How Constructive is Furstenberg's Proof of the Multiple Recurrence Theorem?

Henry Towsner

UCLA

December 11, 2008

(joint work with Jeremy Avigad)

Dynamical Systems

Definition

A dynamical system is a tuple $(X, \mathcal{B}, \mu, T)$ such that:

- $\mathcal{B}$ is a σ-algebra on X
- μ is a σ-additive probability (that is, valued in $[0, 1]$) measure on $\mathcal{B}$
- The T is a measurable, measure-preserving transformations on X

We mostly work with the space of L^2 functions on $(X, \mathcal{B}, \mu, T)$: functions f such that f^2 is integrable. These functions form a Hilbert space.

The Multiple Recurrence Theorem

Theorem

Let $(X, \mathcal{B}, \mu, T)$ be a dynamical system, and let $A \in \mathcal{B}$ with $\mu(A) > 0$. Then there is some d such that

$$\mu\Big(\bigcap_{i \leq k} T^{di} A\Big) > 0.$$

Indeed, the stronger SZ property holds:

Theorem

Let $(X, \mathcal{B}, \mu, T)$ be a dynamical system, and let $A \in \mathcal{B}$ with $\mu(A) > 0$. Then there is some d such that

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{d < N} \mu\Big(\bigcap_{i \leq k} T^{di} A\Big) > 0.$$

Weak Mixing

Definition

A dynamical system is *weak mixing* if for all L^2 functions f, g,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i<N} \left[\int fT^i g d\mu - \int f d\mu \int g d\mu \right]^2 = 0.$$

Theorem

If a dynamical system is weak mixing, multiple recurrence holds.

Compactness

Definition

A function f in a dynamical system is *compact* (or *almost periodic*) if the orbit

$$\{f, Tf, T^2 f, \ldots, T^n f, \ldots\}$$

is totally bounded.

Theorem

If χ_A is compact, multiple recurrence holds for A.

Compact Factors

Theorem

If $(X, \mathcal{B}, \mu, T)$ *is not weak mixing, there is a non-trivial compact function* f*.*

Definition

A *factor* of a dynamical system $(X, \mathcal{B}, \mu, T)$ is given by a σ-algebra $\mathcal{C} \subseteq \mathcal{B}$ closed under T.

The Structure Theorem

Theorem

- *A compact extension of an SZ factor is SZ*
- *A limit of SZ factors is SZ*
- *A weak mixing extension of an SZ factor is SZ*
- *If $(X, \mathcal{B}, \mu, T)$ is not weak mixing relative to a factor $\mathcal{C}$ then there is a non-trivial compact extension contained in $\mathcal{B}$*

Proof of Multiple Recurrence: Take $\mathcal{C}$ to be the maximal factor formed by compact extensions and limits. This factor is SZ, and $(X, \mathcal{B}, \mu, T)$ is weak mixing relative to this factor, therefore $(X, \mathcal{B}, \mu, T)$ is SZ.

Transfinite Induction?

Is it really necessary to use a transfinite induction, and if so, how far do we really need to go?

On the one hand:

Theorem (Beleznay and Foreman)

For every countable α, there is a measure space such that the Furstenberg-Zimmer tower has height α.

Determining whether a measure space is equal to the Furstenberg-Zimmer tower is Π^1_1 complete.

Transfinite Induction?

On the other hand:

Theorem (Furstenberg)

The factor $\mathcal{Y}_k$ is sufficient to calculate the average

$$\liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mu\Big(\bigcap_{i\leq k+2} T^{ni}A\Big).$$

Theorem (Host-Kra, Ziegler)

A factor $\mathcal{Z}_k$, the "k-step nilpotent factor," contained in $\mathcal{Y}_k$ (but, in general, much smaller) is sufficient to calculate the average

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mu\Big(\bigcap_{i\leq k+2} T^{ni}A\Big).$$

Formalizing the Proof

With some coding effort, we can formalize Furstenberg's argument in the theory ID_1.

ID_1 is Peano Arithmetic plus an inductive predicate. That is, there is a set I with the axioms

$$n \in I \leftrightarrow A(n, I)$$

and

$$\forall x[A(x, P) \rightarrow P[x]] \rightarrow \forall x[x \in I \rightarrow P[x]]$$

where $A(x, X)$ is an arithmetic formula where X appears positively.

This theory is equivalent $\mathbf{\Pi^1_1} - \mathbf{CA_0^-}$.

ID_1 with Ordinals

We may reformulate ID_1 to make the ordinal induction explicit. We define a type of ordinals and two predicates $I(\alpha, n)$, $I(< \alpha, n)$ with the axioms

- $\neg I(0, n)$
- $I(< \alpha, n) \leftrightarrow \exists \beta < \alpha I(\beta, n)$
- $A(n, \lambda x.I(< \alpha, x)) \leftrightarrow I(\alpha, n)$
- $A(n, \lambda x.\exists \alpha I(< \alpha, x)) \rightarrow I(< \Omega, n)$

A Dialectica Translation for ID_1

Definition

The theory of functionals OR_1 contains two base types, $\mathbb{N}$ and the type Ω of well-founded trees of integers, and allows quantifiers over $\mathbb{N}$.

Theorem (AT)

If ϕ is provable in ID_1, there is an arithmetic formula

$$\phi^D(\alpha, \beta)$$

such that for every well-ordered tree α, there is a term β such that OR_1 proves $\phi^D(\alpha, \beta)$.

An Example

When we apply this to the tower of compact extensions, we obtain, for each ordinal α, a factor $\mathcal{C}_\alpha$. A crucial statement is that the system is weak mixing relative to the limit. That is, for any f, g,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i<N} \int \left[E(fT^i g \mid \mathcal{C}_{\omega_1}) - E(f \mid \mathcal{C}_{\omega_1}) T^i E(g \mid \mathcal{C}_{\omega_1})\right]^2 d\mu = 0.$$

Formalized in an ordinal version of ID_1, this becomes the statement that for any $\epsilon > 0$, any f, g, and any ordinal α, there is some ordinal β above α such that

$$\exists N \forall m \geq N \frac{1}{m} \sum_{i<m} \int \left[E(fT^i g \mid \mathcal{C}_\beta) - E(f \mid \mathcal{C}_\beta) T^i E(g \mid \mathcal{C}_\beta)\right]^2 d\mu < \epsilon.$$

An Example

The Dialectica translation applied to this statement tells us that for every $\epsilon > 0$, each f, g, and any ordinal functional, $\hat{\alpha} : \mathbb{N} \times (\Omega \to \Omega) \to \Omega$, there is a function $\hat{\beta}$ on ordinals such that

$$\exists N \forall j \exists \gamma \in [\hat{\alpha}(j, \hat{\beta}), \hat{\beta}(\hat{\alpha}(j, \hat{\beta}))] \forall m \geq N$$

$$\frac{1}{m} \sum_{i<m} \int \left[E(fT^i g \mid \mathcal{C}_\gamma) - E(f \mid \mathcal{C}_\gamma) T^i E(g \mid \mathcal{C}_\gamma) \right]^2 d\mu < \epsilon.$$

The functional $\hat{\beta}$ is "relatively constructive." It is allowed arbitrary arithmetic operations, but is limited in the sort of ordinal operations it can make.

Breaking up the Proof

The proof can be broken up into the following pieces:

- For all α, $\mathcal{C}_\alpha$ has the SZ property
- $\mathcal{C}$ is weak mixing relative to $\bigcup_\alpha \mathcal{C}_{<\alpha}$
- Since $\mathcal{C}$ is weak mixing relative to $\bigcup_\alpha \mathcal{C}_{<\alpha}$, $\mathcal{C}$ has the SZ property

The first part is Π_1 on ordinals, and so has no information about the levels of the hierarchy required.

The third part derives an arithmetic (indeed, Π_2) conclusion, so it amounts to applying these Dialecticized statements to *particular* functions $\hat{\alpha}$.

The Particular Functionals

In this case, it turns out that the particular functions $\hat{\alpha}$ we begin with are quite simple (in fact, they are constant), and the operations $\hat{\beta}$ are also fairly simple. In fact, all functionals we needed are, in a strong sense, already closed below the ordinal ω^ω.

The example above becomes

Theorem

For every $f, \epsilon > 0$, and $q(n)$, there is a K such that for every α, there are $p \leq K$, n, and $\beta \in [\alpha, \alpha + \omega^{q(K)})$ with the following property: for every $\gamma \in [\beta, \beta + \omega^{q(p)})$, there is a $\delta \in [\gamma, \gamma + \omega^p)$ such that for every $m \geq n$,

$$\frac{1}{m} \sum_{i<m} \int \left[E(fT^i g \mid \mathcal{C}_\delta) - E(f \mid \mathcal{C}_\delta) T^i E(g \mid \mathcal{C}_\delta)\right]^2 d\mu < \epsilon.$$

Why are the Functionals Weak?

ID_1 contains two induction schemes, the usual one on the integers, and one along the ordinals. It is known that various weakenings of those schemes give much weaker theories.

In particular, one might think that because the statement that $\mathcal{C}_\alpha$ has the SZ property is arithmetic, and this is the only statement proven by induction, the proof contains only arithmetic induction on ordinals. But as an intermediate step, a “stronger” induction—namely, closure under sums—is used.

A Conjecture

Let ϕ be arithmetic with $ID_1 \vdash \phi$. Suppose there is an arithmetic statement ψ such that $ID_1 \vdash \forall x[x \in I \rightarrow \psi(x)]$, and $T + (CI) \vdash \forall x[x \in I \rightarrow \psi(x)] \rightarrow \phi$. Then

$$T + HA(\alpha) \vdash \phi$$

where α is the proof-theoretic ordinal of T and $T + HA(\alpha)$ adds to T a predicate computing α levels of the hyperarithmetic hierarchy.

The Furstenberg Correspondence

Theorem (Szemerédi)

For every $\delta > 0$ and every k, there is an N such that whenever $E \subseteq [1, N]$ such that $\frac{|E|}{N} \geq \delta$, E contains an arithmetic progression of length k.

Theorem (Furstenberg Multiple Recurrence Theorem)

If $(X, \mathcal{B}, \mu, T)$ is a dynamical system and A is a measurable set with positive measure then for any k there is some n such that,

$$\mu(A \cap T^n A \cap \cdots \cap T^{(k-1)n} A) > 0.$$

Theorem (Furstenberg)

Szemerédi's Theorem and the Multiple Recurrence Theorem are equivalent.

The Furstenberg Correspondence

Observe that $[1, N]$ is a dynamical system: take the points to be the set $[1, N]$, let all sets be measurable, let μ be the density $\mu(B) = \frac{|B|}{N}$, and let T be the shift $n \mapsto n + 1 \mod N$.

The Multiple Recurrence Theorem always holds trivially in such systems, since if $\mu(B) > 0$, $\mu(B \cap T^N B \cap \cdots \cap T^{(k-1)N} B) > 0$, since $T^N B = B$.

Informally, Szemerédi's Theorem says that the Multiple Recurrence Theorem holds in finite dynamical systems with bounds (on the size of n, the "gaps" in the progression) independent of the cardinality of the system.

The Furstenberg Correspondence

Suppose Szemerédi's Theorem fails. Then there is a δ, a k, and an infinite sequence of $E_N \subseteq [1, N]$ such that $\frac{|E_N|}{N} \geq \delta$ but E_N contains no arithmetic progression of length k.

If we pass to a nonstandard model, we obtain a nonstandard integer a and a set $E_a \subseteq [1, a]$ such that $st(\frac{|E_a|}{a}) \geq \delta$ but E_a contains no arithmetic progression of length k.

Nonstandard Analysis

The internal sets form a finitely additive measure space with the measure $\mu^{\circ}(A) = st(\frac{|A|}{a})$, and it is not hard to check that, together with the transformation $x \mapsto x + 1$, this is a finitely additive dynamical system.

The Loeb measure construction tells us that we may extend this to a σ-algebra, giving us a true dynamical system. We may now forget about the fact that we used nonstandard analysis and treat this as any other dynamical system.

Nonstandard Analysis

Applying the Multiple Recurrence Theorem, we find an n such that

$$\mu(E_a \cap T^n E_a \cap \cdots \cap T^{kn} E_a) > 0.$$

There must be some point x in this intersection, and so

$$x, T^n x, \ldots, T^{kn} x \in E_a.$$

x is a nonstandard integer, so transfer tells us that, for infinitely many N, there is an x_N such that

$$x_N, x_N + n, \ldots, x_N + kn \in E_N.$$

Notice that this is very uniform, more so than we originally promised: for instance, it would have sufficed if the n given by the Multiple Recurrence Theorem were nonstandard.

Failure of Ergodicity

After we construct the nonstandard model, there's no reason we have to choose $x \mapsto x+1$ as our transformation. Indeed, we could take any nonstandard integer $b < a$, and take T to be the transformation $x \mapsto x + b \mod a$.

Observe that for any finite $A_N \subseteq [1, N]$ and every point x, the average

$$\frac{1}{M} \sum_{i<M} \chi_{A_N}(x+i) \to \mu(A_N).$$

But when we take the correspondant, this may not be true.

Recovering Information

Consider again our nonstandard construction of a dynamical system as being generated by the internal subsets of $[1, a]$. $[1, a]$ is actually a group, so we could obtain $(X, \mathcal{B}, \mu, \{T_g\}_{g\in G})$, a dynamical system acted on by a large group.

Viewed as a group, $[1, a]$ is rather messy (for instance, it is uncountable). But Tao and Ziegler showed that if we select countably many elements from $[1, a]$ at random, to obtain $(X, \mathcal{B}, \mu, T_1, \ldots, T_n, \ldots)$, the average

$$\lim_{N\to\infty} \lim_{M\to\infty} \frac{1}{M^N} \sum_{\vec{i}\in[1,M]^N} \chi_A(T_1^{i_1} \cdots T_N^{i_N} x) = \mu(A).$$