Proof Theory: From the Foundations of Mathematics to Applications in Core Mathematics

Ulrich Kohlenbach

Department of Mathematics

TECHNISCHE UNIVERSITÄT DARMSTADT

Talk at Tianyuan Mathematical Center in Central China and School of Philosophy, Wuhan University, March 9, 2021

Early history

Hilbert's Program:

Establish that uses of higher noneffective/transfinite **(,,ideal") principles** $\mathcal{I}$ in proofs of combinatorial/finitistic **(,,real") propositions** $\mathcal{P}$ can be **eliminated**, at least in principle.

Early history

Hilbert's Program:

Establish that uses of higher noneffective/transfinite **(,,ideal")** **principles $\mathcal{I}$** in proofs of combinatorial/finitistic **(,,real") propositions** $\mathcal{P}$ can be **eliminated**, at least in principle.

If

$$\mathcal{P} \equiv \forall \underline{n} \in \mathbb{N}^k \ (s(\underline{n}) = t(\underline{n}))\,, \text{ where } s, t \in \mathbb{Z}[\underline{x}],$$

this amounts to showing the **consistency** of $\mathcal{I}$ without $\mathcal{I}$ (e.g. using **finitary reasoning** only, Primitive Recursive Arithmetic).

Early history

Hilbert's Program:

Establish that uses of higher noneffective/transfinite **(,,ideal") principles** $\mathcal{I}$ in proofs of combinatorial/finitistic **(,,real") propositions** $\mathcal{P}$ can be **eliminated**, at least in principle.

If

$$\mathcal{P} \equiv \forall \underline{n} \in \mathbb{N}^k \ (s(\underline{n}) = t(\underline{n})), \text{ where } s, t \in \mathbb{Z}[\underline{x}],$$

this amounts to showing the **consistency** of $\mathcal{I}$ without $\mathcal{I}$ (e.g. using **finitary reasoning** only, Primitive Recursive Arithmetic).

Though in this specific form, in general impossible (Gödel), the basic approach is largely correct for existing ordinary mathematics: **proof-theoretic tameness** of ordinary mathematics!

Relative consistency proofs: proof interpretations

Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be theories with languages $\mathcal{L}(\mathcal{T}_1)$ and $\mathcal{L}(\mathcal{T}_2)$.

Relative consistency proofs: proof interpretations

Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be theories with languages $\mathcal{L}(\mathcal{T}_1)$ and $\mathcal{L}(\mathcal{T}_2)$.

- **Interpret** propositions A from $\mathcal{L}(\mathcal{T}_1)$ (inductively over the logical structure of A) by propositions A^I from $\mathcal{L}(\mathcal{T}_2)$.

Relative consistency proofs: proof interpretations

Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be theories with languages $\mathcal{L}(\mathcal{T}_1)$ and $\mathcal{L}(\mathcal{T}_2)$.

- **Interpret** propositions A from $\mathcal{L}(\mathcal{T}_1)$ (inductively over the logical structure of A) by propositions A^I from $\mathcal{L}(\mathcal{T}_2)$.
- **Transform a proof** p of A into a proof p^I of A^I (induction on p).

Relative consistency proofs: proof interpretations

Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be theories with languages $\mathcal{L}(\mathcal{T}_1)$ and $\mathcal{L}(\mathcal{T}_2)$.

- **Interpret** propositions A from $\mathcal{L}(\mathcal{T}_1)$ (inductively over the logical structure of A) by propositions A^I from $\mathcal{L}(\mathcal{T}_2)$.
- **Transform a proof** p of A into a proof p^I of A^I (induction on p).
- If

$$\mathbf{(0 = 1)^I \equiv (0 = 1)}$$

  this reduces the consistency problem of $\mathcal{T}_1$ to that of $\mathcal{T}_2$.

Kreisel's shift of emphasis

G. Kreisel (1951,...): use ***I*** to **extract new information** from interesting proofs of existential statements.

Kreisel's shift of emphasis

G. Kreisel (1951,...): use ***I*** to **extract new information** from interesting proofs of existential statements.

Given: Ineffective proof of some theorem T.

Kreisel's shift of emphasis

G. Kreisel (1951,...): use ***I*** to **extract new information** from interesting proofs of existential statements.

Given: Ineffective proof of some theorem T.

Goal: Additional information on T extracted from a proof of p by computing proof p^I for T^I:

Kreisel's shift of emphasis

G. Kreisel (1951,...): use I to **extract new information** from interesting proofs of existential statements.

Given: Ineffective proof of some theorem T.

Goal: Additional information on T extracted from a proof of p by computing proof p^I for T^I:

- **Quantitative information: effective bounds.**

Kreisel's shift of emphasis

G. Kreisel (1951,...): use ***I*** to **extract new information** from interesting proofs of existential statements.

Given: Ineffective proof of some theorem T.

Goal: Additional information on T extracted from a proof of p by computing proof p^I for T^I:

- **Quantitative information: effective bounds.**
- **Qualitative information: new uniformity results** (relevance pointed out by T. Tao).

Kreisel's shift of emphasis

G. Kreisel (1951,...): use ***I*** to **extract new information** from interesting proofs of existential statements.

Given: Ineffective proof of some theorem T.

Goal: Additional information on T extracted from a proof of p by computing proof p^I for T^I:

- **Quantitative information: effective bounds.**
- **Qualitative information: new uniformity results** (relevance pointed out by T. Tao).

Central methods: Monotone/bounded extensions and variants of Gödel's **Functional ('Dialectica') Interpretation!**

Kreisel's shift of emphasis

G. Kreisel (1951,...): use I to **extract new information** from interesting proofs of existential statements.

Given: Ineffective proof of some theorem T.

Goal: Additional information on T extracted from a proof of p by computing proof p^I for T^I:

- **Quantitative information: effective bounds.**
- **Qualitative information: new uniformity results** (relevance pointed out by T. Tao).

Central methods: Monotone/bounded extensions and variants of Gödel's **Functional ('Dialectica') Interpretation!**

Uses embedding into systems based in **intuitionistic logic** (**Brouwer**).

Monotone Functional interpretation in five minutes

Gödel's **functional interpretation** G is a map ($\mathcal{A}$ formal system of analysis)

$$G : Formula(\mathcal{A}) \to Formula(\mathcal{A}),\ A \mapsto A^G$$

such that

Monotone Functional interpretation in five minutes

Gödel's **functional interpretation** G is a map ($\mathcal{A}$ formal system of analysis)

$$G : Formula(\mathcal{A}) \to Formula(\mathcal{A}), \ A \mapsto A^G$$

such that

- $A^G \equiv \forall \underline{x} \exists \underline{y}\, A_G(\underline{x}, \underline{y})$, where A_G is **quantifier-free**,

Monotone Functional interpretation in five minutes

Gödel's **functional interpretation** G is a map ($\mathcal{A}$ formal system of analysis)

$$G : Formula(\mathcal{A}) \to Formula(\mathcal{A}), \; A \mapsto A^G$$

such that

- $A^G \equiv \forall \underline{x} \exists \underline{y} \, A_G(\underline{x}, \underline{y})$, where A_G is **quantifier-free**,
- For $A \equiv \forall \underline{x} \exists \underline{y} \, A_{qf}(\underline{x}, \underline{y})$ one has $A^G \equiv A$.

Monotone Functional interpretation in five minutes

Gödel's **functional interpretation** G is a map ($\mathcal{A}$ formal system of analysis)

$$G : Formula(\mathcal{A}) \to Formula(\mathcal{A}),\ A \mapsto A^G$$

such that

- $A^G \equiv \forall \underline{x} \exists \underline{y}\, A_G(\underline{x}, \underline{y})$, where A_G is **quantifier-free**,
- For $A \equiv \forall \underline{x} \exists \underline{y}\, A_{qf}(\underline{x}, \underline{y})$ one has $A^G \equiv A$.
- $A \leftrightarrow A^G$ by **quantifier-free choice** in all types

$$\textbf{QF-AC} : \ \forall \underline{a} \exists \underline{b}\, F_{qf}(\underline{a}, \underline{b}) \to \exists \underline{B} \forall \underline{a}\, F_{qf}(\underline{a}, \underline{B}(\underline{a})).$$

Monotone Functional interpretation in five minutes

Gödel's **functional interpretation** G is a map ($\mathcal{A}$ formal system of analysis)

$$G : \mathit{Formula}(\mathcal{A}) \to \mathit{Formula}(\mathcal{A}),\ A \mapsto A^G$$

such that

- $A^G \equiv \forall \underline{x} \exists \underline{y}\, A_G(\underline{x}, \underline{y})$, where A_G is **quantifier-free**,
- For $A \equiv \forall \underline{x} \exists \underline{y}\, A_{qf}(\underline{x}, \underline{y})$ one has $A^G \equiv A$.
- $A \leftrightarrow A^G$ by **quantifier-free choice** in all types

$$\text{QF-AC}: \ \forall \underline{a} \exists \underline{b}\, F_{qf}(\underline{a}, \underline{b}) \to \exists \underline{B} \forall \underline{a}\, F_{qf}(\underline{a}, \underline{B}(\underline{a})).$$

- $\underline{x}, \underline{y}$ are tuples of **functionals of finite type** over $\mathbb{N}, X_1, \ldots, X_n$ (X_i abstract metric or normed spaces).

Monotone Functional interpretation in five minutes

Gödel's **functional interpretation** G is a map ($\mathcal{A}$ formal system of analysis)

$$G : Formula(\mathcal{A}) \to Formula(\mathcal{A}),\ A \mapsto A^G$$

such that

- $A^G \equiv \forall \underline{x} \exists \underline{y}\, A_G(\underline{x}, \underline{y})$, where A_G is **quantifier-free**,
- For $A \equiv \forall \underline{x} \exists \underline{y}\, A_{qf}(\underline{x}, \underline{y})$ one has $A^G \equiv A$.
- $A \leftrightarrow A^G$ by **quantifier-free choice** in all types

$$\textbf{QF-AC}: \ \forall \underline{a} \exists \underline{b}\, F_{qf}(\underline{a}, \underline{b}) \to \exists \underline{B} \forall \underline{a}\, F_{qf}(\underline{a}, \underline{B}(\underline{a})).$$

- $\underline{x}, \underline{y}$ are tuples of **functionals of finite type** over $\mathbb{N}, X_1, \ldots, X_n$ (X_i abstract metric or normed spaces).
- Functional interpretations **extracts uniform bounds** on A^G from a given **proof** of A.

Proof Mining in core mathematics

- During (mainly) the last 20 years this proof-theoretic approach has resulted in **numerous new quantitative results** as well as **qualitative uniformity results** in particular in: nonlinear analysis, fixed point theory, ergodic theory, topological dynamics, approximation theory, convex optimization, abstract Cauchy problems, pursuit-evasion games (≥ 100 papers mostly in specialized journals in the resp. areas or general mathematics journals).

Proof Mining in core mathematics

- During (mainly) the last 20 years this proof-theoretic approach has resulted in **numerous new quantitative results** as well as **qualitative uniformity results** in particular in: nonlinear analysis, fixed point theory, ergodic theory, topological dynamics, approximation theory, convex optimization, abstract Cauchy problems, pursuit-evasion games ($\geq$ 100 papers mostly in specialized journals in the resp. areas or general mathematics journals).
- General **logical metatheorems** explain applications as instances of logical phenomena (K. 2005, Gerhardy/K. 2008, TAMS).

Proof Mining in core mathematics

- During (mainly) the last 20 years this proof-theoretic approach has resulted in **numerous new quantitative results** as well as **qualitative uniformity results** in particular in: nonlinear analysis, fixed point theory, ergodic theory, topological dynamics, approximation theory, convex optimization, abstract Cauchy problems, pursuit-evasion games ($\geq$ 100 papers mostly in specialized journals in the resp. areas or general mathematics journals).
- General **logical metatheorems** explain applications as instances of logical phenomena (K. 2005, Gerhardy/K. 2008, TAMS).
- Some of the logical tools used have been rediscovered in 2007 in special cases by Terence Tao prompted by concrete mathematical needs **"finitary analysis"!**

Goal: Extract effective bounds for $\forall \underline{x} \in \underline{X}\, \exists n \in \mathbb{N}\, A(\underline{x}, n)$-theorems.

Goal: Extract effective bounds for $\forall \underline{x} \in \underline{X}\, \exists n \in \mathbb{N}\, A(\underline{x}, n)$-theorems.

For proofs using **full classical logic** A must be **purely existential!**

Goal: Extract effective bounds for $\forall \underline{x} \in \underline{X} \, \exists n \in \mathbb{N} \, A(\underline{x}, n)$-theorems.

For proofs using **full classical logic** A must be **purely existential!**

For **semi-classical proofs**: A can be of **arbitrary complexity!**

Goal: Extract effective bounds for $\forall \underline{x} \in \underline{X}\, \exists n \in \mathbb{N}\, A(\underline{x}, n)$-theorems.

For proofs using **full classical logic** A must be **purely existential!**
For **semi-classical proofs**: A can be of **arbitrary complexity!**

General **logical metatheorems** guarantee the extractability of effective bounds on '$\exists$' that are **independent** from parameters x from

Goal: Extract effective bounds for $\forall \underline{x} \in \underline{X}\, \exists n \in \mathbb{N}\, A(\underline{x}, n)$-theorems.

For proofs using **full classical logic** A must be **purely existential!**
For **semi-classical proofs**: A can be of **arbitrary complexity!**

General **logical metatheorems** guarantee the extractability of effective bounds on '$\exists$' that are **independent** from parameters x from

- compact metric spaces K (if separability is used) (K.1993,1998) and

Goal: Extract effective bounds for $\forall \underline{x} \in \underline{X}\, \exists n \in \mathbb{N}\, A(\underline{x}, n)$-theorems.

For proofs using **full classical logic** A must be **purely existential!**

For **semi-classical proofs**: A can be of **arbitrary complexity!**

General **logical metatheorems** guarantee the extractability of effective bounds on '$\exists$' that are **independent** from parameters x from

- compact metric spaces K (if separability is used) (K.1993,1998) and
- metrically bounded subsets of abstract spaces X that are not assumed to be separable provided X belongs to a sufficiently uniformly axiomatizable class of spaces (K.2005, Gerhardy/K.2008).

Goal: Extract effective bounds for $\forall \underline{x} \in \underline{X} \, \exists n \in \mathbb{N} \, A(\underline{x}, n)$-theorems.

For proofs using **full classical logic** A must be **purely existential!**

For **semi-classical proofs**: A can be of **arbitrary complexity!**

General **logical metatheorems** guarantee the extractability of effective bounds on '$\exists$' that are **independent** from parameters x from

- compact metric spaces K (if separability is used) (K.1993,1998) and
- metrically bounded subsets of abstract spaces X that are not assumed to be separable provided X belongs to a sufficiently uniformly axiomatizable class of spaces (K.2005, Gerhardy/K.2008).

Examples of such spaces X : metric, normed, Hilbert, uniformly convex uniformly smooth, hyperbolic, CAT(0), abstract L^p and $C(K)$ spaces (but e.g. not separable, strictly convex or smooth spaces).

'Proof Mining': Applied Foundations

- **Translation of a given proof into a new proof of a new result based on intuitionistic logic** while still using **ideal elements**.

‘Proof Mining’: Applied Foundations

- **Translation of a given proof into a new proof of a new result based on intuitionistic logic** while still using **ideal elements**.
- **‘Finitistic’ analysis (relative to the ideal elements)** of the latter by exhibiting explicit constructions hidden in the use of quantifiers.

‘Proof Mining’: Applied Foundations

- **Translation of a given proof into a new proof of a new result based on intuitionistic logic** while still using **ideal elements**.
- **‘Finitistic’ analysis (relative to the ideal elements)** of the latter by exhibiting explicit constructions hidden in the use of quantifiers.
- Finitistic use of ‘ideal elements’ allows one to **replace them by finitary versions** (without the proof noticing the difference!).

‘Proof Mining’: Applied Foundations

- **Translation of a given proof into a new proof of a new result based on intuitionistic logic** while still using **ideal elements**.
- **‘Finitistic’ analysis (relative to the ideal elements)** of the latter by exhibiting explicit constructions hidden in the use of quantifiers.
- Finitistic use of ‘ideal elements’ allows one to **replace them by finitary versions** (without the proof noticing the difference!).
- This in the end gives some form of **finitistic realization of $\forall\exists$-theorems** proven together with an elementary proofs.

‘Proof Mining’: Applied Foundations

- **Translation of a given proof into a new proof of a new result based on intuitionistic logic** while still using **ideal elements**.
- **‘Finitistic’ analysis (relative to the ideal elements)** of the latter by exhibiting explicit constructions hidden in the use of quantifiers.
- Finitistic use of ‘ideal elements’ allows one to **replace them by finitary versions** (without the proof noticing the difference!).
- This in the end gives some form of **finitistic realization of $\forall\exists$-theorems** proven together with an elementary proofs.
- For **semi-intuitionistic proofs** the restriction to $\forall\exists$-theorem can be overcome: **arbitrary logical complexity allowed!**

'Proof Mining': Applied Foundations

- **Translation of a given proof into a new proof of a new result based on intuitionistic logic** while still using **ideal elements**.
- **'Finitistic' analysis (relative to the ideal elements)** of the latter by exhibiting explicit constructions hidden in the use of quantifiers.
- Finitistic use of 'ideal elements' allows one to **replace them by finitary versions** (without the proof noticing the difference!).
- This in the end gives some form of **finitistic realization of $\forall\exists$-theorems** proven together with an elementary proofs.
- For **semi-intuitionistic proofs** the restriction to $\forall\exists$-theorem can be overcome: **arbitrary logical complexity allowed!**
- **Final product: explicit bound** with **ordinary analytical proof** for its correctness (no reference to logic!).

The running theme: convergence statements in analysis

Let (x_n) be a Cauchy sequence in a metric space (X, d), i.e.

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

The running theme: convergence statements in analysis

Let (x_n) be a Cauchy sequence in a metric space (X, d), i.e.

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

is **noneffectively** equivalent to

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

The running theme: convergence statements in analysis

Let (x_n) be a Cauchy sequence in a metric space (X, d), i.e.

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

is **noneffectively** equivalent to

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

Kreisel's **no-counterexample interpretation** or **metastability** (T. Tao).

The running theme: convergence statements in analysis

Let (x_n) be a Cauchy sequence in a metric space (X, d), i.e.

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

is **noneffectively** equivalent to

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

Kreisel's **no-counterexample interpretation** or **metastability** (T. Tao).

A bound $\Phi(k, g)$ on '$\exists n$' in the latter formula is a **rate of metastability**.

Effective full rates of convergence?

- Usually **possible for asymptotic regularity** results

$$d(x_n, T(x_n)) \to 0,$$

even when (x_n) may not converge to a fixed point of T.

Effective full rates of convergence?

- Usually **possible for asymptotic regularity** results

$$d(x_n, T(x_n)) \to 0,$$

  even when (x_n) may not converge to a fixed point of T.

- Possible for (x_n) if sequence converges to **unique** fixed point.

Effective full rates of convergence?

- Usually **possible for asymptotic regularity** results

$$d(x_n, T(x_n)) \to 0,$$

even when (x_n) may not converge to a fixed point of T.

- Possible for (x_n) if sequence converges to **unique** fixed point. Extraction of **modulus of uniqueness** $\Phi : \mathbb{R}^*_+ \to \mathbb{R}^*_+$

$$\forall \varepsilon > 0\, \forall x, y \in X \left(d(x, T(x)), d(y, T(y)) < \Phi(\varepsilon) \to d(x, y) < \varepsilon \right)$$

gives rate of convergence (or – in the noncompact case – existence at all)! Numerous applications in analysis!

Effective full rates of convergence?

- Usually **possible for asymptotic regularity** results

$$d(x_n, T(x_n)) \to 0,$$

even when (x_n) may not converge to a fixed point of T.

- Possible for (x_n) if sequence converges to **unique** fixed point. Extraction of **modulus of uniqueness** $\Phi : \mathbb{R}^*_+ \to \mathbb{R}^*_+$

$$\forall \varepsilon > 0\, \forall x, y \in X\, \big(d(x, T(x)), d(y, T(y)) < \Phi(\varepsilon) \to d(x, y) < \varepsilon \big)$$

gives rate of convergence (or – in the noncompact case – existence at all)! Numerous applications in analysis!

- Possible also in the nonunique case for **Fejér monotone algorithms** if one has a **modulus of metric regularity** (see below).

Applications to Convex Optimization

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

Consider a 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':

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

Previously 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.

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

Consider a Hilbert space H and nonempty closed and convex subsets $\boldsymbol{C_1, \ldots, C_N \subseteq H}$ with metric projections $\boldsymbol{P_{C_i}}$, define $\boldsymbol{T := P_{C_N} \circ \ldots \circ P_{C_1}}$. In 2003 Bauschke proved the 'zero displacement conjecture':

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

Previously 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.

Proof uses abstract theory of maximal monotone operators: Minty's theorem, Brézis-Haraux theorem, Rockafellar's maximal monotonicity and sum theorems, Bruck-Reich theory of strongly nonexpansive mappings, conjugate functions, normal cone operator...).

$(\|T^{n+1}x - T^n x\|)_{n\in\mathbb{N}}$ is **nonincreasing** and hence the conclusion in Bauschke's theorem is of the form $\forall\exists$.

$(\|T^{n+1}x - T^n x\|)_{n\in\mathbb{N}}$ is **nonincreasing** and hence the conclusion in Bauschke's theorem is of the form $\forall\exists$.

Logical metatheorems, therefore, guarantee the extractability a uniform rate of asymptotic regularity which only depends on the error $\varepsilon > 0$, $N \in \mathbb{N}$ and **majorants** for $x \in H$ and $P_{C_1}, \ldots, P_{C_N}$:

$b \geq \|x\|$ and $K \geq \|c_1\|, \ldots, \|c_N\|$ for some **arbitrary** points $c_1 \in C_1, \ldots, c_N \in C_N$:

$$\|P_{C_i}0\| \leq \|c_i\| \leq K.$$

Since the mappings P_{C_i} are nonexpansive, the corollary guarantees a computable $\Phi(\varepsilon, N, b, K)$ s.t. for $b \geq \|x\|$

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

Theorem (K. FoCM 2019)

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

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

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

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

Here $b \geq \|x\|$ and $K \geq \left(\sum_{i=1}^{N} \|c_i\|^2\right)^{\frac{1}{2}}$ for some $(c_1, \ldots, c_N) \in C_1 \times \ldots \times C_N$.

Theorem (K. FoCM 2019)

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

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

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

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

Here $b \geq \|x\|$ and $K \geq \left(\sum_{i=1}^{N} \|c_i\|^2\right)^{\frac{1}{2}}$ for some $(c_1, \ldots, c_N) \in C_1 \times \ldots \times C_N$.

A. Sipoș (2020): Extension to general averaged mappings.

Proximal mappings in Banach spaces

In Bačák/K. J. Convex Analysis 2018, we give a quantitative treatment of the **proximal mapping in uniformly convex Banach spaces** $X, f : X \to (-\infty, \infty]$ proper lsc convex, $\lambda > 0, \Phi$ a Young function:

$$\mathrm{prox}^{\Phi}_{\lambda,f}(x) := \operatorname*{argmin}_{y \in X} \left[f(y) + \frac{1}{\phi(\lambda)} \Phi\left(\|x - y\| \right) \right], \quad x \in X.$$

Proximal mappings in Banach spaces

In Bačák/K. J. Convex Analysis 2018, we give a quantitative treatment of the **proximal mapping in uniformly convex Banach spaces** $X, f : X \to (-\infty, \infty]$ proper lsc convex, $\lambda > 0, \Phi$ a Young function:

$$\mathrm{prox}^{\Phi}_{\lambda,f}(x) := \underset{y\in X}{\mathrm{argmin}} \left[f(y) + \frac{1}{\phi(\lambda)} \Phi\left(\|x-y\|\right)\right], \quad x \in X.$$

Classically: X Hilbert space, $\Phi(t) := \frac{1}{2}t^2, \phi(t) = t$. Then

$$\mathrm{prox}_{\lambda,f}(x) := \underset{y\in X}{\mathrm{argmin}} \left[f(y) + \frac{1}{2\lambda}\|x-y\|^2\right] = (I+\lambda\partial f)^{-1}(x), \quad x \in X,$$

i.e. $\mathrm{prox}_{\lambda,f}$ is the resolvent of the maximally monotone subdifferential ∂f and hence the **metric projection** P_C for closed convex C if f is the **indicator function** of C.

Nonexpansive if X Hilbert space (this characterizes Hilbert spaces).

For the above generalization to arbitrary Young functions and in uniformly convex spaces.

Theorem (Bačák/K., J. Convex Anal. 2018)

For all $z \in X, r > 0, \varepsilon > 0, x, y \in B(z,r)$ and $\lambda \in (0,1]$:

$$\|x-y\| < \delta(\varepsilon) \implies \left\|\mathbf{prox}^{\Phi}_{\lambda,f}(x) - \mathbf{prox}^{\Phi}_{\lambda,f}(y)\right\| < \varepsilon, \text{ where}$$

$$\delta(\varepsilon) := \min\left\{\frac{\varepsilon}{2}, \frac{2}{\phi(R)}\delta_R\left(\frac{\varepsilon}{2}\right)\right\},$$

and $R > 0$ is a constant independent of $\lambda \in (0,1]$ and δ_R is a modulus of uniform convexity of $\Phi \circ \|\cdot\|$ on the ball $B(0,R)$.

Important **qualitative consequence**: modulus does not depend on λ. This is not true for an alternative definition of proximal mappings due to Penot and Ratsimahalo and supports our definition!

The proximal point algorithm

Let $(\lambda_n) \subset (0,1)$ and $A : X \to 2^X$ a set-valued (maximal) monotone operator (X Hilbert space). $\boldsymbol{J_A := (I + A)^{-1}}$ resolvent of $\boldsymbol{A}$.

The proximal point algorithm

Let $(\lambda_n) \subset (0,1)$ and $A : X \to 2^X$ a set-valued (maximal) monotone operator (X Hilbert space). $\boldsymbol{J_A := (I + A)^{-1}}$ resolvent of $\boldsymbol{A}$.

$$\boldsymbol{x_{n+1} := J_{\lambda_n A} x,}\quad \textbf{Proximal Point Algorithm (PPA).}$$

The proximal point algorithm

Let $(\lambda_n) \subset (0,1)$ and $A : X \to 2^X$ a set-valued (maximal) monotone operator (X Hilbert space). $\boldsymbol{J_A := (I+A)^{-1}}$ resolvent of $\boldsymbol{A}$.

$$x_{n+1} := J_{\lambda_n A} x, \quad \textbf{Proximal Point Algorithm (PPA).}$$

Under suitable conditions on $(\lambda_n) \subset (0,\infty)$: $\boldsymbol{(x_n)}$ **converges weakly to a zero** of $\boldsymbol{A}$ (Rockafellar 1976), but not strongly (Güler 1996).

The proximal point algorithm

Let $(\lambda_n) \subset (0,1)$ and $A : X \to 2^X$ a set-valued (maximal) monotone operator (X Hilbert space). $\boldsymbol{J_A := (I + A)^{-1}}$ resolvent of $\boldsymbol{A}$.

$$\boldsymbol{x_{n+1} := J_{\lambda_n A} x,} \quad \textbf{Proximal Point Algorithm (PPA).}$$

Under suitable conditions on $(\lambda_n) \subset (0, \infty)$: $\boldsymbol{(x_n)}$ **converges weakly to a zero** of $\boldsymbol{A}$ (Rockafellar 1976), but not strongly (Güler 1996). Generalizations to Banach spaces (then A accretive) or geodesic spaces.

The proximal point algorithm

Let $(\lambda_n) \subset (0,1)$ and $A : X \to 2^X$ a set-valued (maximal) monotone operator (X Hilbert space). $\boldsymbol{J_A := (I+A)^{-1}}$ resolvent of $\boldsymbol{A}$.

$$\boldsymbol{x_{n+1} := J_{\lambda_n A} x,} \quad \textbf{Proximal Point Algorithm (PPA).}$$

Under suitable conditions on $(\lambda_n) \subset (0,\infty)$: $\boldsymbol{(x_n)}$ **converges weakly to a zero** of $\boldsymbol{A}$ (Rockafellar 1976), but not strongly (Güler 1996).
Generalizations to Banach spaces (then A accretive) or geodesic spaces.
Rates of metastability in the finite dimensional/boundedly compact case:

The proximal point algorithm

Let $(\lambda_n) \subset (0,1)$ and $A : X \to 2^X$ a set-valued (maximal) monotone operator (X Hilbert space). $\boldsymbol{J_A := (I + A)^{-1}}$ resolvent of $\boldsymbol{A}$.

$$x_{n+1} := J_{\lambda_n A} x,$$ **Proximal Point Algorithm (PPA).**

Under suitable conditions on $(\lambda_n) \subset (0, \infty)$: $\boldsymbol{(x_n)}$ **converges weakly to a zero** of $\boldsymbol{A}$ (Rockafellar 1976), but not strongly (Güler 1996). Generalizations to Banach spaces (then A accretive) or geodesic spaces. Rates of metastability in the finite dimensional/boundedly compact case:

- Hilbert space: K./Leuștean/Nicolae. Comm.Contemp. Math. 2018.

The proximal point algorithm

Let $(\lambda_n) \subset (0,1)$ and $A : X \to 2^X$ a set-valued (maximal) monotone operator (X Hilbert space). $\boldsymbol{J_A := (I+A)^{-1}}$ resolvent of $\boldsymbol{A}$.

$$x_{n+1} := J_{\lambda_n A} x, \quad \textbf{Proximal Point Algorithm (PPA).}$$

Under suitable conditions on $(\lambda_n) \subset (0,\infty)$: $\boldsymbol{(x_n)}$ **converges weakly to a zero** of $\boldsymbol{A}$ (Rockafellar 1976), but not strongly (Güler 1996).
Generalizations to Banach spaces (then A accretive) or geodesic spaces.
Rates of metastability in the finite dimensional/boundedly compact case:

- Hilbert space: K./Leuştean/Nicolae. Comm.Contemp. Math. 2018.
- CAT(0) spaces: Leuştean/Sipoş J. Nonlin. Var. Anal. 2018.

The proximal point algorithm

Let $(\lambda_n) \subset (0,1)$ and $A : X \to 2^X$ a set-valued (maximal) monotone operator (X Hilbert space). $\boldsymbol{J_A := (I+A)^{-1}}$ resolvent of $\boldsymbol{A}$.

$$\boldsymbol{x_{n+1} := J_{\lambda_n A} x,} \quad \textbf{Proximal Point Algorithm (PPA).}$$

Under suitable conditions on $(\lambda_n) \subset (0,\infty)$: $\boldsymbol{(x_n)}$ **converges weakly to a zero** of $\boldsymbol{A}$ (Rockafellar 1976), but not strongly (Güler 1996).
Generalizations to Banach spaces (then A accretive) or geodesic spaces.
Rates of metastability in the finite dimensional/boundedly compact case:

- Hilbert space: K./Leuștean/Nicolae. Comm.Contemp. Math. 2018.
- CAT(0) spaces: Leuștean/Sipoș J. Nonlin. Var. Anal. 2018.
- abstract forms of PPA in Hilbert space: Leuștean/Nicolae/Sipoș J. Global Opt. 2018.

The proximal point algorithm

Let $(\lambda_n) \subset (0,1)$ and $A : X \to 2^X$ a set-valued (maximal) monotone operator (X Hilbert space). $\boldsymbol{J_A := (I + A)^{-1}}$ resolvent of $\boldsymbol{A}$.

$$\boldsymbol{x_{n+1} := J_{\lambda_n A} x,}$$ **Proximal Point Algorithm (PPA).**

Under suitable conditions on $(\lambda_n) \subset (0,\infty)$: $\boldsymbol{(x_n)}$ **converges weakly to a zero** of $\boldsymbol{A}$ (Rockafellar 1976), but not strongly (Güler 1996). Generalizations to Banach spaces (then A accretive) or geodesic spaces. Rates of metastability in the finite dimensional/boundedly compact case:

- Hilbert space: K./Leuștean/Nicolae. Comm.Contemp. Math. 2018.
- CAT(0) spaces: Leuștean/Sipoș J. Nonlin. Var. Anal. 2018.
- abstract forms of PPA in Hilbert space: Leuștean/Nicolae/Sipoș J. Global Opt. 2018.
- uniformly convex Banach spaces: K. J. Convex Anal. 2021.

Rates of convergence (without finite dimension/compactness) **possible if** A is **uniformly monotone (accretive)** in a suitable sense:

Rates of convergence (without finite dimension/compactness) **possible if** A is **uniformly monotone (accretive)** in a suitable sense:

- In Hilbert space: Leuștean/Nicolae/Sipoș J. Global Opt. 2018.

Rates of convergence (without finite dimension/compactness) **possible if** A is **uniformly monotone (accretive)** in a suitable sense:

- In Hilbert space: Leuștean/Nicolae/Sipoș J. Global Opt. 2018.
- In suitable Banach spaces and for many other related algorithms: K./Powell: Computers & Mathematics Appl. 2020.

Rates of convergence (without finite dimension/compactness) **possible if** A is **uniformly monotone (accretive)** in a suitable sense:

- In Hilbert space: Leuștean/Nicolae/Sipoș J. Global Opt. 2018.
- In suitable Banach spaces and for many other related algorithms: K./Powell: Computers & Mathematics Appl. 2020.

In general: **strong convergence** (even in infinite dimensional Hilbert spaces) **only for** so-called **Halpern type variant of PPA**:

$$x_{n+1} := \alpha_n u + (1-\alpha_n) J_{\lambda_n A} \quad \textbf{(HPPA)}$$

(necessary conditions: $\lim \alpha_n = 0, \sum \alpha_n = \infty$).

Rates of metastability of HPPA

Rates of metastability of HPPA

- In Hilbert space for $\lim \lambda_n = \lambda \in (0,1)$: Pinto (Thesis June 2019), Leuștean/Pinto Comput. Opt. Appl. 2021.

Rates of metastability of HPPA

- In Hilbert space for $\lim \lambda_n = \lambda \in (0, 1)$: Pinto (Thesis June 2019), Leuştean/Pinto Comput. Opt. Appl. 2021.
- In Hilbert space for $\lim \lambda_n \to \infty$: Pinto (Thesis June 2019), Numer. Funct. Anal. Optim. 2021.

Rates of metastability of HPPA

- In Hilbert space for $\lim \lambda_n = \lambda \in (0,1)$: Pinto (Thesis June 2019), Leuştean/Pinto Comput. Opt. Appl. 2021.
- In Hilbert space for $\lim \lambda_n \to \infty$: Pinto (Thesis June 2019), Numer. Funct. Anal. Optim. 2021.
- In Banach spaces which are uniformly convex and uniformly smooth and for $\inf \lambda_n \geq \lambda > 0$: K. J. Nonlin. Convex Anal. 2020.

Rates of metastability of HPPA

- In Hilbert space for $\lim \lambda_n = \lambda \in (0,1)$: Pinto (Thesis June 2019), Leuștean/Pinto Comput. Opt. Appl. 2021.
- In Hilbert space for $\lim \lambda_n \to \infty$: Pinto (Thesis June 2019), Numer. Funct. Anal. Optim. 2021.
- In Banach spaces which are uniformly convex and uniformly smooth and for $\inf \lambda_n \geq \lambda > 0$: K. J. Nonlin. Convex Anal. 2020.
- In Hilbert space and for $\inf \lambda_n \geq \lambda > 0$ and a general schema covering HPPA as special case: Dinis/Pinto J. Convex Anal. 2021.

Rates of metastability of HPPA

- In Hilbert space for $\lim \lambda_n = \lambda \in (0,1)$: Pinto (Thesis June 2019), Leuștean/Pinto Comput. Opt. Appl. 2021.
- In Hilbert space for $\lim \lambda_n \to \infty$: Pinto (Thesis June 2019), Numer. Funct. Anal. Optim. 2021.
- In Banach spaces which are uniformly convex and uniformly smooth and for $\inf \lambda_n \geq \lambda > 0$: K. J. Nonlin. Convex Anal. 2020.
- In Hilbert space and for $\inf \lambda_n \geq \lambda > 0$ and a general schema covering HPPA as special case: Dinis/Pinto J. Convex Anal. 2021.

The proofs and their resp. minings are very different!

Viscosity approximation method

Consider for bounded convex $C \subset X$, X normed or hyperbolic space, $T : C \to C$ nonexpansive and suitable $(\alpha_n) \subset (0,1), u \in C$ the Halpern iteration

$$w_0 \in C, \quad w_{n+1} = (1-\alpha_n)T(w_n) + \alpha_n u, \tag{1}$$

and its viscosity generalization for contractions ϕ

$$x_0 \in C, \; x_{n+1} = (1-\alpha_n)T(x_n) + \alpha_n \phi(x_n). \tag{2}$$

Viscosity approximation method

Consider for bounded convex $C \subset X$, X normed or hyperbolic space, $T : C \to C$ nonexpansive and suitable $(\alpha_n) \subset (0,1), u \in C$ the Halpern iteration

$$\boldsymbol{w_0 \in C, \quad w_{n+1} = (1-\alpha_n)T(w_n) + \alpha_n u,} \tag{1}$$

and its viscosity generalization for contractions ϕ

$$\boldsymbol{x_0 \in C, \; x_{n+1} = (1-\alpha_n)T(x_n) + \alpha_n \phi(x_n).} \tag{2}$$

Suzuki 2007 showed that if $(w_n(u))$ converges for all u then also (x_n) **and** $\lim x_n$ is the unique fixed point of $P \circ \phi$ for $P(x) := \lim w_n(x)$.

Viscosity approximation method

Consider for bounded convex $C \subset X$, X normed or hyperbolic space, $T : C \to C$ nonexpansive and suitable $(\alpha_n) \subset (0,1), u \in C$ the Halpern iteration

$$w_0 \in C, \quad w_{n+1} = (1-\alpha_n)T(w_n) + \alpha_n u, \tag{1}$$

and its viscosity generalization for contractions ϕ

$$x_0 \in C, \; x_{n+1} = (1-\alpha_n)T(x_n) + \alpha_n \phi(x_n). \tag{2}$$

Suzuki 2007 showed that if $(w_n(u))$ converges for all u then also (x_n) **and** $\lim x_n$ is the unique fixed point of $P \circ \phi$ for $P(x) := \lim w_n(x)$.
K./Pinto 2020 showed that a uniform rate of metastability θ for $(w_n(u)$ translates into $\Psi[\theta]$ with

$$\forall N \in \mathbb{N}\, \forall \varepsilon > 0\, \exists n \in [N, \Psi[\theta](\varepsilon, g, N)]\, \exists k \in [N, n], z \in C$$
$$\left(d(z, w_k(\Phi(z))) \leq \varepsilon \wedge \forall i \in [n, n+g(n)]\, (d(x_i, z) \leq \tfrac{\varepsilon}{2})\right).$$

Fejér monotonicity and regularity

Why are people in optimization not always using the strongly convergent HPPA instead of the weakly convergent PPA?

Fejér monotonicity and regularity

Why are people in optimization not always using the strongly convergent HPPA instead of the weakly convergent PPA?

HPPA - as all commonly used strongly convergent methods - is **not Fejér monotone** while PPA (and many other weakly convergent methods) is:

Fejér monotonicity and regularity

Why are people in optimization not always using the strongly convergent HPPA instead of the weakly convergent PPA?

HPPA - as all commonly used strongly convergent methods - is **not Fejér monotone** while PPA (and many other weakly convergent methods) is:

Definition

A sequence (x_n) in a metric space (X, d) is Fejér monotone w.r.t. a subset $S \subseteq X$ if $\forall n \in \mathbb{N}\, \forall p \in S\ (d(x_{n+1}, p) \leq d(x_n, p))$.

Fejér monotonicity and regularity

Why are people in optimization not always using the strongly convergent HPPA instead of the weakly convergent PPA?

HPPA - as all commonly used strongly convergent methods - is **not Fejér monotone** while PPA (and many other weakly convergent methods) is:

Definition

A sequence (x_n) in a metric space (X, d) is Fejér monotone w.r.t. a subset $S \subseteq X$ if $\forall n \in \mathbb{N}\, \forall p \in S\ (d(x_{n+1}, p) \leq d(x_n, p))$.

Why is this important?

Fejér monotonicity and regularity

Why are people in optimization not always using the strongly convergent HPPA instead of the weakly convergent PPA?

HPPA - as all commonly used strongly convergent methods - is **not Fejér monotone** while PPA (and many other weakly convergent methods) is:

Definition

A sequence (x_n) in a metric space (X, d) is Fejér monotone w.r.t. a subset $S \subseteq X$ if $\forall n \in \mathbb{N}\, \forall p \in S\ (d(x_{n+1}, p) \leq d(x_n, p))$.

Why is this important?

If one has metric regularity one not only gets strong convergence but even a **rate of convergence!**

Moduli of regularity for mappings

In continuous optimization notions of **linear** or **Hölder metric regularity**, **error bounds** and **weak sharp minima** etc. play an important role which can be viewed as (often local forms of) special cases of (see also R.M. Anderson: 'Almost' implies 'Near', TAMS 1986):

Moduli of regularity for mappings

In continuous optimization notions of **linear** or **Hölder metric regularity**, **error bounds** and **weak sharp minima** etc. play an important role which can be viewed as (often local forms of) special cases of (see also R.M. Anderson: 'Almost' implies 'Near', TAMS 1986):

Definition (K./Lopéz-Acedo/Nicolae, Israel J. Math 2019)

Let $zer\, F := \{x \in X : F(x) = 0\} \neq \emptyset$. F is **regular** w.r.t. $zer\, F$ if

$$\forall n \in \mathbb{N} \exists k \in \mathbb{N}\, \forall x \in X(|F(x)| < 2^{-k} \rightarrow \exists z' \in zer\, F(d(x, z') < 2^{-n})).$$

A function $\omega : \mathbb{N} \rightarrow \mathbb{N}$ providing $k = \omega(n)$ is a **modulus of regularity**.

Moduli of regularity for mappings

In continuous optimization notions of **linear** or **Hölder metric regularity**, **error bounds** and **weak sharp minima** etc. play an important role which can be viewed as (often local forms of) special cases of (see also R.M. Anderson: 'Almost' implies 'Near', TAMS 1986):

Definition (K./Lopéz-Acedo/Nicolae, Israel J. Math 2019)

Let $zer\,F := \{x \in X : F(x) = 0\} \neq \emptyset$. F is **regular** w.r.t. $zer\,F$ if

$$\forall n \in \mathbb{N}\, \exists k \in \mathbb{N}\, \forall x \in X \left(|F(x)| < 2^{-k} \to \exists z' \in zer\,F\, (d(x, z') < 2^{-n}) \right).$$

A function $\omega : \mathbb{N} \to \mathbb{N}$ providing $k = \omega(n)$ is a **modulus of regularity**.

This covers fixed point and equilibrium problems as well as zeros of set-valued operators.

Computational use of moduli of regularity

Proposition (K./Lopéz-Acedo/Nicolae Israel J. Math. 2019)

Let $F : X \to \mathbb{R}$ *be with* $zer\ F \neq \emptyset$ *and with modulus of metric regularity* ω. *Let* (x_n) *be a sequence in* X *and* $\psi : \mathbb{N} \to \mathbb{N}$ *be s.t.*

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

where (x_n) *is Fejér monotone w.r.t.* $zer\ F$. *Then* (x_n) *is Cauchy:*

$$\forall k \in \mathbb{N}\, \forall n, \tilde{n} \geq \Phi(k) := \psi(\omega(k+1))\ \left(d(x_n, x_{\tilde{n}}) < 2^{-k}\right)$$

and $$\forall k \in \mathbb{N}\, \forall n \geq \Phi(k)\ \left(dist(x_n, zer\ F) < 2^{-k}\right).$$

If X *is complete and* F *is continuous, then* $\lim x_n \in zer\ F$.

Noncomputability of moduli of metric regularity

Proposition

If F is continuous, X is compact and $zerF \neq \emptyset$, then F has a modulus of regularity.

Noncomputability of moduli of metric regularity

Proposition

If F is continuous, X is compact and $zerF \neq \emptyset$, then F has a modulus of regularity.

In general, there will be no computable moduli of metric regularity:

Proposition (K./López-Acedo/Nicolae Israel J. Math. 2019)

There exists a **computable firmly nonexpansive** mapping $T : [0,1] \to [0,1]$ which has **no computable modulus** of metric regularity ϕ w.r.t. $Fix(T)\,(= zer\,(I - T))$.

Noncomputability of moduli of metric regularity

Proposition

If F is continuous, X is compact and $zerF \neq \emptyset$, then F has a modulus of regularity.

In general, there will be no computable moduli of metric regularity:

Proposition (K./López-Acedo/Nicolae Israel J. Math. 2019)

There exists a **computable firmly nonexpansive** mapping $T : [0,1] \to [0,1]$ which has **no computable modulus** of metric regularity ϕ w.r.t. $Fix(T)\ (= zer\,(I - T))$.

In fact, the cases where one can compute such a modulus are rare. However there are important cases where this is true (connection to o-minimality: tame optimization, Ioffe, Lewis, Bolte, Daniilidis...!)

Applications to Pursuit-evasion games

Let (X,d) be a uniquely geodesic space, $D>0$. $L_0, M_0 \in A$ starting points of the lion L and the man M. After n-steps, M moves to any point M_n s.t. $d(M_n, M_{n+1}) \leq D$ and L moves via the geodesic $[L_n, M_n]$ s.t. $d(L_n, L_{n+1}){=}\min\{D, d(L_n, M_n)\}$.

Let (X, d) be a uniquely geodesic space, $D > 0$. $L_0, M_0 \in A$ starting points of the lion L and the man M. After n-steps, M moves to any point M_n s.t. $d(M_n, M_{n+1}) \leq D$ and L moves via the geodesic $[L_n, M_n]$ s.t. $d(L_n, L_{n+1})=\min\{D, d(L_n, M_n)\}$.

Lopéz-Acedo/Nicolae/Piatek, Geom.Dedicata 2019: if X is a compact uniquely geodesic space with the betweenness property, then **the lion wins** i.e. $\lim d(L_{n+1}, M_n) = 0$ (proof makes iterated use of sequential compactness).

Let (X, d) be a uniquely geodesic space, $D > 0$. $L_0, M_0 \in A$ starting points of the lion L and the man M. After n-steps, M moves to any point M_n s.t. $d(M_n, M_{n+1}) \leq D$ and L moves via the geodesic $[L_n, M_n]$ s.t. $d(L_n, L_{n+1}) = \min\{D, d(L_n, M_n)\}$.

Lopéz-Acedo/Nicolae/Piatek, Geom.Dedicata 2019: if X is a compact uniquely geodesic space with the betweenness property, then **the lion wins** i.e. $\boldsymbol{\lim d(L_{n+1}, M_n) = 0}$ (proof makes iterated use of sequential compactness).

'$\boldsymbol{\lim d(L_{n+1}, M_n) = 0}$' $\in \forall\exists$ since the sequence is nonincreasing!

- **Proof Mining extracts an explicit rate of convergence** if one upgrades 'uniquely geodesic' and 'betweenness property' to **'uniform uniquely geodesic (with modulus)'** and **'uniform betweenness property (with modulus Θ)'**.

- **Proof Mining extracts an explicit rate of convergence** if one upgrades 'uniquely geodesic' and 'betweenness property' to **'uniform uniquely geodesic (with modulus)'** and **'uniform betweenness property (with modulus Θ)'**.
- With these upgrades the assumption of **compactness can be replaced by boundedness** in this particular case!

- **Proof Mining extracts an explicit rate of convergence** if one upgrades ‘uniquely geodesic’ and ‘betweenness property’ to **‘uniform uniquely geodesic (with modulus)’** and **‘uniform betweenness property (with modulus Θ)’**.
- With these upgrades the assumption of **compactness can be replaced by boundedness** in this particular case!
- Even the **uniqueness of geodesics can be dropped**.

- **Proof Mining extracts an explicit rate of convergence** if one upgrades 'uniquely geodesic' and 'betweenness property' to **'uniform uniquely geodesic (with modulus)'** and **'uniform betweenness property (with modulus Θ)'**.
- With these upgrades the assumption of **compactness can be replaced by boundedness** in this particular case!
- Even the **uniqueness of geodesics can be dropped**.
- Proof mining provides an explicit **rate of convergence** which only depends on Θ (in addition to $b \geq \text{diam}(A), D, \varepsilon > 0$).

- **Proof Mining extracts an explicit rate of convergence** if one upgrades 'uniquely geodesic' and 'betweenness property' to **'uniform uniquely geodesic (with modulus)'** and **'uniform betweenness property (with modulus Θ)'**.
- With these upgrades the assumption of **compactness can be replaced by boundedness** in this particular case!
- Even the **uniqueness of geodesics can be dropped**.
- Proof mining provides an explicit **rate of convergence** which only depends on Θ (in addition to $b \geq \text{diam}(A), D, \varepsilon > 0$).
- **Moduli of uniform betweenness** can be **extracted** from **proofs of mere betweenness** for the admissible structures.

- **Proof Mining extracts an explicit rate of convergence** if one upgrades 'uniquely geodesic' and 'betweenness property' to **'uniform uniquely geodesic (with modulus)'** and **'uniform betweenness property (with modulus Θ)'**.
- With these upgrades the assumption of **compactness can be replaced by boundedness** in this particular case!
- Even the **uniqueness of geodesics can be dropped**.
- Proof mining provides an explicit **rate of convergence** which only depends on Θ (in addition to $b \geq \text{diam}(A), D, \varepsilon > 0$).
- **Moduli of uniform betweenness** can be **extracted** from **proofs of mere betweenness** for the admissible structures.

Details in K./López-Acedo/Nicolae in Pacific J. Math. 2021.

Betweenness and uniform betweenness in metric spaces

Definition (Diminnie and White 1981)

Let (X, d) be a metric space. X satisfies the betweenness property if for any distinct points $x, y, z, w \in X$

$$\left.\begin{array}{l} d(x,y) + d(y,z) \leq d(x,z) \\ d(y,z) + d(z,w) \leq d(y,w) \end{array}\right\} \Rightarrow d(x,z) + d(z,w) \leq d(x,w).$$

Betweenness and uniform betweenness in metric spaces

Definition (Diminnie and White 1981)

Let (X, d) be a metric space. X satisfies the betweenness property if for any distinct points $x, y, z, w \in X$

$$\left. \begin{array}{l} d(x,y) + d(y,z) \le d(x,z) \\ d(y,z) + d(z,w) \le d(y,w) \end{array} \right\} \Rightarrow \ d(x,z) + d(z,w) \le d(x,w).$$

For normed spaces, betweenness follows from (but is strictly weaker than) strict convexity. It fails for $(\mathbb{R}^2, \|\cdot\|_\infty), (\mathbb{R}^2, \|\cdot\|_1)$ but holds for some nonstrictly convex spaces.

The functional interpretation upgrades betweenness to (equivalent in the compact case!):

Definition (K., Lopéz-Acedo, Nicolae, to appear in: Pacific J. Math.)

A metric space (X, d) satisfies the uniform betweenness property with modulus $\Theta : (0, \infty)^3 \to (0, \infty)$ if

$$\forall \varepsilon, a, b > 0\, \forall x, y, z, w \in X \left(\left\{ \begin{array}{l} \mathrm{sep}\{x, y, z, w\} \geq a \wedge \mathrm{diam}\{x, y, z, w\} \leq b \\ d(x, y) + d(y, z) \leq d(x, z) + \Theta(\varepsilon, a, b) \\ d(y, z) + d(z, w) \leq d(y, w) + \Theta(\varepsilon, a, b) \end{array} \right\} \begin{array}{c} \\ \\ \\ \Rightarrow d(x, z) + d(z, w) \leq d(x, w) + \varepsilon \end{array} \right).$$

Theorem (K./Lopéz-Acedo/Nicolae, to appear in: Pacific J. Math.)

Let X be a bounded metric space with the uniform betweenness property and $\langle (M_n), (L_n) \rangle$ be a Lion-Man game, speed $D > 0$.
Then the Lion approaches the man arbitrarily close.

Theorem (K./Lopéz-Acedo/Nicolae, to appear in: Pacific J. Math.)

Let X be a bounded metric space with the uniform betweenness property and $\langle (M_n),(L_n)\rangle$ be a Lion-Man game, speed $D>0$.
Then the Lion approaches the man arbitrarily close.

Moreover with

$b \geq \mathrm{diam}(X)$, $\Theta(\varepsilon) := \Theta(\varepsilon,\varepsilon,b)$, $N\in\mathbb{N}$, $b+1<ND$:

$$\forall \varepsilon>0\,\forall n\geq \Omega_{D,b,\Theta}(\varepsilon)\ (d(L_{n+1},M_n)<\varepsilon),$$

where

$$\Omega_{D,b,\Theta}(\varepsilon) = N + N\left\lceil \frac{b}{\Theta^{(N)}(\alpha)}\right\rceil$$

with

$$0<\alpha\leq \min\left\{\frac{1}{N},\frac{D}{2},\frac{\varepsilon}{2}\right\}.$$

Moduli of uniform betweennes

Θ can be **explicitly computed** for L^p $(1 < p < \infty)$ (of order 2 if $1 < p < 2$ and of order p if $2 \leq p < \infty$) and **CAT(κ)-spaces, $\kappa > 0$** (of order 2).

Moduli of uniform betweennes

Θ can be **explicitly computed** for L^p $(1 < p < \infty)$ (of order 2 if $1 < p < 2$ and of order p if $2 \le p < \infty$) and **CAT(κ)-spaces**, $\kappa > 0$ (of order 2).

Low complexity Θ's can also be obtained in a number of **non-uniquely geodesic** normed and metric cases!

Applications to Ergodic Theory

The Mean Ergodic Theorem

X **Hilbert space**, $f : X \to X$ **linear** and $\|f(z)\| \leq \|z\|$ for all $z \in X$.

$$\mathbf{A_n(z) := \frac{1}{n+1} S_n(z),\ \text{where } S_n(z) := \sum_{i=0}^{n} f^{(i)}(z) \quad (n \geq 0)}$$

The Mean Ergodic Theorem

X **Hilbert space**, $f : X \to X$ **linear** and $\|f(z)\| \leq \|z\|$ for all $z \in X$.

$$\mathbf{A_n(z) := \frac{1}{n+1} S_n(z),} \text{ where } \mathbf{S_n(z) := \sum_{i=0}^{n} f^{(i)}(z) \quad (n \geq 0)}$$

Theorem (von Neumann Mean Ergodic Theorem)

For every $z \in X$, the sequence $(A_n(z))_n$ converges.

The Mean Ergodic Theorem

X **Hilbert space**, $f: X \to X$ **linear** and $\|f(z)\| \leq \|z\|$ for all $z \in X$.

$$A_n(z) := \frac{1}{n+1} S_n(z), \text{ where } S_n(z) := \sum_{i=0}^{n} f^{(i)}(z) \quad (n \geq 0)$$

Theorem (von Neumann Mean Ergodic Theorem)

For every $z \in X$, the sequence $(A_n(z))_n$ converges.

Avigad/Gerhardy/Towsner (TAMS 2010):

in general **no computable rate of convergence**.

But: **Explicit bound on metastable version**

(extracted by logical metatheorem)!

The Mean Ergodic Theorem

X **Hilbert space**, $f : X \to X$ **linear** and $\|f(z)\| \le \|z\|$ for all $z \in X$.

$$\mathbf{A_n(z) := \frac{1}{n+1} S_n(z),\ \text{where } S_n(z) := \sum_{i=0}^{n} f^{(i)}(z) \quad (n \ge 0)}$$

Theorem (von Neumann Mean Ergodic Theorem)

For every $z \in X$, the sequence $(A_n(z))_n$ converges.

Avigad/Gerhardy/Towsner (TAMS 2010):

in general **no computable rate of convergence**.

But: **Explicit bound on metastable version**

(extracted by logical metatheorem)!

Theorem (Garrett Birkhoff 1939)

Mean Ergodic Theorem holds for uniformly convex Banach spaces.

Since the proof formalizes in $\mathcal{A}^\omega[X, \|\cdot\|, \eta]$ the following is guaranteed:

Since the proof formalizes in $\mathcal{A}^\omega[X, \|\cdot\|, \eta]$ the following is guaranteed:

One can extract an effective functional Φ in ε, g, b, η s.t. whenever X is a uniformly convex normed space with η being a modulus of uniform convexity and $f : X \to X$ is a nonexpansive linear operator. Then for all $z \in X$ with $\|z\| \leq b$, all $\varepsilon > 0$, all $g : \mathbb{N} \to \mathbb{N}$:

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

Since the proof formalizes in $\mathcal{A}^\omega[X, \|\cdot\|, \eta]$ the following is guaranteed:

One can extract an effective functional Φ in ε, g, b, η s.t. whenever X is a uniformly convex normed space with η being a modulus of uniform convexity and $f : X \to X$ is a nonexpansive linear operator. Then for all $z \in X$ with $\|z\| \le b$, all $\varepsilon > 0$, all $g : \mathbb{N} \to \mathbb{N}$:

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

$f^* := I$ majorizes f: $f(0) = 0$.

Theorem (K./Leuştean, Ergodic Theor. Dynam. Syst. 2009)

X uniformly convex Banach space, η a modulus of uniform convexity and $f : X \to X$ as above, $b > 0$.

Then for all $z \in X$ with $\|z\| \le b$, all $\varepsilon > 0$, all $g : \mathbb{N} \to \mathbb{N}$:

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

Theorem (K./Leuştean, Ergodic Theor. Dynam. Syst. 2009)

X uniformly convex Banach space, η a modulus of uniform convexity and $f : X \to X$ as above, $b > 0$.

Then for all $z \in X$ with $\|z\| \leq b$, all $\varepsilon > 0$, all $g : \mathbb{N} \to \mathbb{N}$:

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

where

$$\Phi(\varepsilon, g, b, \eta) := M \cdot \tilde{h}^{(K)}(0), \text{ with}$$
$$M := \left\lceil \frac{16b}{\varepsilon} \right\rceil, \gamma := \frac{\varepsilon}{16}\eta\left(\frac{\varepsilon}{8b}\right), \quad K := \left\lceil \frac{b}{\gamma} \right\rceil,$$
$$h, \tilde{h} : \mathbb{N} \to \mathbb{N},\ h(n) := 2(Mn + g(Mn)), \quad \tilde{h}(n) := \max_{i \leq n} h(i).$$

Recently generalized by Sipoş to multi-parameter version.

Bounding the number of fluctuations

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

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 our analysis of Birkhoff's proof, Avigad and Rute showed

Theorem (Avigad/Rute, Ergodic Theor. Dynam. Syst. 2015)

$(A_n(z))$ admits at most

$$2\log(M) \cdot \frac{b}{\varepsilon} + \frac{b}{\gamma} \cdot (2\log(2M) \cdot \frac{b}{\varepsilon} + \frac{b}{\gamma}$$

many fluctuations.

Bounding the number of fluctuations

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

Using our analysis of Birkhoff's proof, Avigad and Rute showed

Theorem (Avigad/Rute, Ergodic Theor. Dynam. Syst. 2015)

$(A_n(z))$ admits at most

$$2\log(M) \cdot \frac{b}{\varepsilon} + \frac{b}{\gamma} \cdot (2\log(2M) \cdot \frac{b}{\varepsilon} + \frac{b}{\gamma}$$

many fluctuations.

For the Hilbert space case: first fluctuation bounds by Jones, Ostrovskii, Rosenblatt 1996.

Other applications of proof mining in analysis

- Abstract Cauchy problems: Koutsoukou-Argyrakis/K. JMAA 2015.
- Best approximation theory: Chebycheff (K. NFAO 1993), L^1 (K./Oliva APAL 2003), Chebycheff for polynomials with restricted coefficients (Sipoș Math. Nach. 2021).
- Tauberian theorems (Powell MLQ 2020).
- Financial mathematics (Oliva 2020)

Philosophical conclusion from 'proof mining' enterprise

Philosophical conclusion from 'proof mining' enterprise

- Rather than having 'collapsed' (S. Kripke), **Hilbert's program** (in the broad sense) on the possibility of elimination of ideal elements from proofs of real statements has largely been **vindicated** for ordinary mathematical practice.

Philosophical conclusion from ‘proof mining’ enterprise

- Rather than having ‘collapsed’ (S. Kripke), **Hilbert’s program** (in the broad sense) on the possibility of elimination of ideal elements from proofs of real statements has largely been **vindicated** for ordinary mathematical practice.
- To a large extent this is even true for the narrow interpretation of the program, i.e. the reduction to strict finitism: **primitive recursive arithmetic**:

Philosophical conclusion from ‘proof mining’ enterprise

- Rather than having ‘collapsed’ (S. Kripke), **Hilbert’s program** (in the broad sense) on the possibility of elimination of ideal elements from proofs of real statements has largely been **vindicated** for ordinary mathematical practice.
- To a large extent this is even true for the narrow interpretation of the program, i.e. the reduction to strict finitism: **primitive recursive arithmetic**: **All bounds in this talk were primitive recursive.**

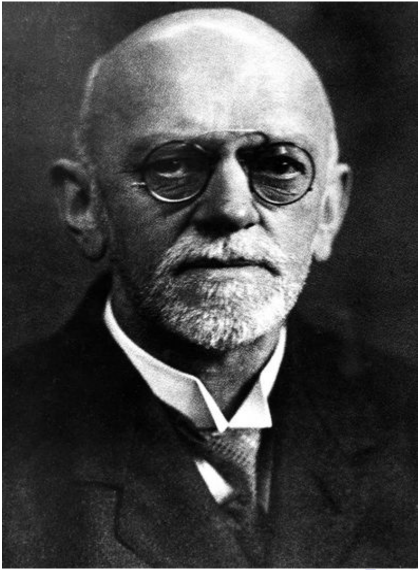
Philosophical conclusion from 'proof mining' enterprise

- Rather than having 'collapsed' (S. Kripke), **Hilbert's program** (in the broad sense) on the possibility of elimination of ideal elements from proofs of real statements has largely been **vindicated** for ordinary mathematical practice.
- To a large extent this is even true for the narrow interpretation of the program, i.e. the reduction to strict finitism: **primitive recursive arithmetic**: **All bounds in this talk were primitive recursive.**
- **Striking exceptions exist** (H. Friedman) but are very rare.

Philosophical conclusion from 'proof mining' enterprise

- Rather than having 'collapsed' (S. Kripke), **Hilbert's program** (in the broad sense) on the possibility of elimination of ideal elements from proofs of real statements has largely been **vindicated** for ordinary mathematical practice.
- To a large extent this is even true for the narrow interpretation of the program, i.e. the reduction to strict finitism: **primitive recursive arithmetic**: **All bounds in this talk were primitive recursive.**
- **Striking exceptions exist** (H. Friedman) but are very rare. **Why?**

David Hilbert (1862-1943)



Surveys

- Kohlenbach, U., Recent progress in proof mining in nonlinear analysis. IFCoLog Journal of Logics and their Applications **10**, Issue 4, pp. 3357-3406 (2017).
- Kohlenbach, U., Proof-theoretic Methods in Nonlinear Analysis. In: Proc. ICM 2018, B. Sirakov, P. Ney de Souza, M. Viana (eds.), Vol. 2, pp. 61-82. World Scientific 2019.