 Pavel Zorin-Kranich, zorin-kranich@uva.nl

The purpose of this note is to show that a compact extension of a measure-preserving dynamical system is isometric (Theorem 22). The main step in the proof is Lemma 20 that shows that every function f on the extension correlates with a generalized eigenfunction unless f is conditionally weakly mixing.

Our tools are Lemma 7 and measurable functional calculus.

We denote by $\pi : (X, \mathscr{X}, \mu, T) \to (Y, \mathscr{Y}, \mu, T)$ a factor map of measure-preserving dynamical systems. We usually identify $\mathscr{Y}$ with a subalgebra of $\mathscr{X}$.

In most of the text we assume that Y is ergodic and T is invertible on X. We also assume that X is a regular measure space, so that we can take advantage of measure disintegration.

We begin with a review of results that we shall use.

Conditional expectation

Definition 1. The *conditional expectation* is the orthogonal projection

$$\mathbb{E}(\cdot|Y) : L^2(X) \to L^2(Y).$$

Lemma 2. *The conditional expectation has the following properties.*

1. $\mathbb{E}(1|Y) = 1$.
2. $\int \mathbb{E}(f|Y) = \int f$ *for* $f \in L^1(X)$.
3. *Let* $f \in L^2(X)$ *and* $F \in \mathscr{Y}$. *Then* $\mathbb{E}(f 1_F|Y) = \mathbb{E}(f|Y)1_F$.
4. *Conditional expectation maps positive functions in* $L^1(X)$ *to positive functions.*
5. $\mathbb{E} : L^\infty(X) \to L^\infty(Y)$ *is a contraction.*
6. $\mathbb{E} : L^1(X) \to L^1(Y)$ *is a contraction.*
7. *Assume that* $f \in L^1(X), g \in L^0(Y)$ *and either* $fg \in L^1(X)$ *or* $f \geq 0, \mathbb{E}(f|Y)g \in L^1(Y)$. *Then*
$$\mathbb{E}(fg|Y) = \mathbb{E}(f|Y)g \text{ in } L^1(Y).$$

This is of course well-known but it is important to have the weakest possible assumptions in (7).

Proof. (1) holds since $1 \in L^2(Y)$.

(2) holds for $f \in L^2(X)$ since $\int \mathbb{E}(f|Y) = \langle \mathbb{E}(f|Y), 1\rangle = \langle f, \mathbb{E}(1|Y)\rangle = \langle f, 1\rangle = \int f$.

For (3) use that the orthogonal projection minimizes the distance.

To show (4) let $0 \leq f \in L^2(X)$ and $F = \{\mathbb{E}(f|Y) < 0\}$. Then $\|f 1_F - 0\| < \|f 1_F - 1_F \mathbb{E}(f|Y)\| = \|f 1_F - \mathbb{E}(f 1_F|Y)\|$, which is a contradiction, unless $F = \emptyset$.

To show (5) use (4) and (1).

Since $\mathbb{E}$ is self-adjoint and $L^\infty(X) \subset L^2(X)$ this implies (6). Thus it can be extended to a contraction $L^1(X) \to L^1(Y)$ by continuity. The properties (2) and (4) continue to hold for $f \in L^1(X)$.

Consider now (7). By linearity we obtain $\mathbb{E}(fg|Y) = \mathbb{E}(f|Y)g$ for $f \in L^2(X)$ and simple functions $g \in L^\infty(Y)$. By density we may weaken the assumption to $f \in L^1(X)$. The monotone convergence theorem shows that the same holds whenever $g \in L^0(Y)$ is such that $fg \in L^1(X)$ and $\mathbb{E}(f|Y)g \in L^1(Y)$.

By (4), the monotone convergence theorem, (2) and monotone convergence theorem again we see that

$$\int |\mathbb{E}(f|Y)g| \leq \int \mathbb{E}(|f||Y)|g| = \lim_k \int \mathbb{E}(|f||Y)|g^k| = \lim_k \int \mathbb{E}(|f||g^k||Y) = \lim_k \int |f||g^k| = \int |fg|,$$

where g^k denotes the truncation of g by k. Hence $fg \in L^1(X)$ implies that $\mathbb{E}(f|Y)g \in L^1(Y)$. Moreover, if $f \geq 0$ then the inequality turns into an equality and we obtain the converse implication. □

Conditional almost periodicity (in measure)

We mostly follow Tao's conventions and exposition [3] [1].

Definition 3. The conditional scalar product of $f, g \in L^2(X)$ is

$$\langle f, g\rangle_{L^2(X|Y)} := \mathbb{E}(f\bar{g}|Y) \in L^1(Y)$$

and the conditional norm of $f \in L^2(X)$ is

$$\|f\|_{L^2(X|Y)} := \langle f, f\rangle_{L^2(X|Y)}^{1/2} = \mathbb{E}(|f|^2|Y)^{1/2} \in L^2(Y).$$

The space $L^2(X|Y)$ is the space of $f \in L^2(X)$ such that $\|\|f\|_{L^2(X|Y)}\|_{L^\infty(Y)}$ is finite.

The space $L^2(X|Y)$ is a $L^\infty(Y)$-module. A *finitely generated submodule* is a set $f_1 L^\infty(Y) + \cdots + f_n L^\infty(Y)$ with some $f_1, \ldots, f_n \in L^2(X|Y)$. A *finitely generated module zonotope* is a set $f_1 B + \cdots + f_n B$ with some $f_1, \ldots, f_n \in L^2(X|Y)$, where $B = B_{L^\infty(Y)}$ is the closed unit ball of $L^\infty(Y)$.

Definition 4. A subset $E \subset L^2(X|Y)$ is said to be conditionally precompact if for every $\varepsilon > 0$ there exists a finitely generated module zonotope $Z \subset L^2(X|Y)$ such that E lies within the ε-neighborhood of Z for the norm $\|\|\cdot\|_{L^2(X|Y)}\|_{L^\infty(Y)}$.

A function $f \in L^2(X|Y)$ is said to be *conditionally almost periodic* if its orbit under T is conditionally precompact.

A function $f \in L^2(X|Y)$ is said to be *conditionally almost periodic in measure* if for every $\varepsilon > 0$ there exists a subset $E \subset Y$ of measure at least $1-\varepsilon$ such that $f 1_E$ is conditionally almost periodic.

The extension $X \to Y$ is called *compact* if every function in $L^2(X|Y)$ is conditionally almost periodic in measure.

Lemma 5. *The subset of c.a.p. functions $CAP(X|Y,T) \subset L^2(X|Y)$ is a shift-invariant $L^\infty(Y)$-module that contains the constants. It is closed for the $L^2(X|Y)$ norm.*

The subset of c.a.p.m. functions $CAPM(X|Y,T) \subset L^2(X|Y)$ is a shift-invariant $L^\infty(Y)$-module that contains the constants. It is the closure of $CAP(X|Y,T)$ in $L^2(X|Y)$ for the $L^2(X)$ norm.

The notation *CAP* and *CAPM* is not standard. We shall also write $CAP(X|Y)$ of just *CAP* instead of $CAP(X|Y,T)$ and analogously for *CAPM* if no confusion can arise. We usually confine ourselves to proving resultss about real-valued c.a.p. or c.a.p.m. functions, the extension to the complex-valued case should not be difficult.

[1] see http://terrytao.wordpress.com/2008/02/27/254a-lecture-13-compact-extensions/

Proof. That the two sets are shift-invariant $L^\infty(Y)$-modules and contain the constants is clear from definitions. Denseness of the former in the latter in the $L^2(X)$ norm also follows from the definition.

The only non-trivial assertion is the closedness of $CAPM(X|Y,T) \subset L^2(X|Y)$ in the $L^2(X)$ norm.

Let f_n be a sequence of c.a.p.m. functions that converges in $L^2(X)$ norm to $f \in L^2(X|Y)$. Let $\varepsilon > 0$ be given and find for each f_n a measurable set $A_n \subset Y$ with measure at least $1-\varepsilon/2^n$ such that $f_n 1_{A_n}$ is c.a.p.

Let $A := \cap_n A_n$. This is a set of measure at least $1-\varepsilon$. Furthermore $f_n 1_A \to f 1_A$ in $L^2(X)$. This implies $\|f_n 1_A - f 1_A\|_{L^2(X|Y)} \to 0$ in $L^2(Y)$. Passing to a subsequence we may assume uniform convergence on a measurable set $B \subset Y$ of measure at least $1-\varepsilon$. Therefore $f_n 1_A 1_B \to f 1_A 1_B$ in $L^2(X|Y)$.

Since the c.a.p. functions form a $L^\infty(Y)$-module we see that the functions $f_n 1_A 1_B$ are c.a.p., so that $f 1_A 1_B$ is also c.a.p. Since ε is arbitrary the function f is c.a.p.m. by definition. □

The main difficulty in dealing with $CAPM$ is that $L^\infty(X)$ is in general not dense in $L^2(X|Y)$ for the norm $\|\|\cdot\|_{L^2(X|Y)}\|_{L^\infty(Y)}$. Example: $Y=[0,1]$, $X = Y\times[0,1]$, $f(y,z) = 1_{z<y} y^{-1/2} \in L^2(X|Y)$ cannot be approximated by bounded functions.

This problem is addressed by Lemma 8. From now on we make the standing assumption that Y is ergodic and T is invertible on X. We also assume that X is a regular measure space, so that we can take advantage of measure disintegration.

Theorem 6 ([1, Theorem 5.8]). *Suppose that X is a regular measure space. Then there exists an essentially unique measurable map $Y \to \mathcal{M}(X)$, $y \mapsto \mu_y$ such that whenever $f \in L^1(X)$ we have $f \in L^1(X,\mu_y)$ and $\int f d\mu_y = \mathbb{E}(f|Y)(y)$ for ν-a.e. y.*

In particular $\int\int f(x) d\mu_y(x) d\nu(y) = \int f d\mu$ whenever $f \in L^1(X)$.

The essential uniqueness of the disintegration $\mu = \int \mu_y d\nu(y)$ readily implies that $\mu_{Ty} = T\mu_y$.

The following technical lemma is implicitely used in [1, (6.14)] and will be useful later on.

Lemma 7. *We have $\mu_{\pi(x)} = \mu_y$ for ν-a.e. $y \in Y$ and μ_y-a.e. $x \in X$.*

Proof. Since X is a regular measure space it suffices to prove $\mu_{\pi(x)}(f) = \mu_y(f)$ for some countable set of bounded continuous functions f (and appropriate y, x). Hence it also suffices to consider just one $f \in L^\infty(X)$. Let $g = \mathbb{E}(f|Y) \in L^\infty(Y)$. Then

$$\begin{aligned}\int\int |g(\pi(x)) - g(y)|^2 d\mu_y(x) d\nu(y) &= \int\int g\bar{g}(\pi(x)) - 2\Re g(\pi(x))\bar{g}(y) + g\bar{g}(y) d\mu_y(x) d\nu(y)\\ &= \int\int (g\bar{g}\circ\pi)(x) - 2\Re(g\circ\pi)(x)\bar{g}(y) + g\bar{g}(y) d\mu_y(x) d\nu(y)\\ &= \int\int g\bar{g}(y) - 2\Re g(y)\bar{g}(y) + g\bar{g}(y) d\nu(y)\\ &= 0\end{aligned}$$

since $\int (h\circ\pi)(x) d\mu_y(x) = \mathbb{E}(h\circ\pi)(y) = h(y)$ for a.e. y whenever $h \in L^\infty(Y)$.

Hence $\mu_{\pi(x)}(f) = g(\pi(x)) = g(y) = \mu_y(f)$ for a.e. y and μ_y-a.e. x. Note that the first equality holds whenever $\pi(x)$ lies outside of a fixed null set, hence x lies outside of a fixed μ-null set. But a μ-null set is μ_y-null for a.e. y. □

Lemma 8. *Let $f \in L^2(X|Y)$ be c.a.p. Then the orbit of f can be approximated by finitely generated module zonotopes spanned by* bounded *functions.*

Proof. Fix $\varepsilon > 0$. Then there exist $f_1, \dots, f_n \in L^2(X|Y)$ such that

$$\mathrm{orb}(f) \subset U_\varepsilon(\sum_i Bf_i).$$

Consider the bounded functions $f_i^k = f_i 1_{\{|f_i| \le k\}}$. Then $|f_i - f_i^k| \to 0$ pointwise monotonously, hence also in $L^2(X)$. Therefore $\mathbb{E}(|f_i - f_i^k|^2 | Y) \to 0$ in $L^1(Y)$ monotonously. Take k so large that $\mathbb{E}(|f_i - f_i^k|^2|Y) < \varepsilon/n$ outside a set F_i of measure at most ε/n for all $i = 1, \dots, n$. Thus we can split $f_i = g_i + \tilde{b}_i$, where $g_i = f_i^k \in L^\infty(X)$ is a good function and $\tilde{b}_i \in L^2(X|Y)$ is a bad function. If we set $b_i = \tilde{b}_i 1_{F_i}$ then $\| \| b_i - \tilde{b}_i \|_{L^2(X|Y)} \|_{L^\infty(Y)} \le \varepsilon/n$, and in particular

$$\mathrm{orb}(f) \subset U_{2\varepsilon}(\sum_i B g_i + \sum_i B b_i).$$

Let $K = Y \setminus \cup_i F_i$, this set has measure at least $1-\varepsilon$. We have $\| \| 1_K T^l f - \sum_i h_i g_i \|_{L^2(X|Y)} \|_{L^\infty(Y)} < 2\varepsilon$ for all l and some $h_i \in B$ depending on l. Here $\| \sum_i h_i g_i \|_{L^\infty(X)} \le kn =: M$. In other words,

$$\| \| 1_{T^l K} f - T^{-l} s_l \|_{L^2(X|Y)} \|_{L^\infty(Y)} < 2\varepsilon$$

for every l and some functions s_l uniformly bounded by M. By ergodicity of (Y, T) the translates $T^l K$ cover Y. Let

$$\tilde{f} = \sum_{l \ge 0} 1_{T^l K} T^{-l} s_l \prod_{0 \le l' < l} 1_{T^l K^\complement}.$$

Then $\|\tilde{f}\|_{L^\infty(X)} \le M$ and $\| \| \tilde{f} - f \|_{L^2(X|Y)} \|_{L^\infty(Y)} \le 2\varepsilon$. In particular we have

$$\mathrm{orb}(\tilde{f}) \subset U_{4\varepsilon}(\sum_i B g_i + \sum_i B b_i).$$

Let now $N = \frac{\sqrt{n}M}{\varepsilon} \sup_i \| \| b_i \|_{L^2(X|Y)} \|_{L^\infty(Y)}$ and $E = \cup_i \{ |b_i| > N \} \subset X$. For a given $n \in \mathbb{Z}$ choose $\gamma_i^n, \beta_i^n \in B$ such that

$$\| \| T^n \tilde{f} - \sum_i \gamma_i^n g_i + \sum_i \beta_i^n b_i \|_{L^2(X|Y)} \|_{L^\infty(Y)} < 4\varepsilon.$$

Now we use the disintegration of measure $\mu = \int_Y \mu_y \nu(y)$. For almost every $y \in Y$ we have

$$\| T^n \tilde{f} - \sum_i \gamma_i^n g_i + \sum_i \beta_i^n b_i \|_{L^2(X, \mu_y)} < 4\varepsilon$$

and

$$\mu_y(E) \le \sum_i \mu_y \{ |b_i| > N \} \le \sum_i \| b_i \|^2_{L^2(X, \mu_y)} / N^2 \le \varepsilon^2 M^{-2}.$$

Therefore

$$\| T^n \tilde{f} - \sum_i \gamma_i^n g_i 1_{E^\complement} + \sum_i \beta_i^n b_i 1_{E^\complement} \|^2_{L^2(X,\mu_y)} \le \| T^n \tilde{f} - \sum_i \gamma_i^n g_i + \sum_i \beta_i^n b_i \|^2_{L^2(X,\mu_y)} + \| T^n \tilde{f} \|^2_{L^\infty(X)} \mu_y(E) \le 16\varepsilon^2 + \varepsilon^2.$$

This shows that

$$\mathrm{orb}(\tilde{f}) \subset U_{5\varepsilon}(\sum_i B g_i 1_{E^\complement} + \sum_i B b_i 1_{E^\complement}),$$

and therefore

$$\mathrm{orb}(f) \subset U_{7\varepsilon}(\sum_i B g_i 1_{E^\complement} + \sum_i B b_i 1_{E^\complement}).$$

This zonotope is generated by bounded functions. Since ε is arbitrary we are done. □

Lemma 9. *Let $f \in CAP(X|Y,T)$. Then every function of the form*

$$\tilde{f}(x) = \begin{cases} a, & f(x) > a, \\ f(x), & b \le f(x) \le a, \\ b, & f(x) < b \end{cases}$$

is also c.a.p.

Proof. Since linear combinations of c.a.p. functions are c.a.p. and constants are c.a.p. it suffices to prove that the positive part of f is c.a.p.

By Lemma 8 we know that for every $\varepsilon > 0$ the orbit of f lies in an ε-neighborhood (with respect to the $L^2(X|Y)$-norm) of a zonotope generated by some bounded functions $f_1,\dots,f_n$. By approximation in the supremum norm we may assume that each f_i is a simple function. Splitting each f_i we may assume that they are characteristic functions with either disjoint or equal supports. Combining identical characteristic functions we may assume that the f_i's are positive multiples of characteristic functions with disjoint supports.

Take $h_i \in B$ such that $||||T^k f - \sum_i h_i f_i||_{L^2(X|Y)}||_{L^\infty(Y)} < \varepsilon$. Then $||||T^k f^+ - \sum_i h_i^+ f_i||_{L^2(X|Y)}||_{L^\infty(Y)} < \varepsilon$. Therefore orb$f^+$ can be approximated by the same zonotope as orb$f$, so that the positive part f^+ is c.a.p. □

Lemma 10. *Let $(g_n) \subset L^\infty(X)$ be a uniformly bounded sequence of c.a.p.m. functions that converges in measure to g, i.e. for every $\varepsilon > 0$ the measure of $\{|g_n - g| > \varepsilon\}$ converges to zero.*

Then g is c.a.p.m.

Proof. Convergence in measure and uniform boundedness imply convergence in $L^2(X)$, so that $\mathbb{E}(|g - g_n|^2|Y) \to 0$ in $L^1(Y)$. Passing to a subsequence we may assume uniform convergence on a set A of arbitrarily large measure $1-\delta$. Shrinking A slightly (by at most δ) we may assume that each $g_n 1_A \in CAP$. Furthermore, $g_n 1_A \to g 1_A$ in $L^2(X|Y)$, hence $g 1_A \in CAP$. Since δ was arbitrary, $g \in CAPM$ by definition. □

Lemma 11. *Let $f \in CAPM$. Then the indicator function of every sub- or superlevel set of f is c.a.p.m.*

Proof. Suppose that f is c.a.p. first. By translation we see that we only need to consider $F = \{f > 0\}$ and $F = \{f \ge 0\}$.

In the former case we define

$$g_n(x) = \begin{cases} 0, & f(x) \le 0, \\ nf(x), & 0 < f(x) < 1/n, \\ 1, & f(x) \ge 1/n. \end{cases}$$

In the latter case we define

$$g_n(x) = \begin{cases} 0, & f(x) \le -1/n, \\ 1 + nf(x), & -1/n < f(x) < 0, \\ 1, & f(x) \ge 0. \end{cases}$$

In either case $g_n \to 1_F$ monotonously and by the monotone convergence theorem also in measure. Moreover $g_n \in CAP$ by Lemma 9. Hence $1_F \in CAPM$ by Lemma 10.

Suppose now that $f \in CAPM$. Let $a \in \mathbb{R}$. Let $\varepsilon > 0$ and A be a set with measure greater than $1 - \varepsilon$ such that $f 1_A \in CAP$.

By the above the indicator function of the set $\{f 1_A > a\}$ is c.a.p.m., hence there exists a set $F \subset \{f 1_A > a\}$ such that 1_F is c.a.p. and $\mu(\{f 1_A > a\} \setminus F)$ is at most ε. But then $\mu(\{f > a\} \Delta F) \leq 2\varepsilon$. Hence 1_F converges to $1_{\{f>a\}}$ in measure as $\varepsilon \to 0$, so that $1_{\{f>a\}}$ is cond.a.p. in measure by Lemma 10.

The proof for $\{f \geq a\}$ is the same. □

Definition 12. Let $\mathscr{X}_{X|Y}$ denote the collection of all measurable subsets $E \subset X$ such that 1_E is c.a.p. in measure.

Since the set of c.a.p.m. functions is an invariant linear subspace of $L^2(X)$ and contains the constants we see that $\mathscr{X}_{X|Y}$ is closed under complementation and translation by T. Lemma 11 shows that $\mathscr{X}_{X|Y}$ is closed under finite unions. Monotone convergence theorem and Lemma 10 show that $\mathscr{X}_{X|Y}$ is closed under countable unions. Clearly we have $X, \emptyset \in \mathscr{X}_{X|Y}$. Hence $\mathscr{X}_{X|Y}$ is a shift-invariant sub-σ-algebra of $\mathscr{X}$. Since every function in $L^\infty(Y)$ is c.a.p. we have $\mathscr{X}_{X|Y} \supset \pi^\# \mathscr{Y}$.

Theorem 13. *A function $f \in L^2(X|Y)$ is conditionally almost periodic in measure if and only if it is $\mathscr{X}_{X|Y}$-measurable.*

Proof. If $f \in L^2(X|Y)$ is c.a.p.m. then it is $\mathscr{X}_{X|Y}$-measurable by Lemma 11.

Conversely, a $\mathscr{X}_{X|Y}$-measurable function in $L^2(X|Y)$ is an $L^2(X)$-limit of simple $\mathscr{X}_{X|Y}$-measurable functions, hence c.a.p.m. by Lemma 5. □

Compact extensions

Furstenberg defines a compact extension in slightly different terms.

Definition 14 ([1, p. 131])**.** The extension $X \to Y$ is called *compact* if for every $f \in L^2(X)$ and every $\varepsilon, \delta > 0$ there exists a Y-measurable set F with measure at least $1-\varepsilon$ and finitely many functions $g_1, \dots, g_k$ such that

$$\min_j ||T^l f 1_F - g_j||_y < \delta$$

for every l and a.e. $y \in Y$. Equivalently, there is a dense subset D of $L^2(X)$ such that for every $f \in D$ and every $\delta > 0$ there exist finitely many functions $g_1, \dots, g_k$ such that

$$\min_j ||T^l f 1_F - g_j||_y < \delta$$

for every l and a.e. $y \in Y$. (See Furstenberg's book for proof of equivalence).

We now show that $(X, \mathscr{X}_{X|Y}, \mu, T)$ is the maximal compact extension of Y inside X.

By Theorem 13 and Lemma 5 we know that $CAP(X|Y)$ is dense in $L^2(\mathscr{X}_{X|Y})$. By Lemma 8 we know that the orbit of every c.a.p. function can be approximated by finitely generated module zonotopes spanned by bounded functions. Taking all possible linear combinations of generators of the above zonotope with coefficients in a sufficiently dense finite subset of the complex unit ball we obtain a finite set of functions for which the last condition in the definition of a compact extension holds.

Conversely, let $f \in L^2(X)$ be any function satisfying the first condition. Then $f 1_{\{||f||_{L^2(X|Y)} < a\}}$ satisfies the same condition for any a. But this function is contained in $L^2(X|Y)$ and is c.a.p.m. (take the zonotope spanned by the g_j's).

By Theorem 13 we have that $f 1_{\{||f||_{L^2(X|Y)} < a\}}$ is $\mathscr{X}_{X|Y}$-measurable. But $f 1_{\{||f||_{L^2(X|Y)} < a\}} \to f$ pointwise a.e. as $a \to \infty$, so that f is also $\mathscr{X}_{X|Y}$-measurable.

Generalized eigenfunctions

Definition 15 ([2, Definition 6.1])**.** A *Y-module* is a closed linear subspace $M \subset L^2(X)$ such that whenever $f \in M$ and $h \in L^0(Y)$ are such that $hf \in L^2(X)$ then $hf \in M$.

A function $f \in L^2(X)$ is called a *generalized eigenfunction* if $\mathrm{orb}(f)$ is contained in a T-invariant finite rank Y-module.

The space of generalized eigenfunctions is denoted by $GE(X|Y,T)$. It is a vector subspace of $L^2(X)$. Furstenberg denoted its closure by $\mathscr{E}(X|Y,T)$ in [2, Definition 6.3]. If $\mathscr{E}(X|Y,T) = L^2(X)$ then the extension $X \to Y$ is called *isometric*, [2, §8].

Lemma 16. *A generalized eigenfunction is an $L^2(X)$-limit of c.a.p.m. functions.*

Proof. By truncation on Y we may assume $f \in L^2(X|Y)$.

Let $f \in L^2(X)$ be a generalized eigenfunction such that $\mathrm{orb}(f)$ is contained in a finite rank T-invariant Y-module M.

Let $F = \{\|f\|_{L^2(X|Y)} > k\}$ and $\tilde f = f 1_F \in L^2(X|Y)$. If k is sufficiently large then $\|f - \tilde f\|_{L^2(X)}$ is small. Furthermore $\mathrm{orb}(\tilde f)$ is contained in M.

Let $f_1,\dots,f_n \in L^2(X)$ be the generators of M. Let $g : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be defined as

$$g(a) := \begin{cases} 1, & a = 0 \\ 1/a, & a > 0 \end{cases}$$

and consider the functions $g_j := g \circ \|f_j\|_{L^2(X|Y)} \in L^0(Y)$. Then $g_j f_j \in L^2(X|Y)$ and these functions also generate M as a Y-module. Hence we may assume that $f_1,\dots,f_n \in L^2(X|Y)$.

By the Gram-Schmidt procedure we may assume that $f_1,\dots,f_n$ are orthogonal w.r.t. the conditional scalar product. Multiplying each f_j by $g_j := g \circ \|f_j\|_{L^2(X|Y)} \in L^0(Y)$ we may assume that $\|f_j\|_{L^2(X|Y)}$ is zero-one-valued.

Splitting each f_j and reassembling we may assume that $\mathrm{supp}\|f_1\|_{L^2(X|Y)} \supset \dots \supset \mathrm{supp}\|f_n\|_{L^2(X|Y)}$.

Let J be the last index such that $\mathrm{supp}\|f_J\|_{L^2(X|Y)} = Y$ and let $F = Y \setminus \mathrm{supp}\|f_{J+1}\|_{L^2(X|Y)}$. Since T is an algebra homomorphism that commutes with $\mathbb{E}(\cdot|Y)$, T is conditionally unitary, and in particular $(Tf_1)|_F,\dots,(Tf_n)|_F$ is an orthonormal set in $M|_F$. (TODO: be careful here) In particular it is an orthonormal set in the J-dimensional space $L^2(X,\mu_y)$ for almost every $y \in F$, hence $\mathrm{supp}\|f_{J+1}\|_{L^2(X|Y)}$ cannot intersect F, hence F is an invariant subset, hence $F = Y$ by ergodicity and $\mathrm{supp}\|f_{J+1}\|_{L^2(X|Y)} = \emptyset$.

Thus we have obtained a conditionally orthonormal generating set of M. If now $T^k \tilde f = \sum_j h_j f_j$ with $h_j \in L^0(Y)$ such that $h_j f_j \in L^2(X)$ then

$$\|T^k \tilde f\|_{L^2(X|Y)}^2 = \sum_j |h_j|^2 \|f_j\|_{L^2(X|Y)}^2 = \sum_j |h_j|^2 \in L^\infty(Y)$$

with a bound that is uniform in k. In particular the orbit of $\tilde k$ is contained in a finitely generated module zonotope (with some scalar multiples of $f_1,\dots,f_n$ as generators). By definition $\tilde f$ is c.a.p.

From the construction of $\tilde f$ we see that f is c.a.p.m. □

Conditional weak mixing

Definition 17. A function $f \in L^2(X|Y)$ is called *conditionally weakly mixing* if

$$C - \lim_n \| \langle T^n f, f \rangle_{L^2(X|Y)} \|_{L^2(Y)}^2 = 0.$$

Here $C-\lim$ stands for the Cesàro limit, i.e. $C-\lim_n a_n = \lim_N \frac{1}{N}\sum_{n=0}^{N-1} a_n$.
The space of c.w.m. functions is denoted by $CWM(X|Y,T)$ (this not a standard notation).

Lemma 18. *Let $f \in CWM(X|Y,T)$ and $g \in L^2(X|Y)$. Then $C-\lim_n \| \langle T^n f, g\rangle_{L^2(X|Y)} \|^2_{L^2(Y)} = 0$.*

This shows in particular that $CWM \subset L^2(X|Y)$ is a linear subspace.

Proof. Note that $\langle g, T^n f\rangle_{L^2(X|Y)}$ is uniformly bounded in $L^\infty(Y)$ by conditional Cauchy-Schwarz.

$$\begin{aligned}\frac{1}{N}\sum_{n=0}^{N-1} \| \langle T^n f, g\rangle_{L^2(X|Y)} \|^2_{L^2(Y)} &= \frac{1}{N}\sum_{n=0}^{N-1} \int_Y \langle g, T^n f\rangle_{L^2(X|Y)} \langle T^n f, g\rangle_{L^2(X|Y)} \\ &= \frac{1}{N}\sum_{n=0}^{N-1} \int_Y \left\langle \langle g, T^n f\rangle_{L^2(X|Y)} T^n f, g\right\rangle_{L^2(X|Y)} \\ &= \left\langle \frac{1}{N}\sum_{n=0}^{N-1} \langle g, T^n f\rangle_{L^2(X|Y)} T^n f, g \right\rangle_{L^2(X)}\end{aligned}$$

We wish to apply the van der Corput lemma. For this it suffices to show that

$$C-\lim_h C-\sup_n \left| \left\langle \langle g, T^n f\rangle_{L^2(X|Y)} T^n f, \left\langle g, T^{n+h} f\right\rangle_{L^2(X|Y)} T^{n+h} f \right\rangle_{L^2(X)} \right| = 0.$$

Indeed, the scalar product can be estimated as follows

$$\begin{aligned}\dots &= \left| \int_X \langle g, T^n f\rangle_{L^2(X|Y)} T^n f \overline{\left\langle g, T^{n+h} f\right\rangle_{L^2(X|Y)}} T^{n+h} \bar{f} \,\mathrm{d}\mu \right| \\ &= \left| \int_Y \langle g, T^n f\rangle_{L^2(X|Y)} \overline{\left\langle g, T^{n+h} f\right\rangle_{L^2(X|Y)}} \mathbb{E}(T^n f T^{n+h} \bar{f} | Y) \mathrm{d}\nu \right| \\ &\leq C \left| \int_Y \mathbb{E}(T^n f T^{n+h} \bar{f} | Y) \mathrm{d}\nu \right| \\ &= C \left| \int_Y \mathbb{E}(f T^h \bar{f} | Y) \mathrm{d}\nu \right| \\ &\leq C \left(\int_Y |\mathbb{E}(f T^h \bar{f} | Y)|^2 \mathrm{d}\nu \right)^{1/2} \\ &= C \| \left\langle f, T^h f\right\rangle_{L^2(X|Y)} \|_{L^2(Y)}.\end{aligned}$$

The convergence to zero in Cesaro sense follows from definition of conditional weakly mixing and uniform boundedness of this sequence. □

Lemma 19. *Let $f \in CWM(X|Y,T)$ and $g \in CAP(X|Y,T)$. Then $f \perp_{L^2(X|Y)} g$.*

Proof. Let $g_1, \dots, g_r \in L^2(X|Y)$ generate a module zonotope Z such that g lies within ε of Z. By Lemma 18 we have

$$C-\lim_n \| \langle T^n f, g_i\rangle_{L^2(X|Y)} \|_{L^2(Y)} = 0$$

for every i. Multiplying g_i by a function in $B_{L^\infty(Y)}$ can only improve the rate of convergence. Therefore

$$C-\lim_n \| \langle T^n f, T^n g\rangle_{L^2(X|Y)} \|_{L^2(Y)} \leq \varepsilon.$$

But the sequence on the left-hand side is identically $\|\langle f,g\rangle_{L^2(X|Y)}\|_{L^2(Y)}$ and $\varepsilon>0$ is arbitrary. Hence $\|\langle f,g\rangle_{L^2(X|Y)}\|_{L^2(Y)}=0$. □

Lemma 20. *If $f\in L^2(X|Y)\setminus CWM(X|Y,T)$ then $f\not\perp_{L^2(X)} GE(X|Y,T)$.*

Proof. Fix a measurable representative for f defined everywhere. By the hypothesis

$$\limsup_n \frac{1}{N}\sum_{n=0}^{N-1}\|\langle T^n f,f\rangle_{L^2(X|Y)}\|^2_{L^2(Y)}>0.$$

Passing to a subsequence we may assume that the limit exists and is positive. Applying the mean ergodic theorem to the fiber product $X\times_Y X$ we obtain a T-invariant function $H\in L^2(X\times_Y X)$ such that $\frac{1}{N}\sum_{n=0}^{N-1}T^n\bar f\otimes T^n f\to H$. This means that

$$\int\int\left|\frac{1}{N}\sum_{n=0}^{N-1}T^n\bar f\otimes T^n f-H\right|^2\mathrm{d}(\mu_y\times\mu_y)\mathrm{d}\nu(y)\to 0.$$

Fix a representative for H. Passing to a subsequence we may assume that

$$\int\left|\frac{1}{N}\sum_{n=0}^{N-1}T^n\bar f\otimes T^n f-H\right|^2\mathrm{d}(\mu_y\times\mu_y)\to 0 \tag{21}$$

for almost every $y\in Y$. In particular $\int|H|^2\mathrm{d}(\mu_y\times\mu_y)\le\|\|f\|_{L^2(X|Y)}\|^2_{L^\infty(Y)}$ uniformly in y. Moreover we may assume that $\bar H(x,x')=H(x',x)$.

For almost every y we can define an operator $S_y g(x):=\int H(x',x)g(x')\mathrm{d}\mu_y(x')$ on $L^2(X,\mu_y)$. By (21) almost every operator S_y is a Hilbert-Schmidt limit of finite rank operators and its Hilbert-Schmidt norm is uniformly bounded by $M=\|\|f\|_{L^2(X|Y)}\|_{L^\infty(Y)}$.

Let $\mathscr{L}^2(X)$ be the space of a.e. defined square integrable function on X (not equivalence classes). Define an operator S on $\mathscr{L}^2(X)$ by the formula $Sg(x)=S_{\pi(x)}g(x)$ [2]. Lemma 7 implies that

$$\begin{aligned}&\int\int|Sg(x)-S_y g(x)|^2\mathrm{d}\mu_y(x)\mathrm{d}\nu(y)\\&\quad=\int\int\left|\int H(x',x)g(x')\mathrm{d}\mu_{\pi(x)}(x')-\int H(x',x)g(x')\mathrm{d}\mu_y(x')\right|^2\mathrm{d}\mu_y(x)\mathrm{d}\nu(y)=0,\end{aligned}$$

so that $Sg=S_y g$ in $L^2(X,\mu_y)$ for a.e. y.

Note that the operator S passes to a bounded linear operator on $L^2(X)$ since

$$\begin{aligned}\|Sg\|^2_{L^2(X)}&=\int|Sg|^2\mathrm{d}\mu\\&=\int\int|Sg(x)|^2\mathrm{d}\mu_y(x)\mathrm{d}\nu(y)\\&=\int\int|S_y g(x)|^2\mathrm{d}\mu_y(x)\mathrm{d}\nu(y)\\&\le\int M^2\int|g(x)|^2\mathrm{d}\mu_y(x)\mathrm{d}\nu(y)\\&=M^2\|g\|^2_{L^2(X)}.\end{aligned}$$

[2] Furstenberg [1, (6.13)] writes $H*g$ for Sg

Using T-invariance of H we check that the operator S commutes with T.

$$\begin{aligned}STg(x) &= S_{\pi(x)}Tg(x)\\ &= \int H(x',x)Tg(x')\mathrm{d}\mu_{\pi(x)}(x')\\ &= \int H(Tx',Tx)g(Tx')\mathrm{d}\mu_{\pi(x)}(x')\\ &= \int H(x'',Tx)g(x'')\mathrm{d}\mu_{\pi(Tx)}(x'')\\ &= Sg(Tx).\end{aligned}$$

Note also that

$$\begin{aligned}\langle Sf,f\rangle_{L^2(X)} &= \int\int H(x',x)f(x')\bar{f}(x)\mathrm{d}\mu_{\pi(x)}(x')\mathrm{d}\mu(x)\\ &= \int\int\int H(x',x)f(x')\bar{f}(x)\mathrm{d}\mu_{\pi(x)}(x')\mathrm{d}\mu_y(x)\mathrm{d}\nu(y)\\ &= \lim_N \int\int\int \frac{1}{N}\sum_{n=0}^{N-1} T^n\bar{f}(x')T^nf(x)f(x')\bar{f}(x)\mathrm{d}\mu_{\pi(x)}(x')\mathrm{d}\mu_y(x)\mathrm{d}\nu(y)\\ &= \lim_N \frac{1}{N}\sum_{n=0}^{N-1}\int |\langle T^nf,f\rangle_{L^2(X|Y)}(y)|^2\mathrm{d}\nu(y)\\ &> 0.\end{aligned}$$

The operators S and S_y are self-adjoint by construction. By the measurable functional calculus there exists a constant $a > 0$ such that $\langle p(S)Sf,f\rangle_{L^2(X)} \neq 0$, where $p = \chi_{[-a,a]^{\complement}}$.

Let p_n be a sequence of polynomials such that $p_n(-a) = 0 = p_n(a)$ and $p_n \to p$ pointwise and boundedly on $[-M,M]$ and uniformly on $[-M,-a-\varepsilon]\cup[-a,a]\cup[a+\varepsilon,M]$ for every ε. Since S_y are self-adjoint Hilbert-Schmidt operators on Hilbert spaces, $\sigma(S_y)\setminus\{0\}$ is discrete. By the continuous functional calculus $p_n(S_y)$ converges in the operator norm topology to the projection onto the linear span of the eigenspaces of S_y with eigenvalues outside $[-a,a]$.

Recall that the Hilbert-Schmidt norm of S_y is uniformly bounded. Therefore the number of eigenspaces to eigenvalues with absolute value at least a is also uniformly bounded. Therefore the rank of $p(S_y)$ is uniformly bounded. Moreover $p_n(S) \to p(S)$ in the strong operator topology by the measurable functional calculus.

Let $g \in \mathscr{L}^2(X)$. For a.e. y and every n we have $p_n(S)g = p_n(S_y)g$ in $L^2(X,\mu_y)$. Here the right-hand side converges in $L^2(X,\mu_y)$. The left-hand side converges in $L^2(X)$, so we can pass to a subsequence such that the convergence is pointwise μ-almost everywhere, hence also pointwise μ_y-a.e. for a.e. y. Therefore the two limits coincide μ_y-a.e. for a.e. y, i.e. $p(S)g = p(S_y)g$ in $L^2(X,\mu_y)$.

The $L^\infty(Y)$-module $p(S)(L^2(X|Y)) \subset L^2(X|Y)$ has finite rank. In fact, using cutoffs in Y we can construct a sequence $(f_n) \subset L^2(X|Y)$ such that $\operatorname{supp}\|p(S)f_n\|_{L^2(X|Y)}$ has maximal measure subject to the condition that $p(S)f_{n+1} \notin \operatorname{lin}(p(S)f_1,\dots,p(S)f_n)$ in $L^2(X,\mu_y)$ for a.e. $y \in \operatorname{supp}\|p(S)f_n\|_{L^2(X|Y)}$. Since the rank of $p(S_y)$ is uniformly bounded we conclude that this sequence must become identically zero after a finite number of terms. A Gram-Schmidt procedure provides a finite sub-orthonormal base.

It remains to be shown that $p(S)(L^2(X|Y))$ is T-invariant. But $p_n(S) \to p(S)$ strongly and each of $p_n(S)$ commutes with T. Hence $p(S)$ commutes with T as an operator on $L^2(X)$.

Hence $p(S)Sf$ is a generalized eigenfunction that correlated with f. □

Theorem 22. *The closures of $GE(X|Y,T)$, $CAP(X|Y)$ and $CAPM(X|Y)$ in $L^2(X)$ coincide and are equal to $L^2(X, \mathscr{X}_{X|Y}, \mu)$ and $\mathscr{E}(X|Y)$.*

In particular every compact extension is isometric.

Proof. The space CAP is a dense subspace of $CAPM$ for the $L^2(X)$ norm by monotone convergence. Both have closure $L^2(X, \mathscr{X}_{X|Y}, \mu)$ by Theorem 13. The inclusion $GE \subset \overline{CAPM}^{L^2(X)}$ is given by Lemma 16.

By Lemma 19 we have $CWM \perp_{L^2(X|Y)} CAP$, hence also $CWM \perp_{L^2(X|Y)} CAPM$, hence also $CWM \perp_{L^2(X)} CAPM$.

Let $0 \neq f \in L^2(X, \mathscr{X}_{X|Y}, \mu)$ and let F be a sublevel set of $|\mathbb{E}(f|Y)|$ such that the truncation $\tilde{f} = f 1_F$ does not vanish identically. Then $\tilde{f} \in CAPM$ by Theorem 13, hence $\tilde{f} \notin CWM$.

By Lemma 20 the function $\tilde{f}$ correlates with some generalized eigenfunction g. But then $g1_F$ is also a generalized eigenfunction and f correlates with $g1_F$.

Hence the orthogonal complement of GE in $L^2(X, \mathscr{X}_{X|Y}, \mu)$ consists only of the zero, and we are done. □

Zimmer's definition of *(relatively) discrete spectrum* [6, II.4] differs from Furstenberg's definition of an isometric extension (Definition 15) in that the finite rank invariant submodules (corresponding to finite dimensional invariant subbundles) that span the space must be mutually orthogonal. But this can be achieved by a variant of Gram-Schmidt procedure.

A structure theorem for extensions with relatively discrete spectrum was announced by Zimmer in [4] and proved in [6, Theorem 4.3]. In [5] he studies towers of isometric extensions. The same structure theorem for isometric extensions was proved by Furstenberg [2].

References

[1] H. Furstenberg, *Recurrence in ergodic theory and combinatorial number theory*, Princeton University Press, Princeton, N.J., 1981. M. B. Porter Lectures. MR603625 (82j:28010) ↑3, 6, 9

[2] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256. MR0498471 (58 #16583) ↑7, 11

[3] T. Tao, *Poincaré's legacies, pages from year two of a mathematical blog. Part II*, American Mathematical Society, Providence, RI, 2009. MR2541289 (2010h:00021) ↑2

[4] R. J. Zimmer, *Extensions of ergodic actions and generalized discrete spectrum*, Bull. Amer. Math. Soc. **81** (1975), 633–636. MR0372160 (51 #8376) ↑11

[5] ———, *Ergodic actions with generalized discrete spectrum*, Illinois J. Math. **20** (1976), no. 4, 555–588. MR0414832 (54 #2924) ↑11

[6] ———, *Extensions of ergodic group actions*, Illinois J. Math. **20** (1976), no. 3, 373–409. MR0409770 (53 #13522) ↑11