

POINTWISE AND L^1 MIXING RELATIVE TO A SUB-SIGMA ALGEBRA

DANIEL J. RUDOLPH

ABSTRACT. We consider two natural definitions for the notion of a dynamical system being mixing relative to an invariant sub σ -algebra \mathcal{H} . Both concern the convergence of

$$|E(f \cdot g \circ T^n | \mathcal{H}) - E(f | \mathcal{H})E(g \circ T^n | \mathcal{H})| \rightarrow 0$$

as $|n| \rightarrow \infty$ for appropriate f and g . The weaker condition asks for convergence in L^1 and the stronger for convergence a.e. We will see that these are different conditions. Our goal is to show that both these notions are robust. As is quite standard we show that one need only consider $g = f$ and $E(f | \mathcal{H}) = 0$, and in this case $|E(f \cdot f \circ T^n | \mathcal{H})| \rightarrow 0$. We will see rather easily that for L^1 convergence it is enough to check an L^2 -dense family. Our major result will be to show the same is true for pointwise convergence, making this a verifiable condition. As an application we will see that if T is mixing then for any ergodic S , $S \times T$ is relatively mixing with respect to the first coordinate sub σ -algebra in the pointwise sense.

1. Introduction

Mixing properties for ergodic measure preserving systems generally have versions “relative” to an invariant sub σ -algebra (factor algebra). For most cases the fundamental theory for the absolute case lifts to the relative case. For example one can say T is relatively weakly mixing with respect to a factor algebra \mathcal{H} if

- (1) $L^2(\mu)$ has no finite dimensional invariant submodules over the subspace of \mathcal{H} -measurable functions, or
- (2) T has no nontrivial factors containing \mathcal{H} that are relatively isometric over \mathcal{H} , or
- (3) the 2-fold relatively independent self-joining of T is ergodic, or
- (4) the 2-fold relatively independent joining of T with any ergodic action having \mathcal{H} as a factor is ergodic (see [4]).

Received April 25, 2003; received in final form March 10, 2004.
2000 *Mathematics Subject Classification*. Primary 37A05, 37A25.

©2004 University of Illinois

A similar situation holds for the K -property in that the following are all known to be equivalent:

- (1) T has \mathcal{H} -relative trivial tail fields in that for any finite partition P

$$\bigcap_{j=0}^{\infty} \left(\bigvee_{n=j}^{\infty} T^{-n}(P) \vee \mathcal{H} \right) = \mathcal{H}.$$

- (2) T has no \mathcal{H} relative Pinsker algebra, i.e., if $h(T, P|\mathcal{H}) = 0$ then $P \subseteq \mathcal{H}$.
- (3) An appropriately formulated version of \mathcal{H} -relative uniform multiple mixing holds (see [1] and [2]).

A similar situation holds for the Bernoulli property and is the content of Thouvenot’s well known theory of relatively Bernoulli actions (see [3]). In particular one has both an \mathcal{H} -relatively finitely determined and \mathcal{H} -relatively very-weakly Bernoulli characterization of those systems where the factor algebra \mathcal{H} has a complimentary factor on which the action is Bernoulli.

It is not surprising that the situation for mixing is murkier. Studying this is the substance of our work here. What mixing relative to a factor algebra \mathcal{H} should say is that for any f and g

$$|E(f \cdot g \circ T^n|\mathcal{H}) - E(f|\mathcal{H})E(g \circ T^n|\mathcal{H})| \rightarrow 0$$

as $|n| \rightarrow \infty$. What is not clear is in what sense. We will show here that two distinct senses are both reasonable, convergence in L^1 and convergence pointwise a.e. For L^1 -convergence it is natural to take functions in L^2 . For convergence pointwise a.e. this may be true but we cannot show it. Rather we work in a smaller Banach space between L^2 and L^∞ .

DEFINITION 1.1. *Set*

$$L_{\mathcal{H}}^{2,\infty} = \{f : E(|f|^2|\mathcal{H}) \in L^\infty(\mu)\}.$$

Notice that if \mathcal{H} is trivial, this is L^2 and if \mathcal{H} is the full algebra, it is L^∞ .

Setting $\|f\|_{2,\infty}^{\mathcal{H}} = \|E(|f|^2|\mathcal{H})^{1/2}\|_\infty$ we turn $L_{\mathcal{H}}^{2,\infty}$ into a Banach space.

The Rokhlin decomposition of μ over the factor \mathcal{H} says we have a family of fiber measures μ_x with

$$\mu = \int \mu_x d\mu$$

and

$$E(f|\mathcal{H})(x) = \int f d\mu_x.$$

Thus saying $f \in L_{\mathcal{H}}^{2,\infty}$ is just saying $f \in L^2(\mu_x)$ for a.e. x and that the $L^2(\mu_x)$ norms of f are uniformly bounded in x . We will move interchangeably between conditional probabilities ($E(f|\mathcal{H})(x)$) and fiber measures ($\int f d\mu_x$) in our calculations. When writing norms or inner products we will add the measure in the subscript (e.g., $\langle f, g \rangle_{\mu_x}$ or $\|f\|_{2,\mu_x}$) to clarify the measure

involved. To give an example, the invariance of the measure μ and the factor \mathcal{H} means that

$$\langle f, g \rangle_{\mu_x} = \langle f \circ T^{-1}, g \circ T^{-1} \rangle_{\mu_{T(x)}}$$

DEFINITION 1.2. We say a measure preserving action T is L^1 -relatively mixing w.r.t. a factor algebra \mathcal{H} if as $|n| \rightarrow \infty$,

$$\|E(f \cdot g \circ T^n | \mathcal{H}) - E(f | \mathcal{H})E(g \circ T^n | \mathcal{H})\|_1 \rightarrow 0$$

for all f and g in L^2 .

DEFINITION 1.3. We say a measure preserving action T is pointwise-relatively mixing w.r.t. a factor algebra \mathcal{H} if as $|n| \rightarrow \infty$,

$$|E(f \cdot g \circ T^n | \mathcal{H}) - E(f | \mathcal{H})E(g \circ T^n | \mathcal{H})| \rightarrow 0$$

pointwise a.s. for all f and g in $L^2_{\mathcal{H}}{}^{\infty}$.

As a first step let's see that these are different notions. The example is a “ T, T^{-1} ” map where T is mixing. Let S be the shift map on $X = \{-1, 1\}^{\mathbb{Z}}$ with uniform Bernoulli measure. Let T be some mixing action on a space (Y, \mathcal{G}, ν) . For $\vec{x} = \{x_i\}_{i=-\infty}^{\infty} \in X$ and $y \in Y$ set $\widehat{T}(\vec{x}, y) = (S(\vec{x}), T^{x_0}(y))$. Let \mathcal{H} be the first-coordinate algebra.

THEOREM 1.4. The map \widehat{T} is L^1 - but not pointwise-mixing relative to the factor \mathcal{H} .

Proof. In proving this we will see that checking convergence on a dense family of functions f and g is sufficient to show relative L^1 -mixing. Letting $s_n(\vec{x}) = \sum_{j=0}^{n-1} x_j$ we get $\widehat{T}^n(\vec{x}, y) = (S^n(\vec{x}), T^{s_n(\vec{x})}(y))$. Fix $f, g \in L^2$ and $1/4 > \varepsilon > 0$. W.l.o.g. we assume $\|f\|_2 = \|g\|_2 = 1$. First find f^1, \dots, f^k and g^1, \dots, g^k all in $L^2(\nu)$ and sets $A_1, \dots, A_k \subseteq X$ so that

$$\left\| \sum_i \chi_{A_i} \otimes f^i - f \right\|_2 < \varepsilon/10 \quad \text{and} \quad \left\| \sum_i \chi_{A_i} \otimes g^i - g \right\|_2 < \varepsilon/10.$$

Now select N_0 so that for all $|n| \geq N_0$ and all $1 \leq i, j \leq k$,

$$\left| \int f^i \otimes g^j \circ T^n - \int f^i \int g^j \right| < \varepsilon/10.$$

Select N so large that for all $n \geq N$ and $B_n = \{\vec{x} : |s_n(\vec{x})| < N_0\}$, $\mu(B_n) < \varepsilon/10$.

Set $\bar{f} = \sum_i \chi_{A_i} \otimes f^i$ and $\bar{g} = \sum_j \chi_{A_j} \otimes g^j$. Now for $n \geq N$ we can calculate

$$\begin{aligned} & \left\| E(f \cdot g \circ \widehat{T}^n | \mathcal{H}) - E(f | \mathcal{H}) E(g \circ \widehat{T}^n | \mathcal{H}) \right\|_1 \\ & \leq \left\| E(\bar{f} \cdot \bar{g} \circ \widehat{T}^n | \mathcal{H}) - E(\bar{f} | \mathcal{H}) E(\bar{g} \circ \widehat{T}^n | \mathcal{H}) \right\|_1 \\ & \quad + 2 \|f - \bar{f}\|_2 \|\bar{g}\|_2 + 2 \|g - \bar{g}\|_2 \|f\|_2 \\ & < \left\| E(\bar{f} \cdot \bar{g} \circ \widehat{T}^n | \mathcal{H}) - E(\bar{f} | \mathcal{H}) E(\bar{g} \circ \widehat{T}^n | \mathcal{H}) \right\|_1 + \varepsilon/2. \end{aligned}$$

Now

$$\begin{aligned} & \left\| E(\bar{f} \cdot \bar{g} \circ \widehat{T}^n | \mathcal{H}) - E(\bar{f} | \mathcal{H}) E(\bar{g} \circ \widehat{T}^n | \mathcal{H}) \right\|_1 \\ & = \sum_{1 \leq i, j \leq k} \left(\int_{A_i \cap S^{-n}(A_j)} \left| \int \bar{f}^i \cdot \bar{g}^j \circ T^{s_n(\vec{x})} \right. \right. \\ & \quad \left. \left. - \int \bar{f}^i \int \bar{g}^j \circ T^{s_n(\vec{x})} \right| d\mu(\vec{x}) \right) \\ & \leq \sum_{1 \leq i, j \leq k} \left(\int_{A_i \cap S^{-n}(A_j) \cap B_n^c} \left| \int \bar{f}^i \cdot \bar{g}^j \circ T^{s_n(\vec{x})} \right. \right. \\ & \quad \left. \left. - \int \bar{f}^i \int \bar{g}^j \circ T^{s_n(\vec{x})} \right| d\mu(\vec{x}) \right) + \varepsilon/4 \\ & \leq \varepsilon/10 + \varepsilon/4 < \varepsilon/8. \end{aligned}$$

Thus \widehat{T} is L^1 -mixing relative to \mathcal{H} .

On the other hand, for μ -a.e. \vec{x} there exists infinitely many values n_i where $s_{n_i}(\vec{x}) = 0$ as the standard symmetric random walk on \mathbb{Z} is recurrent. At such values n_i ,

$$\begin{aligned} & \left| E(f \cdot g \circ \widehat{T}^{n_i} | \mathcal{H})(\vec{x}) - E(f | \mathcal{H})(\vec{x}) E(g \circ \widehat{T}^{n_i} | \mathcal{H})(\vec{x}) \right| \\ & = |E(f \cdot g | \mathcal{H})(\vec{x}) - E(f | \mathcal{H})(\vec{x}) E(g | \mathcal{H})(\vec{x})|, \end{aligned}$$

which of course need not be zero. \square

We have seen now that there do indeed exist actions which are L^1 -mixing relative to a nontrivial factor algebra. The proof of Theorem 1.4 shows that in fact it is sufficient to check L^1 mixing relative to a factor by checking it on a dense subset of functions and in particular a dense set of piecewise constant functions of the form $\sum_i \chi_{A_i} \otimes f^i$ where the A_i are measurable with respect to \mathcal{H} and the f^i are measurable w.r.t. a second coordinate. Hence one can see that if T is mixing then any $S \times T$ is L^1 -mixing relative to its first coordinate algebra. What is not at all evident at this point is that there exist any actions that are pointwise-mixing relative to a nontrivial factor algebra. We are now

ready to begin our work to show that indeed not only do such actions exist but it is as easy to verify it as it is for the L^1 -notion in that verifying it on a dense family is sufficient.

2. Reducing to just one function

We now show that for both the L^1 and *pointwise* notions it is sufficient to show that functions mix against themselves. Notice that for \mathcal{H} measurable functions, one always has \mathcal{H} -relative mixing. These functions play the role of constants. Also notice that the set of functions on which the appropriate limits hold for relative mixing is closed under linear combinations. Hence it is enough to verify the limits for those f and g whose conditional expectations are zero.

DEFINITION 2.1. Set $E^2 = \{f \in L^2 : E(f|\mathcal{H}) = 0\}$ and $E^{2,\infty} = \{f \in L^{2,\infty}_{\mathcal{H}} : E(f|\mathcal{H}) = 0\}$.

COROLLARY 2.2. A map T is L^1 -mixing relative to \mathcal{H} iff for all $f, g \in E^2$, $E(f \cdot g \circ T^n|\mathcal{H}) \rightarrow 0$ in L^1 and is *pointwise-mixing* relative to \mathcal{H} iff for all $f, g \in E^{2,\infty}$, $E(f \cdot g \circ T^n|\mathcal{H}) \rightarrow 0$ *pointwise a.s.*

DEFINITION 2.3. We say $f \in E_2$ (or $E^{2,\infty}$) is L^1 - (or *pointwise*)-self-mixing relative to a factor \mathcal{H} if as $|n| \rightarrow \infty$ we have $E(f \cdot f \circ T^n|\mathcal{H}) \rightarrow 0$ in L^1 (or *pointwise*).

Our goal now is to show that if all f are self-mixing relative to \mathcal{H} then T is mixing relative to \mathcal{H} either in L^1 or *pointwise*. We then want to see that just checking an L^2 dense family, in either case, is sufficient.

All our work is based on the standard trick that one function cannot be dependent of all terms in a series of independent functions. We start with the simplest lemma and walk through the complete argument to (re)familiarize the reader with it.

THEOREM 2.4. Suppose $f \in E^2$ and

$$\lim_{|n| \rightarrow \infty} \|E(f \cdot f \circ T^n|\mathcal{H})\|_1 = 0.$$

Then for all $g \in L^2$ we have

$$\lim_{|n| \rightarrow \infty} \|E(g \cdot f \circ T^n|\mathcal{H})\|_1 = 0.$$

Proof. Suppose not, i.e., there is a $g \in L^2$ with $n_i \nearrow \infty$ and $1 > a > 0$ for which

$$\|E(g \cdot f \circ T^{n_i}|\mathcal{H})\|_1 > a.$$

W.l.o.g. we can assume that for all $i \neq j$ we have

$$\|E(f \cdot f \circ T^{n_j - n_i} | \mathcal{H})\|_1 < \frac{a^2}{2\|g\|_2^2}$$

by dropping to a subsequence of the n_i .

Let $h_i = \text{signum} E(g \cdot f \circ T^{n_i} | \mathcal{H})$, which is an \mathcal{H} measurable functions. Now for all i ,

$$\int E(h_i \cdot g \cdot f \circ T^{n_i} | \mathcal{H}) > a.$$

Thus

$$\begin{aligned} Ia &< \int E \left(\sum_{i=1}^I h_i \cdot g \cdot f \circ T^{n_i} | \mathcal{H} \right) \\ &\leq \int \left(E \left(\left(\sum_{i=1}^I h_i \cdot f \circ T^{n_i} \right)^2 | \mathcal{H} \right) E(g^2 | \mathcal{H}) \right)^{1/2} \\ &\leq \left(\int E \left(\left(\sum_{i=1}^I h_i \cdot f \circ T^{n_i} \right)^2 | \mathcal{H} \right) \right)^{1/2} \left(\int E(g^2 | \mathcal{H}) \right)^{1/2} \\ &= \left(\sum_{1 \leq i, j \leq I} \int (h_i \circ T^{-n_i} h_j \circ T^{n_i}) E(f \cdot f \circ T^{n_j - n_i} | \mathcal{H}) \right)^{1/2} \|g\|_2 \\ &\leq \left(\sum_{1 \leq i, j \leq I} \|E(f \cdot f \circ T^{n_j - n_i} | \mathcal{H})\|_1 \right)^{1/2} \|g\|_2 \\ &\leq \left(I\|f\|_2^2 + I^2 \frac{a^2}{2\|g\|_2^2} \right)^{1/2} \|g\|_2 \leq \sqrt{I}\|f\|_2\|g\|_2 + \frac{Ia}{2}. \end{aligned}$$

Once I is large enough, $Ia \leq \sqrt{I}\|f\|_2\|g\|_2 + Ia/2$ cannot hold. \square

COROLLARY 2.5. *For a measure preserving map T , if there is a dense family of functions f in L^2 which are L^1 -self-mixing relative to a factor algebra \mathcal{H} then T is L^1 -mixing relative to \mathcal{H} .*

We now obtain a result parallel to Theorem 2.4 for pointwise-mixing relative to a factor.

THEOREM 2.6. *Suppose $f \in E^{2,\infty}$ and $E(f \cdot f \circ T^n | \mathcal{H}) \rightarrow 0$ pointwise a.e. Then for all $g \in L_{\mathcal{H}}^{2,\infty}$ we have $E(g \cdot f \circ T^n | \mathcal{H}) \rightarrow 0$ pointwise a.e.*

Proof. Following the pattern set above, suppose this is not true. That is to say, there is a $g \in L_{\mathcal{H}}^{2,\infty}$, an $a > 0$, and a set $B \subset X$ with $\mu(B) > a$ and for

$x \in B$ there are $n_i = n_i(x) \nearrow \infty$ with

$$|E(g \cdot f \circ T^{n_i} | \mathcal{H})(x)| > a.$$

W.l.o.g. we can assume $E(g \cdot f \circ T^{n_i} | \mathcal{H})(x) > a$ by taking a subset of B where this value is of one sign infinitely often, taking this subsequence and replacing g with $-g$ if necessary.

For a.e. $x \in B$ we can write

$$E(g \cdot f \circ T^n | \mathcal{H})(x) = \int g \cdot f \circ T^n d\mu_x$$

where μ_x is the fiber measure of the factor \mathcal{H} at the point x . Refine the sequence $n_i(x)$ inductively for a.e. x so that the successive gaps are large enough that for any $|m| \geq n_{i+1}(x) - n_i(x)$ we will have

$$E(f \cdot f \circ T^m | \mathcal{H})(T^{n_i}(x)) = \int f \cdot f \circ T^m d\mu_{T^{n_i}(x)} < \frac{a^2}{2(\|g\|_{2,\infty}^{\mathcal{H}})^2}.$$

Once more following the template of the previous theorem, for a.e. $x \in B$ we calculate

$$\begin{aligned} Ia &< E\left(\sum_{i=1}^I g \cdot f \circ T^{n_i} | \mathcal{H}\right)(x) \\ &\leq E\left(\left(\left(\sum_{i=1}^I f \circ T^{n_i}\right)^2 | \mathcal{H}\right)(x)\right)^{1/2} \|g\|_{2,\mu_x} \\ &\leq \left(\sum_{i=1}^I \|f\|_{2,\mu_{T^{n_i}(x)}}^2 + \sum_{i \neq j} E(f \circ T^{n_i} f \circ T^{n_j} | \mathcal{H})(x)\right)^{1/2} \|g\|_{2,\infty}^{\mathcal{H}} \\ &\leq \left(I(\|f\|_{2,\infty}^{\mathcal{H}})^2 + \sum_{i \neq j} \int f f \circ T^{n_j - n_i} d\mu_{T^{n_i}(x)}\right)^{1/2} \|g\|_{2,\infty}^{\mathcal{H}} \\ &\leq \sqrt{I} \|f\|_{2,\infty}^{\mathcal{H}} \|g\|_{2,\infty}^{\mathcal{H}} + \frac{Ia}{2}. \end{aligned}$$

Once I is large enough, this cannot hold. □

COROLLARY 2.7. *If for all $f \in E^{2,\infty}$, f is pointwise-self-mixing relative to the factor algebra \mathcal{H} then T is pointwise-mixing relative to \mathcal{H} .*

3. Pointwise self-mixing on an L^2 dense family is enough

To demonstrate pointwise-mixing relative to a factor algebra \mathcal{H} it is enough to check pointwise-self-mixing for functions in the $L^2_{\mathcal{H}}$ unit ball \mathcal{B}_1 of $E^{2,\infty}$. This unit ball is closed in L^2 as the L^2 limit of a sequence of uniformly bounded functions will possess the same bound a.s. Our goal now is to show

that the subset of functions in \mathcal{B}_1 which are pointwise-self-mixing are closed in L^2 . Hence to demonstrate pointwise-mixing relative to \mathcal{H} it will be sufficient to prove pointwise-self-mixing on an L^2 -dense subset of \mathcal{B}_1 .

As we continue we assume that $f \in \mathcal{B}_1 \subseteq E^{2,\infty}$ is fixed. We investigate the failure of f to be pointwise-self-mixing relative to \mathcal{H} .

DEFINITION 3.1. *We say a point $x \in X$ is a -bad if there are $n_i = n_i(x)$, $|n_i| \nearrow \infty$ with*

$$|E(f \cdot f \circ T^{n_i} | \mathcal{H})(x)| = |\langle f, f \circ T^{n_i} \rangle_{\mu_x}| > a.$$

To say that f is not pointwise-self-mixing relative to \mathcal{H} is to say that for some $a > 0$ the set of a -bad points has positive measure. Notice that for fixed a the values $n_i(x)$ can be chosen measurably.

DEFINITION 3.2. *We say a set $B \subseteq X$, $\mu(B) > 0$ is a -very bad if for a.e. $x \in B$ there are $n_i = n_i(x)$, $|n_i| \nearrow \infty$ with*

- (i) $T^{n_i}(x) \in B$ and
- (ii) $|\langle f, f \circ T^{n_i} \rangle_{\mu_x}| > a$.

Note that on an a -very bad set B the values $n_i(x)$ can be chosen measurably.

DEFINITION 3.3. *We say a set $B \subseteq X$, $\mu(B) > 0$ is a -terrible if any subset $B' \subseteq B$ of positive measure is a -very bad.*

Our goal is the following result:

THEOREM 3.4. *If $f \in \mathcal{B}_1 \subseteq E^{2,\infty}$ is not pointwise-self-mixing relative to a factor algebra \mathcal{H} then for some $a > 0$ there is an a -terrible set $B \subseteq X$, $\mu(B) > 0$.*

Before completing the proof of Theorem 3.4 we show how this implies that the pointwise-self-mixing functions in \mathcal{B}_1 are L^2 -closed.

COROLLARY 3.5. *Suppose $f_i \in \mathcal{B}_1 \subseteq E^{2,\infty}$ and each f_i is pointwise-self-mixing relative to \mathcal{H} . Moreover suppose $f_i \rightarrow f$ in L^2 . Then f is also pointwise-self-mixing relative to \mathcal{H} .*

Proof. As all the $f_i \in \mathcal{B}_1$ so is f . Now suppose f is not pointwise-self-mixing. Then for some value $a > 0$ there is an a -terrible set B of positive measure. Choose i so large that

$$\mu(B \cap \{x : \|f - f_i\|_{2,\mu_x} < a/4\}) > 0.$$

Now set $B' = B \cap \{x : \|f - f_i\|_{2,\mu_x} < a/4\}$ and B' will be an a -very bad set. Thus for a.e. $x \in B'$ there are $n_j = n_j(x)$ with $|n_j| \nearrow \infty$ and both

- (i) $T^{n_j(x)}(x) \in B'$ and
- (ii) $\langle f, f \circ T^{n_j} \rangle_{\mu_x} > a$.

But now for a.e. $x \in B'$ we obtain the following conflict.

$$\begin{aligned} a &\leq \limsup_j |\langle f, f \circ T^{n_j} \rangle_{\mu_x} - \langle f_i, f_i \circ T^{n_j} \rangle_{\mu_x}| \\ &= \limsup_j |\langle f - f_i, f \circ T^{n_j} \rangle_{\mu_x} + \langle f_i, (f - f_i) \circ T^{n_j} \rangle_{\mu_x}| \\ &\leq \limsup_j (\|f - f_i\|_{2, \mu_x} \|f\|_{2, \mu_{T^{n_j}(x)}} + \|f - f_i\|_{2, \mu_{T^{n_j}(x)}} \|f_i\|_{2, \mu_x}) \\ &\leq a/2 \quad \text{as both } x \text{ and } T^{n_j}(x) \text{ are in } B'. \end{aligned} \quad \square$$

We now set about proving Theorem 3.4. At the core of this argument is the same basic trick we have used twice before. We begin with some definitions.

DEFINITION 3.6. We say a set G is c -good if for all $x \in G$,

$$\{n \neq 0 : T^n(x) \in G \text{ and } |\langle f, f \circ T^n \rangle_{\mu_x}| \geq c\}$$

is a finite set. We say G is c -very good if for all $x \in G$ this set of integers is empty.

LEMMA 3.7. Any subset of a c -good set is c -good and any c -good set can be partitioned into a countable collection of c -very good sets.

Proof. The first statement is clear. As

$$\{n \neq 0 : T^n(x) \in G \text{ and } |\langle f, f \circ T^n \rangle_{\mu_x}| \geq c\}$$

is finite for $x \in G$ it is bounded. Partition G first according to this bound b , writing $G = \bigcup_{b=1}^\infty G_b$. Now, using the Rokhlin Lemma, one can partition each G_b into subsets which never recur to themselves in time $\leq b$. This provides the desired partition. \square

LEMMA 3.8. If X contains no c -terrible set then X can be partitioned into a countable collection of c -good sets and hence into a countable collection of c -very good sets.

Proof. If a set A is **not** very bad then the set

$$A' = \{x \in A : \#\{n \neq 0 : T^n(x) \in A \text{ and } |\langle f, f \circ T^n \rangle_{\mu_x}| \geq c\} < \infty\}$$

has positive measure. It follows that the set A' must be c -good. Thus if X has no c -terrible subsets, any subset of X of positive measure must contain a c -good subset of positive measure. One now obtains the desired partition by exhaustion. \square

We now make another application of the basic trick.

LEMMA 3.9. Suppose G is $a^2/4$ -very good. Then for a.e. $x \in X$,

$$\{n : T^n(x) \in G \text{ and } |\langle f, f \circ T^n \rangle_{\mu_x}| > a\}$$

has cardinality at most $8/a^2$.

Proof. Suppose $T^{n_1}(x), \dots, T^{n_I}(x) \in G$ and $|\langle f, f \circ T^{n_i} \rangle_{\mu_x}| > a$ for all i . Assume in addition that all the values $\langle f, f \circ T^{n_i} \rangle_{\mu_x}$ are of the same sign. We compute

$$\begin{aligned} Ia &< \left| \langle f, \sum_{i=1}^I f \circ T^{n_i} \rangle_{\mu_x} \right| \\ &\leq \|f\|_{2, \mu_x} \left(\sum_{i=1}^I \|f\|_{2, \mu_{T^{n_i}(x)}}^2 + \sum_{i \neq j} \langle f \circ T^{n_i}, f \circ T^{n_j} \rangle_{\mu_x} \right)^{1/2} \\ &\leq \left(I + I^2 \frac{a^2}{4} \right)^{1/2} \leq \sqrt{I} + I \frac{a}{2}. \end{aligned}$$

This cannot hold once $I > 4/a^2$. Among any collection of $8/a^2$ values n_i a collection of half of them must be of the same sign. \square

The next proposition will complete the proof of Theorem 3.4.

PROPOSITION 3.10. *Let T be an ergodic action, \mathcal{H} an invariant factor algebra and $f \in \mathcal{B}_1 \subseteq E^{2, \infty}$. If X contains no $a^2/4$ -terrible set of positive measure then the set of a -bad points in X has measure zero.*

Proof. As X has no $a^2/4$ terrible sets, we can partition X into G_1, G_2, \dots where each G_i is an $a^2/4$ -very good set. Now suppose that x is an a -bad point, i.e., there are $n_i = n_i(x)$ with $|n_i| \nearrow \infty$ and $|\langle f, f \circ T^{n_i} \rangle_{\mu_x}| > a$. Lemma 3.9 tells us that at most $8/a^2$ of the $T^{n_i}(x)$ can belong to any particular G_j . We conclude that for each a -bad point x the orbit points $T^{n_i(x)}(x)$ must become ever more concentrated in the tail of the sequence of sets G_j .

To continue we argue by contradiction. That is to say assume B is a set of a -bad points with $\mu(B) > 0$ and w.l.o.g. we will assume that $B \subseteq G_1$. (The set of a -bad points must intersect some G_j and we can relabel it to be the first and restrict to the intersection.)

Select a value N so that

$$\mu \left(\bigcup_{j=N}^{\infty} G_j \right) < \frac{a^2}{8} \mu(B).$$

We know that for each $x \in B$ there must be an $n(x)$ with

$$|\langle f, f \circ T^{n(x)} \rangle_{\mu_x}| > a \text{ and } T^{n(x)} \in \bigcup_{j=N}^{\infty} G_j.$$

The value $n(x)$ can be chosen measurably and so we can partition B as $\bigcup_n B_n$ where $B_n = \{x \in B : n(x) = n\}$.

Now note that for each B_n the map $T^n : B_n \rightarrow \bigcup_{j=N}^\infty G_j$ is 1-1 and measure preserving. As $\mu(B) > (8/a^2)\mu(\bigcup_{j=N}^\infty G_j)$, there must be $L > 8/a^2$ points $x_1, x_2, \dots, x_L \in B \subseteq G_1$ where all the $T^{n(x_i)}(x_i)$ are identical, i.e., equal to some $x_0 \in \bigcup_{j=N}^\infty G_j$. But this then gives us one point x_0 with $L > 8/a^2$ images $x_i = T^{-n(x_i)}(x_0)$ all in the same G_1 and with

$$|\langle f, f \circ T^{-n(x_i)} \rangle_{\mu_{x_0}}| > a,$$

which we saw was impossible. □

4. Conclusions

We can now discuss the general class of cocycle extensions, of which T, T^{-1} maps are an example. We begin with the standard description. Let S and T be measure preserving and ergodic transformations of (X, \mathcal{F}, μ) and (Y, \mathcal{G}, ν) , respectively. For any function $n : X \rightarrow \mathbb{Z}$ we can construct the cocycle extension $S_n : X \times Y \rightarrow X \times Y$ by setting $S_n(x, y) = (S(x), T^{n(x)}(y))$. Both the T, T^{-1} maps and direct products are of this form.

Suppose now that T is mixing and let \mathcal{H} be the first coordinate algebra \mathcal{F} . We want to understand in generality when S_n is L^1 - and pointwise-relatively mixing w.r.t. \mathcal{H} . We can so long as n is integrable.

As is usual, let $n(i, x)$ be the cocycle generated by n so that $S_n^j(x, y) = (S^j(x), T^{n(i,x)}(y))$. Whether or not S_n is relatively mixing now depends on the distribution of the functions $n(i, x)$ for i large.

DEFINITION 4.1. *We say the cocycle $n(i, x)$ “spreads in the mean” if for all $N > 0$*

$$\lim_{i \rightarrow \infty} \mu(\{x : |n(i, x)| < N\}) = 0.$$

We say the cocycle $n(i, x)$ “spreads pointwise” if

$$\lim_{i \rightarrow \infty} |n(i, x)| = \infty \text{ a.e.}$$

The proof that T, T^{-1} actions are L^1 -mixing relative to \mathcal{H} applies to any S_n where the cocycle spreads in the mean. We now know that to check pointwise-relative mixing it is sufficient to check an L^2 dense family in \mathcal{B}_1 , in particular to check it on functions of the form $\sum_i \chi_{A_i} \otimes f_i$ as are used in the discussion of T, T^{-1} -maps. We can thus apply that argument to show that if the cocycle n spreads pointwise then S_n is pointwise-relative mixing w.r.t. the first coordinate algebra. We now pull all this together.

PROPOSITION 4.2. *Suppose S and T are as above. Suppose $n : X \rightarrow \mathbb{Z}$ is in $L^1(\mu)$.*

- (a) *If $\int n d\mu \neq 0$ then S_n is pointwise-relative mixing with respect to its first coordinate algebra. In particular $S \times T$ is pointwise-relative mixing with respect to the first coordinate algebra.*

- (b) *If $\int n d\mu = 0$ and the cocycle generated by n spreads in the mean then S_n is L^1 - but not pointwise-relative mixing w.r.t. its first coordinate algebra.*
- (c) *If n generates a cocycle that does not spread in the mean then S_n is not L^1 -relatively mixing w.r.t. its first coordinate algebra.*

Proof. Statement (a) follows from the ergodic theorem in that if n is not of mean zero then the cocycle it generates will spread pointwise. Statement (b) is just a general version of the T, T^{-1} argument given in the introduction once one knows that any mean zero n will generate a recurrent cocycle. The proof of (c) is quite simple and we leave it to the reader. \square

We leave it now to the reader to explore more general cocycle extensions. Certainly there is no problem applying these methods to extensions by larger abelian groups of mixing actions.

A significant motivation for understanding properties relative to an invariant factor \mathcal{H} is to use the orbit transference method of [2] to lift results from actions of \mathbb{Z} to general discrete amenable actions. For this method to work whatever properties of an action are under discussion must have relative versions and these relative versions must be invariant under \mathcal{H} -measurable orbit equivalences. To clarify the picture we state a result that one would wish to prove for general amenable actions. Is it the case that a weakly mixing isometric extension of a mixing action must be mixing? This is known to be true for \mathbb{Z} actions. Moreover the \mathbb{Z} proof does not seem to lift to the amenable case. To apply orbit transference to this question one would need to know a relativized version of this result but only for \mathbb{Z} actions. One must ask whether an \mathcal{H} -relatively weakly mixing \mathcal{H} -relatively isometric extension of an \mathcal{H} -relatively mixing action remains \mathcal{H} -relatively mixing. We have explicitly avoided indicating which type of relative mixing we mean. This question is meaningful for both L^1 - and pointwise-relative mixing. Only the pointwise version though would be useful for the orbit transference method. We will show elsewhere that in fact L^1 -relative mixing w.r.t. a factor \mathcal{H} is not invariant under \mathcal{H} -measurable orbit equivalences. The pointwise notion obviously is. Hence it is necessary to verify the above relativized result on isometric extensions for pointwise-relative mixing. Again to apply the transference method one must know that if T has an absolute property, then $S \times T$ has the relative property w.r.t. the first coordinate algebra. We have settled this piece of the problem here. It appears the above relativized result can be proven for pointwise-mixing relative to \mathcal{H} , giving all the pieces to settle the problem. This work will appear elsewhere if it is in fact correct.

The author wishes to thank Kyewon Park, Mariusz Lemanczyk and Jean-Paul Thouvenot for very helpful conversations on the issues discussed here.

REFERENCES

- [1] M. Rahe, *Relatively finitely determined implies relatively very weak Bernoulli*, Canad. J. Math. **30** (1978), 531–548. MR **81j**:28029
- [2] D. J. Rudolph and B. Weiss, *Entropy and mixing for amenable group actions*, Ann. of Math. (2) **151** (2000), 1119–1150. MR **2001g**:37001
- [3] J.-P. Thouvenot, *Quelques propriétés des systèmes dynamiques qui se décomposent en un produit de deux systèmes dont l'un est un schéma de Bernoulli*, Israel J. Math. **21** (1975), 177–207. MR 53 #3263
- [4] R. J. Zimmer, *Ergodic actions with generalized discrete spectrum*, Illinois J. Math. **20** (1976), 555–588. MR 54 #2924

DEPT. OF MATHEMATICS, UNIV. OF MARYLAND, COLLEGE PARK, MD 20742, USA
E-mail address: djr@math.umd.edu