Metastability in ergodic theory

Jeremy Avigad

Department of Philosophy and Department of Mathematical Sciences
Carnegie Mellon University

(joint work with Philipp Gerhardy, Ksenija Simic, and Henry Towsner)

January, 2009

The mean ergodic theorem

Let $\mathcal{H}$ be a Hilbert space, and let $T : \mathcal{H} \to \mathcal{H}$ be a nonexpansive mapping, i.e. satisfying $\|Tf\| \leq \|f\|$ for every f.

For each $n \geq 1$, define $A_n f$ to be the function $\frac{1}{n} \sum_{i<n} T^i f$.

Theorem (von Neumann / Riesz). The sequence $(A_n f)$ converges in the Hilbert space norm.

Question: can we compute a bound on the rate of convergence of $(A_n f)$ from the initial data (T and f)?

In other words: can we compute a function $r : \mathbb{Q} \to \mathbb{N}$ such that for every rational $\varepsilon > 0$,

$$\|A_m f - A_{r(\varepsilon)} f\| < \varepsilon$$

whenever $m \geq r(\varepsilon)$?

Rates of convergence

Krengel (et al.): convergence can be arbitrarily slow. But computability is a different question.

For example, if $(a_n)_{n \in \mathbb{N}}$ is a sequence of reals that decreases to 0, no matter how slowly, one can compute a bound on the rate of convergence from (a_n).

But there are bounded, computable, decreasing sequences (b_n) of rationals that do not have a computable limit.

There are also computable sequences (c_n) of rationals that converge to 0, with no computable bound on the rate of convergence.

Conclusion: at issue is not the *rate* of convergence, but its *predictability*.

A negative result

Theorem (A-S). There are a computable measure-preserving transformation of $[0,1]$ under Lebesgue measure and a computable characteristic function $f = \chi_A$, such that if $f^* = \lim_n A_n f$, then $\|f^*\|_2$ is not a computable real number.

In particular, f^* is not a computable element of $L^2(\mathcal{X})$, and there is no computable bound on the rate of convergence of $(A_n f)$ in either the L^2 or L^1 norm.

A positive result

Theorem (A-G-T). Let T be a nonexpansive operator on a separable Hilbert space and let f be an element of that space. Let $f^* = \lim_n A_n f$. Then f^*, and a bound on the rate of convergence of $(A_n f)$ in the Hilbert space norm, can be computed from f, T, and $\|f^*\|$.

In particular, if T arises from an ergodic transformation of a measure space, then f^* is computable from T and f.

A more explicit mean ergodic theorem

Even when there is no computable bound on the rate of convergence, there is more information to be had.

The assertion that the sequence $(A_n f)$ converges can be represented as follows:

$$\forall \varepsilon > 0 \, \exists n \, \forall m \geq n \, (\|A_m f - A_n f\| < \varepsilon).$$

This is classically equivalent to the assertion that for any function K,

$$\forall \varepsilon > 0 \, \exists n \, \forall m \in [n, K(n)] \, (\|A_m f - A_n f\| < \varepsilon).$$

A more explicit mean ergodic theorem

Theorem (A-G-T). Let T be any nonexpansive operator on a Hilbert space, let f be any element of that space, and let $\varepsilon > 0$, and let K be any function. Then there is an $n \geq 1$ such that for every m in $[n, K(n)]$, $\|A_m f - A_n f\| < \varepsilon$.

In fact, there is a bound on n that depends only on K and $\rho = \|f\|/\varepsilon$ (and is independent of T).

As special cases, we have the following:

- If $K = n^{O(1)}$, then $n(f, \varepsilon) = 2^{2^{O(\rho^2 \log\log \rho)}}$.
- If $K = 2^{O(n)}$, then $n(f, \varepsilon) = 2^{1}_{O(\rho^2)}$.
- If $K = O(n)$ and T is an isometry, then $n(f, \varepsilon) = 2^{O(\rho^2 \log \rho)}$.

Metastability

We have similar results for the pointwise ergodic theorem.

The central idea: if one is interested in pockets of approximate stability rather than exact limits, one can obtain stronger uniformity and/or computability results.

We called this phenomenon “local stability.”

Terence Tao has used the phrase “metastability.”

The Furstenberg-Zimmer tower

Let $\mathcal{X} = (X, \mathcal{B}, \mu, T)$ be a measure preserving system.

If $\mathcal{Y}$ is any factor of $\mathcal{X}$, let $Z(\mathcal{Y})$ denote the maximal compact (isometric) extension of $\mathcal{Y}$.

Define a transfinite sequence of factors:

- $\mathcal{Y}_0$ is the trivial factor.
- For every α, $\mathcal{Y}_{\alpha+1} = Z(\mathcal{Y}_\alpha)$.
- For every limit λ, $\mathcal{Y}_\lambda$ is the factor generated by $\bigcup_{\alpha<\lambda} \mathcal{Y}_\alpha$.

If $\mathcal{X}$ is separable, the process stabilizes at some countable α. $\mathcal{Y} = \mathcal{Y}_\alpha$ is the *maximal distal factor*.

The Furstenberg-Zimmer structure theorem

Theorem (Furstenberg-Zimmer). $\mathcal{X}$ is weak mixing relative to the maximal distal factor, $\mathcal{Y}$.

Furstenberg observed (and he and Katznelson later spelled out the details) that this can be used to give a very perspicuous proof of Szemerédi's theorem, a statement of ordinary (finitary) combinatorics.

Beleznay and Foreman have shown that for every countable α, there is a separable system such that the tower has height α.

These two facts are striking.

Metastability in the Furstenberg-Zimmer tower

Saying that $\mathcal{X}$ is weak mixing relative to the maximal distal factor, $\mathcal{Y}$, means that for every f and g in $L^\infty(\mathcal{X})$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i<n} \int \left[E(fT^i g \mid \mathcal{Y}) - E(f \mid \mathcal{Y}) E(T^i g \mid \mathcal{Y})\right]^2 d\mu = 0.$$

Theorem. For every f and g in $L^\infty(\mathcal{X})$ and $\varepsilon > 0$, there are m and $\alpha < \omega$ such that for all $n \geq m$,

$$\frac{1}{n} \sum_{i<n} \int \left[E(fT^i g \mid \mathcal{Y}_\alpha) - E(f \mid \mathcal{Y}_\alpha) E(T^i g \mid \mathcal{Y}_\alpha)\right]^2 d\mu < \varepsilon.$$

(This is not hard. Hint: find $\alpha < \omega$ such that $E(f \mid \mathcal{Y}_{\alpha+1}) - E(f \mid \mathcal{Y}_\alpha)$ is sufficiently small.)

Metastability in the Furstenberg-Zimmer tower

In fact, relative weak mixing implies relative weak mixing of all orders: for all k and $f_0, \ldots, f_{k-1}$ in $L^\infty(\mathcal{X})$,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i<n} \int \left(\prod_{l<k} E(T^{ln} f_l \mid \mathcal{Y}) - \prod_{l=1}^{k} T^{ln} E(f_l \mid \mathcal{Y}) \right)^2 = 0.$$

Theorem (A-T). For every k, $f_0, \ldots, f_{k-1}$ in $L^\infty(\mathcal{X})$, and $\varepsilon > 0$, there are m and $\alpha < \omega^{\omega^\omega}$ such that for every $n \geq m$,

$$\frac{1}{n} \sum_{i<n} \int \left(\prod_{l<k} E(T^{ln} f_l \mid \mathcal{Y}_\alpha) - \prod_{l<k} T^{ln} E(f_l \mid \mathcal{Y}_\alpha) \right)^2 < \varepsilon.$$

This fact suffices for the proof of Szemerédi's theorem.

Conclusions

In fact, Furstenberg's original proof shows that for each k, the kth distal factor is characteristic for the averages in question. So we already knew that relative weak mixing and the full transfinite tower are not needed in the proof of Szemerédi's theorem.

But our results provide a general explanation of why the full tower is not needed in the Furstenberg-Katznelson proof.

Goals and future work:

- See what other data can be mined from proofs in ergodic theory and applications to combinatorics and number theory.
- Gain a better understanding of the role that nonconstructive methods play in proofs of concrete or computational results.