996].
If the linearity of the nonexpansive operator T is dropped, then the convergence of (x_n) holds weakly (but in general not strongly, see Genel and Lindenstrauss [1975]) by Baillon's nonlinear ergodic theorem:

Theorem 5.6 (Baillon [1975]). *Let H be a real Hilbert space, $C \subseteq H$ bounded closed and convex and $T : C \to C$ be nonexpansive. Then for every $x_0 \in C$, the sequence of Cesàro means (x_n) converges weakly to a fixed point of T.*

A rate of metastability for the weak Cauchy property is extracted in Kohlenbach [2012].

If one either changes the Cesàro means slightly (or adds some weak form of linearity, see below) one can achieve strong convergence. Consider the so-called Halpern iteration Halpern [1967]: Let $T : C \to C$ be nonexpansive, $x_1 \in C$, $\alpha_n \in [0, 1]$

$$x_{n+1} := \alpha_n x_1 + (1 - \alpha_n) T(x_n) \quad (n \geq 1).$$

In contrast to other iterative schemes such as Krasnoselski-Mann iterations, the Halpern iteration often converges strongly (one reason, though, why it is less used convex optimization is that it is not Fejér monotone; see Kohlenbach, Leuştean, and Nicolae [2018] for explicit rates of metastability from strong convergence proofs based on Fejér monotonicity).
Using a weak compactness argument, Wittmann proved in 1992 the following strong convergence result:

Theorem 5.7 (Wittmann [1992]). *Let H be a real Hilbert space, $C \subseteq H$ closed and convex, $x_0 \in C$ and $Fix(T) \neq \emptyset$. Under suitable conditions on (α_n) (e.g. for $\alpha_n := \frac{1}{n+1}$) (x_n) converges strongly towards the fixed point of T that is closest to x_0.*

Remark 5.8. 1. *Wittmann's theorem is a nonlinear generalization of the Mean Ergodic Theorem: for $\alpha_n := 1/(n+1), C := H$ and linear T, the Halpern iteration coincides with the Cesàro means.*

2. *Another nonlinear generalization of the Mean Ergodic Theorem has been obtained in* Baillon [1976]. *Here one keeps the original Cesàro means but requires that T (in addition to being nonexpansive) is odd (and C is symmetric). This was further generalized in* Wittmann [1990] *from which an explicit rate of metastability was extracted in* Safarik [2012].

Wittmann's result has been generalized to CAT(0)-spaces by Saejung [2010] using Banach limits. Explicit rates of metastability have been extracted in Kohlenbach [2011] (for Hilbert spaces) with an elimination of the use of weak compactness and in Kohlenbach and Leuştean [2012, 2014] (for CAT(0) spaces) with an elimination of the use of Banach limits.

Moreover, one has a quadratic rate of asymptotic regularity $d(x_n, T(x_n)) \to 0$:

$$\forall \varepsilon > 0 \, \forall n \geq \frac{4M}{\varepsilon} + \frac{32M^2}{\varepsilon^2} \; (d(x_n, T(x_n)) < \varepsilon)$$

(See Kohlenbach and Leuştean [2012].) In Leuştean and Nicolae [2016], the proof-theoretic analysis of Saejung's proof has been further generalized to the highly nontrivial case of CAT(κ)-spaces for $\kappa > 0$ producing an explicit rate of metastability even in this context.

6 Proof Mining in Nonlinear Semigroup Theory

Let X be a Banach space, $C \subseteq X$ be a nonempty subset and $\lambda \in (0,1)$.

Definition: A family $\{T(t) : t \geq 0\}$ of nonexpansive mappings $T(t) : C \to C$ is a nonexpansive semigroup if

$$\begin{aligned}
&(i)\; T(s+t) = T(s) \circ T(t) \;\; (s,t \geq 0),\\
&(ii)\; \text{for each } x \in C, \text{ the mapping } t \mapsto T(t)x \text{ is continuous.}
\end{aligned}$$

Theorem 6.1 (Suzuki [2006]). *Let $0 < \alpha < \beta$ such that α/β is irrational. Then any fixed point $p \in C$ of*

$$S := \lambda T(\alpha) + (1-\lambda)T(\beta) : C \to X$$

is a common fixed point of $T(t)$ for all $t \geq 0$.

Let $t \mapsto T(t)x$ be equicontinuous on norm-bounded subsets of C with modulus ω, let f_γ be an effective irrationality measure for $\gamma := \alpha/\beta$, $\Lambda, N, D \in \mathbb{N}$ be s.t. $1/\Lambda \leq \lambda, 1-\lambda$

and $1/N \leq \beta \leq D$. Then one can extract a bound (see Section 3.3) $\Phi(\varepsilon, M, b) := \Phi(\varepsilon, M, b, N, \Lambda, D, f_\gamma, \omega)$ s.t. for all $M, b \in \mathbb{N}$, $p \in C$, $\varepsilon > 0$

$$\|p\| \leq b \wedge \|S(p) - p\| \leq \Phi(\varepsilon, M, b) \to \forall t \in [0, M] \, (\|T(t)p - p\| \leq \varepsilon).$$

The main noneffective tool used in Suzuki's proof is the binary König's lemma WKL and by Remark 2.4 it is guaranteed to have a primitive recursive (in the sense of Kleene) bound Φ. In fact, the bound actually extracted in Kohlenbach and Koutsoukou-Argyraki [2016] is of rather low complexity:

$$\Phi(2^{-m}, M, b) = \frac{2^{-m}}{8(\sum_{i=1}^{\phi(k, f_\gamma)-1} \Lambda^i + 1)(1 + MN)}, \text{ where}$$

$$k := D 2^{\omega_{D,b}(3 + [\log_2(1+MN)] + m) + 1}, \ \phi(k, f) := \max\{2f(i) + 6 \ : \ 0 < i \leq k\}.$$

Example: $\alpha = \sqrt{2}, \beta = 2, \lambda = 1/2$. Then $\Lambda = 2, N = 1, D = 2, f_\gamma(p) = 4p^2$. If C is convex (so that $S : C \to C$) and $x_{n+1} := \frac{1}{2}x_n + \frac{1}{2}Sx_n \in C$ starting from $x_0 \in C$ is a d-bounded Krasnoselski iteration sequence of S one has a quadratic rate of asymptotic regularity $\Psi(\varepsilon, d) := 4d^2/(\pi\varepsilon^2)$ (Baillon and Bruck [1996]) and so

$$\forall n \geq \Psi(\Phi(\varepsilon, M, b), d) \ \forall t \in [0, M] \, (\|T(t)x_n - x_n\| \leq \varepsilon).$$

Nonexpansive semigroups feature prominently - via the Crandall-Liggett formula - in the study of abstract Cauchy problems that are given by accretive set-valued operators. Explicit rates on the asymptotic behavior of solutions have been obtained by our proof-theoretic methods in Kohlenbach and Koutsoukou-Argyraki [2015] and Koutsoukou-Argyraki [2017].

References

David Ariza-Ruiz, Genaro López-Acedo, and Adriana Nicolae (2015). "The asymptotic behavior of the composition of firmly nonexpansive mappings". *J. Optim. Theory Appl.* 167.2, pp. 409–429. MR: 3412444 (cit. on p. 93).

Jeremy Avigad, Philipp Gerhardy, and Henry Towsner (2010). "Local stability of ergodic averages". *Trans. Amer. Math. Soc.* 362.1, pp. 261–288. MR: 2550151 (cit. on pp. 88, 93).

Jeremy Avigad and Jason Rute (2015). "Oscillation and the mean ergodic theorem for uniformly convex Banach spaces". *Ergodic Theory Dynam. Systems* 35.4, pp. 1009–1027. MR: 3345161 (cit. on pp. 88, 89, 94, 95).

Miroslav Bačák and Ulrich Kohlenbach (2018). "On proximal mappings with Young functions in uniformly convex Banach spaces". *J. Convex Analysis* 25 (4) (cit. on p. 90).

Jean-Bernard Baillon (1975). “Un théorème de type ergodique pour les contractions non linéaires dans un espace de Hilbert”. *C. R. Acad. Sci. Paris Sér. A-B* 280.22, Aii, A1511–A1514. MR: 0375009 (cit. on p. 95).

– (1976). “Quelques propriétés de convergence asymptotique pour les contractions impaires”. *C. R. Acad. Sci. Paris Sér. A-B* 283.8, Aii, A587–A590. MR: 0428124 (cit. on p. 96).

Jean-Bernard Baillon and R. E. Bruck (1996). “The rate of asymptotic regularity is $O(1/n)$ $[O(1/\sqrt{}$ In: *Theory and applications of nonlinear operators of accretive and monotone type*. Vol. 178. Lecture Notes in Pure and Appl. Math. Dekker, New York, pp. 51–81. MR: 1386668 (cit. on p. 97).

Heinz H. Bauschke (2003). “The composition of projections onto closed convex sets in Hilbert space is asymptotically regular”. *Proc. Amer. Math. Soc.* 131.1, pp. 141–146. MR: 1929033 (cit. on p. 90).

Heinz H. Bauschke, Jonathan M. Borwein, and Adrian S. Lewis (1997). “The method of cyclic projections for closed convex sets in Hilbert space”. In: *Recent developments in optimization theory and nonlinear analysis (Jerusalem, 1995)*. Vol. 204. Contemp. Math. Amer. Math. Soc., Providence, RI, pp. 1–38. MR: 1442992 (cit. on p. 90).

Garrett Birkhoff (1939). “The mean ergodic theorem”. *Duke Math. J.* 5.1, pp. 19–20. MR: 1546101 (cit. on p. 94).

E. Bombieri and A. J. van der Poorten (1988). “Some quantitative results related to Roth’s theorem”. *J. Austral. Math. Soc. Ser. A* 45.2, pp. 233–248. MR: 951583 (cit. on p. 81).

Eyvind Martol Briseid (2009). “Logical aspects of rates of convergence in metric spaces”. *J. Symbolic Logic* 74.4, pp. 1401–1428. MR: 2583826 (cit. on p. 87).

F. E. Browder and W. V. Petryshyn (1966). “The solution by iteration of nonlinear functional equations in Banach spaces”. *Bull. Amer. Math. Soc.* 72, pp. 571–575. MR: 0190745 (cit. on p. 85).

Ronald E. Bruck (1982). “Random products of contractions in metric and Banach spaces”. *J. Math. Anal. Appl.* 88.2, pp. 319–332. MR: 667060 (cit. on p. 93).

Ronald E. Bruck and Simeon Reich (1977). “Nonexpansive projections and resolvents of accretive operators in Banach spaces”. *Houston J. Math.* 3.4, pp. 459–470. MR: 0470761 (cit. on pp. 90, 91).

Simon Cho (2016). “A variant of continuous logic and applications to fixed point theory”. arXiv: 1610.05397 (cit. on p. 82).

A. Genel and J. Lindenstrauss (1975). “An example concerning fixed points”. *Israel J. Math.* 22.1, pp. 81–86. MR: 0390847 (cit. on p. 95).

Gerhard Gentzen (1936). “Die Widerspruchsfreiheit der reinen Zahlentheorie”. *Math. Ann.* 112.1, pp. 493–565. MR: 1513060 (cit. on p. 80).

Philipp Gerhardy and Ulrich Kohlenbach (2008). "General logical metatheorems for functional analysis". *Trans. Amer. Math. Soc.* 360.5, pp. 2615–2660. MR: 2373327 (cit. on pp. 83, 84).

Kurt Gödel (1941). *In what sense is intuitionistic logic constructive*. Lecture at Yale. First published on pages 189–200 of Gödel [1995] (cit. on pp. 80, 81).

– (1958). "Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes". *Dialectica* 12, pp. 280–287. MR: 0102482 (cit. on pp. 80, 81).

– (1995). *Collected works. Vol. III – Unpublished essays and lectures*. Edited and with a preface by Solomon Feferman. The Clarendon Press, Oxford University Press, New York, pp. xvi+474. MR: 1332489.

Kazimierz Goebel and W. A. Kirk (1983). "Iteration processes for nonexpansive mappings". In: *Topological methods in nonlinear functional analysis (Toronto, Ont., 1982)*. Vol. 21. Contemp. Math. Amer. Math. Soc., Providence, RI, pp. 115–123. MR: 729507 (cit. on p. 83).

Daniel Günzel and Ulrich Kohlenbach (2016). "Logical metatheorems for abstract spaces axiomatized in positive bounded logic". *Adv. Math.* 290, pp. 503–551. MR: 3451931 (cit. on p. 81).

Benjamin Halpern (1967). "Fixed points of nonexpanding maps". *Bull. Amer. Math. Soc.* 73, pp. 957–961. MR: 0218938 (cit. on p. 95).

David Hilbert (1926). "Über das Unendliche". *Math. Ann.* 95.1, pp. 161–190. MR: 1512272 (cit. on p. 80).

W. A. Howard (1973). "Appendix: Hereditarily majorizable functionals of finite type". In: *Metamathematical investigation of intuitionistic arithmetic and analysis*. Ed. by A. S. Troelstra. Vol. 344. Lecture Notes in Math. Springer, Berlin, pp. 454–461. MR: 0469712 (cit. on p. 83).

Roger L. Jones, Iosif V. Ostrovskii, and Joseph M. Rosenblatt (1996). "Square functions in ergodic theory". *Ergodic Theory Dynam. Systems* 16.2, pp. 267–305. MR: 1389625 (cit. on p. 95).

W. A. Kirk (1981/82). "Krasnoselskiĭ's iteration process in hyperbolic space". *Numer. Funct. Anal. Optim.* 4.4, pp. 371–381. MR: 673318 (cit. on p. 83).

Ulrich Kohlenbach (1993). "Effective moduli from ineffective uniqueness proofs. An unwinding of de la Vallée Poussin's proof for Chebycheff approximation". *Ann. Pure Appl. Logic* 64.1, pp. 27–94. MR: 1241250 (cit. on p. 86).

– (2005a). "Some computational aspects of metric fixed-point theory". *Nonlinear Anal.* 61.5, pp. 823–837. MR: 2130066 (cit. on p. 88).

– (2005b). "Some logical metatheorems with applications in functional analysis". *Trans. Amer. Math. Soc.* 357.1, pp. 89–128. MR: 2098088 (cit. on pp. 81, 83, 84).

Ulrich Kohlenbach (2008a). *Applied proof theory: proof interpretations and their use in mathematics*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, pp. xx+532. MR: 2445721 (cit. on pp. 81–84, 86, 88, 90, 94).

– (2008b). "Effective uniform bounds from proofs in abstract functional analysis". In: *New computational paradigms*. Springer, New York, pp. 223–258. MR: 2762087 (cit. on pp. 85, 90).

– (2011). "On quantitative versions of theorems due to F. E. Browder and R. Wittmann". *Adv. Math.* 226.3, pp. 2764–2795. MR: 2739793 (cit. on pp. 89, 96).

– (2012). "A uniform quantitative form of sequential weak compactness and Baillon's nonlinear ergodic theorem". *Commun. Contemp. Math.* 14.1, 20 pp. MR: 2902296 (cit. on pp. 89, 95).

– (2016). "On the quantitative asymptotic behavior of strongly nonexpansive mappings in Banach and geodesic spaces". *Israel J. Math.* 216.1, pp. 215–246. MR: 3556967 (cit. on pp. 88, 89, 91–93).

– (2017). "Recent progress in proof mining in nonlinear analysis". *IFCoLog Journal of Logics and their Applications* 10 (4), pp. 3357–3406 (cit. on pp. 85, 90).

– (2018). "A polynomial rate of asymptotic regularity for compositions of projections in Hilbert space". *Found Comput Math* (cit. on pp. 91, 92).

Ulrich Kohlenbach and Angeliki Koutsoukou-Argyraki (2015). "Rates of convergence and metastability for abstract Cauchy problems generated by accretive operators". *J. Math. Anal. Appl.* 423.2, pp. 1089–1112. MR: 3278188 (cit. on pp. 89, 97).

– (2016). "Effective asymptotic regularity for one-parameter nonexpansive semigroups". *J. Math. Anal. Appl.* 433.2, pp. 1883–1903. MR: 3398796 (cit. on p. 97).

Ulrich Kohlenbach and Branimir Lambov (2004). "Bounds on iterations of asymptotically quasi-nonexpansive mappings". In: *International Conference on Fixed Point Theory and Applications, Valencia 2003*. Yokohama Publ., Yokohama, pp. 143–172. MR: 2144170 (cit. on p. 88).

Ulrich Kohlenbach and Laurenţiu Leuştean (2009). "A quantitative mean ergodic theorem for uniformly convex Banach spaces". *Ergodic Theory Dynam. Systems* 29.6, pp. 1907–1915. MR: 2563097 (cit. on pp. 89, 94, 95).

– (2012). "Effective metastability of Halpern iterates in CAT(0) spaces". *Adv. Math.* 231, pp. 2526–2556. MR: 2970458 (cit. on pp. 89, 96).

– (2014). "Addendum to "Effective metastability of Halpern iterates in CAT(0) spaces" [Adv. Math. 231 (5) (2012) 2526–2556] [MR2970458]". *Adv. Math.* 250, pp. 650–651. MR: 3122180 (cit. on pp. 89, 96).

Ulrich Kohlenbach, Laurenţiu Leuştean, and Adriana Nicolae (2018). "Quantitative results on Fejér monotone sequences". Comm. Contemp. Math 20, 42 pp. (cit. on pp. 88, 89, 93, 95).

Ulrich Kohlenbach, Genaro López-Acedo, and Adriana Nicolae (2017a). "Moduli of regularity and rates of convergence for Fejér monotone sequences". Preprint. arXiv: 1711.02130 (cit. on pp. 89, 90, 93).

– (2017b). "Quantitative asymptotic regularity results for the composition of two mappings". *Optimization* 66.8, pp. 1291–1299. MR: 3665825 (cit. on pp. 89, 93).

Ulrich Kohlenbach and Pavol Safarik (2014). "Fluctuations, effective learnability and metastability in analysis". *Ann. Pure Appl. Logic* 165.1, pp. 266–304. MR: 3111912 (cit. on p. 88).

Daniel Körnlein (2016). "Quantitative Analysis of Iterative Algorithms in Fixed Point Theory and Convex Optimization". PhD thesis. Technische Universität Darmstadt (cit. on pp. 89, 93).

Angeliki Koutsoukou-Argyraki (2017). "Effective rates of convergence for the resolvents of accretive operators". *Numer. Funct. Anal. Optim.* 38.12, pp. 1601–1613. MR: 3740740 (cit. on p. 97).

Georg Kreisel (1951). "On the interpretation of non-finitist proofs. I". *J. Symbolic Logic* 16, pp. 241–267. MR: 0049135 (cit. on pp. 80, 87).

– (1952). "On the interpretation of non-finitist proofs. II. Interpretation of number theory. Applications". *J. Symbolic Logic* 17, pp. 43–58. MR: 0051193 (cit. on pp. 80, 87).

– (1959). "Interpretation of analysis by means of constructive functionals of finite types". In: *Constructivity in mathematics: Proceedings of the colloquium held at Amsterdam, 1957 (edited by A. Heyting)*. Studies in Logic and the Foundations of Mathematics. North-Holland Publishing Co., Amsterdam, pp. 101–128. MR: 0106838 (cit. on p. 80).

– (1982). "Finiteness theorems in arithmetic: an application of Herbrand's theorem for Σ_2-formulas". In: *Proceedings of the Herbrand symposium (Marseilles, 1981)*. Vol. 107. Stud. Logic Found. Math. North-Holland, Amsterdam, pp. 39–55. MR: 757021 (cit. on p. 81).

Laurenţiu Leuştean (2014). "An application of proof mining to nonlinear iterations". *Ann. Pure Appl. Logic* 165, pp. 1484–1500. MR: 3210080 (cit. on p. 89).

Laurenţiu Leuştean and Adriana Nicolae (2016). "Effective results on nonlinear ergodic averages in CAT(κ) spaces". *Ergodic Theory Dynam. Systems* 36.8, pp. 2580–2601. MR: 3570025 (cit. on pp. 89, 96).

Edgar R. Lorch (1939). "Means of iterated transformations in reflexive vector spaces". *Bull. Amer. Math. Soc.* 45, pp. 945–947. MR: 0001460 (cit. on p. 94).

H. Luckhardt (1989). "Herbrand-Analysen zweier Beweise des Satzes von Roth: Polynomiale Anzahlschranken". *J. Symbolic Logic* 54.1, pp. 234–263. MR: 987335 (cit. on p. 81).

Eike Neumann (2015). "Computational problems in metric fixed point theory and their Weihrauch degrees". *Log. Methods Comput. Sci.* 11.4, 44 pp. MR: 3438729 (cit. on p. 87).

Charles Parsons (1972). “On n-quantifier induction”. *J. Symbolic Logic* 37, pp. 466–482. MR: 0325365 (cit. on p. 80).
Simeon Reich and Itai Shafrir (1990). “Nonexpansive iterations in hyperbolic spaces”. *Nonlinear Anal.* 15.6, pp. 537–558. MR: 1072312 (cit. on p. 83).
Satit Saejung (2010). “Halpern’s iteration in CAT(0) spaces”. *Fixed Point Theory Appl.* 2010. Article ID 471781, 13 pp. MR: 2578767 (cit. on p. 96).
Pavol Safarik (2012). “A quantitative nonlinear strong ergodic theorem for Hilbert spaces”. *J. Math. Anal. Appl.* 391.1, pp. 26–37. MR: 2899834 (cit. on pp. 89, 96).
Andrei Sipoş (2017a). “Contributions to Proof Mining”. PhD thesis. Univ. of Bucharest (cit. on pp. 89, 93).
– (2017b). “Effective results on a fixed point algorithm for families of nonlinear mappings”. *Ann. Pure Appl. Logic* 168.1, pp. 112–128. MR: 3564378 (cit. on p. 89).
Craig Smorynski (1977). “The incompleteness theorems”. In: *Barwise, J. (ed.), Handbook of Mathematical Logic*, pp. 821–865 (cit. on p. 79).
Ernst Specker (1949). “Nicht konstruktiv beweisbare Sätze der Analysis”. *J. Symbolic Logic* 14, pp. 145–158. MR: 0031447 (cit. on pp. 85, 86).
Tomonari Suzuki (2006). “Common fixed points of one-parameter nonexpansive semigroups”. *Bull. London Math. Soc.* 38.6, pp. 1009–1018. MR: 2285255 (cit. on p. 96).
Wataru Takahashi (1970). “A convexity in metric space and nonexpansive mappings. I”. *Kōdai Math. Sem. Rep.* 22, pp. 142–149. MR: 0267565 (cit. on p. 83).
Terence Tao (2008a). “Norm convergence of multiple ergodic averages for commuting transformations”. *Ergodic Theory Dynam. Systems* 28.2, pp. 657–688. MR: 2408398 (cit. on pp. 93, 94).
– (2008b). *Structure and randomness*. Pages from year one of a mathematical blog. American Mathematical Society, Providence, RI, pp. xii+298. MR: 2459552 (cit. on pp. 87, 88).
Rainer Wittmann (1990). “Mean ergodic theorems for nonlinear operators”. *Proc. Amer. Math. Soc.* 108.3, pp. 781–788. MR: 1004427 (cit. on p. 96).
– (1992). “Approximation of fixed points of nonexpansive mappings”. *Arch. Math. (Basel)* 58.5, pp. 486–491. MR: 1156581 (cit. on p. 95).

Received 2017-12-05.

Ulrich Kohlenbach
Department of Mathematics
Technische Universität Darmstadt
Schlossgartenstraße 7
64289 Darmstadt
Germany
kohlenbach@mathematik.tu-darmstadt.de