3. Furstenberg's Correspondence Principle

The connection between Ramsey theory and ergodic theory hinges on the Furstenberg Correspondence Principle which we will soon formulate. Recall from Exercise 1.8 that the upper density satisfies $\bar{d}(A) = \bar{d}(A-1)$ for any set $A \subset \mathbb{N}$. Denote by $T : \mathbb{N} \to \mathbb{N}$ the successor map $T : x \mapsto x+1$. Then $A-1$ can be written as $T^{-1}A$. While $\bar{d}$ is not a probability measure (it is not even finitely additive), the tuple $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \bar{d}, T)$ looks a lot like a measure preserving system (by $\mathcal{P}(\mathbb{N})$ we denote the collection of all subsets of $\mathbb{N}$).

Furstenberg Correspondence Principle.

For many "arithmetic purposes", the tuple $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \bar{d}, T)$ behaves like a measure preserving system.

This is, of course a very vague statement, which is why it is a "principle" and not a "theorem". There are several incarnations of this principle as precise statements, but it is good to keep in mind the overarching principle, which can be adapted for different purposes.

Exercise 3.1. *Show that there are sets $A, B \subset \mathbb{N}$ with $\bar{d}(A) = \bar{d}(B) = 1$ but $A \cap B = \emptyset$.*

It is natural to wonder if the problem lies with the definition of upper density itself, and in particular with the lim sup. For instance, if one restricts attention to sets with ***natural density***, defined as the limit $d(A) := \lim_{N\to\infty} \frac{1}{N}|A \cap \{1, \ldots, N\}|$ only for those sets A for which the limit exists, it is clear that the bad examples such as those from Exercise 3.1 can no longer exist. Unfortunately, using natural density leads to problems of a different kind:

Exercise 3.2. *Show that there are sets $A, B \subset \mathbb{N}$ both having natural density but such that $A \cap B$ does not.*

The first instance of the Correspondence Principle was used by Furstenberg to give an ergodic theoretic proof of Szemerédi's theorem (Theorem 1.9), which states that if $A \subset \mathbb{N}$ has $\bar{d}(A) > 0$, then A contains an arithmetic progression of length k for any prescribed $k \in \mathbb{N}$. Note that

$$\exists x, n \in \mathbb{N} : \{x, x+n, \ldots, x+kn\} \subset A \quad \iff \quad \exists n \in \mathbb{N} : A \cap (A-n) \cap \cdots \cap (A-kn) \neq \emptyset.$$

Using the shift T, we can write this as $\exists n \in \mathbb{N} : A \cap T^{-n}A \cap \cdots \cap T^{-kn}A \neq \emptyset$. If we subscribe to the Correspondence Principle, then Szemerédi's theorem becomes the statement that in a measure preserving system, whenever A has positive measure and $k \in \mathbb{N}$, there exists $n \in \mathbb{N}$ such that $A \cap T^{-n}A \cap \cdots \cap T^{-kn}A \neq \emptyset$. Since sets of measure 0 in a measure space might as well be empty, we have shown that Szemerédi's theorem is morally equivalent to the following.

Theorem 3.3 (Furstenberg's multiple recurrence theorem, [6])**.** *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system, let $A \in \mathcal{B}$ have $\mu(A) > 0$ and let $k \in \mathbb{N}$. Then there exists $n \in \mathbb{N}$ such that*

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

It turns out that Theorem 3.3 is indeed equivalent to Szemerédi's theorem, which will be proved using a concrete instance of the Furstenberg Correspondence Principle. Note that for $k = 1$, Theorem 3.3 reduces to Poincaré's Recurrence theorem (Theorem 2.7), which has a fairly simple proof. On the other hand, the case $k = 1$ of Szemerédi's theorem states the even more trivial fact that any set with positive upper density contains a 2-term arithmetic progression.

The proof of Theorem 3.3, which will occupy a few lectures, not only yields a proof of Szemerédi's theorem, but it reveals some deep structural results about *arbitrary* measure preserving systems.

Here is the version of the correspondence principle, formulated in [6], that we will use.

Theorem 3.4 (Correspondence Principle)**.** *Let $E \subset \mathbb{N}$. Then there exist a measure preserving system $(X, \mathcal{B}, \mu, T)$ and a set $A \in \mathcal{B}$ with $\mu(A) = \bar{d}(E)$ such that for any $n_1, \ldots, n_k \in \mathbb{N}$,*

$$\mu\big(A \cap T^{-n_1}A \cap \cdots \cap T^{-n_k}A\big) \leq \bar{d}\big(E \cap (E-n_1) \cap \cdots \cap (E-n_k)\big). \tag{3.1}$$

Proof. It turns out that $X, \mathcal{B}, T$ and A will not depend on E and only μ does. Take $X = \{0,1\}^{\mathbb{N}_0}$, with the product topology (where $\{0,1\}$ has the discrete topology) and the Borel σ-algebra $\mathcal{B}$. Let $T : X \to X$ be the left shift map $T : (x_n)_{n=0}^{\infty} \mapsto (x_{n+1})_{n=0}^{\infty}$ and let A be the cylinder set at 0, described as $A = \{(x_n) \in X : x_0 = 1\}$.

Then let $x \in X$ be the indicator function of E, so that $x_n = 1 \iff n \in E$. For each $N \in \mathbb{N}$, let $\mu_N = \frac{1}{N}\sum_{n=1}^{N} \delta_{T^n x}$ be the empirical measure (here, as usual, we denote by δ_y the Dirac measure (a.k.a the point mass) at y). Find a sequence $(N_k)_{k\in\mathbb{N}}$ such that $\bar{d}(E)=\lim_{k\to\infty}\frac{1}{N_k}|E\cap\{1,\ldots,N_k\}|$. Since X is compact, so is the space of probability measures on X under the weak* topology[3]. Therefore, we may pass to a subsequence of (N_k) (which to simplify notation will still be denoted by (N_k)) so that the limit

$$\mu=\lim_{k\to\infty}\mu_{N_k}$$

exists. It is not hard to show that μ is T-invariant (see Exercise 3.5) so that $(X,\mathcal{B},\mu,T)$ is indeed a measure preserving system. Note that

$$\delta_{T^n x}(A)=1 \iff T^n x\in A \iff x_n=1 \iff n\in E,$$

so $\mu_N(A)=\frac{1}{N}|E\cap\{1,\ldots,N\}|$ and hence $\mu(A)=\bar{d}(E)$. Finally, for any $n_1,\ldots,n_k\in\mathbb{N}$ we have

$$\delta_{T^n x}\big(A\cap T^{-n_1}A\cap\cdots\cap T^{-n_k}A\big)=1 \iff n\in E\cap(E-n_1)\cap\cdots\cap(E-n_k),$$

and hence

$$\begin{aligned}\mu\big(A\cap T^{-n_1}A\cap\cdots\cap T^{-n_k}A\big) &= \lim_{k\to\infty}\frac{1}{N_k}\big|E\cap(E-n_1)\cap\cdots\cap(E-n_k)\cap\{1,\ldots,N_k\}\big|\\ &\le \limsup_{N\to\infty}\frac{1}{N}\big|E\cap(E-n_1)\cap\cdots\cap(E-n_k)\cap\{1,\ldots,N\}\big|\\ &= \bar{d}\big(E\cap(E-n_1)\cap\cdots\cap(E-n_k)\big)\end{aligned}$$

□

Exercise 3.5. *Show that the measure μ constructed in the proof of the Furstenberg Correspondence Principle is T-invariant. [Hint: Using Exercise 2.2, it suffices to show that $\int_X f\,\mathsf{d}\mu=\int_X f\circ T\,\mathsf{d}\mu$ for every $f\in C(X)$.]*

3.1. **Applications of the Correspondence Principle.** The first application, as mentioned in the previous subsection, is to reduce Szemerédi's theorem to the multiple recurrence theorem, which, while a deep result, is purely about ergodic theory. Indeed, given $E\subset\mathbb{N}$ with $\bar{d}(E)>0$, applying the correspondence principle in the form of Theorem 3.4 yields a measure preserving system $(X,\mathcal{B},\mu,T)$ and a set $A\in\mathcal{B}$ with $\mu(A)>0$ and satisfying (3.1). Then, using Theorem 3.3, one can find for any $k\in\mathbb{N}$ a number $n\in\mathbb{N}$ such that $\mu(A\cap T^{-n}A\cap\cdots\cap T^{-kn}A)>0$, which in view of (3.1) implies that $\bar{d}\big(E\cap(E-n)\cap\cdots\cap(E-kn)\big)>0$. Now any $x\in E\cap(E-n)\cap\cdots\cap(E-kn)$ gives rise to an arithmetic progression $\{x,x+n,\cdots,x+kn\}$ contained in E.

As also mentioned above, it turns out that the converse direction is true as well, i.e., taking Szemerédi's theorem as a blackbox, one can easily prove Theorem 3.3 (see Exercise 3.6).

The next application of the correspondence principle is Proposition 2.11, which states that sets of recurrence are the same as intersective sets:

Proof of Proposition 2.11. Suppose first that R is a set of recurrence and let $E\subset\mathbb{N}$ with $\bar{d}(E)>0$. We need to find $n\in R\cap(E-E)$. Applying Theorem 3.4 we get a m.p.s. $(X,\mathcal{B},\mu,T)$ and a set $A\in\mathcal{B}$ with $\mu(A)>0$ satisfying (3.1). Since R is a set of recurrence, there exists $n\in R$ with $\mu(A\cap T^{-n}A)>0$, which in view of (3.1) implies that $\bar{d}(E\cap(E-n))>0$. In particular $E\cap(E-n)$ is non-empty, and if x belongs to it, then $x,x+n\in E$, whence $n=(x+n)-x\in(E-E)\cap R$. We conclude that R is an intersective set.

Next suppose that R is intersective. Let $(X,\mathcal{B},\mu,T)$ be a m.p.s. and let $A\in\mathcal{B}$ with $\mu(A)>0$. For each $x\in X$ let $E_x=\{n\in\mathbb{N}:T^n x\in A\}$. The upper density of E_x is

$$\bar{d}(E_x)=\limsup_{N\to\infty}\frac{1}{N}\big|\{n\in\{1,\ldots,N\}:T^n x\in A\}\big|=\limsup_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}1_A(T^n x).$$

Using Fatou's lemma, we can now estimate the average upper density of E_x:

$$\int_X\bar{d}(E_x)\,\mathsf{d}\mu=\int_X\limsup_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}1_{T^{-n}A}\,\mathsf{d}\mu\ge\limsup_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\int_X 1_{T^{-n}A}\,\mathsf{d}\mu=\mu(A)$$

[3]This follows from combining the Riesz representation theorem for measures with the Banach-Alaoglu theorem and the trivial fact that the constant function 1 has compact support.

Therefore the set $B := \{x \in X : \bar{d}(E_x) > \mu(A)/2\}$ has positive measure, and for each $x \in B$ we can use the fact that R is intersective to find $a_x, b_x \in E_x$ with $a_x - b_x \in R$. Since there only countably many choices for the pairs (a_x, b_x), there exists a pair $(a, b) \in \mathbb{N}^2$ and a positive measure subset $C \subset B$ such that for every $x \in C$ we have $\{a, b\} \subset E_x$ and $n := a - b \in R$. Therefore $C \subset T^{-a}A \cap T^{-b}A = T^{-b}(T^{-n}A \cap A)$ which implies that $\mu(T^{-n}A \cap A) > 0$, and hence that R is a set of recurrence. □

The method used for the second half of the proof can be adapted to show that Szemerédi's theorem implies Furstenberg's Multiple Recurrence theorem.

Exercise 3.6. *Adapting the proof of Proposition 2.11, show that Theorem 1.9 implies Theorem 3.3.*

Exercise 3.7. *Show that, in the proof of Proposition 2.11, the function $x \mapsto \bar{d}(E_x)$ is measurable and hence we can in fact consider its integral.*

Exercise 3.8. *Show that, in the proof of Proposition 2.11, for μ-a.e. $x \in X$ the set E_x has a natural density, i.e., show that the limit $\lim_{N\to\infty} \frac{1}{N}|E_x \cap \{1, \dots, N\}|$ exists.*

Recall Khintchine's theorem (Theorem 2.24). Applying the correspondence principle we obtain the following combinatorial corollary.

Corollary 3.9. *Let $E \subset \mathbb{N}$ have $\bar{d}(E) > 0$. Then the set $E - E$ is syndetic.*

In fact, given sets $E_1, E_2, \dots, E_k \subset \mathbb{N}$ with $\bar{d}(E_i) > 0$ for all i, the intersection $(E_1 - E_1) \cap \cdots \cap (E_k - E_k)$ is syndetic.

Proof. We prove only the second statement, which naturally implies the first one. Let $E_1, \dots, E_k \subset \mathbb{N}$ have all positive upper density. Apply Theorem 3.4 to each of them to get measure preserving systems $(X_i, \mathcal{B}_i, \mu_i, T_i)$ and sets $A_i \in \mathcal{B}_i$ for each $i = 1, \dots, k$ satisfying $\mu_i(A_i) = \bar{d}(E_i) > 0$. Then let $X = \prod_{i=1}^k X_i$, $\mathcal{B} = \bigotimes_{i=1}^k \mathcal{B}_i$, $\mu = \bigotimes_{i=1}^k \mu_i$ and $T : X \to X$ be the map $T(x_1, \dots, x_k) = (T_1x_1, \dots, T_kx_k)$. Let $A = \prod_{i=1}^k A_i \subset X$ and note that $\mu(A) = \mu_1(A_1) \times \cdots \times \mu_k(A_k)$.

In view of Theorem 2.24, the set $R := \{n \in \mathbb{N} : \mu(A \cap T^{-n}A) > 0\}$ is syndetic. Noting that $A \cap T^{-n}A = \prod_{i=1}^k (A_i \cap T_i^{-n}A_i)$ it follows that whenever $n \in R$, for each $i = 1, \dots, k$ we have $\mu_i(A_i \cap T_i^{-n}A_i) > 0$. Using (3.1) it follows that $\bar{d}(E_i \cap (E_i - n)) > 0$ and in particular that $n \in E_i - E_i$ for each i. We conclude that $R \subset (E_1 - E_1) \cap \cdots \cap (E_k - E_k)$ and hence that this intersection is syndetic. □

As an application of this circle of ideas, here is a proof of Schur's theorem (Theorem 1.1) essentially first discovered by Bergelson.

Proof of Theorem 1.1. Let $\mathbb{N} = C_1 \cup \cdots \cup C_r$ be a finite partition (i.e. coloring) of $\mathbb{N}$. After reordering the C_i's if needed we can find $s \in \{1, \dots, r\}$ such that $\bar{d}(C_i) > 0$ for every $i = 1, \dots, s$ and $\bar{d}(C_i) = 0$ for each $i > s$. It follows that the (possibly empty) intersection $E := \bigcup_{i>s} C_i$ has 0 density and in particular is not a syndetic set.

Using Corollary 3.9 it follows that the intersection $(C_1 - C_1) \cap \cdots \cap (C_s - C_s)$ is syndetic, and hence is not contained in E. Therefore there exists $x \in (C_1 - C_1) \cap \cdots \cap (C_s - C_s) \cap (\mathbb{N} \setminus E)$. Say $x \in C_j$; we then have that $j \leq s$, so $x \in C_j - C_j$ as well. Let $z, y \in C_j$ be such that $z - y = x$. It follows that $\{x, y, x+y\} \subset C_j$. □