Proof Mining:
Extraction of Information from Proofs

Ulrich Kohlenbach
Department of Mathematics
Technische Universität Darmstadt

Proof Society 2024: 6th International Summer School and Workshop,
U Birmingham, UK, 9-13 September 2024.

Overview of Content

Lecture I: General Introduction to the Unwinding of Proofs ('Proof Mining')

Overview of Content

Lecture I: General Introduction to the Unwinding of Proofs ('Proof Mining')

Lecture II: Functional Interpretation and Logical Metatheorems

Overview of Content

Lecture I: General Introduction to the Unwinding of Proofs ('Proof Mining')

Lecture II: Functional Interpretation and Logical Metatheorems

Lecture III: Applications in Ergodic Theory, Convex Optimization and Nonlinear Semigroup Theory

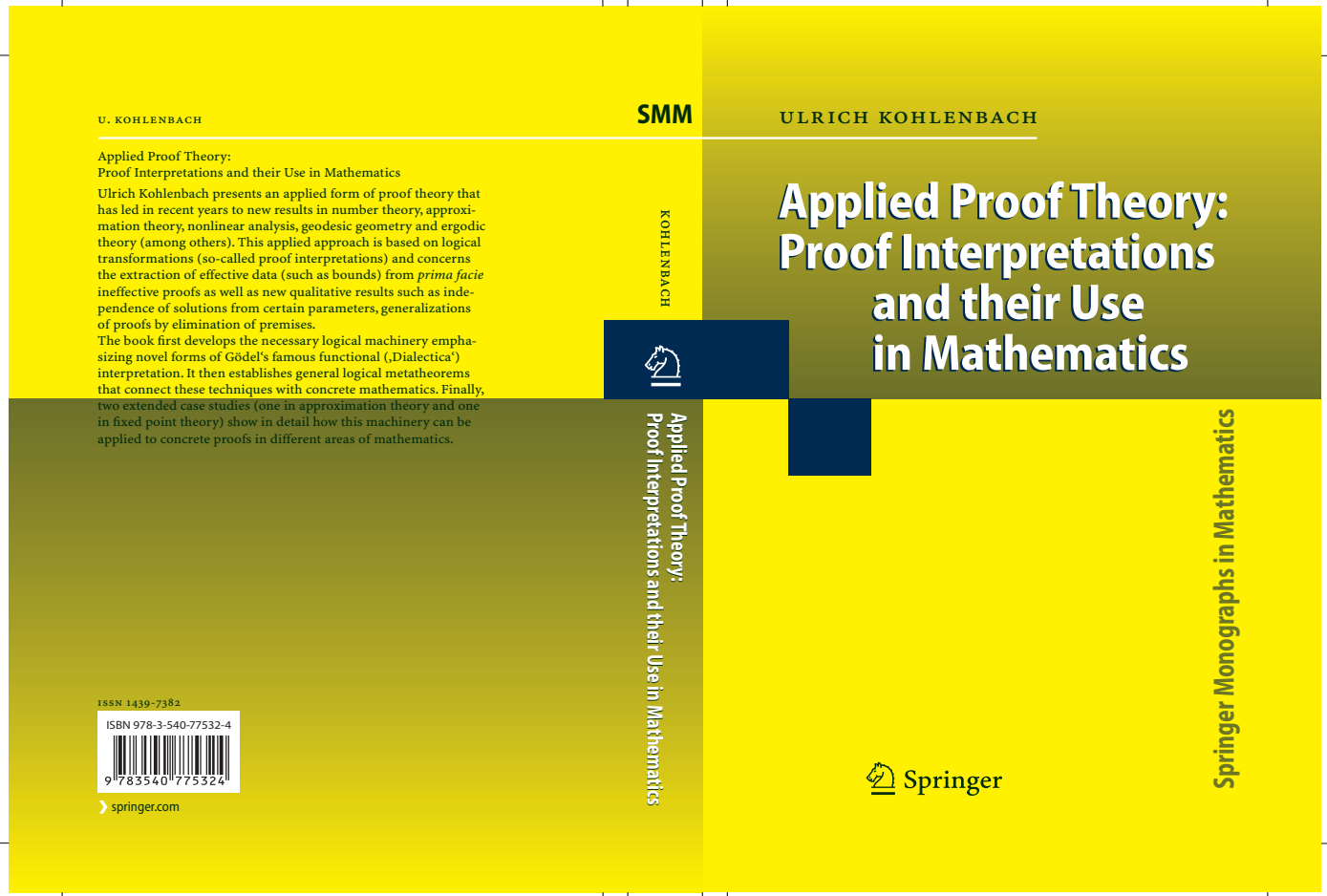
- During (mainly) the last 20 years modern forms of functional interpretations were used to obtain **numerous new quantitative results** as well as **qualitative uniformity results** in particular in: nonlinear analysis, fixed point theory, ergodic theory, topological dynamics, approximation theory, convex optimization, abstract Cauchy problems, probability theory ($\geq$ 200 papers mostly in journals in the resp. areas or general mathematics journals).

- During (mainly) the last 20 years modern forms of functional interpretations were used to obtain **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, probability theory ($\geq$ 200 papers mostly in journals in the resp. areas or general mathematics journals).
- General **logical metatheorems** explain this as instances of logical phenomena (K. APAL 1993, TAMS 2005, Gerhardy/K. TAMS 2008, Leucstean APAL 2014, Günzel/K. Adv. Math. 2016, Engracia/ Ferreira 2020, Sipoș MLQ 2023, K./Pischke Phil. Trans. Royal Soc. 2023, Pischke JML 2023, TAMS 2024, Neri/Pischke 2024.

- During (mainly) the last 20 years modern forms of functional interpretations were used to obtain **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, probability theory ($\geq$ 200 papers mostly in journals in the resp. areas or general mathematics journals).
- General **logical metatheorems** explain this as instances of logical phenomena (K. APAL 1993, TAMS 2005, Gerhardy/K. TAMS 2008, Leucstean APAL 2014, Günzel/K. Adv. Math. 2016, Engracia/ Ferreira 2020, Sipoș MLQ 2023, K./Pischke Phil. Trans. Royal Soc. 2023, Pischke JML 2023, TAMS 2024, Neri/Pischke 2024.
- Some of the logical tools used have been rediscovered in 2007 in special cases by Terence Tao prompted by concrete mathematical needs **'finitary analysis'!**

Main references

2008 book:



Recent surveys

1) U. Kohlenbach: Recent progress in proof mining in nonlinear analysis. IFCoLog J. of Logics and its Applications 10:4, pp. 3357-3406 (2017).

2) U. Kohlenbach: Proof-theoretic methods in nonlinear analysis. Proc. Int. Cong. of Math. - ICM 2018.

Both available at: www.mathematik.tu-darmstadt.de/~kohlenbach

Lecture I

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 $\mathrm{Con}_{\mathcal{T}}$ of $\mathcal{T}$ by finitistic means FIN!

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 $\mathrm{Con}_{\mathcal{T}}$ of $\mathcal{T}$ by finitistic means FIN!

For '**real statements**' φ (i.e. $\varphi \in \Pi^0_1$) this would imply (Hilbert 1927)

$$\mathcal{T} \vdash \varphi \overset{\text{Kreisel}}{\Rightarrow} \mathrm{FIN} + \mathrm{Con}_{\mathcal{T}} \vdash \varphi \overset{\mathrm{FIN} \vdash \mathrm{Con}_{\mathcal{T}}}{\Rightarrow} \mathrm{FIN} \vdash \varphi.$$

- Though in this specific form in general impossible (Gödel), large parts of ordinary mathematics can be carried out in systems which can be proof-theoretically reduced to (and are conservative over) Peano arithmetic PA or even PRA (Feferman, Friedman, Simpson...).

- Though in this specific form in general impossible (Gödel), large parts of ordinary mathematics can be carried out in systems which can be proof-theoretically reduced to (and are conservative over) Peano arithmetic PA or even PRA (Feferman, Friedman, Simpson...).
- While it has been claimed (Tait) that Hilbert's finitism corresponds to primitive recursive arithmetic PRA, Hilbert himself argued in 1925 for the finitistic character of the primitive recursive functionals of finite type (nowadays called Gödel's T) which includes the Ackermann function.

- Though in this specific form in general impossible (Gödel), large parts of ordinary mathematics can be carried out in systems which can be proof-theoretically reduced to (and are conservative over) Peano arithmetic PA or even PRA (Feferman, Friedman, Simpson...).
- While it has been claimed (Tait) that Hilbert's finitism corresponds to primitive recursive arithmetic PRA, Hilbert himself argued in 1925 for the finitistic character of the primitive recursive functionals of finite type (nowadays called Gödel's T) which includes the Ackermann function.
- In his 'Dialectica' paper (1958), Gödel showed that the consistency of Peano Arithmetic PA can be finitistically reduced to that of T.

Proof Theory: from Foundations to Applications

General malaise of consistency proofs:

Proof Theory: from Foundations to Applications

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).

Proof Theory: from Foundations to Applications

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).

- Already Hilbert apparently had broader goal: exhibit the computational content of proofs of concrete statements (e.g. Π_2-sentences) which use ideal elements.

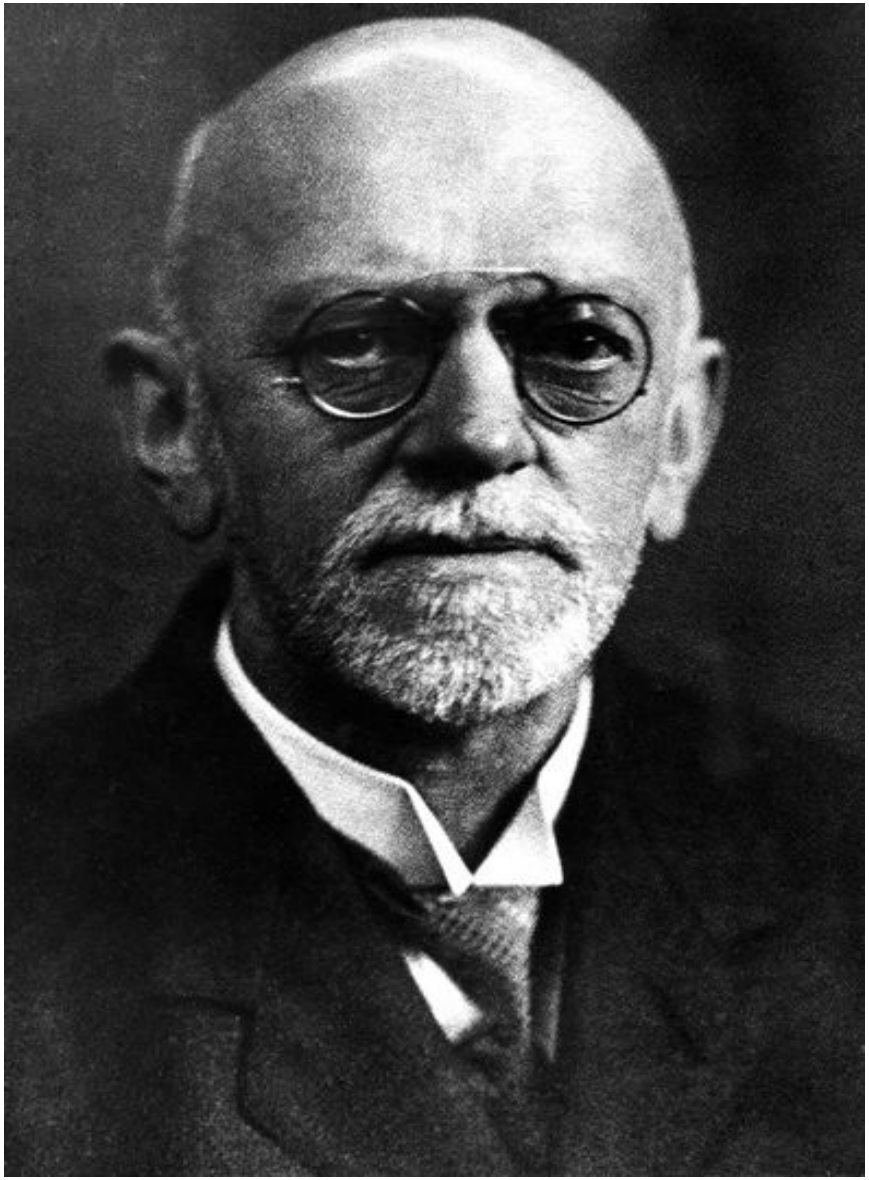
Proof Theory: from Foundations to Applications

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).

- Already Hilbert apparently had broader goal: exhibit the computational content of proofs of concrete statements (e.g. Π_2-sentences) which use ideal elements.
- In the Hilbert school bounds were extracted from proofs in abstract algebra, e.g. by Grete Hermann in her 1925 PhD thesis (supervised by Emmy Noether).

David Hilbert (1862-1943)



- Since the 50's, Kreisel suggested to use proof transformations developed in the course of Hilbert's program to analyze noneffective proofs in mathematics:

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

- Since the 50's, Kreisel suggested to use proof transformations developed in the course of Hilbert's program to analyze noneffective proofs in mathematics:

  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).

- Since the 50's, Kreisel suggested to use proof transformations developed in the course of Hilbert's program to analyze noneffective proofs in mathematics:

  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).

- Early success: Luckhardt's polynomial bounds on Roth's theorem (1985, see below) and Delzell's work (1980-1996) on Hilbert's 17th problem.

Georg Kreisel (1923-2015)

Proof Mining since 1990: focus on proofs in analysis

Input: Noneffective proof P of C

Proof Mining since 1990: focus on proofs in analysis

Input: Noneffective proof P of C

Goal: Additional information on C:

Proof Mining since 1990: focus on proofs in analysis

Input: Noneffective proof P of C

Goal: Additional information on C:

- quantitative information: explicit effective bounds,

Proof Mining since 1990: focus on proofs in analysis

Input: Noneffective proof P of C

Goal: Additional information on C:

- quantitative information: explicit effective bounds,
- qualitative aspects:

Proof Mining since 1990: focus on proofs in analysis

Input: Noneffective proof P of C

Goal: Additional information on C:

- quantitative information: explicit effective bounds,
- qualitative aspects:
  - existence of bound that is independent from certain parameters,

Proof Mining since 1990: focus on proofs in analysis

Input: Noneffective proof P of C

Goal: Additional information on C:

- quantitative information: explicit effective bounds,
- qualitative aspects:
  - existence of bound that is independent from certain parameters,
  - generalizations of proofs: weakening of premises,

Proof Mining since 1990: focus on proofs in analysis

Input: Noneffective proof P of C

Goal: Additional information on C:

- quantitative information: explicit effective bounds,
- qualitative aspects:
  - existence of bound that is independent from certain parameters,
  - generalizations of proofs: weakening of premises,
  - finding 'correct' definitions of concepts.

Numerous case studies in nonlinear analysis show that indeed computational content of low complexity can be extracted even from proofs that use ideal principles:

proof-theoretic tameness of ordinary mathematics!

- The approach proceeds via a **translation of a given proof into a new proof of a new result based on intuitionistic logic** while still using **ideal elements**.

- The approach proceeds via a **translation of a given proof into a new proof of a new result based on intuitionistic logic** while still using **ideal elements**.
- A **'finitistic' analysis (relative to the ideal elements)** of the latter exhbits explicit constructions hidden in the use of quantifiers.

- The approach proceeds via a **translation of a given proof into a new proof of a new result based on intuitionistic logic** while still using **ideal elements**.
- A **'finitistic' analysis (relative to the ideal elements)** of the latter exhbits explicit constructions hidden in the use of quantifiers.
- The finitistic use of the 'ideal elements' **replaces them by finitary versions** (without the proof noticing the difference!).

- The approach proceeds via a **translation of a given proof into a new proof of a new result based on intuitionistic logic** while still using **ideal elements**.
- A **'finitistic' analysis (relative to the ideal elements)** of the latter exhbits explicit constructions hidden in the use of quantifiers.
- The finitistic use of the 'ideal elements' **replaces them by finitary versions** (without the proof noticing the difference!).
- This gives a **finitistic realization of $\forall\exists$-theorems** together with an elementary verification.

- The approach proceeds via a **translation of a given proof into a new proof of a new result based on intuitionistic logic** while still using **ideal elements**.
- A **'finitistic' analysis (relative to the ideal elements)** of the latter exhbits explicit constructions hidden in the use of quantifiers.
- The finitistic use of the 'ideal elements' **replaces them by finitary versions** (without the proof noticing the difference!).
- This gives a **finitistic realization of $\forall\exists$-theorems** together with an elementary verification.
- For **semi-intuitionistic proofs** the restriction to $\forall\exists$-theorem can be overcome: **arbitrary logical complexity allowed!**

A polynomial rate of asymptotic regularity in Bauschke's solution of the 'minimal 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 'minimal 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 'minimal 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 'minimal 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.

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

Theorem (K. FoCM 2018)

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

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

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

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

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

Theorem (K. FoCM 2018)

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

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

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

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

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

A. Sipoș: Extension to general **averaged mappings** (Opt.Lett. 2021) and to **super strongly nonexpansive mappings** (Israel J. Math. 2024).

Naive attempt fails

Naive Attempt: try to extract an explicit computable function realizing (or bounding) '$\exists y$': $\forall x \in \mathbb{N}\, F(x, f(x))$.

Proposition

There exist a sentence $A \equiv \forall x \exists y \forall z\, A_{qf}(x, y, z)$ in the language of arithmetic (A_{qf} quantifier-free and hence decidable), such

- A is **logical valid**,

Naive attempt fails

Naive Attempt: try to extract an explicit computable function realizing (or bounding) '$\exists y$': $\forall x \in \mathbb{N}\, F(x, f(x))$.

Proposition

There exist a sentence $A \equiv \forall x \exists y \forall z\, A_{qf}(x, y, z)$ in the language of arithmetic (A_{qf} quantifier-free and hence decidable), such

- A is **logical valid**,
- there is **no recursive bound f** s.t. $\forall x \exists y \leq f(x) \forall z\, A_{qf}(x, y, z)$.

Naive attempt fails

Naive Attempt: try to extract an explicit computable function realizing (or bounding) '$\exists y$': $\forall x \in \mathbb{N}\, F(x, f(x))$.

Proposition

There exist a sentence $A \equiv \forall x \exists y \forall z\, A_{qf}(x, y, z)$ in the language of arithmetic (A_{qf} quantifier-free and hence decidable), such

- A is **logical valid**,
- there is **no recursive bound f** s.t. $\forall x \exists y \leq f(x) \forall z\, A_{qf}(x, y, z)$.

Proof: Take $A :\equiv \forall x \exists y \forall z \big(T(x, x, y) \vee \neg T(x, x, z)\big)$, where T is the (primitive recursive) Kleene-T-predicate.

Naive attempt fails

Naive Attempt: try to extract an explicit computable function realizing (or bounding) '$\exists y$': $\forall x \in \mathbb{N}\, F(x, f(x))$.

Proposition

There exist a sentence $A \equiv \forall x \exists y \forall z\, A_{qf}(x, y, z)$ in the language of arithmetic (A_{qf} quantifier-free and hence decidable), such

- A is **logical valid**,
- there is **no recursive bound f** s.t. $\forall x \exists y \leq f(x) \forall z\, A_{qf}(x, y, z)$.

Proof: Take $A :\equiv \forall x \exists y \forall z \big(T(x, x, y) \vee \neg T(x, x, z)\big)$, where T is the (primitive recursive) Kleene-T-predicate.

Any bound g on '$\exists y$', i.e. no computable g such that

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

since this would solve the halting problem! □

However, one can obtain such **witness candidates** and bounds (and even realizing function(al)s) for a **weakened version A^H of A**:

However, one can obtain such **witness candidates** and bounds (and even realizing function(al)s) for a **weakened version A^H of A**:

Definition

$\boldsymbol{A \equiv \exists x_1 \forall y_1 \exists x_2 \forall y_2 A_{qf}(x_1, y_1, x_2, y_2)}$. Then the **Herbrand normal form** of A is defined as

$$\mathbf{A^H :\equiv \exists x_1, x_2 A_{qf}(x_1, f(x_1), x_2, g(x_1, x_2))},$$

where $\boldsymbol{f, g}$ are new function symbols, called index functions.

However, one can obtain such **witness candidates** and bounds (and even realizing function(al)s) for a **weakened version A^H of A**:

Definition

$\boldsymbol{A \equiv \exists x_1 \forall y_1 \exists x_2 \forall y_2 A_{qf}(x_1, y_1, x_2, y_2)}$. Then the **Herbrand normal form** of A is defined as

$$\mathbf{A^H :\equiv \exists x_1, x_2 A_{qf}(x_1, f(x_1), x_2, g(x_1, x_2))},$$

where $\boldsymbol{f, g}$ are new function symbols, called index functions.

A and A^H are equivalent with respect to logical validity, i.e.

$$\models \mathbf{A} \Leftrightarrow \models \mathbf{A^H},$$

but are not logically equivalent (but only in the presence of AC).

We now consider again the sentence

$$\mathbf{A} \equiv \forall \mathbf{x}\, \exists \mathbf{y}\, \forall \mathbf{z}\, (\mathbf{P}(\mathbf{x}, \mathbf{y}) \vee \neg \mathbf{P}(\mathbf{x}, \mathbf{z})),$$

We now consider again the sentence

$$\mathbf{A \equiv \forall x\, \exists y\, \forall z\, (P(x,y) \vee \neg P(x,z)),}$$

In contrast to A, the **Herbrand normal form** A^H of A

$$\mathbf{A^H \equiv \exists y \big(P(x,y) \vee \neg P(x, g(y))\big)}$$

allows one to construct a **list of candidates** (uniformly in x, g) for '$\exists y$', namely $(c, g(c))$ for any constant c (also $(x, g(x))$)

We now consider again the sentence

$$\mathbf{A} \equiv \forall \mathbf{x}\, \exists \mathbf{y}\, \forall \mathbf{z}\, (\mathbf{P}(\mathbf{x}, \mathbf{y}) \vee \neg \mathbf{P}(\mathbf{x}, \mathbf{z})),$$

In contrast to A, the **Herbrand normal form** A^H of A

$$\mathbf{A^H} \equiv \exists \mathbf{y} \big(\mathbf{P}(\mathbf{x}, \mathbf{y}) \vee \neg \mathbf{P}(\mathbf{x}, \mathbf{g}(\mathbf{y}))\big)$$

allows one to construct a **list of candidates** (uniformly in x, g) for '$\exists y$', namely $(c, g(c))$ for any constant c (also $(x, g(x))$)

$$\boldsymbol{A^{H,D}} :\equiv \big(\boldsymbol{P(x, c)} \vee \neg \boldsymbol{P(x, g(c))}\big) \vee \big(\boldsymbol{P(x, g(c))} \vee \neg \boldsymbol{P(x, g(g(c)))}\big)$$

We now consider again the sentence

$$\mathbf{A} \equiv \forall \mathbf{x}\, \exists \mathbf{y}\, \forall \mathbf{z}\, (\mathbf{P}(\mathbf{x}, \mathbf{y}) \vee \neg \mathbf{P}(\mathbf{x}, \mathbf{z})),$$

In contrast to A, the **Herbrand normal form** A^H of A

$$\mathbf{A^H} \equiv \exists \mathbf{y} \big(\mathbf{P}(\mathbf{x}, \mathbf{y}) \vee \neg \mathbf{P}(\mathbf{x}, \mathbf{g}(\mathbf{y}))\big)$$

allows one to construct a **list of candidates** (uniformly in x, g) for '$\exists y$', namely $(c, g(c))$ for any constant c (also $(x, g(x))$)

$$\boldsymbol{A^{H,D}} :\equiv \underbrace{\big(\boldsymbol{P}(\boldsymbol{x}, \boldsymbol{c}) \vee \neg \boldsymbol{P}(\boldsymbol{x}, \boldsymbol{g}(\boldsymbol{c}))\big)}_{\in \mathrm{TAUT}} \vee \big(\boldsymbol{P}(\boldsymbol{x}, \boldsymbol{g}(\boldsymbol{c})) \vee \neg \boldsymbol{P}(\boldsymbol{x}, \boldsymbol{g}(\boldsymbol{g}(\boldsymbol{c})))\big)$$

is a tautology.

J. Herbrand's Theorem ('Théorème fondamental', 1930)

Theorem

Let $A \equiv \exists x_1 \forall y_1 \exists x_2 \forall y_2 A_{qf}(x_1, y_1, x_2, y_2)$. Then:
PL $\vdash A$ iff there are terms $s_1, \ldots, s_k, t_1, \ldots, t_n$ (built up out of the constants and variables of A and the **index functions** used for the formation of A^H) such that

$$A^{H,D} :\equiv \bigvee_{i=1}^{k} \bigvee_{j=1}^{n} A_{qf}(s_i, f(s_i), t_j, g(s_i, t_j))$$

is a tautology. $A^{H,D}$ is called a **Herbrand Disjunction**.

J. Herbrand's Theorem ('Théorème fondamental', 1930)

Theorem

Let $A \equiv \exists x_1 \forall y_1 \exists x_2 \forall y_2 A_{qf}(x_1, y_1, x_2, y_2)$. Then:
PL $\vdash A$ iff there are terms $s_1, \ldots, s_k, t_1, \ldots, t_n$ (built up out of the constants and variables of A and the **index functions** used for the formation of A^H) such that

$$A^{H,D} :\equiv \bigvee_{i=1}^{k} \bigvee_{j=1}^{n} A_{qf}(s_i, f(s_i), t_j, g(s_i, t_j))$$

is a tautology. $A^{H,D}$ is called a **Herbrand Disjunction**.
Note that the length of this disjunction is fixed: $k \cdot n$.

J. Herbrand's Theorem ('Théorème fondamental', 1930)

Theorem

Let $\boldsymbol{A \equiv \exists x_1 \forall y_1 \exists x_2 \forall y_2 A_{qf}(x_1, y_1, x_2, y_2)}$. Then:

PL $\boldsymbol{\vdash A}$ iff there are terms $s_1, \ldots, s_k, t_1, \ldots, t_n$ (built up out of the constants and variables of A and the **index functions** used for the formation of A^H) such that

$$\mathbf{A^{H,D} :\equiv \bigvee_{i=1}^{k} \bigvee_{j=1}^{n} A_{qf}\left(s_i, f(s_i), t_j, g(s_i, t_j)\right)}$$

is a tautology. $\boldsymbol{A^{H,D}}$ is called a **Herbrand Disjunction**.

Note that the length of this disjunction is fixed: $k \cdot n$. The terms s_i, t_j can be extracted from a given PL-proof of A.

Herbrand's Theorem continued

Replacing in $A^{H,D}$ all terms '$g(s_i, t_j)$', '$f(s_i)$', by new variables (treating larger terms first) results in another tautological disjunction A^{Dis} s.t. A can be inferred from A by a **direct proof**.

Remark

- For sentences $A \equiv \forall x \exists y \forall z\, A_{qf}(x, y, z)$, A^{Dis} can be written in the form

$$A_{qf}(x, t_1, b_1) \vee A_{qf}(x, t_2, b_2) \vee \ldots \vee A_{qf}(x, t_k, b_k),$$

where the b_i are new variables and t_i does not contain any b_j with $i \leq j$ (used by Luckhardt's analysis of Roth's theorem, see below).

Remark

- For sentences $A \equiv \forall x \exists y \forall z\, A_{qf}(x, y, z)$, A^{Dis} can be written in the form

$$A_{qf}(x, t_1, b_1) \vee A_{qf}(x, t_2, b_2) \vee \ldots \vee A_{qf}(x, t_k, b_k),$$

  where the b_i are new variables and t_i does not contain any b_j with $i \leq j$ (used by Luckhardt's analysis of Roth's theorem, see below).
- Herbrand's theorem immediately extends to first-order theories $\mathcal{T}$ whose non-logical axioms $G_1, \ldots, G_n$ are all purely universal.

Theorem (Roth 1955)

An algebraic irrational number α has only finitely many exceptionally good rational approximations, i.e. for $\varepsilon > 0$ there are only finitely many $q \in \mathbb{N}$ such that

$$\mathbf{R(q) :\equiv q > 1 \wedge \exists! p \in \mathbb{Z} : (p, q) = 1 \wedge |\alpha - pq^{-1}| < q^{-2-\varepsilon}.}$$

Theorem (Roth 1955)

An algebraic irrational number α has only finitely many exceptionally good rational approximations, i.e. for $\varepsilon > 0$ there are only finitely many $q \in \mathbb{N}$ such that

$$\mathbf{R(q) :\equiv q > 1 \wedge \exists! p \in \mathbb{Z} : (p, q) = 1 \wedge |\alpha - pq^{-1}| < q^{-2-\varepsilon}.}$$

Theorem (Luckhardt 1985/89)

The following upper bound on $\#\{q : R(q)\}$ holds:

$$\#\{q : R(q)\} < \frac{7}{3}\varepsilon^{-1} \log N_\alpha + 6 \cdot 10^3 \varepsilon^{-5} \log^2 d \cdot \log(50\varepsilon^{-2} \log d),$$

where $N_\alpha < \max(21 \log 2h(\alpha), 2 \log(1 + |\alpha|))$ and h is the logarithmic absolute homogeneous height and $d = deg(\alpha)$.

Independently: Bombieri and van der Poorten 1988.

Exercise (U. Berger)

Consider open theory $\mathcal{T} := \{\forall x(S(x) \neq 0)\}$ in language with equality, constant 0 and two unary function symbols S, f.

Exercise (U. Berger)

Consider open theory $\mathcal{T} := \{\forall x(S(x) \neq 0)\}$ in language with equality, constant 0 and two unary function symbols S, f.

Proposition

$\mathcal{T} \vdash \exists x\big(f(S(f(x))) \neq x\big)$.

Proof: Suppose that

$$\forall x\big(f(S(f(x))) = x\big),$$

then f is injective, but also (since $S(x) \neq 0$) surjective on $\{x : x \neq 0\}$ and hence non-injective. Contradiction! □

Exercise (U. Berger)

Consider open theory $\mathcal{T} := \{\forall x(S(x) \neq 0)\}$ in language with equality, constant 0 and two unary function symbols S, f.

Proposition

$\mathcal{T} \vdash \exists x\big(f(S(f(x))) \neq x\big)$.

Proof: Suppose that

$$\forall x\big(f(S(f(x))) = x\big),$$

then f is injective, but also (since $S(x) \neq 0$) surjective on $\{x : x \neq 0\}$ and hence non-injective. Contradiction! □

Exercise 1: Analyze the above proof to extract Herbrand terms $s_1, \ldots, s_k, t_1, \ldots, t_n$ s.t.

$$\mathrm{PL} \vdash \Big(\bigwedge_{i=1}^{k} S(s_i) \neq 0\Big) \to \bigvee_{j=1}^{n} \big(f(S(f(t_j))) \neq t_j\big).$$

Limitations

- Techniques work only for restricted formal contexts: mainly purely universal ('algebraic') axioms, restricted use of induction, no higher analytical principles.

Limitations

- Techniques work only for restricted formal contexts: mainly purely universal ('algebraic') axioms, restricted use of induction, no higher analytical principles.
- Require that one can 'guess' the correct Herbrand terms: in general procedure results in proofs of length $2_n^{|P|}$, where $2_{n+1}^k = 2^{2_n^k}$ (n cut complexity).

Towards generalizations of Herbrand's theorem

Allow **functionals $\Phi(x, f)$** instead of just Herbrand terms: Let's consider again the example

$$\mathbf{A} \equiv \forall \mathbf{x} \exists \mathbf{y} \forall \mathbf{z} \big(\mathbf{T(x, x, y)} \vee \neg \mathbf{T(x, x, z)} \big) \big).$$

Towards generalizations of Herbrand's theorem

Allow **functionals $\Phi(x, f)$** instead of just Herbrand terms: Let's consider again the example

$$\mathbf{A} \equiv \forall \mathbf{x} \exists \mathbf{y} \forall \mathbf{z} \big(\mathbf{T}(\mathbf{x}, \mathbf{x}, \mathbf{y}) \vee \neg \mathbf{T}(\mathbf{x}, \mathbf{x}, \mathbf{z}) \big) \big).$$

A^H can be realized by a computable functional of type level 2 which is defined by cases:

$$\Phi(\mathbf{x}, \mathbf{g}) := \begin{cases} c & \text{if } \neg T(x, x, g(c)) \\ g(c) & \text{otherwise.} \end{cases}$$

Towards generalizations of Herbrand's theorem

Allow **functionals $\Phi(x, f)$** instead of just Herbrand terms: Let's consider again the example

$$\mathbf{A} \equiv \forall x \exists y \forall z \big(\mathbf{T}(x, x, y) \vee \neg \mathbf{T}(x, x, z))\big).$$

A^H can be realized by a computable functional of type level 2 which is defined by cases:

$$\Phi(x, g) := \begin{cases} c & \text{if } \neg T(x, x, g(c)) \\ g(c) & \text{otherwise.} \end{cases}$$

From this definition it easily follows that

$$\forall x, g \big(\mathbf{T}(x, x, \Phi(x, g)) \vee \neg \mathbf{T}(x, x, g(\Phi(x, g)))\big).$$

Φ satisfies **G. Kreisel's no-counterexample interpretation!**

If A is not provable in PL but e.g. in PA more **complicated functionals** are needed (Kreisel 1951):

If A is not provable in PL but e.g. in PA more **complicated functionals** are needed (Kreisel 1951):

Let (a_n) be a nonincreasing sequence in $[0,1]$. Then, clearly, (a_n) is convergent and so a Cauchy sequence which we write as:

$$\mathbf{(1)}\ \forall \mathbf{k} \in \mathbb{N} \exists \mathbf{n} \in \mathbb{N} \forall \mathbf{m} \in \mathbb{N} \forall \mathbf{i,j} \in \mathbf{[n;n+m]}\ (|\mathbf{a_i} - \mathbf{a_j}| \leq \mathbf{2^{-k}}),$$

where $[n; n+m] := \{n, n+1, \ldots, n+m\}$.

If A is not provable in PL but e.g. in PA more **complicated functionals** are needed (Kreisel 1951):

Let (a_n) be a nonincreasing sequence in $[0,1]$. Then, clearly, (a_n) is convergent and so a Cauchy sequence which we write as:

$$\mathbf{(1)}\ \forall \mathbf{k} \in \mathbb{N} \exists \mathbf{n} \in \mathbb{N} \forall \mathbf{m} \in \mathbb{N} \forall \mathbf{i,j} \in \mathbf{[n;n+m]}\ (|\mathbf{a_i} - \mathbf{a_j}| \leq \mathbf{2^{-k}}),$$

where $[n; n+m] := \{n, n+1, \ldots, n+m\}$.
Then the (partial) Herbrand normal form of this statement is

$$\mathbf{(2)}\ \forall \mathbf{k} \in \mathbb{N} \forall \mathbf{g} \in \mathbb{N}^{\mathbb{N}} \exists \mathbf{n} \in \mathbb{N} \forall \mathbf{i,j} \in \mathbf{[n;n+g(n)]}\ (|\mathbf{a_i} - \mathbf{a_j}| \leq \mathbf{2^{-k}}).$$

By E. Specker 1949 there exist **computable** such sequences (a_n) even in $\mathbb{Q} \cap [0,1]$ **without computable bound** on '$\exists n$' in (1).

By E. Specker 1949 there exist **computable** such sequences (a_n) even in $\mathbb{Q} \cap [0,1]$ **without computable bound** on '$\exists n$' in (1).

By contrast, there is a **simple (primitive recursive) bound** $\Phi^*(g,k)$ on (2) (also referred to as **'metastability'** by T.Tao):

By E. Specker 1949 there exist **computable** such sequences (a_n) even in $\mathbb{Q} \cap [0,1]$ **without computable bound** on '$\exists n$' in (1).

By contrast, there is a **simple (primitive recursive) bound** $\Phi^*(g,k)$ on (2) (also referred to as **'metastability'** by T.Tao):

Proposition

Let (a_n) be any nonincreasing sequence in $[0,1]$ then

$$\forall k \in \mathbb{N} \forall g \in \mathbb{N}^{\mathbb{N}} \exists n \leq \Phi^*(g,k) \forall i,j \in [n; n+g(n)] \, (|a_i - a_j| \leq 2^{-k}),$$

where

$$\Phi^*(g,k) := \tilde{g}^{(2^k-1)}(0) \text{ with } \tilde{g}(n) := n + g(n).$$

Moreover, there exists an $i < 2^k$ such that n can be taken as $\tilde{g}^{(i)}(0)$.

Remark

The previous result can be viewed as a polished form of a **Herbrand disjunction** of **variable (in k) length**:

$$\bigvee_{i=0}^{2^k-1} \left(|a_{\tilde{g}^{(i)}(0)} - a_{\tilde{g}(\tilde{g}^{(i)}(0))}| \leq 2^{-k}\right).$$

Remark

The previous result can be viewed as a polished form of a **Herbrand disjunction** of **variable (in k) length**:

$$\bigvee_{i=0}^{2^k-1} \left(|a_{\tilde{g}^{(i)}(0)} - a_{\tilde{g}(\tilde{g}^{(i)}(0))}| \leq 2^{-k}\right).$$

Corollary (T. Tao's finite convergence principle)

$$\forall k \in \mathbb{N}, g : \mathbb{N} \to \mathbb{N} \exists M \in \mathbb{N} \forall 1 \geq a_0 \geq \ldots \geq a_M \geq 0 \exists N \in \mathbb{N}$$
$$\left(N + g(N) \leq M \wedge \forall n, m \in [N, N + g(N)](|a_n - a_m| \leq 2^{-k})\right).$$

One may take $M := \tilde{g}^{(2^k)}(0)$.

The Mean Ergodic Theorem

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

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

The Mean Ergodic Theorem

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

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

Theorem (von Neumann Mean Ergodic Theorem)

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

The Mean Ergodic Theorem

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

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

Theorem (von Neumann Mean Ergodic Theorem)

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

Avigad/Gerhardy/Towsner (TAMS 2010):

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

But: **Prim. rec. bound on metastable version!**

The Mean Ergodic Theorem

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

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

Theorem (von Neumann Mean Ergodic Theorem)

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

Avigad/Gerhardy/Towsner (TAMS 2010):

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

But: **Prim. rec. bound on metastable version!**

Theorem (Garrett Birkhoff 1939)

Mean Ergodic Theorem holds for uniformly convex Banach spaces.

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

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

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

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

where

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

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

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

where

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

Special Hilbert case: treated prior by Avigad/Gerhardy/Towsner TAMS 2010 (again based on logical metatheorem).

Some history:

- 2004, first rate of metastability (for the asymptotic regularity of asymptotically nonexpansive mappings) extracted by Kohlenbach/Lambov).
- 2005, rate of metastability for Krasnoselski-Mann iterations of nonexpansive selfmappings $T : X \to X$ of compact hyperbolic spaces X (Kohlenbach).
- 2007, Tao introduced the term 'metastability' in connection with the von Neumann Mean Ergodic Theorem (MET).
- 2007, independently from Tao, the first rate of asymptotic regularity for MET was extracted by Avigad/Gerhardy/Towsner.
- 2008, Kohlenbach/Leustean generalized this with a better bound to uniformly convex Banach spaces which, subsequently, led to oscillation bounds by Avigad/Rute (see below).
- Since then, numerous papers extracting explicit rates of metastability have appeared.

Tao also established (without bound) a uniform version (in a special case) of the Mean Ergodic Theorem as base step for a generalization to commuting families of operators.

Tao also established (without bound) a uniform version (in a special case) of the Mean Ergodic Theorem as base step for a generalization to commuting families of operators.

'We shall establish Theorem 1.6 by "finitary ergodic theory" techniques, reminiscent of those used in [Green-Tao]...' 'The main advantage of working in the finitary setting ... is that the underlying dynamical system becomes extremely explicit'...'In proof theory, this finitisation is known as Gödel functional interpretation...which is also closely related to the Kreisel no-counterexample interpretation'
(T. Tao: Norm convergence of multiple ergodic averages for commuting transformations, Ergodic Theor. and Dynam. Syst. 28, 2008)

No-Counterexample Interpretation (Kreisel 1951)

Recall: for a formula

$$A \equiv \exists x_1 \forall y_1 \ldots \exists x_n \forall y_n \, A_{qf}(x_1, y_1, \ldots, x_n, y_n)$$

we call a tuple of functionals $\underline{\varphi}$ a **solution for the no-counterexample interpretation of** A if $\underline{\varphi}$ provides a witness for A^H

$$\forall f_1, \ldots, f_n \, \exists x_1, \ldots, x_n \, A_{qf}(x_1, f_1(x_1), \ldots, x_n, f_n(x_1, \ldots, x_n)),$$

No-Counterexample Interpretation (Kreisel 1951)

Recall: for a formula

$$A \equiv \exists x_1 \forall y_1 \ldots \exists x_n \forall y_n \, A_{qf}(x_1, y_1, \ldots, x_n, y_n)$$

we call a tuple of functionals $\underline{\varphi}$ a **solution for the no-counterexample interpretation of** A if $\underline{\varphi}$ provides a witness for A^H

$$\forall f_1, \ldots, f_n \, \exists x_1, \ldots, x_n \, A_{qf}(x_1, f_1(x_1), \ldots, x_n, f_n(x_1, \ldots, x_n)),$$

i.e.

$$\forall \underline{f} \, A_{qf}(\varphi_1(\underline{f}), f_1(\varphi_1(\underline{f}), \ldots, \varphi_n(\underline{f}), f_n(\varphi_1(\underline{f}), \ldots, \varphi_n(\underline{f}))).$$

Problems of the no-counterexample interpretation

For principles $F \in \exists\forall\exists$ n.c.i. no longer 'correct'. $C_n := \{0, 1, \ldots, n\}$.

Problems of the no-counterexample interpretation

For principles $F \in \exists\forall\exists$ n.c.i. no longer 'correct'. $C_n := \{0, 1, \ldots, n\}$.

Direct example: Infinitary Pigeonhole Principle (IPP):

$$\forall n \in \mathbb{N}\forall f : \mathbb{N} \to C_n \exists i \leq n \forall k \in \mathbb{N}\exists m \geq k \left(f(m) = i\right).$$

Problems of the no-counterexample interpretation

For principles $F \in \exists\forall\exists$ n.c.i. no longer 'correct'. $C_n := \{0, 1, \ldots, n\}$.

Direct example: Infinitary Pigeonhole Principle (IPP):

$$\forall n \in \mathbb{N}\forall f : \mathbb{N} \to C_n \exists i \leq n \forall k \in \mathbb{N}\exists m \geq k \left(f(m) = i\right).$$

IPP causes arbitrary **primitive recursive complexity**, but $(\mathrm{IPP})^H$

$$\forall n \in \mathbb{N}\forall f : \mathbb{N} \to C_n \forall F : C_n \to \mathbb{N}\exists i \leq n \exists m \geq F(i) \left(f(m) = i\right)$$

Problems of the no-counterexample interpretation

For principles $F \in \exists\forall\exists$ n.c.i. no longer 'correct'. $C_n := \{0, 1, \ldots, n\}$.

Direct example: Infinitary Pigeonhole Principle (IPP):

$$\forall n \in \mathbb{N}\forall f : \mathbb{N} \to C_n \exists i \leq n \forall k \in \mathbb{N}\exists m \geq k \left(f(m) = i\right).$$

IPP causes arbitrary **primitive recursive complexity**, but $(\mathrm{IPP})^H$

$$\forall n \in \mathbb{N}\forall f : \mathbb{N} \to C_n \forall F : C_n \to \mathbb{N}\exists i \leq n \exists m \geq F(i) \left(f(m) = i\right)$$

has trivial n.c.i.-solution for '$\exists i$','$\exists m$':

$M(n, f, F) := \max\{F(i) : i \leq n\}$ and $I(n, f, F) := f(M(n, f, F))$.

Problems of the no-counterexample interpretation

For principles $F \in \exists\forall\exists$ n.c.i. no longer 'correct'. $C_n := \{0, 1, \ldots, n\}$.

Direct example: Infinitary Pigeonhole Principle (IPP):

$$\forall n \in \mathbb{N}\forall f : \mathbb{N} \to C_n \exists i \leq n \forall k \in \mathbb{N}\exists m \geq k \, (f(m) = i).$$

IPP causes arbitrary **primitive recursive complexity**, but $(\mathrm{IPP})^H$

$$\forall n \in \mathbb{N}\forall f : \mathbb{N} \to C_n \forall F : C_n \to \mathbb{N}\exists i \leq n \exists m \geq F(i) \, (f(m) = i)$$

has trivial n.c.i.-solution for '$\exists i$','$\exists m$':

$M(n, f, F) := \max\{F(i) : i \leq n\}$ and $I(n, f, F) := f(M(n, f, F))$.

M, I **do not reflect** true complexity of IPP!

Problems of the no-counterexample interpretation

For principles $F \in \exists\forall\exists$ n.c.i. no longer 'correct'. $C_n := \{0, 1, \ldots, n\}$.

Direct example: Infinitary Pigeonhole Principle (IPP):

$$\forall n \in \mathbb{N}\forall f : \mathbb{N} \to C_n \exists i \leq n \forall k \in \mathbb{N}\exists m \geq k \, (f(m) = i).$$

IPP causes arbitrary **primitive recursive complexity**, but $(\text{IPP})^H$

$$\forall n \in \mathbb{N}\forall f : \mathbb{N} \to C_n \forall F : C_n \to \mathbb{N}\exists i \leq n \exists m \geq F(i) \, (f(m) = i)$$

has trivial n.c.i.-solution for '$\exists i$','$\exists m$':

$M(n, f, F) := \max\{F(i) : i \leq n\}$ and $I(n, f, F) := f(M(n, f, F))$.

M, I **do not reflect** true complexity of IPP!

Related problem: bad behavior w.r.t. modus ponens!

A Modular Approach: Proof Interpretations

- **Interpret** the formulas A in $P : A \mapsto A^{\mathcal{I}}$,

A Modular Approach: Proof Interpretations

- **Interpret** the formulas A in P : $A \mapsto A^{\mathcal{I}}$,
- Interpretation $C^{\mathcal{I}}$ contains the **additional information**,

A Modular Approach: Proof Interpretations

- **Interpret** the formulas A in $P : A \mapsto A^{\mathcal{I}}$,
- Interpretation $C^{\mathcal{I}}$ contains the **additional information**,
- Construct by **recursion on** P a new proof $P^{\mathcal{I}}$ of $C^{\mathcal{I}}$.

A Modular Approach: Proof Interpretations

- **Interpret** the formulas A in $P : A \mapsto A^{\mathcal{I}}$,
- Interpretation $C^{\mathcal{I}}$ contains the **additional information**,
- Construct by **recursion on** P a new proof $P^{\mathcal{I}}$ of $C^{\mathcal{I}}$.
  In particular: solve **modus ponens problem:**

$$\frac{A^I \quad , \quad (A \to B)^I}{B^I}.$$

A Modular Approach: Proof Interpretations

- **Interpret** the formulas A in $P : A \mapsto A^{\mathcal{I}}$,
- Interpretation $C^{\mathcal{I}}$ contains the **additional information**,
- Construct by **recursion on** P a new proof $P^{\mathcal{I}}$ of $C^{\mathcal{I}}$.
  In particular: solve **modus ponens problem:**

$$\frac{A^I \quad , \quad (A \to B)^I}{B^I}.$$

Our approach is based on novel forms and extensions of:

K. Gödel's functional interpretation!

Literature

1) Delzell, C.N., A continuous, constructive solution to Hilbert's 17th problem. Invent. math. **76**, pp. 365-384 (1984).

2) Delzell, C.N., Kreisel's unwinding of Artin's proof-Part I. In: Odifreddi, P., Kreiseliana, pp. 113-246, A K Peters, Wellesley, MA, 1996.

3) Gaspar, J., Kohlenbach, U., On Tao's 'finitary' infinite pigeonhole principle. J. Symb. Logic **75**, pp. 355-371 (2010).

4) Gerhardy, P., Kohlenbach, U., Extracting Herbrand disjunctions by functional interpretation. Arch. Math. Logic. vol. 44, pp. 633-644 (2005).

5) Girard, J.-Y., Proof Theory and Logical Complexity Vol.I.

Studies in Proof Theory. Bibliopolis (Napoli) and Elsevier Science Publishers (Amsterdam) 1987.

6) Kohlenbach, U., Applied Proof Theory: Proof Interpretations and their Use in Mathematics. Springer Monographs in Mathematics. xx+536pp., Springer Heidelberg-Berlin, 2008.

7) Kohlenbach, U., On the no-counterexample interpretation. J. Symbolic Logic **64**, pp. 1491-1511 (1999).

8) Kreisel, G., Finiteness theorems in arithmetic: an application of Herbrand's theorem for Σ_2-formulas. Proc. of the Herbrand symposium (Marseille, 1981), North-Holland (Amsterdam), pp. 39-55 (1982).

9) Kreisel, G., On the interpretation of non-finitist proofs, part I. J. Symbolic Logic **16**, pp.241-267 (1951).

10) Kreisel, G., On the interpretation of non-finitist proofs, part II: Interpretation of number theory, applications. J. Symbolic Logic **17**, pp. 43-58 (1952).

11) Kreisel, G., Macintyre, A., Constructive logic versus algebraization I. In: Troelstra, A.S., van Dalen, D. (eds.), Proc. L.E.J. Brouwer Centenary Symposium (Noordwijkerhout 1981), North-Holland (Amsterdam), pp. 217-260 (1982).

12) Luckhardt, H., Herbrand-Analysen zweier Beweise des Satzes von Roth: Polynomiale Anzahlschranken. J. Symb. Logic **54**, pp. 234-263 (1989).

13) Macintyre, A., The mathematical significance of proof theory. Phil. Trans. R. Soc. A **363**, pp. 2419-2435 (2005).

14) Sipoș, A., On extracting variable Herbrand disjunctions. Studia Logica **110**, pp. 1115-1134 (2022).

15) Tao, T., Soft analysis, hard analysis, and the finite convergence principle. In: 'Structure and Randomness. AMS, 298pp., 2008'.

16) Troelstra, A.S. (ed.) Metamathematical investigation of intuitionistic arithmetic and analysis. Springer LNM **344** (1973).

Lecture II

Detour through intuitionistic systems and higher types

Detour through intuitionistic systems and higher types

- **HA** ('Heyting arithmetic' is defined as Peano arithmetic but with **intuitionistic (constructive) logic**.

Detour through intuitionistic systems and higher types

- **HA** ('Heyting arithmetic' is defined as Peano arithmetic but with **intuitionistic (constructive) logic**.
- **(WE-)HA$^\omega$** is the extension of **HA** to all finite types over $\mathbb{N}$.

Detour through intuitionistic systems and higher types

- **HA** ('Heyting arithmetic' is defined as Peano arithmetic but with **intuitionistic (constructive) logic**.
- **(WE-)HA$^\omega$** is the extension of **HA** to all finite types over $\mathbb{N}$.

Types T: (i) $\mathbb{N} \in \mathsf{T}$, $\quad \rho, \tau \in \mathsf{T} \Rightarrow (\rho \to \tau) \in \mathsf{T}$.

Detour through intuitionistic systems and higher types

- **HA** ('Heyting arithmetic' is defined as Peano arithmetic but with **intuitionistic (constructive) logic**.
- **(WE-)HA$^\omega$** is the extension of **HA** to all finite types over $\mathbb{N}$.

Types T: (i) $\mathbb{N} \in \mathsf{T}$, $\quad \rho, \tau \in \mathsf{T} \Rightarrow (\rho \to \tau) \in \mathsf{T}$.

HA$^\omega$ has λ-abstraction $(\lambda x^\rho . t[x]^\tau)(s^\rho) =_\tau t[s/x]$ and **primitive recursion in all finite types** (Hilbert 1926, Gödel 1958): for $x \in \mathbb{N}$

$$R_\rho(0, y, z) =_\rho y, \; R_\rho(x+1, y, z) =_\rho z(R_\rho xyz, x),$$

where $=_\rho$ is defined as pointwise (extensional) equality.

Detour through intuitionistic systems and higher types

- **HA** ('Heyting arithmetic' is defined as Peano arithmetic but with **intuitionistic (constructive) logic**.
- **(WE-)HA$^\omega$** is the extension of **HA** to all finite types over $\mathbb{N}$.

Types T: (i) $\mathbb{N} \in \mathsf{T}$, $\quad \rho, \tau \in \mathsf{T} \Rightarrow (\rho \to \tau) \in \mathsf{T}$.

HA$^\omega$ has λ-abstraction $(\lambda x^\rho . t[x]^\tau)(s^\rho) =_\tau t[s/x]$ and **primitive recursion in all finite types** (Hilbert 1926, Gödel 1958): for $x \in \mathbb{N}$

$$R_\rho(0, y, z) =_\rho y, \; R_\rho(x+1, y, z) =_\rho z(R_\rho xyz, x),$$

where $=_\rho$ is defined as pointwise (extensional) equality.

PA$^\omega$ $=$**HA$^\omega$** $+ (A \vee \neg A)$.

Detour through intuitionistic systems and higher types

- **HA** ('Heyting arithmetic' is defined as Peano arithmetic but with **intuitionistic (constructive) logic**.
- **(WE-)HA$^\omega$** is the extension of **HA** to all finite types over $\mathbb{N}$.

Types T: (i) $\mathbb{N} \in \mathsf{T}$, $\quad \rho, \tau \in \mathsf{T} \Rightarrow (\rho \to \tau) \in \mathsf{T}$.

HA$^\omega$ has λ-abstraction $(\lambda x^\rho.t[x]^\tau)(s^\rho) =_\tau t[s/x]$ and **primitive recursion in all finite types** (Hilbert 1926, Gödel 1958): for $x \in \mathbb{N}$

$$R_\rho(0, y, z) =_\rho y, \; R_\rho(x+1, y, z) =_\rho z(R_\rho xyz, x),$$

where $=_\rho$ is defined as pointwise (extensional) equality.

PA$^\omega$ $=$**HA$^\omega$** $+ (A \vee \neg A)$.

Only **quantifier-free rule of extensionality** (A_{qf} quantifier-free)

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

$\mathcal{S}^\omega$ full set-theoretic model of **PA$^\omega$**.

In 2007 is het 100 jaar geleden dat L.E.J. Brouwer (1881 – 1966) de stelling van Aristoteles verwierp. Brouwer vond dat een wiskundige stelling pas waar is als er 'Positief Bewijs' is. Brouwer is de grondlegger van de intuïtionistische wiskunde. Naar hem is o.a. de dekpuntstelling van Brouwer vernoemd. Iedere drie jaar reikt het Koninklijk Wiskundig Genootschap de Brouwer medaille uit aan een belangrijk wiskundige. Voor meer informatie: www.knaw.nl
NEDERLAND
€0,44
$\nvdash A \vee \neg A$
Er is positief bewijs!
100 jaar na dato wordt de wiskundige
L.E.J. Brouwer (1881-1966)
geëerd met een eigen postzegel.
TNT post

Proof mining in the intuitionistic case: modified realizability (Kreisel 1959)

To each $A \in \mathcal{L}(\mathrm{HA}^\omega)$ we assign a new formula $\underline{x}\ mr\ A$ ('$\underline{x}$ modified realizes A') inductively by

Proof mining in the intuitionistic case: modified realizability (Kreisel 1959)

To each $A \in \mathcal{L}(\mathrm{HA}^\omega)$ we assign a new formula $\underline{x}\ mr\ A$ ('$\underline{x}$ modified realizes A') inductively by

(i) $\underline{x}\ mr\ A :\equiv A$ with the empty tuple $\underline{x}$, if A is a prime formula.

(ii) $\underline{x}, \underline{y}\ mr\ (A \wedge B) :\equiv \underline{x}\ mr\ A \wedge \underline{y}\ mr\ B$.

(iii) $z^{\mathbb{N}}, \underline{x}, \underline{y}\ mr\ (A \vee B) :\equiv [(z = 0 \to \underline{x}\ mr\ A) \wedge (z \neq 0 \to \underline{y}\ mr\ B)]$.

(iv) $\underline{y}\ mr\ (A \to B) :\equiv \forall \underline{x}(\underline{x}\ mr\ A \to \underline{y}\,\underline{x}\ mr\ B)$.

(v) $\underline{x}\ mr\ (\forall y^\rho A(y)) :\equiv \forall y^\rho(\underline{x}y\ mr\ A(y))$.

(vi) $z^\rho, \underline{x}\ mr\ (\exists y^\rho A(y)) :\equiv \underline{x}\ mr\ A(z)$.

Program extraction by modified realizability

Axiom of choice (in all types)

$$\mathrm{AC}: \quad \forall a^{\alpha}\, \exists b^{\beta}\, F(a,b) \to \exists B^{\rho\to\tau}\, \forall a^{\rho}\, F(a, B(a)).$$

Theorem

From a proof of

$$\mathrm{HA}^{\omega}+\mathrm{AC} \vdash \forall x^{\rho}\, (\neg B(x) \to \exists y^{\tau}\, A(x,y))$$

one can extract by *mr* a primitive recursive functional Φ s.t.

$$\mathcal{S}^{\omega} \models \forall x^{\rho}\, (\neg B(x) \to A(x, \Phi(x)))$$

(A, B, ρ, τ arbitrary).

Towards proofs based on classical logic

Negative Translation N: Translates $\mathbf{PA}^{\omega}$ into $\mathbf{HA}^{\omega}$.

Towards proofs based on classical logic

Negative Translation N: Translates $\mathbf{PA}^\omega$ into $\mathbf{HA}^\omega$.

In particular (for A_{qf} quantifier-free):

$$\mathbf{PA}^\omega \vdash \forall x^\rho \exists y^\tau A_{qf}(x,y) \Rightarrow \mathbf{HA}^\omega \vdash \forall x^\rho \neg\neg y^\tau A_{qf}(x,y).$$

Entrance door for classical logic: Markov's principle $\mathbf{M}^\omega$!

$$\mathbf{M}^\omega: \neg\neg\exists x^\rho A_{qf}(x) \to \exists x^\rho A_{qf}(x), \quad A_{qf} \text{ quantifier-free.}$$

Towards proofs based on classical logic

Negative Translation N: Translates $\mathbf{PA}^\omega$ into $\mathbf{HA}^\omega$.
In particular (for A_{qf} quantifier-free):

$$\mathbf{PA}^\omega \vdash \forall x^\rho \exists y^\tau A_{qf}(x,y) \Rightarrow \mathbf{HA}^\omega \vdash \forall x^\rho \neg\neg y^\tau A_{qf}(x,y).$$

Entrance door for classical logic: Markov's principle $\mathbf{M}^\omega$!

$$\mathbf{M}^\omega: \neg\neg\exists x^\rho A_{qf}(x) \to \exists x^\rho A_{qf}(x), \quad A_{qf} \text{ quantifier-free.}$$

For $\rho = \mathbb{N}$, this has a partial computable solution by unbounded search (no complexity information, Kleene realizability), but no total computable solution of its modified realizability *mr*!

Towards proofs based on classical logic

Negative Translation N: Translates $\mathbf{PA}^\omega$ into $\mathbf{HA}^\omega$.
In particular (for A_{qf} quantifier-free):

$$\mathbf{PA}^\omega \vdash \forall x^\rho \exists y^\tau A_{qf}(x,y) \Rightarrow \mathbf{HA}^\omega \vdash \forall x^\rho \neg\neg y^\tau A_{qf}(x,y).$$

Entrance door for classical logic: Markov's principle $\mathbf{M}^\omega$!

$$\mathbf{M}^\omega: \neg\neg\exists x^\rho A_{qf}(x) \to \exists x^\rho A_{qf}(x), \quad A_{qf} \text{ quantifier-free.}$$

For $\rho = \mathbb{N}$, this has a partial computable solution by unbounded search (no complexity information, Kleene realizability), but no total computable solution of its modified realizability *mr*!

For $\rho \neq \mathbb{N}$: not even unbounded search possible!

Gödel's functional ('Dialectica') interpretation D (Gödel 1941, 1958)

Solution: Don't try to solve $\mathbf{M}^{\omega}$ but eliminate it from proofs!

Gödel's functional ('Dialectica') interpretation D (Gödel 1941, 1958)

Solution: Don't try to solve $\mathbf{M}^\omega$ but eliminate it from proofs!

Gödel's D assigns to each A a formula $A^D \equiv \exists \underline{x} \forall \underline{y}\, A_D(\underline{x}, \underline{y})$ (A_D qf).

Gödel's functional ('Dialectica') interpretation D (Gödel 1941, 1958)

Solution: Don't try to solve $\mathbf{M}^\omega$ but eliminate it from proofs!

Gödel's D assigns to each A a formula $A^D \equiv \exists \underline{x}\, \forall \underline{y}\, A_D(\underline{x}, \underline{y})$ (A_D qf).

Interpretation differs from mr for the clause of '$\to$'

$$(A \to B)^D :\equiv \exists \underline{U}\, \underline{Y}\, \forall \underline{x}\, \underline{v}\, \big(\underbrace{A_D(\underline{x}, \underline{Y}\,\underline{x}\,\underline{v}) \to B_D(\underline{U}\underline{x}, \underline{v})}_{(A \to B)_D :\equiv} \big).$$

Gödel's functional ('Dialectica') interpretation D (Gödel 1941, 1958)

Solution: Don't try to solve $\mathbf{M}^\omega$ but eliminate it from proofs!

Gödel's D assigns to each A a formula $A^D \equiv \exists \underline{x}\, \forall \underline{y}\, A_D(\underline{x}, \underline{y})$ (A_D qf).

Interpretation differs from *mr* for the clause of '$\to$'

$$(A \to B)^D :\equiv \exists \underline{U}\, \underline{Y}\, \forall \underline{x}\, \underline{v}\, \big(\underbrace{A_D(\underline{x}, \underline{Y}\,\underline{x}\,\underline{v}) \to B_D(\underline{U}\underline{x}, \underline{v})}_{(A \to B)_D :\equiv} \big).$$

Then

$$\mathbf{M}^D \equiv \exists x\, \neg\neg A_{qf}(x) \to \exists x\, A_{qf}(x) \quad \text{(trivial since } A_{qf} \text{ decidable)}.$$

Gödel's functional ('Dialectica') interpretation D (Gödel 1941, 1958)

Solution: Don't try to solve $\mathbf{M}^\omega$ but eliminate it from proofs!

Gödel's D assigns to each A a formula $A^D \equiv \exists \underline{x}\, \forall \underline{y}\, A_D(\underline{x}, \underline{y})$ (A_D qf).

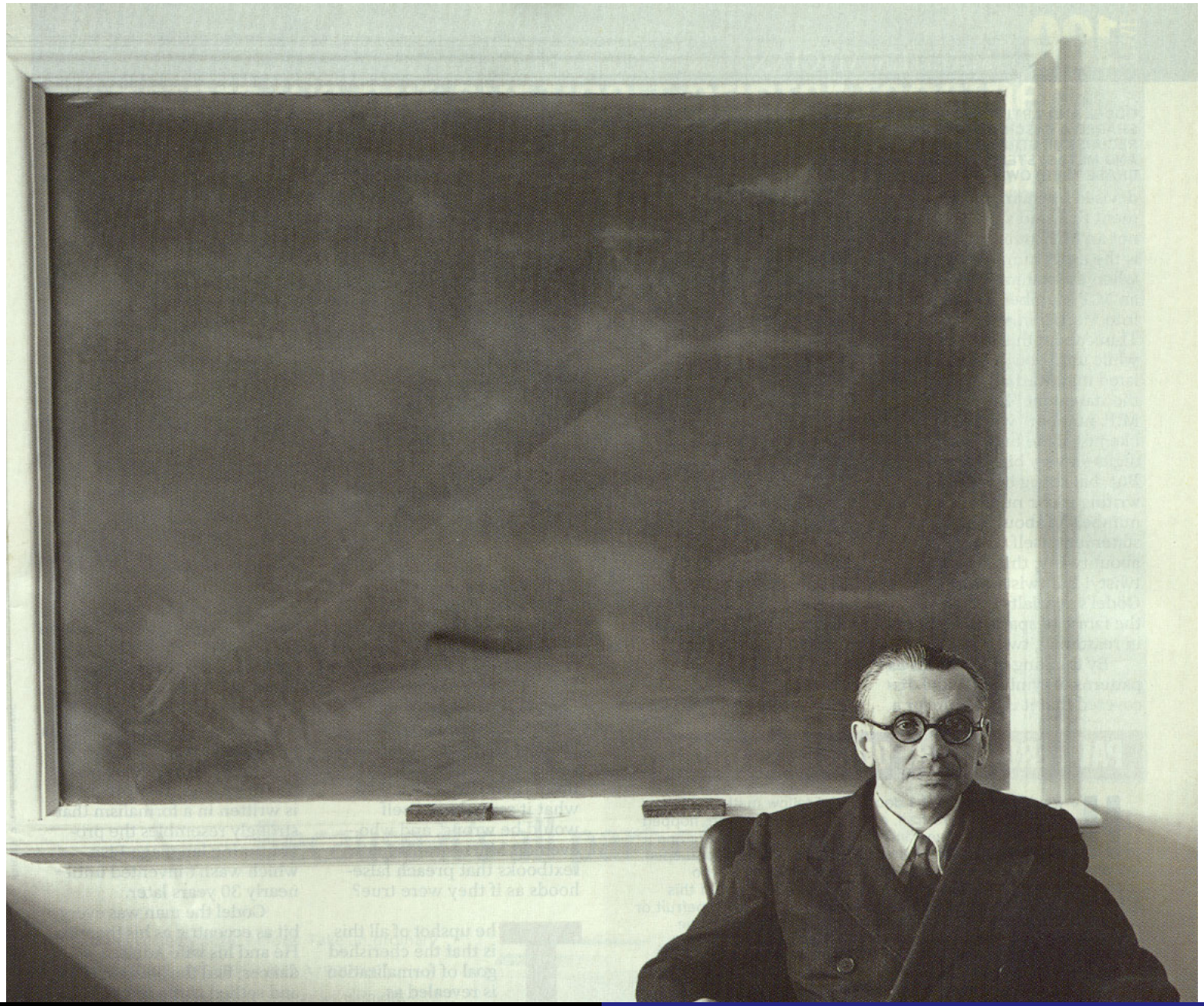
Interpretation differs from *mr* for the clause of '$\to$'

$$(A \to B)^D :\equiv \exists \underline{U}\, \underline{Y}\, \forall \underline{x}\, \underline{v}\, \big(\underbrace{A_D(\underline{x}, \underline{Y}\,\underline{x}\,\underline{v}) \to B_D(\underline{U}\underline{x}, \underline{v})}_{(A \to B)_D :\equiv} \big).$$

Then

$$\mathbf{M}^D \equiv \exists x\, \neg\neg A_{qf}(x) \to \exists x\, A_{qf}(x) \quad \text{(trivial since } A_{qf} \text{ decidable)}.$$

Partial alternative: Friedman-Dragalin A-interpretation after negative translation as intermediate step (variants: Schwichtenberg, Coquand-Hofmann).



Program extraction by D

$$\mathbf{AC}: \ \forall a^{\alpha}\, \exists b^{\beta}\, A(a,b) \to \exists F^{\alpha\to\beta}\, \forall b^{\beta}\, A(a,Fa).$$

Theorem

From a proof of

$$\mathbf{HA}^{\omega}+\mathbf{AC}+\mathbf{M}^{\omega} \ \vdash \forall x^{\rho}\, (\forall u\, B_{qf}(x,u) \to \exists y^{\tau}\, A(x,y))$$

one can extract by D a primitive recursive functionals Φ s.t.

$$\mathcal{S}^{\omega} \models \forall x^{\rho}\, (\forall u\, B_{qf}(x,u) \to \ A(x,\Phi(x)))$$

(A, ρ, τ arbitrary, B_{qf} quantifier-free).

Proofs based on full classical logic

Negative translation N combined with D (i.e. $D \circ N$) gives:

Theorem

From a proof of

$$\mathbf{PA^{\omega}+QF\text{-}AC} \vdash \forall x^{\rho} \left(\forall u\, B_{qf}(x,u) \rightarrow \exists y^{\tau}\, A_{qf}(x,y)\right)$$

one can extract by *Sh* a primitive recursive functionals Φ s.t.

$$\mathcal{S}^{\omega} \models \forall x^{\rho} \left(\forall u\, B_{qf}(x,u) \rightarrow\ A_{qf}(x,\Phi(x))\right)$$

(ρ, τ arbitrary, A_{qf}, B_{qf} quantifier-free, QF-AC restriction of AC to quantifier-free formulas).

Gödel's functional interpretation in five minutes

Gödel's **functional interpretation** D combined with Krivine's **negative translation** N results in an interpretation $Sh = D \circ N$ (Streicher/K.07)

$$A \mapsto A^{Sh} \quad \text{(Shoenfield variant)}$$

such that

Gödel's functional interpretation in five minutes

Gödel's **functional interpretation** D combined with Krivine's **negative translation** N results in an interpretation $Sh = D \circ N$ (Streicher/K.07)

$$A \mapsto A^{Sh} \quad \text{(Shoenfield variant)}$$

such that

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

Gödel's functional interpretation in five minutes

Gödel's **functional interpretation** D combined with Krivine's **negative translation** N results in an interpretation $Sh = D \circ N$ (Streicher/K.07)

$$A \mapsto A^{Sh} \quad \text{(Shoenfield variant)}$$

such that

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

Gödel's functional interpretation in five minutes

Gödel's **functional interpretation** D combined with Krivine's **negative translation** N results in an interpretation $Sh = D \circ N$ (Streicher/K.07)

$$A \mapsto A^{Sh} \quad \text{(Shoenfield variant)}$$

such that

- $A^{Sh} \equiv \forall \underline{x} \exists \underline{y}\, A_{Sh}(\underline{x}, \underline{y})$, where A_{Sh} is **quantifier-free**,
- For $A \equiv \forall \underline{x} \exists \underline{y}\, A_{qf}(\underline{x}, \underline{y})$ one has $A^{Sh} \equiv A$.
- $\underline{x}, \underline{y}$ are tuples of **functionals of finite type** over the base types of the system at hand.

Gödel's functional interpretation in five minutes

Gödel's **functional interpretation** D combined with Krivine's **negative translation** N results in an interpretation $Sh = D \circ N$ (Streicher/K.07)

$$A \mapsto A^{Sh} \quad \text{(Shoenfield variant)}$$

such that

- $A^{Sh} \equiv \forall \underline{x} \exists \underline{y}\, A_{Sh}(\underline{x}, \underline{y})$, where A_{Sh} is **quantifier-free**,
- For $A \equiv \forall \underline{x} \exists \underline{y}\, A_{qf}(\underline{x}, \underline{y})$ one has $A^{Sh} \equiv A$.
- $\underline{x}, \underline{y}$ are tuples of **functionals of finite type** over the base types of the system at hand.
- $\exists Y \forall x\, A_{Sh}(x, Y(x)) \equiv A^{D \circ N}$.

$$A^{Sh} \equiv \forall u \exists x \, A_{Sh}(u, x), \quad B^{Sh} \equiv \forall v \exists y \, B_{Sh}(v, y).$$

$$A^{Sh} \equiv \forall u \exists x\, A_{Sh}(u, x), \quad B^{Sh} \equiv \forall v \exists y\, B_{Sh}(v, y).$$

(Sh1) $P^{Sh} \equiv P \equiv P_{Sh}$ for atomic P

$$A^{Sh} \equiv \forall u \exists x\, A_{Sh}(u, x), \quad B^{Sh} \equiv \forall v \exists y\, B_{Sh}(v, y).$$

(Sh1) $P^{Sh} \equiv P \equiv P_{Sh}$ for atomic P

(Sh2) $(\neg A)^{Sh} \equiv \forall f \exists u\, \neg A_{Sh}(u, f(u))$

$$A^{Sh} \equiv \forall u \exists x\, A_{Sh}(u,x), \quad B^{Sh} \equiv \forall v \exists y\, B_{Sh}(v,y).$$

(Sh1) $P^{Sh} \equiv P \equiv P_{Sh}$ for atomic P

(Sh2) $(\neg A)^{Sh} \equiv \forall f \exists u\, \neg A_{Sh}(u, f(u))$

(Sh3) $(A \vee B)^{Sh} \equiv \forall u, v \exists x, y\, \big(A_{Sh}(u,x) \vee B_{Sh}(v,y)\big)$

$$A^{Sh} \equiv \forall u \exists x\, A_{Sh}(u, x), \quad B^{Sh} \equiv \forall v \exists y\, B_{Sh}(v, y).$$

(Sh1) $P^{Sh} \equiv P \equiv P_{Sh}$ for atomic P

(Sh2) $(\neg A)^{Sh} \equiv \forall f \exists u\, \neg A_{Sh}(u, f(u))$

(Sh3) $(A \vee B)^{Sh} \equiv \forall u, v \exists x, y\, \big(A_{Sh}(u, x) \vee B_{Sh}(v, y)\big)$

(Sh4) $(\forall z\, A)^{Sh} \equiv \forall z, u \exists x\, A_{Sh}(z, u, x)$

$$A^{Sh} \equiv \forall u \exists x\, A_{Sh}(u,x), \quad B^{Sh} \equiv \forall v \exists y\, B_{Sh}(v,y).$$

(Sh1) $P^{Sh} \equiv P \equiv P_{Sh}$ for atomic P

(Sh2) $(\neg A)^{Sh} \equiv \forall f \exists u\, \neg A_{Sh}(u, f(u))$

(Sh3) $(A \vee B)^{Sh} \equiv \forall u, v \exists x, y\, \big(A_{Sh}(u,x) \vee B_{Sh}(v,y)\big)$

(Sh4) $(\forall z\, A)^{Sh} \equiv \forall z, u \exists x\, A_{Sh}(z,u,x)$

(Sh5) $(A \to B)^{Sh} \equiv \forall f, v \exists u, y\, \big(A_{Sh}(u, f(u)) \to B_{Sh}(v,y)\big)$

$$A^{Sh} \equiv \forall u \exists x\, A_{Sh}(u,x), \quad B^{Sh} \equiv \forall v \exists y\, B_{Sh}(v,y).$$

(Sh1) $P^{Sh} \equiv P \equiv P_{Sh}$ for atomic P

(Sh2) $(\neg A)^{Sh} \equiv \forall f \exists u\, \neg A_{Sh}(u, f(u))$

(Sh3) $(A \vee B)^{Sh} \equiv \forall u, v \exists x, y\, \big(A_{Sh}(u,x) \vee B_{Sh}(v,y)\big)$

(Sh4) $(\forall z\, A)^{Sh} \equiv \forall z, u \exists x\, A_{Sh}(z,u,x)$

(Sh5) $(A \to B)^{Sh} \equiv \forall f, v \exists u, y\, \big(A_{Sh}(u, f(u)) \to B_{Sh}(v,y)\big)$

(Sh6) $(\exists z A)^{Sh} \equiv \forall U \exists z, f\, A_{Sh}(z, U(z,f), f(U(z,f)))$

$$A^{Sh} \equiv \forall u \exists x\, A_{Sh}(u,x), \quad B^{Sh} \equiv \forall v \exists y\, B_{Sh}(v,y).$$

(Sh1) $P^{Sh} \equiv P \equiv P_{Sh}$ for atomic P

(Sh2) $(\neg A)^{Sh} \equiv \forall f \exists u\, \neg A_{Sh}(u, f(u))$

(Sh3) $(A \vee B)^{Sh} \equiv \forall u, v \exists x, y\, \big(A_{Sh}(u,x) \vee B_{Sh}(v,y)\big)$

(Sh4) $(\forall z\, A)^{Sh} \equiv \forall z, u \exists x\, A_{Sh}(z,u,x)$

(Sh5) $(A \to B)^{Sh} \equiv \forall f, v \exists u, y\, \big(A_{Sh}(u, f(u)) \to B_{Sh}(v,y)\big)$

(Sh6) $(\exists z A)^{Sh} \equiv \forall U \exists z, f\, A_{Sh}(z, U(z,f), f(U(z,f)))$

(Sh7) $(A \wedge B)^{Sh} \equiv$
$\forall n, u, v \exists x, y\, (n{=}0 \to A_{Sh}(u,x)) \wedge (n{\neq}0 \to B_{Sh}(v,y))$
$\leftrightarrow \forall u, v \exists x, y\, \big(A_{Sh}(u,x) \wedge B_{Sh}(v,y)\big).$

Comments

- The program extraction theorem scales down to weak systems such as RCA_0 (where then Φ is ordinarily prim. rec., Parsons 1971) or of bounded arithmetic (then Φ is basic feasible, Cook/Urquhart 1993).

Comments

- The program extraction theorem scales down to weak systems such as RCA_0 (where then Φ is ordinarily prim. rec., Parsons 1971) or of bounded arithmetic (then Φ is basic feasible, Cook/Urquhart 1993).
- It also scales up all the way to full countable and even dependent choice (including full 2nd order arithmetic), where then Φ is bar recursive: Spector 1962 (Consistency proof for analysis!).

Comments

- The program extraction theorem scales down to weak systems such as RCA_0 (where then Φ is ordinarily prim. rec., Parsons 1971) or of bounded arithmetic (then Φ is basic feasible, Cook/Urquhart 1993).
- It also scales up all the way to full countable and even dependent choice (including full 2nd order arithmetic), where then Φ is bar recursive: Spector 1962 (Consistency proof for analysis!).
- Full extensionality is allowed (since it can be eliminated) if all variables have types ≤ 1.

Connection to no-counterexample interpretation

Let A be a prenex (arithmetical) formula and $A^S, A^{ND}, A^{n.c.i}$ its Skolem, $D \circ N$ and $n.c.i.$ interpretations resp., then

$$\mathbf{HA}^\omega \vdash \boldsymbol{A^S \to A^{ND} \to A^{n.c.i}},$$

but the converse implications in general fail to hold even in PA^ω+QF-AC!

Connection to no-counterexample interpretation

Let A be a prenex (arithmetical) formula and $A^S, A^{ND}, A^{n.c.i}$ its Skolem, $D \circ N$ and $n.c.i.$ interpretations resp., then

$$\mathbf{HA}^\omega \vdash \boldsymbol{A^S \to A^{ND} \to A^{n.c.i}},$$

but the converse implications in general fail to hold even in PA^ω+QF-AC!

- A^S **too strong** (for a computable solution): Specker!

Connection to no-counterexample interpretation

Let A be a prenex (arithmetical) formula and $A^S, A^{ND}, A^{n.c.i}$ its Skolem, $D \circ N$ and $n.c.i.$ interpretations resp., then

$$\mathbf{HA}^\omega \vdash \boldsymbol{A^S \to A^{ND} \to A^{n.c.i}},$$

but the converse implications in general fail to hold even in PA^ω+QF-AC!

- A^S **too strong** (for a computable solution): Specker!
- $A^{n.c.i.}$ **too weak** (see IPP above; modus ponens problem).

Connection to no-counterexample interpretation

Let A be a prenex (arithmetical) formula and $A^S, A^{ND}, A^{n.c.i}$ its Skolem, $D \circ N$ and $n.c.i.$ interpretations resp., then

$$\mathbf{HA}^\omega \vdash A^S \to A^{ND} \to A^{n.c.i},$$

but the converse implications in general fail to hold even in PA^ω+QF-AC!

- A^S **too strong** (for a computable solution): Specker!
- $A^{n.c.i.}$ **too weak** (see IPP above; modus ponens problem).
- A^{ND} **just right**:

$$\mathbf{PA}^\omega\mathbf{+QF\text{-}AC} \vdash A \leftrightarrow A^{ND}.$$

Majorizability

The functionals occurring in functional interpretation have a striking mathematical structure property:

Definition (W.A. Howard 1973)

$$\begin{cases} x^* \gtrsim_{\mathbb{N}} x :\equiv x^* \geq x, \\ x^* \gtrsim_{\rho\to\tau} x :\equiv \forall y^*, y(y^* \gtrsim_\rho y \to x^*(y^*) \gtrsim_\tau x(y)). \end{cases}$$

Read: 'x^* **majorizes** x' for $x^* \gtrsim x$.

Majorizability

The functionals occurring in functional interpretation have a striking mathematical structure property:

Definition (W.A. Howard 1973)

$$\begin{cases} x^* \gtrsim_{\mathbb{N}} x :\equiv x^* \geq x, \\ x^* \gtrsim_{\rho\to\tau} x :\equiv \forall y^*, y(y^* \gtrsim_\rho y \to x^*(y^*) \gtrsim_\tau x(y)). \end{cases}$$

Read: 'x^* **majorizes** x' for $x^* \gtrsim x$.

Proposition (W.A. Howard 1973)

To each closed term t^ρ of PA^ω one can define a closed term t^* s.t.

$$\mathbf{HA}^\omega \vdash t^* \gtrsim_\rho t.$$

Majorizability

The functionals occurring in functional interpretation have a striking mathematical structure property:

Definition (W.A. Howard 1973)

$$\begin{cases} x^* \gtrsim_{\mathbb{N}} x :\equiv x^* \geq x, \\ x^* \gtrsim_{\rho\to\tau} x :\equiv \forall y^*, y(y^* \gtrsim_\rho y \to x^*(y^*) \gtrsim_\tau x(y)). \end{cases}$$

Read: 'x^* **majorizes** x' for $x^* \gtrsim x$.

Proposition (W.A. Howard 1973)

To each closed term t^ρ of PA^ω one can define a closed term t^* s.t.

$$\mathbf{HA}^\omega \vdash t^* \gtrsim_\rho t.$$

Monotone functional interpretation MD (K.96) directly extracts t^*.

Uniform bound extraction by NMD

Let Δ be a set of $\mathcal{S}^\omega$-valid sentences of the form

$$\forall a^\gamma \exists b \leq_\delta ta \forall c^\eta F_{qf}(a, b, c)$$

(γ, δ, η arbitrary, t closed term, F_{qf} quantifier-free).

Uniform bound extraction by NMD

Let Δ be a set of $\mathcal{S}^\omega$-valid sentences of the form

$$\forall a^\gamma \exists b \leq_\delta ta \forall c^\eta F_{qf}(a, b, c)$$

(γ, δ, η arbitrary, t closed term, F_{qf} quantifier-free).

Example of axiom Δ: WKL!

Uniform bound extraction by NMD

Let Δ be a set of $\mathcal{S}^\omega$-valid sentences of the form

$$\forall a^\gamma \exists b \leq_\delta t a \forall c^\eta \, F_{qf}(a,b,c)$$

(γ, δ, η arbitrary, t closed term, F_{qf} quantifier-free).

Example of axiom Δ: WKL!

Theorem (K., JSL 1992)

From a proof of

$$\mathbf{PA^\omega + QF\text{-}AC + \Delta} \vdash \forall x^{\mathbb{N}\to\mathbb{N}} \forall y \leq_\rho s x \exists z^{\mathbb{N}} \, A_{qf}(x,y,z)$$

one can extract by $MD \circ N$ a primitive recursive functionals Φ s.t.

$$\mathcal{S}^\omega \models \forall x^{\mathbb{N}\to\mathbb{N}} \forall y \leq_\rho s x \exists z \leq \Phi(x) \, A_{qf}(x,y,z).$$

(ρ arbitrary, A_{qf} quantifier-free, s closed term).

General logical metatheorems I

- Context: **continuous functions** between constructively represented **Polish spaces**.

General logical metatheorems I

- Context: **continuous functions** between constructively represented **Polish spaces**.
- Uniformity w.r.t. parameters from **compact** Polish spaces.

General logical metatheorems I

- Context: **continuous functions** between constructively represented **Polish spaces**.
- Uniformity w.r.t. parameters from **compact** Polish spaces.
- Extraction of **bounds** from **noneffective** existence proofs.

In the uniform bound extraction theorem, the bound $\Phi(x)$ only depends on $x \in \mathbb{N}^{\mathbb{N}}$ but not on $y \leq_\rho sx$. For $\rho = \mathbb{N} \to \mathbb{N}$ and $s(x) :\equiv 1$ this, in particular, gives **independence from parameters in Cantor space** $2^{\mathbb{N}}$.

In the uniform bound extraction theorem, the bound $\Phi(x)$ only depends on $x \in \mathbb{N}^{\mathbb{N}}$ but not on $y \leq_\rho sx$. For $\rho = \mathbb{N} \to \mathbb{N}$ and $s(x) :\equiv 1$ this, in particular, gives **independence from parameters in Cantor space** $2^{\mathbb{N}}$.

General **Polish spaces** X and **compact metric spaces** K can be **represented** as continuous images of Baire space $\mathcal{B} := \mathbb{N}^{\mathbb{N}}$ resp. Cantor space $\mathcal{C} := 2^{\mathbb{N}}$.

In the uniform bound extraction theorem, the bound $\Phi(x)$ only depends on $x \in \mathbb{N}^{\mathbb{N}}$ but not on $y \leq_\rho sx$. For $\rho = \mathbb{N} \to \mathbb{N}$ and $s(x) :\equiv 1$ this, in particular, gives **independence from parameters in Cantor space** $2^{\mathbb{N}}$.

General **Polish spaces** X and **compact metric spaces** K can be **represented** as continuous images of Baire space $\mathcal{B} := \mathbb{N}^{\mathbb{N}}$ resp. Cantor space $\mathcal{C} := 2^{\mathbb{N}}$.

Polish spaces X are represented as the quotient of the space of fast (e.g. 2^{-n}) convergent Cauchy sequences f of elements of a countable dense subset (f can be viewed as elements in $\mathbb{N}^{\mathbb{N}}$) w.r.t. equivalence relation

$$f =_X g :\equiv \lim_{n\to\infty} d_X(f(n), g(n)) =_{\mathbb{R}} 0.$$

K., 1993-96: P Polish space, K a compact P-space, $A_\exists$ existential.
BA:= **basic arithmetic**, HBC Heine/Borel compactness WKL (SEQ$^-$ restricted sequential compactness, ACA).

K., 1993-96: P Polish space, K a compact P-space, $A_\exists$ existential.
BA:= **basic arithmetic**, HBC Heine/Borel compactness WKL (SEQ$^-$ restricted sequential compactness, ACA).
From a proof

$$\mathbf{BA} + \mathbf{HBC(+SEQ^-)} \vdash \forall \mathbf{x} \in \mathbf{P}\, \forall \mathbf{y} \in \mathbf{K}\, \exists \mathbf{m} \in \mathbb{N}\, \mathbf{A_\exists(x, y, m)}$$

K., 1993-96: P Polish space, K a compact P-space, $A_\exists$ existential.
BA:= **basic arithmetic**, HBC Heine/Borel compactness WKL (SEQ$^-$ restricted sequential compactness, ACA).
From a proof

$$\mathbf{BA} + \mathbf{HBC(+SEQ^-)} \vdash \forall \mathbf{x} \in \mathbf{P}\, \forall \mathbf{y} \in \mathbf{K}\, \exists \mathbf{m} \in \mathbb{N}\, \mathbf{A_\exists(x, y, m)}$$

one can extract a closed term Φ of BA (+iteration)

$$\mathbf{BA\ (+\ IA\ )} \vdash \forall \mathbf{x} \in \mathbf{P}\, \forall \mathbf{y} \in \mathbf{K}\, \exists \mathbf{m} \leq \mathbf{\Phi(f_x)}\, \mathbf{A_\exists(x, y, m).}$$

K., 1993-96: P Polish space, K a compact P-space, $A_\exists$ existential.
BA:= **basic arithmetic**, HBC Heine/Borel compactness WKL (SEQ$^-$ restricted sequential compactness, ACA).
From a proof

$$\mathbf{BA} + \mathbf{HBC(+SEQ^-)} \vdash \forall \mathbf{x} \in \mathbf{P}\, \forall \mathbf{y} \in \mathbf{K}\, \exists \mathbf{m} \in \mathbb{N}\, \mathbf{A_\exists(x, y, m)}$$

one can extract a closed term Φ of BA (+iteration)

$$\mathbf{BA\ (+\ IA\ )} \vdash \forall \mathbf{x} \in \mathbf{P}\, \forall \mathbf{y} \in \mathbf{K}\, \exists \mathbf{m} \leq \mathbf{\Phi(f_x)}\, \mathbf{A_\exists(x, y, m).}$$

Important:

$\Phi(f_x)$ does **not depend** on $y \in K$ but on a **representation** f_x of x!

All conditions necessary: total boundedness and completeness of K, A purely existential, Φ depending on name f_x of x.

Unique existence

P, K Polish, K compact, $f : P \times K \to \mathbb{R}$ (BA-definable).

Unique existence

P, K Polish, K compact, $f : P \times K \to \mathbb{R}$ (BA-definable).

MFI transforms **uniqueness statements**

$$\forall x \in P, y_1, y_2 \in K \Big(\bigwedge_{i=1}^{2} f(x, y_i) =_{\mathbb{R}} 0 \to d_K(y_1, y_2) =_{\mathbb{R}} 0 \Big)$$

into **moduli of uniqueness** $\Phi : \mathbb{Q}^*_+ \to \mathbb{Q}^*_+$

$$\forall x \in P, y_1, y_2 \in K, \varepsilon > 0 \Big(\bigwedge_{i=1}^{2} |f(x, y_i)| < \Phi(x, \varepsilon) \to d_K(y_1, y_2) < \varepsilon \Big).$$

Unique existence

P, K Polish, K compact, $f : P \times K \to \mathbb{R}$ (BA-definable).

MFI transforms **uniqueness statements**

$$\forall x \in P, y_1, y_2 \in K \Big(\bigwedge_{i=1}^{2} f(x, y_i) =_{\mathbb{R}} 0 \to d_K(y_1, y_2) =_{\mathbb{R}} 0 \Big)$$

into **moduli of uniqueness** $\Phi : \mathbb{Q}_+^* \to \mathbb{Q}_+^*$

$$\forall x \in P, y_1, y_2 \in K, \varepsilon > 0 \Big(\bigwedge_{i=1}^{2} |f(x, y_i)| < \Phi(x, \varepsilon) \to d_K(y_1, y_2) < \varepsilon \Big).$$

Let $\widehat{y} \in K$ be the unique root of $f(x, \cdot)$, y_ε an ε-root $|f(x, y_\varepsilon)| < \varepsilon$.

Unique existence

P, K Polish, K compact, $f : P \times K \to \mathbb{R}$ (BA-definable).

MFI transforms **uniqueness statements**

$$\forall x \in P, y_1, y_2 \in K\big(\bigwedge_{i=1}^{2} f(x, y_i) =_{\mathbb{R}} 0 \to d_K(y_1, y_2) =_{\mathbb{R}} 0\big)$$

into **moduli of uniqueness** $\Phi : \mathbb{Q}_+^* \to \mathbb{Q}_+^*$

$$\forall x \in P, y_1, y_2 \in K, \varepsilon > 0\big(\bigwedge_{i=1}^{2} |f(x, y_i)| < \Phi(x, \varepsilon) \to d_K(y_1, y_2) < \varepsilon\big).$$

Let $\widehat{y} \in K$ be the unique root of $f(x, \cdot)$, y_ε an ε-root $|f(x, y_\varepsilon)| < \varepsilon$.
Then

$$d_K(\widehat{y}, y_{\Phi(x,\varepsilon)}) < \varepsilon).$$

Theorem (K./Paulo Oliva, APAL 2003)

Let $dist_1(f, P_n) := \inf_{p \in P_n} \|f - p\|_1$ and ω a modulus of uniform continuity for f.

$$\Psi(\omega, n, \varepsilon) := \min\{\tfrac{c_n \varepsilon}{8(n+1)^2}, \tfrac{c_n \varepsilon}{2} \omega_n(\tfrac{c_n \varepsilon}{2})\}, \text{ where}$$
$$c_n := \tfrac{\lfloor n/2 \rfloor! \lceil n/2 \rceil!}{2^{4n+3}(n+1)^{3n+1}} \text{ and}$$
$$\omega_n(\varepsilon) := \min\{\omega(\tfrac{\varepsilon}{4}), \tfrac{\varepsilon}{40(n+1)^4 \lceil \frac{1}{\omega(1)} \rceil}\}.$$

Theorem (K./Paulo Oliva, APAL 2003)

Let $dist_1(f, P_n) := \inf_{p \in P_n} \|f - p\|_1$ and ω a modulus of uniform continuity for f.

$$\Psi(\omega, n, \varepsilon) := \min\{\tfrac{c_n \varepsilon}{8(n+1)^2}, \tfrac{c_n \varepsilon}{2} \omega_n(\tfrac{c_n \varepsilon}{2})\}, \text{ where}$$
$$c_n := \tfrac{\lfloor n/2 \rfloor ! \lceil n/2 \rceil !}{2^{4n+3}(n+1)^{3n+1}} \text{ and}$$
$$\omega_n(\varepsilon) := \min\{\omega(\tfrac{\varepsilon}{4}), \tfrac{\varepsilon}{40(n+1)^4 \lceil \frac{1}{\omega(1)} \rceil}\}.$$

Then $\forall n \in \mathbb{N}, p_1, p_2 \in P_n$

$$\forall \varepsilon \in \mathbb{Q}_+^* \Big(\bigwedge_{i=1}^{2} (\|f - p_i\|_1 - dist_1(f, P_n) \leq \Psi(\omega, n, \varepsilon)) \to \|p_1 - p_2\|_1 \leq \varepsilon \Big).$$

Theorem (K./Paulo Oliva, APAL 2003)

Let $dist_1(f, P_n) := \inf_{p \in P_n} \|f - p\|_1$ and ω a modulus of uniform continuity for f.

$$\Psi(\omega, n, \varepsilon) := \min\{\tfrac{c_n \varepsilon}{8(n+1)^2}, \tfrac{c_n \varepsilon}{2} \omega_n(\tfrac{c_n \varepsilon}{2})\}, \text{ where}$$
$$c_n := \tfrac{\lfloor n/2 \rfloor! \lceil n/2 \rceil!}{2^{4n+3}(n+1)^{3n+1}} \text{ and}$$
$$\omega_n(\varepsilon) := \min\{\omega(\tfrac{\varepsilon}{4}), \tfrac{\varepsilon}{40(n+1)^4 \lceil \frac{1}{\omega(1)} \rceil}\}.$$

Then $\forall n \in \mathbb{N}, p_1, p_2 \in P_n$

$$\forall \varepsilon \in \mathbb{Q}_+^*\Big(\bigwedge_{i=1}^{2} (\|f - p_i\|_1 - dist_1(f, P_n) \leq \Psi(\omega, n, \varepsilon)) \to \|p_1 - p_2\|_1 \leq \varepsilon \Big).$$

Best Chebycheff approximation: K. 1990, Num.Funct.Optim. 1993.

Best Chebycheff approximation by polynomials with restricted coefficients: K. 1990, Sipoș Math. Nachrichten 2021.

The nonseparable/noncompact case: a simply example

Proposition

Let $(X, \|\cdot\|)$ be a strictly convex normed space and $C \subseteq X$ a convex subset. Then any point $x \in X$ has at most one point $c \in C$ of minimal distance, i.e. $\|x - c\| = \text{dist}(x, C)$.

The nonseparable/noncompact case: a simply example

Proposition

Let $(X, \|\cdot\|)$ be a strictly convex normed space and $C \subseteq X$ a convex subset. Then any point $x \in X$ has at most one point $c \in C$ of minimal distance, i.e. $\|x - c\| = \mathrm{dist}(x, C)$.

Hence: if X is separable and complete and provably strictly convex and C compact, then one can extract a modulus of uniqueness.

The nonseparable/noncompact case: a simply example

Proposition

Let $(X, \|\cdot\|)$ be a strictly convex normed space and $C \subseteq X$ a convex subset. Then any point $x \in X$ has at most one point $c \in C$ of minimal distance, i.e. $\|x - c\| = \mathrm{dist}(x, C)$.

Hence: if X is separable and complete and provably strictly convex and C compact, then one can extract a modulus of uniqueness.

Observation: compactness only used to extract uniform bound on strict convexity (= **modulus of uniform convexity**) from proof of strict convexity.

Assume that X is uniformly convex with modulus η.

Assume that X is uniformly convex with modulus η.

Then for $d \geq \mathrm{dist}(x, C)$ we have the following modulus of uniqueness (K.1990):

Assume that X is uniformly convex with modulus η.

Then for $d \geq \mathrm{dist}(x, C)$ we have the following modulus of uniqueness (K.1990):

$$\Phi(\varepsilon) := \min\left(1, \frac{\varepsilon}{4}, \frac{\varepsilon}{4} \cdot \frac{\eta(\varepsilon/(d+1))}{1-\eta(\varepsilon/(d+1))}\right).$$

Conclusion: neither compactness nor separability required!

Assume that X is uniformly convex with modulus η.

Then for $d \geq \mathrm{dist}(x, C)$ we have the following modulus of uniqueness (K.1990):

$$\Phi(\varepsilon) := \min\left(1, \frac{\varepsilon}{4}, \frac{\varepsilon}{4} \cdot \frac{\eta(\varepsilon/(d+1))}{1-\eta(\varepsilon/(d+1))}\right).$$

Conclusion: neither compactness nor separability required!

In particular: **existence** of solution (for complete X and closed C) from **uniform uniqueness** which in turn stems from **uniform convexity**.

General logical metatheorems II

Many abstract types of metric structures can be added as atoms: metric, hyperbolic, CAT(0), δ-hyperbolic, $\mathbb{R}$-trees, normed, uniformly convex, uniformly smooth, Hilbert, abstract L^p and $C(K)$-spaces, dual spaces, tracial von Neumann algebras, probability spaces, spaces of Rademacher type p : add **new base type** X, all **finite types over** $\mathbb{N}, X$ and a new **constant** d_X representing d etc.

General logical metatheorems II

Many abstract types of metric structures can be added as atoms: metric, hyperbolic, CAT(0), δ-hyperbolic, $\mathbb{R}$-trees, normed, uniformly convex, uniformly smooth, Hilbert, abstract L^p and $C(K)$-spaces, dual spaces, tracial von Neumann algebras, probability spaces, spaces of Rademacher type p : add **new base type** X, all **finite types over** $\mathbb{N}, X$ and a new **constant** d_X representing d etc.

Condition: Defining axioms must have a monotone functional interpretation.

General logical metatheorems II

Many abstract types of metric structures can be added as atoms: metric, hyperbolic, CAT(0), δ-hyperbolic, $\mathbb{R}$-trees, normed, uniformly convex, uniformly smooth, Hilbert, abstract L^p and $C(K)$-spaces, dual spaces, tracial von Neumann algebras, probability spaces, spaces of Rademacher type p : add **new base type** X, all **finite types over** $\mathbb{N}, X$ and a new **constant** d_X representing d etc.

Condition: Defining axioms must have a monotone functional interpretation.

Counterexamples (to extractibility of uniform bounds): for the classes of strictly convex ($\to$ uniformly convex) or separable ($\to$ totally bounded) spaces!

Formal systems for analysis with abstract spaces X

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

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

Formal systems for analysis with abstract spaces X

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

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

$\mathbf{PA}^{\omega,X}$ is the extension of Peano Arithmetic to all types.

$\mathcal{A}^{\omega,X}$**:=**$\mathbf{PA}^{\omega,X}+\mathbf{DC}$, where

DC: axiom of dependent choice for all types

Implies **full comprehension** for numbers (higher order arithmetic).

Formal systems for analysis with abstract spaces X

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

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

$\mathbf{PA}^{\omega,X}$ is the extension of Peano Arithmetic to all types.

$\mathcal{A}^{\omega,X}$**:=**$\mathbf{PA}^{\omega,X}+\mathbf{DC}$, where

DC: axiom of dependent choice for all types

Implies **full comprehension** for numbers (higher order arithmetic).

$\mathcal{A}^{\omega}[X, d, \ldots]$ results by adding constants $d_X, \ldots$ with axioms expressing that $(X, d, \ldots)$ is a nonempty metric, hyperbolic ... space.

A warning concerning equality

Extensionality rule (**only!**):

$$\frac{A_{qf} \to s =_\rho t}{A_{qf} \to r(s) =_\tau r(t)},$$

where only $x =_{\mathbb{N}} y$ primitive equality predicate but for $\rho \to \tau$

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

A novel form of majorization

y, x functionals of types $\rho, \widehat{\rho} := \rho[\mathbb{N}/X]$ and a^X of type X:

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

A novel form of majorization

y, x functionals of types $\rho, \widehat{\rho} := \rho[\mathbb{N}/X]$ and a^X of type X:

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

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

A novel form of majorization

y, x functionals of types $\rho, \widehat{\rho} := \rho[\mathbb{N}/X]$ and a^X of type X:

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

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

Example:

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

A novel form of majorization

y, x functionals of types $\rho, \widehat{\rho} := \rho[\mathbb{N}/X]$ and a^X of type X:

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

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

Example:

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

$f : X \to X$ is **nonexpansive (n.e.)** if $d(f(x), f(y)) \leq d(x, y)$.

Then $\lambda n.n + b \gtrsim^{a}_{X \to X} f$, if $d(a, f(a)) \leq b$.

A novel form of majorization

y, x functionals of types $\rho, \widehat{\rho} := \rho[\mathbb{N}/X]$ and a^X of type X:

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

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

Example:

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

$f : X \to X$ is **nonexpansive (n.e.)** if $d(f(x), f(y)) \leq d(x, y)$.

Then $\lambda n.n + b \gtrsim^{a}_{X \to X} f$, if $d(a, f(a)) \leq b$.

Normed linear case: $a := 0_X$.

Theorem (Gerhardy/K.,Trans.Amer.Math.Soc. 2008)

Let P, K be Polish resp. compact metric spaces, $A_\exists$ $\exists$-formula, $\underline{\tau}$ types in $\{X, \mathbb{N} \to X, X \to X\}$ or a bit more complicated.
If $\mathcal{A}^\omega[X, \|\cdot\|]$ **proves**

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

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

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

Theorem (Gerhardy/K.,Trans.Amer.Math.Soc. 2008)

Let P, K be Polish resp. compact metric spaces, $A_\exists$ $\exists$-formula, $\underline{\tau}$ types in $\{X, \mathbb{N} \to X, X \to X\}$ or a bit more complicated.
If $\mathcal{A}^\omega[X, \|\cdot\|]$ **proves**

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

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

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

For the bounded cases: K. Trans.AMS 2005.

As special case of **general logical metatheorems** due to Gerhardy/K. (Trans. Amer. Math. Soc. 2008) one has:

Corollary (Gerhardy/K., TAMS 2008)

If $\mathcal{A}^\omega[X, \|\cdot\|]$ proves

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

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

$$\forall y \in K \forall z \in X \forall f : X \to X$$
$$\big(f \text{ n.e.} \wedge \|z\|, \|f(0)\| \leq b \to \exists v \leq \Phi(r_x, b)\, A_\exists\big)$$

holds in **all nontrivial normed spaces**.

Admissible classes of functions: uniformly continuous, Lipschitzian, nonexpansive, firmly-, averaged- and strongly nonexpansive mappings; some classes of **discontinuous** functions: pseudo-contractions, maps with Suzuki's condition (E) etc.

Admissible classes of functions: uniformly continuous, Lipschitzian, nonexpansive, firmly-, averaged- and strongly nonexpansive mappings; some classes of **discontinuous** functions: pseudo-contractions, maps with Suzuki's condition (E) etc.

Recently: set-valued accretive operators and their resolvents in Cauchy problems and Optimization: (K./Koutsoukou-Argyraki, K./Powell, K./Pischke, Pischke, Pischke/Powell), gradients and Fenchel conjugates (Pischke), Daniel integrable functions (Neri/Pischke), super strongly nonexpansive mappings (Sipoș).

Literature

1) Avigad, J., Feferman, S., Gödel's functional ('Dialectica') interpretation. In: Buss, S. (ed.), Handbook of Proof Theory, North-Holland, pp. 337-405 (1998).

2) Ferreira, F., Oliva, P., Bounded functional interpretation. Ann. Pure Appl. Logic **135**, pp. 73-112 (2005).

3) Gerhardy, P., Kohlenbach, U., General logical metatheorems for functional analysis, Trans. Amer. Math. Soc. **360** , pp. 2615-2660 (2008).

4) Gödel, K., Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. Dialectica **12**, pp. 280-287 (1958).

5) Günzel, D., Kohlenbach, U., Logical metatheorems for abstract spaces axiomatized in positive bounded logic. Advances in Mathematics **290**, pp. 503-551 (2016).

6) Hilbert, D., Über das Unendliche. Math. Ann. **95**, pp. 161-190 (1926).

7) Howard, W.A., Hereditarily majorizable functionals of finite type. In: Troelstra (ed.), Metamathematical investigation of intuitionistic arithmetic and analysis, pp. 454-461. Springer LNM 344 (1973).

8) Kohlenbach, U., Effective bounds from ineffective proofs in analysis: an application of functional interpretation and majorization. J. Symbolic Logic **57**, pp. 1239–1273 (1992).

9) Kohlenbach, U., Effective moduli from ineffective uniqueness proofs. An unwinding of de La Vallee Poussin's proof for Chebycheff approximation. Annals of Pure and Applied Logic **64**, pp. 27-94 (1993).

10) Kohlenbach, U., New effective moduli of uniqueness and uniform a-priori estimates for constants of strong unicity by logical analysis of known proofs in best approximation theory. Numer.Funct.Anal.and Optimiz. **14**, pp. 581-606 (1993).

11) Kohlenbach, U., Analysing proofs in analysis. In: W. Hodges, M. Hyland, C. Steinhorn, J. Truss, editors, *Logic: from Foundations to Applications. European Logic Colloquium* (Keele, 1993), pp. 225–260, Oxford University Press (1996).

12) Kohlenbach, U., Some logical metatheorems with applications in functional analysis. Trans. Amer. Math. Soc. **357**, pp. 89-128 (2005).

13) Kohlenbach, U., Applied Proof Theory. Springer Monograph in Mathematics. Springer 2008, xix+536pp.

14) Kohlenbach, U., Gödel's functional interpretation and its use in current mathematics. In: Kurt Gödel and the Foundations of Mathematics. Horizons of Truth. Baaz, M. et al. (eds.), Cambridge University Press, pp. 361-398, New York 2011.

15) Kohlenbach, U., Oliva, P., Proof mining in L_1-approximation. Ann. Pure Appl. Logic **121**, pp. 1-38 (2003).

16) Pischke, N., Logical Metatheorems for Accretive and (Generalized) Monotone Set-Valued Operators. J. Math. Logic **24** (2024).

17) Sipoș, Bounds on strong unicity for Chebyshev approximation with bounded coefficients. Math. Nachrichten **294**, pp. 2425-2440 (2021).

18) Spector, C., Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics. In: Recursive function theory, Proceedings of Symposia in Pure Mathematics, vol. 5 (J.C.E. Dekker (ed.)), AMS, Providence, R.I., pp. 1-27 (1962).

Lecture III

As special case of **general logical metatheorems** one has:

Corollary (Gerhardy/K., TAMS 2008)

If $\mathcal{A}^\omega[X, \|\cdot\|]$ proves

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

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

$$\forall y \in K \forall z \in X \forall f : X \to X$$
$$\big(f \text{ n.e.} \wedge \|z\|, \|f(0)\| \leq b \to \exists v \leq \Phi(r_x, b)\, A_\exists\big)$$

holds in **all nontrivial normed spaces**.

Similar for Hilbert spaces and uniformly convex spaces (then bound depends on modulus of convexity).

Applications to Ergodic Theory

The Mean Ergodic Theorem

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

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

The Mean Ergodic Theorem

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

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

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 $z \in X$.

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

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: **Prim. rec. bound on metastable version!**

The Mean Ergodic Theorem

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

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

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: **Prim. rec. bound on metastable version!**

Theorem (Garrett Birkhoff 1939)

Mean Ergodic Theorem holds for uniformly convex Banach spaces.

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

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

X uniformly convex Banach space with modulus η and $f : X \to X$ nonexpansive linear operator. Let $b > 0$. Then there is an effective functional Φ in ε, g, b, η s.t. for all $z \in X$ 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 Birkhoff's proof formalizes in $\mathcal{A}^\omega[X, \|\cdot\|, \eta]$ the following is guaranteed:

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

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

$f^* := id$ 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 $\|z\| \leq b$, all $\varepsilon > 0$, all $g : \mathbb{N} \to \mathbb{N}$:

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

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

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

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 $g : \mathbb{N} \to \mathbb{N}$:

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

where

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

Special Hilbert case: treated prior by Avigad/Gerhardy/Towsner TAMS 2010 (again based on logical metatheorem).

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 $g : \mathbb{N} \to \mathbb{N}$:

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

where

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

Special Hilbert case: treated prior by Avigad/Gerhardy/Towsner TAMS 2010 (again based on logical metatheorem).

Extension to finitely many mappings Sipoş (Pacific J. Math. 2021).

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 $i_1 \leq j_1 \leq \ldots i_k \leq j_k$ s.t. $\|x_{j_n} - x_{i_n}\| \geq \varepsilon$ for $n = 1, \ldots, k$.

As a corollary to 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 $i_1 \leq j_1 \leq \ldots i_k \leq j_k$ s.t. $\|x_{j_n} - x_{i_n}\| \geq \varepsilon$ for $n = 1, \ldots, k$.

As a corollary to 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.

Extensions to Nonlinear Ergodic Theorems

- **Metastability for Baillon's nonlinear ergodic theorem** (**K., Comm.Contemp.Math. 2012**).
- **Metastability for Wittmann's strong nonlinear ergodic theorem on Halpern iterations** (**K., Adv.Math. 2011**).
- **Generalization of the above to CAT(0)-spaces** (Saejung 2010) (**K., Leuştean: Adv.Math. 2012**).
- **Generalization to CAT(κ) spaces** $\kappa > 0$ (**Leuştean, Nicolae Ergodic Theor. Dynam. Syst. 2016**).
- **Metastable version of strong nonlinear ergodic theorem** for maps

$$\forall u, v \in C\ (\|f(u) + f(v)\| \leq \|u + v\|).$$

(covers Baillon's result for n.e. odd operators): **Safarik, J.Math.Anal.Appl. 2012**.

Let X be a uniformly convex Banach space, $C \subseteq X$ nonempty, bounded, closed and convex, $T : C \to C$ nonexpansive.

$$S_n x := \frac{1}{n} \sum_{i=0}^{n-1} T^i x.$$

Let X be a uniformly convex Banach space, $C \subseteq X$ nonempty, bounded, closed and convex, $T : C \to C$ nonexpansive.

$$S_n x := \frac{1}{n} \sum_{i=0}^{n-1} T^i x.$$

Theorem (Bruck 1981)

$$\lim_{n\to\infty} \|T(S_n x) - S_n x\| \to 0 \text{ uniformly in } x \in C.$$

Let X be a uniformly convex Banach space, $C \subseteq X$ nonempty, bounded, closed and convex, $T : C \to C$ nonexpansive.

$$S_n x := \frac{1}{n} \sum_{i=0}^{n-1} T^i x.$$

Theorem (Bruck 1981)

$$\lim_{n\to\infty} \|T(S_n x) - S_n x\| \to 0 \text{ uniformly in } x \in C.$$

Bruck's proof crucially uses a result due to Pisier 1973 that uniformly convex Banach spaces (in fact uniformly nonsquare and even B-convex Banach spaces) have a nontrivial Rademacher type $q > 1$ which implies that there exists a constant C_q s.t. for every finite sequence $X_1, \ldots, X_n$ of independent mean zero Radon random variables in $L_q(X)$

$$\mathbb{E}\left(\left\|\sum_{i=1}^{n} X_i\right\|^q\right)^{1/q} \leq 2C_q \cdot \left(\sum_{i=1}^{n} \mathbb{E}\left(\|X_i\|^q\right)\right)^{1/q}.$$

Let $(X, \|\cdot\|)$ be **uniformly nonsquare** with $\delta \in (0, 1]$ such that

$$\min\left\{\frac{\|x-y\|}{2}, \frac{\|x+y\|}{2}\right\} \leq (1-\delta)\cdot \max\{\|x\|, \|y\|\}$$

holds for all $x, y \in X$.

Let $(X, \|\cdot\|)$ be **uniformly nonsquare** with $\delta \in (0, 1]$ such that

$$\min\left\{\frac{\|x-y\|}{2}, \frac{\|x+y\|}{2}\right\} \leq (1-\delta)\cdot\max\{\|x\|, \|y\|\}$$

holds for all $x, y \in X$.

If X is uniformly convex and η is a modulus of convexity, we may take $\delta := \eta(1)$.

From Pisier's proof one can extract the following explicit constants:

Theorem (Pisier 1973, Freund/K. ASVAO 2022)

Let $(X, \|\cdot\|)$ be a uniformly nonsquare Banach space with $\delta \in (0,1)$ witnessing this property. Define $\lambda := 1 - \delta$. Assume that $\xi \in (0,1)$ is so small and that $p' \in [2, \infty)$ is so large that

$$\frac{1-\xi}{1+2\sqrt{2\xi}} \geq \frac{1}{2}\sqrt{2\lambda^2+2} \quad \text{and} \quad \frac{1}{2^{1/p'}} \geq 1-\xi.$$

Take p with $1 = \frac{1}{p} + \frac{1}{p'}$.

From Pisier's proof one can extract the following explicit constants:

Theorem (Pisier 1973, Freund/K. ASVAO 2022)

Let $(X, \|\cdot\|)$ be a uniformly nonsquare Banach space with $\delta \in (0,1)$ witnessing this property. Define $\lambda := 1 - \delta$. Assume that $\xi \in (0,1)$ is so small and that $p' \in [2, \infty)$ is so large that

$$\frac{1-\xi}{1+2\sqrt{2\xi}} \geq \frac{1}{2}\sqrt{2\lambda^2+2} \quad \text{and} \quad \frac{1}{2^{1/p'}} \geq 1-\xi.$$

Take p with $1 = \frac{1}{p} + \frac{1}{p'}$.
Then - for any $q \in (1,p)$ - X has Rademacher type q with constant

$$C_q = 3 \cdot \frac{2^{1/q}}{2^{(1/q)-(1/p)} - 1}.$$

Analyzing proofs due to Bruck 1979, 1981 one can extract:

Theorem (Freund/K. ASVAO 2022)

Let $(X, \|\cdot\|)$ be uniformly convex with modulus η. $T : C \to C$ nonexpansive, $C \subseteq B_{b/2}(0)$ closed and convex. Assume X is of Rademacher type $q \in (1, 2]$ with constant C_q. Given $\varepsilon \in (0, 1]$, pick $\tilde{p} \in \mathbb{N}$ so large that $2C_q \cdot \tilde{p}^{(1-q)/q} \leq \varepsilon/(9b)$. Consider $p \in \mathbb{N}$ with $p \geq 2b/\delta^2$ for

$$\delta := \xi^{\tilde{p}}\left(\frac{\varepsilon}{9}\right) \quad \text{with} \quad \xi(t) := \frac{t}{12} \cdot \eta\left(\min\left\{2, \frac{t}{b}\right\}\right).$$

For any $\alpha < \xi^{p-1}(\delta^2/2)$ with $0 < \alpha < \varepsilon/3$ we have

$$\forall n \geq \frac{b}{\alpha} \ \forall x \in C \ (\|T(S_n x) - S_n x\| \leq \varepsilon).$$

Theorem (Kobayasi-Miyadera 1980)

Consider $T : C \to C$ as above. Given $x \in C$, assume that the sequences $(\|T^n x - T^{n+i} x\|)_n$ converge uniformly in i. We then have

$$\lim_{n\to\infty} \|y - S_n T^k x\| = 0 \quad \text{uniformly in } k,$$

for some fixed point y of T.

Freund/K. (ETDS 2023): Explicit function Φ which translates any rate A of metastability for the assumption that the sequences $(\|T^n x - T^{n+i} x\|)_n$ converge uniformly in i into a common rate of metastability for the conclusion.

Proof mining involving monotone operators

Maximal monotone operators

Let $(\lambda_n) \subset (0, \infty)$ and $A : X \to 2^X$ a set-valued (maximal) monotone (resp. accretive) operator with$X$ Hilbert (resp. Banach space):

$$\langle x - y, u - v \rangle \geq 0, \ u \in Ax, v \in Ay$$

(in Banach space: normalized duality map instead of scalar product).

Maximal monotone operators

Let $(\lambda_n) \subset (0, \infty)$ and $A : X \to 2^X$ a set-valued (maximal) monotone (resp. accretive) operator with$X$ Hilbert (resp. Banach space):

$$\langle x - y, u - v \rangle \geq 0, \ u \in Ax, v \in Ay$$

(in Banach space: normalized duality map instead of scalar product).

- Pischke JML 2023: maximal monotone/accretive operators $A : X \to 2^X$ and resolvents $J_A : X \to X$, $J_A := (I + A)^{-1}$ can be incorporated into the logic metatheorems **intensionally** via the characteristic function for $graph_A$ (with a weak extensionality rule).

Maximal monotone operators

Let $(\lambda_n) \subset (0, \infty)$ and $A : X \to 2^X$ a set-valued (maximal) monotone (resp. accretive) operator with$X$ Hilbert (resp. Banach space):

$$\langle x - y, u - v \rangle \geq 0, \ u \in Ax, v \in Ay$$

(in Banach space: normalized duality map instead of scalar product).

- Pischke JML 2023: maximal monotone/accretive operators $A : X \to 2^X$ and resolvents $J_A : X \to X$, $J_A := (I + A)^{-1}$ can be incorporated into the logic metatheorems **intensionally** via the characteristic function for $graph_A$ (with a weak extensionality rule).
- Important theorems on such operators such as the **Brezis-Haraux theorem** admit a **tame** proof-theoretic analysis (K./Pischke, Phil. Trans. Royal. Soc. 2023)

- The fundamental relation $(+)\ 0 \in Az \leftrightarrow z =_X J_A z$ (which needs extensionality) holds quantitatively if $0 \in Az$ is replaced by

$$\exists x, \tilde{z} \in X\,(x =_X 0 \wedge z =_X \tilde{z} \wedge x \in A\tilde{z})$$

via the provable intensional certificate for '$\leftarrow$' in $(+)$

$$z - J_A z \in A(J_A z)$$

and the provable

$$x =_X 0 \wedge z =_X \tilde{z} \wedge x \in A\tilde{z} \rightarrow z =_X J_A z.$$

Proximal Point Algorithms

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

Proximal Point Algorithms

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

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

Proximal Point Algorithms

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

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

Always **strong convergence** for **Halpern type variant of PPA**:

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

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

Proximal Point Algorithms

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

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

Always **strong convergence** for **Halpern type variant of PPA**:

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

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

X **uniformly convex and uniformly smooth** and for **$\inf \lambda_n \geq \lambda > 0$**: **effective rate of metastability** extracted from a proof due to Aoyama and Toyoda 2017 (K. J. Nonlin. Convex Anal. 2020).

Proximal Point Algorithms

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

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

Always **strong convergence** for **Halpern type variant of PPA**:

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

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

X **uniformly convex and uniformly smooth** and for **$\inf \lambda_n \geq \lambda > 0$**: **effective rate of metastability** extracted from a proof due to Aoyama and Toyoda 2017 (K. J. Nonlin. Convex Anal. 2020).

That analysis is based on the mining (due to K./Sipoș 2021) of a proof of a deep theorem by Reich for uniformly smooth spaces.

Adaptions made possible by the finitary proof

Our analysis of the HPPA in uniformly convex/uniformly smooth Banach spaces was subsequently adapted to

Adaptions made possible by the finitary proof

Our analysis of the HPPA in uniformly convex/uniformly smooth Banach spaces was subsequently adapted to

- CAT(0)-spaces (Sipoș Comp.Optim.Appl. 2022)

Adaptions made possible by the finitary proof

Our analysis of the HPPA in uniformly convex/uniformly smooth Banach spaces was subsequently adapted to

- CAT(0)-spaces (Sipoș Comp.Optim.Appl. 2022)
- maximally monotone operators $A : X \to 2^{X^*}$ in Banach spaces with **Bregman resolvents** (Pischke/K.2023)

Adaptions made possible by the finitary proof

Our analysis of the HPPA in uniformly convex/uniformly smooth Banach spaces was subsequently adapted to

- CAT(0)-spaces (Sipoș Comp.Optim.Appl. 2022)
- maximally monotone operators $A : X \to 2^{X^*}$ in Banach spaces with **Bregman resolvents** (Pischke/K.2023)
- a **new class** of **uniformly smooth hyperbolic spaces** (Pinto TAMS 2024). That extension of 'uniformly smooth' to the metric setting was motivated by the aforementioned mining of Reich's theorem which has been extended by Pinto to the metric setting.

A rate of asymptotic regularity for the Bregman HPPA

X real Banach space, $f : X \to \mathbb{R}$ supercoercive, bounded and totally convex on bounded sets and uniformly Fréchet differentiable and f^* is the Fenchel conjugate. $A : X \to 2^{X^*}$ maximally monotone with $A^{-1}0 \neq \emptyset$. $(\alpha_n) \subset (0, 1]$ with $\lim \alpha_n = 0, \sum \alpha_n = \infty$, $(r_n) \subset (0, \infty)$ with $\inf r_n > 0$. For $x_0 = x, u \in X$

$$x_{n+1} := \nabla f^*(\alpha_n \nabla f u + (1 - \alpha_n)\nabla f \, Res^f_{r_n} x_n).$$

A rate of asymptotic regularity for the Bregman HPPA

X real Banach space, $f : X \to \mathbb{R}$ supercoercive, bounded and totally convex on bounded sets and uniformly Fréchet differentiable and f^* is the Fenchel conjugate. $A : X \to 2^{X^*}$ maximally monotone with $A^{-1}0 \neq \emptyset$. $(\alpha_n) \subset (0, 1]$ with $\lim \alpha_n = 0, \sum \alpha_n = \infty$, $(r_n) \subset (0, \infty)$ with $\inf r_n > 0$. For $x_0 = x, u \in X$

$$x_{n+1} := \nabla f^*(\alpha_n \nabla f u + (1 - \alpha_n) \nabla f \, Res^f_{r_n} x_n).$$

Pischke/K. 2023:

- In suitable moduli, an effective rate of metastability for the strong convergence of (x_n) towards $P^f_{A^{-1}0}(u)$ is given.

A rate of asymptotic regularity for the Bregman HPPA

X real Banach space, $f : X \to \mathbb{R}$ supercoercive, bounded and totally convex on bounded sets and uniformly Fréchet differentiable and f^* is the Fenchel conjugate. $A : X \to 2^{X^*}$ maximally monotone with $A^{-1}0 \neq \emptyset$. $(\alpha_n) \subset (0, 1]$ with $\lim \alpha_n = 0, \sum \alpha_n = \infty$, $(r_n) \subset (0, \infty)$ with $\inf r_n > 0$. For $x_0 = x, u \in X$

$$x_{n+1} := \nabla f^*(\alpha_n \nabla f u + (1 - \alpha_n) \nabla f \, Res^f_{r_n} x_n).$$

Pischke/K. 2023:

- In suitable moduli, an effective rate of metastability for the strong convergence of (x_n) towards $P^f_{A^{-1}0}(u)$ is given.
- If $r_n \to \infty$, then full effective rate of asymptotic regularity.

A rate of asymptotic regularity for the Bregman HPPA

X real Banach space, $f : X \to \mathbb{R}$ supercoercive, bounded and totally convex on bounded sets and uniformly Fréchet differentiable and f^* is the Fenchel conjugate. $A : X \to 2^{X^*}$ maximally monotone with $A^{-1}0 \neq \emptyset$. $(\alpha_n) \subset (0, 1]$ with $\lim \alpha_n = 0, \sum \alpha_n = \infty$, $(r_n) \subset (0, \infty)$ with $\inf r_n > 0$. For $x_0 = x, u \in X$

$$x_{n+1} := \nabla f^*(\alpha_n \nabla f u + (1 - \alpha_n) \nabla f \, Res^f_{r_n} x_n).$$

Pischke/K. 2023:

- In suitable moduli, an effective rate of metastability for the strong convergence of (x_n) towards $P^f_{A^{-1}0}(u)$ is given.
- If $r_n \to \infty$, then full effective rate of asymptotic regularity.

Pischke 2023: extensions of logical metatheorem which include all the above concepts and cover these applications!

Fejér monotonicity and regularity

Why not always use the strongly convergent HPPA instead of the weakly convergent PPA?

Fejér monotonicity and regularity

Why not always use 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 not always use 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 $F \subseteq X$ if $\forall n \in \mathbb{N} \forall p \in F \ (d(x_{n+1}, p) \leq d(x_n, p))$.

Fejér monotonicity and regularity

Why not always use 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 $F \subseteq X$ if $\forall n \in \mathbb{N} \forall p \in F \ (d(x_{n+1}, p) \leq d(x_n, p))$.

Why is this important?

Fejér monotonicity and regularity

Why not always use 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 $F \subseteq X$ if $\forall n \in \mathbb{N} \forall p \in F \ (d(x_{n+1}, p) \leq d(x_n, p))$.

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:

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:

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

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

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:

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

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

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

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

Covers: fixed point, equilibrium problems, zeros of set-valued maps.

Computational use of moduli of regularity

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

Let $F : X \to \mathbb{R}$ *be with* $\operatorname{zer} F \neq \emptyset$ *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.* $\operatorname{zer} F$. *Then* (x_n) *is Cauchy:*

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

Computational use of moduli of regularity

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

Let $F : X \to \mathbb{R}$ *be with* $\operatorname{zer} F \neq \emptyset$ *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.* $\operatorname{zer} F$. *Then* (x_n) *is Cauchy:*

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

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

Dykstra's algorithm and generalized Fejér monotonicity

Dykstra algorithm: strongly convergent variant of the weakly convergent cyclic projection method to compute points in intersections $C := \bigcap_{i=1}^{m} C_i$ of closed convex subsets in Hilbert space (consistent feasibility problem).

$$\begin{cases} x_0 \in X \\ q_{-(m-1)} = \ldots = q_0 = 0 \end{cases} \text{ and } \begin{cases} x_n := P_{C_{j(n)}}(x_{n-1} + q_{n-m}) \\ q_n := x_{n-1} + q_{n-m} - x_n, \end{cases}$$

where $j_n := (n - 1 \, mod \, m) + 1$.

Dykstra's algorithm and generalized Fejér monotonicity

Dykstra algorithm: strongly convergent variant of the weakly convergent cyclic projection method to compute points in intersections $C := \bigcap_{i=1}^{m} C_i$ of closed convex subsets in Hilbert space (consistent feasibility problem).

$$\begin{cases} x_0 \in X \\ q_{-(m-1)} = \ldots = q_0 = 0 \end{cases} \text{ and } \begin{cases} x_n := P_{C_{j(n)}}(x_{n-1} + q_{n-m}) \\ q_n := x_{n-1} + q_{n-m} - x_n, \end{cases}$$

where $j_n := (n - 1 \, mod \, m) + 1$.

Pinto Zeitschr. f. Anal. Anwend. 2024: by proof mining

- effective rate of metastability,

Dykstra's algorithm and generalized Fejér monotonicity

Dykstra algorithm: strongly convergent variant of the weakly convergent cyclic projection method to compute points in intersections $C := \bigcap_{i=1}^{m} C_i$ of closed convex subsets in Hilbert space (consistent feasibility problem).

$$\begin{cases} x_0 \in X \\ q_{-(m-1)} = \ldots = q_0 = 0 \end{cases} \text{ and } \begin{cases} x_n := P_{C_{j(n)}}(x_{n-1} + q_{n-m}) \\ q_n := x_{n-1} + q_{n-m} - x_n, \end{cases}$$

where $j_n := (n - 1 \, mod \, m) + 1$.

Pinto Zeitschr. f. Anal. Anwend. 2024: by proof mining

- effective rate of metastability,
- Although (x_n) is **not Fejér monotone** (w.r.t. C): effective rate of convergence in the presence of a modulus of regularity.

Dykstra's algorithm and generalized Fejér monotonicity

Dykstra algorithm: strongly convergent variant of the weakly convergent cyclic projection method to compute points in intersections $C := \bigcap_{i=1}^{m} C_i$ of closed convex subsets in Hilbert space (consistent feasibility problem).

$$\begin{cases} x_0 \in X \\ q_{-(m-1)} = \ldots = q_0 = 0 \end{cases} \text{ and } \begin{cases} x_n := P_{C_{j(n)}}(x_{n-1} + q_{n-m}) \\ q_n := x_{n-1} + q_{n-m} - x_n, \end{cases}$$

where $j_n := (n - 1 \, mod \, m) + 1$.

Pinto Zeitschr. f. Anal. Anwend. 2024: by proof mining

- effective rate of metastability,
- Although (x_n) is **not Fejér monotone** (w.r.t. C): effective rate of convergence in the presence of a modulus of regularity.
- This lead to a **new generalized form** of **Fejér monotonicity** which covers this (K./Pinto 2024).

Noncomputability of moduli of metric regularity

Proposition

F continuous, X compact, $zerF \neq \emptyset$: F has a modulus of regularity.

Noncomputability of moduli of metric regularity

Proposition

F continuous, X compact, $zerF \neq \emptyset$: F has a modulus of regularity.

In general, **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

F continuous, X compact, $zerF \neq \emptyset$: F has a modulus of regularity.

In general, **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...!)

Linear rates for the Tikhonov regularization method

Recently, Boţ et al. studied **strongly convergent** iteration methods in optimization using a so-called **Tikhonov regularization term** which has been generalized by Cheval/Leuştean (Optimiz. Method. Softw., to appear) to the **Tikhonov-Mann iteration**

$$\text{(T-M)}: \begin{cases} x_{n+1} = (1-\lambda_n)u_n + \lambda_n T u_n \\ u_n = (1-\beta_n)u + \beta_n x_n, \end{cases}$$

where $C \subseteq X$ is a convex subset of a W-hyperbolic space X, $T : C \to C$ is **nonexpansive** and $(\lambda_n)_{n\in\mathbb{N}}, (\beta_n)_{n\in\mathbb{N}}$ are sequences in $[0,1]$ and $x_0, u \in C$ and is studied in proof mining also by Dinis/Pinto.

Recently, Boţ et al. studied **strongly convergent** iteration methods in optimization using a so-called **Tikhonov regularization term** which has been generalized by Cheval/Leuştean (Optimiz. Method. Softw., to appear) to the **Tikhonov-Mann iteration**

$$\text{(T-M)}: \begin{cases} x_{n+1} = (1-\lambda_n)u_n + \lambda_n T u_n \\ u_n = (1-\beta_n)u + \beta_n x_n, \end{cases}$$

where $C \subseteq X$ is a convex subset of a W-hyperbolic space X, $T : C \to C$ is **nonexpansive** and $(\lambda_n)_{n\in\mathbb{N}}, (\beta_n)_{n\in\mathbb{N}}$ are sequences in $[0,1]$ and $x_0, u \in C$ and is studied in proof mining also by Dinis/Pinto.

While Boţ et al. only considered the case $u := 0$ in Hilbert spaces, (x_n) is defined in arbitrary W-hyperbolic spaces.

A lemma due to Sabach/Shtern 2017 combined with the results obtained by proof mining in the general case gives a **linear** rate of asymptotic regularity for $\lambda_n = \lambda \in (0,1), \beta_n := 1 - \frac{2}{n}$ in general hyperbolic spaces:

A lemma due to Sabach/Shtern 2017 combined with the results obtained by proof mining in the general case gives a **linear** rate of asymptotic regularity for $\lambda_n = \lambda \in (0,1), \beta_n := 1 - \frac{2}{n}$ in general hyperbolic spaces:

Theorem (Cheval/K./Leuștean, JOTA 2023)

Let $K \geq \max\{d(x_0,p), d(u,p)\}$ for some fixed point p of T. Assume that $\lambda_n = \lambda \in (0,1]$, $\beta_1 = 0$ and $\beta_n = 1 - \frac{2}{n}$ for $n \geq 2$, and let (x_n) be the Tikhonov-Mann iteration. Then for all $n \geq 1$,

$$d(x_{n+1}, x_n) \leq \frac{4K}{n+1}, \; d(x_n, Tx_n) \leq \frac{8K}{\lambda n}.$$

(T-M) has been generalized by Dinis/Pinto to two nonexpansive mappings T, U

$$\text{(T-M)}: \begin{cases} x_{n+1} = (1-\lambda_n)u_n + \lambda_n T u_n \\ u_n = (1-\beta_n)u + \beta_n U x_n. \end{cases}$$

(T-M) has been generalized by Dinis/Pinto to two nonexpansive mappings T, U

$$\text{(T-M)}: \begin{cases} x_{n+1} = (1-\lambda_n)u_n + \lambda_n T u_n \\ u_n = (1-\beta_n)u + \beta_n U x_n. \end{cases}$$

Theorem (Dinis/Pinto, SIAM Journal on Optimization 2023)

If U, T have common fixed points, then (x_n) strongly converges to the common fixed point closest to u if X is a CAT(0) space under suitable conditions on the scalars. Moreover, an explicit effective rate of metastability can be given.

(T-M) has been generalized by Dinis/Pinto to two nonexpansive mappings T, U

$$\text{(T-M)}: \begin{cases} x_{n+1} = (1-\lambda_n)u_n + \lambda_n T u_n \\ u_n = (1-\beta_n)u + \beta_n U x_n. \end{cases}$$

Theorem (Dinis/Pinto, SIAM Journal on Optimization 2023)

If U, T have common fixed points, then (x_n) strongly converges to the common fixed point closest to u if X is a CAT(0) space under suitable conditions on the scalars. Moreover, an explicit effective rate of metastability can be given.

Here even the qualitative result (obtained by generalizing the mined proof of a strong convergent theorem in Hilbert space) is new.

(T-M) has been generalized by Dinis/Pinto to two nonexpansive mappings T, U

$$\text{(T-M)}: \begin{cases} x_{n+1} = (1-\lambda_n)u_n + \lambda_n T u_n \\ u_n = (1-\beta_n)u + \beta_n U x_n. \end{cases}$$

Theorem (Dinis/Pinto, SIAM Journal on Optimization 2023)

If U, T have common fixed points, then (x_n) strongly converges to the common fixed point closest to u if X is a CAT(0) space under suitable conditions on the scalars. Moreover, an explicit effective rate of metastability can be given.

Here even the qualitative result (obtained by generalizing the mined proof of a strong convergent theorem in Hilbert space) is new.

For $U := Id$ the result is due to Seajung 2010 and - quantitatively - Schade/K. 2012.

A polynomial rate of asymptotic regularity in Bauschke's solution of the 'minimal 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 'minimal 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 'minimal 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 'minimal 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.

Proof uses abstract theory of maximal monotone operators: Minty's theorem (Zorn's lemma), 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 $b \geq \|x\|$

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

This explains the rate presented in Lecture I!

Proof Mining in Nonlinear Semigroup Theory

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

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

(*i*) $\boldsymbol{T(s+t) = T(s) \circ T(t) \quad (s,t \geq 0)}$,

(*ii*) for each $\boldsymbol{x \in C}$, the mapping $\boldsymbol{t \mapsto T(t)x}$ is continuous.

Proof Mining in Nonlinear Semigroup Theory

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

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

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

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

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

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

$$Fix(S) \subseteq \bigcap_{t \geq 0} Fix(T(t)).$$

General logical metatheorems guarantee:

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

General logical metatheorems guarantee:

let $t \mapsto T(t)x$ be equicontinuous on norm-bounded subsets of X with modulus ω, let f_γ be a **modulus of irrationality** for $\gamma := \alpha/\beta$, $\Lambda, N, D \in \mathbb{N}$ be s.t. $1/\Lambda \leq \lambda, 1-\lambda$ and $1/N \leq \beta \leq D$. Then one can extract a bound $\Phi(\varepsilon, M, d, N, \Lambda, D, f_\gamma, \omega)$ s.t. for all $M, d \in \mathbb{N} \forall p \in C \, \forall \varepsilon > 0$

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

General logical metatheorems guarantee:

let $t \mapsto T(t)x$ be equicontinuous on norm-bounded subsets of X with modulus ω, let f_γ be a **modulus of irrationality** for $\gamma := \alpha/\beta$, $\Lambda, N, D \in \mathbb{N}$ be s.t. $1/\Lambda \leq \lambda, 1-\lambda$ and $1/N \leq \beta \leq D$. Then one can extract a bound $\Phi(\varepsilon, M, d, N, \Lambda, D, f_\gamma, \omega)$ s.t. for all $M, d \in \mathbb{N} \forall p \in C \, \forall \varepsilon > 0$

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

Let $x_{n+1} := \frac{1}{2}x_n + \frac{1}{2}Sx_n$ be a d-bounded Krasnoselski iteration of S with rate of asymptotic regularity Ψ, then

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

General logical metatheorems guarantee:

let $t \mapsto T(t)x$ be equicontinuous on norm-bounded subsets of X with modulus ω, let f_γ be a **modulus of irrationality** for $\gamma := \alpha/\beta$, $\Lambda, N, D \in \mathbb{N}$ be s.t. $1/\Lambda \leq \lambda, 1-\lambda$ and $1/N \leq \beta \leq D$. Then one can extract a bound $\Phi(\varepsilon, M, d, N, \Lambda, D, f_\gamma, \omega)$ s.t. for all $M, d \in \mathbb{N} \forall p \in C \, \forall \varepsilon > 0$

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

Let $x_{n+1} := \frac{1}{2}x_n + \frac{1}{2}Sx_n$ be a d-bounded Krasnoselski iteration of S with rate of asymptotic regularity Ψ, then

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

Note: Suzuki's proof uses WKL: **WKL-elimination!**

Theorem (Koutsoukou-Argyraki/K., J.Math.Anal.Appl. 2016)

Under the previous assumptions:

$$\forall M \in \mathbb{N}\, \forall m \in \mathbb{N}\, \forall n \geq \Phi(M,m)\, \forall t \in [0,M]\, (\|T(t)x_n - x_n\| < 2^{-m})$$

Theorem (Koutsoukou-Argyraki/K., J.Math.Anal.Appl. 2016)

Under the previous assumptions:

$$\forall M \in \mathbb{N}\, \forall m \in \mathbb{N}\, \forall n \geq \Phi(M,m)\, \forall t \in [0,M]\, (\|T(t)x_n - x_n\| < 2^{-m})$$

with

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

where $d \geq \|x_0 - Sx_n\|$ for all n, $k := D2^{\omega_{D,d}(3+[\log_2(1+MN)]+m)+1}$ and

$$\phi(k,f) := \max\{2f(i)+6 \ : \ 0 < i \leq k\}.$$

Theorem (Koutsoukou-Argyraki/K., J.Math.Anal.Appl. 2016)

Under the previous assumptions:

$$\forall M \in \mathbb{N}\, \forall m \in \mathbb{N}\, \forall n \geq \Phi(M,m)\, \forall t \in [0,M]\, (\|T(t)x_n - x_n\| < 2^{-m})$$

with

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

where $d \geq \|x_0 - Sx_n\|$ for all n, $k := D2^{\omega_{D,d}(3+[\log_2(1+MN)]+m)+1}$ and

$$\phi(k,f) := \max\{2f(i)+6 \ : \ 0 < i \leq k\}.$$

Example:

$\alpha = \sqrt{2}, \beta = 2, \lambda = 1/2$. Then $\Lambda = 2, N = 1, D = 2, f_\gamma(p) = 4p^2$.

Semigroups and Cauchy problems

- Semigroups play a central role in the study of Cauchy problems which are given by accretive operators in Banach spaces (Crandal-Liggett, Pazy, Barbu etc. since the 1970's).

Semigroups and Cauchy problems

- Semigroups play a central role in the study of Cauchy problems which are given by accretive operators in Banach spaces (Crandal-Liggett, Pazy, Barbu etc. since the 1970's).
- Proof mining in this context has been carried out by Koutsoukou-Argyraki/K. JMAA 2015, Pinto/Pischke Doc. Math. 2023, Pischke JMAA 2024.

Semigroups and Cauchy problems

- Semigroups play a central role in the study of Cauchy problems which are given by accretive operators in Banach spaces (Crandal-Liggett, Pazy, Barbu etc. since the 1970's).
- Proof mining in this context has been carried out by Koutsoukou-Argyraki/K. JMAA 2015, Pinto/Pischke Doc. Math. 2023, Pischke JMAA 2024.
- Logical metatheorems covering the theory of nonlinear semigroups are due to Pischke 2023.

Pursuit-evasion games: Lion-Man

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

Pursuit-evasion games: Lion-Man

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

Under general conditions on the X one recently has shown highly noneffectively that **the lion wins** i.e. $\mathbf{\lim d(L_{i+1}, M_i) = 0}$.

Pursuit-evasion games: Lion-Man

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

Under general conditions on the X one recently has shown highly noneffectively that **the lion wins** i.e. $\mathbf{\lim d(L_{i+1}, M_i) = 0}$.

Proof mining yields (K./Lopéz-Acedo/Nicolae Pacific J. Math. 2021): simple **rate of convergence** and **vast generalization** of the class of permitted spaces X.

Proof Mining papers which appeared since 2023

- H. Cheval, L. Leuștean: Linear rates of asymptotic regularity for Halpern-type iterations. Mathematics of Computation (2024).
- H. Cheval, U. Kohlenbach, L. Leuștean: On modified Halpern and Tikhonov-Mann iterations. J. Optim. Theory Appl. (2023).
- P. Firmino, L. Leuștean: Quantitative asymptotic regularity of the VAM iteration with error terms for m-accretive operators in Banach spaces. Zeitschrift für Analysis und ihre Anwendungen (2024).
- A. Freund, U. Kohlenbach: Bounds for a nonlinear ergodic theorem for Banach spaces. Ergodic Theory and Dynamical Systems (2023).
- B. Dinis, P. Pinto: Strong convergence for the alternating Halpern-Mann iteration in CAT(0) spaces. SIAM J. Optim. (2023).
- R. Findling, U. Kohlenbach: Rates of convergence and metastability for Chidume's algorithm for the approximation of zeros of accretive operators in Banach spaces. Numerical Functional Analysis and Optimization (2024).

- U. Kohlenbach, N. Pischke: Proof theory and nonsmooth analysis. Phil. Trans. Royal Soc. A (2023).
- L. Leuștean, P. Pinto: Rates of asymptotic regularity for the alternating Halpern-Mann iteration. Optimization Letters (2023).
- M. Neri, T. Powell: computational study of a class of recursive inequalities. Journal of Logic and Analysis (2023).
- L. Păunescu, A. Sipoș: A proof-theoretic metatheorem for tracial von Neumann algebras. Math. Logic Quart (2023).
- P. Pinto: On the finitary content of Dykstra's cyclic projections algorithm. Zeitschrift für Analysis und ihre Anwendungen (2024).
- P. Pinto: Nonexpansive maps in nonlinear smooth spaces. Transactions of the American Mathematical Society (2024).

- P. Pinto, N. Pischke: On computational properties of Cauchy problems generated by accretive operators. Documenta Mathematica (2023).
- N. Pischke: Logical Metatheorems for Accretive and (Generalized) Monotone Set-Valued Operators. J. Math. Logic (2024).
- N. Pischke: Quantitative Results on Algorithms for Zeros of Differences of Monotone Operators in Hilbert Space. Journal of Convex Analysis 30, pp. 295-315 (2023).
- N. Pischke: On Korpelevich's extragradient algorithm. Zeitschrift für Analysis und ihre Anwendungen (2023).
- N. Pischke: Rates of convergence for the asymptotic behavior of second-order Cauchy problems. Journal of Mathematical Analysis and Applications (2024).

- N. Pischke: Proof mining for the dual of a Banach space with extensions for uniformly Fréchet differentiable functions. Transactions of the American Mathematical Society (2024),
- T. Powell: A finitization of Littlewood's Tauberian theorem and an application in Tauberian remainder theory. Ann. Pure Appl. Logic, 174(4): 103231, 2023.
- A. Sipoș: On quantitative metastability for accretive operators. To appear in: Zeitschrift für Analysis und ihre Anwendungen.
- A. Sipoș: The computational content of super strongly nonexpansive mappings and uniformly monotone operators. To appear in: Israel J. Math.

Book, surveys and bibliography

- Kohlenbach, U., Applied Proof Theory: Proof Interpretations and their Use in Mathematics. Springer Monograph in Mathematics, 2008.
- 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.

Bibliography on Proof Mining:

https://sites.google.com/view/nicholaspischke/proof-mining-bibliography