11.5 The infinitary ergodic approach

In this section we discuss some of the ideas underlying Furstenberg's infinitary ergodic approach to Szemerédi's theorem. These arguments are the shortest and most elegant way to prove the theorem, but also require a certain amount of machinery concerning infinite measure spaces. Also it is quite difficult to extract a quantitative bound from these methods. As the techniques here are rather disjoint from those in the rest of this book we shall not provide full details, referring the reader instead to [122]. However, the insights developed here were essential in developing several of the finitary arguments in this chapter, most notably the finitary ergodic proof of Szemerédi's theorem, and the Green–Tao theorem on arithmetic progressions in the primes.

Define a *measure-preserving system* to be a (possibly infinite) space X with a σ-algebra $\mathcal{B}$, a probability measure $\mathbf{P}_X$ on $\mathcal{B}$, and a bijection $T : X \to X$ such that all the powers T^n of T with $n \in \mathbf{Z}$ are measure-preserving, thus $\mathbf{P}_X(T^n A) = \mathbf{P}_X(A)$ for all $A \in \mathcal{B}$. In this infinite setting, a σ-algebra cannot be rigorously viewed as a partition; instead it is a collection of sets closed under countable unions, intersections, and complements, and containing $\emptyset$ and X. We define an expectation $\mathbf{E}_X$ on bounded measurable functions from X to $\mathbf{R}$ in the usual manner, and define a shift operator T^n on such functions by $T^n f(x) := f(T^{-n}x)$. To simplify the notation slightly we shall only work with real-valued functions in this section rather than complex-valued ones.

Example 11.20 (**Circle shift**) Let X be the unit circle $\mathbf{R}/\mathbf{Z}$ with Lebesgue measure, and let T be the shift $Tx = x + \alpha$ for some fixed $\alpha \in \mathbf{R}$. The dynamics of this system depend on whether α is rational or irrational; for instance, in the former case the shift T is periodic, but not in the latter case. However in both cases we have the following almost periodicity property: given any bounded measurable function on X, the shifts $\{T^n f : n \in \mathbf{Z}\}$ are pre-compact in $L^2(X)$. In particular given any ε we have $\|T^n f - f\|_{L^2(X)} \leq \varepsilon$ for infinitely many ε. Because of this property we say that this measure-preserving system is *compact*.

Example 11.21 (**Skew shift**) Let X be the torus $(\mathbf{R}/\mathbf{Z}) \times (\mathbf{R}/\mathbf{Z})$ with Lebesgue measure, and let T be the skew shift $T(x, y) := (x + \alpha, y + x)$ for some fixed $\alpha \in \mathbf{R}$. Note that the orbits $T^n(x, y)$ are linear in n in the x variable, but quadratic in n in the y variable. This system is not compact, but contains a non-trivial compact factor, namely the σ-algebra $\mathcal{B}_0$ consisting of all the sets of the form $A \times (\mathbf{R}/\mathbf{Z})$, where A is Borel measurable in $\mathbf{R}/\mathbf{Z}$. (To put this another way, the $\mathcal{B}_0$-measurable functions are precisely those functions which do not depend on the y variable.) This factor is isomorphic to the circle shift mentioned earlier. It turns out that the skew shift is a *relatively compact extension* of the circle shift, though we will not quantify precisely what this means here except to observe that if f is a smooth function on $(\mathbf{R}/\mathbf{Z}) \times (\mathbf{R}/\mathbf{Z})$, then the orbits $\{T^n f : n \in \mathbf{Z}\}$ form a precompact set on each fiber $\{x = \text{constant}\}$ of $\mathcal{B}_0$, endowed with the obvious one-dimensional measure.

Example 11.22 (Bernoulli shift) Now consider the infinite unit cube $X := [0, 1]^{\mathbf{Z}}$ of infinite binary sequences $(\omega_n)_{n\in\mathbf{Z}}$, with the usual product topology and Borel σ-algebra $\mathcal{B}$. Let $B \subset X$ denote the "cylinder" of sequences where $\omega_0 = 1$, and let T be the shift operator defined by $T^h(\omega_n)_{n\in\mathbf{Z}} := (\omega_{n+h})_{n\in\mathbf{Z}}$. Using the Kolmogorov extension theorem (or Caratheodory's extension theorem and Tychonoff's theorem) we can find a measure $\mathbf{P}$ on X such that

$$\mathbf{P}(T^{h_1}B \cap \cdots \cap T^{h_m}B) = 2^{-m}$$

whenever $h_1, \dots, h_m$ are distinct integers. Informally, one can view this system as the probability space corresponding to an infinite number of coin tosses, one for each integer h; the event $T^h B$ is then the event that the hth coin turns up heads, and the shift operator corresponds to relabeling all of the coins up by 1. The behavior here is completely different from the compact case; indeed, if f is bounded and measurable, and has mean zero, one can show that $\langle T^n f, f\rangle_{L^2(X)} \to 0$ as $n \to \infty$. A system with this property is known as *strongly mixing*.

Furstenberg derived Szemerédi's theorem by proving the following equivalent formulation.

Theorem 11.23 (Furstenberg multiple recurrence theorem) *[121], [125], [122] Let $(X, \mathcal{B}, \mathbf{P}, T)$ be a measure-preserving system, and let $f : X \to \mathbf{R}^+$ be a non-negative bounded measurable function with $\mathbf{E}(f) > 0$. Then for all $k \geq 1$ we have*

$$\liminf_{N\to\infty} \mathbf{E}_{1\leq n\leq N} \mathbf{E}_X f T^n f \cdots T^{(k-1)n} f > 0.$$

It is fairly easy to deduce this theorem from Szemerédi's theorem; we leave this as an exercise. The converse deduction of Szemerédi's theorem from Furstenberg's theorem is a little trickier, requiring some measure-theoretic tools:

Proof of Theorem 10.1 assuming Theorem 11.23 (Sketch) Suppose for contradiction that we can find a set $A \subseteq \mathbf{Z}$ of positive upper progressions containing no progressions of length k. Thus we can find a sequence of integers $N_1, N_2, \dots$ going to infinity such that $\liminf_{j\to\infty} \mathbf{P}_{[-N_j,N_j]}(A) > 0$. Now use the Hahn–Banach theorem to construct a linear functional λ on bounded real-valued sequences $(c_j)_{j=1}^\infty$ such that

$$\liminf_{j\to\infty} c_j \leq \lambda((c_j))_{j=1}^\infty \leq \limsup_{j\to\infty} c_j.$$

Now consider the infinite unit cube $X := [0, 1]^{\mathbf{Z}}$ of infinite binary sequences $(\omega_n)_{n\in\mathbf{Z}}$, with the usual product topology and Borel σ-algebra $\mathcal{B}$. Let $B \subset X$ denote the "cylinder" of sequences where $\omega_0 = 1$, and let T be the shift operator defined by $T^h(\omega_n)_{n\in\mathbf{Z}} := (\omega_{n+h})_{n\in\mathbf{Z}}$. Using the Kolmogorov extension theorem (or Caratheodory's extension theorem and Tychonoff's theorem) we can find a measure $\mathbf{P}$ on X such that

$$\mathbf{P}(T^{h_1}B \cap \cdots \cap T^{h_m}B) = \lambda\big((\mathbf{P}_{[-N_j,N_j]}((A+h_1)\cap\cdots\cap(A+h_m)))_{j=1}^{\infty}\big)$$

for all $h_1, \ldots, h_m \in \mathbf{Z}$. In particular we see that $\mathbf{P}(B) > 0$. By Theorem 11.23 applied to $f = 1_B$ we conclude that $\mathbf{P}(B \cap T^n B \cap \cdots \cap T^{(k-1)n}B)$ for at least one non-zero B, which implies that A contains a progression of length k. □

One can prove the multiple recurrence theorem in a manner similar to that in the previous sections. For instance, there is an analog of the Gowers uniformity norm $\|f\|_{U^d(X)}$, defined inductively for bounded measurable f by $\|f\|_{U^0(X)} := \mathbf{E}_X(f)$ and

$$\|f\|_{U^d(X)} := \lim_{N\to\infty}\left(\mathbf{E}_{1\le n\le N}\|fT^hf\|_{U^{d-1}(X)}^{2^{d-1}}\right)^{1/2^d}.$$

(The existence of this limit is guaranteed by the von Neumann ergodic theorem; see [185].) One can verify that these U^d norms obey properties similar to their finitary counterparts; see [185], with a key distinction that it is now quite possible for a non-zero function f to have a vanishing U^d norm. We have an important analog of the generalized von Neumann theorem (11.8), namely that

$$\lim_{N\to\infty}\mathbf{E}_{1\le n\le N}\mathbf{E}_X f_0 T^n f_1 \cdots T^{(k-1)n} f_{k-1} = 0$$

whenever $f_0, \ldots, f_{k-1}$ are bounded measurable functions with at least one of the f_j having a vanishing U^{k-1} norm. Thus functions with vanishing U^{k-1} norm have a negligible impact on recurrence.

Again, attention now turns towards the obstructions to uniformity. It turns out that in the infinitary setting these obstructions have a rather nice description. Let $U^{k-1}(X)^*$ denote the space of all bounded functions f for which the expression

$$\|f\|_{U^{k-1}(X)^*} := \sup\{|\mathbf{E}_X(fg)| : \|g\|_{U^{k-1}(X)} \le 1\}$$

is finite. It turns out (see [185]) that there exists a unique σ-algebra $\mathbf{Z}_{k-2}$ such that the closure of $U^{k-1}(X)^*$ in the L^2 topology consists precisely of those square-integrable functions which are measurable with respect to $\mathbf{Z}_{k-2}$; the $\mathbf{Z}_{k-2}$ are thus the *universal characteristic factor* for the $U^{k-1}(X)$ norm. As a consequence one can precisely quantify which functions are Gowers uniform of order $k-1$:

$$\|f\|_{U^{k-1}(X)} = 0 \iff \mathbf{E}(f|\mathbf{Z}_{k-2}) = 0.$$

Here the conditional expectation $f \mapsto \mathbf{E}(f|\mathbf{Z}_{k-2})$ is defined as the L^2-orthogonal projection onto the space of $\mathbf{Z}_{k-2}$-measurable functions.

One consequence of the above discussion is that in order to prove the Furstenberg recurrence theorem, it suffices to do so under the additional assumption that f is $\mathbf{Z}_{k-2}$-measurable (because the error $f - \mathbf{E}(f|\mathbf{Z}_{k-2})$ has a vanishing $U^{k-1}(X)$ norm and is hence irrelevant). To do this, it is clearly of importance to understand the factors $\mathbf{Z}_{k-2}$ of $\mathcal{B}$ as much as possible.

The factor $\mathbf{Z}_0$ turns out to be the space of invariant sets in X, i.e. $\mathbf{Z}_0 := \{A \in \mathcal{B} : TA = A\}$. This is essentially the *von Neumann ergodic theorem*, which we leave to the exercises. The factor $\mathbf{Z}_1$ is known as the *Kronecker factor* and is generated by all the almost periodic functions, or equivalently by the eigenfunctions of the shift operator T. The higher factors are more difficult to describe explicitly. However, it can be shown without too much difficulty (see e.g. [121], [185], [236]); a closely related result is in [386]) that each factor $\mathbf{Z}_{d+1}$ is a *relatively compact extension* of the preceding factor $\mathbf{Z}_d$ (in fact, it is the maximal relatively compact extension). What this means is a little bit tricky to describe precisely, but it roughly means that for a dense set of f which are measurable in $\mathbf{Z}_{d+1}$, the orbits $\{T^n f : n \in \mathbf{Z}\}$ are precompact relative to $\mathbf{Z}_d$, which informally means that they are precompact when restricted to each "atom" or "fiber" of $\mathbf{Z}_d$. See [122] for a rigorous formulation of these assertions (which requires the theory of disintegration of measures). Using some tools from measure theory and analysis, as well as a combinatorial argument closely related to the van der Waerden theorem, it was shown in [121], [125] that if the Furstenberg recurrence theorem holds for any factor $\mathbf{Z}_d$, then it also holds for a relatively compact extension $\mathbf{Z}_{d+1}$; this is analogous to Proposition 11.19. This fact, combined with the preceding discussion, yields the Furstenberg recurrence theorem and thus Szemerédi's theorem.

Recently, there has been significant progress by Host–Kra [185] (and subsequently by Ziegler [386]) in understanding the factors $\mathbf{Z}_{k-2}$. (Strictly speaking, Ziegler treats a slight variant $\mathbf{Y}_{k-2}$ of the factors $\mathbf{Z}_{k-2}$; see [236] for a comparison between the two.) It turns out that the factors $\mathbf{Z}_{k-2}$ are isomorphic to the inverse limit of $k-2$*-step nilsystems*, or in other words a system $(G/\Gamma, \mathcal{B}, T, \mathbf{P})$, where G is a nilpotent Lie group of order $k-2$, Γ is a co-compact subgroup of G, $\mathcal{B}$ is the usual Borel algebra, T is a left shift operator $T : x\Gamma \mapsto gx\Gamma$ for some fixed group element $g \in G$, and $\mathbf{P}$ is normalized Haar measure. Thus for instance the circle shift in Example 11.20 is a 1-step nilsystem, whereas the skew shift turns out to be isomorphic to a 2-step nilsystem. These characterizations of $\mathbf{Z}_{k-2}$ are roughly analogous to the "hard" inverse theorems discussed in Section 11.2; see [160] for further discussion of this in the $k = 4$ case. Just as these hard inverse theorems lead to better quantitative results on Szemerédi's theorem, the characterizations of $\mathbf{Z}_{k-2}$ given here lead to stronger recurrence theorems; for instance, they can be used to replace the limit inferior in the Furstenberg recurrence theorem with a limit, and in fact obtain the stronger result that the averages $\mathbf{E}_{1\leq n\leq N}T^n f \cdots T^{(k-1)n} f$ converge in L^2 norm to a non-zero (and somewhat explicitly describable) function. See [185], [386]. A current area of research is to develop and simplify these ergodic theory results (which are currently quite difficult and lengthy to prove) and clarify their connection with the analogous developments in the Fourier-analytic and combinatorial approaches.

The ergodic approach is well suited for establishing stronger combinatorial results than Szemerédi's theorem, several of which have not yet been proven by other means. We describe some of them here.

Theorem 11.24 (Multi-dimensional Szemerédi theorem) *[123] Let $d \geq 1$, and let $A \subset \mathbf{Z}^d$ be such that $\limsup_{N\to\infty} \mathbf{P}_{[-N,N]^d}(A) > 0$. Then for any $v_1, \ldots, v_k \in \mathbf{Z}^d$, there exist infinitely many pairs $(a, r) \in \mathbf{Z}^d \times \mathbf{Z}^+$ such that $a + rv_1, \ldots, a + rv_k \in A$.*

Theorem 11.25 (Polynomial Szemerédi theorem) *[23] Let $P_1, \ldots, P_k : \mathbf{Z} \to \mathbf{Z}$ be polynomials that map the integers to the integers such that $P_1(0) = \cdots = P_k(0) = 0$. Let $A \subset \mathbf{Z}$ have positive upper density. Then there exist infinitely many pairs $(a, r) \in \mathbf{Z}^d \times \mathbf{Z}^+$ such that $a + P_1(r), \ldots, a + P_k(r) \in A$.*

Theorem 11.26 (Density Hales–Jewett theorem) *[124] Let $n \geq 1$ and $0 < \delta \leq 1$. Then there exists an integer $d = d(|A|, \delta) \geq 1$ such that if A is any subset of $[0, n-1]^d$ with cardinality $|A| \geq \delta n^d$, then A contains a proper arithmetic progression $a + [0, n-1] \cdot v$ of length n, for some $a \in [0, n-1]^d$ and $v \in [0, 1]^d$.*

Further refinements include additional structural information on the pairs (a, r) constructed by the above theorems, as well as convergence of various limits; in addition, there is much current work in extending the description of the characteristic factor for the U^k norm and for multiple recurrence to these more complex recurrence theorems. Unfortunately a complete survey of these exciting developments is well beyond the scope of this book.

Exercises

11.5.1 Show that Theorem 11.1 for a fixed k implies Theorem 11.23 for the same value of k.

11.5.2 (Poincaré recurrence theorem) Using only the pigeonhole principle and elementary measure theory, prove Theorem 11.23 in the $k = 2$ case.

11.5.3 (Von Neumann ergodic theorem) Let $(X, \mathcal{B}, \mathbf{P}, T)$ be a measure-preserving system. Show that the spaces $\{f \in L^2(X) : Tf = f\}$ and $\{Tf - f : f \in L^2(X)\}$ are complementary orthogonal subspaces of $L^2(X)$. Use this to conclude that if $\mathbf{Z}_0 := \{A \in \mathcal{B} : TA = A\}$, then $\mathbf{E}_{1 \leq n \leq N} T^n f$ converges in $L^2(X)$ to $\mathbf{E}(f|\mathbf{Z}_0)$ for any $f \in L^2(X)$, and that $\|f\|_{U^1(Z)} = \|\mathbf{E}(f|\mathbf{Z}_0)\|_{L^2(X)}$. Note that these results simplify in the case when the system is *ergodic* (which means that $\mathbf{Z}_0 = \{\emptyset, X\}$), since in that case $\mathbf{E}(f|\mathbf{Z}_0)$ is just $\mathbf{E}_X(f)$. In particular we have $\|f\|_{U^1(Z)} = |\mathbf{E}_X(f)|$ in this case, just as in the finitary case.

11.5.4 (Khintchine's recurrence theorem) Let A be a subset of a measure-preserving system $(X, \mathcal{B}, \mathbf{P}, T)$. Show that for every $\varepsilon > 0$ that there exist infinitely many $n \in \mathbf{Z}$ such that $\mathbf{P}(A \cap T^n A) \geq \mathbf{P}(A)^2 - \varepsilon$. (Hint: obtain lower and upper bounds for $\|\mathbf{E}_{1 \leq n \leq N} 1_{T^n A}\|_{L^2(Z)}$. Alternatively, use the von Neumann ergodic theorem.) Show that the theorem fails if $\mathbf{P}(A)^2 - \varepsilon$ is replaced by $\mathbf{P}(A)^2 + \varepsilon$, regardless of how small $\mathbf{P}(A)$ and ε are. It is natural to then conjecture that $\mathbf{P}(A \cap T^n A \cap \cdots \cap T^{(k-1)nA}) \geq \mathbf{P}(A)^k - \varepsilon$ for infinitely many n; this is true for $k = 1, 2, 3, 4$ under the additional assumption of ergodicity, but fails for $k > 4$, see [22].

11.5.5 Let $(X, \mathcal{B}, \mathbf{P}, T)$ be a compact measure-preserving system (so the orbits $\{T^n f : n \in \mathbf{Z}\}$ are precompact in L^2 whenever f is bounded and measurable). Prove the Furstenberg recurrence theorem in this special case. (Compare with Proposition 10.35 or the $k = 3$ proof of Proposition 11.19.)

11.5.6 Let $(X, \mathcal{B}, \mathbf{P}, T)$ be a weakly mixing measure-preserving system, which means that $\lim_{N \to \infty} \mathbf{E}_{1 \leq n \leq N} |\langle T^n f, f \rangle_{L^2(X)}|^2 = 0$ whenever f is bounded, measurable, and has expectation zero. (This is weaker than *strong mixing*, which demands that $\lim_{n \to \infty} \langle T^n f, f \rangle = 0$ under the same hypotheses.) Show that $\|f\|_{U^{k-1}(X)} = 0$ if and only if $\mathbf{E}_X(f) = 0$, and establish the Furstenberg recurrence theorem in this special case.

11.5.7 Let $(X, \mathcal{B}, \mathbf{P}, T)$ be measure-preserving system, and let f be bounded and measurable. Show that if f is almost periodic (thus the orbit $\{T^n f : n \in \mathbf{Z}\}$ is precompact in $L^2(X)$), then $\mathbf{E}_X(fg) = 0$ whenever g is bounded, measurable, and vanishing in $U^2(X)$ norm. Compare this with Exercise 11.4.8.

11.5.8 Let $(X, \mathcal{B}, \mathbf{P}, T)$ be measure-preserving system. Let $\mathbf{Z}_1$ be the smallest σ-algebra with respect to which all almost periodic functions are measurable. If f is bounded and measurable, show that $\|f\|_{U^2(X)} = 0$ if and only if $\mathbf{E}(f|\mathbf{Z}_1) = 0$. (Hint: the "only if" part follows from the preceding exercise. For the "if" part, construct the dual function $\mathcal{D}_2 f := \lim_{N \to \infty} \mathbf{E}_{-N \leq n \leq N} T^n f \mathbf{E}_X(f T^n f|\mathbf{Z}_0)$, and show that this function is