

# **Proof-theoretic aspects of obtaining a constructive version of the mean ergodic theorem**

Philipp Gerhardy (joint work with J.Avigad, H.Towsner)

Department of Mathematics, University of Oslo

Logic Colloquium '07, Wrocław

# Introduction

'Proof mining' is the subfield of mathematical logic that is concerned with the extraction of additional information from proofs in mathematics and computer science.

G. Kreisel: *What more do we know if we have proved a theorem by restricted means other than if we merely know the theorem is true?*

⇒ Kreisel's unwinding program/proof mining.

# Introduction

Develop methods to unwind the computational content of ineffective proofs, i.e. proofs using full classical logic.

**Quantitative information:** Effective realizers and bounds (algorithms), complexity of realizers and bounds.

**Qualitative information:** Uniformities (bounds independent of certain parameters), weakening/elimination of premises.

# Introduction

**Aims of proof mining:** Classify theorems and proofs from which additional information can be extracted. Carry out case studies, i.e. analyse actual proofs in mathematics.

**Methods of proof mining:** Proof interpretations. Transform a proof  $P$  of a theorem  $A$  into an enriched proof  $P'$  of an equivalent theorem  $A'$  from which the desired information can be read off.

**Applications of proof mining:** E.g. algebra, analysis (fixed point theory and approximation theory), combinatorics, number theory and computer science - and recently: ergodic theory.

# Introduction

## Why proof mining in ergodic theory?

- Ergodic theory often uses abstract, non-computational structures and techniques.
- Ergodic theory is often concerned with asymptotic behaviour of iterative processes  $\Rightarrow$  extraction of bounds.
- Ergodic theory has connections with many other areas of mathematics, e.g. combinatorics and number theory.

# Overview

- Introduction
- **Mean Ergodic Theorem**
- Metatheorems
- Proof Analysis

# Ergodic Theory

Very informal introduction to ergodic theory:

- Take a measure space  $(X, \mathcal{B}, \mu)$ .
- Consider a measure preserving map  $T : X \rightarrow X$ , i.e.  $\mu(TA) = \mu(A)$  for all  $A \in \mathcal{B}$ .
- Study the asymptotic behaviour of  $T^n A$ .

One can study a measure preserving system by studying the Hilbert space  $L_2(X, \mathcal{B}, \mu)$  and the induced operator  $U_T$ .

# Mean Ergodic Theorem

Let  $(X, \langle \cdot, \cdot \rangle)$  be a Hilbert space,  $T : X \rightarrow X$  a nonexpansive mapping and for  $f \in X$  define  $A_n f := \frac{1}{n+1} \sum_{i=0}^n T^i f$ .

**Mean Ergodic Theorem:** The sequence  $A_n f$  converges in the Hilbert space norm.

**Questions:** Does a computable rate of convergence exist? If not, what kind of rates can we obtain? In what way do those depend on  $f$ ,  $T$  and the space  $(X, \langle \cdot, \cdot \rangle)$ ?



# Mean Ergodic Theorem

No full rate of convergence, even for space  $L_2(X, \mathcal{B}, \mu)$  - counterexamples using the halting problem. Full rate of convergence for  $L_2(X, \mathcal{B}, \mu)$  if  $T$  is ergodic.

Consider classically equivalent, no-counterexample version of the mean ergodic theorem:

**Mean Ergodic Theorem (n.c.i.):** For every  $f \in X$ , nonexpansive  $T : X \rightarrow X$ ,  $M : \mathbb{N} \rightarrow \mathbb{N}$  and  $\varepsilon > 0$ , there is an  $n \in \mathbb{N}$  such that  $\|A_m f - A_n f\| \leq \varepsilon$  for all  $m \in [n, M(n)]$ .

From a standard proof of the mean ergodic theorem, we extract effective bounds for this no-counterexample version.

# Overview

- Introduction
- Mean Ergodic Theorem
- **Metatheorems**
- Proof Analysis

# Metatheorems

In what kind of formal system can we prove the mean ergodic theorem? What kind of additional information do the metatheorems predict?

## Definitions:

- Finite types  $\mathbf{T}$ : (i)  $\mathbb{N} \in \mathbf{T}$ , (ii)  $\rho, \sigma \in \mathbf{T} \Rightarrow \rho \rightarrow \sigma \in \mathbf{T}$ .
- $\text{PA}^\omega$ : Peano Arithmetic in all finite types.
- DC: Axiom schema of dependent choice.
- $\mathcal{A}^\omega := \text{PA}^\omega + \text{DC}$  - classical analysis in all finite types.

# Metatheorems

Using a monotone variant of Gödel's functional ('Dialectica') interpretation, one may prove general metatheorems about the extraction of effective, uniform bounds from proofs of  $\forall\exists A_{qf}$  statements in  $\mathcal{A}^\omega$  (U.Kohlenbach).

Large parts of classical analysis can be formalized in  $\mathcal{A}^\omega$ , in particular concerning complete separable metric spaces.

What about e.g. abstract Hilbert spaces  $(X, \langle \cdot, \cdot \rangle)$ ?

# Metatheorems for Hilbert Spaces

Define the following formal system  $\mathcal{A}^\omega[X, \langle \cdot, \cdot \rangle]$ :

- Let  $\mathsf{T}^X$  be the finite types over  $\mathbb{N}$  and new type  $X$ .
- Extend  $\mathcal{A}^\omega$  to finite types  $\mathsf{T}^X$ .
- New constants:  $0_X, 1_X, +_X, \cdot_X$  and  $\langle \cdot, \cdot \rangle_X$ .
- New axioms: defining axioms for Hilbert spaces.

**G./Kohlenbach(TAMS,2005/to appear):** General logical metatheorems about the extraction of effective uniform bounds from proofs in  $\mathcal{A}^\omega[X, \langle \cdot, \cdot \rangle]$  (and similar theories).

# Metatheorems for Hilbert Spaces

**Definition:** We call the following finite types over  $\mathbb{N}, X$  *small*:  
 $\mathbb{N}, X, \mathbb{N} \rightarrow \mathbb{N}, \mathbb{N} \rightarrow X, X \rightarrow X$ .

**Definition:** A formula  $F_{\forall}$ , i.e.  $F_{\exists}$  is a formula  $\forall \underline{x}^{\sigma} F_{qf}(\underline{x})$ , i.e.  $\exists \underline{x}^{\sigma} F_{qf}(\underline{x})$ , with  $F_{qf}$  quantifier-free and all types  $\sigma_i$  small.

**Definition:** Let  $(X, \langle \cdot, \cdot \rangle)$  be a Hilbert space. An operator  $T : X \rightarrow X$  is nonexpansive – short  $T$  n.e. – if

$$\forall f, g \in X (\|Tf - Tg\| \leq \|f - g\|).$$

# Metatheorems for Hilbert Spaces

**Corollary<sup>a</sup>:** Let  $(X, \langle \cdot, \cdot \rangle)$  be a Hilbert space. If

$$\forall f^X, T^{X \rightarrow X}, k^0, M^1 (T \text{ n.e.} \wedge \\ \forall u^0 B_{\forall}(f, T, k, M, u) \rightarrow \exists v^0 C_{\exists}(f, T, k, M, v)),$$

is provable in  $\mathcal{A}^\omega[X, \langle \cdot, \cdot \rangle]$ , then there is a computable  $\varphi : \mathbb{N} \times \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \rightarrow \mathbb{N}$ , so that

$$\forall f^X, T^{X \rightarrow X}, k^0, M^1 (\|f\|, \|f - Tf\| \leq b \wedge T \text{ n.e.} \wedge \\ \forall u^0 \leq \varphi(b, k, M) B_{\forall} \rightarrow \exists v^0 \leq \varphi(b, k, M) C_{\exists})$$

holds in every Hilbert space  $(X, \langle \cdot, \cdot \rangle)$ .

---

<sup>a</sup>I.e., corollary to most general metatheorems in TAMS-paper.

# Metatheorems for Hilbert Spaces

Easy to check that the mean ergodic theorem can be proved in  $\mathcal{A}^\omega[X, \langle \cdot, \cdot \rangle]$  and (the no-counterexample version) has the suitable, logical form for the metatheorem.

Metatheorem predicts (i.e. guarantees) effective bounds depending only on  $\|f\|$ ,  $M$  and  $\varepsilon$ . Bounds are independent of the space  $(X, \langle \cdot, \cdot \rangle)$  and the mapping  $T$ , and uniform on every norm-bounded, *not necessarily compact* ball in  $X$ .

Joint work with J.Avigad and H.Towsner: Extract effective bounds from standard proof of mean ergodic theorem.



# Overview

- Introduction
- Mean Ergodic Theorem
- Metatheorems
- **Proof Analysis**

# Proof Analysis

Standard proof of mean ergodic theorem:

$$U = \overline{\{u - Tu \mid u \in X\}}, \quad V = \{v \in X \mid v = Tv\}, \quad U \perp V.$$

$$\|A_n(u - Tu)\| \leq \frac{2\|u\|}{n}, \quad A_nv = v, \quad X = U \oplus V.$$

Let  $f \in X$  be given, let  $f_U = u - Tu$  be close to the projection of  $f$  on  $U$  and let  $f^* = f - f_U$ .

Since  $f_U$  is close,  $\|A_n f^*\|$  remains stable for all  $n \in \mathbb{N}$ . Let  $N$  be large enough then  $\|A_n f_U\|$  is small for all  $n \geq N$ .

Then  $\|A_n f - A_m f\|$  small for all  $n, m \geq N$ .

# Proof Analysis

Making this argument effective requires:

- Rate of convergence for some sequence  $g_n = u_n - T u_n$  of approximations to projection of  $f$  onto  $U$ .
- Bounds on the norms of elements  $u_n$ .

We will carry out the effective proof as if we had a rate of convergence for  $g_n$ , but will be able to obtain bounds using a slightly weaker notion of convergence for  $g_n$ .

**Recall:** We only want to find an interval  $[n, M(n)]$  where  $A_n f$  is stable (within a given  $\varepsilon$ ).

# Proof Analysis - bounds on $\|u_n\|$

**Observation:** Sufficient to consider projection of  $f$  onto  $U_f = \overline{\{T^i f - T^{i+1} f \mid i = 0, 1, \dots\}}$ , as  $A_n f$  lies in the span of  $\{f, T f, T^2 f, \dots\}$ .

Define

$$u_0 = \frac{\langle f, f - T f \rangle}{\|f - T f\|^2} f, \quad u_{n+1} = u_n + \frac{\langle f - g_n, T^n f - T^{n+1} f \rangle}{\|T^n f - T^{n+1} f\|^2} T^n f,$$

then  $g_n = u_n - T u_n$  approximates projection of  $f$  onto  $U_f$ .

To bound  $\|u_i\|$  we need lower bound on  $\|T^i f - T^{i+1} f\|$ .

# Proof Analysis - bounds on $\|u_n\|$

**Second observation:** If  $\|f - Tf\|$  is small,  $\|A_n f\|$  is stable for a long time (using triangle inequality).

Either  $\|f - Tf\|$  is small enough to get  $\|A_{M(0)} f - f\| \leq \varepsilon$ . Or we have a lower bound on  $\|f - Tf\|$  and thus an upper bound on  $\|u_0\|$ .

Similar, for general  $\|u_n\|$ , i.e. we have a sequence of upper bounds for  $\|u_n\|$  in terms of  $\|f\|$ ,  $\varepsilon$  and  $M$ .

# Proof Analysis - convergence of $g_n$

Let  $u_n$  and  $g_n$  be as before. Assume we have a modulus of convergence for  $g_n$ . For any  $i$

$$\|A_{M(n)}f - A_n f\| \leq \|A_{M(n)}(f - g_i) - A_n(f - g_i)\| + \|A_{M(n)}g_i\| + \|A_n g_i\|.$$

By direct calculation we find a  $\delta > 0$  s.t. if  $\|g_i - g_j\| \leq \delta$  for all  $j > i$ , then  $\|A_{M(n)}(f - g_i) - A_n(f - g_i)\| \leq \varepsilon/2$  for all  $n$ .

Using a modulus of convergence for  $g_n$ , we find such an  $i$ .

Making  $n$  large enough, we get  $\|A_{M(n)}g_i\| + \|A_n g_i\| \leq \varepsilon/2$ .

# Proof Analysis - convergence of $g_n$

For all  $n$ , we have  $\|g_n\| \leq \|g_{n+1}\| \leq \|f\|$ . By the principle of convergence for bounded, monotone sequences, the sequence  $\|g_n\|$  converges (and thus also  $g_n$  converges).

Interpreting (countable) choice to obtain a modulus of convergence for  $g_i$  requires bar-recursion. This would cause very complicated bounds for our constructive mean ergodic theorem.

Inspired by a technique (due to Kohlenbach) to eliminate certain simple instances of choice such as PCM, we observe the following:

# Proof Analysis - convergence of $g_n$

We wrote: “find an  $i$  s.t.  $\|g_j - g_i\| \leq \delta$  for all  $j > i$ ”.

In fact: We only need  $\|g_j - g_i\| \leq \delta$  for a specific  $j$  given in terms of  $i$  and other parameters of the theorem. This yields a sequence of disjoint nonempty intervals  $[i_k, j_k]$ .

Using monotonicity and boundedness of  $\|g_n\|$ , we see that  $\|g_{j_k} - g_{i_k}\| \geq \delta$  only for finitely many  $[i_k, j_k]$ . Thus for one of those intervals we have the result.

Final quirk: The  $j_k$  depend on  $n, M(n)$  - but the number of intervals  $[i_k, j_k]$  we need to consider is independent of  $n$ .



# Proof Analysis - putting it all together

Solution: Define  $i_0 = 0$ ,  $n_k$  in terms of  $i_k$  and  $i_{k+1}$  in terms of  $n_k$  ( $j_k = i_{k+1}$ ):

$$i_0 := 0 \qquad n_k := \lceil \frac{b^2}{\varepsilon^2} \sum_{l=0}^{i_k} M(\frac{2lb}{\varepsilon}) \rceil,$$
$$i_{k+1} := i_k + \lceil \frac{2^{15} M(n_k)^4 b^4}{\varepsilon^4} \rceil.$$

Let  $d = \frac{512b^2}{\varepsilon^2}$  and  $N(b, k, n) = \frac{2n_d b}{\varepsilon}$ , then

$$\forall f^X, T^{X \rightarrow X}, \varepsilon > 0, M : \mathbb{N} \rightarrow \mathbb{N} \exists n \leq N(\|f\| \leq b \wedge T \text{ n.e.} \wedge M(n) > n \rightarrow \|A_{M(n)} f - A_n f\| \leq \varepsilon).$$

# Conclusions

We have obtained a constructive proof of the following:

For every  $f \in X$ ,  $T : X \rightarrow X$  nonexpansive,  $M : \mathbb{N} \rightarrow \mathbb{N}$ ,  $\varepsilon > 0$  there exists an  $n \leq N$  – computable in  $\|f\|$ ,  $M$  and  $\varepsilon$  – such that  $\|A_m f - A_n f\| \leq \varepsilon$  for every  $m \in [n, M(n)]$ .

- Proof theoretic guidelines to put proof (definitions, lemmas, theorem) into suitable logical form.
- Interpretation of lemmas: mostly direct computation.
- Elimination of instance of PCM.
- Extraction of computable bound as predicted by metatheorems.

# References

- Kohlenbach - *Some logical metatheorems with applications to functional analysis*. Trans. Am. Math. Soc. (2005).
- G./Kohlenbach - *General logical metatheorems for functional analysis*. Trans. Am. Math. Soc. (to appear).
- Kohlenbach - *Elimination of Skolem functions for monotone formulas in analysis*. Arch. Math. Log. (1998).
- Avigad/G./Towsner - *Local stability of ergodic averages*. Submitted.