

I Can't Believe It's Not Random!

Joel Moreira's math blog

Furstenberg's Correspondence Theorem

Posted on 19/11/2011

In 1977 Furstenberg gave a new proof of [Szemerédi's theorem](#) using ergodic theory. The first step in that proof was to turn the combinatorial statement into a statement in ergodic theory. Thus Furstenberg created what is now known as Furstenberg's correspondence principle. While this was not (by far) the most difficult part of the proof of Szemerédi's theorem, it was this principle that allowed many generalizations of Szemerédi's theorem to be proved via ergodic theoretical arguments. Most of those generalizations had to wait a long time before seeing a combinatorial proof, and for some, no combinatorial proof was ever found (yet).

In this post I will state and prove the correspondence principle and then I will use it to give another proof of [Sárközy's theorem](#), discussed in my previous post.

— 1. Motivation and a more general setup —

The original correspondence principle was stated for subsets of \mathbb{Z} , but I will consider more general countable commutative semigroups.

This is useful when we want to use this technology to deal with combinatorial properties of (say) the multiplicative semigroup of integers, as well as additive structure on the higher dimension groups \mathbb{Z}^n .

Definition 1 (Følner sequence) Let G be a countable semigroup. A (left) Følner sequence in G is a sequence $\{F_N\}$ where each F_N is a finite subset of G and such that for any $g \in G$ the set $gF_N := \{gx : x \in F_N\}$ satisfies

$$\lim_{N \rightarrow \infty} \frac{|gF_N \cap F_N|}{|F_N|} = 1$$

The standard example is the sequence of intervals $\{[1, N]\}$ in the group \mathbb{Z} (note that the group operation is the addition). We note that not every countable semigroup has a Følner sequence. For instance the free group on two generators has no Følner sequence.

All solvable groups (in particular abelian groups) are amenable. The existence of a Følner sequence in a semigroup allows us to define the notion of density:

Definition 2 (Upper Density) Let G be a countable semigroup admitting a Følner sequence $\{F_N\}$. Then for any subset $A \subset G$ we define its upper density (with respect to $\{F_N\}$) by

$$\bar{d}(A) = \limsup_{N \rightarrow \infty} \frac{|A \cap F_N|}{|F_N|}$$

Note that the limit doesn't exist in general. Also, a set $A \subset G$ can have different densities with respect to different Følner sequences. As an example consider the set $A = \bigcup [n^3, n^3 + n] \subset \mathbb{Z}$. With respect to the Følner sequence $F_N = [N^3, N^3 + N]$ the density of A is 1, but with respect to the standard Følner sequence $G_N = [1, N]$ the density is 0.

The idea of the Correspondence principle is to transfer combinatorial statements into statements about measure preserving systems. For instance, if $A \subset \mathbb{Z}$, to say that A contains an arithmetic progression of length k is the same as to say that for some positive integer r the intersection $A \cap (A - r) \cap (A - 2r) \cap \dots \cap (A - (k - 1)r)$ is non-empty. To

see this note that a is in that intersection if and only if the arithmetic progression $\{a, a + r, a + 2r, \dots, a + (k - 1)r\}$ is contained in A .

Now if we introduce the shift map $T : \mathbb{Z} \rightarrow \mathbb{Z}$ defined by $T(a) = a - 1$, then A contains an arithmetic progression of length k if and only if $A \cap T^r A \cap \dots \cap T^{(k-1)r} A \neq \emptyset$. In many situations we prove that the intersection is non-empty by proving it's actually big (has positive upper density). Now this formulation suggests an ergodic theoretical formulation of the Szemerédi's theorem. First we need to introduce some concepts.

Let (X, \mathcal{B}, μ) be a [probability space](#) and $T : X \rightarrow X$ a probability preserving map, i.e. T is a measurable function such that for each $B \in \mathcal{B}$ we have $\mu(T^{-1}B) = \mu(B)$ (here as usual we define $T^{-1}B := \{x \in X : Tx \in B\}$). The quadruple (X, \mathcal{B}, μ, T) is called a measure preserving system. An ergodic theoretical formulation of Szemerédi's theorem is

Theorem 3 (Furstenberg's Multiple Recurrence theorem) *Let (X, \mathcal{B}, μ, T) be a measure preserving system and let $A \in \mathcal{B}$ have positive measure. Then for all positive integer k there is some positive integer r such that*

$$\mu(A \cap T^{-r}A \cap T^{-2r}A \dots \cap T^{-kr}A) > 0$$

More generally one can consider, for any semigroup G , an action of G on the probability space (X, \mathcal{B}, μ) by probability preserving transformations.

— 2. Furstenberg's Correspondence principle —

We now state the general correspondence principle that, in the particular case $G = \mathbb{Z}$, allows one to derive Szemerédi's theorem from Theorem 3.

Theorem 4 (Furstenberg's Correspondence Principle) *Let G be a countable commutative semigroup (with identity) and admitting a Følner sequence $\{F_N\}$. Let $E \subset G$ be a subset with positive upper density. Then there exists a probability space (X, \mathcal{B}, μ) , a measure preserving G -action T_g on X and a set $A \in \mathcal{B}$ such that $\mu(A) = \bar{d}(E)$ and for any $g_1, \dots, g_k \in G$*

$$\mu(A \cap T_{g_1}^{-1}A \cap \dots \cap T_{g_k}^{-1}A) \leq \bar{d}(E \cap (g_1^{-1}E) \cap \dots \cap (g_k^{-1}E))$$

Proof: Consider the compact space $X_0 = \{0, 1\}^G$ and let $z \in X_0$ be the characteristic function of E (i.e. $z_n = 1$ if $n \in E$ and $z_n = 0$ otherwise). Consider the G action of X_0 given by $(gx)_n = x_{gn}$ for any $x \in X_0$ and $g, n \in G$. Let X be the orbit closure of z , i.e. $X = \overline{\{gz : g \in G\}}$ where the topology on X_0 is the product topology.

If $x \in X$ then for any $g \in G$ we also have $gx \in X$. Let $A \subset X$ be the set of all sequences $x \in X$ such that $x_0 = 1$. We have $gz \in A \iff g \in E$. Let N_j be a sequence such that

$$\lim_{j \rightarrow \infty} \frac{|E \cap F_{N_j}|}{|F_{N_j}|} = \bar{d}(E)$$

Now consider in X the induced topology and the Borel σ -algebra \mathcal{B} . For each positive integer j we define the measure μ_j as the average of the point mass measure at the points gz for $g \in F_{N_j}$. In other words, for a subset $B \in \mathcal{B}$ we set $\mu_j(B) = s/|F_{N_j}|$ where s is the number of elements g of F_{N_j} such that $gz \in B$. We have thus that $\lim_{j \rightarrow \infty} \mu_j(A) = \bar{d}(E)$. On the other hand, all measures μ_j are probability Borel measures, and thus live in the weak-* unit ball of the dual of the space of continuous functions on X $C(X)$. Thus by compactness of that unit ball we have that (possibly after passing to a subsequence) the sequence μ_j converges (in the weak-* topology). Let μ be the limit measure. Then $\mu(A) = \bar{d}(E)$ and for any $g_1, \dots, g_k \in G$ we have

$$\begin{aligned} \bar{d}(E \cap (g_1^{-1}E) \cap \dots \cap (g_k^{-1}E)) &\geq \lim_{j \rightarrow \infty} \frac{|F_{N_j} \cap E \cap (g_1^{-1}E) \cap \dots \cap (g_k^{-1}E)|}{|F_{N_j}|} \\ &= \lim_{j \rightarrow \infty} \mu_j(A \cap T_{g_1}^{-1}A \cap \dots \cap T_{g_k}^{-1}A) \\ &= \mu(A \cap T_{g_1}^{-1}A \cap \dots \cap T_{g_k}^{-1}A) \end{aligned}$$

□

I now show how to use this result to convert the multiple recurrence theorem (theorem 3) into Szemerédi's theorem. Recall that Szemerédi's theorem states that a subset $E \subset \mathbb{Z}$ with positive [upper Banach density](#) contains arbitrarily long arithmetic progressions. If E is such a set then it has positive upper density with respect to some Følner sequence consisting of intervals (cf. definition 2). We can now apply the correspondence theorem to get a measure preserving system (X, \mathcal{B}, μ, T) and a subset $A \in \mathcal{B}$ with $\mu(A) = \bar{d}(E) > 0$.

By theorem 3 for all k there is some r such that $\mu(A \cap T^{-r}A \cap T^{-2r}A \dots \cap T^{-kr}A) > 0$, and from the correspondence theorem we know that

$$\bar{d}(E \cap (E - r) \cap \dots \cap (E - kr)) \geq \mu(A \cap T^{-r}A \cap \dots \cap T^{-kr}A) > 0$$

In particular there is some $a \in E \cap (E - r) \cap \dots \cap (E - kr)$ and so $\{a, a + r, \dots, a + kr\} \subset E$ is the desired arithmetic progression (since k was arbitrary E contains arbitrarily long arithmetic progressions). This finishes the proof of Szemerédi's theorem (assuming theorem 3).

As another simple example of an application of the correspondence principle, I now give the proof that a [recurrent set](#) is also an [intersective set](#), as mentioned in my [previous post](#). Let $R \subset \mathbb{Z}$ be a recurrent set and $E \subset \mathbb{Z}$ be a set of positive upper density (with respect to some Følner sequence). We want to prove that the intersection $R \cap (E - E)$ is non-empty. Applying the correspondence principle we find a measure preserving system (X, \mathcal{B}, μ, T) and $A \in \mathcal{B}$ with $\mu(A) = \bar{d}(E) > 0$. Since R is recurrent, there is some $r \in R$ such that $\mu(A \cap T^{-r}A) > 0$. But then we have $\bar{d}(E \cap (E - r)) > 0$ and if $x \in E \cap (E - r)$ then $x + r \in E$, so both $x, x + r \in E$ and thus $r \in E - E \cap R$.

— 3. Sárközi's theorem, revisited —

In my [last post](#), I proved that any divisible polynomial (I define this below) is a recurrent set. I also observed that that result is a consequence of Sárközi's theorem (which states that divisible polynomials are intersective), because intersective sets are recurrent. In the last section I showed that recurrent sets are intersective, so Sárközi's theorem can be stated as "divisible polynomials are recurrent". More precisely

Theorem 5 (Sárközi/Furstenberg/Kamae-Mendès France) *Let $f \in \mathbb{R}[x]$ be non-zero polynomial such that for any integer n the image $f(n)$ is also an integer. Assume also that f is a divisible polynomial, i.e.*

for all positive integer k there is some $n \in \mathbb{Z}$ such that $k|p(n)$.

Then the set $R = f(\mathbb{Z}) \setminus \{0\}$ is recurrent, i.e. for any measure preserving system (X, \mathcal{B}, μ, T) and any set $A \in \mathcal{B}$ with $\mu(A) > 0$ there is $r \in R$ such that $\mu(A \cap T^{-r}A) > 0$.

Sárközi proved that for polynomials $f \in \mathbb{Z}[x]$ with a zero (i.e. an integer n such that $f(n) = 0$) the set $f(\mathbb{Z})$ is intersective (he first proved the case $f(x) = x^2$) and independently Furstenberg, using his correspondence principle proved the same result for squares (so he actually proved that $\{n^2 : n \geq 1\}$ is a recurrent set, and then from the correspondence principle concluded that it must also be an intersective set). Later Kamae and Mendès France proved the general case when f is just divisible (note that if f has a zero then it is automatically divisible).

I will now provide a proof based on Furstenberg's original argument but using a version of the van der Corput

trick instead of the spectral theorem which would be a (slightly) softer way.

Proposition 6 (van der Corput trick in Hilbert spaces) Let u_n be a bounded sequence in a Hilbert space H . If for each positive integer h we have

$$\limsup_{N \rightarrow \infty} \left| \frac{1}{N} \sum_{n=1}^N \langle u_{n+h}, u_n \rangle \right| = 0$$

then $\frac{1}{N} \sum_{n=1}^N u_n \rightarrow 0$ as $N \rightarrow \infty$.

The proof of theorem 5 illustrates the basic idea behind recurrence theorems (which, thanks to the correspondence theorem can be used to deduce combinatorial results). The first step is to consider the characteristic function 1_A of A instead of the set itself. Then we want to decompose the space (in this case $L^2(X)$) into an orderly component and a pseudo-random component. This is of course a very vague statement, and in each case, to determine what is the orderly component (and what is the pseudo-random) is the important part. Elements in the pseudo-random component should have pseudo-random orbits, in the sense that they somehow average down to 0, and elements in the orderly component can be controlled.

In this case those component are orthogonal subspaces of L^2 . The van der Corput trick will be used to prove that the pseudo-random component does average down to 0, then the orderly part is easy to control.

Proof: of theorem 5

Let $H = L^2(X)$ and let $U : H \rightarrow H$ be given by $(Uf)(x) = f(Tx)$. Then U is a unitary operator (this is called the Koopman operator). Let 1_A be the characteristic function of A (i.e. $1_A(x) = 1$ if $x \in A$ and $1_A(x) = 0$ if $x \notin A$). Then $\mu(T^{-r}A \cap A) = \int_X 1_A(T^r x) 1_A(x) d\mu(x) = \langle U^r 1_A, 1_A \rangle$ and we want to prove that this is positive for some $r \in R$. We will do this by showing that the sum $\sum \langle U^r 1_A, 1_A \rangle$ over some subset of R is non-zero.

By hypothesis, the set $D_k = \{n \in \mathbb{Z} : k|p(n)\}$ is non-empty for all $k \geq 1$, and if $n \in D_k$ then also $n+k \in D_k$ (think $\text{mod } k$) and so $k|p(n+ik)$ for any integer i . It turns out that the average $\sum_{i=1}^N U^{p(n+ik)} 1_A$ gets close to the projection of 1_A in a subspace (which will be the orderly subspace) if N is large enough. The subspace is the rational subspace $H_1 = \overline{\bigoplus_{m \neq 0} \{f \in H : U^m f = f\}}$.

We can decompose $1_A = f + g$ where $f \in H_1$ and g is orthogonal to H_1 . Now we will prove that $\frac{1}{N} \sum_{i=1}^N U^{p(n+ik)} g$ converges to 0, using proposition 6. First note that the function $q : i \mapsto p(n+ik)$ is also a polynomial, we will actually prove that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N U^{q(i)} g = 0 \quad (1)$$

holds for any non-constant polynomial q , using induction on the degree of q . If q has degree 1, say $q(x) = ax + b$, then the [mean ergodic theorem](#) applied to U^a says that the limit in (1) is the projection of $U^b g$ onto the invariant subspace of U^a . Since H_1 is invariant under U , $U^b g$ is orthogonal to H_1 which contains the invariant subspace of U^a . This proves (1) when the degree of q is 1.

Now assume by induction that (1) holds for $q_h(i) := q(i+h) - q(i)$ (which is a polynomial of degree less than q). Let $u_i = U^{q(i)} g$. Then by the induction hypothesis we have for each positive integer h that

$$0 = \limsup_{N \rightarrow \infty} \left| \left\langle \frac{1}{N} \sum_{i=1}^N U^{q_n(i)} g \right\rangle \right| = \limsup_{N \rightarrow \infty} \left| \frac{1}{N} \sum_{i=1}^N \langle u_{i+h}, u_i \rangle \right|$$

Thus by proposition 6 we conclude that (1) holds.

We can now prove that $\sum_{i=1}^N U^{p(n+ik)} 1_A$ gets arbitrarily close to f : fix $\epsilon > 0$ and let f' be such that $U^k f' = f'$ for some non-zero integer k and $\|f - f'\| < \epsilon$ (this can be done because $f \in H_1$). Note that if m is multiple of k then still $U^m f' = f'$ and so $\|U^m f - f\| = \|U^m f - U^m f' + f' - f\| < 2\epsilon$ (because U is unitary). Let $n \in A_k$, taking averages we get $\left\| \frac{1}{N} \sum_{i=1}^N U^{p(n+ik)} f - f \right\| < 2\epsilon$ and so for N large enough, using (1), we also have that $\left\| \frac{1}{N} \sum_{i=1}^N U^{p(n+ik)} 1_A - f \right\| < 3\epsilon$.

Since 00. Putting this together with the last equation we get

$$3\epsilon > \left| \left\langle \frac{1}{N} \sum_{i=1}^N U^{p(n+ik)} 1_A - f, 1_A \right\rangle \right| = \left| \frac{1}{N} \sum_{i=1}^N \langle U^{p(n+ik)} 1_A, 1_A \rangle - c \right|$$

and so for some i we have $|\langle U^{p(n+ik)} 1_A, 1_A \rangle| > c - 3\epsilon$, and so putting ϵ small enough we conclude that $\mu(A \cap T^{p(n+ik)} A) > 0$. \square

EDIT(17/01/2020): I've edited the post, restricting the Furstenberg correspondence principle (Theorem 3) only to commutative semigroups admitting a Følner sequence. Formerly the post contained the (wrong) claim that if a countable semigroup is amenable, then it admits a Følner sequence. Moreover, the proof of Theorem 3 as presented didn't work for non-commutative semigroups.

The Furstenberg Correspondence principle is valid for arbitrary amenable semigroups, but a different proof is required.

Share this:



Like

Be the first to like this.

Related

[Szemerédi's Theorem Part I - Equivalent formulations](#)
In "Combinatorics"

[Szemerédi's Theorem Part II – Overview of the proof](#)
In "Combinatorics"

[Proof of Roth's Theorem using Ergodic Theory](#)
In "Ergodic Theory"

This entry was posted in [Combinatorics](#), [Ergodic Theory](#), [Tool](#) and tagged [Furstenberg](#), [recurrence](#). Bookmark the [permalink](#).

17 Responses to *Furstenberg's Correspondence Theorem*

Pingback: [YAMB](#)

Pingback: [Recurrence theorems | YAMB](#)

Pingback: [Proof of Roth's Theorem using Ergodic Theory | YAMB](#)

Pingback: [Convergence along ultrafilters – part II | YAMB](#)

Pingback: [Sets of nice recurrence | YAMB](#)

Pingback: [YAMB](#)

Pingback: [On \$\{x+y,xy\}\$ patterns in large sets of countable fields | I Can't Believe It's Not Random!](#)

Pingback: [Applications of the coloring trick | I Can't Believe It's Not Random!](#)

Pingback: [Convergence and Recurrence of \$Z\$ actions | I Can't Believe It's Not Random!](#)

Pingback: [Szemerédi's Theorem Part I – Equivalent formulations | I Can't Believe It's Not Random!](#)



isomorphisms says:

14/08/2015 at 7:20 am

Reblogged this on [Human Mathematics](#).

[Reply](#)

Pingback: [Measure preserving actions of affine semigroups and \$\{x+y,xy\}\$ patterns | I Can't Believe It's Not Random!](#)

Pingback: [I Can't Believe It's Not Random!](#)



E says:

16/01/2020 at 3:13 pm

Hi Joel, thanks for your post! It's very informative. I was just wondering if your definition of “amenable” (that there exists a Folner sequence/net) is really sufficient for most working definitions of amenability. For example, commutative semigroups ought to be amenable, and yet the multiplicative semigroup $\mathbb{N} \sqcup \{0\}$ does not admit a Folner sequence because 0 is an absorbing element.

Even if you're given a Folner sequence (or net), constructing a finitely-additive, invariant probability measure on the power set of G doesn't seem to work unless G satisfies some cancellation condition. Absorbing elements seem to screw things up. Even the action $(gx)_n = x_{gn}$ seems to be more of an “anti” action in the sense that $(gh)x = h(gx)$; for groups you could fix this by setting $gx := x \circ \lambda_{g^{-1}}$ where λ_g denotes left-translation. So amenability of semigroups in terms of Folner sequences doesn't seem to be as easy to work with as it is for groups.

[Reply](#)



Joel Moreira says:

17/01/2020 at 6:45 am

Dear E, thank you for your comment!

You're right of course that not every countable amenable semigroup has a Folner sequence (this was a misconception I had at the time I wrote this post), and indeed general (even amenable) semigroups can be much weirder than those embedable into groups.

And you're also right in that the proof presented of the Furstenberg Correspondence Principle only works for commutative semigroups.

I've edited the post to reflect this (perhaps I will later write a new post with a more general and modern proof of the Correspondence principle).

However, strictly speaking, absorbing elements do not prevent invariant means/Folner sequences from existing. Indeed, in the semigroup $\mathbb{N} \sqcup \{0\}$ endowed with multiplication, a “Dirac measure” at 0 is an invariant mean (and we can take as Folner sequence a constant sequence where every F_N consists only of the singleton $\{0\}$). This is of course quite degenerate...

[Reply](#)



E says:

17/01/2020 at 7:31 pm

Thanks Joel for the quick feedback. I appreciate it! I found recently that if G admits a Folner net, then in fact it admits a Folner net (F_λ) in which each member is “cancellable”, meaning $gx=gy \implies x=y$ for all $g \in G$ and $x,y \in F_\lambda$. This is a clever theorem of Gray–Kambites. Once you have this sequence, you can construct a finitely-additive invariant probability measure on the power set of G in the way one would usually do for groups (using an ultralimit in place of a limsup in the definition of the upper density). Unfortunately, without the cancellation property, the ultralimit construction may fail to result in an *invariant* measure, as far as I can tell.



E says:

03/05/2020 at 11:54 pm

I just realized that for $\mathbb{N} \sqcup \{0\}$ under multiplication, the constant sequence $F_n = \{0\}$ is actually a Folner sequence. All commutative semigroups should admit “strong” Folner sequences and are therefore amenable (this is a theorem of Argabright and Wilde).

[Reply](#)

I Can't Believe It's Not Random!

Blog at WordPress.com.