A GENERIC DISTAL TOWER OF ARBITRARY COUNTABLE
HEIGHT OVER AN ARBITRARY INFINITE ERGODIC SYSTEM

ELI GLASNER AND BENJAMIN WEISS

Abstract. We show the existence, over an arbitrary infinite ergodic
Z-dynamical
system, of a generic ergodic relatively distal extension of arbitrary countable rank
and arbitrary infinite compact extending groups (or more generally, infinite quo-
tients of compact groups) in its canonical distal tower.

Contents

Introduction
1
1. Some preliminaries and structure theory
3
2. The strategy of the proof of Theorem 0.1
6
3. A generic cocycle is ergodic
8
4. The proofs of Theorems 2.6 and 2.7
10
5. Some corollaries
20
6. The case of a weakly mixing extension
22
7. A general framework and a master theorem
25
References
27

Introduction

It would be hard to exaggerate the importance and impact of Harry Furstenberg’s
1963 paper “The structure of distal flows”, [3]. In this revolutionary work Furstenberg
started what we call today the “structure theory” of dynamical systems. We recall
that a topologically distal dynamical system (X, T), with X a metric compact space
and T a self homeomorphism, is called distal if the only proximal pairs in X are the
diagonal pairs; i.e. if \( \lim T^{n_i}x = \lim T^{n_i}x' \) for a pair \( x, x' \in X \) and a sequence \( n_i \in \mathbb{Z} \),
then \( x = x' \). Furstenberg’s distal structure theorem asserts that every minimal distal
system (X, T) has, uniquely, a structure of an inverse limit of a family of factors
\( \{ (X_\alpha, T) : \alpha < \eta \} \) directed by a countable ordinal \( \eta \) such that for every \( \alpha < \eta \) the
extension \( X_{\alpha+1} \to X_\alpha \) is a maximal topologically isometric extension.

Whereas in [3] the subject of study is that of “minimal flows”, so in the domain of
topological dynamics, the works [12], [13] and [4] and [5] introduce and prove an anal-
ogous theorem in the context of ergodic theory, called today the Furstenberg-Zimmer
structure theorem for ergodic systems, which is the main tool for Furstenberg’s er-
godic version of Szemerédi’s theorem, [10].

Date: May 14, 2020.
2010 Mathematics Subject Classification. Primary 3720, 37B35.
Key words and phrases. distal systems, structure theory, cocycles.
Roughly speaking, an extension $X \to Y$ of ergodic dynamical systems is \textit{compact} when $X$ is a skew product over $Y$ with fibers all of which have the form of a homogeneous space $K/H$, where $K$ is a compact group and $H < K$ a closed subgroup. In the special case when $H$ is trivial the extension is called a \textit{group extension} and in this case $K$ consists of a compact group of automorphisms of the system $X$ with $Y \cong X/K$. In fact, it turns out that every compact extension is of this form.

In his works \cite{4} and \cite{5} Furstenberg defines an ergodic measure theoretical system $X$ to be \textit{distal} if it is obtained as an iteration of countably many compact extensions, where in the (possibly transfinite construction) at a limit ordinal one takes an inverse limit.

Now W. Parry in his 1967 paper \cite{9} suggested an intrinsic definition of measure distality. He defines a property of measure dynamical systems, called “admitting a separating sieve”, which imitates the intrinsic definition of topological distality as follows:

Let $X = (X, \mathcal{X}, \mu, T)$ be an ergodic system. A sequence $A_1 \supset A_2 \supset \cdots$ of sets in $X$ with $\mu(A_n) > 0$ and $\mu(A_n) \to 0$, is called a \textit{separating sieve} if there exists a subset $X_0 \subset X$ with $\mu(X_0) = 1$ such that for every $x, x' \in X_0$, the condition “for every $n \in \mathbb{N}$ there exists $k \in \mathbb{Z}$ with $T^kx, T^kx' \in A_n$” implies $x = x'$.

In 1976 in two fundamental papers \cite{12}, \cite{13} R. Zimmer developed the theory of distal systems and distal extensions for a general locally compact acting group. He showed that, as in the topologically distal case, systems admitting Parry’s separating sieve are exactly those with generalized discrete spectrum, that is those systems which are exhausted by their Furstenberg tower of compact extensions.

An extension of dynamical systems $\pi : X \to Y$ is called a \textit{relatively weakly mixing extension} when the corresponding relative product $(X \times X, \mu \times \nu, T)$ is ergodic. In particular $X$ is \textit{weakly mixing} when the product system $X \times X$ is ergodic.

The Furstenberg-Zimmer structure theorem says that every ergodic dynamical system has a unique structure as a relative weakly mixing extension of a distal system, and that the latter admits a uniquely defined canonical distal tower. The so called \textit{canonical distal tower} is unique if at each stage one takes the maximal compact extension (within $X$). The height of this tower (a countable ordinal) is called the \textit{rank} of the distal system $X$.

In \cite{2} Beleznay and Foreman show that for every countable ordinal $\eta$ there is an ergodic distal system of rank $\eta$. Our main result in the present work is as follows.

0.1. \textbf{Theorem.} \textit{Given an arbitrary countable ordinal $\eta$ and a transfinite sequence of pairs $\{(K_\alpha, H_\alpha) : \alpha = 0, \text{ and } \alpha < \eta, \text{ } \alpha \text{ a successor ordinal}\}$, where for each $\alpha$ $K_\alpha$ is an infinite compact second countable topological group and $\{e\} \leq H_\alpha < K_\alpha$, a proper closed subgroup of infinite index in $K_\alpha$, with the only requirement that $K_0$ be an infinite monothetic compact group (and $H_0 = \{e\}$), there exists (generically) an ergodic distal system $X$ of rank $\eta$ such that, in its canonical distal tower, for each successor ordinal $\alpha$ the extension $X_\alpha \to X_{\alpha-1}$ is a $K_{\alpha-1}/H_{\alpha-1}$-extension (here $X_0$ is the trivial one point system).}
Note that in [2] the authors prove their result via an explicit inductive construction on the infinite torus $\mathbb{T}^\mathbb{N}$, so that at each stage the extending compact group is $\mathbb{T}$, whereas in our construction the compact groups (and more generally also their quotients), which serve as building blocks for the tower, are arbitrary. Moreover, we show that at each successor ordinal the construction yields a generic extension.

We also note that Theorem 0.1 is in fact the best result one can prove regarding the Furstenberg-Zimmer structure theorem, since the requirement that all the groups $K_\alpha$ be infinite is really necessary, see Remark 4.2 below.

The relative version of the Furstenberg-Zimmer theorem says that for any given extension of ergodic systems $Y \to Z$, there is a diagram $Y \to Y_{rd} \to Z$, where $Y_{rd}$ is the largest relative distal extension of $Z$ in $Y$, with a uniquely defined canonical relatively distal tower, and such that the extension $X \to Y_{rd}$ is relatively weakly mixing. As Theorems 4.4 and 4.9 of the present work are proven over an arbitrary infinite ergodic system $Z$, it follows that our proof of Theorem 0.1 works in the relative case as well, producing a canonical relative distal tower of height $\eta$ over $Z$ (in fact, in view of Theorem 3.1, when $Z$ is infinite ergodic, the assumption on $K_1$ can be relaxed, we only need it to be infinite).

0.2. Theorem. Let $Z$ be an infinite ergodic system. Given an arbitrary countable ordinal $\eta$ and a transfinite sequence of pairs
\[
\{(K_\alpha, H_\alpha) : \alpha = 0, \text{ and } \alpha < \eta, \alpha \text{ a successor ordinal}\},
\]
where for each $\alpha$ $K_\alpha$ is an infinite compact second countable topological group and $\{e\} \leq H_\alpha < K_\alpha$, a proper closed subgroup of infinite index in $K_\alpha$, there exists (generically) an ergodic system $X$ which is relatively distal over $Z$ of rank $\eta$ such that, in its canonical distal tower, for each successor $\alpha$ the extension $X_\alpha \to X_{\alpha-1}$ is a $K_{\alpha-1}/H_{\alpha-1}$-extension (here $X_0$ is the system $Z$).

After a preliminary section, where we introduce the basic definitions (Section 1), we describe in Section 2 the strategy of the proof of the main theorem. Then, as a preliminary result, whose proof will indicate for the reader an essential trait of the general strategy, we show in Section 3, that given an ergodic system $X$ and a compact topological group $G$, the generic cocycle $\phi : X \to G$ induces an ergodic skew product extension $X_\phi \to X$. This is a generalization of a theorem of Jones and Parry [7] where the authors proved this result for an abelian $G$.

We then go on with the main proof, by stages, in Section 4. In Section 5 we draw some corollaries of Theorem 0.1. In Section 6 we prove an analogous statement about generic group extensions over a weakly mixing system (and more generally over a relatively weakly mixing extension). Finally, in the last section (Section 7) we present a general framework for our results and prove a master theorem of which most of our main results are consequences.

1. SOME PRELIMINARIES AND STRUCTURE THEORY

A dynamical system (sometimes also called a $\mathbb{Z}$-action) is a quadruple $X = (X, \mathcal{X}, \mu, T)$, where $(X, \mathcal{X}, \mu)$ is a standard probability space $T$ is an element of the Polish group $\text{Aut}(X, \mu)$ of invertible measure preserving transformation of $(X, \mathcal{X}, \mu)$. When $X$ and
A function \( \tilde{\varphi} : \mathbb{X} \to \mathbb{Y} \) is a factor of \( \mathbb{X} \) (or that \( \mathbb{X} \) is an extension of \( \mathbb{Y} \)) if there is a measurable map \( \pi : \mathbb{X} \to \mathbb{Y} \) such that \( \pi_\ast(\mu) = \nu \) and such that \( \pi(Tx) = T\pi(x) \) for \( \mu \) almost every \( x \in \mathbb{X} \). The map \( \pi \) is called a factor map (or an extension).

The system \( \mathbb{X} \) is ergodic if every \( T \)-invariant set \( A \in \mathbb{X} \) (i.e. \( TA = A \) (mod \( \mu \)) is trivial : \( \mu(A)(1 - \mu(A)) = 0 \)).

Let \( \mathbb{Y} \) be a dynamical system and \( (V, \mathcal{V}, \rho) \) a standard probability space. Let \( S \mapsto S_y \) be a measurable map \( Y \to \text{Aut}(V, \rho) \); then \( S \) defines a cocycle, i.e. a function \( S : \mathbb{Z} \times Y \to \text{Aut}(V, \rho) \),

\[
S(n, y) = \begin{cases} 
S_{T^n-1} \circ \cdots \circ S_{T_y} \circ S_y & \text{for } n \geq 1 \\
id & \text{for } n = 0 \\
S_{T^{-n}} \circ \cdots \circ S_{T_y}^{-1} & \text{for } n < 0.
\end{cases}
\]

We define the skew-product system \( \mathbb{Y} \times_S (V, \rho) \) to be the system \( (Y \times V, \mathcal{V} \otimes \mathcal{V}, \mu \times \rho, T_S) \), where \( T_S(y, v) = (Ty, S_y(v)) \).

In the special case where \( V \) is a compact group and \( \rho \) is its normalized Haar measure, any measurable function \( \phi : Y \to V \) defines a skew product by the formula:

\[
T_\phi(y, v) = (Ty, \phi(y)v), \quad y \in Y, v \in V, \quad \text{and} \\
T_\phi^n(y, v) = (T^n y, \phi_n(y)v), \quad n \in \mathbb{Z}.
\]

Here \( \phi(n, y) = \phi_n(y) = \phi(T^{n-1}y) \cdots \phi(Ty) \cdot \phi(y) \) for \( n > 0 \) and a similar formula for \( n < 0 \).

We have the following basic theorem:

1.1. **Theorem (Rokhlin).** Let \( \pi : \mathbb{X} \to \mathbb{Y} \) be a factor map of dynamical systems with \( \mathbb{X} \) ergodic, then \( \mathbb{X} \) is isomorphic to a skew product over \( \mathbb{Y} \). Explicitly, there exist a standard probability space \( (V, \mathcal{V}, \rho) \) and a measurable map \( S : Y \to \text{Aut}(V, \rho) \) with \( \mathbb{X} \cong Y \times_S (V, \rho) \) = \( (Y \times V, \mathcal{V} \otimes \mathcal{V}, \nu \times \rho, T_S) \), where \( T_S(y, u) = (Ty, S_y(u)) \), and \( \pi(y, v) = y \).

The map \( y \mapsto S_y \) is called the Rokhlin cocycle of the extension \( \pi \).

The topology on \( \text{Aut}(X, \mu) \) is induced by a complete metric

\[
D(S, T) = \sum_{n \in \mathbb{N}} 2^{-n}(\mu(SA_n \triangle TA_n) + \mu(S^{-1}A_n \triangle T^{-1}A_n)),
\]

with \( \{A_n\}_{n \in \mathbb{N}} \) a dense sequence in the measure algebra \( (\mathcal{X}, d_\mu) \), where \( d_\mu(A, B) = \mu(A \triangle B) \). Equipped with this topology \( \text{Aut}(X, \mu) \) is a Polish topological group and we say that the dynamical system \( \mathbb{X} \) is compact if the set \( \{T^n : n \in \mathbb{Z}\} \) is a precompact subgroup of \( \text{Aut}(X, \mu) \).

1.2. **Example.** Let \( K \) be a compact monothetic topological group; i.e. there is a homomorphism \( \phi : \mathbb{Z} \to K \) with a dense image, and we let \( Tx = ax, \ x \in X \), where \( a = \phi(1) \) (so that the image of \( \phi \) is the dense subgroup \( \{a^n : n \in \mathbb{Z}\} \)). With \( \mathcal{K} \) the algebra of Borel subsets of \( K \) and \( \lambda \) is the normalized Haar measure on \( K \), the system \( \mathbb{X} = (K, \mathcal{K}, \lambda, T) \), is an ergodic compact dynamical system.
It turns out that, in fact, every ergodic compact system $X$ has this form. The notion of compactness can now be relativized as follows:

An extension $\pi : X \to Y$, with $X$ ergodic is a compact extension if there is a compact second countable topological group $K$, a closed subgroup $H < K$ and a measurable map (sometimes called a cocycle) $\phi : Y \to K$ such that

$$X \cong Y \times_{\phi} (K/H, \rho) = (Y \times K/H, Y \otimes \mathcal{K}, \nu \times \rho, T_{\phi}),$$

where $\rho$ is the Haar measure on $K/H$ and $T_{\phi}(y, kH) = (Ty, \phi(y)kH)$. The cocycle $\phi$ is minimal if there is no cocycle $\psi : \Gamma \times Y \to K$ cohomologous to $\phi$ with $K_{\psi} \subsetneq K_{\phi}$. where $K_{\phi}$ and $K_{\psi}$ are the closed subgroups of $K$ generated by the ranges of $\phi$ and $\psi$ respectively. (The cocycles $\phi$ and $\psi$ are cohomologous when there is a measurable map $\kappa : Y \to K$ such that $\psi(Ty) = \kappa(Ty)^{-1}\phi(y)\kappa(y)$, $\nu$-a.e.)

1.3. Theorem. Given a compact extension $\pi : X \to Y$ with $X$ ergodic, we can always assume that

$$X \cong Y \times_{\phi} (K/H, \rho) = (Y \times K/H, Y \otimes \mathcal{K}, \nu \times \rho, T_{\phi}),$$

where the cocycle $\phi$ is minimal with $K_{\phi} = K$. The corresponding group extension $\tilde{\pi} : \tilde{X} \to Y$, with $\tilde{X} = Y \times K$, is ergodic and the diagram

$$\begin{array}{ccc}
\tilde{X} &=& \overline{\times}_{\phi} K \\
\downarrow\quad \sigma & & \downarrow \pi \\
Y &=& \overline{\times}_{\phi} K/H \\
\end{array}$$

commutes. Here $\tilde{X} = Y \times_{\phi} K$ is the group skew-product defined by the cocycle $\phi$, i.e. $\tilde{\mu} = \nu \times \tilde{\rho}$ where $\tilde{\rho}$ is Haar measure on $K$, and $\tilde{\mu}$ is ergodic. The map $\sigma : \tilde{X} \to X$ is the quotient map $\sigma(y, k) = (y, kH)$.

For the proof and more details see e.g. [6, Corollary 3.27].

This construction can be iterated and a dynamical system $X$ is called distal if it is an iteration of countably many compact extensions, where in the (possibly transfinite construction) at a limit ordinal one takes an inverse limit. The so called canonical distal tower is unique if at each stage one takes the maximal compact extension (within $X$). The height of this tower (a countable ordinal) is called the rank of the distal system $X$.

An extension of dynamical systems $\pi : X \to Y$ is called a weakly mixing extension when the corresponding relative product $(X \times X, \mu \times \mu, T)$ is ergodic. In particular $X$ is weakly mixing when the product system $X \times X$ is ergodic.

We now can state the following

1.4. Theorem (Furstenberg-Zimmer structure theorem). Every ergodic system $X$ has (uniquely) a largest distal factor $\pi : X \to Y$ and the extension $\pi$ is a weakly mixing one.
In [2] Beleznay and Foreman show that for every countable ordinal \( \eta \) there is an ergodic distal system of rank \( \eta \). We refer e.g. to [6] for more details on structure theory in ergodic theory.

## 2. The strategy of the proof of Theorem 0.1

We begin with a few comments on Theorem 0.1.

### 2.1. Remarks.

1. Whereas in [2] the authors prove their result for the special case where \( K_\alpha = T = \mathbb{R}/\mathbb{Z} \) for all \( \alpha \), in our construction the compact groups (or their quotients) which serve as building blocks for the tower are arbitrary.

2. Furthermore, our constructions yield generic extensions at each successor ordinal.

3. Note however that, at the first stage of the tower, for the infinite monothetic \( K_0 \), the set of topological generators \( K_g = \{ k \in K_0 : \{ k^n : n \in \mathbb{Z} \} = K_0 \} \) is a dense \( G_\delta \) subset of \( K_0 \) iff \( K_0 \) does not admit a nontrivial finite quotient group.

**Proof of the latter remark.** Let \( \{ U_m \}_{m \in \mathbb{N}} \) be a basis for the topology of \( K_0 \). For each \( m \) set

\[ U_m = \{ k \in K_0 : \exists n \in \mathbb{N}, k^n \in U_m \}. \]

Clearly each \( U_m \) is open, and \( K_g = \bigcap_{m \in \mathbb{N}} U_m \). Thus \( K_g \) is always a \( G_\delta \) set. We assume that \( K_0 \) is monothetic so it has at least one generator, say \( k_0 \); so that \( K_g \) is nonempty.

Now if for each \( 0 \neq n \in \mathbb{Z} \), \( k_0^n \) is a topological generator, then \( \{ k_0^n : n \in \mathbb{Z} \} \) is a subset of \( K_g \) and it follows that \( K_g \) is a dense \( G_\delta \) subset of \( K_0 \). Otherwise, there is \( t \geq 2 \) such that the subgroup \( N = \{ k_0^n : n \in \mathbb{Z} \} \) is a proper subgroup of finite index \([N,K_0] = t\). So clearly in this case \( K_g \) is not dense. \( \square \)

### 2.2. Definition.

An extension \( X \to Y \) of dynamical systems (not necessarily ergodic) is **relatively ergodic**, or that \( X \) is relatively ergodic over \( Y \), if every invariant \( L_2(\mu) \) function is \( Y \) measurable.

In the sequel we will repeatedly use the following lemma which is explicitly formulated in [2, Lemma 2.8]. The authors of [2] base their proof on the characterization of compact extensions in [5, Theorem 6.13]. We give here a brief proof based on [4, Theorem 7.1].

### 2.3. Lemma.

Let \( X \) be an ergodic system. Let \( Z \) be a factor of \( X \) and \( Y \) be a compact extension of \( Z \) in \( X \) (i.e. \( X \to Y \to Z \)). Then \( Y \) is the maximal compact extension of \( Z \) in \( X \) iff \( X \times X \) is relatively ergodic over \( Y \times Y \).

**Proof.** Let \( \hat{Y} \) be the maximal compact extension of \( Z \) in \( X \), so that we have the diagram \( X \to \hat{Y} \to Y \to Z \).

Suppose first that the map \( \hat{Y} \to Y \) is not an isomorphism. As this map is a compact extension we can represent \( \hat{Y} \) as a skew product over \( Y \):

\[ \hat{Y} = Y \times_{\phi} K/H, \]
with $K$ a compact group and $H < K$ a closed subgroup and $\phi : Y \to K$ a measurable cocycle. Then, the extension $\bar{Y} \times \bar{Y} \to Y \times Y$ is not relatively ergodic. Indeed, above any ergodic component $W \subset Y \times Y$, we have a nontrivial ergodic decomposition which corresponds to the ergodic decomposition of the diagonal action of $K$ on $K/H \times K/H$. Now, a fortiori, the extension $X \times X \to Y \times Y$ is not relatively ergodic.

For the other direction assume that $\bar{Y} = Y$ and let $f$ be an invariant function in $L_2(X \times X)$. By [4, Theorem 7.1] the function $f$ is a member of the Hilbert space $L_2(\bar{Y} \times \bar{Y}) = L_2(Y \times Y)$. □

2.4. Lemma. Let $W \to X \to Y \to Z$ be a tower of extensions of ergodic systems such that $Y \to Z$ is the maximal compact extension of $Z$ within $X$, and $X \to Y$ is the maximal compact extension of $Y$ within $W$. Then $Y \to Z$ is the maximal compact extension of $Z$ within $W$.

Proof. If the extension $Y \to Z$ is not the maximal compact extension of $Z$ within $W$, then there are functions $f_1, f_2, \ldots, f_k \in L_2(W)$ such that the finite dimensional $L_\infty(Z)$ module $L(f_1, \ldots, f_k) = \{h_1 f_1 + h_2 f_2 + \cdots + h_k f_k : h_i \in L_\infty(Z), i = 1, 2, \ldots, k\}$ is $\Gamma$ invariant, and $L \not\subset L_2(Y)$. Since $L_\infty(Z) \subset L_\infty(Y)$ it follows that $L$ is also a $\Gamma$ invariant finite dimensional $L_\infty(Y)$ module, whence, by the maximality of $Y$ in $W$, we have $L \subset L_2(X)$. Now the maximality of $Y$ in $X$ implies that the $L_\infty(Z)$ module $L$ is in fact a subset of $L_2(Y)$ and this contradiction proves our claim. □

2.5. Lemma. Let $\alpha$ be a countable limit ordinal. Let $X$ be an ergodic distal $\mathbb{Z}$-system built as a tower of height $\alpha$ consisting of group extensions and inverse limits, such that for each ordinal $\beta < \alpha$ the extension $X_{\beta+1} \to X_\beta$ is the maximal compact extension of $X_\beta$ within $X_{\beta+2}$, then for each $\beta < \alpha$ each extension $X_{\beta+1} \to X_\beta$ is the maximal compact extension of $X_\beta$ within $X$.

Proof. In view of Lemma 2.3 what we have to show is that, for each $\beta < \alpha$, the extension $X \times X_\beta \times X_{\beta+1} \to X_{\beta+2}$ is ergodic. Now this extension is an inverse limit of the extensions $X_\eta \times X_\eta \to X_{\beta+1} \times X_{\beta+1}$, where the ordinal $\eta$ ranges over the interval $\beta < \eta < \alpha$. Now applying transfinite induction, using Lemma 2.4 and the fact that an inverse limit of ergodic extensions is an ergodic extension, we conclude the proof. □

In view of the above three lemmas we conclude that in order to prove Theorem 0.1 we only need to prove the following two statements:

2.6. Theorem. Let $\alpha$ be a countable successor ordinal. Let $X$ be an ergodic distal $\mathbb{Z}$-system of rank $\alpha$, where in the canonical tower the final extension $X = X_\alpha \to X_{\alpha-1}$ is a compact extension and let $G$ be a compact second countable group. Then, for a
generic function \( \phi : X \to G \), the corresponding skew-product \((X \times G, T_\phi)\) is an ergodic compact extension of \(X\) which is distal of rank \(\alpha + 1\).

2.7. **Theorem.** Let \(\alpha\) be a countable limit ordinal. Let \(X\) be an ergodic distal \(\mathbb{Z}\)-system of rank \(\alpha\), and let \(G\) be a compact second countable group. Then, for a generic function \(\phi : X \to G\), the corresponding skew-product \((X \times G, T_\phi)\) is an ergodic compact extension of \(X\) which is distal of rank \(\alpha + 1\).

### 3. A GENERIC COCYCLE IS ERGODIC

As a warm up let us first prove a simpler statement which generalises a theorem of Jones and Parry [7], where the authors deal with the case where the extending group \(G\) is a compact second countable abelian group.

So for now let \(Y = (Y, y, \mu, T)\) be an ergodic system, \(G\) a compact second countable topological group, \(\lambda_G\) its normalized Haar measure and \(d_G\) a bi-invariant metric on \(G\). Let \(\mathcal{C} = \mathcal{C}(Y, G)\) be the space of Borel maps \(\phi : Y \to G\), where we identify \(\phi\) and \(\psi\) if they agree \(\mu\)-a.e. We equip \(\mathcal{C}\) with a metric \(d\) as follows:

\[
d(\phi, \psi) = \inf \{ \epsilon : \mu \{ y \in Y : d_G(\phi(y), \psi(y)) > \epsilon \} < \epsilon \}.\]

1. **Claim.** The function \(d\) above defines a complete metric on the space \(\mathcal{C}(Y, G)\). With the induced topology \(\mathcal{C}(Y, G)\) is second countable; i.e. it is a Polish space.

**Proof.** We only check that \(d\) satisfies the triangle inequality. Let \(d(\phi, \psi) = \epsilon_1, d(\psi, \rho) = \epsilon_2\). By definition there are sequences \(\epsilon_i \searrow \epsilon, \eta_i \searrow \eta\) such that

\[
\mu(\{ y : d_G(\phi(y), \psi(y)) > \epsilon_i \}) < \epsilon_i,
\]

\[
\mu(\{ y : d_G(\psi(y), \rho(y)) > \eta_i \}) < \eta_i.
\]

Then

\[
\{ y : d_G(\phi(y), \rho(y)) > \epsilon_i + \eta_i \} \subset \{ y : d_G(\phi(y), \psi(y)) > \epsilon_i \} + \{ y : d_G(\psi(y), \rho(y)) > \eta_i \},
\]

hence

\[
\mu(\{ y : d_G(\phi(y), \rho(y)) > \epsilon_i + \eta_i \}) \leq \mu(\{ y : d_G(\phi(y), \psi(y)) > \epsilon_i \}) + \mu(\{ y : d_G(\psi(y), \rho(y)) > \eta_i \}) < \epsilon_i + \eta_i.
\]

Therefore

\[
d(\phi, \rho) \leq \inf(\epsilon_i + \eta_i) = d(\phi, \psi) + d(\psi, \rho).
\]

\(\square\)

Recall that the **finite full group** of the system \(Y\), denoted by \([T]_f\), is defined as the group of invertible measure preserving transformations \(\tau = \tau_{\sigma, \sigma}\) of the probability space \((Y, \mu, \mu)\) for which there is a finite measurable partition \(\mathcal{P} = \{P_1, \ldots, P_n\}\) of \(Y\) and a function \(\sigma : Y \to \mathbb{Z}\) so that for every \(j\), \(\sigma \upharpoonright P_j = s_j\) is a constant, and \(\tau \upharpoonright P_j = T^{s_j}\).
With $\phi \in \mathcal{C}$ we associate the skew product transformation $T_\phi : Y \times G \rightarrow Y \times G$ which is defined by

$$T_\phi(y, g) = (Ty, \phi(y)g), \quad y \in Y, g \in G.$$ 

We then see that

$$T^n_\phi(y, g) = (T^n y, \phi_n(y)g),$$

where

$$\phi_n(y) = \begin{cases} 
\phi(T^{n-1}y) \cdots \alpha(Ty)\phi(y) & \text{for } n \geq 1 \\
id & \text{for } n = 0 \\
\phi(T^n y)^{-1} \cdots \phi(T^{-1}y)^{-1} & \text{for } n < 0.
\end{cases}$$

For $\tau = \tau_{\phi, \sigma} \in [T]_f$ we denote

$$\phi_{\tau}(y) = \phi_{\sigma(y)}(y), \quad y \in Y.$$

3.1. Theorem. For a generic $\phi \in \mathcal{C}$ the system $Y_\phi = (Y \times G, \mathcal{Y} \times \mathcal{B}_G, \mu \times \lambda_G, T_\phi)$ is ergodic.

Proof. Fix a subset $C \subset Y$, $\mu(C) > 0$, an element $g$ in $G$, a positive small constant $a$, and a positive constant $c_a$ (which will depend on $a$). Define the set

$$U(C, a, c_a, g) \subset \mathcal{C}$$

as the collection of all the functions $\phi \in \mathcal{C}$ with the following property:

There exists an element $\tau \in [T]_f$ such that:

1. $\tau(C) = C$,
2. $\mu(\{y \in C : d(\phi_{\tau}(y), g) < a\}) > c_a \mu(C)$

3.2. Proposition. For a sufficiently small $c_a$, the set $U = U(C, a, c_a, g)$ is open and dense in $\mathcal{C}$.

Proof. Fix $C$ and $a$. It is easy to check that $U$ is an open set. In order to see that it is dense (for a sufficiently small $c_a$) fix $\phi_0 \in \mathcal{C}$ and $\delta > 0$ (to be determined later on). As finite valued functions are dense in $\mathcal{C}$, we may and will assume that $\phi_0(Y)$ is a finite subset of $G$.

Let $B \subset Y$ be a base for a Rokhlin tower in $Y$ of height $2N+1$, so that $\nu(\bigcup_{i=0}^{2N} T^i B_0) > 1 - \delta$ (again, $N$ will be determined later on). Next purify the $Y$-tower according to $\phi_0$ and $C$; i.e. subdivide $B$ into a finite number of subsets $\{B_j\}_{j=1}^Y$ so that for each $j$, on each level of the column of the tower above $B_j$, the function $\phi_0$ is constant, and each level is either contained in $C$ or in $Y \setminus C$.

By the ergodic theorem, for sufficiently large $N$, but for a set of measure $< \delta$, in the remaining columns of the tower, there is (almost) an equal proportion of $C$-levels in the bottom half and top half of the tower (up to $\delta$). In each of the good columns enumerate the $C$-levels, from the bottom up to level $N - 1$ and from level $N$ to the top. Both in the top half and in the bottom half of the tower there are at least $(1 - \delta)N$ levels contained in $C$ ($C$-levels). The transformation $\tau$ is defined by interchanging pairs of $C$-levels in the lower and upper half of the tower, using the appropriate power of $T$. It is defined to be the identity elsewhere.
When we focus on a particular $C$-level, say $i_0$ in the lower half, that maps to a $C$-level $i'_0$ in the upper half of the column, the value of the cocycle corresponding to $\phi_0$, as we move up the tower to level $N - 1$, and the value of the cocycle as we move from level $N + 1$ up to the level $i'_0$, define a pair $(h_1, h_2) \in G \times G$. As $\phi_0$ has only finitely many values the collection $H$ of all pairs $(h, h') \in G \times G$ so obtained is finite. Thus $\tau$ moves from the bottom to the top at least a $\left(\frac{1}{2} - 10\delta\right)$-fraction of $C$.

Let $V$ be $\frac{a}{10}$-ball around $c$. Let $F = \{f_1, \ldots, f_m\} \subset G$ be a finite set with $G = \bigcup_{i=1}^{m} f_i V$. As our metric on $G$ is bi-invariant, for any pair $(h, h') \in H$ there is an element $f \in F$ with $hfh' \in V$. Note that $m = m(a)$ depends only on $a$. Now divide each central level $T^i B_j$ into $m$ sets $\{D_1^j, \ldots, D_m^j\}$ of equal measure, and change $\phi_0$ to $\phi$ by defining $\phi \upharpoonright D_m^j \equiv f_i$, $1 \leq l \leq m$. We can now check and see that, with $c_a = \frac{1}{m} \left(\frac{1}{2} - 10\delta\right)$, we have

$$\mu\left(\{y \in C : d(\phi_\tau(y), g) < a\}\right) > c_a \mu(C).$$

□

In order to finish the prof of Theorem 3.1 we need to show that for a $\phi$ in a dense $G_0$ subset of $\mathcal{C}$ the corresponding system $Y_\phi = (Y \times G, Y \times \mathcal{B}_G, \mu \times \lambda_G, T_\phi)$ is ergodic.

Given $C, a$ and $g$ as above, we define the set $U = U(C, a, c_a, g) \subset \mathcal{C}$ and by Proposition 3.2 we know that, with a suitable constant $c_a$, it is open and dense in $\mathcal{C}$.

Let $\{C_n\}_{n \in \mathbb{N}}$ be a dense collection in the measure algebra $(Y, \mu)$ and set

$$\mathcal{U}(a, c_a, g) = \bigcap_{n \in \mathbb{N}} U(C_n, a, c_a, g).$$

Note that it is here that we use the uniform bound $\mu(\{y \in C : d(\phi_\tau(y), g) < a\}) > c_a \mu(C)$, for now we have that $\phi \in \mathcal{U}(a, c_a, g)$ implies $\phi \in U(C, a, c_a, g)$ for every positive $C \in Y$.

Next set

$$\mathcal{U}(g) = \bigcap_{n \in \mathbb{N}} \mathcal{U}(\frac{1}{n}, c_\frac{1}{n}, g),$$

and finally let

$$\mathcal{C}_0 = \bigcap_{g \in G_0} \mathcal{U}(g),$$

with $G_0 \subset G$ a countable dense subset of $G$.

We now see that every $g \in G$ is an essential value for each $\phi$ in the dense $G_0$ subset $\mathcal{C}_0$ of $\mathcal{C}$. This is a sufficient (and necessary) condition for $T_\phi$ to be ergodic (see e.g. [1, Page 1287]) and the proof of Theorem 3.1 is complete. □

4. The proofs of Theorems 2.6 and 2.7

We deal first with Theorem 2.7. Thus, our system $X$ is assumed to be ergodic and distal of order $\alpha$, with $\alpha$ a limit ordinal. As the distal tower for $X$ is canonical, for each $\beta < \alpha$, the extension $X_{\beta+1} \to X_\beta$ is the maximal compact extension of $X_\beta$ in $X$. In order to show that, for a generic $\phi : X \to G$, the system $(X \times G, T_\phi)$, denoted
by $X_\phi$, is ergodic of rank $\alpha + 1$, what we have to show, in view of Lemma 2.5, is that for each ordinal $\beta < \alpha$, considering the the extension

$$X_\phi \times X_\phi \to X_{\beta + 1} \times X_{\beta + 1}$$

is relatively ergodic. This situation is isolated in the following setup, where we use the notation $Z = X_\beta$ and $Y = X_{\beta + 1}$.

4.1. **Set up.** Let $K$ and $G$ be two arbitrary second countable compact topological groups with normalized Haar measures $\lambda_K$ and $\lambda_G$ respectively. Let $H \leq K$ be a closed subgroup, with its Haar measure $\lambda_H$. We write $\lambda_{K/H}$ for the normalized Haar measure on the homogeneous space $K/H$. Suppose that $Z = (Z, \theta, T)$ is an ergodic $\mathbb{Z}$-system and that $Y = (Y, \nu, T)$ is the skew product

$$Y \cong Z \times_\gamma (K/H, \lambda_{K/H}) = (Z \times K/H, \mathbb{Z} \otimes \mathcal{B}_{K/H}, \theta \times \lambda_{K/H}, T),$$

where $Y = Z \times K/H$, $\nu = \theta \times \lambda_{K/H}$ and $T = T_\gamma$ for a cocycle $\gamma : Z \to K$ which is minimal with $K_\gamma = K$, as in Theorem 1.3, and such that the system $Y$ is ergodic.

We recall that the corresponding group extension $\hat{\pi} : \hat{Y} \to Z$, with $\hat{Y} = Z \times K$, is ergodic and the diagram

$$\begin{array}{ccc}
\hat{Y} = Z \times_\gamma K & \xrightarrow{\hat{\pi}} & Z \\
\downarrow \hat{\pi} & & \downarrow \pi \\
Z & \xleftarrow{\pi} & Y = Y \times_\gamma K/H
\end{array}$$

commutes, where $\hat{Y} = Z \times_\gamma (K, \lambda_K)$ is the group skew-product defined by the cocycle $\gamma$, with $\hat{\nu} = \theta \times \lambda_K$ and $\hat{\nu}$ is ergodic. The map $\sigma : Y \to Y$ is the quotient map $\sigma(y, k) = (y, kH)$.

Suppose further that $X = (X, \mathcal{X}, \mu, T)$ is an ergodic system, $X = (X, \mathcal{X}, \mu, T) \to Y = (Y, \mathcal{Y}, \nu, T)$ a factor map, such that in the diagram $X \to Y \to Z$, the map $Y \to Z$ is the maximal compact extension of $Z$ in $X$. By Rokhlin’s theorem 1.1 we can represent $X$ as $Y \times V$ where $(V, \rho)$ is a probability space, $\mu = \nu \times \rho$ and the transformation $T : X \to X$ is given by a Rokhlin cocycle $S$ from $Y$ to the Polish group $MPT(V, \rho)$ of invertible measure preserving transformations of $(V, \rho)$, with

$$T(y, v) = (Ty, S_y v), \quad y \in Y, v \in V.$$

(Note that we use $T$ for both $X$ and $Y$.)

We let let $\mathcal{C} = \mathcal{C}(X, G)$ be the Polish space of Borel maps $\phi : X \to G$. For $\phi \in \mathcal{C}$ we set $T_\phi : X \times G \to X \times G$ as :

$$T_\phi(x, g) = (Tx, \phi(x)g), \quad (x \in X, g \in G).$$

We let $X_\phi = (X \times G, \mathcal{X} \times \mathcal{B}_G, \mu \times \lambda_G)$ be the corresponding skew product system.

We need to show that, for a generic $\phi \in \mathcal{C}$ the system $X_\phi \times X_\phi$ is a relatively ergodic extension of $Y \times Y$. The ergodic components of $Y \times Y$ are parametrized by $k_0 H \in K/H$ and may be identified with $Z \times H/H_{k_0}$, with $H_{k_0} = k_0 H k_0^{-1} \cap H$, as follows.
The ergodic components of the system $\hat{Y} \times \hat{Y}$ are parametrized by $k_0 \in K$ and have the form $\{(z, k, kk_0) : z \in Z, k \in K\}$. Under $\sigma$ these are mapped onto the sets $\{(z, kH, kk_0H) : z \in Z, k \in K\}$.

Now for $k_1, k_2 \in K$ we have $(z, k_1 H, k_1 k_0 H) = (z, k_2 H, k_2 k_0 H)$ iff

$$k_1^{-1}k_2 \in H \cap k_0 H k_0^{-1} = H_{k_0},$$

and, in particular, $(z, kH, kk_0H) = (z, H, k_0 H)$ iff $k \in H_{k_0}$. Thus $H_{k_0}$ is the stability group of the point $(H, k_0 H)$ under the left action by $K$. It therefore follows that the correspondence

$$(1) \quad (z, kH, kk_0 H) \longleftrightarrow (z, kH_{k_0})$$

is an isomorphism, with the diagonal action:

$$T(z, kH, kk_0 H) = (Tz, \gamma(z) kH, \gamma(z) kk_0 H) \longleftrightarrow (Tz, \gamma(z) kH_{k_0}).$$

Our assumption on the diagram $X \rightarrow Y \rightarrow Z$ implies that $X \times Z$ is an ergodic extension of $Y \times Z$, and thus for a fixed $k_0$, the diagonal transformation

$$(2) \quad T_{k_0}(z, kH, kk_0 H, v_1, v_2) = (Tz, \gamma(z) kH, \gamma(z) kk_0 H, S(z, kH) v_1, S(z, kk_0 H) v_2),$$

is ergodic. We note that all these sets are subsets of the space $Z \times K/H \times K/H \times V \times V$.

Alternatively, under the correspondence (1)

$$T_{k_0}(z, kH_{k_0}, v_1, v_2) = (Tz, \gamma(z) kH_{k_0}, S(z, kH) v_1, S(z, kk_0 H) v_2).$$

is ergodic. Denoting $L_{k_0} = K/H_{k_0}$, we note that this identifies the ergodic component associated with $k_0$ with $Z \times L_{k_0} \times V \times V$. We write $\mu_{k_0}$ for the product measure $\theta \times \lambda_{L_{k_0}} \times \rho \times \rho$ on this space.

We will use the same letter $\mu_{k_0}$ for the corresponding measure on the space (4.1). Thus in this notation we have

$$\theta \times \lambda_{K/H} \times \lambda_{K/H} \times \rho \times \rho = \int_{K/H} \mu_k d\lambda_{K/H}(k).$$

Note that with this representation the transformations $T_k$, as in (2), are all one and the same as the diagonal action of $T$ and we only change the measure, so that $T = T_k$ is ergodic with respect to $\mu_k$.

What we need to show is that for a generic cocycle $\phi : X \rightarrow G$, the cocycle $\psi = \psi_{k_0}$, defined from $Z \times K/H \times K/H \times V \times V$ to $G \times G$ by

$$\psi(z, kH, kk_0 H, v_1, v_2) = (\phi(z, kH, v_1), \phi(z, kk_0 H, v_2)),$$

is ergodic with respect to $\mu_{k_0}$ for $\lambda_{K/H}$ a.e. $k_0 H \in K/H$. 

4.2. Remark. Suppose $Y \to Z$ is a $K$ extension with $K$ a finite group. Then as we noted in the above discussion the ergodic components of the system $Y \times Y$ are parametrized by $k_0 \in K$ and have the form $\{(z, k, kk_0) : z \in Z, k \in K\}$. It then follows that the ergodic component which corresponds to $k_0 = e$, i.e., the component $W = \{(z, k, k) : z \in Z, k \in K\}$, with diagonal action $T(z, k, k) = (Tz, \gamma(z)k, \gamma(z)k)$ has positive measure in the system $Y \times Y$. But then, no matter which $\phi \in \mathcal{C}(Y, G)$ we choose, the extension $Y_\phi \to Y$ can not be such that

$$Y_\phi \times Y_\phi \to Y \times Y.$$ 

is ergodic since that, in particular means that the ergodic component in the system $Y_\phi \times Y_\phi$, which sits above $W$ with diagonal action

$$T_\phi(z, k, k, g_1, g_2) = (Tz, \gamma(z)k, \gamma(z)k, \phi(z, k)g_1, \phi(z, k)g_2) = (Tz, \gamma(z)k, \gamma(z)k, \phi(z, k), \phi(z, k)g_2(g_1)^{-1})g_1,$$

is $G \times G$ saturated, which is impossible.

This simple observation shows why in our Theorem 0.1, as well as the other main results, we have to require that the extending compact groups (or homogeneous quotient) be infinite.

4.3. Remark. The next theorem is the main tool in the proof of Theorem 2.7. However, we note that whereas in Theorem 2.7 we assume that the system $X$ is distal, here we only need the system $Z$ to be ergodic and infinite.

4.4. Theorem. We are given a chain of factors $X \to Y \to Z$ of the ergodic system $X$ such that $Z$ is infinite, $Y \to Z$ is a $K/H$-extension, and $Y \to Z$ is the maximal compact extension of $Z$ in $X$. Then, for a generic $\phi \in \mathcal{C}(X, G)$ the system $X_\phi = (X \times G, \mathcal{B}_G, \mu \times \lambda_G, T_\phi)$, is ergodic and the extension, $Y \to Z$, is the maximal compact extension of $Z$ in $X_\phi$.

Proof. We define

$$U = U(C, a, b, c_a, g_1, g_2) \subset \mathcal{C}(X, G),$$

where $C$ is a given measurable subset of $Z \times K/H \times K/H \times V \times V$, and $U$ is the set of cocycles $\phi : Z \times K/H \times V \to G$ such that for a set $K_0 \subset K$ with $\lambda_K(K_0) > 1 - b$, and all $k_0 \in K_0$, there is an element $\tau \in [T]_f$ such that

$$\mu_{k_0}((z, kH, kk_0H, v_1, v_2) \in C : d(\phi_{\tau}(z, kH, v_1), g_1) < a \quad \& \quad d(\phi_{\tau}(z, kk_0H, v_2), g_2) < a) > c_a \mu_{k_0}(C)^2. \quad (3)$$

4.5. Proposition. For a sufficiently small $c_a$ the set $U(C, a, b, c_a, g_1, g_2)$ is open and dense in $\mathcal{C}(X, G)$.

Proof. Fix $C, a$ and $b$ and consider the corresponding set $U = U(C, a, b, c_a, g_1, g_2)$. It is easy to check that $U$ is an open set. In fact, given $\phi_0 \in U$, for each $n$ let

$$K_n = \{k_0 \in K_0 : \mu_{k_0}((z, kH, kk_0H, v_1, v_2) \in C : d(\phi_{\tau}(z, kH, v_1), g_1) < a \quad \& \quad d(\phi_{\tau}(z, kk_0H, v_2), g_2) < a) \} > c_a \mu_{k_0}(C)^2 + \frac{1}{n}.\]
Since $K_0 = \bigcup_{n \in \mathbb{N}} K_n$, for some $n_0$, $\lambda_K(K_{n_0}) > 1 - b$. If $\phi \in \mathcal{C}$ is such that

$$\mu_k(\{(z, kH, kk_0H, v_1, v_2) \in C : \phi_0(z, kH, v_1) \neq \phi(z, kk_0H, v_2)\}) < \frac{1}{2n_0},$$

then for $k_0 \in K_{n_0}$

$$\mu_k(\{(z, kH, kk_0H, v_1, v_2) \in C : d(\phi_{r}(z, kH, v_1), g_1) < a \& d(\phi_{r}(z, kk_0H, v_2), g_2) < a\}) > c_a \mu_k(E)^2 + \frac{1}{2n_0},$$

and therefore $\phi \in U$.

In order to see that $U$ is dense (for a sufficiently small $c_a$) fix $\phi_0 \in \mathcal{C}$ and $\delta > 0$. We will show that there is a $\phi \in U$ with $\mu(\{(y, v) \in Y \times V : \phi(y, v) \neq \phi_0(y, v)\}) < \delta$. As finite valued functions are dense in $\mathcal{C}$, we may and will assume that $\phi_0(Y \times V)$ is a finite subset of $G$.

**Step 1: Constructing a Rokhlin tower**

In $Z \times K/H \times K/H \times V \times V$ we will take a Rokhlin tower with base $B = B_0 \times K/H \times K/H \times V \times V$ measurable with respect to $Z$, which is a factor of all the ergodic components $\mu_k$, of height $2N + 1$ with $N > \frac{1}{100}$, so that, for every $k \in K$ we have $\mu_k(\bigcup_{j=0}^{2N} T^n_k B) > 1 - \frac{4}{100}$.

Next purify the $Z \times K/H \times K/H \times V \times V$-tower according to $\phi_0$ and $C$; i.e. subdivide $B$ into a finite number of subsets $\{B_j\}_{j=1}^J$ so that for each $j$, on each level of the column of the tower above $B_j$, the function $\phi_0$ is constant, and each level is either contained in $C$ or in $Y \setminus C$. Of course the atoms of the purified tower are no longer $K/H \times K/H \times V \times V$-saturated. We call the columns $\{T^i B_j\}_{i=0}^{2N} \cup \bigcup_{j=N+1}\{T^j_k(B)\}$, the *pure columns*.

Now for almost every $k \in K$ the transformation $T_k$ is ergodic and hence, for some $n_0(k)$, if $N \geq n_0(k)$ we will have that, after purifying the tower with respect to $C$, in most of the pure columns there will be a proportion of roughly $\mu_k(E)N$ of the levels contained in $C$ in both the lower and upper half of the tower. Since the family $\{T_k : k \in K\}$ is measurable with respect to $k$, the index $n_0(k)$ is a measurable function of $k$ and consequently there will be a set $K_0 \subset K$, with $\lambda_K(K_0) > 1 - b/10$, for which this will hold for a fixed $N$ and for all $k \in K_0$. Note that this purification, and therefore also the sets $B_j \subset B$ depend on $k$.

In order to define $\tau$ we will use a random permutation on $\{0, 1, \ldots, N - 1\}$ to interchange the first $N$-levels of the tower with the last $N$ levels as follows. Let $\Omega$ be the sample space of such a random permutation; that is $\Omega = \text{Sym}(N)$ with the uniform measure, and we use $\pi \in \Omega$ to define an involution $\tau$ between

$$\bigcup_{j=0}^{N-1} T_k^j(B) \quad \text{and} \quad \bigcup_{j=N+1}^{2N} T_k^j(B)$$

by mapping $T_k^j(B)$ to $T_k^{\pi(j)+N+1}(B)$ via $T_k^{N+1-j+\pi(j)}$. This $\tau$ is clearly in $[T]_f$.

Set $\gamma_k = \mu_k(C)$. We will show that there is a choice of a random permutation such that the corresponding $\tau$ will map a $\gamma_k$-proportion of $C \cap \bigcup_{j=0}^{N-1} T_k^j(B))$ to $C$, for $k$ in a set $K_1 \subset K_0$ with $\lambda_K(K_1) > 1 - b/20$. To do this we need two lemmas.
Step 2: A lemma about random permutations

4.6. Lemma. For any $0 < \gamma < 1$

$$\Pr\{\pi \in \text{Sym}(N) : |\{i \leq \lfloor \gamma N \rfloor : \pi(i) \leq \lfloor \gamma N \rfloor\}| > \frac{1}{2} \gamma^2 N\} > 1 - \frac{10(1 - \gamma)}{N \gamma^2}.$$  

Proof. Define random variables $Y_i$ for $1 \leq i \leq M := \lfloor \gamma N \rfloor$ by

$$Y_i(\pi) = \begin{cases} 1 & \text{if } \pi(i) \leq M \\ 0 & \text{if } \pi(i) > M. \end{cases}$$

Clearly $\Pr(Y_i = 1) = \varrho := \frac{M}{N}$, and for $i \neq j$

$$\Pr(Y_i Y_j = 1) = \frac{M(M-1)}{N^2}.$$  

We now have

$$\mathbb{E}\left(\left(\frac{1}{M} \sum_{i=1}^{M} (Y_i - \varrho)\right)^2\right) = \frac{1}{M^2} \sum_{i=1}^{M} \mathbb{E}\left((Y_i - \varrho)^2\right) + \frac{1}{M^2} \sum_{i \neq j} \mathbb{E}\left((Y_i - \varrho)(Y_j - \varrho)\right)$$

$$= \frac{\varrho(1 - \varrho)}{M} + \frac{1}{M^2} \sum_{i \neq j} \left(\frac{M(M-1)}{N^2} - \varrho^2\right)$$

$$= \frac{\varrho(1 - \varrho)}{M} - \frac{1}{M^2} \cdot \frac{M}{N^2} < \frac{\varrho(1 - \varrho)}{M}.$$  

It follows by Chebytchef’s inequality that

$$\Pr\left(\left|\frac{1}{M} \sum_{i=1}^{M} (Y_i - \varrho)\right| > \frac{\varrho}{3}\right) \leq \frac{9}{\varrho^2} \cdot \frac{\varrho(1 - \varrho)}{M}$$

and thus

$$\Pr\left(\frac{1}{M} \sum_{i=1}^{M} Y_i - \varrho \geq -\frac{\varrho}{3}\right) \geq 1 - \frac{9}{\varrho^2} \cdot \frac{\varrho(1 - \varrho)}{M},$$

or

$$\Pr\left(\sum_{i=1}^{M} Y_i \geq \frac{2}{3} \varrho M\right) \geq 1 - \frac{9}{\varrho} \cdot \frac{1 - \varrho}{M} \geq 1 - \frac{10(1 - \gamma)}{\gamma^2 N}.$$  

□

Step 3: The $\sqrt{\delta}$ to $\delta$ lemma

4.7. Lemma. Let $\mathcal{P} = \{P_i : i \in I\}$ be a finite partition of a probability space $(\Omega, \mu)$.  

Let $E \subset \Omega$ with $\mu(E) < \delta$, then for

$$I_0 = \{i : \frac{\mu(P_i \cap E)}{\mu(P_i)} > \sqrt{\delta}\},$$

we have then $\sum_{i \in I_0} \mu(P_i) > 1 - \sqrt{\delta}$. 

**Proof.** For \( i \in I_0 \), we have \( \mu(P_i \cap E) > \sqrt{\delta} \mu(P_i) \), whence
\[
\delta > \sum_{i \in I_0} \mu(P_i \cap E) > \sqrt{\delta} \sum_{i \in I_0} \mu(P_i),
\]
so that \( \sqrt{\delta} > \sum_{i \in I_0} \mu(P_i) \).

\[\square\]

**Step 4: Choosing a good \( \tau \)**

We can now apply Lemma 4.6 to each fixed pure column based on \( B_j \), where there are at least a \( \frac{\gamma_{k_0}}{10} \)-proportion of the \( C \)-level in both the upper and lower halves of the column. For such pure column, which will fill most of the tower, with probability \( \geq 1 - O(\frac{1}{N}) \) the random \( \tau^\omega \) will map at least a \( \frac{1}{3} \frac{\gamma_{k_0}^2}{\sqrt{N}} \)-proportion of \( C \) to \( C \), in that pure column.

In the product space \( \Omega \times B \) define the set \( E \) to be the set of pairs \((\omega, (z, kH, kk_0H, v_1, v_2))\) such that \( \tau^\omega \) fails the requirement above in the fiber \( \{T^k_{k_0}(z, kH, kk_0H, v_1, v_2)\}_{l=0}^{2N} \), and normalize the measure on \( B \) to be a probability measure. Then we apply Lemma 4.7, with \( \delta = O(\frac{1}{N}) \) and \( \mathcal{P} \)-the partition of \( \Omega \) into points, to get a set \( \Omega_0 \subset \Omega \) with probability \( \geq 1 - O(\frac{1}{\sqrt{N}}) \) such that for each \( \omega \in \Omega_0 \) the random \( \tau^\omega \) maps \( \frac{1}{3} \frac{\gamma_{k_0}^2}{\sqrt{N}} \)-proportion \( C \) to \( C \), but for a \( O(\frac{1}{\sqrt{N}}) \)-proportion of the pure columns.

This was done for a fixed \( k_0 \in K \), i.e. the set \( \Omega_0 \in \Omega \) depends on \( k_0 \), since the partition of \( B \) into pure columns depends on \( k_0 \). Now we take this \( O(\frac{1}{\sqrt{N}}) \) to be less than \( \frac{b}{100} \) and applying Fubini’s theorem to \( \Omega \times K \), we see that there is a random \( \omega \) and a set \( K_2 \subset K \) with \( \lambda_K(K_2) > 1 - \frac{b}{20} \), such that \( \tau^\omega \) is good for all \( k \in K_2 \). We let \( K_0 = K_1 \cap K_2 \) (where \( K_1 \) is the subset of \( K \) defined in Step 1) so that \( \lambda_K(K_0) \geq 1 - \frac{b}{10} \).

When we focus on a particular \( C \)-level, say \( i_0 \) in the lower half that maps to a \( C \)-level \( i_0' \) in the upper half of the column, the value of the cocycle corresponding to \( \phi_0 \), as we move up the tower to level \( N - 1 \), and the value of the cocycle as we move from level \( N + 1 \) up to the level \( i_0' \), define a pair \((h_1, h_2) \in G \times G\).

Let \( V_1, V_2 \) be \( \frac{a}{10} \)-balls around \( g_1, g_2 \) respectively. As in the proof of Proposition 3.2 there is a finite set \( F \subset G \) of cardinality \( m \) such that for any two pairs \((h_1, h_1')\) and \((h_2, h_2')\) in \( G \times G \) as above there is a pair \((f_1, f_2) \in F \times F\) with
\[
h_1 f_1 h_1' \in V_1, \quad h_2 f_2 h_2' \in V_2.
\]
Note that \( m \) depends only on \( a \).

**Step 5: Defining the cocycle \( \phi \)**

On the space \( Z \times K/H \times K/H \times V \times V \) we now define sequences of partitions as follows. We choose a refining sequence of measurable partitions \( \{\mathcal{P}_i\} \) on \( Z \) which together generate the Borel \( \sigma \)-algebra \( \mathcal{B}_Z \). Also, a refining sequence of measurable partitions \( \{\mathcal{Q}_i\} \) on \( K/H \) which together generate the Borel \( \sigma \)-algebra \( \mathcal{B}_{K/H} \). Finally we let \( \{\mathcal{R}_i\} \) be a sequence of refining partitions of \( V \) which together generate the \( \sigma \)-algebra \( \mathcal{B}_V \). Now consider the corresponding sequence of partitions \( \{\mathcal{P}_i \times \mathcal{Q}_i \times \mathcal{Q}_i \times \mathcal{R}_i \times \mathcal{R}_i\} \) of \( Z \times K/H \times K/H \times V \times V \). By choosing \( l \geq i_0 \) sufficiently large we can assume that each atom \( T^N_k B_j \) is approximated by \( \mathcal{P}_l \times \mathcal{Q}_l \times \mathcal{Q}_l \times \mathcal{R}_l \times \mathcal{R}_l \) up to a set of relative measure \( < \frac{a}{10} \) (i.e. relative to the total measure of \( T^N_k(B) \)).
We let $d_K$ be a bi-invariant metric on $K$ and let $\bar{d}$ be the restriction of the Hausdorff metric on the hyperspace $2^K$ of closed subsets of $K$ to the subspace \( \{ kh : k \in K \} \).

Let $\eta > 0$ be such that for $K_0 = \{ k \in K : d_K(k, H) > \eta \}$ we have $\lambda(K_0) > 1 - b/10$, (this is possible when $\lambda_K(H) = 0$ (iff $[K, H] = \infty$)). We also assume that $l$ is sufficiently large so that

$$\max\{ \text{diam}_\bar{d}(Q) : Q \in \mathcal{Q}_l \} < \eta.$$

Note that, as the metric $d_K$ is invariant, if $k_0 \in K_0$ then for any $Q \in \mathcal{Q}$ and any $kH \in Q$, we have $d(kH, kk_0H) > \eta$, whence $Q \cap Q' = \emptyset$, for any $Q' \in R_{k_0}(Q)$, where $R_{k_0}(kH) = \{ khk_0H : h \in H \}$.

Once $l$ is chosen we choose a suitable index $l'$ so that

$$\max\{ \mu(P) : P \in \mathcal{P}_l \}$$

is small (how small it needs to be will be determined in Step 6).

We can now define our new cocycle $\phi$ by setting it equal to $\phi_0$ outside $T^N(B)$, and on each atom $P \times Q \times R$ of $\mathcal{P}_l \times \mathcal{Q}_l \times \mathcal{R}_l \mid T^N(B_0 \times K/H \times V)$, letting $\phi$ be a random variable $\xi_{P \times Q \times R}$ taking values in $F$, uniformly and independently for distinct atoms (so that now $\phi = \phi^\omega$, $\omega \in \Omega$, the sample space of all these random variables).

Note that for $k_0 \in K_0$, if $(z, kH, v_1)$ and $(z, kk_0H, v_2)$ are in distinct atoms of $\mathcal{P}_l \times \mathcal{Q}_l \times \mathcal{R}_l$, then $\phi^\omega(z, kH, v_1)$ and $\phi^\omega(z, kk_0H, v_2)$ are independent and therefore take on each value of $F \times F$ with probability $1/|F|^2$.

**Step 6: A simple probabilistic lemma**

Here we prove a probabilistic lemma which will serve us in Step 7.

4.8. **Lemma.** Let $p$ be a number in $(0, 1)$ and let $L$ be a positive constant. Suppose that $w_j \geq 0$ and $\sum_{j=1}^n w_j = 1$. Let $X_j$, $j = 1, \ldots, n$ be random variables taking values in $\{0, 1\}$ with $P(X_j = 1) \geq p$. Suppose further that for each $j$ there is a set $I_j$, $|I_j| \leq L$, such that for $i \notin I_j$, $X_i$ and $X_j$ are independent. Set $w = \max_{1 \leq j \leq n} w_j$. Then for $X = \sum_{j=1}^n w_j X_j$ we have

$$P(X \geq p/2) \geq 1 - \frac{4(L + 1)}{p^2} w.$$

**Proof.** Let $\bar{x}_j = P(X = j = 1)$ and define $Y_j = X_j - \bar{x}_j$. Let $\bar{p} = \sum_{j=1}^n w_j \bar{x}_j \geq p$. Then $\sum w_j Y_j = X - \bar{p}$, and if $j \notin I_j$ we have

$$\mathbb{E}(Y_j Y_i) = 0, \quad \text{while} \quad \mathbb{E}(X - p) = 0.$$

Now

$$(X - \bar{p})^2 = \sum_{j=1}^n w_j^2 \bar{x}_j^2 + \sum_{j=1}^n \sum_{i \notin I_j} Y_j Y_i.$$

Since $\mathbb{E}(Y_j^2) \leq 1/4$ we get

$$\mathbb{E}((X - \bar{p})^2) \leq 1/4 \cdot w + w \cdot L \leq w(L + 1).$$

As $\bar{p} \geq p$, our claim follows. \(\square\)

**Step 7: Estimating the probability of success**
Fix \( k_0 \in K_0 \) and fix a pure column based on \( B_j \). Since \( B_j \) is the base of a pure column, either \( \tau \) maps it to a \( C \)-level, or to a level not in \( C \). We will assume that the \( i \)-th level of the \( B_j \) column is mapped by \( \tau \) to a \( C \)-level. The total \( \mu_{k_0} \) measure of these \( C \)-levels is at least \( \frac{1}{6} \mu_{k_0}(C)^2 \) (see Step 4).

Suppose \((z, kH, kk_0, v_1, v_2) \in T^i_{k_0} B_j \) and, it is mapped by \( \tau \) to level \( i' := N + \pi(i) \). Let
\[
    h_1(z, kH, v_1) = \phi_{N-1-i}(z, kH, v_1), \quad h_2(z, kk_0H, v_2) = \phi_{N-1-i}(z, kk_0H, v_2)
\]
and
\[
    h_1'(z, kH, v_1) = \phi_{N-1}(T^{N-i+1}(z, kH), S_{(z,kH)}^{(N-i)} v_1),
    h_2'(z, kk_0H, v_2) = \phi_{N-1}(T^{N-i+1}(z, kk_0H), S_{(z,kk_0H)}^{(N-i)} v_2).
\]

Suppose \( T^{N-i}(z, kH, v_1) \in P \times Q_1 \times R_1 \) and \( T^{N-i}(z, kk_0H, v_2) \in P \times Q_2 \times R_2 \). Then
\[
    \psi_\tau^w(z, kH, kk_0H, v_1, v_2) = (\phi_\tau^w(z, kH, v_1), \phi_\tau^w(z, kk_0H, v_2)) = (h_1'(z, kH, v_1) \cdot \xi_{P \times Q_1 \times R_1}, h_2'(z, kk_0H, v_2) \cdot \xi_{P \times Q_2 \times R_2} \cdot h_2(z, kk_0H, v_2)).
\]

The set \( T^i_{k_0} B_j \) is partitioned into sets according to the atoms \( P, Q_1, Q_2, R_1, R_2 \), and we denote an element of this partition by \( \alpha \). We let \( X_\alpha^w = 1 \) iff \( \psi_\tau^w \) on \( \alpha \) satisfies
\[
    d(\phi_\tau^w(z, kH, v_1), g_1) < a \quad \& \quad d(\phi_\tau^w(z, kk_0H, v_2), g_2) < a.
\]

Now, for \( Q_1 \neq Q_2 \) the corresponding random variables are independent and when we define pieces \( \alpha \) and the corresponding \( X_\alpha \) as before, we get \( \text{Prob}(X_\alpha = 1) \geq \frac{1}{m^2} \).

We now want to apply Lemma 4.8. To do this we need to estimate, for a fixed \( X_\alpha \), the number of \( \beta \) that are not independent of it. Clearly if \( \alpha \) and \( \beta \) correspond to distinct \( P_\tau \) atoms, \( X_\alpha \) and \( X_\beta \) are independent. Therefore the number of “bad” \( \beta \)’s is at most \( L := |Q_1| \cdot |R_1|^2 \). By choosing \( \ell' \) sufficiently large we can get the maximum weight \( w \) in Lemma 4.8 small enough so that we get for all pure columns and their \( C \)-levels, for each fixed \( k \), a set in \( \Omega \) of probability at most \( 1 - b/10 \), for which the set of \((z, kk_0H, v_2), g_2) \) in equation (3) is at least \( \frac{1}{100} \cdot \frac{1}{|P'|^2} \cdot \mu_{k_0}(C)^2 \).

As before, using Fubini’s theorem on the probability space \((\Omega \times K, \text{Prob} \times \lambda_K)\), we find a subset of \( K \) with \( \lambda_K \) measure at least \( 1 - b \) for which equation 3 holds. The constant \( c_a \) is now determined to be \( \frac{1}{100} \cdot \frac{1}{|P'|^2} \).

This concludes the proof of Proposition 4.5. \( \square \)

Now back to the proof of Theorem 4.4. By Lemma 2.3 what we have to show is that the system \( X_\phi \times X_\phi \) is relatively ergodic over \( Y \times Y \). As above we define the set \( U = U(C, a, b, c_\alpha, g_1, g_2) \subset \mathcal{C} \) and by Proposition 4.5 we know that it is open and dense in \( \mathcal{C} \).

Now the ergodic components of \( Y \times Y \) are of the form
\[
    Y_{k_0} = \{(z, kH, kk_0H) : z \in Z, \, k \in K\},
\]
and, as we have seen above, the corresponding ergodic component of $X \times X$ has the form (2)

$$T_{k_0}(z, kH, kk_0H, v_1, v_2) = (Tz, \gamma(z)kH, \gamma(z)kk_0H, S_{(z, kH)}v_1, S_{(z, kk_0H)}v_2),$$

Thus for a fixed $k_0$ we have a cocycle $\psi_{k_0} : Z \times K/H \times K/H \times V \times V \to G \times G$,

$$\psi_{k_0}(z, kH, kk_0H, v_1, v_2) = (\phi(z, kH, v_1), \phi(z, kk_0H, v_2)),$$

which we must show is ergodic.

Fix $a$ and the corresponding $c_a$ and consider the union $K_\infty = \bigcup_n K_n \subset K$, with

$$K_n = \{ k_0 \in K : \mu_{k_0}(\{(z, kH, kk_0H, v_1, v_2) \in C : d(\phi(\tau, z, kH, v_1), g_1) < a \\ & d(\phi(\tau, z, kk_0H, v_2), g_2) < a \}) > c_a \mu_{k_0}(C)^2 \},$$

so that $\lambda_K(K_n) > 1 - \frac{1}{2^n}$. Then $\lambda_K(K_\infty) = 1$, and for every $k_0 \in K_\infty$

$$\mu_{k_0}(\{(z, kH, kk_0H, v_1, v_2) \in C : \exists \tau \in [T]_f \text{ with } d(\phi(\tau, z, kH, v_1), g_1) < a \\
& d(\phi(\tau, z, kk_0H, v_2), g_2) < a \}) > c_a \mu_{k_0}(C)^2.$$

Now with

$$\mathcal{U}(C, a, c_a, g_1, g_2) = \bigcap_n \mathcal{U}(C, a, \frac{1}{2^n}, c_a, g_1, g_2),$$

we have a set $K_\infty = K_\infty(C, a, c_a, g_1, g_2) \subset K$, with $\lambda_K(K_\infty) = 1$ such that for $k \in K_\infty$

$$\mathcal{U}(C, a, c_a, g_1, g_2)$$

is dense $G_0$ in $\mathcal{C}$ (we eliminated the parameter $b$).

Now take an intersection

$$\mathcal{U}(a, c_a, g_1, g_2) = \bigcap_{n \in \mathbb{N}} \mathcal{U}(C_n, a, c_a, g_1, g_2)$$

over a dense collection $\{C_n\}_{n \in \mathbb{N}}$ in the measure algebra of $Z \times K/H \times K/H \times V \times V$

and take the corresponding intersection

$$K_\infty(a, c_a, g_1, g_2) = \bigcap_{n \in \mathbb{N}} K_\infty(C_n, a, c_a, g_1, g_2).$$

Next take

$$\mathcal{U}(g_1, g_2) = \bigcap_{n \in \mathbb{N}} \mathcal{U}(\frac{1}{n}, c_{\frac{1}{n}}, g_1, g_2)$$

and the corresponding intersection

$$K_\infty(g_1, g_2) = \bigcap_{n \in \mathbb{N}} L(\frac{1}{n}, c_{\frac{1}{n}}, g_1, g_2).$$

Finally let

$$\mathcal{U} = \bigcap \{ \mathcal{U}(g_1, g_2) : (g_1, g_2) \in G_0 \times G_0 \},$$

with $G_0 \subset G$ a countable dense subset of $G$, and the corresponding

$$L = \bigcap \{ K_\infty(g_1, g_2) : (g_1, g_2) \in G_0 \times G_0 \}.$$

The existence of the set $L$ demonstrates the fact that for almost all $k \in K$, and every cocycle $\phi$ in the residual set $\mathcal{U}$, every pair $(g_1, g_2) \in G \times G$ is an essential value for the cocycle $\psi_k$. It now follows that the system $X_{\psi_k} \times X_{\psi_k}$ is relatively ergodic over $Y \times Y$

and our proof of Theorem 4.4 is complete. □
The next theorem is actually a special case of Theorem 4.4 (when $X = Y$), however it is convenient for us to formulate it as a separate theorem, in order to use it in the following proof of Theorem 2.6.

4.9. Theorem. Let $Y \to Z$ be an infinite factor of the ergodic system $Y$ such that the extension $Y \to Z$ is a $K/H$-extension. Then for a generic $\phi \in \mathcal{C}(Y, G)$ the system $Y_\phi = (Y \times G, Y \times B_G, \mu \times \lambda_G, T_\phi)$, is ergodic and the projection map, $\pi : Y \to Z$, is the maximal compact extension of $Z$ within $Y_\phi$.

We can now complete the proofs of Theorems 2.6 and 0.1.

Proof of Theorem 2.6. Apply a transfinite induction along the ordinal $\alpha$, using Theorem 4.9 for successor ordinals and, Lemma 2.5 and Theorem 4.4 for limit ordinals. □

Proof of Theorem 0.1. As was explained in Section 2 the combination of Theorems 2.6 and Theorem 2.7 implies Theorem 0.1. □

5. SOME COROLLARIES

5.1. Set up. Let $K$ and $G$ be two arbitrary second countable compact groups. Suppose that $Z = (Z, Z, \nu, R)$ is an ergodic $Z$-system and suppose that $Y = (Y, Y, \mu, T)$, with $\mu = \nu \times \lambda_K$, $\lambda_K$ being the Haar measure on $K$, is a $K$-extension of $Z$. Set $X = Y \times G = Z \times K \times G$ and let $\mathcal{C} = \mathcal{C}(Y, G)$ be the Polish space of Borel maps $\phi : Y \to G$. For $\phi \in \mathcal{C}$ set

$$T_\phi(y, g) = (Ty, \phi(y)g), \quad (y \in Y, g \in G).$$

5.2. Theorem. For any $n \geq 1$ and fixed $(k_1, k_2, \ldots, k_n)$ of distinct elements of $K$ there is a dense $G_\delta$ subset $\mathcal{C}_0 \subset \mathcal{C}(Y, G)$ such that:

1. For $\phi \in \mathcal{C}_0$ the corresponding cocycle

$$\phi_{(k_1, k_2, \ldots, k_n)} : Y \to G^{n+1}$$

defined by

$$\phi_{(k_1, k_2, \ldots, k_n)}(y) = (\phi(y), \phi(yk_1), \ldots, \phi(yk_n))$$

is ergodic.

2. Furthermore, denoting by $Y_{(k_1, k_2, \ldots, k_n)}$ the corresponding ergodic skew product on $Y \times G^{n+1}$, we have that the extension $Y \to Z$ is the largest compact extension of $Z$ in $Y_{(k_1, k_2, \ldots, k_n)}$.

3. Denoting, for $k \in K$, by $\phi_k : Y \to G$ the cocycle $\phi_k(y) = \phi(yk)$, we have that the corresponding skew products $Y_{\phi_{k_i}}$, $i = 0, 1, \ldots, n$, are jointly disjoint over their common factor $Y$.

Proof. (1) and (2): The proof is similar to the proof of Theorem 4.9. The situation here differs in two ways. The first is the fact that we are considering here the groups $K^n$ and $G^{n+1}$ rather than $K$ and $G$. This change however does not cause any new difficulty, the generalization is straightforward. The second way actually makes the
proof much easier since we now have to deal with a single element of the group $K^n$, rather than a subset of $K^n$ with large Haar measure. Following the proof of Theorem 4.9, one formulates a proposition similar to Proposition 4.5, where now the definition of the open set $U$ is simplified, as follows:

Fix a subset $C \subset Y$, $\mu(C) > 0$, elements $g_1, \ldots, g_{n+1}$ in $G$, a positive small constant $a$, and a positive constant $c_a$ (which will depend on $a$). Define the set $U(C, a, g_1, g_2, \ldots, g_{n+1}) \subset C$ as the collection of all the functions $\phi \in C$ with the following property:

(1) $\tau(C) = C$,

(2) With $k_0 = e$,

$$\mu(\{y \in C : \forall j, \ 0 \leq j \leq n, \ d(\phi_\tau(yk_j), g_j+1) < a\}) > c_a \mu(C).$$

We leave the details of the proof to the reader.

(3) Since the skew product system $Y_{(k_1, k_2, \ldots, k_n)}$ is clearly isomorphic to the relative independent product of $\prod_{i=0}^{n} Y_{\phi_i}$, it follows that the latter’s relative product measure is ergodic and thus, by [6, Theorem 3.30], it is the unique invariant measure on $Y_{(k_1, k_2, \ldots, k_n)}$ which projects onto $\mu$.

5.3. Theorem. (1) There is a dense $G_\delta$ subset $C_1 \subset C(Y, G)$ such that for each $\phi \in C_1$ and every $n \geq 1$, there is a dense $G_\delta$ subset $A_n \subset K^n$ with the property that for every $(k_1, k_2, \ldots, k_n) \in A_n$ the corresponding cocycle $\phi_{(k_1, k_2, \ldots, k_n)}$ from $Y$ to $G^{n+1}$ is ergodic.

(2) There is a Cantor subset $K_0 \subset K$ such that for every $n \geq 1$ and every $(k_1, k_2, \ldots, k_n)$ with $k_i \in K_0$, $i = 1, 2, \ldots, n$, the corresponding cocycle $\phi_{(k_1, k_2, \ldots, k_n)}$ from $Y$ to $G^{n+1}$ is ergodic, so that, as in Theorem 5.2(3), the skew products $Y_{\phi_{k_i}}, i = 0, 1, \ldots, n$, are jointly disjoint over their common factor $Y$.

Proof. The first claim follows from the Kuratowski–Ulam theorem (see, for example, [8, Theorem 8.41]). Then the second claim follows by Mycielski’s theorem (see, for example, [8, Theorem 19.1]).

5.4. Theorem. Let $K$ be an infinite monothetic compact second countable topological group with a topological generator $a$ (i.e. $K = \{a^n : n \in \mathbb{Z}\}$). Let $K = (K, \mathcal{B}_K, \lambda_K, T_a)$ be the corresponding ergodic system. Let $G$ be an arbitrary compact second countable topological group. Then for a generic cocycle $\phi \in C(K, G)$ the corresponding skew product system $X_\phi = (K \times G, \mathcal{B}_K \times \mathcal{B}_G, \lambda_K \times \lambda_G, T_\phi)$ is (i) ergodic and (ii) the system $K$ is the largest compact factor of $X_\phi$. In particular then $X_\phi$ is distal of rank 2.

Proof. This is in fact a special case of Theorem 4.9.
6. The case of a weakly mixing extension

6.1. Set up. Let $Y \to Z$ be a nontrivial relatively weakly mixing extension of ergodic systems. Let $G$ be a second countable compact topological group with Haar measure $\lambda_G$. Let $C = C(Y,G)$ be the Polish space of measurable cocycles $\phi : Y \to G$. For $\phi \in C$ we let

$$T_{\phi}(y,g) = (Ty, \phi(x)g), \quad y \in Y, g \in G,$$

and $Y_{\phi} = (Y \times G, Y \times \mathcal{B}_G, \mu \times \lambda_G, T_{\phi})$. We let $\psi : Y \times Y \to G \times G$ be the cocycle $\psi(y_1, y_2) = (\phi(y_1), \phi(y_2))$.

6.2. Theorem. For a generic cocycle $\phi \in C(Y,G)$ and the corresponding skew product transformation $T_{\phi} : Y \times G \to Y \times G$, the extension $Y_{\phi} \to Z$ is relatively weakly mixing. In particular, when $Y$ is weakly mixing (i.e. when $Z$ is the trivial one point system) so is the system $Y_{\phi}$ for a generic cocycle $\phi \in C(Y,G)$.

Proof. Let $\pi : Y \to Z$ be the factor map $Y \to Z$. Our assumption is that the system $W = Y \times Y$ with

$$W = Y \times Y = \{(y_1, y_2) : \pi(y_1) = \pi(y_2)\} \subset Y \times Y,$$

$T(y_1, y_2) = (Ty_1, Ty_2)$, and $\zeta = \mu \times \mu$ the relative product measure over $\pi$, is ergodic.

What we have to show is that for a generic $\phi \in C$ the system $W_{\phi} = (W, \zeta, T_{\phi})$

where

$$W_{\phi} = W \times G \times G = \{(y_1, g_1), (y_2, g_2) : \pi(y_1) = \pi(y_2)\},$$

$\tilde{\zeta} = \zeta \times \lambda_G \times \lambda_G$ and $T_{\phi}(y_1, y_2, g_1, g_2) = (Ty_1, Ty_2, \phi(y_1)g_1, \phi(y_2)g_2)$, is ergodic.

Define, for $C \subset W = Y \times Y$, $a > 0$, $c_a$ a constant depending on $a$, and elements $g_1, g_2 \in G$, a subset $U = U(C, a, c_a, g_1, g_2) \subset C(Y, G)$ as follows

$$U = \{\phi \in C : \exists \tau \in [T \times T], \zeta(\{(y_1, y_2) \in C : \tau(y_1, y_2) \in C, \zeta(\{(y_1, y_2) \in C : \tau(y_1, y_2) \in C, and d(\psi_{\tau}(y_1, y_2), (g_1, g_2)) < a\}) \}

> c_a(\mu \times \mu)(C)\}.$$ 

Here $\tau(y_1, y_2) = (T^{\sigma(y_1, y_2)}y_1, T^{\sigma(y_1, y_2)}y_2)$, for a function $\sigma : W \to Z$.

6.3. Proposition. For a sufficiently small $c_a$ the set $U(C, a, c_a, g_1, g_2)$ is open and dense in $C(Y, G)$.

Proof. It is clear that $U$ is an open set. We will show that it is dense in $C$. So fix $\phi_0 \in C$ and $\delta > 0$, with $\delta \ll \zeta(C)$. Take a set $A \subset X$ with $\mu(A) < \delta/10$. Let $B \subset W$ be a base of a Rokhlin tower of height $3N$, such that (using the ergodic theorem for $T \times T$ on $W$), when the tower is purified with respect to $A \times A$ and $C$, the following holds: the pure columns that satisfy

(i) the number of $C$-levels in the top and bottom thirds of the column are at least

$$\frac{1}{2} \zeta(C) \cdot N,$$
(ii) in the middle third there is at least one level in $A \times A$, fill $(1 - \frac{\delta}{10})$ of $W$.

Now define $\tau$ on good pure columns by exchanging $C$-levels in the lower and upper thirds, and identity elsewhere.

By [11] we can assume that in the diagram $\pi : Y \to Z$, the ergodic systems $Y$ and $Z$ are represented as topological strictly ergodic flows; i.e. $Y$ and $Z$ are metric compact spaces, $T$ acts on both as a homeomorphism under which the topological flows are minimal and uniquely ergodic, and finally that $\pi$ is a continuous map. As we assume that the extension $\pi$ is nontrivial and relatively weakly mixing, it follows that the minimal flow $(Y, T)$ has no finite orbits.

Working with these models we see that the in the system $W$ the space $W = Y \times Z$ is compact metric and the transformation $T \times T : W \to W$ is a homeomorphism.

Given $\eta > 0$ set

$$E = \{(y_1, y_2) \in W : \min_{0 \leq i, j \leq 3N} d(T^i y_1, T^j y_1) > \eta, \min_{0 \leq i, j \leq 3N} d(T^i y_1, T^j y_2) > \eta, \min_{0 \leq i, j \leq 3N} d(T^i y_2, T^j y_2) > \eta\}.$$

It then follows that for a sufficiently small $\eta > 0$ we will have $\zeta(E) > 1 - \frac{\delta}{100}$. We call $E$ the $\eta$-separated set.

Next let $\mathcal{P}$ be a finite partition of $A$ with

$$\max_{P \in \mathcal{P}} \text{diam}(P) < \frac{\eta}{100}.$$

Let $F = \{f_1, f_2, \ldots, f_m\} \subset G$ be a set such that $\bigcup_{i=1}^{m} B(f_i, \frac{\eta}{10}) = G$. We define a random cocycle

$$\phi^\omega(y) = \begin{cases} \phi_0(y), & x \notin A_0 \\ \xi^\omega_P, & y \in P \in \mathcal{P}, \end{cases}$$

where $\xi^\omega_P$ are random variables on $\Omega$ with

$$\text{Prob}(\xi^\omega_P = f_i) = 1/m, \quad f_i \in F$$

and independent for $P \neq P'$. If $y \in A$ and $y \in P$ we write $P = P(y)$ so that $y \in P(y)$.

Now, for $(y_1, y_2)$ in a $C$-level of a pure column $i$, $0 \leq i < N$, which is mapped by $\tau$ to the $C$-level $i'$, $2N < i' \leq 3N$, the cocycle $\psi^\omega_\tau$ takes the following form :

$$\psi^\omega_\tau(y_1, y_2) = \prod_{n=0}^{i'-i-1} \rho^\omega_n(y_1, y_2),$$

where the random variables $\rho^\omega_n$ are defined as follows

$$\rho^\omega_n(y_1, y_2) = \begin{cases} (\xi^\omega_{P(T^n y_1)}, \xi^\omega_{P(T^n y_2)}), & \text{if } T^n y_1 \in A \& T^n y_2 \in A \\ (\phi_0(T^n y_1), \phi_0(T^n y_2)), & \text{if } T^n y_1 \notin A \& T^n y_2 \notin A \\ (\xi^\omega_{P(T^n y_1)}, \phi_0(T^n y_2)), & \text{if } T^n y_1 \notin A \& T^n y_2 \in A \\ (\phi_0(T^n y_1), \xi^\omega_{P(T^n y_2)}), & \text{if } T^n y_1 \notin A \& T^n y_2 \in A. \end{cases}$$
When \((y_1, y_2)\) is in the \(\eta\)-separated set, all \(P(T^ny_1), P(T^ny_2)\) that appear in the product for \(\psi_\tau^\omega\) are distinct and there is at least one \((P(T^ny_1), P(T^ny_2))\) in the product by our assumption that in the middle third, that we have to pass through, there is at least one level in \(A \times A\).

all the \(P_n\) and \(Q_n\) that appear are distinct, and there is at least one \((P_n \times Q_n)\) in the product (4) since we pass through the middle third where there is at least one \((A \times A)\)-level.

It follows that for each such \((y_1, y_2)\)

\[
\text{Prob}(d(\psi_\tau^\omega(x_1, x_2), (g_1, g_2)) < a) \geq \frac{1}{m^2}.
\]

Applying Fubini’s theorem to \(\Omega \times (C \cap (\bigcup_{i=0}^{N-1} (T \times T)^i B))\) we see that there is a choice of \(\omega\) such that

\[
\zeta(\{(y_1, y_2) \in C : \tau(y_1, y_2) \in C, \text{ and } d(\psi_\tau^\omega(y_1, y_2), (g_1, g_2) < a)\}) > \frac{1}{10m^2} \cdot \zeta(C).
\]

We now let \(c_a = \frac{1}{10m^2}\) and observe that this \(\phi^\omega\) is \(\delta\)-close to \(\phi_0\) and lies in \(U\). Thus we have shown that \(U\) is an open and dense subset of \(\mathcal{C}\). \(\square\)

To deduce the statement of Theorem 6.2 we now follow the same procedure that we used in deducing Theorem 4.9 from Proposition 4.5. \(\square\)

We now have the following corollary which is a far reaching strengthening of Theorem 3.1.

**6.4. Theorem.** Let \(Y \to Z\) be a factor map of ergodic systems. Let \(Y \to Y_{rd} \to Z\) be the (relative) Furstenberg-Zimmer structure for the extension \(Y \to Z\). Then for a generic \(\phi \in \mathcal{C}(Y, G)\) we have that \(Y_{\phi} \to Y_{rd} \to Z\) is the (relative) Furstenberg-Zimmer structure for the extension \(Y_{\phi} \to Z\). In particular, when we take \(Z\) to be the trivial one point system, it follows that for every ergodic system \(Y\) with Furstenberg-Zimmer structure \(Y \to Y_d\) (the latter being the largest distal factor \(Y\)), for a generic \(\phi \in \mathcal{C}(Y, G)\) the corresponding system \(Y_{\phi}\) is ergodic with Furstenberg-Zimmer structure \(Y_{\phi} \to Y_d\).

**Proof.** Let \(Y \to Y_{rd} \to Z\) be the Furstenberg-Zimmer structure for the extension \(Y \to Z\); i.e. \(Y_{rd}\) is the largest relative distal extension of \(Z\) in \(Y\), and the extension \(Y \to Y_{rd}\) is relatively weakly mixing. If this latter extension is non-trivial then, by Theorem 6.2, for a generic \(\phi \in \mathcal{C}(Y, G)\), the extension \(Y_{\phi} \to Y_{rd}\) is relatively weakly mixing. On the other hand, if \(Y = Y_{rd}\) then we use Theorem 4.4. \(\square\)

We also have the following theorem:

**6.5. Theorem.** Let \(Y \to Z\) be an infinite factor of the ergodic system \(Y\) and let \(G\) be a compact second countable topological group. Then for a generic \(\phi \in \mathcal{C}(Y, G)\), with \(Y_{\phi} = (Y \times G, \xi \times B_G, \mu \times \lambda_G, T_{\phi})\), the extension \(Y_{\phi} \times Y_{\phi} \to Y \times Y\) is relatively ergodic.
Proof. Let \( \tilde{Z} \to Z \) be the maximal compact extension of \( Z \) in \( Y \). If this factor map is trivial, that is if the extension \( Y \to Z \) is relatively weakly mixing, the assertion holds by Theorem 6.2. Otherwise, by the relative version of the Furstenberg-Zimmer theorem, the extension \( \tilde{Z} \to Z \) is a nontrivial \( K/H \)-extension and the assertion follows from Theorem 4.4. \( \square \)

7. A general framework and a master theorem

7.1. Definition. Given a diagram \( X \to Y \to Z \) with \( X \) an ergodic system we say that \( X \) is \( 2 \)-fold ergodic over \( Y \to Z \) when the extension

\[
X \times X \to Y \times Y
\]

is relatively ergodic.

Given a diagram \( W \to X \to Y \to Z \) with \( W \) an ergodic system such that the condition (5) is satisfied, we will say that the extension \( W \to X \) is \( 2 \)-fold ergodic over \( Y \to Z \). Let \( G \) be a compact second countable topological group. Given a diagram \( X \to Y \to Z \) with \( X \) an ergodic system, \( a \) cocycle \( \phi \in C(X,G) \) is \( 2 \)-fold ergodic over \( Y \to Z \) if the corresponding extension \( X_\phi \to X \) is \( 2 \)-fold ergodic over \( Y \to Z \); i.e. when the extension

\[
X_\phi \times X_\phi \to Y \times Y
\]

is relatively ergodic. This is the same as saying that for the cocycle \( \psi : X \to G \times G \), given by \( \psi(x_1,x_2) = (\phi(x_1),\phi(x_2)) \), the extension

\[
(X \times X)_\psi \to Y \times Y
\]

is relatively ergodic. Using this terminology the assertion of Lemma 2.3 is that when \( X \to Y \to Z \) is a diagram of ergodic systems and \( Y \to Z \) is a compact extension, then \( Y \) is the maximal compact extension of \( Z \) in \( X \) iff the system \( X \) is \( 2 \)-fold ergodic over \( Y \to Z \).

More generally, given a diagram \( W \to X \to Y \to Z \) with \( W \) an ergodic system, we define, for every \( n \geq 2 \), the notions of \( n \)-fold ergodicity of the extension \( W \to X \) over \( Y \to Z \), and of \( \phi \in C(X,G) \) over \( Y \to Z \). E.g. \( X \) is \( 3 \)-fold ergodic over \( Y \to Z \) when the extension

\[
X \times X \times X \to Y \times Y \times Y
\]

is relatively ergodic, and the cocycle \( \phi \in C(X,G) \) is \( 3 \)-fold ergodic over \( Y \to Z \) when the cocycle \( \psi(x_1,x_2,x_3) = (\phi(x_1),\phi(x_2),\phi(x_3)) \) is relatively ergodic over \( Y \times Y \).

It is not hard to see that \( 2 \)-fold ergodicity implies \( n \)-fold ergodicity for all \( n \geq 3 \).

7.2. Lemma. Let \( X \to Y \to Z \) with \( X \) ergodic be given. If the extension \( X \times X \to Y \times Y \) is relatively ergodic then the extension \( X \to Y \) is relatively weakly mixing.
Proof. Suppose $X \to Y$ is not relatively weakly mixing. Then there exists an intermediate factor $X \to E \to Y$ with $E \to Y$ a nontrivial compact extension. It follows that the extension $E \times E \to Y \times Y$ is not relatively ergodic and a fortiori also $X \times X \to Y \times Y$ is not relatively ergodic, contradicting our assumption. \hfill \Box

7.3. Lemma. Let $X \to Y \to Z$ with $X$ ergodic be given. If the extensions $X \to Y$ and $Y \to Z$ are relatively weakly mixing then so is the extension $X \to Y \times Z$.

Proof. Suppose $X \to Z$ is not relatively weakly mixing. Then there exists an intermediate factor $X \to E \to Z$ with $E \to Z$ a nontrivial compact extension. By our assumptions $E \to Z$ is compact and $Y \to Z$ is relatively weakly mixing. It then follows that $Y$ and $E$ are relatively disjoint over their common factor $Z$, and we have $X \to Y \times E \to Y$. (For this we refer e.g. to [6, Theorem 6.27], where this claim is proven in the absolute case; however the same proof works also in the relative case.) Now clearly $(Y \times E) \times (Y \times E)$ is not ergodic, but we also have $X \times X \to (Y \times E) \times (Y \times E)$ and we arrived at a contradiction. \hfill \Box

Combining the 2-fold ergodicity theorems proven so far, we can now state and prove a master theorem on 2-fold ergodicity.

7.4. Theorem. Let $G$ be a compact second countable topological group with normalized Haar measure $\lambda_G$. Let $X \to Y \to Z$ be a chain of factors of the ergodic system $X$ with $Z$ infinite. Suppose further that the extension $X \times X \to Y \times Y$ is relatively ergodic. Then for a generic $\phi \in \mathcal{C}(X,G)$, with $X_\phi = (X \times G, X \times B_G, \mu \times \lambda_G, T_\phi)$, the extension $X_\phi \times X_\phi \to Y \times Y$ is relatively ergodic. Equivalently, the extension $X_\phi \to X$ is 2-fold ergodic over $Y \to Z$.

Proof. Let $Y \to Y_{rd} \to Z$ be the Furstenberg-Zimmer structure for the extension $Y \to Z$. Then the fact that the extension $Y \to Y_{rd}$ is relatively weakly mixing, combined with our assumption imply that also the extension $X \to Y_{rd}$ is relatively weakly mixing; that is, $X_{rd} = Y_{rd}$ and $X \to Y_{rd} \to Z$ is the Furstenberg-Zimmer structure for the extension $X \to Z$.

Indeed, by Lemma 7.2 the extension $X \to Y$ is relatively weakly mixing. Also, by definition the extension $Y \to Y_{rd}$ is relatively weakly mixing. Hence, by Lemma 7.3 the extension $X \to Y_{rd}$ is relatively weakly mixing.

Case 1: Suppose first that $Y_{rd} \to Z$ is nontrivial; then there is a nontrivial intermediate extension $X_\phi \to Z_1 \to Z$, so that $Z_1$ is the maximal compact extension of $Z$ in $X_\phi$. By Lemma 2.3 the extension $X_\phi \times X_\phi \to Z_1 \times Z_1$
is relatively ergodic and, since, for a generic $\phi$, by Theorem 6.4, $Z_1$ is also a factor of $Y$, a fortiori also

$$X_\phi \times X_\phi \to Y \times Y$$

is relatively ergodic. (In fact, if $f \in L^2(X_\phi \times X_\phi)$ is $T$-invariant then, by relative ergodicity, it is $Z_1 \times Z_1$-measurable, and, a fortiori, also $Y \times Y$.)

Case 2: In the remaining case the extension $X_\phi \to Z$ is relatively weakly mixing; i.e. the system $X_\phi \times X_\phi$ is ergodic, and clearly then the extension

$$X_\phi \times X_\phi \to Y \times Y$$

is relatively ergodic.

We now observe that:

- Theorem 4.4 is a special case of Theorem 7.4, when we assume that the extension $Y \to Z$ is a $K/H$-extension.
- Theorem 4.9 is the special case of Theorem 7.4 when we assume that $X = Y$ and that the extension $Y \to Z$ is a $K/H$-extension.
- Theorem 6.2 is obtained from Theorem 7.4 when we take $Z = Y$.
- By taking $X = Y$ in Theorem 7.4 we arrive at Theorem 6.5 whose conclusion can be stated as the claim that the system $Y_\phi$ is 2-fold ergodic over $Y \to Z$.

References

[1] Jon Aaronson and Benjamin Weiss, On Herman's theorem for ergodic, amenable group extensions of endomorphisms, Ergod. Th. & Dynam. Sys. (2004), 24, 1283–1293
[2] Ferenc Beleznay and Matthew Foreman, The complexity of the collection of measure-distal transformations, Ergodic Theory Dynam. Systems 16, (1996), no. 5, 929–962.
[3] Harry Furstenberg, The structure of distal flows, American J. of Math. 85, (1963), 477-515.
[4] Harry Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, J. d’Analyse Math. 31, (1977), 204-256.
[5] Harry Furstenberg, Recurrence in ergodic theory and combinatorial number theory, Princeton university press, Princeton, N.J., 1981.
[6] Eli Glasner, Ergodic theory via joinings, Mathematical Surveys and Monographs, 101, American Mathematical Society, Providence, RI, 2003. xii+384 pp. ISBN: 0-8218-3372-3
[7] Roger Jones and William Parry, Compact abelian group extensions of dynamical systems II, Compositio Mathematica, tome 25, no 2 (1972), p. 135–147.
[8] Alexander S. Kechris, Classical descriptive set theory, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995.
[9] William Parry, Zero entropy of distal and related transformations, in Topological dynamics, J. Auslader and W. Gottschalk, eds., Benjamin, New York, 1967.
[10] Endre Szemerédi, On sets of integers containing no $k$ elements in arithmetic progression, Acta Arith. 27, (1975), 199-245.
[11] Benjamin Weiss, Strictly ergodic models for dynamical systems, Bull. of the Amer. Math. Soc. 13, (1985), 143–146.
[12] Robert J. Zimmer, Extensions of ergodic group actions, Illinois J. Math. 20, (1976), 373-409.
[13] Robert J. Zimmer, *Ergodic actions with generalized discrete spectrum*, Illinois J. Math. 20, (1976), 555-588.

Department of Mathematics, Tel Aviv University, Tel Aviv, Israel

E-mail address: glasner@math.tau.ac.il

Institute of Mathematics, Hebrew University of Jerusalem, Jerusalem, Israel

E-mail address: weiss@math.huji.ac.il