Extensions of probability-preserving systems by measurably-varying homogeneous spaces and applications

by

Tim Austin (Los Angeles, CA)

Abstract. We study a generalized notion of a homogeneous skew-product extension of a probability-preserving system in which the homogeneous space fibres are allowed to vary over the ergodic decomposition of the base. The construction of such extensions rests on a simple notion of 'direct integral' for a 'measurable family' of homogeneous spaces, which has a number of precedents in older literature. The main contribution of the present paper is the systematic development of a formalism for handling such extensions, including non-ergodic versions of the results of Mackey describing ergodic components of such extensions, of the Furstenberg–Zimmer structure theory and of results of Mentzen describing the structure of automorphisms of such extensions when they are relatively ergodic. We then offer applications to two structural results for actions of several commuting transformations: firstly to describing the possible joint distributions of three isotropy factors corresponding to three commuting transformations; and secondly to describing the characteristic factors for a system of double non-conventional ergodic averages. Although both applications are modest in themselves, we hope that they point towards a broader usefulness of this formalism in ergodic theory.

Contents

[133]

1. Introduction. This work is concerned with probability-preserving actions $T : \Gamma \curvearrowright (X,\mu)$ of locally compact second countable topological groups on standard Borel probability spaces. We often denote such an action by (X, μ, T) if the group is understood.

One of the more versatile constructions by which a more complicated system may be built from comparatively simple ingredients is the homogeneous skew-product (see, for example, Examples 2.21 in Glasner [28]). From some given Γ-system (Y, ν, S), a compact group G and a closed subgroup $K \leq G$, and a measurable cocycle $\rho : \Gamma \times Y \to G$ for the action S, we form the system $(Y \times G/K, \nu \otimes m_{G/K}, T)$ by setting

$$T^\gamma(y, gK) := (S^\gamma y, \rho(\gamma, y)gK) \quad \text{for } \gamma \in \Gamma.$$

A well-developed theory of such systems is available in case the base system (Y, ν, S) is ergodic (much of which can be extended to the setting in which ν is only quasi-invariant under S; see [7]). In addition to providing a wealth of example systems, such homogeneous skew-products over ergodic base systems acquire a greater significance through the structure theory developed by Zimmer in [49, 48] and Furstenberg in [22]. This is concerned with the failure of relative weak mixing of extensions (see, for example, Definition 9.22 in Glasner [28]). Relative weak mixing is a strengthening of the condition of relative ergodicity which has numerous consequences for how this extension may be joined to others. An understanding of these consequences and of the ways in which relative weak mixing can fail is crucial to Furstenberg's approach to Szemerédi's Theorem ([22]; see also the excellent treatment in [23]). The core result of Furstenberg and Zimmer is an inverse theorem according to which an extension of ergodic systems fails to be relatively weakly mixing if and only if it contains a non-trivial subextension that can be coordinatized as a homogeneous skew-product: thus, homogeneous skew-products are identified as precisely the 'obstructions' to relative weak mixing within other extensions.

However, in many applications in which this ergodicity of the base system fails, this simple homogeneous skew-product construction is not quite general enough, and the Furstenberg–Zimmer theory outlined above is not available without modification.

In this paper we shall extend the definition of homogeneous skew-product to a more general class of systems by the simple artifice of allowing the fibre G_y/K_y, in addition to the cocycle $\rho(y)$, to vary as a function of the base point $y \in Y$. This leads to a definition of an extension of (Y, ν, S) given as an action of the group on a 'direct integral' of homogeneous space fibres over (Y, ν). It is relatively simple to settle on a definition of 'measurability' for such an assignment of fibres, and to turn this idea into a rigorous definition.
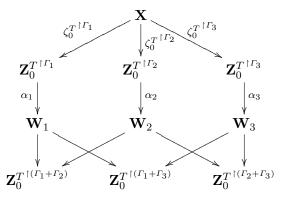
The study of such measurably-varying groups is certainly not new. It is already alluded to during the introduction to Section 6.1 of Guichardet's book [29] in the context of 'measurable current groups' and their representation theory, motivated in turn by considerations from algebraic quantum field theory. A number of more recent works have studied constructions of this nature explicitly. For example, in [15] Conze and Raugi analyze the ergodic decomposition of various σ-finite measures associated to an extension of a non-singular ergodic base transformation by a locally compact non-compact group, invoking for their description a measurably-varying family of subgroups of that fibre group. However, in their setting the need for a measurably-varying family of groups is related to the possibly non-smooth structure of the Borel equivalence relation of conjugacy on the collection of all such subgroups, an issue which disappears upon restriction to the case of compact fibre groups, and so the results that they develop are still rather disconnected from the present paper. Perhaps closest to our present emphasis is the emergence of measurably-varying subgroups of a fixed compact group in the analysis of measurably-varying Mackey groups for certain non-ergodic self-joinings of an ergodic system, which underlies some known approaches to the study of non-conventional ergodic averages; see, in particular, Section 3.2 in Meiri [35], the proof of Lemma 9.2 in Furstenberg and Weiss [27], and Section 6.8 in Ziegler [47].

More generally, a need to extend known machinery for different kinds of ergodic system to their non-ergodic relatives has been felt in other areas: consider, for example, Downarowicz' discussion in [16] of 'assignments' for his study of simplices of invariant measures for topological systems on zero-dimensional metric spaces, and the work of Fisher, Witte-Morris and Whyte on cocycle superrigidity for non-ergodic systems [18]. The careful study of such non-ergodic settings also has many parallels with the analysis of direct integrals in the infinite-dimensional representation theory of locally compact groups or von Neumann and C*-algebras (nicely introduced, for instance, in Arveson [1]), although we will not explore this connection further here.

Notwithstanding the diversity of these previous developments, within structural ergodic theory the treatment of extensions with varying homogeneous space fibres seems to have stayed largely immersed in other analyses, such as those cited above. Although it is intuitively clear that the fundamental structural results for the ergodic case of homogeneous skew-product extensions should admit natural generalizations, it seems that this has not yet been carried out. In fact, after setting up the right definitions we will find that it is largely routine to extend both the results of Mackey ([33]) on the invariant factor of the extended system and also the Furstenberg–Zimmer theory to this setting. We lay out the details of this generalization in the first part of this paper. More interesting is the extension of the results of Mentzen [36] on the possible structure of an automorphism of an isometric extension of ergodic systems: we generalize this by presenting structure theorems for factors and automorphisms of a relatively ergodic extension by measurably-varying compact homogeneous spaces.

Although this generalization is as much a matter of care as new ideas, it pays off by broadening the applicability of the theory of homogeneous skew-products to settings in which an assumption of base ergodicity is unavailable. This arises, in particular, when considering an action of a larger group T : $\Gamma \curvearrowright (X, \mu)$ restricted to some subgroup $\Lambda \leq \Gamma$. Although a routine appeal to the ergodic decomposition can often justify the assumption that T is ergodic overall, if we disintegrate further to guarantee that the restricted action $T^{\restriction \Lambda}$ is ergodic then T^γ for $\gamma \in \Gamma \setminus \Lambda$ need not preserve the resulting disintegrands of μ. As a consequence, if we are concerned with how the Λ-subaction sits within the whole original action, we may be forced to retain a system for which this subaction of Λ is not ergodic.

In this paper we offer two closely-related applications meeting this description. For both cases we specialize to $\Gamma = \mathbb{Z}^d$. These two applications are relatively simple, and are included largely to illustrate the arguments made possible by the formalism described above, but they also exemplify much more general questions on which we suspect these methods will shed light in the future.

Given a $\mathbb{Z}^d$-system $\mathbf{X} = (X, \mu, T)$ we can consider the σ-subalgebra $\Sigma_X^{T^{\restriction \Gamma}}$ of sets left invariant by the subaction of T corresponding to some subgroup $\Gamma \leq \mathbb{Z}^d$. As is standard in the category of standard Borel spaces, this can be identified up to μ-negligible sets with the σ-algebra generated by a factor map $\zeta_0^{T^{\restriction \Gamma}} : \mathbf{X} \to \mathbf{Z}_0^{T^{\restriction \Gamma}}$ to some new system on which the subaction of Γ is trivial. Although individually these new systems can still be quite complicated, a possibly more tractable task is to describe their possible joint distributions within the original system. If $\Gamma, \Lambda \leq \mathbb{Z}^d$ are two subgroups then it is easy to show that $\zeta_0^{T^{\restriction \Gamma}}$ and $\zeta_0^{T^{\restriction \Lambda}}$ are relatively independent over $\zeta_0^{T^{\restriction (\Gamma + \Lambda)}}$, but for three or more subgroups matters become more complicated.

Clearly given three subgroups $\Gamma_1, \Gamma_2, \Gamma_3 \leq \mathbb{Z}^d$ we see that $\zeta_0^{T\upharpoonright\Gamma_1}$ and $\zeta_0^{T\upharpoonright\Gamma_2}$ both contain $\zeta_0^{T\upharpoonright(\Gamma_1+\Gamma_2)}$, and similarly for other pairs, and so a naïve candidate for a generalization of the above result could be that the three isotropy factors $\zeta_0^{T\upharpoonright\Gamma_i}$ are relatively independent over the smaller triple of factors $\zeta_0^{T\upharpoonright(\Gamma_i+\Gamma_j)} \vee \zeta_0^{T\upharpoonright(\Gamma_i+\Gamma_k)}$ (the factor generated by $\zeta_0^{T\upharpoonright(\Gamma_i+\Gamma_j)}$ and $\zeta_0^{T\upharpoonright(\Gamma_i+\Gamma_k)}$ together) as (i, j, k) ranges over permutations of $(1, 2, 3)$. If we denote the target $\mathbb{Z}^d$-system of this joint factor map by $\mathbf{W}_i$ (so this is a joining of $\mathbf{Z}_0^{T\upharpoonright(\Gamma_i+\Gamma_j)}$ and $\mathbf{Z}_0^{T\upharpoonright(\Gamma_i+\Gamma_k)}$) and let $\alpha_i : \mathbf{Z}_0^{T\upharpoonright\Gamma_i} \to \mathbf{W}_i$ be the factor map defined by $\zeta_0^{T\upharpoonright(\Gamma_i+\Gamma_j)} \vee \zeta_0^{T\upharpoonright(\Gamma_i+\Gamma_k)} = \alpha_i \circ \zeta_0^{T\upharpoonright\Gamma_i}$, then these factors are arranged as in the following commutative diagram:

In fact the naïve conjecture that the factors $\zeta_0^{T\upharpoonright\Gamma_i}$ are relatively independent over their further factors α_i is false, but 'not by very much': we will see that it can fail only in a very restricted way. In general, the three factors $\zeta_0^{T\upharpoonright\Gamma_i}$ are relatively independent over some subextensions of these 'natural candidate' factors α_i, and these subextensions can be coordinatized by measurable compact fibre groups subject to certain further restrictions.

Here we will examine this when $d = 3$ and Γ_i is the cyclic subgroup $\mathbb{Z}\mathbf{e}_i$ in the direction of a basis vector $\mathbf{e}_i$, but it seems clear that our methods can be extended both to more general subgroups of Abelian groups and (probably with considerably more work) to larger numbers of subgroups.

THEOREM 1.1 (Joint distributions of three isotropy factors). *Let* $\mathbf{X} = (X, \mu, T)$ *be a* $\mathbb{Z}^3$-*system and write* $T_i := T^{\mathbf{e}_i}$ *for* $i = 1, 2, 3$. *Let* $\mathbf{W}_i$ *be the target of the joint factor map* $\alpha_i := \zeta_0^{T_i, T_j} \vee \zeta_0^{T_i, T_k}$, *where* $\zeta_0^{T_i, T_j} := \zeta_0^{T\upharpoonright(\mathbb{Z}\mathbf{e}_i+\mathbb{Z}\mathbf{e}_j)}$, *let* W_i *be its underlying standard Borel space and let* $T_j|_{\alpha_i}$ *be the restriction of* T_j *to the factor* α_i. *Between the single isotropy factors* $\zeta_0^{T_i} : \mathbf{X} \to \mathbf{Z}_0^{T_i}$ *and the smaller factors* $\alpha_i : \mathbf{X} \to \mathbf{W}_i$ *there are three intermediate factors* $\phi_i \circ \zeta_0^{T_i} : \mathbf{X} \to \mathbf{V}_i$, *where*

$$\mathbf{X} \xrightarrow{\zeta_0^{T_i}} \mathbf{Z}_0^{T_i} \xrightarrow{\phi_i} \mathbf{V}_i \to \mathbf{W}_i,$$
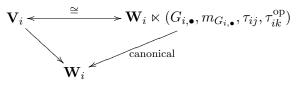
such that

- *the triple of factors $\zeta_0^{T_i}$ is relatively independent over the triple $\phi_i \circ \zeta_0^{T_i}$ under μ;*
- *there exist compact metrizable group data $G_{i,\bullet}$ on W_i invariant under the restriction of the whole action T to the factor space W_i, a cocycle $\tau_{ij} : W_i \to G_{i,\bullet}$ invariant under the restriction of T_k to W_i and a cocycle $\tau_{ik} : W_i \to G_{i,\bullet}$ invariant under the restriction of T_j to W_i such that we can coordinatize the extension $\mathbf{V}_i \to \mathbf{W}_i$ as the extension of $\mathbf{W}_i$ by the measurable compact fibre groups $G_{i,\bullet}$ with the lifted actions defined by*

$$T_j|_{\phi_i \circ \zeta_0^{T_i}}(w_i, g_i) = (T_j|_{\alpha_i}(w_i), \tau_{ij}(w_i) \cdot g_i)$$

and

$$T_k|_{\phi_i \circ \zeta_0^{T_i}}(w_i, g_i) = (T_k|_{\alpha_i}(w_i), g_i \cdot \tau_{ik}(w_i)).$$

We will generally denote a coordinatization of the extension $\mathbf{V}_i \to \mathbf{W}_i$ as above by the commutative diagram

$$\mathbf{V}_i \xleftarrow{\quad\cong\quad} \mathbf{W}_i \ltimes (G_{i,\bullet}, m_{G_{i,\bullet}}, \tau_{ij}, \tau_{ik}^{\mathrm{op}})$$

$$\mathbf{W}_i$$

canonical

where we use the superscript $^{\mathrm{op}}$ to denote a cocycle that acts on fibres by right multiplication, and we have suppressed mention of the transformation T_i since by definition its restriction to $\mathbf{V}_i$ is the identity.

Although our final conclusion here yields measurably-varying fibre groups $G_{i,\bullet}$ that are invariant under the whole action T—and so would be constant if we had assumed that the overall action T is ergodic—the analysis leading to this conclusion will go via homogeneous space fibres of possibly greater variability, for the reason described earlier that at first we will only be able to assume that the fibres are invariant under the subaction T_i.

The same is true of our second application. This is to a special case of the problem of describing the 'minimal characteristic factors' for the non-conventional ergodic averages

$$\frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^{d} f_i \circ T_i^n$$

associated to a d-tuple of commuting actions $T_i : \mathbb{Z} \curvearrowright (X, \mu)$ and functions $f_1, \ldots, f_d \in L^\infty(\mu)$. Let us write $\mathbf{X} = (X, \mu, T)$ for the $\mathbb{Z}^d$-system given by these one-dimensional actions in the coordinate directions.

The question of convergence in $L^2(\mu)$ for such averages was first settled when $d = 2$ by Conze and Lesigne in [12], and since then a number of other

works have addressed other versions or relatives of this question [45, 27, 30], culminating in Host and Kra's detailed analysis of the case in which $T_i = T^i$ for some fixed T in [31] (see also Ziegler [47]) and Tao's recent proof in [42] of convergence for arbitrary d. We direct the reader to [6] for a more detailed discussion of this problem and an alternative proof of convergence.

Here we will consider in the case $d = 2$ an important part of these developments: the theory of 'characteristic factor-tuples' for such averages. In our setting, a pair of factors $\xi_i : \mathbf{X} \to \mathbf{Y}_i$ is 'characteristic' if
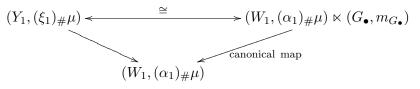
$$\frac{1}{N} \sum_{n=1}^{N} (f_1 \circ T_1^n) \cdot (f_2 \circ T_2^n) - \frac{1}{N} \sum_{n=1}^{N} \mathsf{E}_\mu(f_1 \circ T_1^n \,|\, \xi_1) \cdot \mathsf{E}_\mu(f_2 \circ T_2^n \,|\, \xi_2) \to 0$$

in $L^2(\mu)$ as $N \to \infty$ for any $f_1, f_2 \in L^\infty(\mu)$. Clearly given such a pair of factors, the problem of proving convergence reduces to the case when each f_i is ξ_i-measurable, and this reduction forms an important first step in many of the known proofs of convergence. Although characteristic factors are well-understood in some special cases, the more recent proofs of general convergence in [42, 43, 6] proceed by first heavily modifying the original system and only then asking after the characteristic factors (or their finitary analog in Tao's proof in [42]), and so our knowledge of the characteristic factors of the original system remains incomplete except in some special cases [45, 31, 19, 47]. More is known in the case $d = 2$ from the work of Conze and Lesigne [12], and in addition the following very precise description of the characteristic factors when $d = 2$ has achieved folkloric currency since that work appeared. However, a complete proof seems to be surprisingly subtle, and we shall give such a proof as our second application of our non-ergodic machinery for extensions by homogeneous spaces.

THEOREM 1.2 (Characteristic factors for double non-conventional averages). *Given a $\mathbb{Z}^2$-system $\mathbf{X} = (X, \mu, T_1, T_2)$, let $\mathbf{W}_i$ be the target system of the factor $\alpha_i := \zeta_0^{T_i} \vee \zeta_0^{T_1 T_2^{-1}}$ with underlying standard Borel space W_i. Then $\mathbf{X}$ admits a characteristic pair of factors $\xi_i : \mathbf{X} \to \mathbf{Y}_i$ with underlying standard Borel spaces Y_i that extend the factors $\mathbf{X} \to \mathbf{W}_i$ and can be described as follows: there are*

- *a T-invariant measurable family of compact fibre groups $G_\bullet$,*
- *a T_1-ergodic cocycle $\sigma : \mathbf{Z}_0^{T_1 T_2^{-1}} \to G_\bullet$ that is ergodic for the restricted action of T_1,*
- *and a pair of cocycles $\tau_i : \mathbf{Z}_0^{T_{3-i}} \to G_\bullet$ ergodic for the restricted action of T_i*

such that we can coordinatize these probability spaces as

$$(Y_1, (\xi_1)_\# \mu) \xleftarrow{\quad\cong\quad} (W_1, (\alpha_1)_\# \mu) \ltimes (G_\bullet, m_{G_\bullet})$$

canonical map

$$(W_1, (\alpha_1)_\# \mu)$$

so that the restricted actions are given by

$$\text{restriction of } T_1 \colon (w_1, g) \mapsto \left(T_1|_{\alpha_1}(w_1), \sigma(\zeta_0^{T_1 T_2^{-1}}(w_1)) \cdot g\right),$$

$$\text{restriction of } T_2 \colon (w, g) \mapsto \left(T_2|_{\alpha_1}(w_1), \sigma(\zeta_0^{T_1 T_2^{-1}}(w_1)) \cdot g \cdot \tau_2(\zeta_0^{T_1}(w_1))\right),$$

and similarly

$$(Y_2, (\xi_2)_\# \mu) \xleftarrow{\quad\cong\quad} (W_2, (\alpha_2)_\# \mu) \ltimes (G_\bullet, m_{G_\bullet})$$

canonical map

$$(W_2, (\alpha_2)_\# \mu)$$

with

$$\text{restriction of } T_1 \colon (w_2, g) \mapsto \left(T_1|_{\alpha_2}(w_2), \sigma(\zeta_0^{T_1 T_2^{-1}}(w_2)) \cdot g \cdot \tau_1(\zeta_0^{T_2}(w_2))\right),$$

$$\text{restriction of } T_2 \colon (w_2, g) \mapsto \left(T_2|_{\alpha_2}(w_2), \sigma(\zeta_0^{T_1 T_2^{-1}}(w_2)) \cdot g\right).$$

We suspect that our methods should extend to offer at least some description of characteristic factor-tuples for larger numbers of commuting transformations, although we also suspect that it will become rapidly more complicated.

In summary, the body of this paper is organized as follows.

In Section 2 we recall some definitions and standard results from group theory, measure theory and ergodic theory that we will need later in the paper, and in doing so set up some convenient notation.

Section 3 introduces our definitions of measurable families of homogeneous space data and their direct integrals.

In Section 4 we cover quite briskly the main definitions and results of the non-ergodic Mackey theory, and then in Section 5 we treat similarly the non-ergodic version of the Furstenberg–Zimmer inverse theory.

In Section 6 we pursue a slightly less standard consequence of the Mackey theory, using it first to describe the possible factors and groups of automorphisms of an extension by homogeneous space data, and then translating this into conditions on an automorphism of a base system that it be liftable to an automorphism of an extension. This generalizes the classical work of Mentzen [36] in the case of ergodic systems, and will be important for the applications of the theory that follow.

In Section 7 we present our two applications, to the joint three-fold distributions of isotropy factors and to double characteristic factors.

Finally, in Section 8 we discuss some further possible applications of this machinery.

2. Background and notation

2.1. Measurable functions and probability kernels. We will work exclusively in the category of standard Borel probability spaces (X, Σ_X, μ), and so will often suppress mention of their σ-algebras.

Any Borel map $\phi : X \to Y$ specifies a σ-subalgebra of Σ_X in the form of $\phi^{-1}(\Sigma_Y)$. Two such maps $\phi : X \to Y$ and $\psi : X \to Z$ are *equivalent* if the σ-subalgebras of Σ_X that they generate are equal up to μ-negligible sets, in which case we shall write $\phi \simeq \psi$; this clearly defines an equivalence relation among Borel maps with domain X. As is standard, in the category of standard Borel spaces equivalence classes of such Borel maps are in bijective correspondence with equivalence classes of σ-subalgebras under the relation of equality modulo the σ-ideal of μ-negligible sets. A treatment of these classical issues may be found, for example, in Chapter 2 of Glasner [28].

A measure-preserving Borel map $\pi : (X, \mu) \to (Y, \nu)$ *contains* another such map $\psi : (X, \mu) \to (Z, \theta)$ if $\pi^{-1}(\Sigma_Y) \supseteq \psi^{-1}(\Sigma_Z)$ up to μ-negligible sets. In this case we shall write $\pi \succsim \psi$ or $\psi \precsim \pi$, and sometimes say that ψ is *μ-virtually a function of ϕ* or that it is *μ-virtually $\phi^{-1}(\Sigma_Y)$-measurable*. It is again a classical fact that in the category of standard Borel spaces this notion of containment is equivalent to the existence of a *factorizing* Borel map $\phi : (Y, \nu) \to (Z, \theta)$ with $\psi = \phi \circ \pi$ μ-almost everywhere, and that a measurable analog of the Schroeder–Bernstein Theorem holds: $\pi \simeq \psi$ if and only if in each direction such a ϕ may be chosen that is invertible away from some negligible subsets of the domain and target. It is clear that (up to set-theoretic niceties) this defines a partial order on the class of $\simeq$-equivalence classes of Borel maps out of the given space (X, μ).

Measure-respecting Borel maps from one probability space to another comprise the simplest class of morphisms between such spaces, but in this paper we shall sometimes find ourselves handling also a weaker class of morphisms. Suppose that Y and X are standard Borel spaces. Then by a *probability kernel from Y to X* we understand a function $P : Y \times \Sigma_X \to [0, 1]$ such that

- the map $y \mapsto P(y, A)$ is Σ_Y-measurable for every $A \in \Sigma_X$;
- the map $A \mapsto P(y, A)$ is a probability measure on Σ_X for every $y \in Y$.

Intuitively, such a kernel amounts to a 'randomized map' from Y to X: rather than specify a unique image in X for each point $y \in Y$, it specifies only a probability distribution $P(y, \cdot)$ from which a point of X could be chosen. The first of the above conditions is then the natural sense in which this assignment of a probability distribution is measurable in y; indeed, a

popular alternative definition of probability kernel is as a measurable function from Y to the set $\operatorname{Pr} X$ of Borel probability measures on X. In ergodic theory this notion lies behind that of a 'quasifactor' (which assumes also a certain equivariance of this map): see, for example, Chapter 8 of Glasner [28], where this alternative convention and notation are used. We will write $P : Y \xrightarrow{\mathrm{p}} X$ when P is a probability kernel from Y to X.

Given a kernel $P : Y \xrightarrow{\mathrm{p}} X$ and a probability measure ν on Y, we define the measure $P_{\#}\nu$ on X by

$$P_{\#}\nu(A) := \int_Y P(y, A)\,\nu(\mathrm{d}y);$$

this measure on X can be interpreted as the law of a member of X selected randomly by first selecting a member of Y with law ν and then selecting a member of X with law $P(y, \cdot)$. By analogy with the case of a function between measurable spaces, we will refer to this as the *pushforward* of ν by P. This extends standard deterministic notation: given a measurable function $\phi : Y \to X$, we may associate to it the deterministic probability kernel given by $P(y, \cdot) = \delta_{\phi(y)}$ (the point mass at the image of y under ϕ), and now $P_{\#}\nu$ is the usual pushforward measure $\phi_{\#}\nu$.

Certain special probability kernels naturally serve as adjoints to factor maps, in the sense of the following theorem.

THEOREM 2.1. *Suppose that Y and X are standard Borel spaces, that μ is a probability measure on X and that $\phi : X \to Y$ is a measurable factor map. Then, denoting the pushforward $\phi_{\#}\mu$ by ν, there is a ν-almost surely unique probability kernel $P : Y \xrightarrow{\mathrm{p}} X$ such that $\mu = P_{\#}\nu$ and which represents the conditional expectation with respect to ϕ: for any $f \in L^1(\mu)$, the function*

$$x_1 \mapsto \int_X f(x)\,P(\phi(x_1), \mathrm{d}x)$$

is a version of the μ-conditional expectation of f with respect to $\phi^{-1}(\Sigma_Y)$.

We also write that this P represents the disintegration of μ over ϕ. A general probability kernel $P : Y \xrightarrow{\mathrm{p}} X$ represents the disintegration over ϕ of some measure that pushes forward onto ν if and only if $\int_A P(x, \cdot)\,\nu(\mathrm{d}y)$ and $\int_B P(y, \cdot)\,\nu(\mathrm{d}y)$ are mutually singular whenever $A \cap B = \emptyset$.

Proof. See Theorem 6.3 in Kallenberg [32]. ∎

2.2. Systems, subactions and factors. In this paper we shall spend a great deal of time passing up and down from systems to extensions or factors. Moreover, sometimes one system will appear as a factor of a 'larger' system in *several different ways* (most obviously, when we work with a system that appears under each coordinate projection from some self-joining). For this

reason the notational abuse of referring to one system as a factor of another but leaving the relevant factor map to the understanding of the reader, although popular and useful in modern ergodic theory, seems dangerous here, and we shall carefully avoid it. In its place we substitute the alternative abuse, slightly safer in our circumstances, of often referring only to the factor maps we use, and leaving either their domain or target systems to the reader's understanding. Let us first set up some notation to support this practice.

If Γ is a locally compact second countable topological group, by a Γ-*system* (or, if Γ is clear, just a *system*) we understand a jointly measurable probability-preserving action $T : \Gamma \curvearrowright (X, \mu)$ on a standard Borel probability space. We will often alternatively denote this space and action by (X, μ, T), or by a corresponding single boldface letter such as $\mathbf{X}$. If $\Lambda \leq \Gamma$ we denote by $T^{\restriction \Lambda} : \Lambda \curvearrowright (X, \mu)$ the action defined by $(T^{\restriction \Lambda})^\gamma := T^\gamma$ for $\gamma \in \Lambda$, and refer to this as a *subaction*, and if $\mathbf{X} = (X, \mu, T)$ is a Γ-system we write similarly $\mathbf{X}^{\restriction \Lambda}$ for the system $(X, \mu, T^{\restriction \Lambda})$ and refer to it as a *subaction system*.

A *factor* from one system (X, μ, T) to another (Y, ν, S) is a Borel map $\pi : X \to Y$ with $\pi_\# \mu = \nu$ and $\pi \circ T = S \circ \pi$. Given such a factor, we sometimes write $T|_\pi$ to denote the action S with which T is intertwined by π.

In this paper, given a globally invariant σ-subaglebra in $\mathbf{X}$, a choice of factor $\pi : \mathbf{X} \to \mathbf{Y}$ generating that σ-subalgebra will sometimes be referred to as a *coordinatization* of the σ-subalgebra. Importantly for us, some choices of a coordinatizing factor π may reveal some underlying structure more clearly than others, and so we will sometimes need to pass between coordinatizing factors. Given one coordinatization $\pi : \mathbf{X} \to \mathbf{Y}$ and an isomorphism $\psi : \mathbf{Y} \to \mathbf{Z}$, we shall sometimes refer to the composition $\psi \circ \pi$ as a *recoordinatization* of π. We will also extend this terminology to that of *coordinatizations* and *recoordinatizations* of families of factors of a system in the obvious way in terms of the appropriate commutative diagram of isomorphisms.

Given a Γ-system $\mathbf{X} = (X, \mu, T)$, the σ-algebra Σ_X^T of sets $A \in \Sigma_X$ for which $\mu(A \triangle T^\gamma(A)) = 0$ for all $\gamma \in \Gamma$ is T-invariant, so defines a factor of $\mathbf{X}$. More generally, if Γ is Abelian and $\Lambda \leq \Gamma$ then we can consider the σ-algebra $\Sigma_X^{T^{\restriction \Lambda}}$ generated by all $T^{\restriction \Lambda}$-invariant sets: we refer to this as the Λ-*isotropy factor* and write $\mathbf{Z}_0^{T^{\restriction \Lambda}}$ for some new system that we adopt as the target for a factor map $\zeta_0^{T^{\restriction \Lambda}}$ that generates $\Sigma_X^{T^{\restriction \Lambda}}$, and $Z_0^{T^{\restriction \Lambda}}$ for the standard Borel space underlying $\mathbf{Z}_0^{T^{\restriction \Lambda}}$. Note that in this case the Abelianness condition (or, more generally, the condition that $\Lambda \trianglelefteq \Gamma$) is needed for this to be a globally T-invariant factor. If T_1 and T_2 are two commuting actions of the same

Abelian group Γ on (X, μ) then we can define a third action $T_1 T_2^{-1}$ by setting $(T_1 T_2^{-1})^\gamma := T_1^\gamma T_2^{\gamma^{-1}}$, and in this case we may write $\zeta_0^{T_1 = T_2} : \mathbf{X} \to \mathbf{Z}_0^{T_1 = T_2}$ in place of $\zeta_0^{T_1 T_2^{-1}} : \mathbf{X} \to \mathbf{Z}_0^{T_1 T_2^{-1}}$. If $S \subseteq \Gamma$ and Λ is the group generated by S, we will sometimes write $\mathbf{Z}_0^{T \upharpoonright S}$ in place of $\mathbf{Z}_0^{T \upharpoonright \Lambda}$, and similarly.

An important construction of new systems from old is that of *relatively independent products*. If $\mathbf{Y} = (Y, \nu, S)$ is some fixed system and $\pi_i : \mathbf{X}_i = (X_i, \mu_i, T_i) \to \mathbf{Y}$ is an extension of it for $i = 1, \ldots, k$ then we define the relatively independent product of the systems $\mathbf{X}_i$ over their factor maps π_i to be the system

$$\prod_{\{\pi_1 = \cdots = \pi_k\}} \mathbf{X}_i = \Big(\prod_{\{\pi_1 = \cdots = \pi_k\}} X_i, \bigotimes_{\{\pi_1 = \cdots = \pi_k\}} \mu_i, T_1 \times \cdots \times T_k \Big)$$

where

$$\prod_{\{\pi_1 = \cdots = \pi_k\}} X_i := \{(x_1, \ldots, x_k) \in X_1 \times \cdots \times X_k : \pi_1(x_1) = \cdots = \pi_k(x_k)\},$$

$$\bigotimes_{\{\pi_1 = \cdots = \pi_k\}} \mu_i = \int_Y \bigotimes_{i=1}^k P_i(y, \cdot) \, \nu(\mathrm{d}y)$$

and $P_i : Y \xrightarrow{\mathrm{p}} X_i$ is a probability kernel representing the disintegration of μ_i over π_i. In case $k = 2$ we will write this instead as $\mathbf{X}_1 \times_{\{\pi_1 = \pi_2\}} \mathbf{X}_2$, and in addition if $\mathbf{X}_1 = \mathbf{X}_2 = \mathbf{X}$ and $\pi_1 = \pi_2 = \pi$ then we will abbreviate this further to $\mathbf{X} \times_\pi \mathbf{X}$, and similarly for the individual spaces and measures.

2.3. Measurable selectors. At several points in this paper we need to appeal to some basic results on the existence of measurable selectors, often as a means of making rigorous a selection of representatives of one or another kind of data above the ergodic components of a non-ergodic system.

THEOREM 2.2. *Suppose that (X, Σ_X) and (Y, Σ_Y) are standard Borel spaces, that $A \subseteq X$ is Borel and that $\pi : X \to Y$ is a Borel surjection. Then the image $\pi(A)$ lies in the ν^c-completion of Σ_Y for every Borel probability measure ν on (Y, Σ_Y) with completion ν^c, and for any such ν there is a map $f : B \to A$ with domain $B \in \Sigma_Y$ such that $B \subseteq \pi(A)$, $\nu^c(\pi(A) \setminus B) = 0$ and $\pi \circ f = \mathrm{id}_B$.* ∎

Proof. See, for example, 423O and its consequence 424X(h) in Fremlin [21]. ∎

DEFINITION 2.3 (Measurable selectors). We refer to a map f as given by the above theorem as a *measurable selector* for the set A.

REMARK. We should stress that this is only one of several versions of the 'measurable selector theorem', due variously to von Neumann, Jankow,

Lusin and others. Note in particular that in some other versions a map f is sought that selects points of A for *strictly* all points of $\pi(A)$. In the above generality we cannot guarantee that a strictly-everywhere selector f is Borel, but only that it is Souslin-analytic and hence universally measurable (of course, from this the above version follows at once). On the other hand, if the map $\pi|_A$ is countable-to-one, then a version of the result due to Lusin does guarantee a strictly-everywhere Borel selector f. This version has already played a significant rôle in our corner of ergodic theory in the manipulation of the Conze–Lesigne equations (see, for example, [12, 27, 10]), and so we should be careful to distinguish it from the above. A thorough account of all these different results and their proofs can be found in Sections 423, 424 and 433 of Fremlin [21].

In the right circumstances it is possible to strengthen Theorem 2.2 to obtain a Borel selector that is invariant under a group of transformations, by making use of a coordinatization of the invariant factor.

PROPOSITION 2.4. *Suppose that (X, Σ_X) and (Y, Σ_Y) are standard Borel spaces, $A \subseteq X$ is Borel and $\pi : X \to Y$ is a surjective Borel map, and in addition that $T : \Gamma \curvearrowright (X, \Sigma_X)$ is a jointly measurable action of a locally compact second countable group such that π is a factor map, so $\pi \circ T^\gamma = S^\gamma \circ \pi$ for some jointly measurable action $S : \Gamma \curvearrowright (Y, \Sigma_Y)$, and that A is T-invariant. Then for any S-invariant probability measure ν on (Y, Σ_Y) with completion ν^c there are an S-invariant set $B \in \Sigma_Y$ such that $B \subseteq \pi(A)$ and $\nu^c(\pi(A) \setminus B) = 0$ and an S-invariant map $f : B \to A$ such that $\pi \circ f = \mathrm{id}_B$.*

Proof. Let $f_0 : B_0 \to A$ be an ordinary measurable selector as given by Theorem 2.2, and let ν be any S-invariant probability measure on (Y, Σ_Y). This B_0 must be ν-almost S-invariant, simply because $\pi(A)$ is T-invariant and $\nu^c(B \bigtriangleup \pi(A)) = 0$. Using local compactness and second countability, let $(F_i)_{i \geq 1}$ be a countable compact cover of Γ, and also let m_Γ be a left-invariant Haar measure on Γ. From the joint measurability of T it follows that the set

$$B := \big\{ y \in Y : m_\Gamma \{ \gamma \in \Gamma : S^\gamma(y) \in Y \setminus B_0 \} = 0 \big\}$$
$$= \bigcap_{i \geq 1} \big\{ y \in Y : m_\Gamma \{ \gamma \in F_i : S^\gamma(y) \in B_0 \} = m_\Gamma(F_i) \big\}$$

is Borel, T-invariant and satisfies $\nu(B_0 \bigtriangleup B) = 0$.

We now let $\zeta : (Y, \Sigma_Y, \nu) \to (Z, \Sigma_Z, \theta)$ be any coordinatization of the invariant factor Σ_Y^T; it is easy to see that this may be chosen so that there exists some $C \in \Sigma_Z$ such that $B = \zeta^{-1}(C)$. We can now use B and ζ to 'tidy up' our original selector f_0. Indeed, by the S-invariance of ζ and the fact that for every $y \in B$ we have $S^\gamma(y) \in B_0$ for some (indeed, almost all) $\gamma \in \Gamma$,

we must have $\zeta(B_0) \supseteq C$. Therefore by applying the ordinary Measurable Selector Theorem a second time we can find a Borel subset $D \in \Sigma_Z$ with $D \subseteq C$ and $\theta(C \setminus D) = 0$ and a Borel section $\eta : D \to B_0$ such that $\zeta \circ \eta = \mathrm{id}_D$; and so now replacing B_0 with B and the map f_0 with $f : y \mapsto f_0(\eta(\zeta(y)))$ completes the proof. ■

DEFINITION 2.5 (Invariant measurable selectors). We refer to a map f as given by the above proposition as a *T-invariant measurable selector* for the set A.

2.4. Background from group theory. We collect here some standard group-theoretic definitions and results for future reference.

DEFINITION 2.6 (Core). If G is a group and $H \leq G$ we denote by $\mathrm{Core}_G(H)$ the *core of H in G*: the largest subgroup of H that is normal in G. It is clear that this exists and equals $\bigcap_{g \in G} g^{-1}Hg$. If G is compact and H is closed then so is $\mathrm{Core}_G(H)$.

If $\mathrm{Core}_G(H) = \{1_G\}$ we shall write that H is *core-free* in G.

DEFINITION 2.7 (Full one-dimensional projections; slices). Given two groups G_1 and G_2 and a subgroup $M \leq G_1 \times G_2$, and writing $\pi_i : G_1 \times G_2 \to G_i$ for the two coordinate projections, we say that M has *full one-dimensional projections* if $\pi_i(M) = G_i$ for $i = 1, 2$.

We refer to the subgroups

$$H_1 := \pi_1(M \cap (G_1 \times \{1_{G_2}\})) \quad \text{and} \quad H_2 := \pi_2(M \cap (\{1_{G_1}\} \times G_2))$$

as the *first* and *second slices* of M respectively.

It is a classical observation of Goursat (see, for example, Section 1.6 of Schmidt [40]) that M has full one-dimensional projections and trivial first and second slices if and only if it is the graph of an isomorphism $\Phi : G_2 \to G_2$. If the slices are non-trivial, we do at least have the following.

LEMMA 2.8. *If $M \leq G_1 \times G_2$ has full one-dimensional projections then its slices satisfy $H_i \trianglelefteq G_i$ for $i = 1, 2$.*

Proof. By symmetry it suffices to treat the case $i = 1$. Let $r_1 \in G_1$. Since $\pi_1(M) = G_1$ we can find $r_2 \in G_2$ such that $(r_1, r_2) \in M$. It is now easy to check that

$$
\begin{aligned}
r_1 H_1 &= \{g \in G_1 : (r_1^{-1}g, e) \in M\} = \{g \in G_1 : (r_1, r_2)(r_1^{-1}g, e) \in M\} \\
&= \{g \in G_1 : (g, r_2) \in M\} = \{g \in G_1 : (gr_1^{-1}, e)(r_1, r_2) \in M\} \\
&= \{g \in G_1 : (gr_1^{-1}, e) \in M\} = H_1 r_1.
\end{aligned}
$$

Since r_1 was arbitrary, H_1 is normal, as required. ■

Given a compact group G or one of its homogeneous spaces G/H we shall always consider it endowed with its usual Borel structure and Haar probability measure, which we shall denote by m_G or $m_{G/H}$.

If U is a compact metrizable group then we write $\text{Clos}\, U$ for its collection of closed subsets endowed with the Vietoris topology (see, for example, 2.7.20, 3.12.27 and 4.5.23 of Engelking [17]; as is standard, this is also compact and metrizable) and the associated standard Borel structure and $\text{Lat}\, U \subseteq \text{Clos}\, U$ for the further Veitoris-closed subfamily of closed sub*groups* with its induced standard Borel structure. In this setting of subgroups of compact metrizable groups, the Vietoris topology is easily seen to coincide with the Fell topology and the Chabauty topology, both of which also commonly appear in the study of lattices of closed subgroups; see Subsection 2.1 of Conze and Raugi [15] and the references given there. This topology and Borel structure can be understood in terms of Haar measures in the following standard way.

LEMMA 2.9. *The Vietoris topology and measurable structure on* $\text{Lat}\, U$ *coincide with the pullbacks of the vague topology and measurable structure under the Haar-measure map* $H \mapsto \mu_H$. ∎

3. Direct integrals of homogeneous-space data. In this section we give the rigorous definition of a 'direct integral' of measurably-varying homogeneous spaces and of the lifted transformation acting on it, and establish some of their elementary properties. We build such an extension X as a union of different fibres G_y/K_y above each $y \in Y$, the fibre actually depending only on $\zeta_0^S(y) \in Z_0^S$, and we extend S to an action T on X using a cocycle constrained to lie at (almost) every point in the relevant fibre. We enforce a suitable measurable structure by drawing G_y and K_y from among the compact subgroups of some fixed 'repository' group, subject to the condition $K_y \le G_y$, measurably for the Vietoris measurable structure on such subgroups.

DEFINITION 3.1 (Measurable homogeneous space data). Let Y be a standard Borel space and U a fixed compact metrizable group. By *measurable compact group data on* Y *with fibre repository* U we understand a map $Y \to \text{Lat}\, U : y \mapsto G_y$ that is measurable for the Vietoris Borel structure on $\text{Lat}\, U$. We shall usually denote such a map by $G_\bullet$, and will often omit explicit mention of the fibre repository U. More generally, by *measurable compact homogeneous space data on* Y *with fibre repository* U we understand a pair $(G_\bullet, K_\bullet)$ of measurable compact group data with repository U such that $K_y \le G_y$ for every y. We shall usually denote this pair instead by $G_\bullet/K_\bullet$, and think of it as a measurable assignment of the compact homogeneous space G_y/K_y to each point $y \in Y$.

DEFINITION 3.2 (Direct integral of measurable homogeneous space data). Given a standard Borel probability space (Y, ν) and measurable compact homogeneous space data as above, we shall define their *direct integral* to be the subset

$$\{(y, gK_y) : y \in Y, \ g \in G_y\} \subseteq Y \times \operatorname{Clos} U,$$

which we denote by $Y \ltimes G_\bullet / K_\bullet$. This is easily verified to be standard Borel for the relevant product measurable structure, and we will always assume it to be endowed with the restriction of that measurable structure.

On this space we define the *direct integral* measure $\nu \ltimes m_{G_\bullet / K_\bullet}$ by

$$\nu \ltimes m_{G_\bullet / K_\bullet}(A) := \int\limits_Y \delta_y \otimes m_{G_y / K_y}(A \cap (\{y\} \times G_y / K_y)) \, \nu(\mathrm{d}y).$$

Given another measurable assignment of subgroup data $H_\bullet \leq G_\bullet$, we define analogously the direct integrals $(Y \ltimes H_\bullet \backslash G_\bullet, \nu \ltimes m_{H_\bullet \backslash G_\bullet})$ of the spaces of right-cosets and $(Y \ltimes (H_\bullet \backslash G_\bullet / K_\bullet), \nu \ltimes m_{H_\bullet \backslash G_\bullet / K_\bullet})$ of the spaces of double cosets.

REMARK. We will rarely remark again on the assumption that the fibre repository U be metrizable, but this will always be implicit. This ensures that the above construction keeps us within the category of standard Borel spaces (and it will be a natural consequence of the non-ergodic Furstenberg–Zimmer theory applied to such spaces), and will occasionally be important for proofs (such as in Lemma 3.7 below). One could attempt to construct an extended theory that allows non-metrizable fibre groups and works instead in the larger category of perfect measure spaces (see 342K of Fremlin [20]), but we will not do so here.

DEFINITION 3.3 (Cocycle-sections). Suppose that (Y, ν), U and $G_\bullet / K_\bullet$ are as above, that Γ is a locally compact second countable group and that $S : \Gamma \curvearrowright (Y, \nu)$, and suppose further that the group data $y \mapsto G_y$ and $y \mapsto K_y$ are S-invariant. Then a *cocycle-section* of $G_\bullet$ over S is a measurable cocycle $\rho : \Gamma \times Y \to U$ over S such that $\rho(\gamma, y) \in G_y$ for every $\gamma \in \Gamma$ and $y \in Y$. We shall denote such a cocycle-section by $\rho : \Gamma \times Y \to G_\bullet$.

REMARK. Note that in the setting of a general locally compact second countable group Γ, the definition that ρ be a cocycle over S demands only that $\rho(\gamma_1 \gamma_2, y) = \rho(\gamma_1, S^{\gamma_2} y) \cdot \rho(\gamma_2, y)$ for ν-almost every $y \in Y$ for strictly every γ_1 and γ_2 (see, for instance, Section 4.2 of Zimmer [50]), where the negligible set of 'bad' y is allowed to vary with (γ_1, γ_2); and that by convention two cocycles are equivalent if they agree ν-almost surely for strictly every γ. In view of this, we lose no generality in asking that $\rho(\gamma, y) \in G_y$ for strictly every y and γ, rather than for almost every y for strictly every γ, since in the latter case we may simply adjust ρ to equal 1 on the Borel set

where it falls outside the specified repository, and this changes each $\rho(\gamma, \cdot)$ on only a ν-negligible set for strictly every γ.

Finally, we can define our class of extensions.

DEFINITION 3.4 (Extensions by measurable homogeneous space data). Suppose that $\mathbf{Y} = (Y, \nu, S)$ and $G_\bullet/K_\bullet$ are as above, that the group data $y \mapsto G_y$ and $y \mapsto K_y$ are S-invariant and that $\rho : \Gamma \times Y \to G_\bullet$ is a cocycle-section over S. Then the *extension of* $\mathbf{Y}$ *by the data* $(G_\bullet/K_\bullet, \rho)$ is the action T of Γ on $(Y \ltimes G_\bullet/K_\bullet, \nu \ltimes m_{G_\bullet/K_\bullet})$ given by

$$T^\gamma(y, gK_y) := (S^\gamma y, \rho(\gamma, y)gK_y);$$

it is routine to verify that this is measurable and measure-preserving.

We will often denote this extended system by $\mathbf{Y} \ltimes (G_\bullet/K_\bullet, m_{G_\bullet/K_\bullet}, \rho)$. It clearly admits $\mathbf{Y}$ as a factor simply by projecting out the fibre coordinate; we will refer to this as the *canonical* factor map. The data $G_\bullet/K_\bullet$ and cocycle-section ρ are together *relatively ergodic* if the extension $\mathbf{Y} \ltimes (G_\bullet/K_\bullet, m_{G_\bullet/K_\bullet}, \rho) \to \mathbf{Y}$ through the canonical map is relatively ergodic.

REMARKS. 1. In light of the Peter–Weyl Theorem (treated in most standard texts on compact group representations, such as in Section III.3 of Bröcker and tom Dieck [11]) all compact metrizable groups can be realized isomorphically, albeit highly non-uniquely, as closed subgroups of a suitably large direct product of unitary groups, say $U := \prod_{n \geq 1} \mathrm{U}(n)^\mathbb{N}$. This suggests that such a direct product should suffice as a compact repository for all purposes, and indeed this can be proved with just a little work; however, this result seems to contribute little to the theory, and so we will not present it here. Note, however, that it is also precisely such direct products of unitary groups that will emerge naturally as repositories in the non-ergodic Furstenberg–Zimmer inverse theory of Section 5 below.

2. In view of the condition that $G_\bullet$ and $K_\bullet$ are S-invariant, given a coordinatization $\zeta_0^S : Y \to Z_0^S$ of the S-isotropy factor we could alternatively work with compact measurable group data defined initially as functions on the space $Z_0^{\tilde{S}}$ and then lifted through ζ_0^S. We will occasionally use this alternative description when it is notationally convenient.

The following related definition will also occasionally be useful.

DEFINITION 3.5 (Opposite extensions by measurable group data). If $\mathbf{Y}$, $G_\bullet$ and ρ are as above, then they also define an extended action T on $(Y \ltimes G_\bullet, \nu \ltimes m_{G_\bullet})$ by

$$T^\gamma(y, g) := (S^\gamma y, g\rho(\gamma, y)^{-1});$$

this is the *opposite extension of* $\mathbf{Y}$ *by the data* $(G_\bullet, \rho)$, and we will denote this T by $S \ltimes \rho^{\mathrm{op}}$.

REMARK. In fact we always have

$$\mathbf{Y} \ltimes (G_\bullet, \rho) \xleftarrow{\quad \cong \quad} \mathbf{Y} \ltimes (G_\bullet, \rho^{\mathrm{op}})$$

$$\text{canonical} \searrow \qquad \swarrow \text{canonical}$$

$$\mathbf{Y}$$

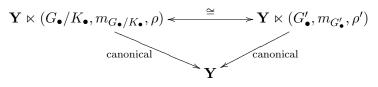through the fibrewise isomorphism $(y, g) \mapsto (y, g^{-1})$. The use of opposite extensions will matter to us in situations where we have two different actions on the extended space, one by a cocycle and one by an opposite cocycle.

Before leaving this section, it is worth noting one way in which some redundancy in the above definition can be removed.

DEFINITION 3.6. Homogeneous space data $G_\bullet/K_\bullet$ over (Y, ν) is *core-free* if K_y is core-free in G_y almost everywhere.

LEMMA 3.7. *Suppose* $\mathbf{Y} := (Y, \nu, S)$ *is a* Γ*-system, that* $G_\bullet/K_\bullet$ *are measurable* S*-invariant homogeneous space data on* Y *with repository* U *and that* $\rho : \Gamma \times Y \to G_\bullet$ *is a cocycle-section over* S*. If in addition the group* K_y *is normal in* G_y *for* ν*-almost every* y*, then there are a fibre repository* U'*, measurable* S*-invariant group data* G' *on* Y *and a measurable* S*-invariant family of isomorphisms* $\Psi_y : G_y/K_y \to G'_y$ *such that the map* $(y, gK_y) \mapsto (y, \Psi_y(gK_y))$ *defines an isomorphism of extensions*

$$\mathbf{Y} \ltimes (G_\bullet/K_\bullet, m_{G_\bullet/K_\bullet}, \rho) \xleftarrow{\quad \cong \quad} \mathbf{Y} \ltimes (G'_\bullet, m_{G'_\bullet}, \rho')$$

$$\text{canonical} \searrow \qquad \swarrow \text{canonical}$$

$$\mathbf{Y}$$

with $(\gamma, y) \mapsto \rho'(\gamma, y) := \Psi_y(\rho(\gamma, y)) : \Gamma \times Y \to G'$.

Proof. This rests on the construction of the new fibre repository for the quotient groups $G_\bullet/K_\bullet$. For $y \in Y$ let $\mathfrak{H}_y \leq L^2(m_U)$ be the separable Hilbert subspace of square-integrable functions on U invariant under left rotation by K_y. This is an S-invariant measurable family of separable Hilbert spaces (in the sense familiar from the analysis of group representations and von Neumann algebras; see, for instance, Mackey [34]), and so we can partition Y into S-invariant measurable subsets $A_1, A_2, \ldots, A_\infty$ and for each $n \in \mathbb{N} \cup \{\infty\}$ select an S-invariant measurable family of isomorphisms $\Phi_y : \mathfrak{H}_y \to \mathfrak{H}'_n$ for $y \in A_n$, where $\mathfrak{H}'_n$ is some fixed n-dimensional reference complex Hilbert space and $\mathfrak{H}'_1 \leq \mathfrak{H}'_2 \leq \cdots \leq \mathfrak{H}'_\infty$.

Now let $\pi_y : G_y \curvearrowright \mathfrak{H}_\infty$ for $y \in A_n$ be the representation that results from first restricting the left-regular representation of G_y on $L^2(m_U)$ to $\mathfrak{H}_y$ (which is possible when $K_\bullet \trianglelefteq G_\bullet$, hence almost everywhere), then composing

with Φ_y to obtain a representation on $\mathfrak{H}'_n$ and finally extending this to act on $\mathfrak{H}'_\infty$ by acting trivially on $\mathfrak{H}'_\infty \ominus \mathfrak{H}'_n$.

This defines an S-invariant measurable family of representations π_y of G_y for $y \in Y$ outside some ν-conegligible subset, and such that $K_y = \ker \pi_y$ almost surely. Next, it is easy to see that the decomposition of π_y into finite-dimensional representations given by the Peter–Weyl Theorem is measurable in y (for example, since they may recovered as the spectral projections of each of a countable dense subfamily of all the measurable selections over y of π_y-invariant compact operators on $\mathfrak{H}_\infty$). Hence this decomposition gives a measurable family of continuous homomorphic embeddings $G_\bullet \to \prod_{n \geq 1} \mathrm{U}(n)^{\mathbb{N}}$ with kernels $K_\bullet$, and so letting $G'_\bullet$ be the image group data of these embeddings they define a measurable family of isomorphisms $\Psi_\bullet$ such that defining ρ' as above and applying $\Psi_\bullet$ fibrewise on $Y \ltimes G_\bullet/K_\bullet$ gives the desired isomorphism of extensions. ∎
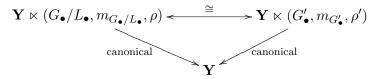
COROLLARY 3.8. *If* $\mathbf{X} = \mathbf{Y} \ltimes (G_\bullet/K_\bullet, m_{G_\bullet/K_\bullet}, \rho)$ *is an extension by homogeneous space data, then it is isomorphic (as an extension of* $\mathbf{Y}$ *through the canonical map) to an extension by core-free homogeneous space data.*

Proof. Let U be the repository and let $L_y := \bigcap_{g \in G_y} g^{-1} K_y g$ be the pointwise core of K_y in G_y. First observe that for any $u \in U$ the set

$$\left\{ (G, K) : u \in \bigcap_{g \in G} g^{-1} K g \right\} = \{ (G, K) : K \ni g u g^{-1} \; \forall g \in G \}$$

is open in $(\mathrm{Lat}\, U)^2$, since if (K, G) does not lie in this set then there are a closed set $V_1 \subseteq U$ with non-empty interior and an open set $V_2 \subseteq U$ such that $K \cap V_1 = \emptyset$, $G \cap V_2 \neq \emptyset$ and $V_2 u V_2^{-1} \subseteq V_1$. It follows that the map $(\mathrm{Lat}\, U)^2 \to \mathrm{Lat}\, U : (G, K) \mapsto \bigcap_{g \in G} g^{-1} K g$ is measurable, and hence that $L_\bullet$ is measurable group data.

Now by the preceding lemma we can select a measurable family of embeddings of the groups G_y/L_y into a suitably modified repository to obtain an isomorphism of systems

$$\mathbf{Y} \ltimes (G_\bullet/L_\bullet, m_{G_\bullet/L_\bullet}, \rho) \xleftarrow{\quad \cong \quad} \mathbf{Y} \ltimes (G'_\bullet, m_{G'_\bullet}, \rho')$$

<div style="text-align:center">canonical   canonical</div>

$$\mathbf{Y}$$

corresponding to a continuous group isomorphism $G_y/L_y \to G'_y$ at almost every y. Under these isomorphisms the subgroups $K_y \leq G_y$ correspond measurably to some $K'_y \leq G'_y$ so that $(G'_y, K'_y) \cong (G_y/L_y, K_y L_y/L_y)$, and so observing from its definition that $L_y K_y/L_y$ is always core-free in G_y/L_y, this completes the proof. ∎

4. Mackey theory in the non-ergodic setting. We will now move on to a more detailed analysis of extensions by homogeneous space data, and more specifically of their invariant factors and relatively ergodic measures. Many of the ideas that follow are nearly direct translates to our setting of those of Mackey in the case of an ergodic base system, and we will follow quite closely their treatment in Section 3.5 of Glasner [28].

In fact, more is true: earlier work on multiple recurrence and non-conventional ergodic averages has already encountered the possibility of a measurably-varying Mackey group within an extension of a non-ergodic base system by a fixed overall group. This technicality arises in the work of Meiri [35] on correlation sequences arising from probability-preserving systems, of Furstenberg and Weiss [27] on certain polynomial non-conventional ergodic averages and more recently in Ziegler's approach in [47] to convergence of linear non-conventional averages for powers of a single transformation. For example, during the analysis in [27] a homogeneous skew-product extension of ergodic systems $(X, \mu, T) = (Y, \nu, S) \ltimes (G, m_G, \rho)$ is three-fold joined to itself, to give a measure on X^3 that is invariant for a transformation of the form $T^r \times T^s \times T^t$ but which is not ergodic for that transformation. This system is now coordinatized as an extension of an action on Y^3 by the group and cocycle $(G^3, (\rho^{(r)}, \rho^{(s)}, \rho^{(t)}))$, but since the base is no longer ergodic the description of the ergodic components of the overall system requires the possibility that the Mackey group can vary among the closed subgroups of G^3 (a possibility that is then discounted by an argument showing that they are all actually conjugate, and so may in fact be taken to be constant; we shall see a similar trick in Subsection 7.4 below).

The only extra subtlety for which we must allow here is that the overall group $G_\bullet$ now also varies measurably. This will require us to work rather harder in setting up the proof, although the overall idea is very similar to those mentioned above. For this reason, although we have included complete proofs here, we refer the reader to these other sources, and also Section 3.5 of Glasner [28], for relevant background.
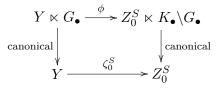
4.1. Ergodic decompositions and Mackey group data. Mackey theory describes the invariant factor of an extension $\mathbf{X} = \mathbf{Y} \ltimes (G_\bullet, m_{G_\bullet}, \rho)$ in terms of the invariant factor of $\mathbf{Y}$ and the data $(G_\bullet, m_{G_\bullet}, \rho)$ of the extension. Here it will prove convenient to treat $G_\bullet$ as varying over the factor space Z_0^S of ergodic components, lifted to Y for the purpose of defining the extended system.

THEOREM 4.1 (Mackey Theorem in the non-ergodic case). *Suppose that* $\mathbf{X}$ *is the group-data extension* $\mathbf{Y} \ltimes (G_\bullet, m_{G_\bullet}, \rho)$ *and that* $\zeta_0^S : \mathbf{Y} \to \mathbf{Z}_0^S$ *is a coordinatization of the base isotropy factor, and let* $\pi : X = Y \ltimes G_\bullet \to Y$ *and* $\theta : Z_0^S \ltimes G_\bullet \to Z_0^S$ *be the canonical factor maps. Then there are measurable*

subgroup data $K_\bullet \leq G_\bullet$ on Z_0^S, a T-invariant map

$$\phi : (X, \mu) \to (Z_0^S, \zeta_{0\,\#}^S \nu) \ltimes (K_\bullet \backslash G_\bullet, m_{K_\bullet \backslash G_\bullet})$$

and a section $b : Y \to G_\bullet$ such that

(1) *ϕ coordinatizes $\mathbf{Z}_0^T$ and the following diagram commutes:*

$$
\begin{array}{ccc}
Y \ltimes G_\bullet & \xrightarrow{\ \ \phi\ \ } & Z_0^S \ltimes K_\bullet \backslash G_\bullet \\
\text{canonical}\Big\downarrow & & \Big\downarrow \text{canonical} \\
Y & \xrightarrow{\ \ \zeta_0^S\ \ } & Z_0^S
\end{array}
$$

(2) *$\phi(y, g) = (\zeta_0^S(y), K_{\zeta_0^S(y)} b(y)g)$ for μ-almost every (y, g);*

(3) *the cocycle-section $(\gamma, y) \mapsto b(S^\gamma y)\rho(\gamma, y)b(y)^{-1}$ takes a value in $K_{\zeta_0^S(y)}$ for ν-almost every y for every γ;*

(4) *(Conjugate minimality) if $K_\bullet' \leq G_\bullet$ is another measurable assignment of compact subgroup data on Z_0^S and $b' : Y \to G_\bullet$ another section such that the cocycle-section $(\gamma, y) \mapsto b'(S^\gamma y)\rho(\gamma, y)b'(y)^{-1}$ takes a value in $K_{\zeta_0^S(y)}'$ for ν-almost every y for every γ, then there is a section $c : Z_0^S \to G_\bullet$ such that*

$$c(s) \cdot K_s' \cdot c(s)^{-1} \geq K_s$$

for $(\zeta_0^S)_\# \nu$-almost every s;

(5) *if $P : Z_0^S \xrightarrow{\text{p}} Y$ is a version of the disintegration of ν over ζ_0^S, then the probability kernel $(s, K_s g') \xmapsto{\text{p}} P(s, \cdot) \ltimes m_{b(\bullet)^{-1} K_s g'}$ is a version of the disintegration of μ over ϕ.*

REMARK. Clearly with hindsight we can take the property (2) above as defining ϕ; the point, however, is that we will obtain $K_\bullet$ and ϕ first and then show that ϕ takes this form for some b.

The proof of this theorem will require some initial constructions and an enabling lemma concerning the measurable selection of generic points.

First let $\zeta_0^T : \mathbf{X} \to \mathbf{Z}_0^T$ be some coordinatization of the isotropy factor of the large system with the property that $\zeta_0^S \circ \pi$ factorizes through the natural factor $\xi : \mathbf{Z}_0^T \to \mathbf{Z}_0^S$ in the sense of the commutative diagram

$$
\begin{array}{ccc}
\mathbf{X} & \xrightarrow{\ \ \zeta_0^T\ \ } & \mathbf{Z}_0^T \\
\pi\Big\downarrow & & \Big\downarrow \xi \\
\mathbf{Y} & \xrightarrow{\ \ \zeta_0^S\ \ } & \mathbf{Z}_0^S
\end{array}
$$

(it is a standard fact that this is possible; see, for example, Section 2.2 of Glasner [28]).

We need a formal way to work with the action by right multiplication of the fibre G_y on itself (the difficulty being, of course, that global constructions based on this pointwise-varying action need to be kept measurable). To this end we define the map

$$\tau_0 : X \times_{\{\zeta_0^S \circ \pi = \theta\}} (Z_0^S \ltimes G_\bullet) \to X : ((y, g'), (s, g)) \mapsto (y, g'g);$$

intuitively, $\tau_0(x, (s, g))$ gives the image of x under right multiplication by g, which is an element of the group G_s over $s = \zeta_0^S(\pi(x))$ that acts on the fibre above $\pi(x)$. This map τ_0 can be well-defined only for those tuples with $s = \zeta_0^S(\pi(x))$, hence the need for the relative self-product in the specification of its domain; however, with this restriction in place it is easily seen to be measurable.

We can now work with the domain of τ_0 as a system in its own right under the action $T \times \mathrm{id}_{Z_0^S \ltimes G_\bullet}$. It is clear that the measure $\mu \otimes_{\{\zeta_0^S \circ \pi = \theta\}}$ $((\zeta_0^S)_\# \nu \ltimes m_{G_\bullet})$ is $(T \times \mathrm{id}_{Z_0^S \ltimes G_\bullet})$-invariant and that τ_0 itself is a factor map from the resulting system onto (X, μ, T).

From τ_0 we now define the composition

$$\tau : X \times_{\{\zeta_0^S \circ \pi = \theta\}} (Z_0^S \ltimes G_\bullet) \xrightarrow{\tau_0} X \xrightarrow{\zeta_0^T} Z_0^T.$$

Heuristically this assigns to the pair $(x, (s, g))$ the ergodic component of (X, μ, T) that contains the image of x under the right multiplication by g acting on its fibre. Of course, the whole point is that different points within a single fibre of π will generally lie in different ergodic components, and this map τ reports on this dependence.

Informally, the proof of Theorem 4.1 now proceeds by selecting a representative $p(s) \in X$ above each $s \in Z_0^S$ and then defining K_s to be the subgroup of those $g \in G_s$ such that the T-ergodic component of $p(s)$ does not change upon right multiplication by g inside the π-fibre of s: that is, such that $\tau(p(s), (s, g)) = \zeta_0^T(p(s))$. We need p to select points that are sufficiently 'generic' in the fibres above s, in the sense made precise by the following lemma.

LEMMA 4.2. *In the setting of Theorem 4.1, we can find a Borel measurable section* $p : Z_0^S \to X$ *of the factor map* $\zeta_0^S \circ \pi$ *such that the probability kernel*

$$P(s, A) := m_{G_s}\{g \in G_s : \tau(p(s), (s, g)) \in A\}, \quad A \in \Sigma_{Z_0^T},$$

is a version of the disintegration of $(\zeta_0^T)_\# \mu$ *over* $\xi : Z_0^T \to Z_0^S$.

Proof. We first define a probability kernel $P' : X \xrightarrow{\mathrm{p}} Z_0^T$ by

$$P'((y, g'), A) := m_{G_{\zeta_0^S(y)}}\{g \in G_{\zeta_0^S(y)} : \tau((y, g'), (\zeta_0^S(y), g)) \in A\}.$$

Intuitively this takes a point $x = (y, g')$, replaces it with an average over the fibre $\pi^{-1}\{y\}$, regarded as a copy of the fibre group $G_{\zeta_0^S(y)}$, and then pushes the resulting probability measure down to Z_0^T.

It is easy to see that this satisfies the measurability conditions of a probability kernel, since we have

$$P'((y, g'), A) = P''((y, g'), (\zeta_0^T)^{-1}(A))$$

where $P'' : X \xrightarrow{\mathrm{p}} X$ is the manifestly measurable probability kernel

$$P''((y, g'), \cdot) := (\delta_y \otimes m_{G_{\zeta_0^S(y)}})(\cdot).$$

In addition the above definition implies that

$$P'((y, g'), \cdot) = P'((y, g'g), \cdot) = P'((S^\gamma y, \rho(\gamma, y)g'), \cdot)$$

for any $g \in G_{\zeta_0^S(y)}$ and $\gamma \in \Gamma$, firstly in view of the averaging over $G_{\zeta_0^S(y)}$ and secondly because we take the image under the invariant function ζ_0^T. Therefore the function $x \mapsto P'(x, \cdot)$, regarded as a Borel map from X to the space of Borel probability measures on Z_0^T with its usual Borel structure, both factorizes through π and is then S-invariant. Therefore there is some $P : Z_0^S \xrightarrow{\mathrm{p}} Z_0^T$ such that $P' = P \circ \zeta_0^S \circ \pi$, μ-almost surely. This P must be a version of the disintegration of $(\zeta_0^T)_\# \mu$ over ξ.

Setting

$$A_s := \{x \in X : \zeta_0^S(\pi(x)) = s \text{ and } P'(x, \cdot) = P(s, \cdot)\},$$

this is now the set of 'generic points' above s from which we need to select $p(s)$. From the above relation between P' and P we deduce that $P(s, A_s) = 1$ for $(\zeta_0^S)_\# \nu$-almost every s. In addition, since A_s is the section above s of the Borel set

$$A := \{(s, x) \in Z_0^S \times X : \ \zeta_0^S(\pi(x)) = s \text{ and } P'(x, \cdot) = P(s, \cdot)\},$$

by the Measurable Selector Theorem 2.2 we can choose a measurable selector p of $\zeta_0^S \circ \pi$ such that $(s, p(s)) \in A$ for $(\zeta_0^S)_\# \nu$-almost every s. This selector now has the properties claimed. ∎

Proof of Theorem 4.1. (1) Given the measurable selector of the above lemma, define

$$K_s := \{g \in G_s : \tau(p(s), (s, g)) = \zeta_0^T(p(s))\}.$$

It is clear from the definition of τ that $\tau(p(s), (s, g)) = \zeta_0^T(p(s))$ if and only if $\tau(p(s), (s, gg')) = \tau(p(s), (s, g'))$ for every $g' \in G_s$, and hence K_s is a closed subgroup of G_s for almost every s.

Also, we have

$$A := \{(s, g) \in Z_0^S \times G_\bullet : g \in K_s\}$$
$$= \{(s, g) \in Z_0^S \times G_\bullet : \tau(p(s), (s, g)) = \zeta_0^T(p(s))\},$$

and so this is a Borel subset of $Z_0^S \times G_\bullet$. Letting U be the repository, it follows that for any open $V_1, \ldots, V_k \subseteq U$ and closed $W \subseteq U$ we have

$$\{s \in Z_0^S : K_s \cap V_i \neq \emptyset \ \forall i \leq k \ \& \ K_s \cap W = \emptyset\}$$
$$= \theta\Big(A \cap \bigcap_{i \leq k}(Z_0^S \times V_i)\Big)\backslash\theta(A \cap (Z_0^S \times W))$$

(recalling that θ is the projection onto the first coordinate), and so from the Measurable Selector Theorem 2.2 this is a universally measurable set. Allowing $V_1, \ldots, V_k$ and $U \setminus W$ to run over all finite strings drawn from some countable collection of open subsets of U that generates the whole topology, we deduce that the map $s \mapsto K_s$ is universally measurable, and so after modifying it on a ν-negligible set if necessary we may assume it is Borel.

Finally, it also follows from the fact that $\tau(p(s), (s, gg')) = \tau(p(s), (s, g'))$ for every $g' \in G_s$ and $g \in K_s$ that the map $(s, g') \mapsto \tau(p(s), (s, g'))$ virtually factorizes through the canonical factor

$$Z_0^S \ltimes G_\bullet \to Z_0^S \ltimes K_\bullet\backslash G_\bullet$$

to leave a map

$$\alpha : Z_0^S \ltimes K_\bullet\backslash G_\bullet \to Z_0^T,$$

and that this is injective away from some negligible set, since if $\alpha(s_1, K_{s_1}g_1) = \alpha(s_2, K_{s_2}g_2)$ then $s_1 = \xi(\alpha(s_1, g_1)) = \xi(\alpha(s_2, g_2)) = s_2$ and now

$$\tau(p(s_1), (s_1, g_1)) = \tau(p(s_1), (s_1, (g_1 g_2^{-1})g_2)) = \tau(p(s_1), (s_1, g_2)),$$

which implies that $g_1 g_2^{-1} \in K_{s_1}$ provided s_1 did not lie in the negligible subset of Z_0^S on which we modified $K_\bullet$ above. By another appeal to Theorem 2.2 the map α has a Borel virtual inverse, say

$$\beta : Z_0^T \to Z_0^S \ltimes K_\bullet\backslash G_\bullet.$$

Now $\phi := \beta \circ \zeta_0^T$ is a coordinatization of Z_0^T for which the desired diagram is commutative.

(2) From the construction of β and another measurable selection there is a measurable map $u : Y \to K_\bullet\backslash G_\bullet$ such that

$$\phi(y, g) = \phi\big(\tau_0\big((y, 1_{G_{\zeta_0^S(y)}}), (\zeta_0^S(y), g)\big)\big) = (\zeta_0^S(y), u(y)g);$$

composing u with a measurable selector $Y \ltimes K_\bullet\backslash G_\bullet \to Y \ltimes G_\bullet$ gives a measurable map $b : Y \to G_\bullet$ such that $\phi(y, g) = \big(\zeta_0^S(y), K_{\zeta_0^S(y)}b(y)g\big)$ almost everywhere.

(3) Since ϕ is T-invariant we have

$$\big(\zeta_0^S(y), K_{\zeta_0^S(y)}b(S^\gamma y)\rho(\gamma, y)1_{G_{\zeta_0^S(y)}}\big)$$

$$= \big(\zeta_0^S(S^\gamma y), K_{\zeta_0^S(S^\gamma y)}b(S^\gamma y)\rho(\gamma, y)1_{G_{\zeta_0^S(S^\gamma y)}}\big) = \phi\big(S^\gamma y, \rho(\gamma, y)1_{G_{\zeta_0^S(y)}}\big)$$

$$= \phi(T(y, g)) = \phi(y, g) = \big(\zeta_0^S(y), K_{\zeta_0^S(y)}b(y)1_{G_{\zeta_0^S(y)}}\big)$$

ν-almost surely for any $\gamma \in \Gamma$, from which the required cohomology condition follows at once.

(4) Suppose that $K'_\bullet$ and b' have the asserted properties, and let U be the overall repository for our compact group data. Then the map

$$Y \ltimes G_\bullet \to \operatorname{Clos} U : (y, g) \mapsto K'_{\zeta_0^S(y)}b'(y)g$$

is a measurable map into a standard Borel space, and it is μ-almost surely T-invariant from the property that $b'(S^\gamma y)\rho(\gamma, y)b'(y)$ lies almost surely in $K'_{\zeta_0^S(y)}$. Therefore, because ϕ coordinatizes the T-invariant factor, we know that there is a Borel map $f : Z_0^S \ltimes (K_\bullet \backslash G_\bullet) \to \operatorname{Clos} U$ such that

$$K'_{\zeta_0^S(y)}b'(y)g = f\big(\zeta_0^S(y), K_{\zeta_0^S(y)}b(y)g\big)$$

for μ-almost every (y, g). In particular, it follows that for ν-almost every y, it is the case that for $m_{G_{\zeta_0^S(y)}}$-almost every $g \in G_{\zeta_0^S(y)}$ and $m_{b(y)^{-1}K_{\zeta_0^S(y)}b(y)}$-almost every $h \in b(y)^{-1}K_{\zeta_0^S(y)}b(y)$ we have

$$f\big(\zeta_0^S(y), K_{\zeta_0^S(y)}b(y)g\big) = f\big(\zeta_0^S(y), K_{\zeta_0^S(y)}b(y)hg\big),$$

and hence

$$K'_{\zeta_0^S(y)}b'(y)g = K'_{\zeta_0^S(y)}b'(y)hg \;\Rightarrow\; K'_{\zeta_0^S(y)}b'(y) = K'_{\zeta_0^S(y)}b'(y)h.$$

Since (g, h) were chosen arbitrarily from a Haar-conegligible subset of $G_{\zeta_0^S(y)} \times b(y)^{-1}K_{\zeta_0^S(y)}b(y)$, it follows that for μ-almost every y we have

$$b(y)b'(y)^{-1} \cdot K'_{\zeta_0^S(y)} \cdot (b(y)b'(y)^{-1})^{-1} \geq K_{\zeta_0^S(y)}.$$

This tells us that the set

$$\{(s, c) \in Z_0^S \ltimes G_\bullet : c \cdot K'_s \cdot c^{-1} \geq K_s\}$$

(which is Borel for Borel versions of the measurable assignments $K_\bullet$ and $K'_\bullet$) has non-empty fibre above $(\zeta_0^S)_\#\nu$-almost every $s \in Z_0^S$, and so letting c be a measurable selector for this set completes the argument.

(5) Finally, observe that for almost every $(s, K_s g')$ that parameterizes a T-ergodic component we have $\phi(y, g) = (s, K_s g')$ if and only if $\zeta_0^S(y) = s$ and $K_s b(y)g = K_s g'$, hence if and only if $g \in b(y)^{-1}K_s g'$. From this the last conclusion follows at once. ∎

DEFINITION 4.3 (Mackey data). We refer to the measurable group data K_s given by the above theorem as *Mackey group data* of ρ over (Y, ν, S), and to the section b as a *Mackey section* (note that in general the Mackey group data is not unique, but is so up to S-invariant conjugacy, by part (4) of the theorem).

From the above result for extensions by group data we can easily generalize to extensions by homogeneous space data.

COROLLARY 4.4. *Suppose that* $(X, \mu, T) = (Y, \nu, S) \ltimes (G_\bullet/H_\bullet, m_{G_\bullet/H_\bullet}, \rho)$, $\zeta_0^S : Y \to Z_0^S$ *a coordinatization of the base isotropy factor and* $P : Z_0^S \xrightarrow{\mathrm{p}} Y$ *a version of the disintegration of* ν *over* ζ_0^S. *Then there are subgroup data* $K_\bullet \leq G_\bullet$ *and a cocycle-section* $b : Y \to G_\bullet$ *such that the factor map*

$$\phi : X \to Z_0^S \ltimes (K_\bullet \backslash G_\bullet / H_\bullet) : (y, gH_{\zeta_0^S(y)}) \mapsto (\zeta_0^S(y), K_{\zeta_0^S(y)} b(y) g H_{\zeta_0^S(y)})$$

is a coordinatization of the isotropy factor $\zeta_0^T : X \to Z_0^T$, *and the probability kernel*

$$(s, K_s g' H_s) \xrightarrow{\mathrm{p}} P(s, \cdot) \ltimes m_{b(\bullet)^{-1} K_s g' H_s / H_s}$$

is a version of the disintegration of μ *over* ϕ, *where for any subset* $S \subseteq G_s$ *we write* $S/H_s := \{gH_s : g \in S\}$.

Proof. Let $(X', \mu', T') := (Y, \nu, S) \ltimes (G_\bullet, \rho)$ and $\pi : (X', \mu', T') \to (X, \mu, T)$ the covering factor map, and now let $K_\bullet$, b and ϕ' be given by applying Theorem 4.1 to the canonical factor map $\pi' : (X', \mu', T') \to (Y, \nu, S)$. Let ϕ be the map given by the above formula.

Since ϕ is manifestly T-invariant we need only show that it recoordinatizes the whole of Z_0^T, and not a properly smaller factor. If $f \in L^\infty(\mu)$ is T-invariant then $f \circ \pi$ is T'-invariant, and so factorizes through the map ϕ' given by Theorem 4.1. It follows that f virtually factorizes through ϕ. ∎

4.2. More general lifted measures on homogeneous space extensions. Theorem 4.1 describes the components of the ergodic decomposition of $\mu \ltimes m_{G_\bullet}$ under $S \ltimes \rho$, but in fact the same ideas can be used to describe *all* relatively $(S \ltimes \rho)$-ergodic lifts of ν. This stronger result, and its corollary for extensions by homogeneous space data, will be important for both the inverse theory to be developed in the next section and the study of automorphisms of extensions in the section after that.

We first set up some simple enabling results concerning the collection of lifts of a given probability measure on an extension by homogeneous space data to measures on a covering extension by group data.

DEFINITION 4.5 (Lift topology). Suppose that $S : \Gamma \curvearrowright (Y, \nu)$, $G_\bullet/H_\bullet$ is an S-invariant measurable homogeneous space data with repository U, $\rho : \Gamma \times Y \to G_\bullet$ is a cocycle-section and μ is an $(S \ltimes \rho)$-invariant probability

measure on $Y \ltimes G_\bullet/H_\bullet$ that lifts μ, and let Q be the set of all further lifts of μ to $(S \ltimes \rho)$-invariant probability measures on $Y \ltimes G_\bullet$. We define the *lift topology* on Q as the weakest topology with respect to which the evaluation functionals

$$\mu' \mapsto \int\limits_{Y \ltimes G_\bullet} 1_A(y) \cdot f(g) \, \mu'(\mathrm{d}(y,g))$$

are continuous for all $A \in \Sigma_Y$ and continuous functions $f : U \to \mathbb{R}$.

The following is now routine.

LEMMA 4.6. *Under the lift topology Q is a non-empty compact convex set.*

Proof. It is easy to witness one member of Q (and so see that it is non-empty):

$$\mu' := \int\limits_{Y \ltimes G_\bullet/H_\bullet} m_{gH_y} \, \mu(\mathrm{d}(y,gH_y)).$$

Convexity is obvious, so we need only verify compactness. However, Q is a closed subset of the larger convex set Q_0 containing all lifts of μ to $Y \times U$, and this set is easily seen to be a closed subset of the unit ball of $L^\infty(\mu; \mathcal{M}(U))$ (where $\mathcal{M}(U)$ is the Banach space of signed measures on U) in the weak* topology. Now the Banach–Alaoglu Theorem tells us that this larger set Q_0 is compact and the proof is complete. ∎

REMARK. Of course, it is easy to construct examples of general extensions of Borel actions $(X, T) \to (Y, S)$ for which a given invariant probability measure on Y has no invariant extension to X (indeed, with $Y = \{*\}$ any Borel action on X with no invariant probability measure will do). The nonemptiness assertion of the above lemma, though simple, is very much a consequence of the isometric structure of the extensions in question.

REMARK. An alternative route to topologizing Q (to be found, for example, in Furstenberg's paper [22] and Glasner's book [28]) is to choose a coordinatization of (Y, ν, S) as a homeomorphic action on a compact space with an invariant Borel probability measure, and then simply introduce the usual vague topology on our convex set of lifts. It is not hard to see that the resulting topology is the same; we have chosen the present approach only because it seems more intrinsic.

We can now approach the main results of this subsection.

PROPOSITION 4.7. *Suppose that $S : \Gamma \curvearrowright (Y, \nu)$, that $G_\bullet$ are Z_0^S-measurable group data and $\rho : \Gamma \times Y \to G_\bullet$ is a cocycle-section over S and that X is the space $Y \ltimes G_\bullet$ but equipped with some unknown $(S \ltimes \rho)$-invariant and relatively ergodic lift μ of ν. Then there are subgroup data $K_\bullet \leq G_\bullet$ and a section $b : Y \to G_\bullet$ such that $\mu = \nu \ltimes m_{b(\bullet)^{-1}K_\bullet}$.*

REMARK. Once again, the case with S ergodic is classical: it can be found as Theorem 3.26 in Glasner [28].

Proof. Let $T := S \ltimes \rho$ and consider again the extension of the Borel system $T : \Gamma \curvearrowright X$ given by

$$\tau_0 : X \times_{\{\zeta_0^S \circ \pi = \theta\}} (Z_0^S \ltimes G_\bullet) \to X.$$

In addition, let $P : Z_0^S \xrightarrow{\mathrm{p}} X$ be a version of the disintegration of the unknown lift μ over $\zeta_0^S \circ \pi$. Now let

$$K_s' := \{g \in G_s : \tau_0(P(s, \cdot), (s, g)) = P(s, \cdot)\};$$

this is a closed subgroup of G_s that is universally measurable in s by just the same argument as in part (1) of the proof of Theorem 4.1, and is such that

$$\tau_0(P(s, \cdot), (s, gg')) = \tau_0(P(s, \cdot), (s, g'))$$

whenever $g' \in G_s$ and $g \in K_s'$. Adjusting $K_\bullet'$ on a negligible subset of Y so that it is Borel, we still obtain that the composed kernel $(s, g) \xmapsto{\mathrm{p}} \tau_0(P(s, \cdot), (s, g))$ virtually factorizes through the canonical factor $Z_0^S \ltimes G_\bullet \to Z_0^S \ltimes K_\bullet' \backslash G_\bullet$ to a kernel $P' : Z_0^S \ltimes K_\bullet' \backslash G_\bullet \xrightarrow{\mathrm{p}} X$ such that $P'((s, K_s'), \cdot) = P(s, \cdot)$.

Now define another kernel $P'' : Z_0^S \xrightarrow{\mathrm{p}} X$ by

$$P''(s, A) := \int\limits_{K_s \backslash G_s} P'((s, K_s'g), A) \, \mathrm{d}(K_s'g), \quad A \in \Sigma_X.$$

This is an 'averaged out' version of P'. It is also clearly S-invariant, satisfies $\pi_\# P_\#''(\zeta_0^S)_\# \nu = \nu$, and now also satisfies

$$\tau_0(P''(s, \cdot), (s, g)) = P''(s, \cdot) \quad \text{for all } g \in G_s;$$

hence $P_\#''(\zeta_0^S)_\# \nu$ must simply be equal to $\nu \ltimes m_{G_\bullet}$.

Now, the measure $P'((s, K_s'g), \cdot)$ is T-ergodic for almost every (s, g) (since it is a fibrewise right translate of $P(s, \cdot)$ by some fixed element of G_s), and so it follows that the integral

$$\nu \ltimes m_{G_\bullet} = P_\#'' \zeta_0^S{}_\# \nu = \int\limits_{Z_0^S} \int\limits_{K_s' \backslash G_s} P'((s, K_s'g), \cdot) \, \mathrm{d}(K_s'g) \, \nu(\mathrm{d}s)$$

is a version of the ergodic decomposition of $\nu \ltimes m_{G_\bullet}$. By Theorem 4.1, K_s' must be a version of the Mackey group data for ρ over (Y, ν, S), and hence by part (3) of that theorem it follows that there is a section $b' : Y \to G_\bullet$ such that $\rho'(\gamma, y) := b'(S^\gamma y) \cdot \rho(\gamma, y) \cdot b'(y)^{-1}$ takes values in $K_{\zeta_0^S(y)}'$ almost surely.

Now we finish the proof simply by applying the fibrewise recoordinatizing isomorphism $\psi : (y, g) \mapsto (y, b'(y)g)$ from our original system (X, μ, T) to
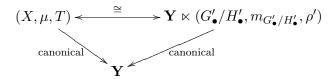
the system $(X, \psi_\# \mu, S \ltimes \rho')$. This map ψ must carry the $(S \ltimes \rho)$-ergodic decomposition of $\nu \ltimes m_{G_\bullet}$ to its $(S \ltimes \rho')$-ergodic decomposition, and hence each of the measures $P'((s, K'_s g), \cdot)$ to ergodic measures supported on the disjoint sets $\{s\} \times K'_s g'$. Since ρ' almost surely takes values in K_s and these components must integrate up to $\nu \ltimes m_{G_\bullet}$, it follows that $\psi_\# P_\# (\zeta_0^S)_\# \nu = \nu \ltimes m_{K'_\bullet g(\bullet)}$ for some $g : Z_0^S \to G_\bullet$, and now applying ψ^{-1} to this equation and replacing $K_\bullet := g(\bullet)^{-1} K'_\bullet g(\bullet)$ and $b := b' \cdot g$ gives the result. ∎

Once again, the result for extensions by group data implies a version for extensions by homogeneous space data through lifting to a covering group extension and then descending again, just as for Corollary 4.4. The proof is essentially the same, and so we omit it here.

COROLLARY 4.8. *Suppose that* $S : \Gamma \curvearrowright (Y, \nu)$, $H_\bullet \leq G_\bullet$ *are* Z_0^S-*measurable group data and* $\rho : \Gamma \times Y \to G_\bullet$ *is a cocycle-section over* S *and* X *is the space* $Y \ltimes G_\bullet / H_\bullet$ *but equipped with some unknown* $(S \ltimes \rho)$-*invariant and relatively ergodic lift* μ *of* ν. *Then there are subgroup data* $K_\bullet \leq G_\bullet$ *on* Z_0^S *and a section* $b : Y \to G_\bullet$ *such that* $\mu = \nu \ltimes m_{b(\bullet)^{-1} K_\bullet H_\bullet / H_\bullet}$. ∎

Arguing exactly as in the classical case of an ergodic base system by replacing some given group data $G_\bullet$ with the Mackey group data $K_\bullet$ and recoordinatizing (see Corollary 3.27 in Glasner [28]), we obtain the following corollary.

COROLLARY 4.9. *Given a* Γ-*system* $\mathbf{Y} = (Y, \nu, S)$, *measurable* S-*invariant homogeneous space data* $G_\bullet / K_\bullet$ *over* Y *and a cocycle-section* $\rho : \Gamma \times Y \to G_\bullet$, *and defining* $X := Y \ltimes G_\bullet / K_\bullet$ *and* $T := S \ltimes \rho$, *any* $(S \ltimes \rho)$-*relatively ergodic lift* μ *of* ν *admits a recoordinatization*

$$(X, \mu, T) \xleftarrow{\quad \cong \quad} \mathbf{Y} \ltimes (G'_\bullet / H'_\bullet, m_{G'_\bullet / H'_\bullet}, \rho')$$

$$\text{canonical} \searrow \qquad \swarrow \text{canonical}$$

$$\mathbf{Y}$$

so that the implicit covering group extension $\mathbf{Y} \ltimes (G'_\bullet, m_{G'_\bullet}, \rho') \to \mathbf{Y}$ *is also relatively ergodic.* ∎

During the development of the inverse theory of the next section we will use the preceding results in conjunction with the following elementary lemma.

LEMMA 4.10. *Suppose that* $\pi : (X, \mu, T) \to (Y, \nu, S)$ *is a relatively ergodic extension of a not necessarily ergodic system, and that*

$$(X, \mu, T) \xrightarrow{\pi_{(n)}} (Z_{(n)}, \mu_{(n)}, T_{(n)}) \xrightarrow{\xi_{(n)}} (Y, \nu, S)$$

for $n = 1, 2, \ldots$ is a sequence of intermediate extensions that are all coordinatizable as extensions by homogeneous-space data. Then the resulting joint extension

$$(Z_{(1)} \times Z_{(2)} \times \cdots, (\pi_{(1)} \vee \pi_{(2)} \vee \cdots)_\# \mu, T|_{\pi_{(1)} \vee \pi_{(2)} \vee \cdots}) \to (Y, \nu, S)$$

is also coordinatizable as an extension by homogeneous-space data.

REMARK. This is a straightforward extension of Lemma 8.4 in Furstenberg [22].

Proof. We know that for each $n \geq 1$ there are some Z_0^S-measurable homogeneous-space data $G_{n,\bullet}/H_{n,\bullet}$, a cocycle-section $\rho_n : \Gamma \times Y \to G_{n,\bullet}$ and a coordinatizing isomorphism

$$\alpha_n : (Z_{(n)}, \mu_{(n)}, T_{(n)}) \overset{\cong}{\to} (Y, \nu, S) \ltimes (G_{n,\bullet}/H_{n,\bullet}, m_{G_{n,\bullet}/H_{n,\bullet}}, \rho_n).$$

Each $\alpha_n \circ \pi_{(n)}$ is a factor map of X that gives a recoordinatization of the factor associated to $\pi_{(n)}$ and takes the form (π, θ_n) for a suitable map $\theta_n : X \to G_{n,\bullet}/H_{n,\bullet}$. Now we simply set $G_\bullet := \prod_{n \geq 1} G_{n,\bullet}$, $H_\bullet := \prod_{n \geq 1} H_{n,\bullet}$, $\rho := (\rho_n)_{n \geq 1}$ and $\theta := (\theta_n)_{n \geq 1}$; it is clear that the map (π, θ) gives the desired recoordinatization of $(Z_{(1)} \times Z_{(2)} \times \cdots, (\pi_{(1)} \vee \pi_{(2)} \vee \cdots)_\# \mu, T|_{\pi_{(1)} \vee \pi_{(2)} \vee \cdots})$ as an extension of Y by homogeneous-space data. ∎

5. Relative weak non-mixing and isometric extensions of non-ergodic systems. In this section we shall recount the main results of our non-ergodic version of the Furstenberg–Zimmer inverse theory. Although it seems that these non-ergodic analogs do not formally follow from their ergodic predecessors, their proofs largely follow the original arguments of Furstenberg and Zimmer, with a few judicious invocations of measurable selectors along the way. For this reason our presentation here, as in the preceding section, will be quite terse. The original papers [49, 48] and [22] remain clear and thorough references for the classical results in the presence of ergodicity, and we direct the reader to these for many of the original ideas.

The theory developed by Furstenberg and Zimmer considers an ergodic extension of an ergodic system $\mathbf{Y} = (Y, \nu, S)$. Given such an extension $\pi : \mathbf{X} \to \mathbf{Y}$, this theory gives an account of the possible failure of ergodicity of the relatively independent self-joining $\mathbf{X} \times_\pi \mathbf{X}$ (that is, of the 'relative weak mixing' of $\mathbf{X}$ over π): it turns out that this occurs if and only if the extension contains a non-trivial subextension that can be coordinatized as a homogeneous skew-product. It is this result that we shall presently extend by dropping the assumption that S be ergodic, and by working instead with extension by (possibly variable) homogeneous space data. Note, however, that we will continue to assume *relative* ergodicity of $\mathbf{X} \to \mathbf{Y}$: if this fails then the arguments that follow derail quite quickly, and the best account of the structure of the extension that can be given in this case seems to result from

simply considering the relatively invariant subextension first, and then working with the remaining (necessarily relatively ergodic) extension over that.

In fact, here as in the ergodic setting just a little extra work will show that once the failure of relative weak mixing is understood in this way, the same structures account for the non-ergodicity of other relatively independent joinings: given two extensions $\pi_i : \mathbf{X}_i = (X_i, \mu_i, T_i) \to \mathbf{Y}_i$, $i = 1, 2$, a joining ν of $\mathbf{Y}_1$ and $\mathbf{Y}_2$ and a lift μ of ν to a joining of $\mathbf{X}_1$ and $\mathbf{X}_2$ under which the copies of these two factors are relatively independent over the copies of $\mathbf{Y}_1$ and $\mathbf{Y}_2$, then this larger joining μ can fail to be $(T_1 \times T_2)$-ergodic only if each of the extensions $\mathbf{X}_i \to \mathbf{Y}_i$ contains a non-trivial subextension that is coordinatizable as a homogeneous skew-product, and the homogeneous spaces and cocycles of this skew-product are suitably related to each other. The result above simply corresponds to the case $\mathbf{X}_1 = \mathbf{X}_2$, $\mathbf{Y}_1 = \mathbf{Y}_2$ and λ the diagonal self-joining of the smaller system.

5.1. Generalized eigenfunctions and finite-rank modules. Key to the reduction from the failure of relative weak mixing to non-trivial extensions by homogeneous space data are the notions of finite rank modules and isometric extensions. These definitions are taken almost unchanged from the papers of Furstenberg [22] and Zimmer [48].

DEFINITION 5.1 (Modules over factors and their rank). If $\pi : (X, \mu) \to (Y, \nu)$ is an extension of standard Borel probability spaces and $\mathfrak{M}$ is a closed subspace of $L^2(\mu)$, then we shall refer to $\mathfrak{M}$ as a π-*module* if $(h \circ \pi) \cdot f \in \mathfrak{M}$ whenever $f \in \mathfrak{M}$ and $h \in L^\infty(\nu)$.

If $\int_Y^\oplus \mathfrak{H}_y \, \nu(\mathrm{d}y)$ is the direct integral decomposition of $L^2(\mu)$ over π, then subordinate to this we may form the direct integral decomposition $\int_Y^\oplus \mathfrak{M}_y \, \nu(\mathrm{d}y)$ of $\mathfrak{M}$ over π; and now we shall write that $\mathfrak{M}$ has *rank r over π* if $r \in \{1, 2, \ldots, \infty\}$ is minimal such that $\dim \mathfrak{M}_y \leq r$ for ν-almost every y. If $r < \infty$ we shall write that $\mathfrak{M}$ has *finite π-rank*.

DEFINITION 5.2 (Isometric extension). A system extension $\pi : \mathbf{X} \to \mathbf{Y}$ is *isometric* if $L^2(\mu)$ is generated as π-module by its finite rank T-invariant π-submodules.

The following first step towards representing finite rank modules will be crucial.

LEMMA 5.3 (Orthonormal basis for a module). *Suppose that $\mathfrak{M}$ is a rank-r π-module for some $r < \infty$. Then there is a tuple $\phi_1, \ldots, \phi_r$ of functions in $\mathfrak{M}$, none of them vanishing everywhere, and a measurable function $R : Y \to \{1, \ldots, r\}$ such that $\mathsf{E}_\mu(\phi_i \cdot \overline{\phi_j} \,|\, \pi) = \delta_{i,j} 1_{\{i,j \leq R\}}$ and*

$$(L^\infty(\nu) \circ \pi) \cdot \phi_1 + \cdots + (L^\infty(\nu) \circ \pi) \cdot \phi_r$$

is L^2-dense in $\mathfrak{M}$.

Proof. Let $R(y) := \dim \mathfrak{M}_y$ and decompose Y as $\bigcup_{s \leq r} \{y : R(y) = s\}$. It is easy to check that R must be measurable, and the existence of suitable ϕ_i now follows just as in the classical case by considering the cells $\{y : R(y) = s\}$ separately: see Lemma 9.4 of Glasner [28]. It is clear that none of the ϕ_i can vanish everywhere, else the module $\mathfrak{M}$ would actually have rank at most $r - 1$, contradicting our assumptions. ■

DEFINITION 5.4. We refer to the function R above as the *local rank function* of the module $\mathfrak{M}$.

The following lemma is trivial, but it will prove convenient to be able to call on it explicitly.

LEMMA 5.5. *If* $(X, \mu) \xrightarrow{\pi} (Z, \theta) \xrightarrow{\pi'} (Y, \nu)$ *is a tower of probability-preserving maps and* $\mathfrak{M} \leq L^2(\mu)$ *is a* $(\pi' \circ \pi)$-*module of rank* r, *then* $\overline{L^\infty(\theta) \cdot \mathfrak{M}}$ *is a* π-*module of rank at most* r. ■

5.2. The non-ergodic Furstenberg–Zimmer inverse theorem. We now present our non-ergodic extension of the Furstenberg–Zimmer inverse theorem. This follows naturally from two separate propositions.

PROPOSITION 5.6 (From relative non-weak-mixing to finite rank modules). *If* $\pi : \mathbf{X} \to \mathbf{Y}$ *is relatively ergodic but not relatively weakly mixing, then* $L^2(\mu)$ *contains a non-trivial* T-*invariant finite-rank* π-*submodule.*
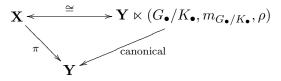
Proof. This first proposition is proved just as in the ergodic case, so we shall only sketch its proof, referring the reader to Chapter 9 in Glasner [28] or Section 7 of Furstenberg [22] for a more careful treatment. Form the relatively independent self-product $\mathbf{X} \times_\pi \mathbf{X}$ with its natural coordinate projections $\pi_1, \pi_2 : \mathbf{X} \times_\pi \mathbf{X} \to \mathbf{X}$. Note that by construction these maps quotient to give a single factor copy of $\mathbf{Y}$ through π, and so up to $(\mu \otimes_\pi \mu)$-almost-everywhere equality of functions we have $L^2(\nu) \circ \pi \circ \pi_1 = L^2(\nu) \circ \pi \circ \pi_2$. The key idea is to choose a $(T \times T)$-invariant function H that lies in $L^2(\mu \otimes_\pi \mu) \setminus (L^2(\nu) \circ \pi \circ \pi_1)$, and then define from it a bounded operator A on $L^2(\mu)$ by

$$A\psi := \mathsf{E}_{\mu \otimes_\pi \mu}\big((\psi \circ \pi_2) \cdot H \,\big|\, \pi_1\big).$$

It is easy to see that this cannot be the identity if $H \notin L^2(\nu) \circ \pi \circ \pi_1$. Replacing $H(x, x')$ by either $H(x, x') + \overline{H(x', x)}$ or $H(x, x') - \mathrm{i}\overline{H(x', x)}$ if necessary, we may also arrange that A be self-adjoint. In defining this A we have produced a 'relative Hilbert–Schmidt operator', acting as a Hilbert–Schmidt operator separately on each fibre $\mathfrak{H}_y$ of the Hilbert space direct integral decomposition $L^2(\mu) = \int_Y^\oplus \mathfrak{H}_y \, \nu(\mathrm{d}y)$. It now admits a spectral decomposition into finite-dimensional eigenspaces in each of these fibres, relative over the factor π. In principle the list of corresponding eigenvalues can vary with $y \in Y$, but it is standard (see Section 9.3 of Glasner [28]) that they do

so measurably and are S-invariant. Now we can simply make a measurable selection of one of these non-zero eigenvalues over each $\zeta_0^S(y) \in Z_0^S$ and associate to that eigenvalue its corresponding finite-dimensional eigenspace $\mathfrak{M}_y$, to produce a non-trivial π-submodule $\mathfrak{M}$ of $L^2(\mu)$ such that $\dim \mathfrak{M}_y$ is S-invariant and almost surely finite, and now truncating this by retaining the non-trivial space $\mathfrak{M}_y$ only on the S-invariant set $\{y : \dim \mathfrak{M}_y < M\}$ for some sufficiently large M gives a true finite-rank module. Its T-invariance follows immediately from that of H and hence of A, completing the proof. ∎

Our machinery of direct integrals of homogeneous spaces becomes necessary for the second stage of the argument.

PROPOSITION 5.7 (Coordinatization of isometric extensions). *If π : $\mathbf{X} \to \mathbf{Y}$ is a relatively ergodic extension for which $L^2(\mu)$ is generated by its T-invariant finite-rank π-submodules, then there are Z_0^S-measurable homogeneous space data $G_\bullet/K_\bullet$ with fibre repository $U := \prod_{n \geq 1} \mathrm{U}(n)^{\mathbb{N}}$ and a cocycle-section $\rho : \Gamma \times Y \to G_\bullet$ such that*

$$\mathbf{X} \xleftarrow{\;\;\cong\;\;} \mathbf{Y} \ltimes (G_\bullet/K_\bullet, m_{G_\bullet/K_\bullet}, \rho)$$
$$\pi \searrow \qquad \swarrow \text{canonical}$$
$$\mathbf{Y}$$

Proof. Let $\mathfrak{M}^{(n)}$ for $n \geq 1$ be a sequence of finite-rank π-modules generating $L^2(\mu)$; since $L^2(\mu)$ is separable we need only countably many.

Suppose $\mathfrak{M}^{(n)}$ has rank r and let $\phi_1, \ldots, \phi_r$ be an orthonormal basis for it as guaranteed by Lemma 5.3. It is easy to check that $\phi_j \circ T^\gamma$, $j \leq r$, also form an orthonormal basis for $\mathfrak{M}^{(n)}$ for each $\gamma \in \Gamma$, and so we can find a measurable cocycle $\Phi : \Gamma \times Y \to \mathrm{U}(R(\bullet))$ with values in the finite-dimensional unitary group such that

$$(\phi_1|_{\pi^{-1}(S^\gamma y)} \circ T^\gamma, \ldots, \phi_{R(y)}|_{\pi^{-1}(S^\gamma y)} \circ T^\gamma)$$
$$= \Phi(\gamma, y)(\phi_1|_{\pi^{-1}(y)}, \ldots, \phi_{R(y)}|_{\pi^{-1}(y)})$$

in $\mathfrak{H}_y \times \cdots \times \mathfrak{H}_y$ for ν-almost every y (indeed, this equation serves as the definition of Φ and witnesses its measurability as a function of (γ, y)). From this it follows that $\sum_{j \leq r} |\phi_j(x)|^2$ is T-invariant, and by relative ergodicity can therefore be factorized through π. It must therefore equal $R(\pi(x))$ almost everywhere by the relative orthonormality of the ϕ_j.

Letting

$$\phi : x \mapsto \frac{1}{R(\pi(x))}(\phi_1(x), \ldots, \phi_{R(y)}(x)),$$

it follows that $x \mapsto (\pi(x), \phi(x))$ is a map $X \to Y \ltimes \mathrm{S}^{2R(\bullet)-1}$ (where S^{2r-1} denotes the unit sphere in $\mathbb{C}^r$) which intertwines T with $S \ltimes \Phi$. Letting

$\xi : Y \to S^{2R(\bullet)-1}$ be a measurable selection (in this case we could take it to be constant on the level sets of R) and then noting that $U(R(\bullet)) \curvearrowright S^{2R(\bullet)-1} \cong U(R(\bullet))/\text{Stab}_{U(R(\bullet))}(\xi(\bullet))$, it follows that (π, ϕ) coordinatizes a subextension of $\mathbf{X} \to \mathbf{Y}$ as the homogeneous space data extension $S \ltimes \Phi \curvearrowright Y \ltimes U(R(\bullet))/\text{Stab}_{U(R(\bullet))}(\xi(\bullet))$ carrying some invariant measure. Finally, by Corollary 4.9 we know that this can be adjusted to a genuine extension by homogeneous space data carrying the associated direct integral measure.

Writing this subextension as $\mathbf{X} \xrightarrow{\pi^{(n)}} \mathbf{Z}^{(n)} \xrightarrow{\pi^{(n)\prime}} \mathbf{Y}$, it is clear that $L^2(\pi^{(n)}_{\#}\mu)\circ\pi^{(n)} \geq \mathfrak{M}^{(n)}$. Hence the target system of the factor $\pi^{(1)}\vee\pi^{(2)}\vee\cdots$ contains every $\mathfrak{M}^{(n)}$, and so must be equivalent to the whole system $\mathbf{X}$. Finally, Lemma 4.10 assures us that this can still be coordinatized as an extension by homogeneous space data, completing the proof. ∎

We should also check the converse of the preceding proposition in the non-ergodic setting.

LEMMA 5.8. *A relatively ergodic extension by compact homogeneous space data* $\mathbf{X} := \mathbf{Y} \ltimes (G_\bullet/K_\bullet, m_{G_\bullet/K_\bullet}, \rho)$ *with canonical factor map* $\pi : \mathbf{X} \to \mathbf{Y}$ *is generated by its finite-rank π-submodules.*

Proof. Let U be the compact fibre repository for the data $G_\bullet$, and let $\mathfrak{H}_y := L^2(m_{G_y/K_y})$ for $y \in Y$; it is easy to check that this defines a measurable family of Hilbert spaces. Let $L : G_y \curvearrowright \mathfrak{H}_y$ be the left-regular unitary action. Now for any continuous function ψ on U the associated operators

$$A_y := \int\limits_{G_y} \psi(g)L_g \, m_{G_\bullet}(\mathrm{d}g)$$

clearly form a measurable family in y and are each a compact operator on $\mathfrak{H}_y$. Moreover, their eigenspaces are L_{G_y}-invariant, and so different measurable selections of these eigenspaces now combine to form finite-rank π-submodules of $L^2(\mu) \cong \int_Y^\oplus \mathfrak{H}_y \, \nu(\mathrm{d}y)$ that are T-invariant. Finally choosing a single sequence $(\psi_n)_{n\geq 1}$ of continuous mollifiers on U (that is, of continuous functions on U such that $m_U\llcorner\psi_n \to \delta_{1_U}$ in the vague topology as $n \to \infty$), all possible measurable selections of eigenspaces of their associated compact operators together generate the whole of $L^2(\mu)$, completing the proof. ∎

Given an arbitrary relatively ergodic extension $\pi : \mathbf{X} \to \mathbf{Y}$ that is not relatively weakly mixing, Proposition 5.6 guarantees that its subextension generated by all finite-rank π-submodules is non-trivial; and now applying Proposition 5.7 to this subextension immediately gives our full version of the inverse theorem. Combined with the 'direct' result of Lemma 5.8 this gives the following.

THEOREM 5.9 (Furstenberg–Zimmer Theorem in the non-ergodic setting). *Suppose that the extension* $\pi : \mathbf{X} \to \mathbf{Y}$ *is relatively ergodic. Then*

there is a unique maximal subextension $\mathbf{X} \to \mathbf{Z} \to \mathbf{Y}$ *that can be coordinatized as an extension of* $\mathbf{Y}$ *by homogeneous space data, it equals the subextension generated by all finite-rank T-invariant π-modules, and π fails to be relatively weakly mixing for T if and only if this subextension contains π strictly.* ∎

DEFINITION 5.10 (Maximal isometric subextension). The subextension $\mathbf{Z} \to \mathbf{Y}$ given by the preceding theorem is the *maximal isometric subextension* of $\pi : \mathbf{X} \to \mathbf{Y}$.

REMARK. Let us digress to locate the need for relative ergodicity in the above arguments. This occurred during the proof of Proposition 5.7 when we argued that our orthonormal basis $\phi_1, \ldots, \phi_r$ must have $\sum_{j \leq r} |\phi_j(x)|^2$ measurable with respect to π in view of its T-invariance, and from this that we could synthesize from this basis a map $X \to \mathrm{S}^{2R(\bullet)-1}$ which together with π would lead to an explicit coordinatization by homogeneous space data.

It is clear that given an extension that is not relatively ergodic, we can still derive some structural consequences from the failure of relative weak mixing as follows. A simple check shows that the argument that converts a non-trivial invariant function on $\mathbf{X} \times_\pi \mathbf{X}$ to a finite-rank module over π does not require relative weak mixing, and so we can still sensibly define the subextension $\mathbf{X} \xrightarrow{\xi} \mathbf{Z} \xrightarrow{\alpha} \mathbf{Y}$ generated by all finite-rank π-modules. Next, any function on X that is actually T-invariant clearly defines a rank-1 such module, and so ξ certainly contains the factor $\zeta_0^T \vee \pi$. Now, in addition, any finite-rank module $\mathfrak{M}$ over the factor π gives a finite-rank module $\overline{L^\infty((\zeta_0^T \vee \pi)_{\#}\mu) \cdot \mathfrak{M}}$ over $\zeta_0^T \vee \pi$. In light of this and the inverse theory for the relatively ergodic case, $\xi : \mathbf{X} \to \mathbf{Z}$ must actually be contained in the maximal isometric subextension of the joint factor map $\zeta_0^T \vee \pi$. Letting $\mathbf{W}$ be the target system of $\zeta_0^T \vee \pi$, the maximal isometric subextension of the relatively ergodic extension $\zeta_0^T \vee \pi$ now *is* coordinatizable as an extension by compact homogeneous space data, and from the results of the next section it will follow that the resulting subextension $\mathbf{Z} \to \mathbf{W}$ is also coordinatizable by compact homogeneous space data. It may, however, be properly contained in the maximal isometric subextension of $\zeta_0^T \vee \pi$, since there may be finite-rank modules over $\zeta_0^T \vee \pi$ that cannot be obtained as above from finite-rank modules over π.

We suspect that more can be said in general about which subextensions of $\zeta_0^T \vee \pi$ can be obtained from finite-rank modules over π, but we will not explore this matter further here. Note that in several previous works, such as Furstenberg and Katznelson's proof of multidimensional multiple recurrence in [24] and their later applications of similar ideas to prove other results in density Ramsey theory in [25] and [26], this very concrete analysis in terms

of extensions by homogeneous space data is avoided altogether. In its place is used a much softer property of extension called 'relative compactness', which is also a consequence of its being generated by finite-rank submodules and turns out to be enough to enable a proof of the relevant multiple recurrence results without the more precise information offered by a coordinatization. We will not explore relations with this idea further here, but refer the reader to Furstenberg's book [23] for a treatment of such arguments.

We can now extend the following definition from the ergodic-base case.

DEFINITION 5.11 (Distal extensions and distal towers). Given a system extension $\pi : \mathbf{X} \to \mathbf{Y}$ we note that $\zeta_0^T \vee \pi$ coordinatizes the maximal factor of $\mathbf{X}$ that is relatively invariant over π, and now write $\zeta_{1/\pi}^T : \mathbf{X} \to \mathbf{Z}_1^T(\mathbf{X}/\pi)$ to denote any choice of factor map that coordinatizes the maximal isometric subextension of $\zeta_0^T \vee \pi$. In general we define recursively an increasing transfinite sequence of factors $\zeta_{\eta/\pi}^T : \mathbf{X} \to \mathbf{Z}_{\eta/\pi}^T(\mathbf{X}/\pi)$ indexed by all ordinals η by letting $\zeta_{\eta+1/\pi}^T : \mathbf{X} \to \mathbf{Z}_{\eta+1/\pi}^T(\mathbf{X}/\pi)$ denote any choice of factor map that coordinatizes the maximal isometric subextension of $\zeta_{\eta/\pi}^T$ for each η, and letting $\zeta_{\eta/\pi}^T := \bigvee_{\kappa < \eta} \zeta_{\eta/\pi}^T$ when η is a limit ordinal. Note that for any fixed system $\mathbf{X}$ there must be some ordinal $\leq \omega_1$ at which this tower stabilizes. If $\mathbf{Y}$ is a trivial system $(\{*\}, \delta_*, \mathrm{id}_{\{*\}})$ we simplify this notation to $\zeta_\eta^T : \mathbf{X} \to \mathbf{Z}_\eta^T$.

The extension $\pi : \mathbf{X} \to \mathbf{Y}$ is *distal* if $\bigvee_\eta \zeta_{\eta/\pi}^T \simeq \mathrm{id}_X$. We refer to $\mathbf{Z}_\eta^T(\mathbf{X}/\pi) \to \mathbf{Y}$ as the *maximal η-step distal subextension* of π, and to the totally ordered collection of all of these subextensions as the *distal tower* of π.

As suggested at the beginning of this section, our inverse theorem can quite easily be extended to account for relative non-ergodicity of more general relatively independent self-joinings. The standard proof of this in the case of ergodic base (Theorem 9.21 in Glasner [28]), which reduces this situation to that of the self-joining treated above, does not rely on the ergodicity of the base system and so carries over essentially unchanged. We only state the result here.

THEOREM 5.12. *Suppose that $\pi_i : \mathbf{X}_i \to \mathbf{Y}_i$ are relatively ergodic extensions for $i = 1, \ldots, n$ and that ν is a joining of $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ forming the system $\mathbf{Y} = (Y, \nu, S) := (Y_1 \times \cdots \times Y_n, \nu, S_1 \times \cdots \times S_n)$. Suppose further that $\mathbf{X} = (X, \mu, T)$ is similarly a joining of $\mathbf{X}_1, \ldots, \mathbf{X}_n$ that extends ν through the coordinatewise factor map $\pi : \mathbf{X} \to \mathbf{Y}$ assembled from the π_i, and such that under μ the coordinate projections $\alpha_i : \mathbf{X} \to \mathbf{X}_i$ are relatively independent over the tuple of further factors $\pi_i \circ \alpha_i$. Then the intermediate factor*

map

$$\zeta_{1/\pi_1}^{T_1} \vee \cdots \vee \zeta_{1/\pi_n}^{T_n} : \mathbf{X} \to \mathbf{Z}$$

whose target $\mathbf{Z}$ *is a joining of the systems* $\mathbf{Z}_1^{T_i}(\mathbf{X}_i/\pi_i)$ *is equivalent to* $\zeta_{1/\pi}^{T}$: $\mathbf{X} \to \mathbf{Z}_1^{T}(\mathbf{X}/\pi)$. *In particular, it contains the relatively invariant extension* $\mathbf{Z}_0^{T} \vee \mathbf{Y} \to \mathbf{Y}$, *which may be non-trivial.* ∎

We will call on this version of the inverse theorem when we come to our applications in Section 7.

6. Factors and automorphisms of isometric extensions. In this subsection we examine the possible forms of factors and automorphisms of extensions by homogeneous space data, using as our main tool the non-ergodic Mackey theory of Section 4.

6.1. Some more notation. We first need to set up some additional notation that will help us to describe algebraic transformations between the fibres G_y/K_y of our extensions.

Our first important convention is that given a Polish group U and two compact subgroups $G, G' \le U$, we identify a continuous isomorphism $\Phi : G \xrightarrow{\cong} G'$ with its graph $\{(g, \Phi(g)) : g \in G\} \le U \times U$. In view of the results of Subsection 2.4, this sets up a bijective correspondence between continuous isomorphisms and compact subgroups $M \in \text{Lat}(U \times U)$ with the property that M has first and second projections equal to G and G' respectively and $M \cap (U \times \{1_U\}) = M \cap (\{1_U\} \times U) = \{1_{U \times U}\}$. We write $\text{Isom}(G, G')$ for this collection of isomorphisms, and interpret composition $\Phi \circ \Phi'$ for $\Phi \in \text{Isom}(G, G')$ and $\Phi' \in \text{Isom}(G', G'')$ in the obvious way. Note that after fixing a complete separable metric on U, the resulting strong topology on $\text{Isom}(G, G')$ is in general strictly stronger than the Vietoris topology inherited from $\text{Lat}(U \times U)$, but the resulting Borel structures are the same. Exactly similarly we can also interpret the collection $\text{Hom}(G, G')$ of homomorphisms $G \to G'$, $\text{Aut}\,G = \text{Isom}(G, G)$ and $\text{End}\,G = \text{Hom}(G, G)$ as collections of subgroups of $U \times U$.

This identification made, the necessary definition follows very naturally.

DEFINITION 6.1 (Homomorphisms sections). Suppose that (Y, ν) is a standard Borel probability space and that $G_{i,\bullet}$, $i = 1, 2$, are two different measurable families of compact group data over Y with compact metrizable repositories U and V respectively. Then a *homomorphism section* associated to this data is a map $y \mapsto \Phi_y : Y \to \text{Lat}(U \times V)$ that is Borel for the Vietoris measurable structure on $\text{Lat}(U \times V)$ and is such that $\Phi_y \in \text{Hom}(G_{1,y}, G_{2,y})$ for ν-almost every $y \in Y$. We will sometimes denote this situation by $\Phi : Y \to \text{Hom}(G_{1,\bullet}, G_{2,\bullet})$.

We extend this definition in the obvious way to epimorphism sections, isomorphism sections and automorphism sections.

The benefit of formulating our notion of 'isomorphism section' as above will become clear shortly, when we use the Mackey theory in a self-joining of a given extension to produce a measurable family of compact subgroups of $U \times U$ that we can then immediately reinterpret as an isomorphism cocycle in this sense. In doing this we will also benefit from having the following notation.

DEFINITION 6.2 (Extensions by epimorphism sections). If $\beta : (Y_1, \nu_1) \to (Y_2, \nu_2)$ is a probability-preserving map, $G_{i,\bullet}/H_{i,\bullet}$ is measurable compact homogeneous space data on Y_i for $i = 1, 2$ and $\Phi : Y \to \mathrm{Hom}(G_{1,y}, G_{2,\beta(y)})$ is a measurable epimorphism section such that $\Phi_y(H_{1,y}) \subseteq H_{2,\beta(y)}$ almost surely then we write $\beta \ltimes \Phi_\bullet$ for the probability-preserving map

$$Y_1 \ltimes G_{1,\bullet}/H_{1,\bullet} \to Y_2 \ltimes G_{2,\bullet}/H_{2,\bullet} : (y, gH_{1,y}) \mapsto (\beta(y), \Phi_y(g)H_{2,\beta(y)}).$$

More generally, if $\rho : Y \to G_{2,\beta(\bullet)}$ is a section then we will sometimes write $L_{\rho(y)}$ (resp. $R_{\rho(y)}$) for the translation of $G_{2,\beta(\bullet)}$ by left rotation by $\rho(y)$ (resp. right rotation by $\rho(y)$), and define

$$\beta \ltimes (L_{\rho(\bullet)} \circ \Phi_\bullet)|_{H_{2,\beta(\bullet)}}^{H_{1,\bullet}} : (y, gH_{1,\bullet}) \mapsto (\beta(y), \rho(y) \cdot \Phi_y(g)H_{2,\beta(y)}).$$

In case $H_{i,\bullet} \equiv \{1_{G_{i,\bullet}}\}$ we simplify this to

$$\beta \ltimes (L_{\rho(\bullet)} \circ \Phi_\bullet) : (y, g) \mapsto (\beta(y), \rho(y) \cdot \Phi_y(g)).$$

Note that if $(Y_1, \nu_1) = (Y_2, \nu_2)$ then $\beta \ltimes L_{\rho(\bullet)} = \beta \ltimes \rho$ and $\beta \ltimes R_{\rho(\bullet)} = \beta \ltimes \rho^{\mathrm{op}}$ in the original notation of Section 3. We introduce the above class of fibrewise transformations, together with their new notation, to help us describe more explicitly certain maps and joinings between isometric extensions, in the sense of the following definition, and to help differentiate them from those isometric extensions themselves, among which these new maps serve as morphisms.

DEFINITION 6.3 (Fibrewise automorphism and affine recoordinatizations). Suppose that $\mathbf{Y} = (Y, \nu, S)$ is a Γ-system, $G_\bullet/K_\bullet$ are S-invariant compact homogeneous space data over Y and $\sigma : \Gamma \times Y \to G_\bullet$ is an ergodic cocycle-section for S, and that $\Phi_\bullet : Y \to \mathrm{Aut}(G_\bullet)$ is an S-invariant automorphism section and $\rho : Y \to G_\bullet$ a section. Then the map $R := \mathrm{id}_Y \ltimes (L_{\rho(\bullet)} \circ \Phi_\bullet)|_{\Phi_\bullet(K_\bullet)}^{K_\bullet}$ defines a recoordinatization

$$\mathbf{Y} \ltimes (G_\bullet/K_\bullet, m_{G_\bullet/K_\bullet}, \sigma) \xleftarrow{\quad R \quad} \mathbf{Y} \ltimes (G_\bullet/K'_\bullet, m_{G_\bullet/K'_\bullet}, \sigma')$$

$$\text{canonical} \searrow \qquad \swarrow \text{canonical}$$

$$\mathbf{Y}$$

with $K'_\bullet := \Phi_\bullet(K_\bullet)$ and

$$\sigma'(\gamma, y) := \rho(S^\gamma y) \cdot \Phi_y(\sigma(\gamma, y)) \cdot \rho(y)^{-1}.$$

We refer to such a recoordinatization as a *fibrewise affine recoordinatization*. If $\rho \equiv 1_{G_\bullet}$ then it is a *fibrewise automorphism recoordinatization*.

REMARKS. 1. The condition that $\Phi_\bullet$ be S-invariant is needed in order that R intertwine $S \ltimes \sigma$ with another cocycle extension; without this condition the new transformation will in general still involve also a non-trivial automorphism-valued cocycle.

2. If $G_\bullet$ and $\Phi_\bullet$ are as above then we have $\beta \ltimes (L_{\rho(\bullet)} \circ \Phi_\bullet) = \beta \ltimes (\mathrm{Co}_{\rho(\bullet)} \circ \Phi_\bullet \circ R_{\rho(\bullet)})$, where $\mathrm{Co}_{\rho(y)} \in \mathrm{Aut}(G_y)$ is the inner automorphism of conjugation by $\rho(y)$. It follows that we can always choose between expressing a fibrewise affine recoordinatization by using fibrewise left rotations or fibrewise right rotations (or even some mixture of the two!). In general we will prefer to write fibrewise affine recoordinatizations with rotation part acting on the left, since this is the form in which the S-invariance of the automorphism part $\Phi_\bullet$ is directly visible, but occasionally this ability to conjugate between the two forms will give some useful flexibility in how we write a fibrewise affine recoordinatization.

Note that transformations of the form $\beta \ltimes (L_{\rho(\bullet)} \circ \Phi_\bullet) \curvearrowright U \times V$ for compact Abelian groups U and V, β a rotation of U, $\rho : U \to V$ and $\Phi : U \to \mathrm{Aut}\, V$ are already objects of study in ergodic theory: examples arise from explicitly coordinatizing a flow on a two-step solvmanifold (see, for example, Auslander, Green and Hahn [2] or Starkov [41] for background), with rather special conditions on the cocycles ρ and Φ that result (although working explicitly with such coordinatizations would probably be a cumbersome way to handle solvflows). More generally, it is possible that an enlarged theory of extensions by compact homogeneous space data could be constructed to treat actions lifted by affine-valued cocycle comprising both fibrewise rotations and automorphisms. Certainly, $\mathbb{Z}^d$-actions by affine maps on compact groups have recently begun to receive greater attention from ergodic theorists (see Schmidt [39], in particular), and it would be interesting to know how far the results that are now known for such systems could be 'relativized'. In this paper, however, fibrewise automorphisms and affine maps will play a strictly auxiliary rôle.

6.2. The structure theorems. The main results of this section amount to structure theorems for factors and automorphisms of isometric extensions, generalizing the classical result (and also the proof) of Mentzen [36] in the ergodic case (see also earlier work of Newton [37], and compare with the classical argument of Veech in [44] for his condition for an extension to be coordinatizable as a compact group skew-product).

THEOREM 6.4 (Relative Factor Structure Theorem). *Suppose that $\mathbf{Y}_i =$ (Y_i, ν_i, S_i) for $i = 1, 2$ are Γ-systems and $\beta : \mathbf{Y}_1 \to \mathbf{Y}_2$ is a factor map, that $G_{i,\bullet}/H_{i,\bullet}$ are S_i-invariant core-free homogeneous space data on Y_i and that $\sigma_i : \Gamma \times Y_i \to G_{i,\bullet}$ are ergodic cocycle-sections for the action S_i, and let $\mathbf{X}_i = (X_i, \mu_i, T_i) := \mathbf{Y}_i \ltimes (G_{i,\bullet}/H_{i,\bullet}, m_{G_{i,\bullet}/H_{i,\bullet}}, \sigma_i)$ with canonical factor $\xi_i : \mathbf{X}_i \to \mathbf{Y}_i$. Suppose further that β admits extension to a factor map $\alpha : \mathbf{X}_1 \to \mathbf{X}_2$:*

$$
\begin{array}{ccc}
\mathbf{X}_1 & \xrightarrow{\ \alpha\ } & \mathbf{X}_2 \\
{\scriptstyle \xi_1} \downarrow & & \downarrow {\scriptstyle \xi_2} \\
\mathbf{Y}_1 & \xrightarrow{\ \beta\ } & \mathbf{Y}_2
\end{array}
$$

Then there are an S_1-invariant measurable family of epimorphisms $\Phi_\bullet :$ $G_{1,\bullet} \to G_{2,\beta(\bullet)}$ such that $\Phi_\bullet(H_{1,\bullet}) \subseteq H_{2,\beta(\bullet)}$ almost surely and a section $\rho : Y_1 \to G_{2,\beta(\bullet)}$ such that $\alpha = \beta \ltimes (L_{\rho(\bullet)} \circ \Phi_\bullet)|_{H_{2,\bullet}}^{H_{1,\bullet}}$, and then

$$
\sigma_2(\gamma, \beta(y)) = \rho(S_1^\gamma y) \cdot \Phi_y(\sigma_1(\gamma, y)) \cdot \rho(y)^{-1}
$$

for ν_1-almost all y and all $\gamma \in \Gamma$.

REMARK. When S_1 is ergodic this tells us that any extension of β to a factor map of $\mathbf{X}_1$ must take the form of fibrewise application of a *fixed* group automorphism and then left multiplication by a cocycle: this is the result of Mentzen [36].

Proof. We will deduce this by considering the joining of $\mathbf{X}_1$ and $\mathbf{X}_2$ defined by the graph of α, proceeding in two steps.

STEP 1. Suppose first that $H_{i,\bullet} \equiv \{1_{G_{i,\bullet}}\}$ for $i = 1, 2$. Set $\lambda :=$ $(\mathrm{id}_{X_1}, \alpha)_\# \mu_1$; this is a $(T_1 \times T_2)$-invariant probability measure on

$$
X_1 \times_{\{\beta \circ \xi_1 = \xi_2\}} X_2 \cong Y_1 \ltimes (G_{1,\bullet} \times G_{2,\beta(\bullet)}),
$$

which is an extension of $\mathbf{Y}_1$ via the natural factor map. Let us denote by π_i, $i = 1, 2$, the two coordinate projections $X_1 \times_{\{\beta \circ \xi_1 = \xi_2\}} X_2 \to X_i$, and, slightly abusively, also the coordinate projections $G_{1,\bullet} \times G_{2,\beta(\bullet)} \to G_{1,\bullet}, G_{2,\beta(\bullet)}$. Since this is a graph joining, the first coordinate projection almost surely determines the second, and so the joined system that results is actually isomorphic to $\mathbf{X}_1$; it follows that λ is $(T_1 \times T_2)$-relatively ergodic over $\mathbf{Y}_1$.

Applying Proposition 4.7 we obtain some S_1-invariant Mackey group data $M_\bullet \leq G_{1,\bullet} \times G_{2,\beta(\bullet)}$ on Y and a section $b : Y_1 \to G_{1,\bullet} \times G_{2,\beta(\bullet)}$ such that $\lambda = \nu \ltimes m_{b(\bullet)^{-1}M_\bullet}$. The measure λ must project onto μ_i under π_i, and so

$$
\pi_{1\#}(\nu \ltimes m_{b(\bullet)^{-1}M_\bullet}) = \nu \ltimes m_{\pi_1(b(\bullet))^{-1}\pi_1(M_\bullet)} = \nu \ltimes m_{G_{1,\bullet}}
$$

and hence $\pi_1(M_\bullet) = G_{1,\bullet}$ almost surely, and similarly $\pi_2(M_\bullet) = G_{2,\beta(\bullet)}$ almost surely.

On the other hand, if $L_{1,\bullet}$ and $L_{2,\bullet}$ are the first and second slices of $M_\bullet$, then we see that the coordinate factors π_1 and π_2 are actually relatively independent over the further canonical factors

$$(X_1, \mu_1, T_1) \to (Y_1, \nu_1, S_1) \ltimes (G_{1,\bullet}/L_{1,\bullet}, m_{G_{1,\bullet}/L_{1,\bullet}}, \sigma_1)$$

and

$$(Y_1, \nu_1, S_1) \otimes_{\{\beta = \xi_2\}} (X_2, \mu_2, T_2)$$
$$\to (Y_1, \nu_1, S_1) \ltimes (G_{2,\beta(\bullet)}/L_{2,\bullet}, m_{G_{2,\beta(\bullet)}/L_{2,\bullet}}, \sigma_2 \circ \beta).$$

Of course, under a graphical joining of a factor map such as λ the second coordinate projection is almost surely determined by the first, so we must have $L_{2,\bullet} \equiv \{1_{G_{2,\beta(\bullet)}}\}$ almost surely. Combined with the property of having full projections, this shows that $M_\bullet$ is almost surely the graph of a measurably-varying epimorphism $\Phi_\bullet : G_{1,\bullet} \to G_{2,\beta(\bullet)}$.

Also, since $M_\bullet$ has full one-dimensional projections we can multiply b by some $M_\bullet$-valued cocycle if necessary to assume that $b(\bullet) = (1_{G_{1,\bullet}}, \rho(\bullet))$. We can now simply read off that when $\alpha(y, g) = (\beta(y), g')$, then almost surely

$$(g, g') = (m, \rho(y)\Phi_y(m))$$

for some $m \in G_{1,y}$. Hence we must have $m = g$, and so $g' = \rho(y)\Phi_y(g)$: that is, $\alpha = \beta \ltimes (L_{\rho(\bullet)} \circ \Phi_\bullet)$, as required. Finally, with this expression in hand it is immediate to check that the commutative diagram relating α, β, S and T is equivalent to the requirement that

$$\rho(S^\gamma(y)) \cdot \Phi_y(\sigma_1(\gamma, y)) = \sigma_2(\gamma, \beta(y)) \cdot \rho(y)$$

for ν_1-almost every y and all $\gamma \in \Gamma$, which rearranges into the equation stated.

STEP 2. Now consider the case of general core-free $H_{i,\bullet}$ and let λ be the graphical self-joining given previously, so as in Step 1 the system $(X_1 \times_{\{\beta \circ \xi_1 = \xi_2\}} X_2, \lambda, T_1 \times T_2)$ has a natural factor isomorphic to $\mathbf{Y}_1$ and is isomorphic to $\mathbf{X}_1$ through the first coordinate projection, which also virtually determines the second coordinate projection.

Now in addition let $\tilde{\mathbf{X}}_i \to \mathbf{X}_i$ be the implied covering group data extensions, and let $\tilde{\lambda}$ be any relatively ergodic lift of λ to an invariant joining of $\tilde{\mathbf{X}}_1$ and $\tilde{\mathbf{X}}_2$. Arguing as in Step 1 now gives Mackey group data $M_\bullet \leq G_{1,\bullet} \times G_{2,\beta(\bullet)}$ and a section $b : Y_1 \to G_{1,\bullet} \times G_{2,\beta(\bullet)}$ such that

$$\tilde{\lambda} = \nu_1 \ltimes m_{b(\bullet)^{-1} M_\bullet (H_{1,\bullet} \times H_{2,\beta(\bullet)})}.$$

Since the cocycles σ_i are ergodic, it follows as before that $M_\bullet$ has full one-dimensional projections. The condition that the first coordinate almost surely determine the second under λ becomes more subtle. Firstly, it requires that the second slice $L_{2,\bullet}$ of $M_\bullet$ satisfy $L_{2,\bullet} \cdot H_{2,\beta(\bullet)} = H_{2,\beta(\bullet)}$; but on

the other hand Lemma 2.8 tells us that $L_{2,\bullet} \trianglelefteq G_{2,\beta(\bullet)}$, and by the core-free assumption $H_{2,\beta(\bullet)}$ does not contain any non-trivial normal subgroup, so in fact we must still have $L_{2,\bullet} \equiv \{1_{G_{2,\beta(\bullet)}}\}$. It follows that $M_\bullet$ still defines the graph of an epimorphism $\Phi_\bullet : G_{1,\bullet} \to G_{2,\beta(\bullet)}$. Secondly, this same condition on the determination of the second coordinate requires that $\Phi_\bullet$ have a well-defined quotient between the spaces $G_{1,\bullet}/H_{i,\bullet}$ and $G_{2,\beta(\bullet)}/H_{2,\beta(\bullet)}$, and hence that $\Phi_\bullet(H_{1,\bullet}) \subseteq H_{2,\beta(\bullet)}$ almost surely.

This leads as before to the expression $\beta \ltimes (L_{\rho(\bullet)} \circ \Phi_\bullet)|^{H_{1,\bullet}}_{H_{2,\beta(\bullet)}}$ for α, and the cocycle equation also follows as before from the condition that $\tilde{\lambda}$ is $(\tilde{T}_1 \times \tilde{T}_2)$-invariant (equivalent to the intertwining property in the previous case). ∎

Specializing the above now gives a structure theorem for groups of automorphisms of an extension.

THEOREM 6.5 (Relative Automorphism Structure Theorem). *Suppose that* $\mathbf{Y} = (Y, \nu, S)$ *is a* Γ-*system, that* $G_\bullet/H_\bullet$ *are* S-*invariant core-free homogeneous space data on* Y *and that* $\sigma : \Gamma \times Y \to G_\bullet$ *is an ergodic cocycle-section for the action* S, *and let* $\mathbf{X} = (X, \mu, T) := \mathbf{Y} \ltimes (G_\bullet/H_\bullet, m_{G_\bullet/H_\bullet}, \sigma)$. *Suppose further that* Λ *is a discrete group and* $R : \Lambda \curvearrowright (X, \mu)$ *is another action that commutes with* T *and respects the canonical factor* $\pi : \mathbf{X} \to \mathbf{Y}$ *(so it defines an action of* Λ *by automorphisms of the extension* π). *Then for each* $h \in \Lambda$ *there are an* S-*invariant measurable family of isomorphisms* $\Phi_{h,\bullet} : G_\bullet \to G_{R|^h_\pi(\bullet)}$ *such that* $\Phi_{h,\bullet}(H_\bullet) = H_{R|^h_\pi(\bullet)}$ *almost surely and a section* $\rho_h : Y \to G_{R|^h_\pi(\bullet)}$ *such that*

$$R^h = R|^h_\pi \ltimes (L_{\rho_h(\bullet)} \circ \Phi_{h,\bullet})|^{H_\bullet}_{H_{R|^h_\pi(\bullet)}}$$

for each $h \in \Lambda$, *and then*

- *we have*

$$\sigma(\gamma, R|^h_\pi(y)) = \rho_h(S^\gamma y) \cdot \Phi_{h,y}(\sigma(\gamma, y)) \cdot \rho_h(y)^{-1}$$

  for ν-*almost all* y *and all* $\gamma \in \Gamma$ *and* $h \in \Lambda$, *and*
- *we have*

$$\Phi_{h_1 h_2, y} = \Phi_{h_1, R|^{h_2}_\pi(y)} \circ \Phi_{h_2, y}$$

  and

$$\rho_{h_1 h_2}(y) = \rho_{h_1}(R|^{h_2}_\pi(y)) \cdot \Phi_{h_1, R|^{h_2}_\pi(y)}(\rho_{h_2}(y))$$

  for ν-*almost all* y *and all* $h_1, h_2 \in \Lambda$.

Proof. Consider first R^h for some fixed $h \in \Lambda$. Treating R^h as a factor map of $\mathbf{X}$ and applying Theorem 6.4 gives immediately the representation of R^h as

$$R|^h_\pi \ltimes (L_{\rho_h(\bullet)} \circ \Phi_{h,\bullet})|^{H_\bullet}_{H_{R|^h_\pi(\bullet)}}$$

for some section ρ_h and measurable family of continuous epimorphisms $\Phi_{h,\bullet}$, and now the condition that R^h actually be equivalent to id_X (that is, it is not a proper factor map) gives that $\Phi_{h,\bullet}$ is an isomorphism and $\Phi_\bullet(H_\bullet) = H_{R|_\pi(\bullet)}$ almost surely.

This establishes the existence of ρ_h and $\Phi_{h,\bullet}$, and also the first of the two additional conclusions above. To deduce the second we need only compare the resulting coordinatizations of each side of the equation $R^{h_1 h_2} = R^{h_1} \circ R^{h_2}$ defining R as a Λ-action: substituting from above this becomes

$$R|_\pi^{h_1 h_2} \ltimes (L_{\rho_{h_1 h_2}(\bullet)} \circ \Phi_{h_1 h_2,\bullet})|_{H_{R|_\pi^{h_1 h_2}(\bullet)}}^{H_\bullet}$$

$$= (R|_\pi^{h_1} \circ R|_\pi^{h_2}) \ltimes (L_{\rho_{h_1}(R|_\pi^{h_2}(\bullet))} \circ \Phi_{h_1, R|_\pi^{h_2}(\bullet)} \circ L_{\rho_{h_2}(\bullet)} \circ \Phi_{h_2,\bullet})|_{H_{R|_\pi^{h_1 h_2}(\bullet)}}^{H_\bullet},$$

and so we must have

$$\rho_{h_1 h_2}(y) \cdot \Phi_{h_1 h_2,y}(g) \cdot H_{R|_\pi^{h_1 h_2}(y)} = \rho_{h_1 h_2}(y) \cdot \Phi_{h_1 h_2,y}(g H_y)$$

$$= \rho_{h_1}(R|_\pi^{h_2}(y)) \cdot \Phi_{h_1, R|_\pi^{h_2}(y)}(\rho_{h_2}(y) \cdot \Phi_{h_2,y}(g H_y))$$

$$= \rho_{h_1}(R|_\pi^{h_2}(y)) \cdot \Phi_{h_1, R|_\pi^{h_2}(y)}(\rho_{h_2}(y)) \cdot \Phi_{h_1, R|_\pi^{h_2}(y)}(\Phi_{h_2,y}(g)) \cdot H_{R|_\pi^{h_1 h_2}(y)}$$

for all $g \in G_y$ for ν-almost every $y \in Y$. Since $H_{R|_\pi^{h_1 h_2}(y)}$ is core-free in G_y almost surely, the validity of this equation for all $g \in G_y$ implies the two parts of the second additional conclusion above, completing the proof. $\blacksquare$

REMARK. It should be possible to enhance the above theorem further by allowing an arbitrary locally compact second countable group Λ and imposing suitable continuity assumptions on the assignments $h \mapsto \rho_h$ of measurable sections and $h \mapsto \Phi_{h,\bullet}$ of measurable families of isomorphisms. The additional arguments required seem to be more fiddly than enlightening, however, and so we leave the details to the interested reader.

Although Theorem 6.5 shows that the S-ergodic fibre systems above the points $\zeta_0^S(y) \in Z_0^S$ and $\zeta_0^S(R|_\pi^h y) \in Z_0^S$ are isomorphic for all $h \in \Lambda$ for almost every $y \in Y$, it need not follow that these fibre systems are almost all isomorphic to a single model system. The following simple example has long been a part of ergodic-theoretic folklore.

EXAMPLE. Let $(Y, \nu) := (\mathbb{T}^2, m_{\mathbb{T}^2})$, and form the direct integral space $X := Y \ltimes \mathbb{T}^2$ with constant fibre $\mathbb{T}^2$ (so this is really just the direct product $Y \times \mathbb{T}^2$) and measure $\mu := m_{\mathbb{T}^2} \otimes m_{\mathbb{T}^2}$. Define $T : \mathbb{Z} \curvearrowright X$ by $T(y, z) := (y, y + z)$, so overall (X, μ) is the direct integral of the individual Kronecker systems $(\mathbb{T}^2, m_{\mathbb{T}^2}, R_y)$, writing R_y for the rotation by $y \in \mathbb{T}^2$.

In addition, suppose that $S \curvearrowright \mathbb{T}^2$ is any ergodic toral automorphism. Then it is easy to check that $S \times S$ commutes with T; in particular, it carries fibres of the obvious factor map $X \to Y$ onto fibres, and so acts

as an automorphism of the fibre system $(\mathbb{T}^2, m_{\mathbb{T}^2}, R_y)$ onto $(\mathbb{T}^2, m_{\mathbb{T}^2}, R_{Sy})$. However, the fibre systems $(\mathbb{T}^2, m_{\mathbb{T}^2}, R_y)$ are not almost all isomorphic for different y: the map

$$y \mapsto (\mathbb{T}^2, m_{\mathbb{T}^2}, R_y)/\sim_{\text{Isomorphism}}$$

is not almost surely constant, even though it is invariant under the ergodic transformation S. This is possible because the isomorphism equivalence relation on the space of all Kronecker systems (suitably interpreted as pairs comprising a monothetic compact metrizable subgroup of a suitable fixed repository and a distinguished element for the rotation) is non-smooth.

QUESTION 6.6. *Can an example be found for which the group fibres G_y themselves are not almost all continuously isomorphic above each $R|_\pi$-ergodic component of ν?*

This may relate to the work of Conze and Raugi [15] on the behaviour of measurable families of (not necessarily compact) groups related by measurable cocycles, but we have not been able to answer the above as a direct corollary of their work.

The following question may also be related to the above:

QUESTION 6.7. *Can an example be found in which for* no *coordinatization of the extension is it possible that each Φ_y can be extended from G_y to an automorphism of the whole repository group U?*

The following corollaries concerning the extendability of automorphisms will also prove useful later, and may be of some independent interest.

COROLLARY 6.8 (Condition for lifting an automorphism to a group data extension). *An action R of Λ by automorphisms of $\mathbf{Y}$ can be lifted to a Λ-action by automorphisms of an ergodic group data extension $\mathbf{Y} \ltimes (G_\bullet, m_{G_\bullet}, \sigma)$ if and only if for every $h \in \Lambda$ the cocycle $\Gamma \times Y \to G_\bullet \times G_{R(\bullet)}$: $(\gamma, y) \mapsto (\sigma(\gamma, y), \sigma(\gamma, R^h y))$ has relativized Mackey group data over $\mathbf{Y}$ that is the graph of an isomorphism almost everywhere, and in this case any such extended action is of the form $h \mapsto R^h \ltimes (L_{\rho_h(\bullet)} \circ \Phi_{h,\bullet})$ for some families of sections $\rho_h : Y \to G_\bullet$ and S-invariant cocycles $\Phi_{h,\bullet} : Y \to \mathrm{Isom}(G_\bullet, G_{R^h(\bullet)})$ and Φ_h is unique up to composition with an arbitrary S-invariant inner automorphism cocycle.* ∎

COROLLARY 6.9 (Automorphisms can always be lifted to core-free ergodic covering group extensions). *Suppose that $\mathbf{Y}$ is a Γ-system, $G_\bullet/H_\bullet$ are S-invariant core-free homogeneous space data and $\sigma : \Gamma \times Y \to G_\bullet$ is an ergodic cocycle-section. Set $\mathbf{X} := \mathbf{Y} \ltimes (G_\bullet/H_\bullet, m_{G_\bullet/H_\bullet}, \sigma)$ and $\tilde{\mathbf{X}} := \mathbf{X} \ltimes (G_\bullet, m_{G_\bullet}, \sigma)$. Then any action of a discrete group by automorphisms of the canonical extension $\mathbf{X} \to \mathbf{Y}$ lifts to an action by automorphisms of the tower $\tilde{\mathbf{X}} \to \mathbf{X} \to \mathbf{Y}$.*

Proof. Let $\pi : \mathbf{X} \to \mathbf{Y}$ and $\tilde{\pi} : \tilde{\mathbf{X}} \to \mathbf{Y}$ be the canonical factor maps and suppose that R is an automorphism of the extension $\pi : \mathbf{X} \to \mathbf{Y}$. Theorem 6.5 allows us to write R explicitly as $R|_\pi \ltimes (L_{\rho(\bullet)} \circ \Phi_\bullet)|^{H_\bullet}_{H_{R|_\pi(\bullet)}}$ for some $\rho : Y \to G_{R|_\pi(\bullet)}$ and S-invariant $\Phi : Y \to \text{Isom}(G_\bullet, G_{R|_\pi(\bullet)})$ such that $\Phi_y(H_y) = H_{R|_\pi(y)}$ almost surely; and, having done this, we have that graph$(\Phi_\bullet)$ is the Mackey group data of the cocycle-section $(\gamma, y) \mapsto (\sigma(\gamma, y), \sigma(\gamma, Ry))$, so that

$$\sigma(\gamma, R(y)) = b(S^\gamma y) \cdot \Phi_y(\sigma(\gamma, y)) \cdot b(y)^{-1}$$

almost surely. This equation immediately tells us that we can lift R to the transformation $\tilde{R} := R|_\pi \ltimes (L_{\rho(\bullet)} \circ \Phi_\bullet)$ on $\tilde{X}$, and that this still commutes with $\tilde{T} = S \ltimes \sigma$. Given a whole Λ-action of automorphisms R^h, applying this argument to each $h \in \Lambda$ individually and considering the consistency equations promised by Theorem 6.5 shows that the lifted maps still define a Λ-action, and hence completes the proof. ∎

7. Applications.

In this section we offer two closely related applications of the theory developed above.

We first study the possible joint distribution of the isotropy factors $\zeta_0^{T_i}$ corresponding to three commuting transformations T_1, T_2 and T_3. This will require some quite careful analysis in terms of Mackey group data, cocycles, and representations given by the Relative Automorphism Structure Theorem. We will then show that this analysis can also be brought to bear on a detailed description of characteristic factors of the double non-conventional ergodic averages associated to a pair of commuting transformations (see, for example, [6] and the references listed there).

Throughout this section we specialize to the setting of $\Gamma := \mathbb{Z}^d$, and will write $\mathbf{e}_1, \ldots, \mathbf{e}_d$ for its standard basis.

7.1. Application to joint distributions of isotropy factors.

For a generic $\mathbb{Z}^d$-action on a fixed atomless (X, μ) the isotropy factors $\zeta_0^{T^{\restriction \Lambda}} : \mathbf{X} \to \mathbf{Z}_0^{T^{\restriction \Lambda}}$ corresponding to subgroups $\Lambda \leq \mathbb{Z}^d$ are all trivial (indeed, it is a classical result that a generic such action is totally weakly mixing). However, if they are not all trivial then they generate a sublattice of the lattice of all factors of (X, μ, T) that can exhibit some quite rich structure.

Letting $T_i := T^{\mathbf{e}_i}$, we will here consider only the further sublattice generated by the isotropy factors $\zeta_0^{T_I} := \zeta_0^{T_{i_1}, \ldots, T_{i_r}}$ corresponding to the possible choices of subset $I := \{i_1, \ldots, i_r\} \subseteq [d]$, where $[d] := \{1, \ldots, d\}$.

Clearly in general the action of each T_j for $j \in [d] \setminus I$ on the sets of $\Sigma_X^{T_I}$ can still be quite arbitrary, and so we cannot hope to say anything about the structure of each isotropy factor as a system in its own right. Instead we will focus on their *joint* distribution within the original system.

EXAMPLE. Let (X, μ, T_1, T_2) be the $\mathbb{Z}^2$-system $(\mathbb{T}^2, \text{Haar}, R_{(\alpha,0)}, R_{(0,\alpha)})$, where R_q denotes the rotation of the compact Abelian group $\mathbb{T}^2$ by an element $q \in \mathbb{T}^2$ and we choose $\alpha \in \mathbb{T}$ irrational. In this case we have natural coordinatizations

$$\zeta_0^{T_i} : X \to \mathbb{T} : (t_1, t_2) \mapsto t_{3-i},$$

and similarly, since $T_1 T_2 = R_{(\alpha,\alpha)}$,

$$\zeta_0^{T_1 T_2} : X \to \mathbb{T} : (t_1, t_2) \mapsto t_1 - t_2.$$

It follows that in this example any two of $\zeta_0^{T_1}$, $\zeta_0^{T_2}$ and $\zeta_0^{T_1 T_2}$ are independent, but also that any two of them generate the whole system (and so overall independence fails).

In this section we will employ the general machinery of non-ergodic isometric extensions and the non-ergodic Furstenberg–Zimmer and Mackey theories to describe this joint distribution in the case $d = 3$. It will turn out that these factors are always relatively independent outside certain special 'obstruction' factors, which are in turn only a little more general than the above example.

THEOREM 7.1. *Suppose that* $T_i : \mathbb{Z} \curvearrowright (X, \mu)$, $i = 1, 2, 3$, *are three commuting actions. Then*

(1) *The triple of factors* $\zeta_0^{T_1,T_2}$, $\zeta_0^{T_1,T_3}$, $\zeta_0^{T_2,T_3}$ *is relatively independent over* ζ_0^{T}.

(2) *The triple of factors* $\zeta_0^{T_1}$, $\zeta_0^{T_2}$, $\zeta_0^{T_3}$ *is relatively independent over the further triple of factors*

$$\zeta_0^{T_1} \wedge (\zeta_0^{T_2} \vee \zeta_0^{T_3}), \quad \zeta_0^{T_2} \wedge (\zeta_0^{T_3} \vee \zeta_0^{T_1}), \quad \zeta_0^{T_3} \wedge (\zeta_0^{T_1} \vee \zeta_0^{T_2}).$$

THEOREM 7.2. *We have*

$$\zeta_0^{T_3} \wedge (\zeta_0^{T_1} \vee \zeta_0^{T_2}) \succsim \zeta_0^{T_1,T_3} \vee \zeta_0^{T_2,T_3},$$

and the extension of systems

$$(\zeta_0^{T_1,T_3} \vee \zeta_0^{T_2,T_3})|_{\zeta_0^{T_3} \wedge (\zeta_0^{T_1} \vee \zeta_0^{T_2})} : (\zeta_0^{T_3} \wedge (\zeta_0^{T_1} \vee \zeta_0^{T_2}))(\mathbf{X}) \to \zeta_0^{T_1,T_3} \vee \zeta_0^{T_2,T_3}(\mathbf{X})$$

can be coordinatized as the group extension

$$(\zeta_0^{T_1,T_3} \vee \zeta_0^{T_2,T_3})(\mathbf{X}) \ltimes (G_{3,\bullet}, m_{G_{3,\bullet}}, (\tau_{3,1} \circ \zeta_0^{T_2,T_3}), (\tau_{3,2} \circ \zeta_0^{T_1,T_3})^{\text{op}}, 1)$$

$$\downarrow \text{canonical}$$

$$\zeta_0^{T_1,T_3} \vee \zeta_0^{T_2,T_3}(\mathbf{X})$$

for some $T|_{\zeta_0^{T_1,T_3} \vee \zeta_0^{T_2,T_3}}$-*invariant compact group data* $G_{3,\bullet}$ *and cocycle-sections* $\tau_{3,1} : Z_0^{T_2,T_3} \to G_{3,\bullet}$ *and* $\tau_{3,2} : Z_0^{T_1,T_3} \to G_{3,\bullet}$ *and similarly for the extension* $(\zeta_0^{T_i,T_j} \vee \zeta_0^{T_i,T_k})|_{\zeta_0^{T_i} \wedge (\zeta_0^{T_j} \vee \zeta_0^{T_k})}$ *for any other permutation* i, j, k *of*

the indices $1, 2, 3$ (*in general with different group data* $G_{i,\bullet}$ *and cocycle-sections* $\tau_{i,j}$).

These two results together amount to Theorem 1.1 of the Introduction. Heuristically, they assert that the joint distribution of the factor-triple $(\zeta_0^{T_i})_{i=1}^3$ fails to be relatively independent over the natural candidate factor-triple $(\zeta_0^{T_i,T_j} \vee \zeta_0^{T_i,T_k})_{i=1}^3$ only up to single isometric extensions, and give fairly explicit coordinatizations of those extensions.

It seems likely that these results could be extended with only routine modifications to treat a triple of commuting actions $T_i : \Gamma_i \curvearrowright (X, \mu)$ of other locally compact second countable groups Γ_i (the key feature being that the different actions commute). We have restricted to a triple of $\mathbb{Z}$-actions for notational simplicity.

On the other hand, although we naturally expect Theorem 1.1 to be a special case of a result for larger numbers of commuting transformations (or actions), the analysis of the corresponding isotropy factors based on the Furstenberg–Zimmer and Mackey theories becomes quickly much more complicated, and we shall not pursue this generalization any further at present.

Observe also that while Theorem 7.2 describes the structure of each system $(\zeta_0^{T_i} \wedge (\zeta_0^{T_j} \vee \zeta_0^{T_k}))(\mathbf{X})$ as an extension of $\zeta_0^{T_i,T_j} \vee \zeta_0^{T_i,T_k}(\mathbf{X})$ (which, by Theorem 7.1, is itself just a relatively independent joining of $\mathbf{Z}_0^{T_i,T_j}$ and $\mathbf{Z}_0^{T_i,T_k}$ over $\mathbf{Z}_0^T$), it does not describe the joint distribution of the factor maps $\zeta_0^{T_i} \wedge (\zeta_0^{T_j} \vee \zeta_0^{T_k})$. This would require a further analysis, using the relative independence of the isotropy factors $\mathbf{Z}_0^{T_i,T_j}$ over $\mathbf{Z}_0^T$ to understand first the joint distribution of the $\zeta_0^{T_i,T_j} \vee \zeta_0^{T_i,T_k}(\mathbf{X})$ and then working upwards, and would proceed using very similar ideas to those below but with relatively smaller returns; we omit the details.

Our basic approach rests on an appeal to the Furstenberg–Zimmer inverse theory to reduce the problem to the study of certain isometric extensions, followed by a detailed analysis of the possible structure of an associated Mackey group to obtain finer information about these extensions.

This strategy is already well-established in the literature from studies of other questions working under more restrictive ergodicity assumptions. Indeed, Furstenberg's original paper [22] developing an ergodic-theoretic approach to Szemerédi's Theorem, for which much of the above-mentioned machinery was originally developed, uses similar ideas to analyze the structure of a certain self-joining of a given ergodic $\mathbb{Z}$-system en route to the proof of multiple recurrence. That paper has since led to a considerably more detailed study of the 'non-conventional ergodic averages' that appear in this connexion, which we will revisit in the next section [12, 13, 14, 45, 27, 31, 47].

In addition, Rudolph has given in [38] an analysis of a different question rather more closely related to the study of isotropy factors: he obtains a description of the possible eigenfunctions of the product system $S \times T \curvearrowright (Y \times X, \nu \otimes \mu)$ built from ergodic transformations $S \curvearrowright (Y, \nu)$ and $T \curvearrowright (X, \mu)$, effectively using for the earlier stages of his work a special case of the analysis to be given below applied to this latter product system.

We depart from these previous works in our use of the non-ergodic versions of the basic machinery. In this more general setting we will find that the resulting structures are considerably more complex, even though the description of the extensions ultimately obtained in Theorem 7.2 involves only data $G_{i,\bullet}$ that are invariant for the whole action T. In particular, many of these older works have ultimately reduced their subjects to the study of factors that lie in a very special class of systems, the 'pronilsystems' (see, in particular, Rudolph's work [38] and the papers of Host and Kra [31] and Ziegler [47] on non-conventional averages). Already in the cases considered in this section we find that we must look beyond that class.

7.2. Reduction to compositions of isotropy factors. We first prove the (rather simpler) Theorem 7.1, by effecting a quite general reduction of the problem to the study of certain composed isotropy factors.

LEMMA 7.3. *Suppose that* $T : \Gamma \curvearrowright (X, \mu)$ *is a probability-preserving action of a locally compact second countable amenable group Γ and that $\pi : (X, \mu, T) \to (Y, \nu, S)$ is a factor. Then ζ_0^T and π are relatively independent over $\zeta_0^S \circ \pi = \pi|_{\zeta_0^T} \circ \zeta_0^T$.*

Proof. Let $(I_N)_{N \geq 1}$ be a left Følner sequence in Γ. If $A \in \pi^{-1}(\Sigma_Y)$ and $B \in \Sigma_X$ is T-invariant, then

$$\mu(A \cap B) = \lim_{N \to \infty} \int_X \left(\frac{1}{m_\Gamma(I_N)} \int_{I_N} 1_{T^\gamma(A)} \, m_\Gamma(\mathrm{d}\gamma) \right) \cdot 1_B \, \mathrm{d}\mu$$

$$= \int_X \mathsf{E}_\mu(1_A \,|\, \zeta_0^S \circ \pi) \cdot 1_B \, \mathrm{d}\mu = \int_X \mathsf{E}_\mu(1_A \,|\, \zeta_0^S \circ \pi) \cdot \mathsf{E}_\mu(1_B \,|\, \zeta_0^S \circ \pi) \, \mathrm{d}\mu,$$

where the middle equality follows from the mean ergodic theorem. ∎

Proof of Theorem 7.1. (1) If $A \in \Sigma_X$ is T_1- and T_2-invariant, $B \in \Sigma_X$ is T_1- and T_3-invariant and $C \in \Sigma_X$ is T_2- and T_3-invariant then averaging first under T_3 gives

$$\int_X 1_A \cdot 1_B \cdot 1_C \, \mathrm{d}\mu = \int_X \mathsf{E}_\mu(1_A \,|\, \zeta_0^{T_3}) \cdot 1_B \cdot 1_C \, \mathrm{d}\mu = \int_X \mathsf{E}_\mu(1_A \,|\, \zeta_0^{T_1,T_2,T_3}) \cdot 1_B \cdot 1_C \, \mathrm{d}\mu,$$

and now averaging under T_2 gives

$$\int_X \mathsf{E}_\mu(1_A \,|\, \zeta_0^{T_1,T_2,T_3}) \cdot 1_B \cdot 1_C \,\mathrm{d}\mu$$

$$= \int_X \mathsf{E}_\mu(1_A \,|\, \zeta^{T_1,T_2,T_3}) \cdot \mathsf{E}_\mu(1_B \,|\, \zeta_0^{T_2}) \cdot 1_C \,\mathrm{d}\mu$$

$$= \int_X \mathsf{E}_\mu(1_A \,|\, \zeta_0^{T_1,T_2,T_3}) \cdot \mathsf{E}_\mu(1_B \,|\, \zeta_0^{T_1,T_2,T_3}) \cdot 1_C \,\mathrm{d}\mu$$

$$= \int_X \mathsf{E}_\mu(1_A \,|\, \zeta_0^{T_1,T_2,T_3}) \cdot \mathsf{E}_\mu(1_B \,|\, \zeta_0^{T_1,T_2,T_3}) \cdot \mathsf{E}_\mu(1_C \,|\, \zeta_0^{T_1,T_2,T_3}) \,\mathrm{d}\mu;$$

concatenating these equalities gives the result.

(2) This follows similarly. For this proof let $\psi_i := \zeta_0^{T_i} \wedge (\zeta_0^{T_j} \vee \zeta_0^{T_k})$. If $A \in \Sigma_X$ is T_1-invariant, $B \in \Sigma_X$ is T_2-invariant and $C \in \Sigma_X$ is T_3-invariant then Lemma 7.3 applied to the action T_i and the factors $\zeta_0^{T_i}$ and $\pi := \zeta_0^{T_j} \vee \zeta_0^{T_k}$ gives that these are relatively independent over ψ_i, and hence that

$$\int_X 1_A \cdot 1_B \cdot 1_C \,\mathrm{d}\mu = \int_X \mathsf{E}_\mu(1_A \,|\, \psi_1) \cdot 1_B \cdot 1_C \,\mathrm{d}\mu$$

$$= \int_X \mathsf{E}_\mu(1_A \,|\, \psi_1) \cdot \mathsf{E}_\mu(1_B \,|\, \psi_2) \cdot 1_C \,\mathrm{d}\mu$$

$$= \int_X \mathsf{E}_\mu(1_A \,|\, \psi_1) \cdot \mathsf{E}_\mu(1_B \,|\, \psi_3) \cdot \mathsf{E}_\mu(1_C \,|\, \psi_3) \,\mathrm{d}\mu,$$
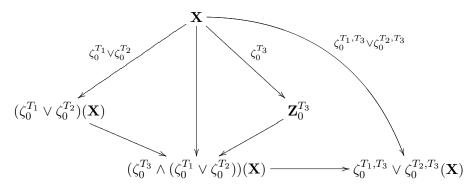
as required. ∎

7.3. Some isometric extensions and their associated Mackey data. To prove Theorem 7.2 (and so complete the proof of Theorem 1.1) we need to understand the structure of the composite factors $\zeta_0^{T_i} \wedge (\zeta_0^{T_j} \vee \zeta_0^{T_k})$ as extensions of $\zeta_0^{T_i,T_j} \vee \zeta_0^{T_i,T_k}$; most of our work will go into this. As in the statement of the theorem we will treat the case $(i,j,k) = (3,1,2)$, the others being analogous. We will first obtain some isometricity for the extensions

$$(\zeta_0^{T_1,T_3} \vee \zeta_0^{T_2,T_3})|_{\zeta_0^{T_3} \wedge (\zeta_0^{T_1} \vee \zeta_0^{T_2})} : (\zeta_0^{T_3} \wedge (\zeta_0^{T_1} \vee \zeta_0^{T_2}))(\mathbf{X}) \to \zeta_0^{T_1,T_3} \vee \zeta_0^{T_2,T_3}(\mathbf{X}),$$

and will then see a gradual extraction of finer and finer properties of these isometric extensions from an analysis of the associated Mackey data, with an occasional recoordinatization of the extensions where necessary.

The various isotropy factors stand related as in the following commutative diagram (where some of the obvious maps have not been named):

$$\mathbf{X}$$

with arrows labeled:

$\zeta_0^{T_1} \vee \zeta_0^{T_2}$

$\zeta_0^{T_3}$

$\zeta_0^{T_1,T_3} \vee \zeta_0^{T_2,T_3}$

$$(\zeta_0^{T_1} \vee \zeta_0^{T_2})(\mathbf{X}) \qquad \mathbf{Z}_0^{T_3}$$

$$(\zeta_0^{T_3} \wedge (\zeta_0^{T_1} \vee \zeta_0^{T_2}))(\mathbf{X}) \longrightarrow \zeta_0^{T_1,T_3} \vee \zeta_0^{T_2,T_3}(\mathbf{X})$$

It will prove helpful to introduce some more notation. For $i = 1, 2$ let $\alpha_i := \zeta_0^{T_i,T_3} \vee \zeta_0^{T_1,T_2}$ and $\mathbf{W}_i = (W_i, (\alpha_i)_\# \mu, T|_{\alpha_i})$ be its target system.

LEMMA 7.4. *We have*

$$\zeta_0^{T_3} \wedge (\zeta_0^{T_1} \vee \zeta_0^{T_2}) \precsim (\zeta_{1/\alpha_1}^{T_3} \wedge \zeta_0^{T_1}) \vee (\zeta_{1/\alpha_2}^{T_3} \wedge \zeta_0^{T_2}).$$

Proof. By Lemma 7.3 the system $(\zeta_0^{T_1} \vee \zeta_0^{T_2})(\mathbf{X})$ is a relatively independent joining of $\mathbf{Z}_0^{T_1}$ and $\mathbf{Z}_0^{T_2}$ over their further factors $\zeta_0^{T_1,T_2}|_{\zeta_0^{T_1}}$ and $\zeta_0^{T_1,T_2}|_{\zeta_0^{T_2}}$. On the other hand the $T_3|_{\zeta_0^{T_1} \vee \zeta_0^{T_2}}$-invariant functions on $(\zeta_0^{T_1} \vee \zeta_0^{T_2})(\mathbf{X})$ are all virtually measurable with respect to the maximal subextension of

$$\zeta_0^{T_1,T_2}|_{\zeta_0^{T_1} \vee \zeta_0^{T_2}} : (\zeta_0^{T_1} \vee \zeta_0^{T_2})(\mathbf{X}) \to \mathbf{Z}_0^{T_1,T_2}$$
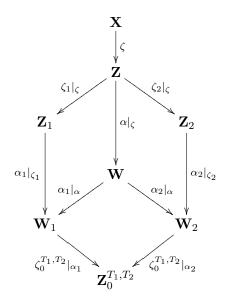
that is isometric for the restricted action of T_3, and so Theorem 5.12 implies that this in turn is contained in $(\zeta_{1/\alpha_1}^{T_3} \wedge \zeta_0^{T_1}) \vee (\zeta_{1/\alpha_2}^{T_3} \wedge \zeta_0^{T_2})$, as required. ∎

At this point we will introduce some new notation for the basic systems and factor maps under study. In addition to lightening the presentation, this will make our main technical results simultaneously relevant to this and the next section and so minimize the duplication of effort.

We have defined $\alpha_i := \zeta_0^{T_i,T_3} \vee \zeta_0^{T_1,T_2}$ with target $\mathbf{W}_i$ above. We define also $\zeta_i := \zeta_{1/\alpha_i}^{T_3} \wedge \zeta_0^{T_i}$ and let $\mathbf{Z}_i$ be its target system, and we let $\mathbf{Z}$ be the target of $\zeta := \zeta_1 \vee \zeta_2$ (a joining of $\mathbf{Z}_1$ and $\mathbf{Z}_2$) and $\mathbf{W}$ be the target of $\alpha := \alpha_1 \vee \alpha_2$ (a joining of $\mathbf{W}_1$ and $\mathbf{W}_2$). As usual the choice of these target systems is arbitrary up to isomorphism, but in this case it is natural (and notationally convenient) to pick $\mathbf{W}$ to be

$$(W_1 \times W_2, (\alpha_1 \vee \alpha_2)_\# \mu, T|_{\alpha_1} \times T_{\alpha_2}),$$

since we will often want to discuss separately the two coordinates of a point $(w_1, w_2) \in W$. These factors are now arranged as shown:

$$
\begin{array}{c}
\mathbf{X}
\end{array}
$$

where $\mathbf{W}$ and $\mathbf{Z}$ are actually generated by all of their exhibited factors, and the factors on left- and right-hand sides are relatively independent over their factor maps to $\mathbf{Z}_0^{T_1,T_2}$.

In this picture the transformations $T_i = T^{\mathbf{e}_i}$ have the following properties:

- T_i restricts to the identity on $\mathbf{Z}_i$ and the factors beneath it, while acting relatively ergodically on the extension $\zeta_0^{T_1,T_2}|_{\zeta_{3-i}} : \mathbf{Z}_{3-i} \to \mathbf{Z}_0^{T_1,T_2}$, for $i = 1, 2$;
- the extensions $\zeta_0^{T_1,T_2}|_{\alpha_i} : \mathbf{W}_i \to \mathbf{Z}_0^{T_1,T_2}$ are relatively invariant for the restriction of T_3, and the extensions $\alpha_i|_{\zeta_i} : \mathbf{Z}_i \to \mathbf{W}_i$ are relatively ergodic and isometric for the restriction of T_3.

Our goal is to identify the T_3-invariant factor of $\mathbf{Z}$ (which we know is also the overall T_3-invariant factor by the above lemma).

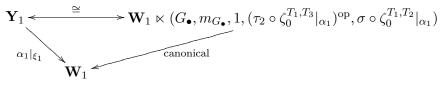In these terms we can now state our main technical result.

PROPOSITION 7.5. *In the situation described above, there are interme-diate factors*

$$
\mathbf{Z}_i \xrightarrow{\xi_i|_{\zeta_i}} \mathbf{Y}_i \xrightarrow{\alpha_i|_{\xi_i}} \mathbf{W}_i
$$

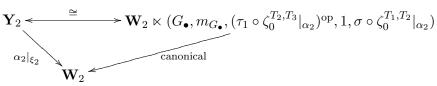*factorizing $\alpha_i|_{\zeta_i}$ such that there are T-invariant compact group data $G_\bullet$ and cocycle-sections*

$\sigma : Z_0^{T_1,T_2} \to G_\bullet$ *that is $T_3|_{\zeta_0^{T_1,T_2}}$-relatively ergodic,*

$\tau_1 : Z_0^{T_2,T_3} \to G_\bullet$ *that is $T_1|_{\zeta_0^{T_2,T_3}}$-relatively ergodic and*

$\tau_2 : Z_0^{T_2,T_3} \to G_\bullet$ *that is $T_2|_{\zeta_0^{T_1,T_3}}$-relatively ergodic*

so that we can coordinatize

$$\mathbf{Y}_1 \xleftarrow{\quad\cong\quad} \mathbf{W}_1 \ltimes (G_\bullet, m_{G_\bullet}, 1, (\tau_2 \circ \zeta_0^{T_1,T_3}|_{\alpha_1})^{\mathrm{op}}, \sigma \circ \zeta_0^{T_1,T_2}|_{\alpha_1})$$

with $\alpha_1|_{\xi_1}$ down to $\mathbf{W}_1$ and *canonical* to $\mathbf{W}_1$.

and

$$\mathbf{Y}_2 \xleftarrow{\quad\cong\quad} \mathbf{W}_2 \ltimes (G_\bullet, m_{G_\bullet}, (\tau_1 \circ \zeta_0^{T_2,T_3}|_{\alpha_2})^{\mathrm{op}}, 1, \sigma \circ \zeta_0^{T_1,T_2}|_{\alpha_2})$$

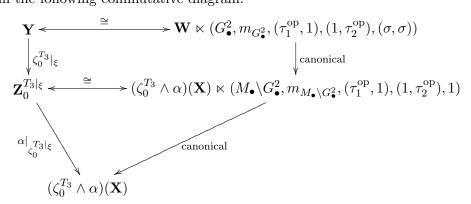with $\alpha_2|_{\xi_2}$ down to $\mathbf{W}_2$ and *canonical* to $\mathbf{W}_2$.

and such that the T_3-invariant factor $\zeta_0^{T_3} \wedge (\zeta_0^{T_1} \vee \zeta_0^{T_2})$ is contained in $\xi_1 \vee \xi_2$.
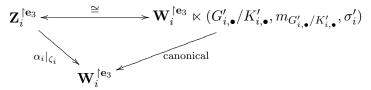
Proof of Theorem 7.2 from Proposition 7.5. This now follows simply by unpacking the new notation. Let $\xi := \xi_1 \vee \xi_2$ with target $\mathbf{Y}$ (a joining of $\mathbf{Y}_1$ and $\mathbf{Y}_2$). We know that ξ_1 and ξ_2 (as factors of $\zeta_0^{T_1}$ and $\zeta_0^{T_2}$) are relatively independent over their further factors $\zeta_0^{T_1,T_2}|_{\xi_i}$, $i = 1, 2$, and hence certainly over the intermediate factors α_1 and α_2, and so the coordinatizations of the extensions $\alpha_i|_{\xi_i} : \mathbf{Y}_i \to \mathbf{W}_i$ by group data given by Proposition 7.5 combine to give a coordinatization of $\mathbf{Y}$ by the group data $G_\bullet^2$ and the combined cocycles. We now observe that the restriction of T_3 to this group data extension is described by the diagonal cocycle-section (σ, σ) corresponding to the $T_3|_{\zeta_0^{T_1,T_2}}$-ergodic cocycle-section σ, and so we can simply deduce that the Mackey group data can be taken to be the diagonal subgroup $M_\bullet \cong \{(g, g) : g \in G_\bullet\}$, and now the associated Mackey section is trivial by symmetry. This leads to the coordinatization of the T_3-invariant factor $\zeta_0^{T_3} \wedge (\zeta_0^{T_1} \vee \zeta_0^{T_2}) = \zeta_0^{T_3} \wedge \zeta = \zeta_0^{T_3} \wedge \xi$ as given by the location of $\mathbf{Z}_0^{T_3|\xi}$ in the following commutative diagram:

$$\mathbf{Y} \xleftarrow{\quad\cong\quad} \mathbf{W} \ltimes (G_\bullet^2, m_{G_\bullet^2}, (\tau_1^{\mathrm{op}}, 1), (1, \tau_2^{\mathrm{op}}), (\sigma, \sigma))$$

with $\zeta_0^{T_3}|_\xi$ down and *canonical* down, to

$$\mathbf{Z}_0^{T_3|\xi} \xleftarrow{\quad\cong\quad} (\zeta_0^{T_3} \wedge \alpha)(\mathbf{X}) \ltimes (M_\bullet \backslash G_\bullet^2, m_{M_\bullet \backslash G_\bullet^2}, (\tau_1^{\mathrm{op}}, 1), (1, \tau_2^{\mathrm{op}}), 1)$$

with $\alpha|_{\zeta_0^{T_3|\xi}}$ and *canonical* down to

$$(\zeta_0^{T_3} \wedge \alpha)(\mathbf{X})$$

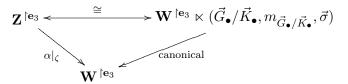(where we have suppressed the need to lift τ_i through $\zeta_0^{T_{3-i},T_3}$). Now by

simply observing that the quotient $M_\bullet \backslash G_\bullet^2$ is canonically bijective with $G_\bullet$ under the map $M_\bullet \cdot (g_1, g_2) \leftrightarrow g_1^{-1} \cdot g_2$ and applying this bijection fibrewise, the restricted action of T_3 on $\zeta_0^{T_3|\xi}$ is of course trivial and the restricted actions of T_1 and T_2 turn into the respective left and right actions by the cocycles τ_1 and τ_2 asserted in Theorem 7.2 (where some additional subscripts '$_3$' from the statement of that theorem have also been suppressed). ∎

We will prove Proposition 7.5 in several steps. First observe that since the extensions $\alpha_i|_{\zeta_i} : \mathbf{Z}_i \to \mathbf{W}_i$ are isometric and relatively ergodic for $T_3|_{\zeta_i}$, the non-ergodic Furstenberg–Zimmer theory of Section 5 enables us to pick coordinatizations by core-free homogeneous space data for the $(\mathbb{Z}\mathbf{e}_3)$-subactions

$$\mathbf{Z}_i^{\restriction \mathbf{e}_3} \xleftarrow{\quad\cong\quad} \mathbf{W}_i^{\restriction \mathbf{e}_3} \ltimes (G'_{i,\bullet}/K'_{i,\bullet}, m_{G'_{i,\bullet}/K'_{i,\bullet}}, \sigma'_i)$$

$$\alpha_i|_{\zeta_i} \searrow \qquad \swarrow \text{canonical}$$

$$\mathbf{W}_i^{\restriction \mathbf{e}_3}$$

(recall that $\mathbf{Z}_i^{\restriction \mathbf{e}_3}$ denotes the subaction system given by retaining only the action through T of the one-dimensional subgroup $\mathbb{Z}\mathbf{e}_3 \leq \mathbb{Z}^3$), where we may also choose the cocycle sections σ'_i to be ergodic.

Since ζ_1 and ζ_2 (like $\zeta_0^{T_1}$ and $\zeta_0^{T_2}$) are relatively independent over $\zeta_0^{T_1,T_2}$ under μ, we can combine the above two coordinatizations to give

$$\mathbf{Z}^{\restriction \mathbf{e}_3} \xleftarrow{\quad\cong\quad} \mathbf{W}^{\restriction \mathbf{e}_3} \ltimes (\vec{G}_\bullet/\vec{K}_\bullet, m_{\vec{G}_\bullet/\vec{K}_\bullet}, \vec{\sigma})$$

$$\alpha|_\zeta \searrow \qquad \swarrow \text{canonical}$$

$$\mathbf{W}^{\restriction \mathbf{e}_3}$$

where $\vec{G}_\bullet := G'_{1,\pi_1(\bullet)} \times G'_{2,\pi_2(\bullet)}$, $\vec{K}_\bullet := K'_{1,\pi_1(\bullet)} \times K'_{2,\pi_2(\bullet)}$ and $\vec{\sigma} := (\sigma'_1 \circ \pi_1, \sigma'_2 \circ \pi_2)$, and we here write π_i for the obvious factor map $W \to W_i$.

Of course, we do not know that the restrictions of T_1 and T_2 to the factors $\alpha_i|_{\zeta_i} : \mathbf{Z}_i \to \mathbf{W}_i$ are isometric, and so we have no similar coordinatization of these transformations using homogeneous space data and cocycles. We will appeal instead to the Relative Automorphism Structure Theorem 6.5 to describe them in terms of cocycles and fibrewise automorphisms.

First, however, an appeal to Corollary 4.4 gives a first step towards the more explicit description of the $T_3|_\xi$-invariant factor in terms of the above coordinatizations:

PROPOSITION 7.6. *There are $T_3|_\alpha$-invariant Mackey group data $M'_\bullet \leq \vec{G}_\bullet$ on W and a measurable section $\vec{b} : W \to \vec{G}_\bullet$ such that the factor map*

$$Z \to Z_0^{T_3|\alpha} \ltimes (M'_\bullet \backslash \vec{G}_\bullet / \vec{K}_\bullet) :$$

$$(w, g\vec{K}_w) \mapsto \left((\zeta_0^{T_1,T_3} \vee \zeta_0^{T_2,T_3})(w), M'_w \cdot \vec{b}(w) \cdot g \cdot \vec{K}_w \right)$$

coordinatizes the $T_3|_\xi$-invariant factor of $\mathbf{Z}$. ∎

The remainder of our work will go into analyzing this Mackey group data $M'_\bullet$ and section b' to deduce properties of the data $(G'_{i,\bullet}/K'_{i,\bullet}, \sigma'_i)$ that gave rise to them, and eventually reduce them to the special form promised by Proposition 7.5.

We now prove two technical lemmas that will underly our subsequent analysis, and which it seems easiest to introduce separately.

LEMMA 7.7. *Suppose that (X, μ, T) is a $\mathbb{Z}$-system, $G_\bullet$ is T-invariant measurable compact group data on X and $\sigma : X \to G_\bullet$ a cocycle-section, and that (Y, ν) is another standard Borel probability space. Suppose further that λ is a $(T \times \mathrm{id}_Y)$-invariant joining of μ and ν. If $M_\bullet \leq G_\bullet$ is the Mackey group data of $G_\bullet$ and σ over (X, μ, T) and $\pi : X \times Y \to X$ is the coordinate projection, then the Mackey group data $N_\bullet$ of $G_{\pi(\bullet)}$ and $\sigma \circ \pi$ over $(X \times Y, \lambda, T \times \mathrm{id}_Y)$ is given by $M_{\pi(\bullet)}$ (up to a T-invariant measurable choice of conjugates) λ-almost surely.*

REMARK. It is easy to see that $N_\bullet \leq M_{\pi(\bullet)}$; the point of this proposition is that if we adjoin to (X, μ, T) a system on a new space (Y, ν) for which the action is trivial, then the Mackey group data does not become any smaller.

Proof. We know from Section 4 that there is a section $b : X \times Y \to G_{\pi(\bullet)}$ such that $b(Tx, y)^{-1} \cdot \sigma(x) \cdot b(x, y) \in N_{(x,y)}$ for λ-almost every (x, y).

Let A be the λ-conegligible subset of $X \times Y$ where this coboundary condition obtains, and let

$$B_0 := \{(x, y) \in A : \text{some conjugate of } N_{(x,y)} \text{ is properly contained in } M_x\};$$

this is easily seen to be Borel and λ-almost $(T \times \mathrm{id}_Y)$-invariant, and so writing $B := \bigcap_{n \in \mathbb{Z}} T^n(B_0)$ we see that $\lambda(B) = \lambda(B_0)$ and that B is strictly $(T \times \mathrm{id}_Y)$-invariant. It will suffice to show that B is λ-negligible, so suppose otherwise. Then by Proposition 2.4 there are a non-negligible T-invariant subset $C \in \Sigma_X$ and a T-invariant measurable selector $\eta : C \to Y$ such that $(x, \eta(x)) \in B$ almost surely. We deduce that $b(Tx, \eta(x))^{-1} \cdot \sigma(x) \cdot b(x, \eta(x)) \in N_{(x,\eta(x))}$ for every $x \in C$ with $N_{(x,\eta(x))}$ properly contained in some (clearly measurably-varying) conjugate of the Mackey group data M_x, contradicting the conjugate-minimality of this latter that was proved in part (4) of Theorem 4.1. ∎

COROLLARY 7.8. *In the notation set up earlier in this section, we have*

$$\{g_1 : \text{there exists } g_2 \in G'_{2,\bullet} \text{ such that } (g_1, g_2) \in M'_\bullet\} = G'_{1,\bullet}$$

almost surely, and similarly for the projection of $M'_\bullet$ onto $G'_{2,\bullet}$.

Proof. We give the argument for $i = 1$. Simply observe that the extension $\mathbf{W}^{\restriction \mathbf{e}_3} \to \mathbf{W}_1^{\restriction \mathbf{e}_3}$ is relatively invariant, and so the Mackey group data for our coordinatization of the extension $\alpha|_{\zeta_1 \vee \alpha_2} : (\zeta_1 \vee \alpha_2)(\mathbf{Z})^{\restriction \mathbf{e}_3} \to \mathbf{W}^{\restriction \mathbf{e}_3}$ must simply be lifted from the Mackey group data for $\alpha_1|_{\zeta_1} : \mathbf{Z}_1^{\restriction \mathbf{e}_3} \to \mathbf{W}_1^{\restriction \mathbf{e}_3}$ downstairs. Since the former is clearly equal to the given one-dimensional projection of $M'_\bullet$, and the cocycle section σ'_1 is assumed to be ergodic, this completes the proof. ∎

The next properties of $M'_\bullet$ that we deduce require a little more work. We begin with a useful group-theoretic lemma.

LEMMA 7.9 (Deconstructing a relation between two group correspondences). *Suppose that G_1, G_2 are compact groups and that $M_1, M_2 \le G_1 \times G_2$ are two subgroups that both have full one-dimensional projections, and let their one-dimensional slices be*

$$L_{1,1} := \{g \in G_1 : (g, 1_{G_2}) \in M_1\}, \quad L_{1,2} := \{g \in G_2 : (1_{G_1}, g) \in M_1\}$$

and similarly $L_{2,1}$, $L_{2,2}$. Suppose further that $\Phi_i : G_i \xrightarrow{\cong} G_i$ and $h_i, k_i \in G_i$ for $i = 1, 2$ satisfy

$$(h_1, h_2) \cdot (\Phi_1 \times \Phi_2)(M_1) \cdot (k_1, k_2) = M_2.$$

Then $\Phi_i(L_{1,i}) = L_{2,i}$ for $i = 1, 2$.

Proof. Suppose first that $(g, 1_{G_2}) \in L_{1,1}$. Then the given equation tells us that

$$(h_1 \cdot \Phi_1(g) \cdot k_1, h_2 \cdot k_2) = (m_1, m_2)$$

for some $m_1, m_2 \in M_2$, and in this case $m_2 = h_2 \cdot k_2$ does not depend on g. Since the above must certainly hold if $g = 1_{G_1}$, applying it also for any other g and differencing gives

$$(h_1 \cdot \Phi_1(g) \cdot k_1) \cdot (h_1 \cdot \Phi_1(1_{G_1}) \cdot k_1)^{-1} = h_1 \cdot \Phi_1(g) \cdot h_1^{-1} \in L_{2,1},$$

so $\Phi_1(L_{1,1}) \subseteq h_1^{-1} \cdot L_{2,1} \cdot h_1$. An exactly symmetric argument gives the reverse inclusion, so in fact $\Phi_1(L_{1,1})$ is a conjugate of $L_{2,1}$. However, since M_1 and M_2 have full one-dimensional projections, by Lemma 2.8 it follows that in fact $\Phi_1(L_{1,1}) = L_{2,1}$, as required. The case of the other coordinate is similar. ∎

LEMMA 7.10. *If $H_{i,\bullet} \le G'_{i,\pi_i(\bullet)}$ are the one-dimensional slices of $M'_\bullet$, then*

(1) $H_{i,(w_1,w_2)}$ *$\alpha_\# \mu$-almost surely depends only on w_i, so after modifying on a negligible set we may write it as H_{i,w_i};*

(2) *under the above coordinatizations, for $i = 1, 2$ the map*

$$W_i \ltimes (G'_{i,\bullet}/K'_{i,\bullet}) \to W_i \ltimes (G'_{i,\bullet}/(H_{i,\bullet}K'_{i,\bullet})) :$$
$$(w_i, gK'_{i,w_i}) \mapsto (w_i, gH_{i,w_i}K'_{i,w_i})$$

defines a factor for the whole $\mathbb{Z}^3$-action T (that is, it is respected by T_1 and T_2 as well as T_3).

Proof. By symmetry it suffices to treat the case $i = 1$ for the first conclusion and $i = 2$ for the second (it will turn out that these come together). First deduce from Corollary 7.8 and Lemma 2.8 that in fact $H_{1,(w_1,w_2)} \trianglelefteq G'_{1,w_1}$ for almost every (w_1, w_2).

We will use the presence of the additional transformations of the factors ζ_1 and ζ_2 given by T_1. Of course, T_1 just restricts to the identity transformation on ζ_1. On the other hand, since $T_1|_{\zeta_2}$ commutes with the transformation $T_3|_{\zeta_2}$ which is relatively ergodic for the extension $\mathbf{Z}_2 \to \mathbf{W}_2$, the Relative Automorphism Structure Theorem 6.5 allows us to express

$$T_1|_{\zeta_2} \cong T_1|_{\alpha_2} \ltimes (L_{\rho'(\bullet)} \circ \Phi'_\bullet)|_{K'_{2,T_1|_{\alpha_2}(\bullet)}}^{K'_{2,\bullet}}$$

for some $\rho' : W_2 \to G'_{2,\bullet}$ and some $T_3|_{\alpha_2}$-invariant section $\Phi'_\bullet : W_2 \to \mathrm{Isom}(G'_{2,\bullet}, G'_{2,T_1|_{\alpha_2}(\bullet)})$ satisfying $\Phi'_\bullet(K'_{2,\bullet}) = K'_{2,T_1|_{\alpha_2}(\bullet)}$ almost surely.

Now observe from Corollary 6.9 that using the above expression and its partner for $T_2|_{\zeta_1}$ we may extend all three transformations $T_j|_{\zeta_i}$, $j = 1, 2, 3$, to the covering group extension $\tilde{\alpha}_i : \tilde{\mathbf{Z}}_i \to \mathbf{Z}_i \xrightarrow{\alpha_i|_{\zeta_i}} \mathbf{W}_i$ arising from our core-free homogeneous-space-data coordinatization of $T_3|_{\zeta_i}$, and that these extensions retain commutativity and all the relative invariance, ergodicity and isometricity properties listed above. Form the relatively independent joining

$$\tilde{\mathbf{Z}} = \tilde{\mathbf{Z}}_1 \otimes_{\{\zeta_0^{T_1,T_2}|_{\alpha_1} \circ \tilde{\alpha}_1 = \zeta_0^{T_1,T_2}|_{\alpha_2} \circ \tilde{\alpha}_2\}} \tilde{\mathbf{Z}}_2$$

with the coordinate projection factors back onto $\tilde{\mathbf{Z}}_1$ and $\tilde{\mathbf{Z}}_2$; with the resulting factor map onto $\mathbf{W}$ it now defines a covering group extension of $\alpha|_\zeta : \mathbf{Z} \to \mathbf{W}$ whose Mackey data are still $M'_\bullet$ and b' (by our initial construction of these). Moreover, these new factors $\tilde{\mathbf{Z}}_i$ and $\tilde{\mathbf{Z}}$ are located in a commutative diagram with the factors $\mathbf{W}_i$ and $\mathbf{W}$ just as we saw previously for $\mathbf{Z}_i$ and $\mathbf{Z}$ (except now not all as factors of the original overall system $\mathbf{X}$, but of some extended overall system).

It follows that for the purpose of proving this proposition, we may work with these covering group extensions throughout without disrupting the final conclusions; or, equivalently, that it suffices to treat the case in which the core-free kernels $K'_\bullet$ are trivial. Let us therefore make this assumption for the rest of this proof so as to lighten notation.

Given this assumption, consider the condition that $T_1|_\zeta$ respect $\zeta_0^{T_3}|_\zeta$ in terms of the above expression for $T_1|_{\zeta_2}$ and the Mackey data. First, since $M'_\bullet$ has full one-dimensional projections we may take the Mackey section $\vec{b}$ of Proposition 7.6 to be of the form $\vec{b}(w) = (1_{G_{1,w_1}}, b'(w))$. Now the above condition requires, in particular, that $\zeta_0^{T_3}|_\zeta(T_1|_\zeta(z))$ almost surely depend only on $\zeta_0^{T_3}|_\zeta(z)$ for $z \in Z$; and on the other hand, in terms of the above Mackey description, writing points of Z as (w, g_1, g_2) we know that $\zeta_0^{T_3}|_\zeta(w, g_1, g_2) = \zeta_0^{T_3}|_\zeta(w, g'_1, g'_2)$ if and only if

$$M'_w \cdot (1, b'(w)) \cdot (g_1, g_2) = M'_w \cdot (1, b'(w)) \cdot (g'_1, g'_2)$$
$$\Leftrightarrow (g_1, g_2) \in (1, b'(w)^{-1}) \cdot M'_w \cdot (1, b'(w)) \cdot (g'_1, g'_2).$$

Therefore the above relation between $T_1|_\zeta$ and $\zeta_0^{T_3}|_\zeta$ simply asserts that for $\alpha_\#\mu$-almost every $(w_1, w_2) \in W$, for Haar-almost every $(g'_1, g'_2) \in G'_{1,w_1} \times G'_{2,w_2}$ there is some $(g''_1, g''_2) \in G'_{1,w_1} \times G'_{2,T_1|_{\alpha_2}(w_2)}$ such that

$$(\mathrm{id}_{G'_{1,w_1}} \times (L_{\rho'(w_2)} \circ \Phi'_{w_2}))\big((1, b'(w_1, w_2)^{-1}) \cdot M'_{(w_1, w_2)} \cdot (1, b'(w_1, w_2)) \cdot (g'_1, g'_2)\big)$$
$$= (1, b'(w_1, T_1|_{\alpha_2}(w_2))^{-1}) \cdot M'_{(w_1, T_1|_{\alpha_2}(w_2))} \cdot (1, b'(w_1, T_1|_{\alpha_2}(w_2))) \cdot (g''_1, g''_2),$$

or, rearranging, that

$$\big(1_{G'_{1,w_1}}, b'(w_1, T_1|_{\alpha_2}(w_2))\rho'(w_2)\Phi'_{w_2}(b'(w_1, w_2)^{-1})\big)(\mathrm{id}_{G'_{1,w_1}} \times \Phi'_{w_2})(M'_{(w_1, w_2)})$$
$$\cdot \big(g'_1(g''_1)^{-1}, \Phi'_{w_2}(b'(w_1, w_2)g'_2)(b'(w_1, T_1|_{\alpha_2}(w_2))g''_2)^{-1}\big) = M'_{(w_1, T_1|_{\alpha_2}(w_2))}.$$

The two desired conclusions now follow from applying Lemma 7.9 to this equation for the two coordinate projections onto G'_{1,w_1} and G'_{2,w_2}. Under the first coordinate projection we obtain

$$H_{1,(w_1, w_2)} = \mathrm{id}_{G'_{1,w_1}}(H_{1,(w_1, w_2)}) = H_{1,(w_1, T_1|_{\alpha_2}(w_2))},$$

so $H_{1,(w_1, w_2)}$ is a $T_1|_\alpha$-invariant subgroup of G'_{1,w_1}, and so recalling that $T_1|_\alpha$ is relatively ergodic on the extension $\alpha_1|_\alpha : W \to W_1$ we deduce that $H_{1,(w_1, w_2)}$ is virtually a function of w_1 alone, as required for conclusion (1).

For the second coordinate projection we obtain

$$\Phi'_{w_2}(H_{2,(w_1, w_2)}) = H_{2,(w_1, T_1|_{\alpha_2}(w_2))}.$$

In view of the conclusion (1) obtained above we can simplify this to

$$\Phi'_{w_2}(H_{2,w_2}) = H_{2,T_1|_{\alpha_2}(w_2)},$$

and now this is precisely the condition given by the Relative Automorphism Structure Theorem 6.5 for $T_1|_{\zeta_2}$ to respect the given map as a factor map. Since it is clear that the given map defines a factor for the restrictions of T_2 (since this acts trivially on the whole of $\mathbf{Z}_2$) and T_3 (since this acts on this extension by a $G_{2,\bullet}$-valued cocycle-section, and so our factor map is simply

the fibrewise quotient by the subgroup $H_{i,\bullet}$, recalling our assumption that $K'_{i,\bullet}$ is almost surely trivial), this completes the proof. ∎

In view of the above result, we are now able to define our desired intermediate factors $\mathbf{Z}_i \xrightarrow{\xi_i|_{\zeta_i}} \mathbf{Y}_i \xrightarrow{\alpha_i|_{\xi_i}} \mathbf{W}_i$ by the commutative diagrams

$$
\begin{array}{ccc}
\mathbf{Z}_1 & \xleftarrow{\quad\cong\quad} & \mathbf{W}_1 \ltimes (G'_{1,\bullet}/K'_{1,\bullet}, m_{G'_{1,\bullet}/K'_{1,\bullet}}, 1_{G'_{1,\bullet}}, (L_{\rho'_1(\bullet)} \circ \Phi'_{1,\bullet}), \sigma'_1) \\
\downarrow{\scriptstyle \xi_1|_{\zeta_1}} & & \downarrow{\scriptstyle \text{canonical}} \\
\mathbf{Y}_1 & \xleftarrow{\quad\cong\quad} & \mathbf{W}_1 \ltimes (G_{1,\bullet}/K_{1,\bullet}, m_{G_{1,\bullet}/K_{1,\bullet}}, 1_{G_{1,\bullet}}, (L_{\rho_1(\bullet)} \circ \Phi_{1,\bullet}), \sigma_1) \\
\downarrow{\scriptstyle \alpha_1|_{\xi_1}} & \swarrow{\scriptstyle \text{canonical}} & \\
\mathbf{W}_1 & &
\end{array}
$$

and similarly for $\mathbf{Y}_2$, where $G_{i,\bullet} := G'_{i,\bullet}/H_{i,\bullet}$, $K_{i,\bullet} := (H_{i,\bullet}K'_{i,\bullet})/H_{i,\bullet}$ and ρ_i, $\Phi_{i,\bullet}$ and σ_i are the appropriate quotients or restrictions of ρ'_i, $\Phi'_{i,\bullet}$ and σ'_i: part (1) above gives that $H_{i,\bullet}$ is correctly defined as a function on W_i, and part (2) gives that the above diagram defines a factor map for our whole $\mathbb{Z}^3$-action.

The important feature of these new smaller extensions $\mathbf{Y}_i \to \mathbf{W}_i$ is that the Mackey group data $M_\bullet$ of their joining under $\mathbf{X}$ takes a particularly simple form: after having quotiented out the one-dimensional slices $H_{i,\bullet}$, it is almost surely the graph of a continuous isomorphism. Indeed, $M_\bullet$ is clearly obtained from $M'_\bullet$ simply by quotienting out the normal subgroup data $H_{1,\pi_1(\bullet)} \times H_{2,\pi_2(\bullet)}$, and from the definition of $H_{i,\bullet}$ it follows that $M_\bullet$ has full one-dimensional projections and trivial one-dimensional slices almost everywhere, and so defines almost everywhere the graphs of some measurably-varying $T_3|_\alpha$-invariant isomorphisms $\Psi_{(w_1,w_2)} : G_{1,w_1} \xrightarrow{\cong} G_{2,w_2}$. Henceforth we will refer to these as the *Mackey isomorphisms*. In addition we set $b := b' \cdot H_{2,\pi_2(\bullet)}$, so that $(1_{G_{1,\pi_1(\bullet)}}, b)$ is a Mackey section of the extension $\mathbf{Y}^{\restriction \mathbf{e}_3} \to \mathbf{W}^{\restriction \mathbf{e}_3}$ associated to the choice of Mackey group data $M_\bullet$.

On the other hand, from the description given in Proposition 7.6 it follows that the T_3-invariant factor is actually contained in the join of these smaller isometric extensions $\alpha_i|_{\xi_i} : \mathbf{Y}_i \to \mathbf{W}_i$, and so it will suffice to study these new factors. The remaining steps of this subsection will give a recoordinatization of these new factors into the form required by Proposition 7.5.

COROLLARY 7.11. *We have* $\zeta_0^T|_{\alpha_1}(w_1) = \zeta_0^T|_{\alpha_2}(w_2)$ *for* $\alpha_{\#}\mu$*-almost every* (w_1, w_2)*, and there are compact group data* $G_\bullet$ *invariant for the whole action* T *such that we can recoordinatize the extensions* $\alpha_i|_{\xi_i} : \mathbf{Y}_i \to \mathbf{W}_i$ *so that* $G_{1,w_1} = G_{\zeta_0^T|_{\alpha_1}(w_1)} = G_{2,w_2}$ *for* $\alpha_{\#}\mu$*-almost every* (w_1, w_2)*.*

Proof. The first assertion is clear from the definitions.

By $T_3|_{\alpha_i}$-invariance the groups G_{i,w_i} actually depend only on $z_i :=$ $\zeta_0^{T_3}|_{\alpha_i}(w_i) \in Z_0^{T_i,T_3}$, and similarly the isomorphism $\Psi_{(w_1,w_2)}$ depends only on the image (z_1, z_2) of (w_1, w_2). In addition, by Theorem 7.1 the coordinates z_1, z_2 of this image are relatively independent over $\zeta_0^T|_{\alpha_1}(w_1) = \zeta_0^T|_{\alpha_2}(w_w)$ under $\alpha_{\#}\mu$.

Now let $P : Z_0^T \xrightarrow{\mathrm{p}} Z_0^{T_2,T_3}$ be a probability kernel representing the disintegration of $(\zeta_0^{T_2,T_3})_{\#}\mu$ over $\zeta_0^T|_{\zeta_0^{T_2,T_3}}$. For almost every $z_1 \in Z_0^{T_1,T_3}$ we can choose a measurable family of isomorphisms $\Theta_{2,z_2} : G_{2,z_2} \xrightarrow{\cong} G_{1,z_1}$ defined for $P(\zeta_0^T|_{\zeta_0^{T_1,T_3}}(z_1), \cdot)$-almost every $z_2 \in Z_0^{T_2,T_3}$, because the Mackey isomorphisms themselves witness that these almost surely exist. Making a measurable selection of such a z_1 in each fibre of $Z_0^{T_1,T_3} \to Z_0^T$, we now take this family as defining a fibrewise isomorphism recoordinatization of our initial homogeneous-space-data coordinatization of $\alpha_2|_{\xi_2} : \mathbf{Y}_2 \to \mathbf{W}_2$ obtained above. This has the effect of adjusting to a coordinatization in which the covering group of the homogeneous space fibre over $z_2 \in Z_2$ is $P(\zeta_0^T|_{\zeta_0^{T_1,T_3}}(z_1), \cdot)$-almost everywhere equal to G_{1,z_1}, and with the kernel of the homogeneous space fibre given by $\Theta_{2,z_2}(K_{2,z_2}) \leq G_{1,z_1}$.

In particular, the covering group data of this new coordinatization depends only on $\zeta_0^T|_{\zeta_0^{T_1,T_3}}(z_1) = \zeta_0^T|_{\alpha_1}(w_1) = \zeta_0^T|_{\alpha_2}(w_2)$. Exactly similarly we can now recoordinatize $\alpha_1|_{\xi_1} : \mathbf{Y}_1 \to \mathbf{W}_1$ to have covering fibre groups also depending only on $\zeta_0^T|_{\alpha_1}(w_1) = \zeta_0^T|_{\alpha_2}(w_2)$. Since both these recoordinatizations are by fibrewise isomorphisms that are invariant for the relevant restrictions of T_3, the new coordinatizations of these extensions that result are still given as cocycle-section extensions for these restrictions of T_3. Finally, in this new coordinatization the measurable family of Mackey isomorphisms $\Psi_{(w_1,w_2)}$ clearly shows that after one more fibrewise recoordinatization by a T-invariant isomorphism we are left with the same T-invariant group data $G_\bullet$ everywhere. ∎

Now let us re-apply the Relative Automorphism Structure Theorem 6.5 to write

$$T_1|_{\xi_2} = T_1|_{\alpha_2} \ltimes (L_{\rho_1(\bullet)} \circ \Phi_{1,\bullet})|_{K_{2,T_1|_{\alpha_2}(\bullet)}}^{K_{2,\bullet}}$$

and

$$T_2|_{\xi_1} = T_2|_{\alpha_1} \ltimes (L_{\rho_2(\bullet)} \circ \Phi_{2,\bullet})_{H_1,T_2|_{\alpha_1}(\bullet)}^{H_1,\bullet}$$

where now Φ_{i,w_i} is $T_3|_{\alpha_{3-i}}$-invariant and takes values in $\mathrm{Aut}\, G_{\zeta_0^T|_{\alpha_i}(w_i)}$ for $i = 1, 2$. In addition, we recall the notation $\mathrm{Co}_{\rho(\bullet)}$ for the fibrewise automorphism of some measurable group data $G_\bullet$ given by fibrewise conjugation by a section ρ of $G_\bullet$.

PROPOSITION 7.12. *The extensions* $\alpha_i|_{\xi_i} : \mathbf{Y}_i \to \mathbf{W}_i$ *can be recoordinatized by fibrewise affine transformations so that*

(1) *there are cocycles* $\tau_i : W_{3-i} \to G_{\zeta_0^T|_{\alpha_i}(\bullet)}$ *such that* $L_{\rho_i(\bullet)} \circ \Phi_{i,\bullet} = R_{\tau_i(\bullet)}$;

(2) *the Mackey isomorphisms are trivial:* $\Psi_\bullet \equiv \mathrm{id}_{G_{\zeta_0^T|_\alpha(\bullet)}}$;

(3) *the Mackey section is trivial:* $b \equiv 1_{G_{\zeta_0^T|_\alpha(\bullet)}}$;

(4) *the cocycle* τ_i *is invariant under* $T_3|_{\alpha_i}$ *and the cocycle* σ_i *is invariant under* $T_{3-i}|_{\alpha_i}$.

Proof. This will follow from a careful consideration of the commutativity conditions relating the expressions for our three transformations on $\mathbf{Y}_i$. We make our recoordinatizations in two steps, the first by fibrewise automorphisms and the second by fibrewise rotations. We will construct these so as to guarantee the asserted properties of the ingredients ρ and Φ, and will then find that the asserted forms of Ψ and b are an immediate consequence.

First observe that just as in the proof of Lemma 7.10, we may lift all of our commuting transformations $T_j|_{\alpha_i}$ to the covering group data extensions of $\alpha_i|_{\xi_i} : \mathbf{Y}_i \to \mathbf{W}_i$, and have that $M_\bullet$ and b will still be Mackey data of their relatively independent joining over $\mathbf{Z}_0^{T_1,T_2}$, and therefore if we effect our desired fibrewise recoordinatizations on these covering group data extensions then simply quotienting will give the desired recoordinatizations of $\alpha_i|_{\xi_i} : \mathbf{Y}_i \to \mathbf{W}_i$. As in Lemma 7.10, this argument reduces our work to the special case when $K_\bullet \equiv \{1_{G_\bullet}\}$.

The remainder of our work breaks into five steps.

STEP 1. We consider the case $i = 1$. First recall our earlier expression of the fact that $T_1|_\xi$ respects $\zeta_0^{T_3}|_\xi$: for $\alpha_\#\mu$-almost every $(w_1, w_2) \in W$, setting $s := \zeta_0^T|_{\alpha_1}(w_1)$, for Haar-almost any $g' \in G_s$ there is some $g'' \in G_s$ for which

$$(\mathrm{id}_{G_s} \times (L_{\rho_1(w_2)} \circ \Phi_{1,w_2}))\big((1, b(w_1, w_2)^{-1}) \cdot M_{(w_1, w_2)} \cdot (1, b(w_1, w_2)) \cdot (1, g')\big)$$
$$= (1, b(w_1, T_1|_{\alpha_2}(w_2))^{-1}) \cdot M_{(w_1, T_1|_{\alpha_2}(w_2))} \cdot (1, b(w_1, T_1|_{\alpha_2}(w_2))) \cdot (1, g'').$$

We can rewrite this condition in terms of the Mackey isomorphisms to give

$$\rho_1(w_2) \cdot (\Phi_{1,w_2} \circ L_{b(w_1,w_2)^{-1}} \circ \Psi_{(w_1,w_2)})(\bullet) \cdot \Phi_{1,w_2}(b(w_1, w_2) \cdot g')$$
$$= b(w_1, T_1|_{\alpha_2}(w_2))^{-1} \cdot \Psi_{(w_1, T_1|_{\alpha_2}(w_2))}(\bullet) \cdot b(w_1, T_1|_{\alpha_2}(w_2)) \cdot g'',$$

and now if we write $\tilde{\Psi}_\bullet := \mathrm{Co}_{b(\bullet)^{-1}} \circ \Psi_\bullet$ and $\tilde{\Phi}_{1,\bullet} := \mathrm{Co}_{\rho_1(\bullet)} \circ \Phi_{1,\bullet}$ this in turn becomes

$$(\tilde{\Phi}_{1,w_2} \circ \tilde{\Psi}_{(w_1,w_2)})(\bullet) \cdot \rho_1(w_2) \cdot \Phi_{1,w_2}(g') = \tilde{\Psi}_{(w_1, T_1|_{\alpha_2}(w_2))}(\bullet) \cdot g''$$

(so we have simply shifted all the 'translation' parts of our affine transfor-

mations over to the right). Finally, this now clearly requires that

$$\tilde{\Phi}_{1,w_2} = \tilde{\Psi}_{(w_1,T_1|_{\alpha_2}(w_2))} \circ \tilde{\Psi}^{-1}_{(w_1,w_2)}$$

almost everywhere: the automorphisms graphed by the Mackey group data have themselves become a coboundary for the automorphism-valued cocycle $\tilde{\Phi}_{1,\bullet}$.

We will refer to the above as the 'automorphism coboundary equation' for the remainder of this proof. The condition that this hold for almost every (w_1, w_2) also gives non-trivial information on the automorphisms $\tilde{\Psi}_{(w_1,w_2)}$ for different w_1, since w_1 is absent from the left-hand side. Since the extension $\alpha_2|_\alpha : \mathbf{W} \to \mathbf{W}_2$ is relatively invariant for the restrictions of T_1 and T_3, using Proposition 2.4 we can therefore choose a T_1-invariant measurable selector $\eta : W_2 \to W_1$ so that

$$\tilde{\Phi}_{1,w_2} = \tilde{\Psi}_{(\eta(w_2),T_1|_{\alpha_2}(w_2))} \circ \tilde{\Psi}^{-1}_{(\eta(w_2),w_2)}$$

holds almost surely and so witnesses that $\tilde{\Phi}_{1,w_2}$ is a coboundary in $\mathrm{Aut}(G_s)$ for the transformation $T_1|_{\alpha_2} : W_2 \to W_2$.

Naïvely we should now like to use the cocycle $\tilde{\Psi}_{\eta(\bullet),\bullet}$ to make a fibrewise automorphism recoordinatization of the extension $\alpha_2|_{\xi_2} : \mathbf{Y}_2 \to \mathbf{W}_2$ so that the first of our automorphism-valued coboundary equations above gives a simplification of $\tilde{\Phi}_{1,w_2}$. However, this idea runs into difficulties because the new isomorphisms $\tilde{\Psi}_\bullet$, unlike $\Psi_\bullet$, are not necessarily T_3-invariant, and so applying them fibrewise may disrupt the coordinatization of $T_3|_{\xi_2}$ as acting by rotations.

STEP 2. The best we can do at this stage is to apply fibrewise the automorphisms $\Psi^{-1}_{(\eta(w_2),w_2)}$ to our coordinatization of $\alpha_2|_{\xi_2}$. This gives some improvement: in the resulting new coordinatization of this extension, our automorphism coboundary equation above now reads

$$\tilde{\Phi}_{1,w_2} = \mathrm{Co}^{-1}_{b(\eta(w_2),T_1|_{\alpha_2}(w_2))} \circ \mathrm{Co}_{b(\eta(w_2),w_2)} = \mathrm{Co}_{b(\eta(w_2),T_1|_{\alpha_2}(w_2))^{-1} \cdot b(\eta(w_2),w_2)}.$$

Recalling that $\tilde{\Phi}_{1,\bullet} = \mathrm{Co}_{\rho_1(\bullet)} \circ \Phi_{1,\bullet}$ this unravels to give

$$\Phi_{1,w_2} = \mathrm{Co}^{-1}_{\rho_1(w_2)} \circ \mathrm{Co}_{b(\eta(w_2),T_1|_{\alpha_2}(w_2))^{-1} \cdot b(\eta(w_2),w_2)}$$
$$= \mathrm{Co}_{\rho_1(w_2)^{-1} \cdot b(\eta(w_2),T_1|_{\alpha_2}(w_2))^{-1} \cdot b(\eta(w_2),w_2)},$$

so we conclude, in particular, that the automorphism-valued cocycle $\Phi_{1,\bullet}$ takes values in the compact subgroup of inner automorphisms, and so we may represent it as $\mathrm{Co}_{\theta(\bullet)}$ for some $T_3|_{\alpha_2}$-invariant section $\theta : W_2 \to G_{\zeta_0^T|_{\alpha_2}(\bullet)}$. By writing out the above automorphism coboundary equation in terms of θ

it becomes

$$\mathrm{Co}_{\theta(w_2)} = \mathrm{Co}^{-1}_{\rho_1(w_2)} \circ \mathrm{Co}_{b(\eta(w_2), T_1|_{\alpha_2}(w_2))^{-1} \cdot b(\eta(w_2), w_2)}$$

$$= \mathrm{Co}_{\rho_1(w_2)^{-1} \cdot b(\eta(w_2), T_1|_{\alpha_2}(w_2))^{-1} \cdot b(\eta(w_2), w_2)},$$

and so if we now substitute into our original expression for $T_1|_{\xi_2}$ we obtain

$$T_1|_{\xi_2} = T_1|_{\alpha_2} \ltimes (L_{\rho_1(\bullet)} \circ \mathrm{Co}_{\theta(\bullet)}) = T_1|_{\alpha_2} \ltimes (R_{\rho_1(\bullet)} \circ \mathrm{Co}_{\rho_1(\bullet) \cdot \theta(\bullet)})$$

$$= T_1|_{\alpha_2} \ltimes (R_{\rho_1(\bullet)} \circ \mathrm{Co}_{b(\eta(\bullet), T_1|_{\alpha_2}(\bullet))^{-1} \cdot b(\eta(\bullet), \bullet)}).$$

It follows that if we now make a second fibrewise recoordinatization of $\alpha_2|_{\xi_2} : \mathbf{Y}_2 \to \mathbf{W}_2$, this time by rotating each fibre copy of G_s from the left by $b(\eta(w_2), w_2)$ (which virtually depends only on $s = \zeta_0^T|_{\alpha_2}(w_2)$), we are left with a resulting coordinatization of $T_1|_{\xi_2}$ in the desired form of an opposite action:

$$T_1|_{\xi_2} = T_1|_{\alpha_2} \ltimes R_{\tau_1(\bullet)}$$

where

$$\tau_1(\bullet) := b(\eta(\bullet), \bullet)^{-1} \cdot b(\eta(\bullet), T_1|_{\alpha_2}(\bullet)) \cdot \rho_1(\bullet).$$

Of course, both of the above recoordinatizations can be repeated analogously for the extension $\alpha_1|_{\xi_1} : \mathbf{Y}_1 \to \mathbf{W}_1$ to put $T_2|_{\xi_1}$ into a similar right-multiplicative form.

It follows that in the coordinatizations of these extensions that we have now obtained, $T_3|_{\xi_i}$ is still in the form of a cocycle-section extension $T_3|_{\alpha_i} \ltimes \sigma_i$ (with a modified cocycle-section σ_i) and $T_1|_{\xi_2}$ and $T_2|_{\xi_1}$ are in the desired right-multiplicative form, as for part (1) of the proposition.

STEP 3. We now 'invert' the above implication to discover what consequences these improved coordinatizations imply for the data $\Psi_\bullet$ and $b_\bullet$.

Recall that our first automorphism cocycle equation held for almost all (w_1, w_2), before we chose the measurable selector η, and so in our latest coordinatization this tells us that

$$\mathrm{Co}_{\tau_1(w_2)} = \tilde{\Psi}_{(w_1, T_1|_{\alpha_2 \circ \xi_2}(w_2))} \circ \tilde{\Psi}^{-1}_{(w_1, w_2)}.$$

(Note that our first fibrewise recoordinatization above by automorphisms rendered the cocycle $\Psi_\bullet$ inner at $(\alpha_2)_\# \mu$-almost all the points $(\eta(w_2), w_2)$, which depend on our choice of measurable selector η, but we have not yet seen that this cocycle is inner for almost all (w_1, w_2) as a result of this recoordinatization, hence our need to go back to the above form of this equation for this stage of the argument.)

It follows that the class $\tilde{\Psi}_{(w_1, w_2)} \circ \mathrm{Inn}(G_s) \in \mathrm{Out}(G_s)$ is invariant under the action of $\mathrm{id}_{W_1} \times T_1|_{\alpha_2} = T_1|_\alpha$, and it follows similarly that it is invariant under $T_2|_\alpha$. On the other hand, we have $\tilde{\Psi}_{(w_1, w_2)} \circ \mathrm{Inn}(G_s) = \Psi_{(w_1, w_2)} \circ \mathrm{Inn}(G_s)$ and this latter is clearly invariant under $T_3|_\alpha$, since it arises from the Mackey group data. Therefore it is actually $T|_\alpha$-invariant, and so since

$\mathrm{Inn}(G_s) \trianglelefteq \mathrm{Aut}(G_s)$ is compact, and so the resulting space of equivalence classes $\mathrm{Out}(G_s)$ is smooth, it is almost surely equal to $\Psi_s \circ \mathrm{Inn}(G)$ for some Borel map $\Psi_\bullet : Z_0^T \to \mathrm{Aut}(G_\bullet)$. Therefore one last fibrewise automorphism recoordinatization of $\alpha_2|_{\xi_2} : \mathbf{Y}_2 \to \mathbf{W}_2$ by $\Psi_{\zeta_0^T|_{\alpha_2}(\bullet)}$ (which still does not disrupt any of the properties guaranteed previously, provided we replace ρ_1 and τ_1 with $\Psi_{\zeta_0^T|_{\alpha_2}(\bullet)}(\rho_1)$ and $\Psi_{\zeta_0^T|_{\alpha_2}(\bullet)}(\tau_1)$) now gives Mackey group data of the form

$$M_\bullet \equiv \{(g, g_0(\bullet)gg_0(\bullet)^{-1}) : g \in G_{\zeta_0^T|_{\alpha}(\bullet)}\}$$

for some $T_3|_\alpha$-invariant section $g_0 : W \to G_{\zeta_0^T|_{\alpha}(\bullet)}$, and now we can simply adjust the Mackey section b so that $g_0 \equiv 1$, and so the Mackey group data can be taken to be the diagonal subgroup almost everywhere.

STEP 4. Having removed all the non-trivial outer automorphisms and adjusted the joining Mackey group data, our automorphism coboundary equation has now simplified down to

$$\tilde{\Phi}_{1,w_2} = \mathrm{Co}_{\tau_1(\bullet)} \circ \mathrm{Co}_{\tau_1(\bullet)}^{-1} = \mathrm{id} = \mathrm{Co}_{b(w_1, T_1|_{\alpha_2}(w_2))^{-1} \cdot b(w_1, w_2)},$$

and hence we deduce that $b(\bullet) \cdot \mathrm{C}(G_{\zeta_0^T|_{\alpha}(\bullet)})$, where $\mathrm{C}(G_{\zeta_0^T|_{\alpha}(\bullet)})$ is the centre of $G_{\zeta_0^T|_{\alpha}(\bullet)}$, is $T_1|_\alpha$-invariant, and similarly that it is $T_2|_\alpha$-invariant. Making another measurable selection and recoordinatizing each fibre of $\alpha_2|_{\xi_2} : \mathbf{Y}_2 \to \mathbf{W}_2$ by a left rotation by $b(\eta(\bullet), \bullet)$ therefore preserves the structure of $T_1|_\alpha$ as an *opposite* rotation (since the resulting additional cocycle $b(w_1, T_1|_{\alpha_2}(w_2))^{-1} \cdot b(w_1, w_2)$ acting on the left takes values in $\mathrm{C}(G_s)$, and so may in fact be taken to act on either side); and after making this recoordinatization we find that the Mackey section has also trivialized.

STEP 5. Finally, let us look back at the relation between σ_1 and σ_2 that is implied by the cocycle equation satisfied by the Mackey data given by part (3) of Theorem 4.1 in light of this newly simplified Mackey group and section: this now becomes simply that

$$\sigma_2(w_2) = \sigma_1(w_1)$$

$\alpha_\# \mu$-almost surely, and hence in this coordinatization it follows that each σ_i virtually depends only on $\zeta_0^{T_1,T_2}|_{\alpha_i}$, or, equivalently, is $T_{3-i}|_{\alpha_i}$-invariant. Given this, the condition that $T_1|_{\xi_2}$ and $T_2|_{\xi_2}$ commute simply reads that for almost every $w_2 \in W_2$ we have

$$\begin{aligned}
\sigma_2(w_2) \cdot g \cdot \tau_1(T_3|_{\alpha_2}(w_2)) &= \sigma_2(T_1|_{\alpha_2}(w_2)) \cdot g \cdot \tau_1(w_2) \\
&= \sigma_2(w_2) \cdot g \cdot \tau_1(w_2) \quad \forall g \in G_s,
\end{aligned}$$

and so we must also have that τ_1 is $T_3|_{\alpha_2}$-invariant, and similarly that τ_2 is $T_3|_{\alpha_1}$-invariant. This completes the proof. ■

The recoordinatization of the preceding proposition leaves only one detail remaining for the proof of Proposition 7.5.

COROLLARY 7.13. *In our homogeneous-space-data coordinatizations of* $\alpha_i|_{\xi_i} : \mathbf{Y}_i \to \mathbf{W}_i$ *the core-free kernels* $K_{i,\bullet}$ *are almost surely trivial.*

Proof. As remarked at the beginning of the preceding proof, we can lift to the covering group extensions and make the adjustments of Proposition 7.12 there, and they will then quotient back down to well-defined recoordinatizations of the original extensions, because at each stage we have only applied either fibrewise automorphism or fibrewise left rotations. From these we have obtained expressions

$$T_1|_{\xi_2} = T_1|_{\alpha_2} \ltimes R_{\tau_1}$$

and similarly for $T_2|_{\xi_1}$ at the level of the covering group extensions, and so for these to have well-defined quotient it is necessary that for almost every w_1 all two-sided cosets $g \cdot K_{2,w_2} \cdot \tau_1(w_2)$ for $g \in G_{\zeta_0^T|_{\alpha_2}(w_2)}$ actually be left cosets of K_{2,w_2}. This, in turn, requires that τ_1 almost surely take values in the normalizer $\mathrm{N}_{G_{\zeta_0^T|_{\alpha_2}(w_2)}}(K_{2,w_2})$, which is a closed measurably-varying subgroup of $G_{\zeta_0^T|_{\alpha_2}(w_2)}$.

Now we recall that T_1 restricts to a *relatively ergodic* action on the extension $\alpha_2|_{\xi_2} : \mathbf{Y}_2 \to \mathbf{W}_2$—a condition we have not exploited so far—and so we must have

$$\mathrm{N}_{G_{\zeta_0^T|_{\alpha_2}(w_2)}}(K_{2,w_2}) \cdot K_{2,w_2} = G_{\zeta_0^T|_{\alpha_2}(w_2)}$$

almost surely, for otherwise the homogeneous space fibres of the extension $\alpha_2|_{\xi_2} : \mathbf{Y}_2 \to \mathbf{W}_2$ would decompose into cosets of the closed subgroups $\mathrm{N}_{G_{\zeta_0^T|_{\alpha_2}(w_2)}}(K_{2,w_2}) \cdot K_{2,w_2}$ to give additional non-trivial invariant sets under the restriction $T_1|_{\alpha_2}$. However, since

$$\mathrm{N}_{G_{\zeta_0^T|_{\alpha_2}(w_2)}}(K_{2,w_2}) \supseteq K_{2,w_2},$$

this requires in fact that

$$\mathrm{N}_{G_{\zeta_0^T|_{\alpha_2}(w_2)}}(K_{2,w_2}) = G_{\zeta_0^T|_{\alpha_2}(w_2)}$$

almost surely, and since K_{2,w_2} is core-free this is possible only if $K_{2,w_2} = \{1\}$ almost surely. An exactly similar argument treats $K_{2,\bullet}$. ∎

As remarked previously, this completes the proof of Proposition 7.5. ∎

7.4. Application to characteristic factors. We will finish this section by offering a second application of our machinery (although in truth it is largely a corollary of the above).

Since Furstenberg's ergodic-theoretic proof of Szemerédi's Theorem in [22] and his extension with Katznelson of this result to the multi-

dimensional setting in [24], considerable interest has been attracted by the 'non-conventional' ergodic averages

$$\frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^{d} f_i \circ T_i^n$$

associated to a commuting d-tuple of probability-preserving transformations $T_1, \ldots, T_d : \mathbb{Z} \curvearrowright (X, \mu)$, that emerge naturally in the course of those proofs. That these averages converge in $L^2(\mu)$ as $N \to \infty$ in the case $d = 2$ was first shown by Conze and Lesigne in [12], and their result has since been extended in many directions [13, 14, 45, 30, 31, 46], culminating in the first proof of the fully general case by Tao in [42]. We refer the reader to those papers and to [6] for more thorough motivation and historical discussion of this problem.

Conze and Lesigne's proof of convergence is comparatively soft, using only quite weak structural information about the above averages to show that they converge (in particular, using only the structure of certain finite-rank modules rather than their concrete coordinatizations). More recently, other convergence results for non-conventional ergodic averages have been based on a similar but more detailed analysis, resting on the notion of a 'characteristic tuple of factors'. A tuple of factors $\xi_i : \mathbf{X} \to \mathbf{Y}_i$ is *characteristic* if

$$\frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^{d} f_i \circ T_i^n - \frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^{d} \mathsf{E}_\mu(f_i \,|\, \xi_i) \circ T_i^n \to 0$$

in $L^2(\mu)$ as $N \to \infty$ for any $f_1, \ldots, f_d \in L^\infty(\mu)$. Starting with the Conze–Lesigne proof, most convergence proofs in this area require at some stage the identification of a characteristic tuple of factors (or a suitable finitary analog of them in the case of Tao's proof) on which the restricted actions of each T_i take simplified forms, so that the right-hand averages above can be analyzed to prove convergence more easily.

A precise description of these characteristic factors in the special case when $T_i = T^i$ for some fixed ergodic transformation T has now been given in terms of the special class of 'pronilsystems' in work of Host and Kra [31] (see also the subsequent approach of Ziegler [47]). Frantzikinakis and Kra have extended this description to more general commuting tuples subject to some additional ergodicity assumption in [19], but a description for arbitrary tuples of commuting transformations, without those ergodicity assumptions, seems to be more difficult. Indeed, it may be that no comparably clean and useful description is available in the general case. However, at least when $d = 2$ a reasonably simple coordinatization of a characteristic pair of factors seems to have been folklore knowledge in ergodic theory for some time, and in this subsection we will show how our theory enables a careful proof of it.

THEOREM 7.14 (Characteristic factors for double non-conventional averages). *Given a $\mathbb{Z}^2$-system $\mathbf{X} = (X, \mu, T_1, T_2)$, let $\mathbf{W}_i$ be the target system of the joined factor $\zeta_0^{T_i} \vee \zeta_0^{T_1=T_2}$. Then $\mathbf{X}$ admits a characteristic pair of factors $\xi_i : \mathbf{X} \to \mathbf{Y}_i$ that extend the factors $\zeta_0^{T_i} \vee \zeta_0^{T_1=T_2}$ and can be described as follows: there are T-invariant compact group data $G_\bullet$, a $T_1|_{\zeta_0^{T_1=T_2}}$-ergodic cocycle $\sigma : \mathbf{Z}_0^{T_1=T_2} \to G$, and a pair of $T_i|_{\zeta_0^{T_{3-i}}}$-ergodic cocycles $\tau_i : \mathbf{Z}_0^{T_{3-i}} \to G$ such that we can coordinatize*

$$(Y_1, (\xi_1)_{\#}\mu) \xleftarrow{\quad\cong\quad} (W_1, (\alpha_1)_{\#}\mu) \ltimes (G_\bullet, m_{G_\bullet})$$

$$\alpha_1|_{\xi_1} \searrow \qquad \swarrow \text{canonical}$$

$$(W_1, (\alpha_1)_{\#}\mu)$$

with

$$T_1|_{\xi_1} \cong T_1|_{\alpha_1} \ltimes \sigma \circ \zeta_0^{T_1=T_2}|_{\alpha_1} \quad and \quad T_2|_{\xi_1} \cong T_2|_{\alpha_1} \ltimes (L_{\sigma \circ \zeta_0^{T_1=T_2}|_{\alpha_1}} \circ R_{\tau_2 \circ \zeta_0^{T_1}|_{\alpha_1}}),$$

and similarly

$$(Y_2, (\xi_2)_{\#}\mu) \xleftarrow{\quad\cong\quad} (W_2, (\alpha_2)_{\#}\mu) \ltimes (G_\bullet, m_{G_\bullet})$$

$$\alpha_2|_{\xi_2} \searrow \qquad \swarrow \text{canonical}$$

$$(W_2, (\alpha_2)_{\#}\mu)$$

with

$$T_1|_{\xi_2} \cong T_1|_{\alpha_2} \ltimes (L_{\sigma \circ \zeta_0^{T_1=T_2}|_{\alpha_2}} \circ R_{\tau_1 \circ \zeta_0^{T_2}|_{\alpha_2}}) \quad and \quad T_2|_{\xi_2} \cong T_2|_{\alpha_2} \ltimes \sigma \circ \zeta_0^{T_1=T_2}|_{\alpha_2}.$$

REMARKS. 1. The form of the coordinatizations given above with one action extended by a cocycle and the other by an opposite cocycle, similarly to Theorem 7.2, is a special feature of the case of two commuting transformations. It would be possible to replace it with a coordinatization by homogeneous space data (but not group data) in which both extensions are by cocycles acting on fibres on the left, for example by enlarging $G_\bullet$ to $G_\bullet \times G_\bullet$, quotienting by the diagonal subgroup $\{(g, g) : g \in G_\bullet\}$ and having ρ rotate $G_\bullet \times G_\bullet$ only in the first coordinate and σ_1, σ_2 only in the second. It is presumably this more canonical but more fiddly representation, if any, that would admit generalization to larger numbers of commuting transformations.

2. The above result describes the possible structures of the two characteristic factors individually, but some opacity remains as to how they can be joined inside $\mathbf{X}$. While we suspect that the methods of the present section can be brought to bear on this question also, we will not pursue this analysis in detail here.

The first steps of our analysis, which are essentially contained in Conze and Lesigne [12] (as well as many subsequent papers; see, for example, Furstenberg and Weiss [27] for a nice treatment of this stage of the proof), give control over the asymptotic behaviour of our averages in terms of a certain two-fold self-joining of **X**. We will then complete the proof essentially by re-applying Proposition 7.5 to certain factors of that self-joining.

We observe from the mean ergodic theorem that

$$\int_X \frac{1}{N} \sum_{n=1}^N (f_1 \circ T_1^n)(f_2 \circ T_2^n) \, \mathrm{d}\mu = \int_X f_1 \cdot \frac{1}{N} \sum_{n=1}^N (f_2 \circ (T_2 T_1^{-1})^n) \, \mathrm{d}\mu$$

$$\to \int_X f_1 \cdot \mathsf{E}_\mu(f_1 \,|\, \zeta_0^{T_1=T_2}) \, \mathrm{d}\mu = \int_{X^2} f_1 \otimes f_2 \, \mathrm{d}\mu^{\mathrm{F}}$$

where μ^{F} is the *Furstenberg self-joining*, which in this case equals $\mu \otimes_{\zeta_0^{T_1=T_2}} \mu$ (it has a much more complicated structure for larger numbers of commuting transformations which is not yet well understood; see [6, 3] for further discussion of this matter). It is easy to check that μ^{F} is invariant under the lifted transformations $T_1^{\times 2}$ and $T_2^{\times 2}$, and also under the *diagonal transformation* $\vec{T} := T_1 \times T_2$.

This self-joining now helps control our averages through the following consequence of the van der Corput estimate (for which see, for example, Bergelson [8]).

LEMMA 7.15. *If the pair of factors $\xi_i : \mathbf{X} \to \mathbf{Y}_i$ is such that $\xi_i \succsim \zeta_0^{T_1=T_2}$ and $\zeta_0^{\vec{T}} \precsim \xi_1 \times \xi_2$ then this pair is characteristic.*

Proof. This follows from a routine application of the van der Corput estimate. Suppose that $f_1, f_2 \in L^\infty(\mu)$; clearly by symmetry and iterating our argument, it suffices to prove that

$$\frac{1}{N} \sum_{n=1}^N (f_1 \circ T_1^n)(f_2 \circ T_2^n) - \frac{1}{N} \sum_{n=1}^N (\mathsf{E}_\mu(f_1 \,|\, \xi_1) \circ T_1^n)(f_2 \circ T_2^n) \to 0$$

in $L^2(\mu)$ as $N \to \infty$, and hence (taking the difference of the two sides above) that

$$\frac{1}{N} \sum_{n=1}^N (f_1 \circ T_1^n)(f_2 \circ T_2^n) \to 0$$

in $L^2(\mu)$ if $\mathsf{E}_\mu(f_1 \,|\, \xi_1) = 0$.

Letting $F_n := (f_1 \circ T_1^n)(f_2 \circ T_2^n)$, by the van der Corput estimate this will follow if we show that

$$\frac{1}{M} \sum_{h=1}^M \frac{1}{N} \sum_{n=1}^N \langle F_n, F_{n+h} \rangle \to 0$$

as $N \to \infty$ and then $M \to \infty$. Now we simply compute

$$\frac{1}{N} \sum_{n=1}^{N} \langle F_n, F_{n+h} \rangle = \frac{1}{N} \sum_{n=1}^{N} \int_X (f_1 \circ T_1^n)(f_2 \circ T_2^n)(f_1 \circ T_1^{n+h})(f_2 \circ T_2^{n+h}) \, \mathrm{d}\mu$$

$$= \int_X \frac{1}{N} \sum_{n=1}^{N} ((f_1 \cdot (f_1 \circ T_1^h)) \circ T_1^n) \cdot ((f_2 \cdot (f_2 \circ T_2^h)) \circ T_2^n) \, \mathrm{d}\mu$$

$$\to \int_{X^2} (f_1 \otimes f_2)((f_1 \circ T_1^h) \otimes (f_2 \circ T_2^h)) \, \mathrm{d}\mu^{\mathrm{F}},$$

so that if we now average also in h this converges by the mean ergodic theorem to

$$\int_{X^2} (f_1 \circ \pi_1) \cdot (f_2 \circ \pi_2) \cdot g \, \mathrm{d}\mu^{\mathrm{F}}$$

for some $\vec{T}$-invariant function g. Finally, this last integral is zero if we have $\mathsf{E}_{\mu^{\mathrm{F}}}(f_1 \circ \pi_1 \,|\, \pi_2 \vee \zeta_0^{\vec{T}}) = 0$, and this follows from our assumptions and the relative independence of π_1 and π_2 over $\zeta_0^{T_1=T_2} \circ \pi_1$ under μ^{F}. $\blacksquare$

Since on the other hand Theorem 5.12 tells us that $\zeta_0^{\vec{T}} \precsim (\zeta_{1/\zeta_0^{T_1=T_2}}^{T_1} \circ \pi_1) \vee (\zeta_{1/\zeta_0^{T_1=T_2}}^{T_2} \circ \pi_2)$, we can deduce the following at once.

COROLLARY 7.16 (Reduction to isometric extensions of isotropy factors). *There is a characteristic pair of factors satisfying* $\xi_i \precsim \zeta_{1/\zeta_0^{T_1=T_2}}^{T_i}$ *for* $i = 1, 2$. $\blacksquare$

We can now present the factors introduced above as another instance of the situation described before Proposition 7.5, and have that proposition do the heavy lifting we need again here.

To see this, we define a system of *three* commuting transformations on the Furstenberg self-joining. Let $\mathbf{X}^{\mathrm{F}}$ be the $\mathbb{Z}^3$-system $(X^2, \mu^{\mathrm{F}}, S_1, S_2, S_3)$ obtained by setting $S_1 := T_1^{\times 2} \vec{T}^{-1} = \mathrm{id}_X \times (T_1 T_2^{-1})$, $S_2 := T_2^{\times 2} \vec{T}^{-1} = (T_2 T_1^{-1}) \times \mathrm{id}_X$ and $S_3 := \vec{T}$. We observe directly from the definition of μ^{F} that for $i = 1, 2$ the coordinate projection $\pi_i : X^2 \to X$ is equivalent to $\zeta_0^{S_i}$.

Now let $\mathbf{W}_i$ be $(\zeta_0^{T_1=T_2} \vee \zeta_0^{T_i})(\mathbf{X})$ for $i = 1, 2$ and let $\mathbf{Z}_i$ be the target of the maximal subextension of $\zeta_0^{T_1=T_2} \vee \zeta_0^{T_i} : \mathbf{X} \to \mathbf{W}_i$ that is isometric for the restriction of T_i. Let
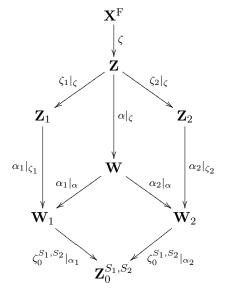
$$\alpha_i := (\zeta_0^{T_1=T_2} \vee \zeta_0^{T_i}) \circ \pi_i = \zeta_0^{S_1, S_2} \vee \zeta_0^{S_i, S_3} : \mathbf{X}^{\mathrm{F}} \to \mathbf{W}_i$$

and

$$\zeta_i := \zeta_{1/\alpha_i}^{T_i} \circ \pi_i : \mathbf{X}^{\mathrm{F}} \to \mathbf{Z}_i,$$

and let $\zeta : \mathbf{X}^{\mathrm{F}} \to \mathbf{Z}$ and $\alpha : \mathbf{X}^{\mathrm{F}} \to \mathbf{W}$ be the joinings $\zeta_1 \vee \zeta_2$ and $\alpha_1 \vee \alpha_2$ respectively.

It is now routine to check from the basic results above that these data satisfy the same conditions as were needed for Proposition 7.5 with $\mathbf{X}^{\mathrm{F}}$ in place of $\mathbf{X}$ and S_i in place of T_i: these factors are once again arranged as in the commutative diagram

where we observe easily from the structure of $\mu^{\mathrm{F}} = \mu \otimes_{\zeta_0^{T_1=T_2}} \mu$ that $\zeta_0^{S_1=S_2} \simeq \zeta_0^{T_1=T_2} \circ \pi_1 \simeq \zeta_0^{T_1=T_2} \circ \pi_2$. In addition, the transformations S_i enjoy the following properties:

- S_i restricts to the identity on $\mathbf{Z}_i$ and the factors beneath it, while acting relatively ergodically on the extension $\zeta_0^{S_1,S_2}|_{\zeta_{3-i}} : \mathbf{Z}_{3-i} \to \mathbf{Z}_0^{S_1,S_2}$, for $i = 1, 2$;
- the extensions $\zeta_0^{S_1,S_2}|_{\alpha_i} : \mathbf{W}_i \to \mathbf{Z}_0^{S_1,S_2}$ are relatively invariant for the restriction of S_3, and the extensions $\alpha_i|_{\zeta_i} : \mathbf{Z}_i \to \mathbf{W}_i$ are relatively ergodic and isometric for the restriction of S_3.
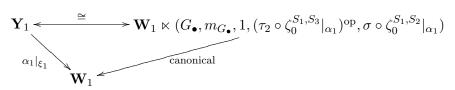
We can therefore apply Proposition 7.5 to these systems and maps to deduce the following.

PROPOSITION 7.17. *There are intermediate factors* $\mathbf{Z}_i \xrightarrow{\xi_i|_{\zeta_i}} \mathbf{Y}_i \xrightarrow{\alpha_i|_{\xi_i}} \mathbf{W}_i$ *factorizing* $\alpha_i|_{\zeta_i}$ *such that there are S-invariant compact group data* $G_\bullet$ *and cocycle-sections*

$\sigma : Z_0^{S_1,S_2} \to G_\bullet$ *that is* $S_3|_{\zeta_0^{S_1,S_2}}$-*relatively ergodic,*

$\tau_1 : Z_0^{S_2,S_3} \to G_\bullet$ *that is* $S_1|_{\zeta_0^{S_2,S_3}}$-*relatively ergodic and*

$\tau_2 : Z_0^{S_2,S_3} \to G_\bullet$ *that is* $S_2|_{\zeta_0^{S_1,S_3}}$*-relatively ergodic*

so that we can coordinatize the actions of the transformations S *as*

$$\mathbf{Y}_1 \xleftarrow{\quad\cong\quad} \mathbf{W}_1 \ltimes (G_\bullet, m_{G_\bullet}, 1, (\tau_2 \circ \zeta_0^{S_1,S_3}|_{\alpha_1})^{\mathrm{op}}, \sigma \circ \zeta_0^{S_1,S_2}|_{\alpha_1})$$

$$\alpha_1|_{\xi_1} \searrow \quad \mathbf{W}_1 \quad \swarrow \text{canonical}$$

and

$$\mathbf{Y}_2 \xleftarrow{\quad\cong\quad} \mathbf{W}_2 \ltimes (G_\bullet, m_{G_\bullet}, (\tau_1 \circ \zeta_0^{S_2,S_3}|_{\alpha_2})^{\mathrm{op}}, 1, \sigma \circ \zeta_0^{S_1,S_2}|_{\alpha_2})$$

$$\alpha_2|_{\xi_2} \searrow \quad \mathbf{W}_2 \quad \swarrow \text{canonical}$$

and such that the $(S_3 = \vec{T})$*-invariant factor of* $\mathbf{X}^{\mathrm{F}}$ *is contained in* $\xi_1 \vee \xi_2$. ∎

Proof of Theorem 7.14. Again this follows simply by unpacking the notation of the above result: the tower of factors $\alpha_i|_{\xi_i} : \mathbf{Y}_i \to \mathbf{W}_i$ of $\mathbf{X}^{\mathrm{F}}$ are all actually contained within the ith coordinate projection $\pi_i : \mathbf{X}^{\mathrm{F}} \to \mathbf{X}_i$, and by definition we have $T_i = (S_i S_3)|_{\pi_i}$ for $i = 1, 2$, and therefore the above coordinatization of the action of S restricted to the tower of factors $\alpha_1|_{\xi_1} : \mathbf{Y}_1 \to \mathbf{W}_1$ converts into a coordinatization description of $T_1|_{\xi_1}$ as

$$(w, g) \mapsto (T_1|_{\alpha_1}(w), \sigma(\zeta_0^{T_1 = T_2}(w)) \cdot g)$$

and of $T_2|_{\xi_1}$ as

$$(w, g) \mapsto (T_1|_{\alpha_1}(w), \sigma(\zeta_0^{T_1 = T_2}(w)) \cdot g \cdot \tau_2(\zeta_0^{T_1}(w))),$$

(since $\zeta_0^{S_1,S_3} \simeq \zeta_0^{T_1} \circ \pi_1$) and similarly for $T_1|_{\xi_2}$ and $T_2|_{\xi_2}$. This completes the proof. ∎

REMARKS. 1. In fact, it is relatively easy to see by checking functions f_1, f_2 that are constructed from measurable selections of representative functions on the compact fibre groups $G_\bullet$ that the characteristic pair of factors ξ_i that we have now isolated is minimal, in that any other characteristic pair ξ_1', ξ_2' satisfies $\xi_i \precsim \xi_i'$.

2. The results of the preceding subsection also give a precise picture of the $\vec{T}$-invariant factor of $\mathbf{X}^{\mathrm{F}}$ in terms of a diagonal Mackey group and trivial Mackey section for the joining of the above coordinatizations of $\mathbf{Y}_1$ and $\mathbf{Y}_2$ inside $\mathbf{X}^{\mathrm{F}}$; we omit these details here.

8. Further questions. This paper leaves open the obvious question of how to generalize the analysis of Section 7 to describe in similar detail

- the possible joint distributions among a larger collection of isotropy factors $\zeta_0^{T^{\restriction_{\Gamma_i}}} : \mathbf{X} \to \mathbf{Z}_0^{T^{\restriction_{\Gamma_i}}}$ for $\Gamma_1, \ldots, \Gamma_d \leq \mathbb{Z}^d$;
- the possible structures of characteristic factors (and, relatedly, Furstenberg self-joinings) for larger commuting tuples of transformations (or commuting actions of some other fixed group).

On the one hand, it seems likely that the machinery of extensions by measurably-varying compact homogeneous spaces will be quite essential to any further developments in this area. On the other, I suspect that even the next cases to consider in the natural hierarchy (joint distributions of four isotropy factors, or characteristic factors for three commuting transformations) become much more complicated, and it may be in general too much to ask for the kind of precision that we obtained in Theorems 1.1 and 1.2.

There is an alternative viewpoint on questions such as these that may be more tractable. Instead of asking about exact joint distributions or characteristic factors for an initially given system, if we allow ourselves the freedom to pass to any extension of that system, matters sometimes improve considerably. Indeed, the new proofs of convergence for linear non-conventional averages in [6] and of Furstenberg and Katznelson's associated multidimensional multiple recurrence theorem in [3] both relied on procedures for passing from an initially given system to some extension in which the relevant characteristic factors and their joint distributions could be described much more simply.

The constructions of those papers were abstract enough to work without any of the machinery of homogeneous-space-data extensions. However, for further applications of this idea, in particular to the problem of convergence of related 'polynomial non-conventional averages' such as

$$\frac{1}{N} \sum_{n=1}^{N} (f_1 \circ T_1^{n^2})(f_1 \circ T_1^{n^2} T_2^n)$$

(discussed, for example, by Bergelson and Leibman in [9]), it seems likely that these more delicate tools will be necessary. In the forthcoming works [4, 5] we will make such an analysis allowing ourselves to pass to extensions, focusing on what improved characteristic factors can be found while retaining some given algebraic relations among the transformations involved, and will then use this to prove convergence of the above polynomial averages in $L^2(\mu)$ as $N \to \infty$.

In addition to these quite specialized applications, let us also mention that there seem to be further issues on the general behaviour of extensions by homogeneous space data to be explored. For example, in [7] it is shown that some of the machinery of Furstenberg and Zimmer concerning finite-rank modules of an extension can be extended to the setting in which the exten-

sion is relatively finite measure-preserving, but the base, while ergodic, is only assumed to be non-singular (that is, its measure is only quasi-invariant). In that paper this machinery is needed for the proof of a result about the lifting of the 'multiplier property' through certain kinds of extension, and this will not require that these general results on finite-rank modules be pushed very far. However, it might be interesting to examine whether that development can be easily recovered without the assumption of ergodicity of the base, using a version of the formalism of the present paper.

References

[1] W. Arveson, *An Invitation to C^*-Algebras*, Springer, 1976.

[2] L. Auslander, L. Green, and F. Hahn, *Flows on Homogeneous Spaces*, Ann. of Math. Stud. 53, Princeton Univ. Press, Princeton, NJ, 1963.

[3] T. Austin, *Deducing the multidimensional Szemerédi Theorem from an infinitary removal lemma*, J. Anal. Math., to appear.

[4] —, *Pleasant extensions retaining algebraic structure, I*, arXiv.org:0905.0518.

[5] —, *Pleasant extensions retaining algebraic structure, II*, arXiv.org:0910.0907.

[6] —, *On the norm convergence of nonconventional ergodic averages*, Ergodic Theory Dynam. Systems 30 (2009), 321–338.

[7] T. Austin and M. Lemańczyk, *Relatively finite measure-preserving extensions and lifting multipliers by Rokhlin cocycles*, J. Fixed Point Theory Appl. 6 (2009), 115–131.

[8] V. Bergelson, *Weakly mixing PET*, Ergodic Theory Dynam. Systems 7 (1987), 337–349.

[9] V. Bergelson and A. Leibman, *Set-polynomials and polynomial extension of the Hales–Jewett theorem*, Ann. of Math. (2) 150 (1999), 33–75.

[10] V. Bergelson, T. Tao, and T. Ziegler, *An inverse theorem for the uniformity seminorms associated with the action of $\mathbb{F}_p^\infty$*, arXiv.org:0901.2602, 2009.

[11] T. Bröcker and T. tom Dieck, *Representations of Compact Lie Groups*, Springer, 1985.

[12] J.-P. Conze et E. Lesigne, *Théorèmes ergodiques pour des mesures diagonales*, Bull. Soc. Math. France 112 (1984), 143–175.

[13] —, —, *Sur un théorème ergodique pour des mesures diagonales*, in: Probabilités, Publ. Inst. Rech. Math. Rennes 1987, Univ. Rennes I, Rennes, 1988, 1–31.

[14] —, —, *Sur un théorème ergodique pour des mesures diagonales*, C. R. Acad. Sci. Paris Sér. I Math. 306 (1988), 491–493.

[15] J.-P. Conze and A. Raugi, *On the ergodic decomposition for a cocycle*, Colloq. Math. 117 (2009), 121–156.

[16] T. Downarowicz, *Minimal models for noninvertible and not uniquely ergodic systems*, Israel J. Math. 156 (2006) 93–110.

[17] R. Engelking, *General Topology*, 2nd ed., Sigma Ser. Pure Math. 6, Heldermann, Berlin, 1989.

[18] D. Fisher, D. W. Morris, and K. Whyte, *Nonergodic actions, cocycles and superrigidity*, New York J. Math. 10 (2004), 249–269.

[19] N. Frantzikinakis and B. Kra, *Convergence of multiple ergodic averages for some commuting transformations*, Ergodic Theory Dynam. Systems 25 (2005), 799–809.

[20] D. H. Fremlin, *Measure Theory, Volume 3: Measure Algebras*, Torres Fremlin, Colchester, 2004.

[21] —, *Measure Theory, Volume 4: Topological Measure Theory*, Torres Fremlin, Colchester, 2005.

[22] H. Furstenberg, *Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Anal. Math. 31 (1977), 204–256.

[23] —, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton Univ. Press, Princeton, NJ, 1981.

[24] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi Theorem for commuting transformations*, J. Anal. Math. 34 (1978), 275–291.

[25] —, —, *An ergodic Szemerédi theorem for IP-systems and combinatorial theory*, ibid. 45 (1985), 117–168.

[26] —, —, *A density version of the Hales–Jewett theorem*, ibid. 57 (1991), 64–119.

[27] H. Furstenberg and B. Weiss, *A mean ergodic theorem for $\frac{1}{N} \sum_{n=1}^{N} f(T^n x) g(T^{n^2} x)$*, in: V. Bergelson et al. (eds.), Convergence in Ergodic Theory and Probability, de Gruyter, Berlin, 1996, 193–227.

[28] E. Glasner, *Ergodic Theory via Joinings*, Amer. Math. Soc., Providence, 2003.

[29] A. Guichardet, *Symmetric Hilbert Spaces and Related Topics*, Springer, 1972.

[30] B. Host and B. Kra, *Convergence of Conze–Lesigne averages*, Ergodic Theory Dynam. Systems 21 (2001), 493–509.

[31] —, —, *Nonconventional ergodic averages and nilmanifolds*, Ann. of Math. 161 (2005), 397–488.

[32] O. Kallenberg, *Foundations of Modern Probability*, 2nd ed., Springer, New York, 2002.

[33] G. W. Mackey, *Ergodic theory and virtual groups*, Math. Ann. 166 (1966), 187–207.

[34] —, *The Theory of Unitary Group Representations*, Univ. of Chicago Press, Chicago, IL 1976,

[35] D. Meiri, *Generalized correlation sequences*, Master's thesis, Tel Aviv Univ.; http://taalul.com/David/Math/ma.pdf, 1990.

[36] M. K. Mentzen, *Ergodic properties of group extensions of dynamical systems with discrete spectra*, Studia Math. 101 (1991), 19–31.

[37] D. Newton, *On canonical factors of ergodic dynamical systems*, J. London Math. Soc. (2) 19 (1979), 129–136.

[38] D. J. Rudolph, *Eigenfunctions of $T \times S$ and the Conze–Lesigne algebra*, in: Ergodic Theory and its Connections with Harmonic Analysis (Alexandria, 1993), London Math. Soc. Lecture Note Ser. 205, Cambridge Univ. Press, Cambridge, 1995, 369–432.

[39] K. Schmidt, *Dynamical Systems of Algebraic Origin*, Progr. Math. 128, Birkhäuser, Basel, 1995.

[40] R. Schmidt, *Subgroup Lattices of Groups*, de Gruyter Exp. Math. 14, de Gruyter, Berlin, 1994.

[41] A. N. Starkov, *Dynamical Systems on Homogeneous Spaces*, Transl. Math. Monogr. 190, Amer. Math. Soc., Providence, RI, 2000,

[42] T. Tao, *Norm convergence of multiple ergodic averages for commuting transformations*, Ergodic Theory Dynam. Systems 28 (2008) 657–688.

[43] H. Towsner, *Convergence of diagonal ergodic averages*, Ergodic Theory Dynam. Systems 29 (2009), 1309–1326.

[44] W. A. Veech, *A criterion for a process to be prime*, Monatsh. Math. 94 (1982), 335–341.

[45] Q. Zhang, *On convergence of the averages* $(1/N)\sum_{n=1}^{N} f_1(R^n x)f_2(S^n x)f_3(T^n x)$, ibid. 122 (1996), 275–300.

[46] T. Ziegler, *A non-conventional ergodic theorem for a nilsystem*, Ergodic Theory Dynam. Systems 25 (2005), 1357–1370.

[47] —, *Universal characteristic factors and Furstenberg averages*, J. Amer. Math. Soc. 20 (2007), 53–97.

[48] R. J. Zimmer, *Ergodic actions with generalized discrete spectrum*, Illinois J. Math. 20 (1976), 555–588.

[49] —, *Extensions of ergodic group actions*, Illinois J. Math. 20 (1976), 373–409.

[50] —, *Ergodic Theory and Semisimple Groups*, Monogr. Math. 81, Birkhäuser, Basel, 1984.

Tim Austin
Department of Mathematics
University of California
Los Angeles, CA 90095-1555, U.S.A.
E-mail: timaustin@math.ucla.edu
http://www.math.ucla.edu/~timaustin