Applications of Proof Theory to Core Mathematics: Recent Developments

Ulrich Kohlenbach

Department of Mathematics

TECHNISCHE UNIVERSITÄT DARMSTADT

Logical Perspectives 2021, Moscow, June 10, 2021

Background: David Hilbert's Program

Background: David Hilbert's Program

Since the 19th century **noneffective (set-theoretic) principles** became increasingly important.

Background: David Hilbert's Program

Since the 19th century **noneffective (set-theoretic) principles** became increasingly important.

The issue of their legitimacy led Hilbert to the program:

Background: David Hilbert's Program

Since the 19th century **noneffective (set-theoretic) principles** became increasingly important.

The issue of their legitimacy led Hilbert to the program:

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

Background: David Hilbert's Program

Since the 19th century **noneffective (set-theoretic) principles** became increasingly important.

The issue of their legitimacy led Hilbert to the program:

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

In particular: Show the consistency of $\mathcal{I}$ by finitistic means.

Background: David Hilbert's Program

Since the 19th century **noneffective (set-theoretic) principles** became increasingly important.

The issue of their legitimacy led Hilbert to the program:

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

In particular: Show the consistency of $\mathcal{I}$ by finitistic means.

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!

General malaise of consistency proofs:

General malaise of consistency proofs:

'To one who has faith, no explanation is necessary. To one without faith, no explanation is possible' (after St Thomas Aquinas).

General malaise of consistency proofs:

'To one who has faith, no explanation is necessary. To one without faith, no explanation is possible' (after St Thomas Aquinas).

Shift of emphasis (G. Kreisel $\geq$ 1951): use proof theory to **extract new information** from proofs of **existential** statements.

General malaise of consistency proofs:

'To one who has faith, no explanation is necessary. To one without faith, no explanation is possible' (after St Thomas Aquinas).

Shift of emphasis (G. Kreisel $\geq$ 1951): use proof theory to **extract new information** from proofs of **existential** statements.

'What more do we know if we have proved a theorem by restricted means than if we merely know that it is true?' (G. Kreisel)

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).

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).
- 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 (≥ 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"!**

Logical Metatheorems based on Functional Interpretations

Let P be a **concrete** Polish space (e.g. $\mathbb{R}$ or $C[0,1]$) and K be a concrete compact metric space (e.g. $[0,1]^d$), X some **abstract** metric structure.

Logical Metatheorems based on Functional Interpretations

Let P be a **concrete** Polish space (e.g. $\mathbb{R}$ or $C[0,1]$) and K be a concrete compact metric space (e.g. $[0,1]^d$), X some **abstract** metric structure.

Goal: Effective bounds from noneffective proofs of $\forall \underline{x} \in P, K, X, X^X, X^{\mathbb{N}} \ldots \exists n \in \mathbb{N}\, A(\underline{x}, n)$-theorems.

Logical Metatheorems based on Functional Interpretations

Let P be a **concrete** Polish space (e.g. $\mathbb{R}$ or $C[0,1]$) and K be a concrete compact metric space (e.g. $[0,1]^d$), X some **abstract** metric structure.

Goal: Effective bounds from noneffective proofs of

$\forall \underline{x} \in P, K, X, X^X, X^{\mathbb{N}} \ldots \exists n \in \mathbb{N}\, A(\underline{x}, n)$-theorems.

Restriction: Because of classical logic: in general A **existential**.

Logical Metatheorems based on Functional Interpretations

Let P be a **concrete** Polish space (e.g. $\mathbb{R}$ or $C[0,1]$) and K be a concrete compact metric space (e.g. $[0,1]^d$), X some **abstract** metric structure.

Goal: Effective bounds from noneffective proofs of $\forall \underline{x} \in P, K, X, X^X, X^{\mathbb{N}} \ldots \exists n \in \mathbb{N}\, A(\underline{x}, n)$-theorems.

Restriction: Because of classical logic: in general A **existential**.

If A is existential, then general **logical metatheorems** (K. TAMS 2005, Gerhardy/K. TAMS 2008) guarantee the extractability of effective bounds on '$\exists$' that are **independent** from parameters $\underline{x}$ from

Logical Metatheorems based on Functional Interpretations

Let P be a **concrete** Polish space (e.g. $\mathbb{R}$ or $C[0,1]$) and K be a concrete compact metric space (e.g. $[0,1]^d$), X some **abstract** metric structure.

Goal: Effective bounds from noneffective proofs of
$\forall \underline{x} \in P, K, X, X^X, X^{\mathbb{N}} \ldots \exists n \in \mathbb{N}\, A(\underline{x}, n)$-theorems.

Restriction: Because of classical logic: in general A **existential**.

If A is existential, then general **logical metatheorems** (K. TAMS 2005, Gerhardy/K. TAMS 2008) guarantee the extractability of effective bounds on '$\exists$' that are **independent** from parameters $\underline{x}$ from

- **compact** metric spaces K (**if separability is used**),

Logical Metatheorems based on Functional Interpretations

Let P be a **concrete** Polish space (e.g. $\mathbb{R}$ or $C[0,1]$) and K be a concrete compact metric space (e.g. $[0,1]^d$), X some **abstract** metric structure.

Goal: Effective bounds from noneffective proofs of
$\forall \underline{x} \in P, K, X, X^X, X^{\mathbb{N}} \ldots \exists n \in \mathbb{N}\, A(\underline{x}, n)$-theorems.

Restriction: Because of classical logic: in general A **existential**.

If A is existential, then general **logical metatheorems** (K. TAMS 2005, Gerhardy/K. TAMS 2008) guarantee the extractability of effective bounds on '$\exists$' that are **independent** from parameters $\underline{x}$ from

- **compact** metric spaces K (**if separability is used**),
- metrically bounded subsets of suitable **abstract** metric structures X,

Logical Metatheorems based on Functional Interpretations

Let P be a **concrete** Polish space (e.g. $\mathbb{R}$ or $C[0,1]$) and K be a concrete compact metric space (e.g. $[0,1]^d$), X some **abstract** metric structure.

Goal: Effective bounds from noneffective proofs of
$\forall \underline{x} \in P, K, X, X^X, X^{\mathbb{N}} \ldots \exists n \in \mathbb{N}\, A(\underline{x}, n)$-theorems.

Restriction: Because of classical logic: in general A **existential**.

If A is existential, then general **logical metatheorems** (K. TAMS 2005, Gerhardy/K. TAMS 2008) guarantee the extractability of effective bounds on '$\exists$' that are **independent** from parameters $\underline{x}$ from

- **compact** metric spaces K (**if separability is used**),
- metrically bounded subsets of suitable **abstract** metric structures X,
- only depend on **majorizing data** in the case if unbounded subsets.

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

$$\boldsymbol{d(x_n, T(x_n)) \to 0,}$$

  even when (x_n) may not converge to a fixed point of $\boldsymbol{T}$.
- 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** $\Phi : \mathbb{R}^*_+ \to \mathbb{R}^*_+$

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

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

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

Applications to Pursuit-evasion games

Pursuit-Evasion Games: The Lion and Man Game

The lion and man problem, going back to R. Rado, is one of the most challenging pursuit-evasion games. **Littlewood's Miscellany** it is described as follows:

> *A lion and a man in a closed circular arena have equal maximum speeds. What tactics should the lion employ to be sure of his meal?*

Pursuit-Evasion Games: The Lion and Man Game

The lion and man problem, going back to R. Rado, is one of the most challenging pursuit-evasion games. **Littlewood's Miscellany** it is described as follows:

> *A lion and a man in a closed circular arena have equal maximum speeds. What tactics should the lion employ to be sure of his meal?*

Very similar problems have appeared under different names in the literature (e.g. the robot and the rabbit or the cop and the robber).

Pursuit-Evasion Games: The Lion and Man Game

The lion and man problem, going back to R. Rado, is one of the most challenging pursuit-evasion games. **Littlewood's Miscellany** it is described as follows:

> *A lion and a man in a closed circular arena have equal maximum speeds. What tactics should the lion employ to be sure of his meal?*

Very similar problems have appeared under different names in the literature (e.g. the robot and the rabbit or the cop and the robber).

The analysis of the game is closely tied to the geometric structure of the domain where the game is played.

Pursuit-Evasion Games: The Lion and Man Game

The lion and man problem, going back to R. Rado, is one of the most challenging pursuit-evasion games. **Littlewood's Miscellany** it is described as follows:

> *A lion and a man in a closed circular arena have equal maximum speeds. What tactics should the lion employ to be sure of his meal?*

Very similar problems have appeared under different names in the literature (e.g. the robot and the rabbit or the cop and the robber).

The analysis of the game is closely tied to the geometric structure of the domain where the game is played.

This fact, as well as the potential applications in different fields such as robotics, biology and random processes.

We focus on a **discrete-time** equal-speed game and **ε-capture**.

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. $d(M_n, M_{n+1}) \le D$ and L moves via the geodesic $[L_n, M_n]$ s.t. $d(L_n, L_{n+1}) = \min\{D, d(L_n, M_n)\}$.

López-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).

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. $d(L_n, L_{n+1})=\min\{D, d(L_n, M_n)\}$.

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

'$\mathbf{\lim d(L_{n+1}, M_n) = 0}$' $\in \forall\exists$ 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 $b \geq \mathrm{diam}(A), D, \varepsilon > 0$).

- **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 $b \geq \mathrm{diam}(A), D, \varepsilon > 0$).
- **Moduli of uniform betweenness** can be **extracted** from **proofs of mere betweenness** for the admissible structures.

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

Applications to the Proximal Point Algorithm

Proximal mappings in Hilbert space

Let H be a real Hilbert space. $f : H \to (-\infty, \infty]$ proper lsc convex.
The **proximal mapping** $\mathrm{prox}_f : H \to H$ is defined (for $\lambda > 0$) by

$$\mathrm{prox}_f(x) := \operatorname*{argmin}_{y \in H} \left[f(y) + \frac{1}{2}\|x - y\|^2 \right].$$

Proximal mappings in Hilbert space

Let H be a real Hilbert space. $f : H \to (-\infty, \infty]$ proper lsc convex. The **proximal mapping** $\mathrm{prox}_f : H \to H$ is defined (for $\lambda > 0$) by

$$\mathrm{prox}_f(x) := \operatorname*{argmin}_{y \in H} \left[f(y) + \frac{1}{2}\|x - y\|^2 \right].$$

Fact: $Fix(\mathrm{prox}_f) = \mathrm{argmin}\, f.$

Proximal mappings in Hilbert space

Let H be a real Hilbert space. $f : H \to (-\infty, \infty]$ proper lsc convex. The **proximal mapping** $\mathrm{prox}_f : H \to H$ is defined (for $\lambda > 0$) by

$$\mathrm{prox}_f(x) := \operatorname*{argmin}_{y \in H} \left[f(y) + \frac{1}{2} \|x - y\|^2 \right].$$

Fact: $Fix(\mathrm{prox}_f) = \mathrm{argmin}\, f$.

Example: Let $C \subseteq H$ be nonempty, closed and convex and

$$\iota_C : H \to [0, \infty],\ x \mapsto \begin{cases} 0, & \text{if } x \in C \\ \infty, & \text{otherwise.} \end{cases}$$

its **indicator function**, then prox_{ι_C} is the metric projection onto C.

Monotone operators

A set-valued mapping $A \subseteq H \to 2^H$ is **monotone** if

$$\forall (x,u),(y,v) \in gr(A)\ (\langle x-y, u-v\rangle \geq 0)\,.$$

Monotone operators

A set-valued mapping $A \subseteq H \to 2^H$ is **monotone** if

$$\forall (x,u),(y,v) \in gr(A)\ (\langle x-y, u-v\rangle \geq 0).$$

If A is monotone then the **resolvent**

$$J_A : R(I+A) \to D(A),\ x \mapsto (I+A)^{-1}(x)$$

is single-valued and **firmly nonexpansive**, i.e. for
$T := J_A, D := R(I+A)$

$$\forall x,y \in D\ \left(\|Tx-Ty\|^2 + \|(I-T)x-(I-T)y\|^2 \leq \|x-y\|^2\right).$$

Monotone operators

A set-valued mapping $A \subseteq H \to 2^H$ is **monotone** if

$$\forall (x, u), (y, v) \in gr(A) \ (\langle x - y, u - v \rangle \geq 0) \, .$$

If A is monotone then the **resolvent**

$$J_A : R(I + A) \to D(A), \ x \mapsto (I + A)^{-1}(x)$$

is single-valued and **firmly nonexpansive**, i.e. for $T := J_A, D := R(I + A)$

$$\forall x, y \in D \ \left(\| Tx - Ty \|^2 + \| (I - T)x - (I - T)y \|^2 \leq \| x - y \|^2 \right) .$$

A is **maximally monotone** if is has no proper monotone extension. In this case $R(I + A) = H$.

The Proximal Point Algorithm I

Example: Let f be as before. Then the **subdifferential** of f

$$\partial f : H \to 2^H : \ x \mapsto \{u \in H : \forall y \in H\,(\langle y - x, u\rangle + f(x) \leq f(y))\}$$

is a maximally monotone operator.

The Proximal Point Algorithm I

Example: Let f be as before. Then the **subdifferential** of f

$$\partial f : H \to 2^H : \; x \mapsto \{u \in H : \forall y \in H \, (\langle y - x, u\rangle + f(x) \leq f(y)\}$$

is a maximally monotone operator.

Facts: If A is monotone and $\lambda > 0$, then

$$Fix(J_{\lambda A}) = zer\ A.$$

The Proximal Point Algorithm I

Example: Let f be as before. Then the **subdifferential** of f

$$\partial f : H \to 2^H : \; x \mapsto \{u \in H : \forall y \in H\,(\langle y - x, u\rangle + f(x) \leq f(y))\}$$

is a maximally monotone operator.

Facts: If A is monotone and $\lambda > 0$, then

$$Fix(J_{\lambda A}) = zer\, A.$$

For f as above and $A := \partial f$ we have $\mathrm{prox}_f = J_{\partial f}$ and

$$\operatorname{argmin} f = Fix(J_{\lambda \partial f}) = zer\, \partial f.$$

The Proximal Point Algorithm I

Example: Let f be as before. Then the **subdifferential** of f

$$\partial f : H \to 2^H : \ x \mapsto \{u \in H : \forall y \in H\,(\langle y - x, u\rangle + f(x) \le f(y)\}$$

is a maximally monotone operator.

Facts: If A is monotone and $\lambda > 0$, then

$$Fix(J_{\lambda A}) = zer\, A.$$

For f as above and $A := \partial f$ we have $\mathrm{prox}_f = J_{\partial f}$ and

$$\mathrm{argmin}\, f = Fix(J_{\lambda \partial f}) = zer\, \partial f.$$

Let $(\lambda_n) \subset (0, \infty)$ and A maximally monotone, then the **Proximal Point Algorithm (PPA)** is defined by

$$x_{n+1} := J_{\lambda_n A}(x_n), \ x_0 \in H.$$

The proximal point algorithm II

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

The proximal point algorithm II

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

Under suitable conditions on $(\lambda_n) \subset (0, \infty)$: (x_n) **converges weakly to a zero** of A (Martinet 1970, Rockafellar 1976), but not strongly (Güler 1996).

The proximal point algorithm II

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

Under suitable conditions on $(\lambda_n) \subset (0, \infty)$: **(x_n) converges weakly to a zero** of ***A*** (Martinet 1970, Rockafellar 1976), but not strongly (Güler 1996).

Generalizations to Banach spaces (then A accretive) or geodesic spaces.

The proximal point algorithm II

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

Under suitable conditions on $(\lambda_n) \subset (0, \infty)$: **(x_n) converges weakly to a zero** of ***A*** (Martinet 1970, Rockafellar 1976), but not strongly (Güler 1996).

Generalizations to Banach spaces (then A accretive) or geodesic spaces.

In general **no effective rates of convergence** already for $\mathbb{R}$ (Neumann 2015).

The proximal point algorithm II

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

Under suitable conditions on $(\lambda_n) \subset (0, \infty)$: (x_n) **converges weakly to a zero** of A (Martinet 1970, Rockafellar 1976), but not strongly (Güler 1996).

Generalizations to Banach spaces (then A accretive) or geodesic spaces.

In general **no effective rates of convergence** already for $\mathbb{R}$ (Neumann 2015).

Rates of metastability in the finite dimensional/boundedly compact case:

The proximal point algorithm II

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

Under suitable conditions on $(\lambda_n) \subset (0, \infty)$: **(x_n) converges weakly to a zero** of A (Martinet 1970, Rockafellar 1976), but not strongly (Güler 1996).

Generalizations to Banach spaces (then A accretive) or geodesic spaces.

In general **no effective rates of convergence** already for $\mathbb{R}$ (Neumann 2015).

Rates of metastability in the finite dimensional/boundedly compact case:

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

The proximal point algorithm II

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

Under suitable conditions on $(\lambda_n) \subset (0, \infty)$: (x_n) **converges weakly to a zero** of A (Martinet 1970, Rockafellar 1976), but not strongly (Güler 1996).

Generalizations to Banach spaces (then A accretive) or geodesic spaces.

In general **no effective rates of convergence** already for $\mathbb{R}$ (Neumann 2015).

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.

The proximal point algorithm II

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

Under suitable conditions on $(\lambda_n) \subset (0, \infty)$: (x_n) **converges weakly to a zero** of A (Martinet 1970, Rockafellar 1976), but not strongly (Güler 1996).

Generalizations to Banach spaces (then A accretive) or geodesic spaces.

In general **no effective rates of convergence** already for $\mathbb{R}$ (Neumann 2015).

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.
- abstract forms of PPA in Hilbert space: Leuştean/Nicolae/Sipoş J. Global Opt. 2018.

The proximal point algorithm II

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

Under suitable conditions on $(\lambda_n) \subset (0, \infty)$: (x_n) **converges weakly to a zero** of A (Martinet 1970, Rockafellar 1976), but not strongly (Güler 1996).

Generalizations to Banach spaces (then A accretive) or geodesic spaces.

In general **no effective rates of convergence** already for $\mathbb{R}$ (Neumann 2015).

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: Leuştean/Nicolae/Sipoş 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**:

$$x_{n+1} := \alpha_n u + (1-\alpha_n) J_{\lambda_n A} x_n, \quad u, x_0 \in H \quad \textbf{(HPPA)}$$

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

Rates of metastability of HPPA

Rates of metastability of HPPA

Again in general **no effective rates of convergence**, **but rates of metastability**:

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

Rates of metastability of HPPA

Again in general **no effective rates of convergence**, **but rates of metastability**:

- 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.

Rates of metastability of HPPA

Again in general **no effective rates of convergence**, **but rates of metastability**:

- 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

Again in general **no effective rates of convergence**, **but rates of metastability**:

- 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.
- 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

Again in general **no effective rates of convergence**, **but rates of metastability**:

- 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.
- 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!

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

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

A function $\omega : \mathbb{N} \to \mathbb{N}$ 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 $F : X \to \overline{\mathbb{R}}$ with $zer\,F := \{x \in X : F(x) = 0\} \neq \emptyset$.
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 also covers fixed point and equilibrium problems.

Computational use of moduli of regularity

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

Let $F : X \to \overline{\mathbb{R}}$ 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 $\lim x_n \in zer\ F$.

If X is **compact** and T is continuous a modulus of regularity always exists:

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 (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.* $Fix(T)$.

Noncomputability of 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. $Fix(T)$.*

This can be recasted in terms of reverse mathematics:

Proposition (K. Computability 2019)

Over RCA_0, the statement that every continuous function $T : [0,1] \to [0,1]$ has modulus of metric regularity is equivalent to ACA.

Comment: The mere uniform metric regularity of continuous selfmaps of compact spaces is already provable in WKL_0.

Noncomputability of 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).*

This can be recasted in terms of reverse mathematics:

Proposition (K. Computability 2019)

Over RCA_0, the statement that every continuous function $T : [0,1] \to [0,1]$ has modulus of metric regularity is equivalent to ACA.

Comment: The mere uniform metric regularity of continuous selfmaps of compact spaces is already provable in WKL_0.

Example of a modulus of regularity

In fact, the cases where one can compute such a modulus are rare.

However there are important cases where this is true!

Example of a modulus of regularity

In fact, the cases where one can compute such a modulus are rare. However there are important cases where this is true!

Closed convex $C_1, C_2 \subseteq \mathbb{R}^n$: consider Douglas-Rachford operator

$$T_{C_1,C_2} := \frac{1}{2}(I + R_{C_2} R_{C_1}), \text{ where } R_{C_i} := 2P_{C_i} - I.$$

Example of a modulus of regularity

In fact, the cases where one can compute such a modulus are rare. However there are important cases where this is true!

Closed convex $C_1, C_2 \subseteq \mathbb{R}^n$: consider Douglas-Rachford operator

$$T_{C_1,C_2} := \frac{1}{2}(I + R_{C_2} R_{C_1}), \text{ where } R_{C_i} := 2P_{C_i} - I.$$

Borwein/Li/Tam SIAM 2017: if C_1, C_2 are convex semialgebraic sets in $\mathbb{R}^n$ with $0 \in C_1 \cap C_2$ which can be described by polynomials on $\mathbb{R}^n$ of degree > 1 then for given $r > 0$ T_{C_1,C_2} has modulus of regularity (w.r.t. $Fix(T_{C_1,C_2})$ on $B(0,r)$)

$$\omega(\varepsilon) := (\varepsilon/\mu)^\gamma$$

for all $x \in B_b(0)$ for suitable $\mu > 0, \gamma > 1$.

Example of a modulus of regularity

In fact, the cases where one can compute such a modulus are rare. However there are important cases where this is true!

Closed convex $C_1, C_2 \subseteq \mathbb{R}^n$: consider Douglas-Rachford operator

$$T_{C_1,C_2} := \frac{1}{2}(I + R_{C_2} R_{C_1}), \text{ where } R_{C_i} := 2P_{C_i} - I.$$

Borwein/Li/Tam SIAM 2017: if C_1, C_2 are convex semialgebraic sets in $\mathbb{R}^n$ with $0 \in C_1 \cap C_2$ which can be described by polynomials on $\mathbb{R}^n$ of degree > 1 then for given $r > 0$ T_{C_1,C_2} has modulus of regularity (w.r.t. $Fix(T_{C_1,C_2})$ on $B(0,r)$)

$$\omega(\varepsilon) := (\varepsilon/\mu)^{\gamma}$$

for all $x \in B_b(0)$ for suitable $\mu > 0, \gamma > 1$. There interesting connections to **o-minimality** which gave rise to **'tame optimization'** (Bolte, Daniilidis, Lewis, Ioffe,...), see e.g.

A.D. Ioffe: In invitation to tame optimization. Siam J. Optimiz. 2009.

Applications in Nonconvex Optimization

The monotonicity of ∂f is due to the convexity assumption on f.

The monotonicity of ∂f is due to the convexity assumption on f.

To treat **nonconvex-nonconcave min-max optimization** one has to consider **generalizations of monotone operators**.

The monotonicity of ∂f is due to the convexity assumption on f.

To treat **nonconvex-nonconcave min-max optimization** one has to consider **generalizations of monotone operators**.

Definition (Bauschke/Moursi/Wang 2020; Combettes/Pennanen 2004)

Let $\rho \in \mathbb{R}$. $A : H \to 2^H$ is called ρ-comonotone if

$$\forall (x,u),(y,v) \in gr(A)\,(\langle x-y, u-v\rangle \geq \rho \|u-v\|^2).$$

The monotonicity of ∂f is due to the convexity assumption on f.

To treat **nonconvex-nonconcave min-max optimization** one has to consider **generalizations of monotone operators**.

Definition (Bauschke/Moursi/Wang 2020; Combettes/Pennanen 2004)

Let $\rho \in \mathbb{R}$. $A : H \to 2^H$ is called ρ-comonotone if

$$\forall (x, u), (y, v) \in gr(A) \, (\langle x - y, u - v \rangle \geq \rho \|u - v\|^2).$$

For $\rho < 0$ this **generalizes** the concept of monotonicity.

The monotonicity of ∂f is due to the convexity assumption on f.

To treat **nonconvex-nonconcave min-max optimization** one has to consider **generalizations of monotone operators**.

Definition (Bauschke/Moursi/Wang 2020; Combettes/Pennanen 2004)

Let $\rho \in \mathbb{R}$. $A : H \to 2^H$ is called ρ-comonotone if

$$\forall (x, u), (y, v) \in gr(A) \left(\langle x - y, u - v \rangle \geq \rho \|u - v\|^2 \right).$$

For $\rho < 0$ this **generalizes** the concept of monotonicity.

Recently (arXiv Oct.2020), Diakonikolas/Daskalakis/Jordan considered this and even more general forms in the context of nonconvex-nonconcave min-max optimization and machine learning!

Uniform strong nonexpansivity of families of functions

Our proof mining of convergence results on the PPA and the HPPA show that these results essentially only need use (though implicitly) that $(J_{\gamma_n}A)$ has a **common modulus of strong nonexpansivity** (SNE-modulus):

Uniform strong nonexpansivity of families of functions

Our proof mining of convergence results on the PPA and the HPPA show that these results essentially only need use (though implicitly) that $(J_{\gamma_n}A)$ has a **common modulus of strong nonexpansivity** (SNE-modulus):

Definition (Bruck/Reich 1977, K. 2016)

$C \subseteq X$ subset of some Banach space X. $T : C \to X$ is **strongly nonexpansive** with **SNE-modulus** $\omega : (0,\infty)^2 \to (0,\infty)$ if

$\forall b, \varepsilon > 0\, \forall x, y \in C$

$\|x-y\| \leq b \wedge \|x-y\| - \|Tx - Ty\| < \omega(b,\varepsilon) \to \|(x-y)-(Tx-Ty)\| < \varepsilon.$

If A is monotone (in Hilbert space) or accretive (in Banach space), then $J_\gamma A$ is firmly nonexpansive (purely universal condition).

If A is monotone (in Hilbert space) or accretive (in Banach space), then $J_\gamma A$ is firmly nonexpansive (purely universal condition).

Proposition (K. Israel J. Math. 2016)

If X is uniformly convex with modulus η and $T : C \to X$ is firmly nonexpansive, then T is SNE with modulus

$$\omega_\eta(b,\varepsilon) = \frac{1}{4}\eta(\varepsilon/b)\cdot\varepsilon.$$

In **Hilbert space** $\omega(b,\varepsilon) := \frac{1}{16b}\varepsilon^2$.

Proposition (K. Optimization Letters 2021)

Let H be a real Hilbert space and $(\gamma_n) \subset (0, \infty), \gamma > 0$ be such that $\gamma_n \geq \gamma > 0$ for all $n \in \mathbb{N}$. Let $\rho \in (-\frac{\gamma}{2}, 0]$ and $A \subseteq H \times H$ be ρ-comonotone.

Proposition (K. Optimization Letters 2021)

Let H be a real Hilbert space and $(\gamma_n) \subset (0, \infty), \gamma > 0$ be such that $\gamma_n \geq \gamma > 0$ for all $n \in \mathbb{N}$. Let $\rho \in (-\frac{\gamma}{2}, 0]$ and $A \subseteq H \times H$ be ρ-comonotone.

Then $J_{\gamma_n A} : R(I + \gamma_n A) \to D(A)$ is strongly nonexpansive with **common SNE-modulus**

$$\omega_\alpha(b, \varepsilon) := \frac{1-\alpha}{4b\alpha} \cdot \varepsilon^2, \text{ where } \alpha := \frac{1}{2((\rho/\gamma)+1)} \in (0,1).$$

The proof uses crucially a recent result by Bauschke/Moursi/Wang 2020, that J_A is an averaged map whenever A is $> -\frac{1}{2}$ comonotone.

Proposition (K. Optimization Letters 2021)

Let H be a real Hilbert space and $(\gamma_n) \subset (0, \infty), \gamma > 0$ be such that $\gamma_n \geq \gamma > 0$ for all $n \in \mathbb{N}$. Let $\rho \in (-\frac{\gamma}{2}, 0]$ and $A \subseteq H \times H$ be ρ-comonotone.

Then $\boldsymbol{J_{\gamma_n A} : R(I + \gamma_n A) \to D(A)}$ is strongly nonexpansive with **common SNE-modulus**

$$\boldsymbol{\omega_\alpha(b, \varepsilon) := \frac{1-\alpha}{4b\alpha} \cdot \varepsilon^2, \text{ where } \alpha := \frac{1}{2((\rho/\gamma)+1)} \in (0,1).}$$

The proof uses crucially a recent result by Bauschke/Moursi/Wang 2020, that J_A is an averaged map whenever A is $> -\frac{1}{2}$ comonotone.

Averaged maps in uniformly convex Banach spaces are also strongly nonexpansive (Bruck/Reich 1977).

Proposition (K. Optimization Letters 2021)

Let H be a real Hilbert space and $(\gamma_n) \subset (0,\infty), \gamma > 0$ be such that $\gamma_n \geq \gamma > 0$ for all $n \in \mathbb{N}$. Let $\rho \in (-\frac{\gamma}{2}, 0]$ and $A \subseteq H \times H$ be ρ-comonotone.

Then $\boldsymbol{J_{\gamma_n A} : R(I + \gamma_n A) \to D(A)}$ is strongly nonexpansive with **common SNE-modulus**

$$\boldsymbol{\omega_\alpha(b,\varepsilon) := \frac{1-\alpha}{4b\alpha} \cdot \varepsilon^2, \text{ where } \alpha := \frac{1}{2((\rho/\gamma)+1)} \in (0,1).}$$

The proof uses crucially a recent result by Bauschke/Moursi/Wang 2020, that J_A is an averaged map whenever A is $> -\frac{1}{2}$ comonotone.

Averaged maps in uniformly convex Banach spaces are also strongly nonexpansive (Bruck/Reich 1977).

Hilbert space: proper generalization of the firmly nonexpansive mappings.

Proposition (K. Optimization Letters 2021)

Let H be a real Hilbert space and $(\gamma_n) \subset (0,\infty), \gamma > 0$ be such that $\gamma_n \geq \gamma > 0$ for all $n \in \mathbb{N}$. Let $\rho \in (-\frac{\gamma}{2}, 0]$ and $A \subseteq H \times H$ be ρ-comonotone.

Then $J_{\gamma_n A} : R(I + \gamma_n A) \to D(A)$ is strongly nonexpansive with **common SNE-modulus**

$$\omega_\alpha(b,\varepsilon) := \frac{1-\alpha}{4b\alpha} \cdot \varepsilon^2, \text{ where } \alpha := \frac{1}{2((\rho/\gamma)+1)} \in (0,1).$$

The proof uses crucially a recent result by Bauschke/Moursi/Wang 2020, that J_A is an averaged map whenever A is $> -\frac{1}{2}$ comonotone.

Averaged maps in uniformly convex Banach spaces are also strongly nonexpansive (Bruck/Reich 1977).

Hilbert space: proper generalization of the firmly nonexpansive mappings.

SNE-modulus for averaged maps in Hilbert space: Sipos 2020.

Results on PPA and HPPA in Hilbert space for ρ-comonotone operators

- Rate of metastability for the convergence of the PPA in the boundedly compact case.

Results on PPA and HPPA in Hilbert space for ρ-comonotone operators

- Rate of metastability for the convergence of the PPA in the boundedly compact case.
- Rates of convergence of the PPA in the general case if one has a modulus of regularity.

Results on PPA and HPPA in Hilbert space for ρ-comonotone operators

- Rate of metastability for the convergence of the PPA in the boundedly compact case.
- Rates of convergence of the PPA in the general case if one has a modulus of regularity.
- Rate of metastability for the convergence of HPPA in the general case together with quantitative information of the limit being a zero of A.

Theorem (K. Optimization Letters 2021)

Let $A \subseteq H \times H$ be ρ-comonotone, $(\gamma_n), \gamma, \rho$ as before. Assume that $\overline{D(A)} \subseteq \bigcap_{n=0}^{\infty} R(I + \gamma_n A)$ is boundedly compact and $x_0 \in \overline{D(A)}$.

Theorem (K. Optimization Letters 2021)

Let $A \subseteq H \times H$ be ρ-comonotone, $(\gamma_n), \gamma, \rho$ as before. Assume that $\overline{D(A)} \subseteq \bigcap_{n=0}^{\infty} R(I + \gamma_n A)$ is boundedly compact and $x_0 \in \overline{D(A)}$. Then (x_n) strongly converges to a zero of A. Moreover,

$$(*) \left\{ \begin{array}{l} \forall k \in \mathbb{N}\, \forall g \in \mathbb{N}^{\mathbb{N}}\, \exists n \leq \Psi(k, g, \beta)\, \forall i, j \in [n, n + g(n)] \\ \left(\|x_i - x_j\| \leq \frac{1}{k+1} \text{ and } x_i \in \tilde{F}_k \right), \end{array} \right.$$

where

$$\tilde{F}_k := \bigcap_{i \leq k} \left\{ x \in \overline{D(A)} \, : \, \|x - J_{\gamma_i A} x\| \leq \frac{1}{k+1} \right\}$$

and β is a modulus of total boundedness for $\overline{D(A)} \cap \overline{B}(0, M)$, where $\overline{B}(0, M) := \{x \in H : \|x\| \leq M\}$, with $M \geq b + \|p\|$ and $b \geq \|x_0 - p\|$ for some $p \in zer\, A$.

Here $\Psi(k,g,\beta) := \Psi_0(P,k_0,g)$, with

$$\begin{cases} \Psi_0(0,k_0,g) := 0 \\ \Psi_0(n+1,k_0,g) := \Phi\left(\chi^M_{k,g}\left(\Psi_0(n,k_0,g), 4k_0+3\right)\right), \end{cases}$$

and

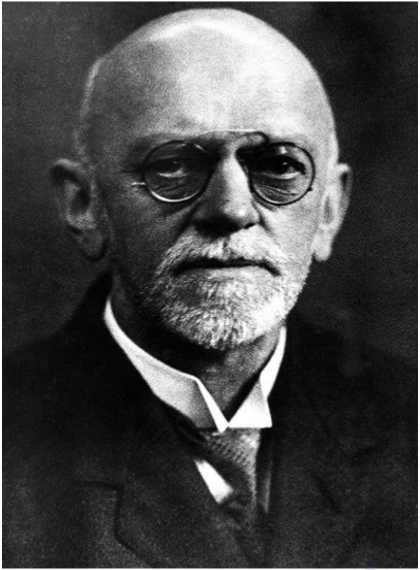
$$\chi_{k,g}(n,r) := \max\{2k+1, \chi(n,g(n),r)\},\ \chi^M_{k,g}(n,r) := \max_{i\le n}\{\chi_{k,g}(i,r)\},$$
$$P := \beta\left(4k_0+3\right),\ k_0 = 2k+1\ \ \chi(n,m,r) := \max\{n+m-1, m(r+1)\}$$
$$\Phi(k) := \left\lceil \frac{b}{\omega_\alpha(b,((k+1)C_k)^{-1})} \right\rceil + 1,\ \ C_k \ge 2 + \frac{\gamma_i}{\gamma} \text{ for all } i \le k.$$

Theorem (K. Optimization Letters 2021)

Let A and $(\gamma_n), \gamma, \rho, b$ be as above and assume that $\overline{D(A)} \subseteq \bigcap_{n=0}^{\infty} R(I + \gamma_n A)$. If A has a modulus ϕ of regularity (suitable adapted for the set-valued case) w.r.t *zer* A and $\overline{B}(p, b)$, then **without compactness assumption** (x_n) converges to a zero $z := \lim x_n$ of A with rate of convergence

$$\boldsymbol{\xi(\varepsilon, \gamma, b) := \left\lceil \frac{b}{\omega_\alpha\left(b, \phi(\varepsilon/2) \cdot \gamma\right)} \right\rceil + 2.}$$

David Hilbert (1862-1943)



- H.H. Bauschke, W.A. Moursi, X. Wang, Generalized monotone operators and their averaged resolvents. Math. Programming Ser. B, to appear.
- J. Diakonikolas, C. Daskalakis, M.I. Jordan, Efficient methods for structured nonconvexnonconcave Min-Max optimization. arXiv:2011.00364.
- B. Dinis, P. Pinto, Quantitative results on the multi-parameters proximal point algorithm. J. Convex Anal. 28, 23 pp, (2021).
- U. Kohlenbach, Proof-theoretic Methods in Nonlinear Analysis. In: Proc. ICM 2018, Vol. 2, pp. 61-82. World Scientific 2019.
- U. Kohlenbach, On the reverse mathematics and Weihrauch complexity of moduli of regularity and uniqueness. Computability vol. 8, pp. 377-387 (2019).
- U. Kohlenbach, Quantitative analysis of a Halpern-type Proximal Point Algorithm for accretive operators in Banach spaces. J. Nonlinear Convex Anal. 21, pp. 2125-2138 (2020).

- U. Kohlenbach, Quantitative results on the Proximal Point Algorithm in uniformly convex Banach spaces. J. Convex Anal. 28, pp. 11-18 (2021).
- U. Kohlenbach, On the Proximal Point Algorithm and its Halpern-type variant for generalized monotone operators in Hilbert space. Optimization Letters, to appear 2021.
- U. Kohlenbach, L. Leuștean, A. Nicolae, Quantitative results on Fejér monotone sequences. Comm. Contemp. Math. 20 (2018), 42pp.
- U. Kohlenbach, G. López-Acedo, A. Nicolae, Moduli of regularity and rates of convergence for Fejer monotone sequences. Israel J. Math. vol. 232, pp. 261-297 (2019).
- U. Kohlenbach, G. López-Acedo, A. Nicolae, A uniform betweenness property in metric spaces and its role in the quantitative analysis of the 'Lion-Man' game. Pacific J. Math. Vol. 310, pp. 181-212 (2021).

- L. Leuștean, A. Nicolae, A. Sipoș, An abstract proximal point algorithm. J. Global Optim. 72, pp. 553-577 (2018).
- L. Leuștean, P. Pinto, Quantitative results on the Halpern type proximal point algorithm. Computational Optim. Appl., pp. 101-125 (2021).
- L. Leuștean, A. Sipoș, Effective strong convergence of the proximal point algorithm in CAT(0) spaces. J. Nonlinear Variational Anal. 2, pp. 219-228 (2018).
- P. Pinto, Proof mining with the bounded functional interpretation. PhD Thesis, Universidade de Lisboa, 143pp., 2019.
- P. Pinto, A rate of metastability for the Halpern type Proximal Point Algorithm. Numer. Funct. Anal. Optim. 42, pp. 320-343, (2021).