ISRAEL JOURNAL OF MATHEMATICS **158** (2007), 41–63
DOI: 10.1007/s11856-007-0003-0

AMENABLE GROUPS, STATIONARY MEASURES AND PARTITIONS WITH INDEPENDENT ITERATES

BY

Boris Begun*

Institute of Mathematics, Hebrew University
91904 Jerusalem, Israel
e-mail: begun@math.huji.ac.il

AND

Andrés del Junco

Department of Mathematics, University of Toronto
Toronto M5S 2E4, Canada
e-mail: deljunco@math.toronto.edu

ABSTRACT

Suppose a discrete amenable group G acts freely on a probability space $(X, \mathcal{B}, \mu)$ and $\{g_i\}$ is any mixing sequence of group elements, that is $\mu(g_i^{-1}A \cap B) \to \mu(A)\mu(B)$ for all $A, B \in \mathcal{B}$. Then given any finite partition P and $\epsilon > 0$ there is a subsequence $\{h_j\}$ of $\{g_i\}$ and a partition P' differing from P on a set of measure less than ϵ such that the partitions $\{gP : g \in \mathrm{IP}'\{h_j\}\}$ are jointly independent, where $\mathrm{IP}'\{h_j\}$ denotes the set

$$\{e_G\} \cup \{h_{j_k} h_{j_{k-1}} \cdots h_{j_1} : j_1 < j_2 < \cdots < j_k\}$$

consisting of the identity of G together with all finite products of the $\{h_j\}$ taken with indices in decreasing order.

Corollary: *Suppose $\{T_i\}$ is a mixing sequence of pairwise commuting automorphisms of $(X, \mathcal{B}, \mu)$ and P is a finite partition. Then for any $\epsilon > 0$ there are a $Q \overset{\epsilon}{\sim} P$ and a subsequence $\{T_{i_j}\}$ such that the partitions $\{TQ : T \in \mathrm{IP}'\{T_{i_j}\}\}$ are (jointly) independent.*

The main result depends on an extension result for certain stationary stochastic processes indexed by a finite subset of G which is of independent interest. It says roughly that if K is a finite subset of G and

* The Research was conducted while the first author was a postdoctoral fellow at the University of Toronto. He thanks the University for its hospitality.
Received April 8, 2005

$\{X_g\}_{g\in K}$ is a stochastic process which is (right) stationary and sufficiently close (depending on the size of K) to being i.i.d., then $\{X_g\}_{g\in K}$ has a stationary extension $\{X_g\}_{g\in G}$ indexed by all of G.

We also give characterizations of weak, mild and strong mixing of the action of a discrete amenable group G in terms of the density in the space of all finite partitions of the class of finite partitions with independent iterates under appropriate types of subsets of G.

1. Introduction

In [K] Krengel proved that if T is a weakly mixing measure-preserving automorphism of a probability space $(X, \mathcal{B}, \mu)$, P is a finite partition of X and $\epsilon > 0$, then there is an infinite sequence $0 = n_1 < n_2 < \cdots$ and a partition P' within ϵ of P (in the usual partition metric) such that the iterates $\{T^{n_i}P'\}_{i\in\mathbb{N}}$ form a jointly independent sequence of partitions, and that this property in fact characterizes weak mixing of T. Krengel actually used the assumption of two-fold weak mixing but shortly thereafter Furstenberg [F1] showed that weak mixing implies weak mixing of all orders.

In [J,R,W] Krengel's result was extended in two directions. Firstly, it was generalized to actions of certain discrete Abelian groups G, still using the assumption of two-fold weak mixing (which follows from weak mixing for some groups, e.g. $\mathbb{Z}^d$). The proof of this result more or less followed Krengel's approach. Secondly, in the case $G = \mathbb{Z}$, that is the case of a single automorphism T, it was shown that if the sequence $\{n_i\}$ is mixing for T, that is $\mu(T^{-n_i}A \cap B) \to \mu(A)\mu(B)$ for all measurable A and B, then for any finite partition P and $\epsilon > 0$ there are a P' within ϵ of P and a subsequence $\{n_{i_j}\}$ such that the iterates $\{T^{n_{i_j}}P'\}$ are jointly independent. (In fact this second result was proved in a stronger "IP-version" which we will explain later.)

For the second result two-fold mixing is not available: to our knowledge it is not known whether

$$\mu(A \cap T^{-n_i}B) \to \mu(A)\mu(B) \quad \forall A, B$$

implies

$$\mu(A \cap T^{-n_i}B \cap T^{-2n_i}C) \to \mu(A)\mu(B)\mu(C) \quad \forall A, B, C.$$

Of course when the sequence $\{n_i\}$ is just $\mathbb{N}$ itself this is a special case of a notorious open question posed by Rohlin more than fifty years ago. Instead of two-fold mixing [J,R,W] used an extension result for approximately independent stationary stochastic processes indexed by finite subsets of $\mathbb{Z}$. This sort of extension result is of independent interest and provides a very natural approach to the problem. The extension result in [J,R,W] was obtained only for $G = \mathbb{Z}$ and was also limited in another way which we will explain later. In this paper we remove both these limitations, proving a general extension result, Theorem 2 below, for approximately independent stationary stochastic processes indexed by a finite subset of a countable discrete amenable group G. This extension result together with the Rohlin lemma for discrete amenable groups [O,W] allows us to obtain a general form of Krengel's theorem for free actions of a discrete amenable group, Theorem 1 below.

Throughout this paper G will denote a countable discrete group with identity element e_G and $(X, \mathcal{B}, \mu)$ a Lebesgue probability space. $\mathcal{A}ut(X, \mathcal{B}, \mu)$ will denote the group of measure-preserving automorphisms of $(X, \mathcal{B}, \mu)$. An action of G on $(X, \mathcal{B}, \mu)$ is a homomorphism $T : G \to \mathcal{A}ut(X, \mathcal{B}, \mu)$. We will write T_g for $T(g)$ and gx for $T_g x$. The action is called **free** if for all $g \in G \backslash \{e_G\}$ we have $\mu\{x : gx = x\} = 0$. A sequence $\{g_i\} \subset G$ is called **mixing** (for the given action) if $\mu(g_i^{-1} A \cap B) \to \mu(A)\mu(B)$ for all measurable A and B. By a **finite partition** of X we mean a measurable mapping P from X to a finite index set A. If Q: $X \to A$ is another partition we write $Q \overset{\epsilon}{\sim} P$ if $\mu\{x : Px \neq Qx\} < \epsilon$.

To state Theorem 1 we need the notion of an IP-set: if $\{g_i\}$ is any sequence of elements of G then we let

$$\mathrm{IP}\{g_i\} = \{g_{i_k} g_{i_{k-1}} \cdots g_{i_1} : i_1 < \cdots < i_k\}$$

and

$$\mathrm{IP}'\{g_i\} = \{e_G\} \cup \mathrm{IP}\{g_i\}.$$

We call a set of this form the IP-set or IP$'$-set generated by $\{g_i\}$. In the non-Abelian setting one can define either left or right IP-sets and the definition we have given is of a left IP-set. This is the natural choice for us as our actions are left actions.

THEOREM 1: *Suppose a discrete amenable group G acts freely on $(X, \mathcal{B}, \mu)$ and $\{g_i\}$ is a sequence of elements of G which is mixing for this action. Then for any finite partition P of X and $\epsilon > 0$ there are a partition $Q \overset{\epsilon}{\sim} P$ and a subsequence $\{g_{i_j}\}$ such that the partitions $\{gP : g \in \mathrm{IP}'\{g_{i_j}\}\}$ are (jointly) independent.*

The extension theorem which is the basis of our proof of Theorem 1 is the following. We defer precise definitions of the terminology for the moment.

THEOREM 2: *Given $\alpha > 0$ and $k \in \mathbb{N}$ there is a $\delta = \delta(\alpha, k) > 0$ such that for any finite set A, any countable discrete amenable group G, any $H \subset G$ with $|H| \leq k$ and any stationary probability measure γ on A^H the following holds: if γ is δ-independent and its one-dimensional marginal p satisfies $\min_{a \in A} p(a) > \alpha$ then γ extends to a stationary probability measure λ on A^G, that is the projection of γ on A^H is λ.*

For now, suffice it to say that "stationary" means, roughly, invariant under the action of G by right translation on A^G, in so far as this is meaningful, that δ-independence of γ means that γ is δ-close to the product of its marginals and that the precise meaning of "δ-close" is unimportant as the cardinality of A^H is bounded by α^{-k}. In fact we shall prove a more general result, Theorem 2′ in §2, which applies to any countable discrete G and which easily implies Theorem 2.

In [J,R,W] Theorem 2 was proved in the case $G = \mathbb{Z}$ and with the stronger assumption that the diameter of H, rather than just its cardinality, is bounded by k.

Theorem 2 is of independent interest and has applications other than our use of it here. It is an interesting and probably hard problem to say anything about which stationary measures have stationary extensions. Stationarity alone is not sufficient for Theorem 2, even when $G = \mathbb{Z}$ and $A = \{0, 1\}$. As an example take γ on $\{0, 1\}^{\{0,1,3\}}$ to be the distribution of random variables X_0, X_1, X_3 such that $\operatorname{dist} X_i = (\frac{1}{2}, \frac{1}{2})$ for each i, $X_1 = X_0$ and X_3 is independent of X_0. If it were possible to interpolate X_2 in a stationary way then we would have $X_0 = X_1 = X_2 = X_3$, contradicting the independence of X_0 and X_3. The rigidity of exact equality in this example may be relaxed somewhat by having X_1 equal to X_0 only with probability $1 - \epsilon$. Of course in case H is an interval in $\mathbb{Z}$ every stationary measure is extendable, via the well-known Markov extension.

When $G = \mathbb{Z}^2$, Theorem 2 fails for stationary measures in general, even when $H = \{0, 1\}^2$. Take γ to be the uniform measure on the four configurations

$$\begin{matrix} 1 & 1 \\ 0 & 1 \end{matrix} \qquad \begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix} \qquad \begin{matrix} 0 & 0 \\ 1 & 1 \end{matrix} \qquad \begin{matrix} 1 & 0 \\ 1 & 0 \end{matrix}$$

It is straightforward to check that γ is stationary. However, the subshift of finite type defined by allowing only these four arrays to appear in an infinite configuration is empty, indeed there are no 3×3 arrays compatible with the four given arrays, so certainly γ can have no stationary extension. Again the rigidity can be loosened a bit, for example by taking a convex combination of γ with a stationary product measure.

In light of Theorem 1 it is natural to ask the following question. Suppose $\{T_i\}$ is a mixing sequence of automorphisms of $(X, \mathcal{B}, \mu)$, P is a finite partition and $\epsilon > 0$. Can one then find a $Q \overset{\epsilon}{\sim} P$ and a subsequence $\{T_{i_j}\}$ such that the partitions $\{TQ : T \in \mathrm{IP}'\{T_{i_j}\}\}$ are (jointly) independent? It is tempting to conjecture that the answer is yes in complete generality. Theorem 1 tells us that if the subgroup G of the automorphism group of $(X, \mathcal{B}, \mu)$ generated by $\{T_i\}$ is amenable and acts freely on X then the answer is yes. If the group generated is not amenable then the techniques of this paper are of no use. It is less clear whether the lack of freeness is a serious obstacle. In the case of an abelian group G freeness of the action is automatic by ergodicity, which follows from the existence of a mixing sequence, so the following result is a corollary of Theorem 1.

Theorem 3: *Suppose $\{T_i\}$ is a mixing sequence of pairwise commuting automorphisms of $(X, \mathcal{B}, \mu)$ and P is a finite partition. Then for any $\epsilon > 0$ there are a $Q \overset{\epsilon}{\sim} P$ and a subsequence $\{T_{i_j}\}$ such that the partitions $\{TQ : T \in \mathrm{IP}'\{T_{i_j}\}\}$ are (jointly) independent.*

Using Theorem 1 we can also prove the following characterizations of weak, mild and strong mixing, generalizing results of [J,R,W]. In the following theorems G is always a countable discrete amenable group, $T\colon G \to \mathcal{A}ut(X, \mathcal{B}, \mu)$ is a free action and $(X, \mathcal{B}, \mu)$ is assumed to be a separable measure space. For $E \subset G$ we will say P is independent along E if the iterates $\{gP\}_{g \in E}$ are jointly independent.

Theorem 4: *T is weakly mixing if and only if given any set $\Sigma \subset G$ which has positive density with respect to some Følner sequence, the partitions which are independent along some* IP′*-set with generators in Σ are dense in the set of all finite partitions.*

Theorem 5: *T is mildly mixing if and only if given any* IP′*-set $\Sigma \subset G$, the collection of finite partitions which are independent along some* IP′*-set $\bar{\Sigma} \subset \Sigma$ is dense in the set of all finite partitions.*

Theorem 6: *T is mixing if and only if given any infinite set $\Sigma \subset G$, the partitions which are independent along some* IP′*-set with generators in Σ are dense in the set of all finite partitions.*

[J,R,W] also contains a relativized version of Krengel's theorem, relative to the Kronecker factor of the automorphism T. We have not worked this out for a general G but very likely there is a result of this sort.

Finally, let us mention a result of Bergelson and Rudolph [Be,Ru] which is related to our main result.

THEOREM (Bergelson, Rudolph): *Suppose $G = \bigoplus_{n=1}^{\infty} F$ is a countable direct sum of copies of a finite field F and $g \mapsto T_g$ is any weakly mixing action of G. Then G has a subgroup H, isomorphic to G, whose action is Bernoulli, in the sense that there is a σ-algebra $\mathcal{F} \subset \mathcal{B}$ such that the σ-algebras $\{g\mathcal{F}\}_{g \in H}$ are jointly independent and generate $\mathcal{B}$ up to null sets.*

In this setting it is easy to construct partitions with independent iterates because the group in question is locally finite, an increasing union of finite subgroups. For finite subgroups one can obtain exact Rohlin towers and thereby control the behaviour of partitions along these subgroups exactly. The more difficult aspect of the Bergelson–Rudolph result is arranging for generation. Very likely something like the Bergelson–Rudolph result is true for any locally finite group G, although H would be simply an infinite subgroup of G, no longer isomorphic to G. In general, such a result cannot hold. For example, for $G = \mathbb{Z}$, if the action of a subgroup is Bernoulli then the full action is Bernoulli (see [O], p. 39). However, one can ask whether weaker results hold. For example, if the action is weakly mixing can one find an infinite partition or σ-algebra with infinitely many independent iterates? Can this be arranged with iterates in an IP-set? Can one arrange for these iterates to generate?

2. Extension of stationary measures

A **finite partition** of a measurable space $(X, \mathcal{B})$ is a measurable map P from X to a finite index set A. The sets $p = P^{-1}(a)$ are the **atoms** of P and we write $p \in P$. If $(X, \mathcal{B})$ carries a probability measure μ then $\operatorname{dist} P$ (or $\operatorname{dist}_\mu P$ if the measure needs to be emphasized) denotes the measure $\mu \circ P^{-1}$ on A. If $B \in \mathcal{B}$ then $\operatorname{dist}_\mu(P|B)$ refers to the restriction $P|_B$ and the normalized measure μ_B. We write $P \subset Q$ if each atom of P is a union of atoms of Q.

If P and Q are finite partitions of a probability space (Y, ν) then we will say P is **δ-independent** of Q whenever $\|\operatorname{dist}(P|q) - \operatorname{dist} P\|_\infty < \delta$ for all $q \in Q$. (We are identifying a measure μ on a finite set A with the vector $\{\mu(a)\}_{a \in A}$.) When P is δ-independent of Q we will write $P \perp_\delta Q$ (or $P \perp_{\delta,\nu} Q$). Note that this definition differs from a more standard one which uses the l_1-norm and allows a small exceptional set of q's. We will make use of the fact that if $P \perp_\delta Q$ and $Q \supset R$ then $P \perp_\delta R$.

Let G be a countable discrete group. A (right) **Følner sequence** in G is a sequence $\{F_n\}$ of finite subsets of G such that $\lim_n |F_n|^{-1}|F_n g \Delta F_n| = 0$ for each $g \in G$. G is said to be **amenable** if it has a Følner sequence. We recall that if G is amenable then it also has a left Følner sequence.

Now suppose A is a finite alphabet and G is a countable discrete group. We will be working with measures on A^H for various subsets $H \subset G$ and we use the usual product Borel structure on A^H. Mostly H will be finite. Whenever $I \subset H \subset G$ we will denote by P^I the projection map from A^H to A^I, which is a finite partition of A^H when I is finite. For $H \subset G$ and $g \in G$ we define the right translation $\mathcal{R}_g$: $A^H \to A^{Hg}$ by $\mathcal{R}_g(x)(hg) = x(h)$ for $h \in H$. We will also view $\mathcal{R}_g$ as mapping measures on A^H to measures on A^{Hg} in the usual way. If γ is a measure on A^H and $I \subset H$ we write γ_I for the projection of γ onto A^I. If p is a probability measure on A then p^H will denote the product measure on A^H.

If H and H' are subsets of G we will say a probability measure γ on A^H is **consistent** with γ' on $A^{H'}$ if the projections of γ and γ' onto $A^{H \cap H'}$ coincide. We say γ is **compatible** with γ' if γ is consistent with $\mathcal{R}_g\gamma'$ for all $g \in G$ and we say γ on A^H is **stationary** if γ is compatible with itself. For H finite we will say γ on A^H is **δ-independent** whenever there is an indexing $H = \{h_1, \ldots, h_n\}$ such that $P^{\{h_i\}} \perp_{\delta,\gamma} P^{\{h_1,\ldots,h_{i-1}\}}$ for all $i = 2, \ldots, n$. If γ is a stationary probability measure on A^H then it has a unique one-dimensional marginal p, a probability measure on A, and one easily sees that δ-independence of γ implies $\|\gamma - p^H\|_\infty < (|H| - 1)\delta$. Note also that if γ is δ-independent then so is γ_I for any $I \subset H$. Here is the general version of our extension theorem, which applies to any countable discrete group. As we have already remarked, the precise definition of δ-independence in the statement of Theorem $2'$ is unimportant but our definition is convenient for the proof of the result.

THEOREM $2'$: *Given $\alpha > 0$ and $k \in \mathbb{N}$ there is a $\delta = \delta(\alpha, k) > 0$ such that the following holds: if A is a finite set, G is any countable discrete group, K is a subset of G with $|K| \le k$ and γ is any stationary probability measure on A^K such that γ is δ-independent and its one-dimensional marginal p satisfies $\min_{a\in A} p(a) > \alpha$ then there is a probability measure λ on A^G which is compatible with γ.*

COROLLARY (Theorem 2): *If the group G in Theorem $2'$ is amenable then the measure γ in the conclusion may be taken to be stationary.*

Proof: Since every right translate of the λ which Theorem $2'$ gives is compatible with γ we may simply average the right translates of λ over the sets in a Følner sequence and take any weak-* limit point to obtain a stationary λ' which is also compatible with γ. ∎

For the proof of Theorem 2′ we will need an abstract extension result for signed measures on A^I where I is simply an abstract finite set, not to be thought of as a subset of G for the time being. We remark that the notions of product measure, projection and consistency remain meaningful for such signed measures. For J and K subsets of I and λ a signed measure on A^K, $\pi_J\lambda$ will denote the projection of λ onto $A^{J\cap K}$. We adopt the convention that $\pi_\emptyset\lambda$ is the scalar $\lambda(A^K)$.

Suppose A and I are finite sets, $I_1, I_2, \ldots, I_k$ are subsets of I whose union is I and ρ is a signed measure on A^I. Then, letting $\rho_j = \pi_{I_j}\rho$ and $\mathcal{K} = \{1, \ldots, k\}$ the family (vector) of measures $\{\rho_j\}_{j\in\mathcal{K}}$ is a pairwise consistent family of signed measures. The following lemma asserts that any consistent family of signed measures arises in this way as the family of projections of a signed measure on A^I. Its proof is more or less the same as the proof of Proposition 1.2 of [J,R,W]. We include it here for completeness, as the statements of the two results are different.

LEMMA 2.1: *Let V denote the real vector space of all signed measures on A^I and W the real vector space of all pairwise consistent families $\{\rho_j\}_{j=1}^k$, ρ_j being a signed measure on A^{I_j}. Then the linear map Π: $V \to W$ defined by $\Pi\rho = \{\pi_{I_j}\rho\}_{j=1}^k$ is surjective.*

Proof: Before beginning the proof we make two observations. First, if J and K are disjoint subsets of I, m is a signed measure on A^J and p is a probability measure on A^K, then $\pi_J(m \times p) = m$. Secondly, assuming again that J and K are disjoint, for any signed measures m_1 and m_2 on A^J and A^K and any subset $L \subset I$ we have $\pi_L(m_1 \times m_2) = \pi_L m_1 \times \pi_L m_2$.

Write $\mathcal{K} = \{1, \ldots, k\}$ and for a non-empty subset $\mathcal{J} \subset \mathcal{K}$ let $I_{\mathcal{J}} = \bigcap_{j\in\mathcal{J}} I_j$. Suppose $\{\rho_j\}_{j\in\mathcal{K}} \in W$ and let $\check{\rho}_{\mathcal{J}}$ denote the common projection of the measures ρ_j onto $A^{I_{\mathcal{J}}}$. Fix an arbitrary probability measure q on A and let $\rho_{\mathcal{J}}$ denote the product of the measure $\check{\rho}_{\mathcal{J}}$ with the product measure $q^{I\setminus I_{\mathcal{J}}}$ on the remaining co-ordinates in I. Similarly for $i \in \mathcal{K}$ let $\rho^i_{\mathcal{J}}$ denote the product of the measure $\check{\rho}_{\{i\}\cup\mathcal{J}}$ with the product measure $q^{I_i\setminus I_{\{i\}\cup\mathcal{J}}}$ on the remaining co-ordinates in I_i. As a result of these definitions we have

$$\pi_{I_i}(\rho_{\mathcal{J}}) = \rho^i_{\mathcal{J}} = \rho^i_{\{i\}\cup\mathcal{J}}.$$

Now define

$$\rho = \sum_{\mathcal{J}\subset\mathcal{K},\mathcal{J}\neq\emptyset} (-1)^{|\mathcal{J}|+1}\rho_{\mathcal{J}}.$$

Fix an i between 1 and k; we must show $\pi_{I_i}\rho = \rho_i$. To this end rewrite

$$\begin{aligned}\rho = &\sum_{\mathcal{J}\subset\mathcal{K},\mathcal{J}\neq\phi,i\in\mathcal{J}} (-1)^{|\mathcal{J}|+1}\rho_{\mathcal{J}} + \sum_{\mathcal{J}\subset\mathcal{K},\mathcal{J}\neq\phi,i\notin\mathcal{J}} (-1)^{|\mathcal{J}|+1}\rho_{\mathcal{J}}\\ =&\rho_{\{i\}} + \sum_{\mathcal{J}\neq\phi,i\notin\mathcal{J}} (-1)^{|\mathcal{J}|}\rho_{\{i\}\cup\mathcal{J}} + \sum_{\mathcal{J}\neq\phi,i\notin\mathcal{J}} (-1)^{|\mathcal{J}|+1}\rho_{\mathcal{J}}.\end{aligned}$$

Now applying π_{I_i} to both sides of this equation and using

$$\pi_{I_i}(\rho_{\mathcal{J}}) = \rho^i_{\{i\}\cup\mathcal{J}} = \pi_{I_i}(\rho_{\{i\}\cup\mathcal{J}})$$

we obtain $\pi_{I_i}(\rho) = \pi_{I_i}\rho_{\{i\}} = \rho_i$. ∎

We will also need the following lemma.

LEMMA 2.2: *Suppose V and W are finite-dimensional real normed vector spaces and $\Pi: V \to W$ is any surjective linear map. Then there is a constant $C = C(\Pi)$ with the following property: if $v \in V$ and $w = \Pi(v) \neq 0$ then Π has a right inverse B ($\Pi B = id_W$) such that $Bw = v$ and $\|B\| \leq C\frac{\|v\|}{\|w\|}$.*

Proof: Before beginning the proof we make two observations. First, as all norms on a finite dimensional vector space are equivalent, the conclusion of the lemma persists if we replace the norm of either V or W by a different one. Also note that if $I: W \to W'$ is any isomorphism then the conclusion of the lemma persists if Π is replaced by $I \circ \Pi$, as both I and I^{-1} are bounded.

In view of these remarks we may start by assuming that V is Euclidean. Let $E = \ker\Pi$ and $F = E^{\perp}$. Then $I = \Pi|_F$ is an isomorphism and replacing Π by $I^{-1}\circ\Pi$ we obtain the orthogonal projection onto F. Consequently, we shall assume at the outset that Π itself is the orthogonal projection onto F. Let $Z = F \cap \{w\}^{\perp}$, so we have $F = Z \oplus \mathbb{R}w$, an orthogonal direct sum. Define $B(z + tw) = z + tv$ for $z \in Z$, $t \in \mathbb{R}$. Certainly B is a right inverse of Π and $Bw = v$. Moreover

$$\frac{\|B(z+tw)\|^2}{\|z+tw\|^2} = \frac{\|z/t\|^2 + \|v\|^2}{\|z/t\|^2 + \|w\|^2} \leq \frac{\|v\|^2}{\|w\|^2},$$

since $\|w\| = \|\Pi v\| \leq \|v\|$. So, in the special case that we have reduced to we may take $C = 1$. ∎

Theorem 2′ is an easy consequence of the following proposition. We set $\beta = \alpha^{k^2}/2k^2$.

Proposition 2.3: *If $\delta = \delta(\alpha, k)$ is sufficiently small then for any finite $I \subset G$, any probability measure λ on A^I which is β-independent and compatible with γ and any $g_0 \in G \backslash I$ there is an extension λ' of λ to a measure on $A^{I \cup \{g_0\}}$ which is again β-independent and compatible with γ.*

We will assume, without loss of generality, that $g_0 = e = e_G$, the identity element of G. For each $g \in K$ let

$$\begin{aligned} S_g &= Kg^{-1} \cap (I \cup \{e\}), \\ S &= \bigcup \{S_g : g \in K\}, \\ \mu_g &= \mathcal{R}_{g^{-1}} \gamma_{S_g g}, \\ R_g &= S_g \backslash \{e\} = Kg^{-1} \cap I, \\ R &= \bigcup \{R_g : g \in K\} = S \backslash \{e\}. \end{aligned}$$

We are going to construct a probability measure σ on A^S such that

$$\pi_R \sigma = \pi_R \lambda, \tag{1}$$

$$\pi_{S_g} \sigma = \mu_g \quad \forall g \in K \tag{2}$$

and

$$P^{\{e\}} \perp_{\beta,\sigma} P^R. \tag{3}$$

Let us assume for the moment that we have already constructed σ satisfying (1), (2) and (3). Because of (1) we can then define λ' to be the relative product measure $\lambda' = \lambda \times_{A^R} \sigma$, that is

$$\lambda'(x) = \frac{\lambda(P^I x) \sigma(P^S x)}{\lambda(P^R x)}, \quad \forall x \in A^{I \cup \{e\}}.$$

λ' is certainly an extension of λ and of σ. (2) ensures that λ' is compatible with γ, since any translate of K which contains e yields one of the S_g upon intersection with $I \cup \{e\}$ and any other translate of K is taken care of by the compatibility of λ with γ. Moreover, the definition of λ' together with (3) implies that for all $X \in P^{\{e\}}$, $Y \in P^R$ and $Z \in P^{I \backslash R}$ we have $\lambda'(X|Y \cap Z) = \sigma(X|Y) \overset{\beta}{\sim} p(X)$. This means that $P^{\{e\}} \perp_{\beta,\lambda'} P^I$, and since λ is β-independent it follows that λ' is also β-independent, completing the proof of the proposition.

We now proceed to construct σ. For each $g \in K$ let

$$\nu_g = \pi_{R_g}(\mu_g) = \pi_{R_g} \lambda_R.$$

(The second equality follows from the definition of the μ_g and the compatibility of λ with γ.) $\{\nu_g\}_{g\in K}$ is a pairwise consistent family of measures as its components are all marginals of the single measure λ_R. We let V denote the vector space of all signed measures on A^R and W the space of all consistent families $\{\rho_g\}_{g\in K}$, where ρ_g is a signed measure on A^{R_g} for each $g \in K$, both spaces endowed with the l_∞-norm. We let Π: $V \to W$ denote the projection map, that is $\Pi(\rho) = \{\pi_{R_g}\rho\}_{g\in K}$ for each $\rho \in V$, so we have $\Pi\lambda_R = \{\nu_g\}_{g\in K}$. Lemma 2.1 tells us that Π is surjective so by Lemma 2.2 Π has a right inverse B such that

$$B(\{\nu_g\}_{g\in K}) = \lambda_R$$

and

$$\|B\| \le C\|\lambda_R\|/\|\{\nu_g\}_{g\in K}\|,$$

where $C = C(\Pi)$ is a constant depending on Π. However, since $R \subset KK^{-1}$, Π is completely determined by specifying the subsets R_g, $g \in K$, of KK^{-1}, of which there are at most k, so there are no more than $(2^{k^2})^k$ possibilities for the projection Π. This means that we may take C to be a constant depending only on k. Since λ_R is a probability measure we have $\|\lambda_R\|_\infty \le 1$. Moreover, the hypothesis $p(a) > \alpha$ ensures that $|A| < \alpha^{-1}$. Since each ν_g is a probability measure on a set of cardinality at most $|A|^k$, we have

$$\|\{\nu_g\}_{g\in K}\|_\infty \ge |A|^{-k} \ge \alpha^k,$$

so we obtain an absolute bound $\|B\| \le C\alpha^{-k} = C'$.

We view each $y \in A^{R\cup\{e\}} = A^S$ as a pair $y = (x, a), x \in A^R$, $a \in A$ and use a similar convention for the sets A^{S_g} (recall $S_g = R_g \cup \{e\}$). With this convention if σ is a signed measure on A^S then $\sigma(\cdot, a)$ is a measure on A^R and $\mu_g(\cdot, a)$ is a measure on A^{R_g} for each $g \in G$. We now define σ by specifying that

$$\sigma(\cdot, a) = B(\{\mu_g(\cdot, a)\}_{g\in K}) \quad \forall a \in A.$$

Note that the consistency of the family $\{\mu_g(\cdot, a)\}_{g\in K}$ follows from the consistency of the family $\{\mu_g\}_{g\in K}$, which in turn follows from the facts that all the μ_g are translates of projections of γ and γ is stationary.

We then have

$$\begin{aligned}\pi_R\sigma &= \sum_{a\in A}\sigma(\cdot,a)\\ &= \sum_{a\in A} B(\{\mu_g(\cdot,a)\}_{g\in K})\\ &= B\left(\sum_{a\in A}\{\mu_g(\cdot,a)\}_{g\in K}\right) \quad \text{(since } B \text{ is linear)}\\ &= B(\{\nu_g\}_{g\in K}) = \pi_R\lambda,\end{aligned}$$

establishing (1). To check (2) observe that for each $a\in A$ and $k\in K$ we have

$$(\pi_{S_k}\sigma)(\cdot,a) = \pi_{R_k}(\sigma(\cdot,a)) = \pi_{R_k}B(\{\mu_h(\cdot,a)\}_{h\in K}) = \mu_k(\cdot,a),$$

since $\Pi B = \mathrm{id}_W$.

It remains to check that σ is non-negative and satisfies (3). Note that each μ_g is δ-independent since it is a translate of a marginal of γ, and the same holds for each ν_g. Now recall that

$$\|\mu_g - p^{S_g}\|_\infty \le (|S_g|-1)\delta < k\delta$$

and again the same goes for ν_g. As a result we have

$$(\mu_g(x,a)) \overset{k\delta}{\sim} p(a)p^{R_g}(x) \overset{k\delta}{\sim} p(a)\nu_g(x) \quad \forall g\in K,\ x\in A^{R_g},\ a\in A.$$

This means (since we are using l_∞-norms) that

$$\{\mu_g(\cdot,a)\}_{g\in K} \overset{2k\delta}{\sim} p(a)\{\nu_g\}_{g\in K},$$

so using $\|B\|\le C'$ we obtain

$$\sigma(\cdot,a) \overset{2C'k\delta}{\sim} p(a)B(\{\nu_g\}_{g\in K}) = p(a)\lambda_R. \tag{4}$$

Since λ_R is β-independent and $|R|\le k^2$, for each $x\in A^R$ we have

$$\lambda_R(x) \overset{k^2\beta}{\sim} p^R(x) \ge \alpha^{k^2}.$$

Recalling that $\beta = \alpha^{k^2}/2k^2$ it follows that $\lambda_R(x)\ge \alpha^{k^2}/2$ for all $x\in A^R$. Combining this with (4) and taking $\delta < \beta\alpha^{k^2}/4C'k$ we get

$$\left|\frac{\sigma(x,a)}{\pi_R\lambda(x)} - p(a)\right| \le \frac{4C'k\delta}{\alpha^{k^2}} < \beta \quad \forall x\in A^R,\ a\in A. \tag{5}$$

Since $p(a) > \alpha$ and $\beta \leq \alpha/2$ this shows that σ takes only positive values so it is a probability measure. (5) also shows that $P^{\{e\}} \perp_{\beta,\sigma} P^R$, establishing (3) and concluding the proof of Proposition 2.3. ∎

Theorem 2′ now follows from Proposition 2.3 by enumerating the elements of G which are not in K in any way and then adjoining them to K one at a time.

3. Partitions with independent iterates

In this section we will be concerned with a free action $g \mapsto T_g$ of a discrete amenable group G on $(X, \mathcal{B}, \mu)$. We will usually write $T_g(x) = gx$. If $P: X \to A$ is a finite partition and L is a subset of G we define the map $P^L: X \to A^L$ by $P^L(x)(l) = P(lx)$ for $l \in L$. When L is finite P^L is again a finite partition. If $L \subset G$ is infinite then the measures $\operatorname{dist} P^{L'}$, $L' \subset L$ finite, are mutually consistent and thus define a measure on A^L (endowed with the standard product Borel structure). We denote this measure by $\operatorname{dist} P^L$, just as for finite L. Note that $\mathcal{R}_g \circ P^L = P^{Lg} \circ T_{g^{-1}}$, which implies that $\operatorname{dist} P^L$ is a stationary measure on A^L for any L (finite or not). (This is the reason for defining stationarity as right invariance.) We define the partition gP to be $P \circ T_{g^{-1}}$ in accordance with the fact that $g\{P = a\} = \{P \circ T_{g^{-1}} = a\}$.

If P and Q are finite partitions of X taking values in A such that $\mu\{x : Px \neq Qx\} \leq \epsilon$ we will write $P \overset{\epsilon}{\sim} Q$. For probability measures λ and ν on a finite set A we write $\lambda \overset{\epsilon}{\sim} \nu$ whenever $\|\lambda - \nu\|_1 \leq \epsilon$. (Take note that we are not using the l_∞-norm here as we did in §2.) We will use the notation $\eta_1 = \flat(\eta)$ generically for any positive function η_1 of a positive real number η which goes to zero as η goes to zero.

The proof of the following lemma is straight-forward.

Lemma 3.1: *If P is a partition of X, $X^* \subset X$ and $\mu(X^*) > 1 - \delta$ then* $\operatorname{dist}(P|X^*) \overset{2\delta}{\sim} \operatorname{dist} P$.

In this section we will be exploiting the left amenability of G as opposed to the right amenability used in §2. Following [O,W], p. 24, if K and L are finite subsets of G such that $e_G \in K$ we say that L is (K, η)-invariant if

$$|\{g \in G : Kg \cap L \neq \emptyset \text{ and } Kg \cap L^c \neq \emptyset\}| < \eta |L|.$$

Since $e_G \in K$ the set above, which may be viewed as the K-boundary of L, can also be written as

$$\Big(L \setminus \bigcap_{k \in K} k^{-1} L\Big) \cup \Big(L^c \setminus \bigcap_{k \in K} k^{-1} L^c\Big).$$

Although we conform to the terminology of [O,W] we will mostly use only the fact that whenever L is (K,η)-invariant then the part of the K-boundary of L which lies inside L is small:

$$|\bigcap_{k\in K} k^{-1}L| > (1-\eta)|L|.$$

If L is a finite subset of G, $E \subset X$ and $\{lE : l \in L\}$ are pairwise disjoint then the pair (L,E), is called a (Rohlin) **tower**. According to [O,W] a family $\{E_i\}_{i\in I}$ of subsets of a measure space (X,μ) is called η-**disjoint** if there are subsets $F_i \subset E_i$ of relative measure greater than $1-\eta$ such that the sets $\{F_i\}_{i\in I}$ are (pairwise) disjoint. By an η-**quasi-tower** we mean a pair (L,E) where L is a finite subset of G and E is a measurable subset of X such that the sets $\{lE : l \in L\}$ are η-disjoint. If $\{F^l\}_{l\in L}$ are the large pairwise disjoint subsets of the sets $\{lE\}$ which are implicit in this definition we will also refer to $(L,E,\{F^l\}_{l\in L})$ as an η-quasi-tower. When $L\subset G$ and $E\subset X$ we write $LE = \bigcup_{l\in L} lE$.

The following proposition is a variant of the strong Rohlin lemma for amenable groups [O,W], which replaces many η-disjoint towers (M,E_j) with a single quasi-tower (L,E) where $E=\bigcup_j E_j$ and L is a large subset of M. Let us say that a sequence $(M_1,\dots,M_n)$ of finite subsets of G is δ-invariant if each M_i is $(M_{i-1}M_{i-1}^{-1},\delta)$-invariant. The statement of the strong Rohlin lemma for discrete amenable groups in [O,W] (II§2 Theorem 6) contains the hypothesis that such a sequence is "sufficiently invariant". A reading of the proof of I§2 Theorem 6 of [O,W] shows that "sufficiently invariant" should be construed as meaning that n is sufficiently large and that $(M_1,\dots,M_n)$ is δ-invariant for a sufficiently small δ (depending on n).

Proposition 3.2: *Given $\eta>0$ there are an $n\in\mathbb{N}$ and a $\delta>0$ with the following property: For any discrete amenable G, any δ-invariant sequence $(M_1,\dots,M_n)$ in G, any free action T of G on $(X,\mathcal{B},\mu)$ and any finite partition Q of X there are subsets L_i of M_i, $i=1,\dots,n$ and subsets E_i of X, $i=1,\dots,n$ such that*

$$|L_i| > (1-\eta)|M_i| \quad \forall i,$$

$$(L_i,E_i) \textit{ is an } \eta\textit{-quasi-tower} \quad \forall i,$$

$$\textit{the sets } L_iE_i, i=1,\dots,n \textit{ are pairwise disjoint},$$

$$\mu\Big(\bigcup_{i=1}^{n} L_iE_i\Big) > 1-\eta$$

and

$$\operatorname{dist}(Q|E_i) = \operatorname{dist} Q \quad \forall i.$$

Proof: Fix a small η' to be specified later. If n is sufficiently large and $(M_1, \ldots, M_n)$ is a δ-invariant sequence with δ sufficiently small then the strong Rohlin lemma ([O,W] II§2 Theorem 6, see also the remark after the proof) gives us sets

$$E_{ij} \subset X, \quad i = 1, \ldots, n, \ j = 1, \ldots, J$$

such that

$$(M_i, E_{ij}) \text{ is a tower for each } i \text{ and } j, \tag{1}$$

$$M_i E_{ij} \cap M_{i'} E_{i'j'} = \emptyset \quad \text{for } i \neq i', \tag{2}$$

$$\{M_i E_{ij} : j = 1, \ldots, J\} \text{ is } \eta'\text{-disjoint for each } i, \tag{3}$$

$$\mu\Big(\bigcup_{i,j} M_i E_{ij}\Big) > 1 - \eta' \tag{4}$$

and

$$\operatorname{dist}\Big(Q|\bigcup_j E_{ij}\Big) = \operatorname{dist} Q \quad \text{for each } i. \tag{5}$$

Fix an i between 1 and n. For each j we have a subset F_{ij} of $M_i E_{ij}$ of relative measure greater than $1-\eta'$ such that the $\{F_{ij}\}_{j=1}^J$ are disjoint. Denote $F_{ij} \cap l E_{ij}$ by F_{ij}^l, observe that the sets $\{F_{ij}^l : j = 1, \ldots, J, \ l \in M_i\}$ are pairwise disjoint and let $F_i^l = \bigcup_j F_{ij}^l$. Let c_i denote $\sum_j \mu(E_{ij})$. We have

$$\sum_{l \in M_i} \mu(F_{ij}^l) = \mu(F_{ij}) > (1-\eta')\mu(M_i E_{ij}) = (1-\eta')|M_i|\mu(E_{ij}) \quad \text{for each } j.$$

Summing over j and reversing the order of summation in the resulting double sum we obtain

$$\sum_{l \in M_i} \mu(F_i^l) > (1-\eta') c_i |M_i|. \tag{6}$$

Letting

$$L_i = \{l \in M_i : \mu(F_i^l) > (1-\sqrt{\eta'})c_i\} \tag{7}$$

(6) implies that

$$|L_i| > (1-\sqrt{\eta'})|M_i|. \tag{8}$$

Letting $E_i = \bigcup_j E_{ij}$, for $l \in L_i$ we have

$$\mu(F_i^l) > (1-\sqrt{\eta'})c_i \geq (1-\sqrt{\eta'})\mu(E_i). \tag{9}$$

(9) tells us that each (L_i, E_i) is an η-quasi-tower if we require $\eta' < \eta^2$. Moreover (8), (9) and (4) imply that

$$\begin{aligned}\mu(\bigcup_i L_i E_i) &\geq \sum_i \sum_{l\in M_i} \mu(F_i^l)\\ &> \sum_i \sum_{l \in L_i} (1-\sqrt{\eta'})(1-\sqrt{\eta'})\mu(lE_i)\\ &> (1-\sqrt{\eta'})(1-\sqrt{\eta'})(1-\eta') > 1-\eta,\end{aligned}$$

if η' is sufficiently small. Finally $\operatorname{dist}(Q|E_i) = \operatorname{dist} Q$ is just equation (5). ∎

Next we describe the process of defining a partition on a quasi-tower by assigning names to points in the base. Suppose $(L, E, \{F^l\}_{l\in L})$ is an η-quasi-tower and λ is a probability measure on A^L. Choose any partition $\xi\colon E \to A^L$ such that $\operatorname{dist}\xi = \lambda$ and define a partition P on $F = \bigcup_{l\in L} F^l$ by letting $P(lx) = \xi(x)(l)$ for $l \in L$ and $x \in E$, whenever $lx \in F^l$. We describe this process as "copying λ onto $(L, E, \{F^l\}_{l\in L})$". Keeping this notation we then have the following result.

Lemma 3.3: *Suppose K is a finite subset of G, $e \in K$ and L is (K,η)-invariant. Let $L' = \bigcap_{k\in K} k^{-1}L$, so $|L'| > (1-\eta)|L|$. Then for each $l \in L'$, P^K is defined on $F'^l = \bigcap_{k\in K} k^{-1}F^{kl}$ and this set has relative measure greater than $1-|K|\eta$ in lE. If λ is stationary then for each $l \in L'$, $\operatorname{dist}(P^K|F'^l) \overset{\eta_1}{\sim} \lambda_K$, $\eta_1 = \flat(\eta)$.*

Proof: The only thing requiring proof is the last assertion of the lemma. Observe that for $l \in L'$, $y = lx \in F'^l$ and $k \in K$ we have

$$P^K(lx)(k) = P(klx) = \xi(x)(kl) = (\mathcal{R}_{l^{-1}}\pi_{Kl}\xi(x))(k),$$

so $P^K(y) = \mathcal{R}_{l^{-1}}\pi_{Kl}\xi(x)$ and hence

$$\operatorname{dist}(P^K|F'^l) = \mathcal{R}_{l^{-1}}\pi_{Kl}\operatorname{dist}(\xi|l^{-1}F'^l). \tag{1}$$

Now since $\mu(l^{-1}F'^l) > (1-|K|\eta)\mu(lE)$ Lemma 3.1 implies that

$$\operatorname{dist}(\xi|l^{-1}F'^l) \overset{\flat(\eta)}{\sim} \operatorname{dist}\xi = \lambda. \tag{2}$$

Finally, (1) and (2) imply that

$$\operatorname{dist}(P^K|F'^l) \overset{\flat(\eta)}{\sim} \mathcal{R}_{l^{-1}}\pi_{Kl}\lambda = \pi_K\lambda$$

as desired. ∎

We will prove Theorem 1 by iteration of the following proposition.

PROPOSITION 3.4: *Given $\alpha > 0$, $k \in \mathbb{N}$ and $\epsilon > 0$ there is a $\delta > 0$ with the following property: Suppose K is a subset of G, $e \in K$, $|K| \le k$ and $P\colon X \to A$ is a partition with* $\operatorname{dist} P = p$ *such that* $\operatorname{dist} P^K \overset{\delta}{\sim} p^K$ *and $p(a) > \alpha$ for all $a \in A$. Then given $\delta' > 0$ there is a partition $Q \overset{\epsilon}{\sim} P$ such that* $\operatorname{dist} Q = p$ *and* $\operatorname{dist} Q^K \overset{\delta'}{\sim} p^K$.

Proof: We start with the observation that if we have achieved the desired requirements for Q, except for possibly $\operatorname{dist} Q = p$, then this last condition is easily achieved as well by a slight further perturbation which will have only a small effect on the other two conditions. Thus we shall ignore the requirement $\operatorname{dist} P = \operatorname{dist} Q$.

We will work with the hypothesis that $\operatorname{dist} P^K \overset{\delta}{\sim} p^K$ and specify in the course of the argument how small δ needs to be. Let $\lambda = \operatorname{dist} P^G$ so $\lambda_K = \operatorname{dist} P^K$. Define the (signed) measure σ on A^K via the formula

$$p^K = \frac{\epsilon}{2}\sigma + \Big(1 - \frac{\epsilon}{2}\Big)\lambda_K.$$

Since ϵ is fixed we will have $\sigma \overset{\delta_1}{\sim} p^K$, $\delta_1 = \flat(\delta)$, and in particular σ is a probability measure if δ, and hence δ_1, is sufficiently small. (Here we use the fact that $p^K(\mathbf{a}) > \alpha^{|K|} \ge \alpha^k$ for all $\mathbf{a} \in A^K$ and that α and k are fixed.) Moreover, σ is stationary, since p^K and λ_K are both stationary. Now if δ, and hence δ_1, is sufficiently small, then by Theorem 2, σ extends to a stationary measure ρ on A^G.

Given a small η to be specified later we wish to find finite sets

$$L_i \subset G, \quad i = 1, \dots, n,$$

and sets

$$E_i \subset X, \quad i = 1, \dots, n$$

such that

(1) $$L_i \text{ is } (K,\eta)\text{-invariant for each } i,$$

(2) $$(L_i, E_i) \text{ is an } \eta\text{-quasi-tower for each } i,$$

(3) $$L_i E_i \cap L_{i'} E_{i'} = \emptyset \quad \text{for } i \ne i',$$

(4) $$\mu\Big(\bigcup_i L_i E_i\Big) > 1 - \eta$$

and

$$\text{(5)} \qquad \operatorname{dist}(P^{L_i}|E_i) = \operatorname{dist} P^{L_i} = \lambda_{L_i}.$$

Proposition 3.2 ensures that we can do this: given a sufficiently invariant $(M_1, \ldots, M_n)$ and taking Q in Proposition 3.2 to be $P^{\cup_i M_i}$ we get sets L_i and E_i as specified in Proposition 3.2. Since $L_i \subset \bigcup M_i$, (5) is immediate. As for (1), if we take each M_i to be (K, η')-invariant and also ensure that $|L_i| > (1-\eta')|M_i|$ then it follows that L_i is $(K, (|K|+2)\frac{|M_i|}{|L_i|}\eta')$-invariant so we need only take $\eta' = \frac{1}{2}\eta(|K|+2)^{-1}$.

For each i and $l \in L_i$ let F_i^l be a subset of lE_i of relative measure greater than $1-\eta$ such that the sets $\{F_i^l : l \in L_i\}$ are disjoint. For each i choose a subset B_i of E_i of relative measure $1-\epsilon/2$ such that $\operatorname{dist}(P^{L_i}|B_i) = \operatorname{dist}(P^{L_i}|E_i)$ and let $A_i = E_i \backslash B_i$. Let

$$A_i^l = F_i^l \cap lA_i, \quad A_i'^l = \bigcap_{k \in K} k^{-1} A_i^{kl}$$

and

$$B_i^l = F_i^l \cap lB_i, \quad B_i'^l = \bigcap_{k \in K} k^{-1} B_i^{kl}.$$

Define Q on $\bigcup_{i,l} A_i^l$ by copying ρ_{L_i} on $(L_i, E_i, \{A_i^l\}_{l \in L})$ for each i. On $X \backslash \bigcup_{i,l} A_i^l$ define $Q(x) = P(x)$. According to Lemma 3.3 for each i and $l \in L_i'$ we have

$$\text{(6)} \qquad \operatorname{dist}(Q^K | A_i'^l) \overset{\eta_1}{\sim} \rho_K = \sigma$$

and

$$\text{(7)} \qquad \operatorname{dist}(Q^K | B_i'^l) \overset{\eta_1}{\sim} \lambda_K,$$

where $\eta_1 = \flat(\eta)$.

We have

$$\text{(8)} \qquad \operatorname{dist}(Q^K | F_i'^l) = a_i^l \operatorname{dist}(Q^K | A_i'^l) + b_i^l \operatorname{dist}(Q^K | B_i'^l),$$

where a_i^l and b_i^l denote the relative measures of $A_i'^l$ and $B_i'^l$ in $F_i'^l$. Since a_i^l and b_i^l are close (within $\flat(\eta)$) to $\epsilon/2$ and $1-\epsilon/2$, (6), (7) and (8) imply that

$$\operatorname{dist}\left(Q^K | \bigcup_{i,l} F_i'^l\right) \overset{\eta_2}{\sim} \frac{\epsilon}{2}\sigma + \left(1 - \frac{\epsilon}{2}\right)\lambda_K = p^K,$$

where $\eta_2 = \flat(\eta)$. Finally, $\bigcup_{i,l} F_i'^l$ has large (more than $1-\flat(\eta)$) measure in X so our result follows from Lemma 3.1. ∎

Corollary 3.5: *In Proposition 3.4, δ' may be replaced by 0, that is, the conclusion becomes* $\operatorname{dist} Q^K = p^K$.

Proof: Iterate Proposition 3.4 to arrive at a limiting partition with the desired properties. ∎

Theorem 1: *Suppose a discrete amenable group G acts freely on $(X, \mathcal{B}, \mu)$ and $\{g_i\}$ is a sequence of elements of G which is mixing for this action. Then for any finite partition P of X and $\epsilon > 0$ there are a partition $Q \overset{\epsilon}{\sim} P$ and a subsequence $\{g_{i_j}\}$ such that the partitions $\{gP : g \in \mathrm{IP}'\{g_{i_j}\}\}$ are (jointly) independent.*

Proof: To begin observe that for any $\mathbf{a} \in A^L$ we have

$$\bigcap_{g \in L} g^{-1}\{x : P(x) = \mathbf{a}(g)\} = \{x : P^L(x) = \mathbf{a}\}.$$

It follows that the independence or almost independence of a family $\{gP : g \in L\}$ corresponds to the independence or almost independence of $\operatorname{dist} P^{L^{-1}}$.

Now suppose that we have already found $Q_k \overset{\epsilon}{\sim} P$ and $g_{i_1}, \ldots, g_{i_k}$ so that, denoting $L = \mathrm{IP}'(\{g_{i_j}\}_{j=1}^k)$, the partitions $\{gQ_k : g \in L\}$ are independent. This means that $\operatorname{dist} Q_k^{(L^{-1})} = p^{(L^{-1})}$ and by translation we also have that $\operatorname{dist} Q_k^{L^{-1}g^{-1}} = p^{L^{-1}g^{-1}}$ for any $g \in G$. Given $\eta > 0$ we can use the mixing property of $\{g_i\}$ to find $g_{i_{k+1}}$ so that the partitions $Q_k^{L^{-1}}$ and $g_{i_{k+1}}^{-1} Q_k^{L^{-1}}$ are almost independent of each other. By our remark above this means that for any given $\eta > 0$ we can ensure that

$$\operatorname{dist} Q_k^{L^{-1} \cup L^{-1} g_{i_{k+1}}^{-1}} \overset{\eta}{\sim} \operatorname{dist} Q_k^{L^{-1}} \times \operatorname{dist} Q_k^{L^{-1} g_{i_{k+1}}^{-1}} = p^{L^{-1} \cup L^{-1} g_{i_{k+1}}^{-1}}.$$

By Corollary 3.5 we can then find $Q_{k+1} \overset{\flat(\eta)}{\sim} Q_k$ such that $\operatorname{dist} Q_k^{L^{-1} \cup L^{-1} g_{i_{k+1}}^{-1}} = p^{L^{-1} \cup L^{-1} g_{i_{k+1}}^{-1}}$, which means that the family $\{gQ_{k+1} : g \in L \cup g_{i_{k+1}} L\}$ is independent. Moreover, if η is sufficiently small we will still have $Q_{k+1} \overset{\epsilon}{\sim} P$.

Iterating the process we have just described yields a sequence $\{g_{i_k}\}$ and a convergent sequence $\{Q_k\}$ whose limiting partition Q satisfies the conclusion of the theorem. ∎

Theorem 3: *Suppose $\{T_i\}$ is a mixing sequence of pairwise commuting automorphisms of $(X, \mathcal{B}, \mu)$ and P is a finite partition. Then for any $\epsilon > 0$ there are a $Q \overset{\epsilon}{\sim} P$ and a subsequence $\{T_{i_j}\}$ such that the partitions $\{TQ : T \in \mathrm{IP}'\{T_{i_j}\}\}$ are (jointly) independent.*

Proof: Let G be the abelian group generated by $\{T_i\}$, acting on $(X, \mathcal{B}, \mu)$ in the obvious way. This action is ergodic since it contains a mixing sequence and it is free since for $T \in G$ the set of fixed points of T is G-invariant and hence has measure either 0 or 1, so theorem 1 applies. ∎

Next we proceed to the characterizations of weak, mild and strong mixing, Theorems 4, 5 and 6. Recall that in these theorems the measure space is assumed to be separable. We refer the reader to the papers of Dye [D] and Bergelson–Rosenblatt [Be,Ro] for a discussion of weak mixing of general group actions. If G is amenable with Følner sequence $\{F_n\}$ and $E \subset G$ the **density** $\mathrm{d}(E)$ of E with respect to $\{F_n\}$ is defined as $\lim_{n\to\infty} |F_n|^{-1}|F_n \cap E|$, if the limit exists. When the dependence on the choice of Følner sequence needs to be made explicit we write $\mathrm{d}_{\{F_n\}}$. The **upper density** of E, denoted $\bar{\mathrm{d}}(E)$, is defined by replacing lim by lim sup. Both $\mathrm{d}_{\{F_n\}}$ and $\bar{\mathrm{d}}_{\{F_n\}}$ are invariant with respect to left translation. Whenever T is an action we will denote by U the corresponding unitary action of G on $\mathcal{H} = L_0^2(\mu)$, the orthocomplment of the constants. Of the many characterizations of weak mixing we shall use two. The first is that T is weakly mixing if and only if for any Følner sequence $\{F_n\}$, any measurable sets A and B and any $\epsilon > 0$ we have

$$\mathrm{d}_{\{F_n\}}\{g \in G : |\mu(g^{-1}A \cap B) - \mu(A)\mu(B)| < \epsilon\} = 1.$$

The second is that T is weakly mixing if and only if it has no non-trivial compact factors. A function $f \in L_2(\mu)$ is **compact** for the action if the orbit Gf is pre-compact in the norm topology of $L_2(\mu)$ and a factor algebra $\mathcal{F}$ of the action is compact if 1_A is compact for each $A \in \mathcal{F}$.

THEOREM 4: *The G-action T is weakly mixing if and only if, given any set $\Sigma \subset G$ which has positive density with respect to some Følner sequence, the partitions which are independent along some* IP′*-set with generators in Σ are dense in the set of all finite partitions.*

Proof: For the "only if" statement, suppose Σ has positive density with respect to the Følner sequence $\{F_n\}$ and let d denote $\mathrm{d}_{\{F_n\}}$. Since T is weakly mixing we have that

$$\mathrm{d}\{g \in G : |\mu(g^{-1}A \cap B) - \mu(A)\mu(B)| < \epsilon\} = 1.$$

The intersection of finitely many such sets still has density one and intersecting further with Σ still leaves us with a set of positive density, in particular a non-empty set. This means that we can find elements of Σ which mix more and more sets better and better. Using a countable dense family of measurable sets and a diagonalisation argument we obtain a mixing sequence in Σ and then we apply Theorem 1.

For the "if" direction, suppose T is not weak mixing so there is a measurable A such that the orbit $G1_A$ is pre-compact. Then for any $\epsilon > 0$ there is a finite $K \subset G$ such that $K1_A$ is ϵ-dense in $G1_A$. Letting $\Sigma = \{g \in G : \|g1_A - 1_A\|_2 < \epsilon\}$ it follows that $K\Sigma = G$. Choosing any left Følner sequence $\{F_n\}$ and letting $\bar{\mathrm{d}}$ denote $\bar{\mathrm{d}}_{\{F_n\}}$, we must have $\bar{\mathrm{d}}(k\Sigma) > 0$ for some $k \in K$ and hence also $\bar{\mathrm{d}}(\Sigma) > 0$. By passing to a subsequence of $\{F_n\}$ we may assume that $\mathrm{d}(\Sigma) > 0$. However, if ϵ is sufficiently small, depending on the measure of A, the definition of Σ clearly prevents the existence of a $Q \overset{\epsilon}{\sim} \{A, A^c\}$ and a $g \in \Sigma$ such that Q and gQ are independent. This contradiction concludes the proof. ∎

The definition of mild mixing which we shall adopt is that T is **mildly mixing** if and only if it has no non-constant rigid functions in $L_2(\mu)$. $f \in L_2(\mu)$ is called **rigid** if there is a sequence $g_i \to \infty$ such that $\|g_i f - f\| \to 0$. It is easy to see that the existence of a non-constant rigid function is equivalent to the existence of a non-constant rigid characteristic function 1_A.

We refer the reader to [F2] for the theory of mild mixing. See also [F,K], [S] and [S,W]. The result we need, Lemma 3.6 below, does not appear in the literature in the general non-abelian setting. However, the development of the aspects of the non-abelian theory of mild mixing which we need is virtually identical to the development in the abelian setting. We will briefly indicate the changes that need to be made.

Let $\mathcal{F}$ denote the family of all finite non-empty subsets of $\mathbb{N}$. For α and β in $\mathcal{F}$ write $\beta > \alpha$ if $\min\beta > \max\alpha$. An **$\mathcal{F}$-sequence** is a mapping $\varphi : \mathcal{F} \to X$, where X is any set. If X is a topological space and $\varphi : \mathcal{F} \to X$ then $\mathrm{IP} - \lim \varphi(\alpha) = x$ means that for every neighborhood U of x there is an $\alpha \in \mathcal{F}$ such that $\varphi(\beta) \in U$ $\forall \beta > \alpha$. Let us say that φ': $\mathcal{F} \to X$ is an $\mathcal{F}$-subsequence of φ if there are $\alpha_1 < \alpha_2 < \cdots \in \mathcal{F}$ such that $\varphi'\{i_k, \ldots, i_1\} = \varphi(\alpha_{i_k} \cup \cdots \cup \alpha_{i_1})$. (The usual definition, see [F2], only requires that the α_i be disjoint.) As a consequence of Hindman's Theorem, if X is compact then every $\mathcal{F}$-sequence φ has an IP-convergent $\mathcal{F}$-subsequence φ. This is theorem 2.1 of [F2] and the proof in [F2] in fact gives an $\mathcal{F}$-subsequence in the stronger sense which we are requiring. If G is a semi-group, a **homomorphism** is any $\mathcal{F}$-sequence φ: $\mathcal{F} \to G$ such that $\varphi(\beta \cup \alpha) = \varphi(\beta)\varphi(\alpha)$ whenever $\beta > \alpha$. This amounts to saying that

$$\varphi\{i_k > \cdots > i_1\} = \varphi\{i_k\} \cdots \varphi\{i_1\}.$$

In [F2] the condition $\beta > \alpha$ in the definition of a homomorphism is replaced by $\beta \cap \alpha = \emptyset$ but it is clear that in the abelian case the two definitions are equivalent. Clearly a subset $E \subset G$ is an IP-set if and only if it is a homomorphic image of $\mathcal{F}$.

LEMMA 3.6: *If T is mildly mixing then any homomorphism $\varphi\colon \mathcal{F} \to G$ has an $\mathcal{F}$-subsequence φ' such that $\mathrm{IP} - \lim \mu(T_{\varphi'(\alpha)}A \cap B) = \mu(A)\mu(B)$ for all measurable A and B.*

Proof: We will use Lemma 4.2 of [F2]. Its proof in the non-abelian setting is identical. By that lemma there is an $\mathcal{F}$-subsequence φ' such that $\mathrm{IP} - \lim U_{\varphi'(\alpha)} = P$ exists in the weak operator topology and is an orthogonal projection. Evidently any function $f \in E = PL_0^2(\mu)$ satisfies $\mathrm{IP} - \lim U_{\varphi'(\alpha)} f = f$ in the weak topology of $\mathcal{H}$ and hence also in the norm topology. This means that f is a rigid function for T so f must be zero. Since $f \in E$ was arbitrary we have $P = 0$ on E, which implies our lemma. ∎

THEOREM 5: *T is mildly mixing if and only if given any IP'-set Σ, the collection of partitions which are independent along some IP'-set $\bar{\Sigma} \subset \Sigma$ is dense in the set of all finite partitions.*

Proof: First suppose Σ is an IP-set so $\Sigma = \{e\} \cup \varphi(\mathcal{F})$ for some homomorphism $\varphi\colon \mathcal{F} \to G$. Choose φ' as in Lemma 3.6 and note that $\varphi'(\mathcal{F}) \subset \Sigma$. Then in particular $\varphi'\{n\}_{n\in\mathbb{N}}$ is a mixing sequence for T. Thus given a partition P we may apply Theorem 1 to yield a subsequence $\{n_i\}$ and a perturbation Q of P such that the iterates $\{gQ : g \in \Sigma'\}$ are independent, where Σ' denotes the IP'-set generated by $\{\varphi'(n_i) : i \in I\}$. Clearly $\mathrm{IP}\{\varphi'(n_i)\} \subset \varphi'(\mathcal{F}) \subset \varphi(\mathcal{F})$ so $\bar{\Sigma} = \mathrm{IP}'\{\varphi'(n_i)\} \subset \Sigma$ and we are done.

For the only if direction suppose T has a rigid set A with $0 < \mu(A) < 1$, say $g_i^{-1}A \to A$. It follows that for any $\epsilon > 0$ we can find a subsequence $\{g_{i_j}\}$ so that $gA \stackrel{\epsilon}{\sim} A$ for all $g \in \Sigma = \mathrm{IP}\{g_{i_j}\}$. If ϵ is sufficiently small (depending on $\mu(A)$) this clearly precludes the possibility of an ϵ perturbation Q of the partition $\{A, A^c\}$ such that $gQ \perp Q$ for even a single g in Σ. ∎

The definition of mixing for T is simply that $\lim_{g\to\infty} \mu(g^{-1}A \cap B) = \mu(A)\mu(B)$ so the following theorem is an immediate consequence of Theorem 1.

THEOREM 6: *T is mixing if and only if given any infinite set Σ, the partitions which are independent along some IP'-set with generators in Σ are dense in the set of all finite partitions.*

We remark that in Theorem 4, 5 or 6 the IP′-set generated by $\{g_i\}$ can be replaced by the sequence $\{g_i\}$ itself, thereby strengthening the "if" direction of each theorem.

References

[Be,Ro] V. Bergelson and J. Rosenblatt, *Mixing actions of groups,* Illinois Journal of Mathematics **32** (1988), 65–80.

[Be,Ru] V. Bergelson and D. J. Rudolph, *Weakly mixing actions of F^∞ have infinite subgroup actions which are Bernoulli,* in *Dynamical Systems* (College Park, MD, 1986–87), Lecture Notes in Mathematics **1342**, Springer, Berlin, 1988, pp. 7–22.

[D] H. A. Dye, *On the ergodic mixing theorem,* Transactions of the American Mathematical Society **118** (1965), 123–130.

[F1] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions,* Journal d'Analyse Mathématique **31** (1977), 204–256.

[F2] H. Furstenberg, *IP-systems in ergodic theory,* in *Conference in Modern Analysis and Probability* (Richard Beals, Anatole Beck, Alexandra Bellow and Arshag Hajian eds.), New Haven, Conn., 1982, Contemporary Mathematics **26**, American Mathematical Society, Providence, RI, 1984, pp. 131–148.

[F,K] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for IP-systems and combinatorial theory,* Journal d'Analyse Mathématique **45** (1985), 117–168.

[J,R,W] A. del Junco, K. Reinhold and B. Weiss, *Partitions with independent iterates along IP-sets,* Ergodic Theory and Dynamical Systems **19** (1999), 447–473.

[K] U. Krengel, *Weakly wandering vectors and weakly independent partitions,* Transactions of the American Mathematical Society **164** (1972), 199–226.

[O] D. S. Ornstein, *Ergodic Theory, Randomness and Dynamical Systems,* Yale Mathematical Monographs, No. 5, Yale University Press, New Haven, Conn.-London, 1974.

[O,W] D. S. Ornstein and B. Weiss, *Entropy and isomorphism theorems for actions of amenable groups,* Journal d'Analyse Mathématique **48** (1987), 1–141.

[S] K. Schmidt, *Asymptotic properties of unitary representations and mixing,* Proceedings of the London Mathematical Society (3) **48** (1984), 445–460.

[S,W] K. Schmidt and P. Walters, *Mildly mixing actions of locally compact groups,* Proceedings of the London Mathematical Society (3) **45** (1982), 506–518.