Computing uniform (metastable) rates of convergence from the statement of the theorem alone

Jason Rute

Pennsylvania State University

New England Recursion and Definability Seminar 8.0
October 17, 2015

Slides available at
www.personal.psu.edu/jmr71/

(Updated on October 17, 2015.)

Convergence

Convergence theorems

Monotone convergence principle

Let $(c_n)_{n\in\mathbb{N}}$ be a bounded nondecreasing sequence of reals in the interval $[a,b]$. Then $(c_n)_{n\in\mathbb{N}}$ converges.

Mean ergodic theorem

Let $(X,\|\cdot\|)$ be a reflexive Banach space with a nonexpansive linear transformation $T\colon X\to X$,

$$T(ax+y)=aT(x)+T(y) \quad \text{and} \quad \|T(x)\|\leqslant\|x\|.$$

Then for $c\in X$, the ergodic averages $\frac{1}{n}\sum_{k<n}T^k(c)$ converge.

Our setup

- Let $\mathcal{M}$ be a complete metric space (X, d) with possible additional structure.
- Let $(c_n)_{n \in \mathbb{N}}$ be a distinguished sequence in X.
- Let P be a property that could hold of $(\mathcal{M}, (c_n)_{n \in \mathbb{N}})$.

Convergence theorem template

If P holds of the pair $(\mathcal{M}, (c_n)_{n \in \mathbb{N}})$, then c_n converges.

Question

For which properties P is the rate of convergence

- **uniform**—exists single rate for all pairs $(\mathcal{M}, (c_n)_{n \in \mathbb{N}})$ satisfying P?
- **computable**—rate is computable uniformly from $(\mathcal{M}, (c_n)_{n \in \mathbb{N}})$?
- **computably uniform**—exists single computable uniform rate?

Bait-and-switch

- This talk is not about usual Cauchy rates of convergence...

Metastable convergence

Three ways to say converge

The following are all equivalent ways to say that $(c_n)_{n\in\mathbb{N}}$ converges.

- $(c_n)_{n\in\mathbb{N}}$ is Cauchy **(contains a lot of information, but not very uniform)**

$$\underbrace{\forall \varepsilon > 0\ \exists m \in \mathbb{N}}_{\text{rate of convergence}}\ \forall n, n' \geqslant m\ d(c_n, c_{n'}) < \varepsilon.$$

- $(c_n)_{n\in\mathbb{N}}$ has finitely many ε-jumps

$$\underbrace{\forall \varepsilon > 0\ \exists n \in \mathbb{N}}_{\text{rate of convergence}}\ \forall u_0 \leqslant v_0 \leqslant u_1 \leqslant v_1 \leqslant \ldots \leqslant v_{n-1} \leqslant v_{n-1}$$

$$\exists k \in [0, n-1]\ d(c_{u_k}, c_{v_k}) < \varepsilon.$$

  - (Similar to upcrossing bounds and variational bounds.)
- $(c_n)_{n\in\mathbb{N}}$ is metastable **(very uniform, but contains little information)**

$$\underbrace{\forall \varepsilon > 0\ \forall F : \mathbb{N} \to \mathbb{N}\ \exists m \in \mathbb{N}}_{\text{rate of convergence}}\ \forall n, n' \in [m, F(m)]\ d(c_n, c_{n'}) < \varepsilon.$$

Syntactic side of metastability

- Metastability is an application of Gödel's Dialectica interpretation:

$$\forall \varepsilon > 0 \quad \exists m \in \mathbb{N} \quad \forall n, n' \geqslant m \quad d(c_n, c_{n'}) < \varepsilon$$
$$\forall \varepsilon > 0 \quad \exists m \in \mathbb{N} \quad \forall k\, \forall n, n' \in [m, k] \quad d(c_n, c_{n'}) < \varepsilon$$
$$\forall \varepsilon > 0 \quad \forall F \in \mathbb{N}^{\mathbb{N}} \quad \exists m \in \mathbb{N} \quad \forall n, n' \in [m, F(m)] \quad d(c_n, c_{n'}) < \varepsilon$$
$$\forall \varepsilon > 0 \quad \forall F \in \mathbb{N}^{\mathbb{N}} \quad \exists \ell\, \exists m \leqslant \ell \quad \forall n, n' \in [m, F(m)] \quad d(c_n, c_{n'}) < \varepsilon$$
$$\exists \Psi \in (\mathbb{Q}^{+} \times \mathbb{N}^{\mathbb{N}})^{\mathbb{N}} \quad \forall \varepsilon > 0 \quad \forall F \in \mathbb{N}^{\mathbb{N}} \quad \exists m \leqslant \Psi(\varepsilon, F) \quad \forall n, n' \in [m, F(m)] \quad d(c_n, c_{n'}) < \varepsilon$$

- A metastable rate $\Psi(\varepsilon, F)$ is always computable from ε, F, and $\{c_n\}_{n \in \mathbb{N}}$.
- This idea works for other Π^0_3 statements.

Example: Monotone convergence principle

Monotone convergence principle

Let $(c_n)_{n\in\mathbb{N}}$ be a bounded nondecreasing sequence of reals in the interval $[a,b]$. Then $(c_n)_{n\in\mathbb{N}}$ converges.

Metastable monotone convergence principle

Let $(c_n)_{n\in\mathbb{N}}$ be a bounded nondecreasing sequence of reals in the interval $[a,b]$. Then $(c_n)_{n\in\mathbb{N}}$ converges with a metastable rate of convergence

$$\Psi(\varepsilon,F) = F^{\lfloor (b-a)/\varepsilon \rfloor}(0),$$

that is the iterated function $F^n(0) = F \dots F(0)$ for $n = \lfloor (b-a)/\varepsilon \rfloor$.

Example: Mean ergodic theorem

Mean ergodic theorem

Let $(X, \|\cdot\|)$ be a reflexive Banach space with a nonexpansive linear transformation $T\colon X \to X$. Then for $c \in X$, the ergodic averages $\frac{1}{n}\sum_{k<n} T^k(c)$ converge.

No metastable bounds for reflexive spaces (Avigad, Rute)

There is a reflexive Banach space $(X, \|\cdot\|)$ such that there is no uniform metastable rate of convergence for all x in the unit ball of X.

Example: Mean ergodic theorem

Mean ergodic theorem (uniformly convex spaces)

Let $(X, \|\cdot\|)$ be a uniformly convex Banach space with modulus $\eta(\varepsilon)$,

$$\|x\|, \|y\| \leqslant 1 \wedge \|x-y\| \geqslant \varepsilon \quad \to \quad \|(x+y)/2\| \leqslant 1-\eta(\varepsilon).$$

a nonexpansive linear transformation $T\colon X \to X$. Then for $c \in X$, the ergodic averages $\frac{1}{n}\sum_{k<n} T^k(c)$ converge.

Metastable mean ergodic theorem (Kohlenbach, Leuştean 2009)

Let $(X, \|\cdot\|)$ be a uniformly convex Banach space with modulus $\eta(\varepsilon)$, let T be a nonexpansive linear transformation. Let $x \in X$. Then $\frac{1}{n}\sum_{k<n} T^k(x)$ converges with a metastable rate of convergence of

$$\Psi(\varepsilon, F) = \tilde{F}^k(0), \quad \text{where} \quad k = O(\rho \log \rho \cdot \eta(1/(8\rho))^{-1}), \quad \rho = \|x\|/\varepsilon,$$

and $\tilde{F}$ is a function that grows slightly faster than F.

Why metastability?

- Analysis
  - Any rate of convergence is better than no rate.
  - Rates of metastable convergence are more uniform.
  - Better rates may not be known (or even possible?).
  - May give simplest or most accessible proof of convergence.
  - An alternative to nonstandard analysis.
  - Example: Tao's ergodic theorem for multiple commuting averages.
- Logic
  - Metastable rates are computable.
  - Metastable convergence theorems are constructive.
  - Proof theoretic methods exist to extract metastable bounds: proof mining.
  - Closely connected to ultraproducts and nonstandard analysis.
  - **Uniform metastable rates can be computed from the statement of the theorem alone! (This talk.)**

Results: prior and new

Let P be a property that could hold of an arbitrary metric structure with a distinguished subsequence: $\mathbb{X} = (X, d, \ldots, \{c_n\}_{n\in\mathbb{N}})$. Consider a theorem:

(∗) If P holds of $\mathbb{X}$ then c_n converges.

- Kohlenbach. *Some logical metatheorems with applications to functional analysis.* Trans AMS, 2004.
  - If (1) the theorem (∗) is provable in $A^\omega[X,d]$,[1] and
  - (2) P is expressible by a ∀-formula (basically a Π^0_1 property), then
  - there there exists a computable uniform metastable rate of convergence (uniformly extractable from the proof of (∗) in $A^\omega[X,d]$).
- Avigad, Iovino. *Ultraproducts and metastability.* NYJM, 2013.
  - If $\mathcal{C} = \{\mathbb{X} : \mathbb{X} \text{ satisfies } P\}$ is closed under ultraproducts, then
  - there is a uniform rate of metastable convergence.
- R. (This talk)
  - If P is axiomitizable by a set of sentences Σ in continuous logic, then
  - there is a uniform rate of metastable convergence computable from Σ.

[1] $A^\omega[X,d]$ is a type theory extending $\mathsf{PA} + \mathsf{DC}$ with a type for X and axioms for the metric d.

Main result

Continuous logic

- Continuous logic is a logic for dealing with "metric structures."
- There have been many variants over the years.
  - Many-valued logics (Łukasiewicz)
  - Continuous Logic (Chang and Keisler)
  - Positive Bounded Logic (Henson and Iovino)
  - Continuous Logic (Ben Yaacov)
- We will use the variant is due to Ben Yaacov.
- Ben Yaacov's version very much resembles first-order logic!
  - Compactness, completeness, Löwenheim-Skolem, ultraproducts, etc.
- See survey article *Model theory for metric structures* for a good introduction.

First-order logic vs. Continuous first-order logic

	First-Order Logic	Continuous First-Order Logic
Universe	set $(M,=)$	comp. bdd. metric space (M,d) $\quad (d(x,y) \leqslant 1)$
Truth values	T and F	$[0,1]$ $\quad (0 = \mathit{true}, 1 = \mathit{false})$
Func. symbol	symbol f $\quad$ (arity n)	symbol f $\quad$ (arity n and mod. of cont. $\delta(\varepsilon)$)
Functions	$f^{\mathcal{M}} : M^n \to M$	$f^{\mathcal{M}} : M^n \to M$ $\quad$ (obeys mod. of cont. $\delta(\varepsilon)$)
Rel. symbol	symbol R $\quad$ (arity n)	symbol R $\quad$ (arity n and mod. of cont. $\delta(\varepsilon)$)
Relations	$R^{\mathcal{M}} : M^n \to \{T,F\}$	$R^{\mathcal{M}} : M^n \to [0,1]$ $\quad$ (obeys mod. of cont. $\delta(\varepsilon)$)
Connectives	$\odot : \{T,F\}^n \to \{T,F\}$	$\odot : [0,1]^n \to [0,1]$ $\quad$ (continuous)
Sufficient con.	$\bot$, $\vee$, $\wedge$, $\to$	1, min, max, $\dot{-}$, $x \mapsto x/2$
Quantifiers	$\exists x\ \varphi(x)$, $\forall x\ \varphi(x)$	$\min_x \varphi(x)$, $\max_x \varphi(x)$
Formulas	$\varphi(\bar{x})$	$\varphi(\bar{x})$
Statements	Sentence: φ	Conditions: $[\varphi = 0]$, $[\varphi > 0]$
Evaluation	$\varphi^{\mathcal{M}} : M^n \to \{T,F\}$	$\varphi^{\mathcal{M}} : M^n \to [0,1]$
Satisfaction	$\mathcal{M} \models \varphi$ $\quad$ iff $\quad \varphi^{\mathcal{M}} = T$	$\mathcal{M} \models [\varphi = 0]$ $\quad$ iff $\quad \varphi^{\mathcal{M}} = 0$
Axioms/rules	complete, comp. list	complete, computable list
Provability	$\Sigma \vdash \varphi$	$\Sigma \vdash [\varphi > 0]$ $\quad$ (Σ set of conditions $[\psi = 0]$)

Main theorem

Theorem (R.)

Let $\mathcal{L}$ be a computable signature containing constants $\{c_n\}_{n\in\mathbb{N}}$.
Let Σ be a set of $\mathcal{L}$-conditions (using connectives 1, min, max, $\dot{-}$, $\cdot/2$).
Assume $(c_n^{\mathcal{M}})_{n\in\mathbb{N}}$ converges for all $\mathcal{M} \models \Sigma$.
Then there is a uniform rate of metastability for $(c_n^{\mathcal{M}})_{n\in\mathbb{N}}$ computable in Σ.

Proof.

- Fix $F\colon \mathbb{N} \to \mathbb{N}$ and $\varepsilon = 2^{-k}$, our goal is to find ℓ such that

$$\exists m < \ell \, \forall n, n' \in [m, F(m)] \; d^{\mathcal{M}}(c_n^{\mathcal{M}}, c_{n'}^{\mathcal{M}}) < \varepsilon.$$

- The quantifiers are bounded, so this can be written as the $\mathcal{L}$-condition

$$\underbrace{\max_{m<\ell} \; \min_{n,n' \in [m,F(m)]} \left(1/2^k \dot{-} d(c_n, c_{n'})\right)}_{\varphi_\ell} > 0.$$

Main theorem (cont.)

Proof.

- Fix $F\colon \mathbb{N} \to \mathbb{N}$ and $\varepsilon = 2^{-k}$, our goal is to find ℓ such that for all $\mathcal{M} \models \Sigma$,

$$\exists m \leqslant \ell \, \forall n, n' \in [m, F(m)] \; d^{\mathcal{M}}(c_n^{\mathcal{M}}, c_{n'}^{\mathcal{M}}) < \varepsilon.$$

- The quantifiers are bounded, so this can be written as the $\mathcal{L}$-condition

$$\underbrace{\max_{m<\ell} \min_{n,n' \in [m,F(m)]} \left(1/2^k \dot{-} d(c_n, c_{n'})\right)}_{\varphi_\ell} > 0.$$

- Notice φ_ℓ is computable in F, k, and ℓ.
- For all $\mathcal{M} \models \Sigma$, since $(c_n^{\mathcal{M}})_{n\in\mathbb{N}}$ converges, there exists ℓ s.t. $\mathcal{M} \models [\varphi_\ell > 0]$.
- By the compactness theorem, there is some ℓ such $\Sigma \models [\varphi_\ell > 0]$.
- By the completeness theorem, $\Sigma \vdash [\varphi_\ell > 0]$ for some ℓ.
- Compute ℓ by searching for a finite proof of $\Sigma \vdash [\varphi_\ell > 0]$ for some ℓ. □

Examples / Applications

Monotone Convergence

Monotone convergence principle

Let $(c_n)_{n\in\mathbb{N}}$ be a bounded nondecreasing sequence of reals in the interval $[0,1]$. Then $(c_n)_{n\in\mathbb{N}}$ converges.

MCP in Continuous Logic

- Let $\mathcal{L}$ include $\{c_n\}_{n\in\mathbb{N}}$ in addition to a constant for each dyadic rational.
- Let Σ be sentences which
  - give the distances between all dyadic rationals,
  - specify that all points are distance $\leqslant 2^{-j}$ from $\{0,2^{-j},\ldots,i2^{-j},\ldots,1\}$.

$$\min\left\{d(x,0),d(x,2^{-j}),\ldots,d(x,i2^{-j}),\ldots,d(x,1)\right\} \leqslant 2^{-j}.$$

  - say that $(c_n)_{n\in\mathbb{N}}$ is nondecreasing, $\quad d(0,c_n) \leqslant d(0,c_{n+1})$.
- This axiomatizes a nondecreasing sequence in $[0,1]$.
- Hence, there is a uniform computable rate of metastable convergence.

Mean Ergodic Theorem

Mean ergodic theorem

Let $(X, \|\cdot\|)$ be a uniformly convex Banach space with a nonexpansive linear transformation $T\colon X \to X$. Then for $c \in X$, the ergodic averages $\frac{1}{n}\sum_{k<n} T^k(c)$ converge.

Mean ergodic theorem in Continuous Logic

- The language of the unit ball of a Banach space includes $f_{q,p}(x,y) = qx + py$ for each pair of rationals $|p| + |q| \leqslant 1$, the norm $\|\cdot\|$ as a unary relation, and 0. This can be axiomatized.
- Let $\mathcal{L}$ be this in addition to the function T and the constants $\{c_n\}$.
- Since T is nonexpansive, we know the modulus of continuity.
- Linearity, and nonexpansivity can all be axiomatized.
- Uniform convexity (for a fixed modulus $\eta(\varepsilon)$) can be axiomitized.
- Add axioms defining c_n: $\quad c_n = \frac{1}{n}\sum_{k<n} T^k(c_0)$.

Mean Ergodic Theorem

Mean ergodic theorem

Let $(X, \|\cdot\|)$ be a reflexive Banach space with a nonexpansive linear transformation $T\colon X \to X$. Then for $c \in X$, the ergodic averages $\frac{1}{n}\sum_{k<n} T^k(c)$ converge.

Reflexivity not axiomitizable

Unlike uniform convexity, reflexivity is not axiomitizable in continuous logic.

Examples / Applications

Comparing the Avigad and Iovino result

- Let $\mathcal{C}$ be a class of structures $(X, d, \ldots, \{c_n\}_{n\in\mathbb{N}})$ which is closed under ultraproducts such that $\{c_n\}_{n\in\mathbb{N}}$ converges for all structures in the class.
- Avigad and Iovino showed that the rate of metastable convergence is uniform.

- Let $\mathcal{D}$ be $\mathcal{C}$ closed under elementary equivalence.
- If $\mathcal{M} \equiv \mathcal{N}$ then

$$d^{\mathcal{M}}(c_m^{\mathcal{M}}, c_n^{\mathcal{M}}) = d^{\mathcal{N}}(c_m^{\mathcal{N}}, c_n^{\mathcal{N}})$$

  hence $\{c_n\}_{n\in\mathbb{N}}$ still converges for all structures in $\mathcal{D}$.
- Moreover, $\mathcal{D}$ is now an axiomitizable class with axioms Σ.
- By my result, the rate of metastability is uniform **and computable from Σ.**

Comparing the Kohlenbach result

- Kohlenbach's result requires either
  - Proving the theorem in $A^\omega[X,d]$, or
  - using proof mining techniques to extract a bound.
- Whereas, my result only requires that the theorem is true and that property P of the space is axiomitizable in continuous logic.

- Nonetheless, Kohlenbach's result has many advantages:
  - It extracts computable bounds from the proof by a bar recursive functional.
  - If the proof avoids DC, then the bounds are primitive recursive in the sense of Gödel.
  - If the proof uses only Σ^0_1-IA and WKL, then the bounds are primitive recursive.
  - The bounds are actually formulas (although complicated looking).
  - Can be applied to discontinuous functions.
- Guenzel and Kohlenbach recently extended this work to be compatible with positive bounded logic.

Closing Thoughts

Summary

- There are many compatible logical tools for investigating theorems in analysis:
  - type theory and the dialectic interpretation,
  - ultraproducts, and
  - continuous logic.
- Usual (computability theoretic, proof theoretic, model theoretic) methods from first order logic extend nicely to continuous logic.
- There is a lot of potential to investigate computable continuous model theory; it is a nice merger of computable analysis and computable model theory.

- Can continuous logic be applied to proof mining in a useful way?
- Can logic methods be used to study other types of rates of convergence?

Thank You!

These slides will be available on my webpage:

http://www.personal.psu.edu/jmr71/

Or just Google™ me, “Jason Rute”.

P.S. I am on the job market.