NONCONVENTIONAL ERGODIC AVERAGES AND NILMANIFOLDS

BERNARD HOST AND BRYNA KRA

Abstract. We study the L^2-convergence of two types of ergodic averages. The first is the average of a product of functions evaluated at return times along arithmetic progressions, such as the expressions appearing in Furstenberg's proof of Szemerédi's Theorem. The second average is taken along cubes whose sizes tend to $+\infty$. For each average, we show that it is sufficient to prove the convergence for special systems, the *characteristic factors*. We build these factors in a general way, independent of the type of the average. To each of these factors we associate a natural group of transformations and give them the structure of a nilmanifold. From the second convergence result we derive a combinatorial interpretation for the arithmetic structure inside a set of integers of positive upper density.

1. Introduction

1.1. **The averages.** A beautiful result in combinatorial number theory is Szemerédi's Theorem, which states that a set of integers with positive upper density contains arithmetic progressions of arbitrary length. Furstenberg [F77] proved Szemerédi's theorem via an ergodic theorem:

Theorem (Furstenberg)**.** *Let $(X, \mathcal{X}, \mu, T)$ be a measure preserving probability system and let $A \in \mathcal{X}$ be a set of positive measure. Then for every integer $k \geq 1$,*

$$\liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mu\big(A\cap T^{-n}A\cap T^{-2n}A\cap\cdots\cap T^{-kn}A\big)\ >0\ .$$

It is natural to ask about the convergence of these averages, and more generally about the convergence in $L^2(\mu)$ of the averages of products of bounded functions along an arithmetic progression of length k for an arbitrary integer $k \geq 1$. We prove:

Theorem 1.1. *Let $(X, \mathcal{X}, \mu, T)$ be an invertible measure preserving probability system, $k \geq 1$ be an integer, and let f_j, $1 \leq j \leq k$, be k bounded measurable functions on X. Then*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f_1(T^n x) f_2(T^{2n}x)\dots f_k(T^{kn}x) \tag{1}$$

exists in $L^2(X)$.

The case $k = 1$ is the standard ergodic theorem of von Neumann. Furstenberg [F77] proved this for $k = 2$ by reducing to the case where X is an ergodic rotation and using the Fourier transform to prove convergence. The existence of

Date: November 7, 2003.

limits for $k = 3$ with an added hypothesis that the system is totally ergodic was shown by Conze and Lesigne in a series of papers ([CL84], [CL87] and [CL88]) and in the general case by Host and Kra [HK01]. Ziegler [Zie02b] has shown the existence in a special case when $k = 4$.

If one assumes that T is weakly mixing, Furstenberg [F77] proved that for every k the limit (1) exists and is constant. However, without the assumption of weak mixing one can easily show that the limit need not be constant and proving convergence becomes much more difficult. Nonconventional averages are those for which even if the system is ergodic, the limit is not necessarily constant. This is the case for $k \geq 3$ in Equation (1).

Some related convergence problems have also been studied by Bourgain [Bo89] and Furstenberg and Weiss [FW96].

We also study the related average of the product of $2^k - 1$ functions taken along *combinatorial cubes* whose sizes tend to $+\infty$. The general formulation of the theorem is a bit intricate and so for clarity we begin by stating a particular case, which was proven in [HK04].

Theorem. *Let $(X, \mathcal{X}, \mu, T)$ be an invertible measure preserving probability system and let f_j, $1 \leq j \leq 7$, be 7 bounded measurable functions on X. Then the averages over $(m, n, p) \in [M, M'] \times [N, N'] \times [P, P']$ of*

$$f_1(T^m x) f_2(T^n x) f_3(T^{m+n} x) f_4(T^p x) f_5(T^{m+p} x) f_6(T^{n+p} x) f_7(T^{m+n+p} x)$$

converge in $L^2(\mu)$ as $M' - M, N' - N$ and $P' - P$ tend to $+\infty$.

Notation. For an integer $k > 0$, let $V_k = \{0, 1\}^k$. The elements of V_k are written without commas or parentheses. For $\epsilon = \epsilon_1 \epsilon_2 \ldots \epsilon_k \in V_k$ and $n = (n_1, n_2, \ldots, n_k) \in \mathbb{Z}^k$, we write

$$\epsilon \cdot n = \epsilon_1 n_1 + \epsilon_2 n_2 + \cdots + \epsilon_k n_k \ .$$

We use $\mathbf{0}$ to denote the element $00\ldots0$ of V_k and set $V_k^* = V_k \setminus \{\mathbf{0}\}$.

We generalize the above theorem to higher dimensions and show:

Theorem 1.2. *Let $(X, \mathcal{X}, \mu, T)$ be an invertible measure preserving probability system, $k \geq 1$ be an integer, and let f_ϵ, $\epsilon \in V_k^*$, be $2^k - 1$ bounded functions on X. Then the averages*

$$\prod_{i=1}^{k} \frac{1}{N_i - M_i} \cdot \sum_{n \in [M_1, N_1) \times \cdots \times [M_k, N_k)} \prod_{\epsilon \in V_k^*} f_\epsilon(T^{\epsilon \cdot n} x) \tag{2}$$

converge in $L^2(X)$ as $N_1 - M_1, N_2 - M_2, \ldots, N_k - M_k$ tend to $+\infty$.

When restricting Theorem 1.2 to the indicator function of a measurable set, we have the following lower bound for these averages:

Theorem 1.3. *Let $(X, \mathcal{X}, \mu, T)$ be an invertible measure preserving probability system and let $A \in \mathcal{X}$. Then the limit of the averages*

$$\prod_{i=1}^{k} \frac{1}{N_i - M_i} \cdot \sum_{n \in [M_1, N_1) \times \cdots \times [M_k, N_k)} \mu\Big(\bigcap_{\epsilon \in V_k} T^{\epsilon \cdot n} A \Big)$$

exists and is greater than or equal to $\mu(A)^{2^k}$ when $N_1 - M_1, N_2 - M_2, \ldots, N_k - M_k$ tend to $+\infty$.

For $k = 1$, Khintchine [K34] proved the existence of the limit along with the associated lower bound, for $k = 2$ this was proven by Bergelson [Be00], and for $k = 3$ by the authors in [HK04].

1.2. **Combinatorial Interpretation.** We recall that the upper density $\overline{d}(A)$ of a set $A \subset \mathbb{N}$ is defined to be

$$\overline{d}(A) = \limsup_{N\to\infty} \frac{1}{N} |A \cap \{1, 2, \dots, N\}| \ .$$

Furstenberg's Theorem as well as Theorem 1.3 have combinatorial interpretations for subsets of $\mathbb{N}$ with positive upper density. Furstenberg's Theorem is equivalent to Szemerédi's Theorem. In order to state the combinatorial counterpart of Theorem 1.3 we recall the definition of a syndetic set.

Definition 1.4. Let Γ be an abelian group. A subset E of Γ is *syndetic* if there exists a finite subset D of Γ such that $E + D = \Gamma$.

When $\Gamma = \mathbb{Z}^d$, this definition becomes:
A subset E of $\mathbb{Z}^d$ is *syndetic* if there exist an integer $N > 0$ such that

$$E \cap \big([M_1, M_1 + N] \times [M_2, M_2 + N] \times \dots \times [M_k, M_k + N]\big) \neq \emptyset$$

for every $M_1, M_2, \dots, M_k \in \mathbb{Z}$.

When A is a subset of $\mathbb{Z}$ and m is an integer, we let $A + m$ denote the set $\{a + m : a \in A\}$. From Theorem 1.3 we have:

Theorem 1.5. *Let $A \subset \mathbb{Z}$ with $\overline{d}(A) > \delta > 0$ and let $k \geq 1$ be an integer. The set of $n = (n_1, n_2, \dots, n_k) \in \mathbb{Z}^k$ so that*

$$\overline{d}\Big(\bigcap_{\epsilon\in V_k} (A + \epsilon \cdot n)\Big) \geq \delta^{2^k}$$

is syndetic.

Both the averages along arithmetic progressions and along cubes are concerned with demonstrating the existence of some arithmetic structure inside a set of positive upper density. Moreover, an arithmetic progression can be seen inside a cube with all indices n_j equal. However, the end result is rather different. In Theorem 1.5, we have an explicit lower bound that is optimal, but it is impossible to have any control over the size of the syndetic constant, as can be seen with elementary examples such as rotations. This means that this result does not have a finite version. On the other hand, Szemerédi's Theorem can be expressed in purely finite terms, but the problem of finding the optimal lower bound is open.

1.3. **Characteristic factors.** The method of *characteristic factors* is classical since Furstenberg's work [F77], even though this term only appeared explicitly more recently [FW96]. For the problems we consider, this method consists in finding an appropriate factor of the given system, referred to as the characteristic factor, so that the limit behavior of the averages remains unchanged when each function is replaced by its conditional expectation on this factor. Then it suffices to prove the convergence when this factor is substituted for the original system, which is facilitated when the factor has a "simple" description.

We follow this general strategy, with the difference that we focus more on the procedure of building characteristic factors than on the particular type of average currently under study. A standard method for finding characteristic factors is an iterated use of the van der Corput Lemma, with the number of steps increasing with the complexity of the averages. For each system and each integer k, we build a factor in a way that reflects k successive uses of the van der Corput Lemma. This factor is almost automatically characteristic for averages of the same "complexity". For example, the k-dimensional average along cubes has the same characteristic factor as the average along arithmetic progressions of length $k-1$. Our construction involves the definition of a "cubic structure" of order k on the system (see Section 3), meaning a measure on its 2^kth Cartesian power. Roughly speaking, the factor we build is the smallest possible factor with this structure (see Section 4).

The bulk of the paper (Sections 5–10), and also the most technical portion, is devoted to the description of these factors. The initial idea is natural: For each of these factors we associate the group of transformations which preserve the natural cubic structure alluded to above (Section 5). This group is nilpotent. We then conclude (Theorems 10.3 and 10.5) that for a sufficiently large (for our purposes) class of systems, this group is a Lie group and acts transitively on the space. Therefore, the constructed system is a *nilsystem*. In Section 11, we show that the cubic structure alluded to above has a simple description for these systems.

Given this construction, we return to the original average along arithmetic progressions in Section 12 and along cubes in Section 13 and show that the characteristic factors of these averages are exactly those which we have constructed. A posteriori, the role played by the nilpotent structure is not surprising: for a k-step nilsystem, the $(k+1)$st term T^kx of an arithmetic progression is constrained by the first k terms $x, Tx, \dots, T^{k-1}x$. A similar property holds for the combinatorial structure considered in Theorem 1.2.

Convergence then follows easily from general properties of nilmanifolds. Finally, we derive a combinatorial result from the convergence theorems.

1.4. **Open questions.** There are at least two possible generalizations of Theorem 1.1. The first one consists in substituting integer valued polynomials $p_1(n)$, $p_2(n)$, $\dots$, $p_k(n)$ for the linear terms $n, 2n, \dots, kn$ in the averages (1). With an added hypothesis, either that the system is totally ergodic or that all the polynomials have degree > 1, we prove convergence of these polynomial averages in [HK02]. The case that the system is not totally ergodic and at least one polynomial is of degree one and some other has higher degree remains open.

Another more ambitious generalization is to consider commuting transformations $T_1, T_2, \dots, T_k$ instead of $T, T^2, \dots, T^k$. Characteristic factors for this problem are unknown.

The question of convergence almost everywhere is completely different and can not be addressed by the methods of this paper.

1.5. **About the organization of the paper.** We begin (Section 2) by introducing the notation relative to 2^k-Cartesian powers. We have postponed to four appendices some definitions and results needed, which do not have a natural place in the main text. Appendix A deals with properties of Polish groups and Lie groups, Appendix B with nilsystems, Appendix C with cocycles and Appendix D with the van der Corput Lemma. Most of the results presented in these Appendices are classical.

2. General notation

2.1. **Cubes.** Throughout, we use 2^k-Cartesian powers of spaces for an integer $k > 0$ and need some shorthand notation.

Let X be a set. For an integer $k \geq 0$, we write $X^{[k]} = X^{2^k}$. For $k > 0$, we use the sets V_k introduced above to index the coordinates of elements of this space, which are written $\mathbf{x} = (x_\epsilon : \epsilon \in V_k)$.

When f_ϵ, $\epsilon \in V_k$, are 2^k real or complex valued functions on the set X, we define a function $\bigotimes_{\epsilon \in V_k} f_\epsilon$ on $X^{[k]}$ by

$$\bigotimes_{\epsilon \in V_k} f_\epsilon(\mathbf{x}) = \prod_{\epsilon \in V_k} f_\epsilon(x_\epsilon) \ .$$

When $\phi \colon X \to Y$ is a map, we write $\phi^{[k]} \colon X^{[k]} \to Y^{[k]}$ for the map given by $\big(\phi^{[k]}(\mathbf{x})\big)_\epsilon = \phi(x_\epsilon)$ for $\epsilon \in V_k$.

We often identify $X^{[k+1]}$ with $X^{[k]} \times X^{[k]}$. In this case, we write $\mathbf{x} = (\mathbf{x}', \mathbf{x}'')$ for a point of $X^{[k+1]}$, where $\mathbf{x}', \mathbf{x}'' \in X^{[k]}$ are defined by

$$x'_\epsilon = x_{\epsilon 0} \text{ and } x''_\epsilon = x_{\epsilon 1}$$

for $\epsilon \in V_k$ and $\epsilon 0$ and $\epsilon 1$ are the elements of V_{k+1} given by

$$(\epsilon 0)_j = (\epsilon 1)_j = \epsilon_j \text{ for } 1 \leq j \leq k \ ; \ (\epsilon 0)_{k+1} = 0 \text{ and } (\epsilon 1)_{k+1} = 1 \ .$$

The maps $\mathbf{x} \mapsto \mathbf{x}'$ and $\mathbf{x} \mapsto \mathbf{x}''$ are called the projections on the first and second side, respectively.

It is convenient to view V_k as indexing the set of vertices of the cube of dimension k, making the use of the geometric words 'side', 'face', and 'edge' for particular subsets of V_k natural. More precisely, for $0 \leq \ell \leq k$, J a subset of $\{1, \dots, k\}$ with cardinality $k - \ell$ and $\eta \in \{0,1\}^J$, the subset

$$\alpha = \{\epsilon \in V_k : \epsilon_j = \eta_j \text{ for every } j \in J\}$$

of V_k is called a *face of dimension ℓ of* V_k, or more succinctly, an ℓ-face. Thus V_k has one face of dimension k, namely V_k itself. It has $2k$ faces of dimension $k-1$, called the *sides*, and has $k2^{k-1}$ faces of dimension 1, called *edges*. It has 2^k sides of dimension 0, each consisting in one element of V_k and called a *vertex*. We often identify the vertex $\{\epsilon\}$ with the element ϵ of V_k.

Let α be an ℓ-face of V_k. Enumerating the elements of α and of V_ℓ in lexicographic order gives a natural bijection between α and V_ℓ. This bijection maps the faces of V_k included in α to the faces of V_ℓ. Moreover, for every set X, it induces a map from $X^{[k]}$ onto $X^{[\ell]}$. We denote this map by $\xi^{[k]}_{X,\alpha}$, or $\xi^{[k]}_\alpha$ when there is no ambiguity about the space X. When α is any face, we call it a *face projection* and when α is a side, we call it a *a side projection*. This is a natural generalization of the projections on the first and second sides.

The symmetries of the cube V_k play an important role in the sequel. We write $\mathcal{S}_k$ for the group of bijections of V_k onto itself which maps every face to a face (of the same dimension, of course). This group is isomorphic to the group of the 'geometric cube' of dimension k, meaning the group of isometries of $\mathbb{R}^k$ preserving the unit cube. It is spanned by digit permutations and reflections, which we now define.

Definition 2.1. Let τ be a permutation of $\{1, \dots, k\}$. The permutation σ of V_k given for $\epsilon \in V_k$ by

$$\big(\sigma(\epsilon)\big)_j = \epsilon_{\tau(j)} \quad \text{for } 1 \leq j \leq k$$

is called a *digit permutation*.

Let $i \in \{1, \dots k\}$. The permutation σ of V_k given for $\epsilon \in V_k$ by

$$\big(\sigma(\epsilon)\big)_j = \epsilon_j \text{ when } j \neq i \quad\text{and}\quad \big(\sigma(\epsilon)\big)_i = 1 - \epsilon_i$$

is called a *reflection*.

For any set X, the group $\mathcal{S}_k$ acts on $X^{[k]}$ by permuting the coordinates: for $\sigma \in \mathcal{S}_k$, we write $\sigma_*\colon X^{[k]} \to X^{[k]}$ for the map given by

$$\big(\sigma_*(x)\big)_\epsilon = x_{\sigma(\epsilon)} \text{ for every } \epsilon \in V_k\ .$$

When σ is a digit permutation (respectively, a reflection) we also call the associated map σ_* a digit permutation (respectively, a reflection).

2.2. **Probability spaces.** In general, we write (X, μ) for a probability space, omitting the σ-algebra. When needed, the σ-algebra of the probability space (X, μ) is written $\mathcal{X}$. By a *system*, we mean a probability space (X, μ) endowed with an invertible, bi-measurable, measure preserving transformation $T\colon X \to X$ and we write the system as (X, μ, T).

For a system (X, μ, T), we use the word *factor* with two different meanings: it is either a T-invariant sub-σ-algebra $\mathcal{Y}$ of $\mathcal{X}$ or a system (Y, ν, S) and a measurable map $\pi\colon X \to Y$ such that $\pi\mu = \nu$ and $S \circ \pi = \pi \circ T$. We often identify the σ-algebra $\mathcal{Y}$ of Y with the invariant sub-σ-algebra $\pi^{-1}(\mathcal{Y})$ of $\mathcal{X}$.

All locally compact groups are implicitly assumed to be metrizable and endowed with their Borel σ-algebras. Every compact group G is endowed with its Haar measure, denoted by m_G.

We write $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. We call a compact abelian group isomorphic to $\mathbb{T}^d$ for some integer $d \geq 0$ a *torus*, with the convention that $\mathbb{T}^0$ is the trivial group.

Let G be a locally compact abelian group. By a *character* of G we mean a continuous group homomorphism from G to either the torus $\mathbb{T}$ or the circle group $\mathcal{S}^1$. The characters of G form a group $\widehat{G}$ called the dual group of G. We use either additive or multiplicative notation for $\widehat{G}$.

For a compact abelian group Z and $t \in Z$, we write (Z, t) for the probability space (Z, m_Z), endowed with the transformation given by $z \mapsto tz$. A system of this kind is called a *rotation*.

3. Construction of the measures

Throughout this section, (X, μ, T) denotes an ergodic system.

3.1. **Definition of the measures.** We define by induction a $T^{[k]}$-invariant measure $\mu^{[k]}$ on $X^{[k]}$ for every integer $k \geq 0$.

Set $X^{[0]} = X$, $T^{[0]} = T$ and $\mu^{[0]} = \mu$. Assume that $\mu^{[k]}$ is defined. Let $\mathcal{I}^{[k]}$ denote the $T^{[k]}$ invariant σ-algebra of $(X^{[k]}, \mu^{[k]}, T^{[k]})$. Identifying $X^{[k+1]}$ with $X^{[k]} \times X^{[k]}$ as explained above, we define the system $(X^{[k+1]}, \mu^{[k+1]}, T^{[k+1]})$ to be the relatively independent joining of two copies of $(X^{[k]}, \mu^{[k]}, T^{[k]})$ over $\mathcal{I}^{[k]}$.

This means that when f_ϵ, $\epsilon \in V_{k+1}$, are bounded functions on X,

$$\int_{X^{[k+1]}} \bigotimes_{\epsilon\in V_{k+1}} f_\epsilon \, d\mu^{[k+1]} = \int_{X^{[k]}} \mathbb{E}\Big(\bigotimes_{\eta\in V_k} f_{\eta 0} \big| \mathcal{I}^{[k]}\Big) \mathbb{E}\Big(\bigotimes_{\eta\in V_k} f_{\eta 1} \big| \mathcal{I}^{[k]}\Big) \, d\mu^{[k]} \ . \tag{3}$$

Since (X, μ, T) is ergodic, $\mathcal{I}^{[1]}$ is the trivial σ-algebra and $\mu^{[1]} = \mu \times \mu$. If (X, μ, T) is weakly mixing, then by induction $\mu^{[k]}$ is the 2^k Cartesian power $\mu^{\otimes 2^k}$ of μ for $k \geq 1$.

We now give an equivalent formulation of the definition of these measures.

Notation. For an integer $k \geq 1$, let

$$\mu^{[k]} = \int_{\Omega_k} \mu_\omega^{[k]} \, dP_k(\omega) \tag{4}$$

denote the ergodic decomposition of $\mu^{[k]}$ under $T^{[k]}$.

Then by definition

$$\mu^{[k+1]} = \int_{\Omega_k} \mu_\omega^{[k]} \times \mu_\omega^{[k]} \, dP_k(\omega) \ . \tag{5}$$

We generalize this formula. For $k, \ell \geq 1$, the concatenation of an element α of V_k with an element β of V_ℓ is the element $\alpha\beta$ of $V_{k+\ell}$. This defines a bijection of $V_k \times V_\ell$ onto $V_{k+\ell}$ and gives the identification

$$\big(X^{[k]}\big)^{[\ell]} = X^{[k+\ell]} \ .$$

Lemma 3.1. *Let $k, \ell \geq 1$ be integers and for $\omega \in \Omega_k$, let $(\mu_\omega^{[k]})^{[\ell]}$ be the measure built from the ergodic system $(X^{[k]}, \mu_\omega^{[k]}, T^{[k]})$ in the same way that $\mu_\omega^{[k]}$ was built from (X, μ, T). Then*

$$\mu^{[k+\ell]} = \int_{\Omega_k} (\mu_\omega^{[k]})^{[\ell]} \, dP_k(\omega) \ .$$

Proof. By definition, $\mu_\omega^{[k]}$ is a measure on $X^{[k]}$ and so $(\mu_\omega^{[k]})^{[\ell]}$ is a measure on $(X^{[k]})^{[\ell]}$, which we identify with $X^{[k+\ell]}$. For $\ell = 1$ the formula is Equation (5). By induction assume that it holds for some $\ell \geq 1$. Let $\mathcal{J}_\omega$ denote the invariant σ-algebra of the system $\big((X^{[k]})^{[\ell]}, (\mu_\omega^{[k]})^{[\ell]}, (T^{[k]})^{[\ell]}\big) = (X^{[k+\ell]}, (\mu_\omega^{[k]})^{[\ell]}, T^{[k+\ell]})$.

Let f and g be two bounded functions on $X^{[k+\ell]}$. By the Pointwise Ergodic Theorem, applied for both the system $(X^{[k+\ell]}, \mu^{[k+\ell]}, T^{[k+\ell]})$ and $(X^{[k+\ell]}, (\mu_\omega^{[k]})^{[\ell]}, T^{[k+\ell]})$, for almost every ω the conditional expectation of f on $\mathcal{I}^{[k+\ell]}$ (for $\mu^{[k+\ell]}$) is equal $(\mu_\omega^{[k]})^{[\ell]}$-almost everywhere to the conditional expectation of f on $\mathcal{J}_\omega$ (for $(\mu_\omega^{[k]})^{[\ell]}$).

As the same holds for g, we have

$$\begin{aligned}\int f\otimes g\,d\mu^{[k+\ell+1]}&\\ &=\int_{X^{[k+\ell]}}\mathbb{E}(f\mid\mathcal{I}^{[k+\ell]})\cdot\mathbb{E}(g\mid\mathcal{I}^{[k+\ell]})\,d\mu^{[k+\ell]}\\ &=\int_{\Omega_k}\Big(\int_{X^{[k+\ell]}}\mathbb{E}(f\mid\mathcal{I}^{[k+\ell]})\cdot\mathbb{E}(g\mid\mathcal{I}^{[k+\ell]})\,d(\mu_\omega^{[k]})^{[\ell]}\Big)\,dP_k(\omega)\\ &=\int_{\Omega_k}\Big(\int_{X^{[k+\ell]}}\mathbb{E}(f\mid\mathcal{J}_\omega)\cdot\mathbb{E}(g\mid\mathcal{J}_\omega)\,d(\mu_\omega^{[k]})^{[\ell]}\Big)\,dP_k(\omega)\\ &=\int_{\Omega_k}\Big(\int_{X^{[k+\ell+1]}}f\otimes g\,d(\mu_\omega^{[k]})^{[\ell+1]}\Big)\,dP_k(\omega)\ ,\end{aligned}$$

where the last identity uses the definition of $(\mu_\omega^{[k]})^{[\ell+1]}$. This means that $\mu^{[k+\ell+1]}=\int_\Omega(\mu_\omega^{[k]})^{[\ell+1]}\,dP_k(\omega)$. □

3.2. **The case** $k=1$**.** By using the well known ergodic decomposition of $\mu^{[1]}=\mu\times\mu$, these formulas can be written more explicitly for $k=1$. The Kronecker factor of the ergodic system (X,μ,T) is an ergodic rotation and we denote it by $(Z_1(X),t_1)$, or more simply (Z_1,t_1). Let μ_1 denote the Haar measure of Z_1, and $\pi_{X,1}$ or π_1, denote the factor map $X\to Z_1$. For $s\in Z_1$, let $\mu_{1,s}$ denote the image of the measure μ_1 under the map $z\mapsto(z,sz)$ from Z_1 to Z_1^2. This measure is invariant under $T^{[1]}=T\times T$ and is a self-joining of the rotation (Z_1,t_1). Let μ_s denote the relatively independent joining of μ over $\mu_{1,s}$. This means that for bounded functions f and g on X,

$$\int_{Z\times Z}f(x_0)g(x_1)\,d\mu_s(x_0,x_1)=\int_Z\mathbb{E}(f\mid\mathcal{Z}_1)(z)\,\mathbb{E}(g\mid\mathcal{Z}_1)(sz)\,d\mu_1(z)\tag{6}$$

where we view the conditional expectations relative to $\mathcal{Z}_1$ as functions defined on Z_1.

It is a classical result that the invariant σ-algebra $\mathcal{I}^{[1]}$ of $(X\times X,\mu\times\mu,T\times T)$ consists in the sets of the form

$$\big\{(x,y)\in X\times X:\pi_1(x)-\pi_1(y)\in A\big\}\ ,$$

where $A\subset Z_1$. From this, it is not difficult to deduce that the ergodic decomposition of $\mu\times\mu$ under $T\times T$ can be written as

$$\mu\times\mu=\int_{Z_1}\mu_s\,d\mu_1(s)\ .\tag{7}$$

In particular, for μ_1-almost every s, the measure μ_s is ergodic for $T\times T$. By Lemma 3.1, for an integer $\ell>0$ we have

$$\mu^{[\ell+1]}=\int_{Z_1}(\mu_s)^{[\ell]}\,d\mu_1(s)\ .\tag{8}$$

Formula (5) becomes

$$\mu^{[2]}=\int_{Z_1}\mu_s\times\mu_s\,d\mu_1(s)\ .$$

When f_ϵ, $\epsilon \in V_2$, are 4 bounded functions on X, writing $\tilde{f}_\epsilon = \mathbb{E}(f_\epsilon \mid \mathcal{Z}_1)$ and viewing these functions as defined on Z_1, by Equation (6) we have

$$\begin{aligned}(9)\quad & \int_{X^4} \bigotimes_{\epsilon \in V_2} f_\epsilon \, d\mu^{[2]} \\ &= \iiint_{Z_1^3} \tilde{f}_{00}(z)\tilde{f}_{10}(z+s_1)\tilde{f}_{01}(z+s_2)\tilde{f}_{11}(z+s_1+s_2)\, d\mu_1(z)\, d\mu_1(s_1)\, d\mu_1(s_2)\ .\end{aligned}$$

The projection under $\pi_1^{[2]}$ of $\mu^{[2]}$ on $Z_1^{[2]}$ is the Haar measure $\mu_1^{[2]}$ of the closed subgroup

$$\{(z, z+s_1, z+s_2, z+s_1+s_2) : z, s_1, s_2 \in Z_1\}$$

of $Z_1^{[2]} = Z_1^4$. We can reinterpret Formula (9): the system $(X^{[2]}, \mu^{[2]}, T^{[2]})$ is a joining of 4 copies of (X, μ, T), which is relatively independent with respect to the corresponding 4-joining $\mu_1^{[2]}$ of Z_1.

3.3. The side transformations.

Definition 3.2. If α is a face of V_k with $k \geq 1$, let $T_\alpha^{[k]}$ denote the transformation of $X^{[k]}$ given by

$$(T_\alpha^{[k]}\mathbf{x})_\epsilon = \begin{cases} T(x_\epsilon) & \text{for } \epsilon \in \alpha \\ x_\epsilon & \text{otherwise} \end{cases}$$

and we call this transformation a *face transformation*. When α is a side of V_k, we call $T_\alpha^{[k]}$ a *side transformation*.

The sides are faces of dimension $k-1$ and we denote the group spanned by the side transformations by $\mathcal{T}_{k-1}^{[k]}$. The subgroup spanned by those $T_\alpha^{[k]}$ where α is a side not containing $\mathbf{0}$ is denoted by $\mathcal{T}_*^{[k]}$.

We note that $\mathcal{T}_{k-1}^{[k]}$ contains $T^{[k]}$ and is spanned by $T^{[k]}$ and $\mathcal{T}_*^{[k]}$.

Lemma 3.3. *For an integer $k \geq 1$, the measure $\mu^{[k]}$ is invariant under the group $\mathcal{T}_{k-1}^{[k]}$ of side transformations.*

Proof. We proceed by induction. For $k = 1$ there are only two transformations, $\mathrm{Id} \times T$ and $T \times \mathrm{Id}$, and $\mu^{[1]} = \mu \times \mu$ is invariant under both.

Assume that the result holds for some $k \geq 1$. We consider first the side $\alpha = \{\epsilon \in V_{k+1} : \epsilon_{k+1} = 0\}$. Identifying $X^{[k+1]}$ with the Cartesian square of $X^{[k]}$, we have $T_\alpha^{[k+1]} = T^{[k]} \times \mathrm{Id}^{[k]}$. Since $T^{[k]}$ leaves each set in $\mathcal{I}^{[k]}$ invariant, by the definition (3) of $\mu^{[k+1]}$, this measure is invariant under $T_\alpha^{[k]}$. The same method gives the invariance under $T_{\alpha'}^{[k]}$, where α' is the side opposite from α.

Any other side β of V_{k+1} can be written as $\gamma \times \{0,1\}$ for some side γ of V_k. Under the identification of $X^{[k+1]}$ with $X^{[k]} \times X^{[k]}$, we have $T_\beta^{[k+1]} = T_\gamma^{[k]} \times T_\gamma^{[k]}$. By the inductive hypothesis, the transformation $T_\gamma^{[k]}$ leaves the measure $\mu^{[k]}$ invariant. Furthermore, it commutes with $T^{[k]}$ and so commutes with the conditional expectation on $\mathcal{I}^{[k]}$. By the definition (3) of $\mu^{[k+1]}$, this measure is invariant under $T_\beta^{[k+1]}$. □

Notation. Let $\mathcal{J}^{[k]}(X) = \mathcal{J}^{[k]}$ denote the σ-algebra of sets on $X^{[k]}$ that are invariant under the group $\mathcal{T}_*^{[k]}$.

Proposition 3.4. *On* $(X^{[k]}, \mu^{[k]})$*, the* σ*-algebra* $\mathcal{J}^{[k]}$ *coincides with the* σ*-algebra of sets depending only on the coordinate* $\mathbf{0}$*.*

Proof. If α is a side not containing $\mathbf{0}$, then $(T^{[k]}_\alpha \mathbf{x})_{\mathbf{0}} = x_{\mathbf{0}}$ for every $\mathbf{x} \in X^{[k]}$. Thus a subset of $X^{[k]}$ depending only on the coordinate $\mathbf{0}$ is obviously invariant under the group $\mathcal{T}^{[k]}_*$ and so belongs to $\mathcal{J}^{[k]}$.

We prove the converse inclusion by induction. For $k = 1$, $X^{[1]} = X^2$, the group $\mathcal{T}^{[k]}_*$ contains $\mathrm{Id} \times T$ and the result is obvious.

Assume the result holds for some $k \geq 1$. Let F be a bounded function on $X^{[k+1]}$ that is measurable with respect to the σ-algebra $\mathcal{J}^{[k+1]}$. Write $\mathbf{x} = (\mathbf{x}', \mathbf{x}'')$ for a point of $X^{[k+1]}$, where $\mathbf{x}', \mathbf{x}'' \in X^{[k]}$. Since $(X^{[k+1]}, \mu^{[k+1]}, T^{[k+1]})$ is a self joining of $(X^{[k]}, \mu^{[k]}, T^{[k]})$, the function $F(\mathbf{x}) = F(\mathbf{x}', \mathbf{x}'')$ on $X^{[k+1]}$ can be approximated in $L^2(\mu^{[k+1]})$ by finite sums of the form

$$\sum_i F_i(\mathbf{x}') G_i(\mathbf{x}'') \ ,$$

where F_i and G_i are bounded functions on $X^{[k]}$. Since $T^{[k+1]}_{k+1} = \mathrm{Id}^{[k]} \times T^{[k]}$ is one of the side transformations of $X^{[k+1]}$, it leaves F invariant and by passing to ergodic averages, we can assume that each of the functions G_i is invariant under $T^{[k]}$. Thus, by the construction of $\mu^{[k+1]}$, for all i, $G_i(\mathbf{x}') = G_i(\mathbf{x}'')$ for $\mu^{[k+1]}$-almost every $(\mathbf{x}', \mathbf{x}'')$. Therefore the above sum is equal $\mu^{[k+1]}$-almost everywhere to a function depending only on $\mathbf{x}'$. Passing to the limit, there exists a bounded function H on $X^{[k]}$ such that $F(\mathbf{x}) = H(\mathbf{x}')$ $\mu^{[k+1]}$-almost everywhere.

Under the natural embedding of V_k in V_{k+1} given by the first side, each side of V_k is the intersection of a side of V_{k+1} with V_k. Since F is invariant under $\mathcal{T}^{[k+1]}_*$, H is also invariant under $\mathcal{T}^{[k]}_*$ and thus is measurable with respect to $\mathcal{J}^{[k]}$. By the induction hypothesis, H depends only on the $\mathbf{0}$ coordinate. □

Corollary 3.5. $(X^{[k]}, \mu^{[k]})$ *is ergodic for the group of side transformations* $\mathcal{T}^{[k]}_{k-1}$*.*

Proof. A subset A of $X^{[k]}$ invariant under the group $\mathcal{T}^{[k]}_{k-1}$ is also invariant under the group $\mathcal{T}^{[k]}_*$. Thus its characteristic function is equal almost everywhere to a function depending only on the $\mathbf{0}$ coordinate. Since A is invariant under $T^{[k]}$, this last function is invariant under T and so is constant. □

Since the side transformations commute with $T^{[k]}$, they induce measure preserving transformations on the probability space (Ω_k, P_k) introduced in (4), which we denote by the same symbols. From the last Corollary, this immediately gives:

Corollary 3.6. (Ω_k, P_k) *is ergodic under the action of the group* $\mathcal{T}^{[k]}_*$*.*

3.4. **Symmetries.**

Proposition 3.7. *The measure* $\mu^{[k]}$ *is invariant under the transformation* σ_* *for every* $\sigma \in \mathcal{S}_k$*.*

We note that σ_* commutes with $T^{[k]}$ for every $\sigma \in \mathcal{S}_k$.

Proof. First we show by induction that $\mu^{[k]}$ is invariant under reflections.

For $k = 1$ the map $(x_0, x_1) \mapsto (x_1, x_0)$ is the unique reflection and it leaves the measure $\mu^{[1]} = \mu \times \mu$ invariant.

Assume that for some integer $k \geq 1$, the measure $\mu^{[k]}$ is invariant under all reflections. For $1 \leq j \leq k+1$, let R_j be the reflection of $X^{[k+1]}$ corresponding to the digit j. If $j < k+1$, R_j can be written $S_j \times S_j$, where S_j is the reflection of $X^{[k]}$ for the digit j. Since $\mu^{[k]}$ is invariant under S_j, by construction $\mu^{[k+1]}$ is invariant under R_j. The reflection R_{k+1} simply exchanges the two sides of $X^{[k+1]}$ and by construction of the measures, it leaves the measure $\mu^{[k+1]}$ invariant.

Next we show that $\mu^{[k]}$ is invariant under digit permutations. For $k = 1$ there is no nontrivial digit permutation and so nothing to prove. For $k = 2$, there is one nontrivial digit permutation, the map $(x_{00}, x_{01}, x_{10}, x_{11}) \mapsto (x_{00}, x_{10}, x_{01}, x_{11})$. By Formula (9), $\mu^{[2]}$ is invariant under this map.

Assume that for some integer $k \geq 2$, the measure $\mu^{[k]}$ is invariant under all digit permutations. The group of permutations of $\{1, \ldots, k, k+1\}$ is spanned by the permutations leaving $k+1$ fixed and the transposition $(k, k+1)$ exchanging k and $k+1$.

Consider first the case of a permutation of $\{1, \ldots, k, k+1\}$ leaving $k+1$ fixed. The corresponding transformation R of $X^{[k+1]} = X^{[k]} \times X^{[k]}$ can be written as $S \times S$, where S is a digit permutation of $X^{[k]}$ and so leaves $\mu^{[k]}$ invariant. By construction, $\mu^{[k+1]}$ is invariant under R.

Next consider the case of the transformation R of $X^{[k+1]}$ associated to the permutation $(k, k+1)$. Using the ergodic decomposition of Formula (4) of $\mu^{[k-1]}$ and Equation (5) for $k-1$ the measure $(\mu_\omega^{[k-1]})^{[2]}$ (as a measure on $(X^{[k-1]})^{[2]}$) is invariant by the transposition of the two digits. Thus, when we consider the same measure as a measure on $X^{[k+1]}$, it is invariant under R. Taking the integral, $\mu^{[k+1]}$ is invariant under R. Therefore $\mu^{[k+1]}$ is invariant under all digit permutations. □

Corollary 3.8. *The image of $\mu^{[k]}$ under any side projection $X^{[k]} \to X^{[k-1]}$ is $\mu^{[k-1]}$.*

Proof. By construction of $\mu^{[k]}$, the result holds for the side projection associated to the side $\{\epsilon \in V_k : \epsilon_k = 0\}$ of V_k. The result for the other side projections follows immediately from Proposition 3.7. □

3.5. **Some seminorms.** We define and study some seminorms on $L^\infty(\mu)$. When X is $\mathbb{Z}/N\mathbb{Z}$ for some integer $N > 0$ and is endowed with the transformation $n \mapsto n+1$ mod N, these seminorms are the same as those used by Gowers in [G01], although the contexts are very different.

For simplicity, we mostly consider real valued functions.

Fix $k \geq 1$. For a bounded function f on X, by the definition (3) of $\mu^{[k]}$:

$$\int_{X^{[k]}} \prod_{\epsilon \in V_k} f(x_\epsilon)\, d\mu^{[k]}(\mathbf{x}) = \int_{X^{[k-1]}} \Big(\mathbb{E}\Big(\prod_{\eta \in V_{k-1}} f(x_\eta) \mid \mathcal{I}^{[k-1]}\Big)\Big)^2 d\mu^{[k-1]} \geq 0$$

and so we can define

(10)
$$|||f|||_k = \Big(\int \bigotimes_{\epsilon \in V_k} f\, d\mu^{[k]}\Big)^{1/2^k} = \Big(\int_{X^{[k-1]}} \Big(\mathbb{E}\Big(\prod_{\eta \in V_{k-1}} f(x_\eta) \mid \mathcal{I}^{[k-1]}\Big)\Big)^2 d\mu^{[k-1]}\Big)^{1/2^k}.$$

Lemma 3.9. *Let $k \geq 1$ be an integer.*

(1) When f_ϵ, $\epsilon \in V_k$, are bounded functions on X,

$$\Bigl|\int \bigotimes_{\epsilon\in V_k} f_\epsilon\, d\mu^{[k]}\Bigr| \leq \prod_{\epsilon\in V_k} \|\!|f_\epsilon|\!\|_k \ .$$

(2) $\|\!|\cdot|\!\|_k$ is a seminorm on $L^\infty(\mu)$.
(3) For a bounded function f, $\|\!|f|\!\|_k \leq \|\!|f|\!\|_{k+1}$.

Proof. (1) Using the definition of $\mu^{[k]}$, the Cauchy-Schwarz inequality and again using definition of $\mu^{[k]}$,

$$\begin{aligned}
&\Bigl(\int \bigotimes_{\epsilon\in V_k} f_\epsilon\, d\mu^{[k]}\Bigr)^2 \\
&\qquad\leq \Bigl\|\mathbb{E}\bigl(\bigotimes_{\eta\in V_{k-1}} f_{\eta 0}|\mathcal{I}^{[k-1]}\bigr)\Bigr\|^2_{L^2(\mu^{[k-1]})}\cdot \Bigl\|\mathbb{E}\bigl(\bigotimes_{\eta\in V_{k-1}} f_{\eta 1}|\mathcal{I}^{[k-1]}\bigr)\Bigr\|^2_{L^2(\mu^{[k-1]})} \\
&\qquad=\Bigl(\int \bigotimes_{\epsilon\in V_k} g_\epsilon\, d\mu^{[k]}\Bigr)\cdot\Bigl(\int \bigotimes_{\epsilon\in V_k} h_\epsilon\, d\mu^{[k]}\Bigr)
\end{aligned}$$

where the functions g_ϵ and h_ϵ are defined for $\eta \in V_{k-1}$ by $g_{\eta 0} = g_{\eta 1} = f_{\eta 0}$ and $h_{\eta 0} = h_{\eta 1} = f_{\eta 1}$. For each of these two integrals, we permute the digits $k-1$ and k and then use the same method. Thus $\bigl(\int \bigotimes_{\epsilon\in V_k} f_\epsilon\, d\mu^{[k]}\bigr)^4$ is bounded by the product of 4 integrals. Iterating this procedure k times, we have the statement.
(2) The only nontrivial property is the subadditivity of $\|\!|\cdot|\!\|_k$. Let f and g be bounded functions on X. Expanding $\|\!|f+g|\!\|^{2^k}$, we get the sum of 2^k integrals. Using part (1) to bound each of them, we have the subadditivity.
(3) For a bounded function f on X,

$$\|\!|f|\!\|_{k+1}^{2^{k+1}} = \Bigl\|\mathbb{E}\bigl(\bigotimes_{\eta\in V_k} f \mid \mathcal{I}^{[k]}\bigr)\Bigr\|^2_{L^2(\mu^{[k]})} \geq \Bigl(\int \bigotimes_{\eta\in V_k} f\, d\mu^{[k]}\Bigr)^2 = \|\!|f|\!\|_k^{2^{k+1}} \ .$$

□

From part (1) of this Lemma, and the definition (3) of $\mu^{[k+1]}$, we have:

Corollary 3.10. *Let $k \geq 1$ be an integer and let f_ϵ, $\epsilon \in V_k$, be bounded functions on X. Then*

$$\Bigl\|\mathbb{E}\bigl(\bigotimes_{\epsilon\in V_k} f_\epsilon \mid \mathcal{I}^{[k]}\bigr)\Bigr\|_{L^2(\mu^{[k]})} \leq \prod_{\epsilon\in V_k} \|\!|f_\epsilon|\!\|_{k+1} \ .$$

In a few cases we also need the seminorm for complex valued function and so introduce notation for its definition. Write $C\colon \mathbb{C} \to \mathbb{C}$ for the conjugacy map $z \mapsto \bar{z}$. Thus $C^m z = z$ for m even and is $\bar{z}$ for m odd. The definition of the seminorm becomes

$$\|\!|f|\!\|_k = \Bigl(\int \bigotimes_{\epsilon\in V_k} C^{|\epsilon|} f\, d\mu^{[k]}\Bigr)^{1/2^k} \ . \tag{11}$$

Similar properties, with obvious modifications, hold for this seminorm.

4. Construction of factors

4.1. **The marginal** $(X^{[k]^*}, \mu^{[k]^*})$**.** We continue to assume that (X, μ, T) is an ergodic system, and let $k \geq 1$ be an integer.

We consider the $2^k - 1$-dimensional marginals of $\mu^{[k]}$. For simplicity, we consider first the marginal obtained by 'omitting' the coordinate $\mathbf{0}$. The other cases are similar.

Recall that $V_k^* = V_k \setminus \{\mathbf{0}\}$. Consider a point $\mathbf{x} \in X^{[k]}$ as a pair $(x_{\mathbf{0}}, \tilde{x})$, with $x_{\mathbf{0}} \in X$ and $\tilde{x} = (x_\epsilon\ ; \epsilon \in V_k^*) \in X^{[k]^*}$. Let $\mu^{[k]^*}$ denote the measure on $X^{[k]^*}$, which is the image of $\mu^{[k]}$ under the natural projection $\mathbf{x} \mapsto \tilde{x}$ from $X^{[k]}$ onto $X^{[k]^*}$.

We recall that $(X^{[k]}, \mu^{[k]})$ is endowed with the measure preserving action of the groups $\mathcal{T}_*^{[k]}$ and $\mathcal{T}_{k-1}^{[k]}$. The first action is spanned by the transformations $T_\alpha^{[k]}$ for α a side not containing $\mathbf{0}$ and the second action is spanned by $T^{[k]}$ and $\mathcal{T}_*^{[k]}$. By Corollary 3.5, $\mu^{[k]}$ is ergodic for the action of $\mathcal{T}_{k-1}^{[k]}$.

All the transformations belonging to $\mathcal{T}_{k-1}^{[k]}$ factor through the projection $X^{[k]} \to X^{[k]^*}$ and induce transformations of $X^{[k]^*}$ preserving $\mu^{[k]^*}$. This defines a measure preserving action of the group $\mathcal{T}_{k-1}^{[k]}$ and of its subgroup $\mathcal{T}_*^{[k]}$ on $X^{[k]^*}$. The measure $\mu^{[k]^*}$ is ergodic for the action of $\mathcal{T}_{k-1}^{[k]^*}$.

On the other hand, all the transformations belonging to $\mathcal{T}_{k-1}^{[k]}$ factor through the projection $\mathbf{x} \mapsto x_{\mathbf{0}}$ from $X^{[k]}$ to X, and induce measure preserving transformations of X. The transformation $T^{[k]}$ induces the transformation T on X, and each transformation belonging to $\mathcal{T}_*^{[k]}$ induces the trivial transformation on X. This defines a measure preserving ergodic action of the group $\mathcal{T}_{k-1}^{[k]}$ on X, with a trivial restriction to the subgroup $\mathcal{T}_*^{[k]}$.

Thus we can consider (in a second way) $\mu^{[k]}$ as a joining between two systems. The first system is $(X^{[k]^*}, \mu^{[k]^*})$, and the second (X, μ), both endowed with the action of the group $\mathcal{T}_{k-1}^{[k]}$.

Let $\mathcal{I}^{[k]^*}$ denote the σ-algebra of $T^{[k]}$-invariant sets of $(X^{[k]^*}, \mu^{[k]})$ and $\mathcal{J}^{[k]^*}$ denote the σ-algebra of subsets of $X^{[k]^*}$ which are invariant under the action of $\mathcal{T}_*^{[k]}$.

4.2. **The definition of the factors** Z_k**.** Let $A \subset X^{[k]^*}$ belong to the σ-algebra $\mathcal{J}^{[k]^*}$. A is invariant under the action of the group $\mathcal{T}_*^{[k]}$ and thus the subset $X \times A$ of $X^{[k]}$ is invariant under $\mathcal{T}_*^{[k]}$. By Proposition 3.4, this set depends only on the first coordinate. This means that there exists a subset B of X with $X \times A = B \times X^{[k]^*}$, up to a subset of $X^{[k]}$ of $\mu^{[k]}$-measure zero. That is,

$$\mathbf{1}_A(\tilde{x}) = \mathbf{1}_B(x_{\mathbf{0}}) \text{ for } \mu^{[k]}\text{-almost every } \mathbf{x} = (x_{\mathbf{0}}, \tilde{x}) \in X^{[k]}\ . \tag{12}$$

It is immediate that if a subset A of $X^{[k]^*}$ satisfies Equation (12) for some $B \subset X$, then it is invariant under $\mathcal{T}_*^{[k]}$ and thus measurable with respect to $\mathcal{J}^{[k]^*}$. Moreover, the subsets B of X corresponding to a subset $A \in \mathcal{J}^{[k]^*}$ in this way form a sub-σ-algebra of $\mathcal{X}$. We define:

Definition 4.1. For an integer $k \geq 1$, $\mathcal{Z}_{k-1}(X)$ is the σ-algebra of subsets B of X for which there exists a subset A of $X^{[k]^*}$ so that Equation (12) is satisfied.

In the sequel, we often identify the σ-algebras $\mathcal{Z}_{k-1}(X)$ and $\mathcal{J}^{[k]^*}(X)$, by identifying a subset B of X belonging to $\mathcal{Z}_{k-1}(X)$ with the corresponding set $A \in \mathcal{J}^{[k]^*}$.

The σ-algebra $\mathcal{Z}_{k-1}$ is invariant under T and so defines a factor of (X,μ,T) that we write $(Z_{k-1}(X),\mu_k,T)$, or simply (Z_{k-1},μ_k,T) or even Z_{k-1}. The factor map $X\mapsto Z_{k-1}(X)$ is written $\pi_{X,k-1}$ or π_{k-1}.

As $X^{[1]^*}=X$, the σ-algebra $\mathcal{J}^{[1]}$ is trivial and $Z_0(X)$ is the trivial factor.

We have already used the notation $Z_1(X)$ for the Kronecker factor and we check now that the two definitions of $Z_1(X)$ coincide. For the moment, let Z denote the Kronecker factor of X and let $\pi\colon X\to Z$ be the natural projection. By Formula (9), we have $\mu^{[2]^*}=\mu\times\mu\times\mu$ and $\mathcal{J}^{[2]^*}$ is the algebra of sets which are invariant under $T\times\mathrm{Id}\times T$ and $\mathrm{Id}\times T\times T$. By classical arguments, $\mathcal{J}^{[2]^*}$ is measurable with respect to $\mathcal{Z}\times\mathcal{Z}\times\mathcal{Z}$, and more precisely $\mathcal{J}^{[2]^*}=\Phi^{-1}(\mathcal{Z})$, where the map $\Phi\colon X^{[2]^*}\to Z$ is given by $\Phi(x_{01},x_{10},x_{11})=\pi(x_{01})-\pi(x_{10})+\pi(x_{11})$. But $\mu^{[2]}$ is concentrated on the set $\{\mathbf{x}:x_{00}=\Phi(\tilde{x})\}$. This is exactly the situation described above, with $\mathcal{Z}_1=\mathcal{Z}$.

Lemma 4.2. *For an integer $k\geq 1$, $(X^{[k]},\mu^{[k]})$ is the relatively independent joining of (X,μ) and $(X^{[k]^*},\mu^{[k]^*})$ over $\mathcal{Z}_{k-1}$ when identified with $\mathcal{J}^{[k]^*}$.*

Proof. Let f be a bounded function on X and g be a bounded function on $X^{[k]^*}$. Since $\mu^{[k]}$ is invariant under the group $\mathcal{T}^{[k]}_{k-1}$, for integers $n_1,n_2,\dots,n_k$ we have

$$\int f(x_{\mathbf{0}})g(\tilde{x})\,d\mu^{[k]}(\mathbf{x})=\int f(x_{\mathbf{0}})g\big((T_1^{[k]})^{n_1}(T_2^{[k]})^{n_2}\dots(T_k^{[k]})^{n_k}\tilde{x}\big)\,d\mu^{[k]}(\mathbf{x})\ ,$$

where $T_1^{[k]},T_2^{[k]},\dots,T_k^{[k]}$ denote the k generators of $\mathcal{T}^{[k]}_*$. Thus, by averaging and taking the limit

$$\begin{aligned}\int f(x_{\mathbf{0}})g(\tilde{x})\,d\mu^{[k]}(\mathbf{x})&=\int f(x_{\mathbf{0}})\mathbb{E}(g\mid\mathcal{J}^{[k]^*})(\tilde{x})\,d\mu^{[k]}(\mathbf{x})\\&=\int\mathbb{E}(f\mid\mathcal{Z}_{k-1})(x_{\mathbf{0}})\mathbb{E}(g\mid\mathcal{J}^{[k]^*})(\tilde{x})\,d\mu^{[k]}(\mathbf{x})\ .\end{aligned}\tag{13}$$

□

Lemma 4.3. *Let f be a bounded function on X. Then*

$$\mathbb{E}(f\mid\mathcal{Z}_{k-1})=0\iff|\!|\!|f|\!|\!|_k=0\ .$$

Proof. Assume that $\mathbb{E}(f\mid\mathcal{Z}_{k-1})=0$. By Equation (13) applied with $g(\tilde{x})=\prod_{\epsilon\in V_k^*}f(x_\epsilon)$, we have $|\!|\!|f|\!|\!|_k=0$ by definition (10) of the seminorm.

Conversely, assume that $|\!|\!|f|\!|\!|_k=0$. By Lemma 3.9, for every choice of f_ϵ, $\epsilon\in V_k^*$,

$$\int_{X^{[k]}}f(x_{\mathbf{0}})\prod_{\epsilon\in V_k^*}f_\epsilon(x_\epsilon)\,d\mu^{[k]}(\mathbf{x})=0\ .$$

By density, the function $\mathbf{x}\mapsto f(x_{\mathbf{0}})$ is orthogonal in $L^2(\mu^{[k]})$ to every function defined on $X^{[k]^*}$, and in particular to every function measurable with respect to $\mathcal{J}^{[k]^*}$. But this means that f is orthogonal in $L^2(\mu)$ to every $\mathcal{Z}_{k-1}$-measurable function and so $\mathbb{E}(f\mid\mathcal{Z}_{k-1})=0$. □

Corollary 4.4. *The factors $Z_k(X)$, $k\geq 1$, form an increasing sequence of factors of X.*

4.3. **Taking factors.** Let $p\colon (X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,T)$ be a factor map. We can associate to Y the space $Y^{[k]}$ and the measure $\nu^{[k]}$ in the same way that $X^{[k]}$ and $\mu^{[k]}$ are associated to X in Section 3. This induces a natural map $p^{[k]}\colon X^{[k]}\to Y^{[k]}$, commuting with the transformations $T^{[k]}$ and the group $\mathcal{T}_{k-1}^{[k]}$.

Lemma 4.5. *Let $p\colon (X,\mu,T)\to(Y,\nu,T)$ be a factor map and let $k\geq 1$ be an integer.*

1. *The map $p^{[k]}\colon (X^{[k]},\mu^{[k]},T^{[k]})\to(Y^{[k]},\nu^{[k]},T^{[k]})$ is a factor map.*
2. *For a bounded function f on Y, $\|\!|f|\!\|_k=\|\!|f\circ p|\!\|_k$, where the first seminorm is associated to Y and the second one to X.*

Proof. (1) Clearly $p^{[k]}$ commutes with the transformation $T^{[k]}$ and so it suffices to show that the image of $\mu^{[k]}$ under $p^{[k]}$ is $\nu^{[k]}$. We prove this statement by induction. The result is obvious for $k=0$ and so assume it holds for some $k\geq 0$. Let f_ϵ, $\epsilon\in V_k$, be bounded functions on Y. Since $p^{[k]}$ is a factor map, it commutes with the operators of conditional expectation on the invariant σ-algebras and we have

$$\mathbb{E}\Big(\big(\bigotimes_{\epsilon\in V_k} f_\epsilon\big)\circ p^{[k]}\big|\mathcal{I}^{[k]}(X)\Big)=\mathbb{E}\Big(\bigotimes_{\epsilon\in V_k} f_\epsilon\big|\mathcal{I}^{[k]}(X)\Big)\circ p^{[k]}\ .$$

The statement for $k+1$ follows from the definitions of the measures $\mu^{[k+1]}$ and $\nu^{[k+1]}$.

(2) This follows immediately from the first part and the definitions of the seminorms. $\square$

Proposition 4.6. *Let $p\colon (X,\mu,T)\to(Y,\nu,T)$ be a factor map and let $k\geq 1$ be an integer. Then $p^{-1}(\mathcal{Z}_{k-1}(Y))=\mathcal{Z}_{k-1}(X)\cap p^{-1}(\mathcal{Y})$.*

Using the identification of the σ-algebras $\mathcal{Y}$ and $p^{-1}(\mathcal{Y})$, this formula is then written

$$\mathcal{Z}_{k-1}(Y)=\mathcal{Z}_{k-1}(X)\cap\mathcal{Y}\ .$$

Proof. For $k=1$ there is nothing to prove. Let $k\geq 2$ and let $p^{[k]^*}\colon X^{[k]^*}\to Y^{[k]^*}$ denote the natural map. By Lemma 4.5, it is a factor map. Let f be a bounded function on X that is measurable with respect to $p^{-1}(\mathcal{Z}_{k-1}(Y))$. Then $f=g\circ p$ for some function g on Y which is measurable with respect to $\mathcal{Z}_{k-1}(Y)$. There exists a function F on $Y^{[k]^*}$, measurable with respect to $\mathcal{J}^{[k]^*}$, so that $g(y_{\mathbf{0}})=F(\tilde{y})$ for $\nu^{[k]}$-almost every $\mathbf{y}=(y_{\mathbf{0}},\tilde{y})\in Y^{[k]}$. Thus $g\circ p(x_{\mathbf{0}})=F\circ p^{[k]^*}(\tilde{x})$ for $\mu^{[k]}$-almost every $\mathbf{x}=(x_{\mathbf{0}},\tilde{x})\in X^{[k]}$ and the function $f=g\circ p$ is measurable with respect to $\mathcal{Z}_{k-1}(X)$. We have $p^{-1}(\mathcal{Z}_{k-1}(Y))\subset\mathcal{Z}_{k-1}(X)\cap p^{-1}(\mathcal{Y})$.

Conversely, assume that f is a bounded function on X, measurable with respect to $\mathcal{Z}_{k-1}(X)\cap p^{-1}(\mathcal{Y})$. Then $f=g\circ p$ for some g on Y. Write $g=g'+g''$, where g' is measurable with respect to $\mathcal{Z}_{k-1}(Y)$ and $\mathbb{E}(g''\mid\mathcal{Z}_{k-1}(Y))=0$. By the first part, $g'\circ p$ is measurable with respect to $\mathcal{Z}_{k-1}(X)$. Using Lemma 4.3 and Part (2) of Lemma 4.5, $\|\!|g''|\!\|_k=0$ and so $\|\!|g''\circ p|\!\|_k=0$ and $\mathbb{E}(g''\circ p\mid\mathcal{Z}_{k-1}(X))=0$. Since $f=g'\circ p+g''\circ p$ is measurable with respect to $\mathcal{Z}_{k-1}(X)$, we have $g''\circ p=0$. Thus $g''=0$ and g is measurable with respect to $\mathcal{Z}_{k-1}(Y)$. $\square$

4.4. **The factor $Z_\ell^{[k]}$ of $X^{[k]}$.** We apply this to the factors $Z_\ell=Z_\ell(X)$ of X. For integers $k,\ell\geq 1$, $(Z_\ell^{[k]},\mu_\ell^{[k]},T^{[k]})$ is the 2^k-dimensional system associated to (Z_ℓ,μ_ℓ,T) in the same way that $(X^{[k]},\mu^{[k]},T^{[k]})$ is associated to (X,μ,T). The map $\pi_\ell^{[k]}\colon X^{[k]} \to Z_\ell^{[k]}$ is a factor map. We have $\mathcal{Z}_k(Z_\ell(X)) = \mathcal{Z}_k(X) \cap \mathcal{Z}_\ell(X)$. Since the sequence $\{\mathcal{Z}_k\}$ is increasing, we have

$$\mathcal{Z}_k(Z_\ell(X)) = \begin{cases} \mathcal{Z}_k(X) & \text{if } k \leq \ell \\ \mathcal{Z}_\ell(x) & \text{otherwise} . \end{cases} \tag{14}$$

Proposition 4.7. *Let $k \geq 1$ be an integer.*

(1) *As a joining of 2^k copies of (X,μ), $(X^{[k]},\mu^{[k]})$ is relatively independent over the joining $(Z_{k-1}^{[k]},\mu_{k-1}^{[k]})$ of 2^k copies of (Z_{k-1},μ_{k-1}).*
(2) *Z_k is the smallest factor Y of X so that the σ-algebra $\mathcal{I}^{[k]}$ is measurable with respect to $Y^{[k]}$.*

Proof. (1) The statement is equivalent to showing whenever f_ϵ, $\epsilon \in V_k$, are bounded functions on X,

$$\int_{X^{[k]}} \bigotimes_{\epsilon\in V_k} f_\epsilon \, d\mu^{[k]} = \int_{Z_{k-1}^{[k]}} \bigotimes_{\epsilon\in V_k} \mathbb{E}(f_\epsilon \mid \mathcal{Z}_{k-1}) \, d\mu_{k-1}^{[k]} . \tag{15}$$

It suffices to show that

$$\int_{X^{[k]}} \bigotimes_{\epsilon\in V_k} f_\epsilon \, d\mu^{[k]} = 0 \tag{16}$$

whenever $\mathbb{E}(f_\eta \mid \mathcal{Z}_{k-1}) = 0$ for some $\eta \in V_k$. By Lemma 4.3, if $\mathbb{E}(f_\eta \mid \mathcal{Z}_{k-1}) = 0$, we have that $|||f_\eta|||_k = 0$. Lemma 3.9 implies equality (16).

(2) Let f_ϵ, $\epsilon \in V_k$, be bounded functions on X. We claim that

$$\mathbb{E}\Big(\bigotimes_{\epsilon\in V_k} f_\epsilon \mid \mathcal{I}^{[k]}\Big) = \mathbb{E}\Big(\bigotimes_{\epsilon\in V_k} \mathbb{E}(f_\epsilon \mid \mathcal{Z}_k) \ \mid \mathcal{I}^{[k]}\Big) . \tag{17}$$

As above, it suffices to show this holds when $\mathbb{E}(f_\eta \mid \mathcal{Z}_k) = 0$ for some $\eta \in V_k$. By Lemma 4.3, this condition implies that $|||f_\eta|||_{k+1} = 0$. By Corollary 3.10, the left hand side of Equation (17) is equal to zero and the claim follows.

Every bounded function on $X^{[k]}$ which is measurable with respect to $\mathcal{I}^{[k]}$ can be approximated in $L^2(\mu^{[k]})$ by finite sums of functions of the form $\mathbb{E}(\bigotimes_{\epsilon\in V_k} f_\epsilon \mid \mathcal{I}^{[k]})$ where f_ϵ, $\epsilon \in V_k$, are bounded functions on X. By Equation (17), one can assume that these functions are measurable with respect to $\mathcal{Z}_k$. In this case, $\bigotimes_{\epsilon\in V_k} f_\epsilon$ is measurable with respect to $\mathcal{Z}_k^{[k]}$ (recall that $\pi_k^{[k]}\colon X^{[k]} \to Z_k^{[k]}$ is a factor map by Part (1) of Lemma 4.5). Since this σ-algebra is invariant under $T^{[k]}$, $\mathbb{E}(\bigotimes_{\epsilon\in V_k} f_\epsilon \mid \mathcal{I}^{[k]})$ is also measurable with respect to $\mathcal{Z}_k^{[k]}$. Therefore $\mathcal{I}^{[k]}$ is measurable with respect to $\mathcal{Z}_k^{[k]}$.

We use induction to show that $\mathcal{Z}_k$ is the smallest factor of X with this property. For $k = 0$, $\mathcal{I}^{[0]}$ and $\mathcal{Z}_0$ are both the trivial factor of X and there is nothing to prove. Let $k \geq 1$ and assume that the result holds for $k-1$.

Let Y be a factor of X so that $\mathcal{I}^{[k]}$ is measurable with respect to $\mathcal{Y}^{[k]}$. For any bounded function f on X with $\mathbb{E}(f \mid \mathcal{Y}) = 0$, we have to show that $\mathbb{E}(f \mid \mathcal{Z}_k) = 0$.

By projecting on the first 2^{k-1} coordinates, $\mathcal{I}^{[k-1]}$ is measurable with respect to $\mathcal{Y}^{[k-1]}$. By the induction hypothesis, $\mathcal{Y} \supset \mathcal{Z}_{k-1}$. Since $\mu^{[k]}$ is a relatively independent joining over $Z_{k-1}^{[k]}$, it is a relatively independent joining over $Y^{[k]}$.

This implies that when f_ϵ, $\epsilon \in V_k$, are bounded functions on X, we have

$$\mathbb{E}(\bigotimes_{\epsilon\in V_k} f_\epsilon \mid \mathcal{Y}^{[k]}) = \bigotimes_{\epsilon\in V_k} \mathbb{E}(f_\epsilon \mid \mathcal{Y}) \ .$$

We apply this with $f_\epsilon = f$ for all ϵ. The function $\mathbf{x} \mapsto \prod_{\epsilon\in V_k} f(x_\epsilon)$ has zero conditional expectation with respect to $\mathcal{Y}^{[k]}$. By hypothesis, it has zero conditional expectation with respect to $\mathcal{I}^{[k]}$. By the definition (10) of the seminorm, $|||f|||_{k+1} = 0$ and by Lemma 4.3, $\mathbb{E}(f \mid \mathcal{Z}_k) = 0$. □

4.5. **More about the marginal $\mu^{[k]^*}$.** The results of this Subsection are used only in Section 13, in the study of the second kind of averages.

Lemma 4.8. *Let $k \geq 2$ and f_ϵ, $\epsilon \in V_k$, be 2^k bounded functions on X. If there exists $\eta \in V_k$ so that f_η is measurable with respect to $\mathcal{Z}_{k-2}$ and if there exists $\zeta \in V_k$ so that $\mathbb{E}(f_\zeta \mid \mathcal{Z}_{k-2}) = 0$, then $\int \bigotimes_{\epsilon\in V_k} f_\epsilon \, d\mu^{[k]} = 0$.*

Proof. If $\eta = \zeta$, then $f_\eta = f_\zeta = 0$ and the result is obvious.

Consider first the case that (η, ζ) is an edge of V_k. Without loss of generality, we can assume that for some j, $\eta_j = 0$ and $\zeta_j = 1$ and that $\eta_i = \zeta_i$ for $i \neq j$. We proceed as in the proof of Lemma 3.9, but stop the iteration of the Cauchy-Schwarz inequality one step earlier. This gives a bound of $(\int \bigotimes_{\epsilon\in V_k} f_\epsilon \, d\mu^{[k]})^{2^{k-1}}$ by a product of 2^{k-1} integrals, with one of them being

$$\int \prod_{\substack{\epsilon\in V_k\\ \epsilon_j=0}} f_\eta(x_\epsilon) \cdot \prod_{\substack{\epsilon\in V_k\\ \epsilon_j=1}} f_\zeta(x_\epsilon) \, d\mu^{[k]}(\mathbf{x})$$

$$= \int \mathbb{E}(\bigotimes_{\epsilon\in V_{k-1}} f_\eta \mid \mathcal{I}^{[k-1]}) \cdot \mathbb{E}(\bigotimes_{\epsilon\in V_{k-1}} f_\zeta \mid \mathcal{I}^{[k-1]}) \, d\mu^{[k-1]} \ .$$

The conditional expectation with respect to $\mathcal{I}^{[k-1]}$ commutes with the conditional expectation with respect to $\mathcal{Z}_{k-2}^{[k-1]}$. The function $\bigotimes_{\epsilon\in V_{k-1}} f_\eta$ is measurable with respect to $\mathcal{Z}_{k-2}^{[k-1]}$ and thus the first conditional expectation in the above integral is measurable with respect to this factor. Since $\mu^{[k-1]}$ is relatively independent over $\mathcal{Z}_{k-2}^{[k-1]}$, we have $\mathbb{E}(\bigotimes_{\epsilon\in V_{k-1}} f_\zeta \mid \mathcal{Z}_{k-2}^{[k-1]}) = 0$ and the conditional expectation with respect to $\mathcal{Z}_{k-2}^{[k-1]}$ of the second term in the integral is 0. Therefore the integral is zero.

Now consider the general case. Choose a sequence $\eta = \eta_1, \eta_2, \dots, \eta_m = \zeta$ in V_k so that $(\eta_\ell, \eta_{\ell+1})$ is an edge for each ℓ. Make a series of changes in the integral $\int \bigotimes_{\epsilon\in V_k} f_\epsilon \, d\mu^{[k]}$, substituting successively $\mathbb{E}(f_{\eta_2} \mid \mathcal{Z}_{k-2})$ for f_{η_2}, $\mathbb{E}(f_{\eta_3} \mid \mathcal{Z}_{k-2})$ for f_{η_3}, ..., and $\mathbb{E}(f_{\eta_m} \mid \mathcal{Z}_{k-2})$ for $f_{\eta_m} = f_\zeta$. By the previous case, each of these substitutions leaves the value of the integral unchanged. After the last substitution, the integral is obviously 0. □

Proposition 4.9.
(1) *For every integer $k \geq 2$, the measure $\mu^{[k]^*}$ is the relatively independent joining of $2^k - 1$ copies of μ over $\mathcal{Z}_{k-2}^{[k]^*}$*
(2) *For every integer $k \geq 1$, the σ-algebra $\mathcal{I}^{[k]^*}$ is measurable with respect to $\mathcal{Z}_{k-1}^{[k]^*}$*
(3) *For every integer $k \geq 1$, the σ-algebra $\mathcal{J}^{[k]^*}$ is measurable with respect to $\mathcal{Z}_{k-1}^{[k]^*}$.*

Proof. (1) Let f_ϵ, $\epsilon \in V_k^*$, be bounded functions on X and assume that $\mathbb{E}(f_\zeta \mid \mathcal{Z}_{k-2}) = 0$ for some $\zeta \in V_k^*$. Set $f_{\mathbf{0}} = 1$. By Lemma 4.8,

$$\int \bigotimes_{\epsilon \in V_k^*} f_\epsilon \, d\mu^{[k]^*} = \int \bigotimes_{\epsilon \in V_k} f_\epsilon \, d\mu^{[k]} = 0 \ .$$

(2) Let f_ϵ, $\epsilon \in V_k^*$, be bounded functions on X and assume that $\mathbb{E}(f_\zeta \mid \mathcal{Z}_{k-1}) = 0$ for some $\zeta \in V_k^*$. Define $f_{\mathbf{0}} = 1$ and 2^k functions on X by $g_{\epsilon 0} = g_{\epsilon 1} = f_\epsilon$ for $\epsilon \in V_k$. Then

$$\begin{aligned} \int \mathbb{E}\Big(\bigotimes_{\epsilon \in V_k^*} f_\epsilon \mid \mathcal{I}^{[k]^*}\Big)^2 \, d\mu^{[k]^*} &= \int \mathbb{E}\Big(\bigotimes_{\epsilon \in V_k} f_\epsilon \mid \mathcal{I}^{[k]}\Big)^2 \, d\mu^{[k]} \\ &= \int \bigotimes_{\eta \in V_{k+1}} g_\eta \, d\mu^{[k+1]} = 0 \end{aligned}$$

by Lemma 4.8, and the result follows.
(3) Let f_ϵ, $\epsilon \in V_k^*$, be bounded functions on X and assume that $\mathbb{E}(f_\zeta \mid \mathcal{Z}_{k-1}) = 0$ for some $\zeta \in V_k^*$. By definition of the factor $\mathcal{Z}_{k-1}$, there exists a bounded function $f_{\mathbf{0}}$ on X, measurable with respect to $\mathcal{Z}_{k-1}$, with

$$f_{\mathbf{0}}(x_{\mathbf{0}}) = \mathbb{E}\Big(\prod_{\epsilon \in V_k^*} f_\epsilon(x_\epsilon) \mid \mathcal{J}^{[k]^*}\Big)(\tilde{x}) \text{ for } \mu^{[k]} \text{ almost every } \mathbf{x} = (x_{\mathbf{0}}, \tilde{x}) \ .$$

As the measure $\mu^{[k]}$ is relatively independent with respect to $\mathcal{Z}_{k-1}$ and $\mathbb{E}(f_\zeta \mid \mathcal{Z}_{k-1}) = 0$ we have

$$\begin{aligned} 0 = \int \prod_{\epsilon \in V_k} f_\epsilon(x_\epsilon) \, d\mu^{[k]}(\mathbf{x}) &= \int f_{\mathbf{0}}(x_{\mathbf{0}}) \mathbb{E}\Big(\prod_{\epsilon \in V_k^*} f_\epsilon(x_\epsilon) \mid \mathcal{J}^{[k]^*}\Big)(\tilde{x}) \, d\mu^{[k]}(x_{\mathbf{0}}, \tilde{x}) \\ &= \int \Big|\mathbb{E}\Big(\prod_{\epsilon \in V_k^*} f_\epsilon(x_\epsilon) \mid \mathcal{J}^{[k]^*}\Big)(\tilde{x})\Big|^2 \, d\mu^{[k]^*}(\tilde{x}) \end{aligned}$$

and the result follows. □

4.6. **Systems of order k.** By Corollary 4.4, the factors $Z_k(X)$ form an increasing sequence of factors of X.

Definition 4.10. An ergodic system (X, μ, T) is *of order k* for an integer $k \geq 0$ if $X = Z_k(X)$.

A system might not be of order k for any integer $k \geq 1$, but we show that any system contains a factor of order k for any integer $k \geq 1$. These factors may all be the trivial system, for example if X is weakly mixing. By Equation (14), a system of order k is also of order ℓ for any integer $\ell > k$. Moreover, for an ergodic system X and any integer k, the factor $Z_k(X)$ is a system of order k.

Systems of order 1 are ergodic rotations, while systems of order 2 are ergodic quasi-affine systems (see [HK01]).

Proposition 4.11. (1) *A factor of a system of order k is of order k.*
(2) *Let X be an ergodic system and Y be a factor of X. If Y is a system of order k, then it is a factor of $Z_k(X)$.*
(3) *An inverse limit of a sequence of systems of order k is of order k.*

Properties (1) and (2) make it natural to refer to $Z_k(X)$ as the *maximal factor of order k* of X.

Proof. The first two assertions follow immediately from Proposition 4.6.

Let $X = \varprojlim X_i$ be an inverse limit of a family of systems of order k and let f be a bounded function on X. If f is measurable with respect to $\mathcal{X}_j$ for some j, then (using the same notation as above) by Definition 4.1 there exists a function F on $X^{[k]^*}$ such that $f(x_{\mathbf{0}}) = F(\tilde{x})$ $\mu^{[k]}$-almost everywhere. By density, the same result holds for any bounded function on X and the result follows by using Definition 4.1 once again. □

Using the characterization of $Z_k(X)$ in Lemma 4.3, we have:

Corollary 4.12. *An ergodic system (X, μ, T) is of order k if and only if $\|\|f\|\|_{k+1} \neq 0$ for every non-zero bounded function f on X .*

5. A group associated to each ergodic system

In this Section, we associate to each ergodic system X a group $\mathcal{G}(X)$ of measure preserving transformations of X. The most interesting case will be when X is of order k for some k. Our ultimate goal is to show that for a large class of systems of order k, the group $\mathcal{G}(X)$ is a nilpotent Lie group and acts transitively on X (Theorems 10.1 and 10.5).

Definition 5.1. Let (X, μ, T) be an ergodic system. We write $\mathcal{G}(X)$ or $\mathcal{G}$ for the group of measure preserving transformations $x \mapsto g \cdot x$ of X which satisfy for every integer $\ell > 0$ the property:

($\mathcal{P}_\ell$) The transformation $g^{[\ell]}$ of $X^{[\ell]}$ leaves the measure $\mu^{[\ell]}$ invariant and acts trivially on the invariant σ-algebra $\mathcal{I}^{[\ell]}(X)$.

$\mathcal{G}(X)$ is endowed with the topology of convergence in probability. This means that when $\{g_n\}$ is a sequence in $\mathcal{G}$ and $g \in \mathcal{G}$, we have $g_n \to g$ if and only if $\mu(g_i \cdot A \,\Delta\, g \cdot A) \to 0$ for every $A \subset X$. An equivalent condition is that for every $f \in L^2(\mu)$, $f \circ g_n \to f \circ g$ in $L^2(\mu)$.

The last condition of $\mathcal{P}_\ell$ means that the transformation $g^{[\ell]}$ leaves each set in $\mathcal{I}^{[\ell]}$ invariant, up to a $\mu^{[\ell]}$-null set.

We begin with a few remarks. Let (X, μ, T) be an ergodic system.

ı) The transformation T itself belongs to $\mathcal{G}(X)$.

ıı) $\mathcal{G}(X)$ is a Polish group.

ııı) Let $p\colon (X, \mu, T) \to (Y, \nu, S)$ be a factor map. Let $g \in \mathcal{G}(X)$ be such that g maps $\mathcal{Y}$ to itself. In other words, there exists a measure preserving transformation $h : y \mapsto h \cdot y$ of Y, with $h \circ p = p \circ g$. For every ℓ, the map $p^{[\ell]}\colon (X^{[\ell]}, \mu^{[\ell]}, T^{[\ell]}) \to (Y^{[\ell]}, \nu^{[\ell]}, S^{[\ell]})$ is a factor map by Lemma 4.5, part (1). Thus the measure $\nu^{[\ell]}$ is invariant under $h^{[\ell]}$. As the inverse image of the σ-algebra $\mathcal{I}^{[\ell]}(Y)$ under $p^{[\ell]}$ is included in $\mathcal{I}^{[\ell]}(X)$, the transformation $h^{[\ell]}$ acts trivially on $\mathcal{I}^{[\ell]}(Y)$. Thus $h \in \mathcal{G}(Y)$.

ıν) Let g be a measure preserving transformation of X satisfying $(\mathcal{P}_\ell)$ for some ℓ and let $k < \ell$ be an integer. We choose a k-face f of V_ℓ, and write as usual $\xi_f^{[\ell]}\colon X^{[\ell]} \to X^{[k]}$ for the associated projection. The image of $\mu^{[\ell]}$ by $\xi_f^{[\ell]}$ is $\mu^{[k]}$ and we have $T^{[k]} \circ \xi_f^{[\ell]} = \xi_f^{[\ell]} \circ T^{[\ell]}$ thus ${\xi_f^{[\ell]}}^{-1}(\mathcal{I}^{[k]}) \subset \mathcal{I}^{[\ell]}$. It follows immediately that g satisfies $(\mathcal{P}_k)$. Thus Property $(\mathcal{P}_\ell)$ implies Property $(\mathcal{P}_k)$ for $k < \ell$.

5.1. **General properties.**

Lemma 5.2. *Let (X, μ, T) be an ergodic system. Then for any $k \geq 0$, every $g \in \mathcal{G}(X)$ maps the σ-algebra $\mathcal{Z}_k = \mathcal{Z}_k(X)$ to itself and thus induces a measure preserving transformation of Z_k, belonging to $\mathcal{G}(Z_k)$.*

Notation. We write $p_k g : x \mapsto p_k g \cdot x$ for this transformation. The map $p_k : \mathcal{G}(X) \to \mathcal{G}(Z_k)$ is clearly a continuous group homomorphism.

Proof. Let $g \in \mathcal{G}$ and $k \geq 0$ be an integer. Let f be a bounded function on X with $\mathbb{E}(f \mid \mathcal{Z}_k) = 0$. By Lemma 4.3 and the definition (10) of the seminorm,

$$0 = |\!|\!| f |\!|\!|_{k+1}^{2^{k+1}} = \int_{X^{[k+1]}} \bigotimes_{\epsilon \in V_{k+1}} f \, d\mu^{[k+1]} = \int_{X^{[k+1]}} \bigotimes_{\epsilon \in V_{k+1}} f \circ g \, d\mu^{[k+1]} .$$

Since $g^{[k+1]}$ leaves the measure $\mu^{[k+1]}$ invariant, we have $|\!|\!| f \circ g |\!|\!|_{k+1} = 0$ and $\mathbb{E}(f \circ g \mid \mathcal{Z}_k) = 0$. By using the same argument with g^{-1} substituted for g, we have that $\mathbb{E}(f \circ g \mid \mathcal{Z}_k) = 0$ implies $\mathbb{E}(f \mid \mathcal{Z}_k) = 0$. We deduce that $g \cdot \mathcal{Z}_k = \mathcal{Z}_k$. Thus g induces a transformation of Z_k. By Remark *iii*) above, this transformation $p_k g$ belongs to $\mathcal{G}(Z_k)$. $\square$

Notation. Let G be a group. Let $k \geq 1$ be an integer and let α be a face of V_k. Analogous to the definition of the side transformations, for $g \in G$ we also write $g_\alpha^{[k]}$ for the element of $G^{[k]}$ given by

$$\big(g_\alpha^{[k]}\big)_\epsilon = g \text{ if } \epsilon \in \alpha \ ; \ \big(g_\alpha^{[k]}\big)_\epsilon = 1 \text{ otherwise.}$$

When G acts on a space X, we write also $g_\alpha^{[k]}$ for the transformation of $X^{[k]}$ associated to this element of $G^{[k]}$: For $\mathbf{x} \in X^{[k]}$,

$$\big(g_\alpha^{[k]} \cdot \mathbf{x}\big)_\epsilon = \begin{cases} g \cdot x_\epsilon & \text{if } \epsilon \in \alpha \\ x_\epsilon & \text{otherwise.} \end{cases}$$

Lemma 5.3. *Let (X, μ, T) be an ergodic system and let $0 \leq \ell < k$ be integers. For a measure preserving transformation $g : x \mapsto g \cdot x$ of X, the following are equivalent:*

1. *For any ℓ-face α of V_k, the transformation $g_\alpha^{[k]}$ of $X^{[k]}$ leaves the measure $\mu^{[k]}$ invariant and maps the σ-algebra $\mathcal{I}^{[k]}$ to itself.*
2. *For any $(\ell+1)$-face β of V_{k+1} the transformation $g_\beta^{[k+1]}$ leaves the measure $\mu^{[k+1]}$ invariant.*
3. *For any $(\ell+1)$-face γ of V_k the transformation $g_\gamma^{[k]}$ leaves the measure $\mu^{[k]}$ invariant and acts trivially on the σ-algebra $\mathcal{I}^{[k]}$.*

Proof. We note first that if any one of these properties holds for a face, then by permuting the coordinates, it holds for any face of the same dimension.
(1) $\Longrightarrow$ (2). Let α be an ℓ-face of V_k. The transformation $g_\alpha^{[k]}$ preserves the measure $\mu^{[k]}$ and the σ-algebra $\mathcal{I}^{[k]}$, thus commutes with the conditional expectation on this σ-algebra. For any bounded function F on $X^{[k]}$, we have $\mathbb{E}(F \mid \mathcal{I}^{[k]}) \circ g_\alpha^{[k]} =$

$\mathbb{E}(F \circ g_\alpha^{[k]} \mid \mathcal{I}^{[k]})$. So, for bounded functions F', F'' on $X^{[k]}$,

$$\begin{aligned}
\int_{X^{[k+1]}} &(F' \otimes F'') \circ (g_\alpha^{[k]} \times g_\alpha^{[k]})\, d\mu^{[k+1]} \\
&= \int_{X^{[k]}} \mathbb{E}(F' \circ g_\alpha^{[k]} \mid \mathcal{I}^{[k]}) \cdot \mathbb{E}(F'' \circ g_\alpha^{[k]} \mid \mathcal{I}^{[k]})\, d\mu^{[k]} \\
&= \int_{X^{[k]}} \mathbb{E}(F' \mid \mathcal{I}^{[k]}) \circ g_\alpha^{[k]} \cdot \mathbb{E}(F'' \mid \mathcal{I}^{[k]}) \circ g_\alpha^{[k]}\, d\mu^{[k]} \\
&= \int_{X^{[k]}} \mathbb{E}(F' \mid \mathcal{I}^{[k]}) \cdot \mathbb{E}(F'' \mid \mathcal{I}^{[k]})\, d\mu^{[k]} \\
&= \int_{X^{[k+1]}} F' \otimes F''\ d\mu^{[k+1]}
\end{aligned}$$

and the measure $\mu^{[k+1]}$ is invariant under $g_\alpha^{[k]} \times g_\alpha^{[k]}$. But this transformation is $g_\beta^{[k+1]}$ for some $(\ell+1)$-face β of V_{k+1} and so Property (2) follows.

(2) $\Longrightarrow$ (3). Let γ be an $(\ell+1)$-face of V_k. Under the bijection between V_k and the first k-face of V_{k+1}, γ corresponds to an $(\ell+1)$-face β of V_{k+1}. Under the usual identification of $X^{[k+1]}$ with $X^{[k]} \times X^{[k]}$, we have $g_\beta^{[k+1]} = g_\gamma^{[k]} \times \mathrm{Id}^{[k]}$. Since the measure $\mu^{[k+1]}$ is invariant under $g_\beta^{[k+1]}$ and each of its projections on $X^{[k]}$ is equal to $\mu^{[k]}$, this last measure is invariant under $g_\gamma^{[k]}$. For a bounded function F on $X^{[k]}$, measurable with respect to $\mathcal{I}^{[k]}$, we have

$$\begin{aligned}
\|F\|^2_{L^2(\mu^{[k]})} &= \int F \otimes F\, d\mu^{[k+1]} = \int (F \otimes F) \circ g_\beta^{[k+1]}\, d\mu^{[k+1]} \\
&= \int (F \circ g_\gamma^{[k]}) \otimes F\, d\mu^{[k+1]} = \int \mathbb{E}(F \circ g_\gamma^{[k]} \mid \mathcal{I}^{[k]}) \cdot F\, d\mu^{[k]}\ .
\end{aligned}$$

Thus $\mathbb{E}(F \circ g_\gamma^{[k]} \mid \mathcal{I}^{[k]}) = F$ and $F \circ g_\gamma^{[k]} = F$. Property (3) is proven.

(3) $\Longrightarrow$ (1). Let α be an ℓ-face of V_k and let γ be an $(\ell+1)$-face of V_k. Since $g_\gamma^{[k]}$ acts trivially on $\mathcal{I}^{[k]}$, by using the definition of the conditional expectation we have $\mathbb{E}(F \circ g_\gamma^{[k]} \mid \mathcal{I}^{[k]}) = \mathbb{E}(F \mid \mathcal{I}^{[k]})$ for any bounded function F on $X^{[k]}$. By the definition of the measure $\mu^{[k+1]}$, this measure is invariant under $g_\gamma^{[k]} \times \mathrm{Id}^{[k]}$. But this transformation is equal to $g_\beta^{[k+1]}$ for some $(\ell+1)$-face β of V_{k+1}. By permuting coordinates, the measure $\mu^{[k+1]}$ is invariant under $g_\beta^{[k+1]}$ for every $(\ell+1)$-face β of V_{k+1}. As the transformation $g_\alpha^{[k]} \times g_\alpha^{[k]}$ is a transformation of this kind, it leaves the measure $\mu^{[k+1]}$ invariant. By projection, the measure $\mu^{[k]}$ is invariant under $g_\alpha^{[k]}$.

Let F be a bounded function on $X^{[k]}$, measurable with respect to $\mathcal{I}^{[k]}$. Then

$$\begin{aligned}
&\|\mathbb{E}(F \circ g_\alpha^{[k]} \mid \mathcal{I}^{[k]})\|^2_{L^2(\mu^{[k]})} = \\
&\qquad \int (F \circ g_\alpha^{[k]}) \otimes (F \circ g_\alpha^{[k]})\, d\mu^{[k+1]} = \int (F \otimes F) \otimes (g_\alpha^{[k]} \times g_\alpha^{[k]})\, d\mu^{[k+1]} \\
&= \int F \otimes F\, d\mu^{[k+1]} = \|\mathbb{E}(F \mid \mathcal{I}^{[k]})\|^2_{L^2(\mu^{[k]})} = \|F\|^2_{L^2(\mu^{[k]})} = \|F \circ g_\alpha^{[k]}\|^2_{L^2(\mu^{[k]})}
\end{aligned}$$

and this means that $F \circ g_\alpha^{[k]}$ is measurable with respect to $\mathcal{I}^{[k]}$. $\square$

By applying this Lemma with $\ell = k-1$ we get some characterizations of the group $\mathcal{G}(X)$:

Corollary 5.4. *Let (X,μ,T) be an ergodic system and $g : x \mapsto g\cdot x$ a measure preserving transformation of X. The following are equivalent:*

1. *For every integer $k>0$ and every side α of V_k the measure $\mu^{[k]}$ is invariant under $g_\alpha^{[k]}$.*
2. *For every integer $k>0$ and every side α of V_k, the measure $\mu^{[k]}$ is invariant under $g_\alpha^{[k]}$ and this transformation maps the σ-algebra $\mathcal{I}^{[k]}$ to itself.*
3. *$g \in \mathcal{G}(X)$.*

By an automorphism of the system (X,μ,T), we mean a measure preserving transformation of X that commutes with T.

Lemma 5.5. *Let (X,μ,T) be an ergodic system. Then every automorphism of X belongs to $\mathcal{G}(X)$.*

Moreover, if $g : x\mapsto g\cdot x$ is an automorphism of X acting trivially on $Z_\ell(X)$ for some integer $\ell\geq 0$, then for every integer $k>0$ the measure $\mu^{[\ell+k]}$ is invariant under $g_\alpha^{[\ell+k]}$ for every $(k-1)$-face α of $V_{\ell+k}$.

Proof. Let g be an automorphism of X as in the second part of the Lemma. We use the formula (4) for $\mu^{[\ell+1]}$ and the expression given by Lemma 3.1 for $\mu^{[\ell+k]}$:

$$\mu^{[\ell+1]} = \int_{\Omega_{\ell+1}} \mu_\omega^{[\ell+1]}\, dP_{\ell+1}(\omega) \text{ and } \mu^{[\ell+k]} = \int_{\Omega_{\ell+1}} \bigl(\mu_\omega^{[\ell+1]}\bigr)^{[k-1]}\, dP_{\ell+1}(\omega)\ .$$

As $\mu^{[\ell+1]}$ is relatively independent over $Z_\ell^{[\ell+1]}$ and g acts trivially on Z_ℓ, we get that the measure $\mu^{[\ell+1]}$ is invariant under $g_\epsilon^{[\ell+1]}$ for any vertex $\epsilon\in V_{\ell+1}$. As the transformation $g_\epsilon^{[\ell+1]}$ commutes with $T^{[\ell+1]}$, it induces a measure preserving transformation h of $\Omega_{\ell+1}$. Moreover, for $P_{\ell+1}$-almost every $\omega\in\Omega_{\ell+1}$, the image of $\mu_\omega^{[\ell+1]}$ under $g_\epsilon^{[\ell+1]}$ is $\mu_{h\cdot\omega}^{[\ell+1]}$. It follows that the measure $\mu^{[\ell+k]}$ is invariant under the transformation $g_\epsilon^{[\ell+1]}\times\cdots\times g_\epsilon^{[\ell+1]}$ (2^{k-1} times). But this transformation is $g_\alpha^{[\ell+k]}$, for some $(k-1)$-face α of $V_{k+\ell}$.

The second part of the Lemma follows by permutation of coordinates. The first part of the Lemma follows from the second part with $\ell=0$ and Corollary 5.4. $\square$

5.2. **Faces and commutators.** We need some algebraic preliminaries.

Definition 5.6. Let G be a Polish group written with multiplicative notation. For every integer $k\geq 0$, $G^{[k]}$ is endowed with the product topology. For $0\leq\ell\leq k$, we write $G_\ell^{[k]}$ for the *closed* subgroup of $G^{[k]}$ spanned by

$$\{g_\alpha^{[k]} : g\in G \ \text{ and } \alpha \text{ is an } \ell\text{-face of } V_k\}\ . \tag{18}$$

Thus $G_0^{[k]} = G^{[k]}$ and $G_k^{[k]}$ is the diagonal subgroup $\{(g,g,\dots,g) : g\in G\}$ of $G^{[k]}$. We call $G_{k-1}^{[k]}$ the *side subgroup* and $G_1^{[k]}$ the *edge subgroup* of $G^{[k]}$.

For $j\geq 0$, $G^{(j)}$ denotes the *closed* jth iterated commutator subgroup of G (see Appendix A). Thus $G^{(0)}=G$, $G^{(1)}=G'$ is the closed commutator subgroup of G, and so on.

Lemma 5.7. *Let G be a Polish group. For integers $0\leq j<k$, the jth iterated commutator subgroup of $G_{k-1}^{[k]}$ contains $(G^{(j)})_{k-j-1}^{[k]}$.*

Actually equality holds, but we omit the proof as this fact is not needed.

Proof. For $g, h \in G$ and faces α, β of V_k, an immediate computation gives

$$\left[g_\alpha^{[k]}; h_\beta^{[k]}\right] = [g; h]_{\alpha\cap\beta}^{[k]} \ . \tag{19}$$

For $j = 0$ the statement of the Lemma is trivial. For $j > 0$ the statement is proved by induction. Every $(k-j-1)$-face γ of V_k can be written as the intersection of a side α and a $(k-j)$-face β. By using Equation (19) we get the result. □

Corollary 5.8. *Let (X, μ, T) be an ergodic system and $\mathcal{G} = \mathcal{G}(X)$. Then, for integers $0 \leq j < k$, any $g \in \mathcal{G}^{(j)}$ and any $(k-j-1)$-face α of V_k, the map $g_\alpha^{[k]}$ leaves the measure $\mu^{[k]}$ invariant and maps the σ-algebra $\mathcal{I}^{[k]}$ to itself.*

Proof. Let $k \geq 1$ and $\mathcal{H}$ be the subgroup of $\mathcal{G}^{[k]}$ consisting of the transformations $\mathbf{g} = (g_\epsilon : \epsilon \in V_k)$ of $X^{[k]}$ that leave the measure $\mu^{[k]}$ invariant and maps the σ-algebra $\mathcal{I}^{[k]}$ to itself. By Corollary 5.4, $\mathcal{H}$ contains the side group $\mathcal{G}_{k-1}^{[k]}$. By Lemma 5.7, $\mathcal{H}$ contains $(\mathcal{G}^{(j)})_{k-j-1}^{[k]}$ for $0 \leq j < k$. □

Corollary 5.9. *If (X, μ, T) is a system of order k, then the group $\mathcal{G}(X)$ is k-step nilpotent.*

Proof. Let $g \in \mathcal{G}^{(k)}$. By Corollary 5.8, for any vertex $\epsilon \in V_{k+1}$, the measure $\mu^{[k+1]}$ is invariant under $g_\epsilon^{[k+1]}$. Let f be a bounded function on X. Then

$$|||f \circ g - f|||_{k+1}^{2^{k+1}} = \int \prod_{\epsilon \in V_{k+1}} \big(f(g \cdot x_\epsilon) - f(x_\epsilon)\big)\, d\mu^{[k+1]} \ .$$

All 2^{k+1} integrals obtained by expanding the right side of this equality are equal up to sign and so this expression is zero. By Corollary 4.12, $f = f \circ g$ so g acts trivially on X, thus is the identity element of $\mathcal{G}$. The group $\mathcal{G}^{(k)}$ is trivial. □

Corollary 5.10. *Let (X, μ, T) be a system of order k and u an automorphism of X inducing the trivial transformation on $Z_{k-1}(X)$. Then u belongs to the center of $\mathcal{G}(X)$.*

Proof. u belongs to $\mathcal{G}(X)$ by Lemma 5.5. Let $g \in \mathcal{G}$. Let ϵ be a vertex of V_{k+1}. We choose an edge α and a side β of V_{k+1} with $\epsilon = \alpha \cap \beta$. By Lemma 5.5, $\mu^{[k+1]}$ is invariant under $u_\alpha^{[k+1]}$. By Corollary 5.4 this measure is invariant under $g_\beta^{[k+1]}$. Thus this measure is invariant under $[u_\alpha^{[k+1]}; g_\beta^{[k+1]}] = [u; g]_\epsilon^{[k+1]}$. We conclude as in the proof of the preceding Corollary that $[u; g]$ is the identity. □

6. Relations between consecutive factors

We study here the relations between the factors $Z_{k-1}(X)$ of a given ergodic system (X, μ, T). For each integer $k > 1$, $Z_k(X)$ is an extension of $Z_{k-1}(X)$. We show first that this extension is isometric, then that it is an extension by a compact abelian group. We then describe this extension more completely.

6.1. **Isometric extensions.** We recall (see [FW96]) that an ergodic isometric extension W of a system (Y,μ,S) can be written $(Y\times G/H,\mu\times\lambda,S)$ where:

- G is a compact (metrizable) group and H is a closed subgroup.
- $\lambda=m_{G/H}$ is the Haar measure on G/H. That is, λ is the unique probability measure on G/H which is invariant under the action of G by left translations. It is also the image of the Haar measure m_G of G under the natural projection $G\mapsto G/H$.
- $\rho=Y\to G$ is a cocycle and $S\colon Y\times G/H\to Y\times G/H$ is given by $S(y,u)=(Ty,\rho(y)u)$, where the left action of G on G/H is written $(g,u)\mapsto gu$.

Without loss, we can reduce to the case that the action of G on G/H is faithful, meaning that H does not contain any nontrivial normal subgroup of G. Moreover, we can assume that the the cocycle $\rho\colon Y\to G$ is ergodic, meaning that the system $(Y\times G,\mu\times m_G,T_\rho)$ is ergodic. As usual, $T_\rho(y,g)=(Ty,\rho(y)g)$.

To every $g\in G$ we associate a measure preserving transformation $x\mapsto g\cdot x$ of W by

$$g\cdot(y,u)=(y,gu)\ .$$

We also denote this transformation by g.

Any factor of $W=Y\times G/H$ over Y has the form $Y\times G/L$, for some closed subgroup L of G containing H. In particular, the action of $g\in G$ on W induces a measure preserving transformation on this factor, written with the same notation.

Lemma 6.1. *Let $W=Y\times G/H$ be an ergodic isometric extension of Y so that the corresponding extension $Y\times G$ is ergodic. Then, for every $g\in G$, $g^{[1]}=g\times g$ acts trivially on the invariant σ-algebra $\mathcal{I}^{[1]}(W)$ of $W\times W$.*

Proof. Let T denote the transformation on W. Consider the factor K of W spanned by Y and the Kronecker factor $Z_1(W)$ of W. Then K is an extension of Y by a compact abelian group. Therefore, $K=Y\times G/L$ for some closed subgroup L of G containing H and containing the commutator subgroup G' of G. Thus, for any $g\in G$, the action of g on K commutes with T and it induces an automorphism of the Kronecker factor $Z_1(K)=Z_1(W)$.

But an automorphism of an ergodic rotation is itself a rotation. By the description in Section 3.2 of the invariant sets of $W\times W$, the result follows. ☐

6.2. Z_k **is an abelian group extension of** Z_{k-1}**.**

Lemma 6.2. *Let (X,μ,T) be an ergodic system and let $k\geq 2$ an integer. Then Z_k is an isometric extension of Z_{k-1}.*

Proof. Let Y be the maximal isometric extension of Z_{k-1} which is a factor of X (see [FW96]).

We consider $(X^{[k]},\mu^{[k]},T^{[k]})$ as a joining of (X,μ,T) and $(X^{[k]^*},\mu^{[k]^*},T^{[k]^*})$ and recall that this joining is relatively independent with respect to the common factor $\mathcal{Z}_{k-1}=\mathcal{I}^{[k]^*}$ of these two systems. It is then classical that the invariant σ-algebra $\mathcal{I}^{[k]}$ of $(X^{[k]},\mu^{[k]},T^{[k]})$ is measurable with respect to $\mathcal{Y}\otimes\mathcal{X}^{[k]^*}$.

Let f be a bounded function on X with $\mathbb{E}(f\mid\mathcal{Y})=0$. Write F for the function $\mathbf{x}\mapsto\prod_{\epsilon\in V_k}f(x_\epsilon)$ on $X^{[k]}$. Since $\mu^{[k]}$ is relatively independent with respect to $\mathcal{Z}_{k-1}^{[k]}$ and $\mathcal{Y}\supset\mathcal{Z}_{k-1}$, F has zero conditional expectation on the σ-algebra $\mathcal{Y}\otimes\mathcal{X}^{[k]^*}$ and so zero conditional expectation on $\mathcal{I}^{[k]}$. With the usual identification of $X^{[k+1]}$ with the Cartesian square of $X^{[k]}$, we have $\int_{X^{[k+1]}}F(\mathbf{x}')F(\mathbf{x}'')\,d\mu^{[k+1]}(\mathbf{x}',\mathbf{x}'')=0$. That is, $\|\!|f|\!\|_{k+1} = 0$ by definition of this seminorm and $\mathbb{E}(f \mid \mathcal{Z}_k) = 0$ by Lemma 4.3. Therefore $\mathcal{Z}_k \subset \mathcal{Y}$. ☐

Proposition 6.3. *Let (X, μ, T) be a system of order $k \geq 2$.*

(1) *X is a compact abelian group extension of Z_{k-1}, written $X = Z_{k-1} \times U$, where U is a compact abelian group.*
(2) *For every $u \in U$ and every edge α of V_k, the transformation $u_\alpha^{[k]}$ acts trivially on $\mathcal{I}^{[k]}$.*

Proof. By Lemma 6.2, X is an isometric extension of Z_{k-1} and so we can write $X = Z_{k-1} \times (G/H)$, where G is a compact group and H a closed subgroup. As in Section 6.1 we write $\rho\colon Z_{k-1} \to G$ for the cocycle defining this extension and let λ denote the Haar measure of G/H.

Since $\mu^{[k]}$ is relatively independent with respect to $Z_{k-1}^{[k]}$, this measure is invariant under the map $g_\epsilon^{[k]}$ for any $g \in G$ and any $\epsilon \in V_k$. A fortiori, it is invariant under $g_\alpha^{[k]}$ for any $g \in G$ and any edge α of V_k.

Claim: *For any $g \in G$ and any edge α of V_k, the transformation $g_\alpha^{[k]}$ acts trivially on $\mathcal{I}^{[k]}$.*

Consider the ergodic decompositions of $\mu^{[k-1]}$ and $\mu_{k-1}^{[k-1]}$ as in Section 3.1. Since $\mathcal{I}^{[k-1]}$ is measurable with respect to $\mathcal{Z}_{k-1}^{[k-1]}$, these decompositions can be written as

$$\mu^{[k-1]} = \int_{\Omega_{k-1}} \mu_\omega^{[k-1]}\, dP_{k-1}(\omega) \ \text{ and } \ \mu_{k-1}^{[k-1]} = \int_{\Omega_{k-1}} \mu_{k-1,\omega}^{[k-1]}\, dP_{k-1}(\omega)\ ,$$

where $\mu_{k-1,\omega}^{[k-1]}$ is the projection of $\mu_\omega^{[k-1]}$ on $Z_{k-1}^{[k-1]}$.

By Part (1) of Proposition 4.7, $(X^{[k-1]}, \mu^{[k-1]}, T^{[k-1]})$ is the relatively independent joining of 2^{k-1} copies of (X, T, μ) over $Z_{k-1}^{[k-1]}$. Thus we can identify $X^{[k-1]}$ with $Z_{k-1}^{[k-1]} \times \big(G^{[k-1]}/H^{[k-1]}\big)$. The measure $\mu^{[k-1]}$ is the product of $\mu_{k-1}^{[k-1]}$ by the 2^{k-1}-power $\lambda^{\otimes[k-1]}$ of λ, which is the Haar measure of $G^{[k-1]}/H^{[k-1]}$ and $X^{[k-1]}$ is the isometric extension of $Z_{k-1}^{[k-1]}$ given by the cocycle $\rho^{[k-1]}\colon Z_{k-1}^{[k-1]} \to G^{[k-1]}$.

So for almost every $\omega \in \Omega_{k-1}$, the system $(X^{[k-1]}, \mu_\omega^{[k-1]}, T^{[k-1]})$ is an isometric extension of $(Z_{k-1}^{[k-1]}, \mu_{k-1,\omega}^{[k-1]}, T^{[k-1]})$, with fiber $G^{[k-1]}/H^{[k-1]}$.

Let $g \in G$ and let $\epsilon \in V_{k-1}$ be a vertex. Since $g_\epsilon^{[k-1]}$ belongs to $G^{[k-1]}$, by Lemma 6.1 the transformation $g_\epsilon^{[k-1]} \times g_\epsilon^{[k-1]}$ of $X^{[k]} = X^{[k-1]} \times X^{[k-1]}$ acts trivially on the $T^{[k]} = T^{[k-1]} \times T^{[k-1]}$ invariant σ-algebra of $(X^{[k]}, \mu_\omega^{[k-1]} \times \mu_\omega^{[k-1]}, T^{[k]})$.

We recall (see Formula 5) that

$$\mu^{[k]} = \int_{\Omega_{k-1}} \mu_\omega^{[k-1]} \times \mu_\omega^{[k-1]}\, dP(\omega)\ .$$

Thus $g_\epsilon^{[k-1]} \times g_\epsilon^{[k-1]}$ acts trivially on the invariant σ-algebra $\mathcal{I}^{[k]}$. But $g_\epsilon^{[k-1]} \times g_\epsilon^{[k-1]}$ is equal to $g_\alpha^{[k]}$ for some edge α of V_k. The claim follows by permuting the coordinates.

Claim: *G is abelian.*

Let $g, h \in G$, and let ϵ be a vertex of V_{k+1}. Choose two edges α and β of V_{k+1} with $\alpha \cap \beta = \epsilon$. By Equation (19), $\big[g_\alpha^{[k]}; h_\beta^{[k]}\big] = [g; h]_\epsilon^{[k]}$. By the first step and Lemma 5.3, the transformations $g_\alpha^{[k+1]}$ and $h_\beta^{[k+1]}$ preserve the measure $\mu^{[k+1]}$, thus also the transformation $[g;h]_\epsilon^{[k+1]}$. As this holds for every vertex ϵ, we conclude as in the proof of Corollary 5.9 that $[g;h]$ acts trivially on X. This means that $[g;h]=1$ and so G is abelian.

By our hypotheses the group H is trivial, and the proof is complete. □

6.3. **Description of the extension.**

Notation. For $k\geq 1$ and $\epsilon\in V_k$, we write

$$|\epsilon|=\epsilon_1+\epsilon_2+\cdots+\epsilon_k \text{ and } s(\epsilon)=(-1)^{|\epsilon|}\ .$$

Let X be a set, U an abelian group written with additive notation and $f\colon X\to U$ a map. For every $k\geq 1$, we define a map $\Delta^k f\colon X^{[k]}\to U$ by:

$$\Delta^k f(\mathbf{x})=\sum_{\epsilon\in V_k} s(\epsilon)f(x_\epsilon)\ .$$

In particular, Δf is the map defined on X^2 by $\Delta f(x',x'')=f(x')-f(x'')$. We have similar notation when the group is written with multiplicative notation.

Proposition 6.4. *Let (X,μ,T) be a system of order $k\geq 2$. By Proposition 6.3, X is an extension of Z_{k-1} by a compact abelian group U for some cocycle $\rho\colon Z_{k-1}\to U$. Then*

(1) *$\Delta^k\rho\colon Z_{k-1}^{[k]}\to U$ is a coboundary (see Appendix C.2) of the system $(Z_{k-1}^{[k]},\mu_{k-1}^{[k]},T^{[k]})$, meaning that there exists $F\colon Z_{k-1}^{[k]}\to U$ with*

$$\Delta^k\rho=F\circ T^{[k]}-F\ . \tag{20}$$

(2) *The σ-algebra $\mathcal{I}^{[k]}(X)$ is spanned by the σ-algebra $\mathcal{I}^{[k]}(Z_{k-1})$ and the map $\Phi\colon X^{[k]}\to U$ given by*

$$\Phi(\mathbf{y},\mathbf{u})=F(\mathbf{y})-\sum_{\epsilon\in V_k}s(\epsilon)u_\epsilon \tag{21}$$

for $\mathbf{y}\in Z_{k-1}^{[k]}$ and $\mathbf{u}\in U^{[k]}$ where we have identified X with $Z_{k-1}\times U$ and $X^{[k]}$ with $Z_{k-1}^{[k]}\times U^{[k]}$.

Proof. Here we consider characters of U as homomorphisms from U to the circle group $\mathcal{S}^1$, written with multiplicative notation.

(1) Let $\chi\in\widehat{U}$. Define the function ψ on $X=Z_{k-1}\times U$ by $\psi(y,u)=\chi(u)$ and the function Ψ on $X^{[k]}=Z_{k-1}^{[k]}\times U^{[k]}$ by

$$\Psi(\mathbf{y},\mathbf{u})=\chi\big(\sum_{\epsilon\in V_k}s(\epsilon)u_\epsilon\big) \text{ for } \mathbf{y}\in Y^{[k]} \text{ and } \mathbf{u}\in U^{[k]}\ .$$

Since X is of order k, $|\!|\!|\psi|\!|\!|_{k+1}\neq 0$ by Corollary 4.12 and $\mathbb{E}(\Psi\mid\mathcal{I}^{[k]})\neq 0$ by Lemma 4.3.

Let J be the linear map from $L^2(\mu_{k-1}^{[k]})$ to $L^2(\mu^{[k]})$ given by

$$Jf(\mathbf{y},\mathbf{u})=f(\mathbf{y})\Psi(\mathbf{y},\mathbf{u}) \text{ for } f\in L^2(\mu_{k-1}^{[k]}),\ \mathbf{y}\in Z_{k-1}^{[k]} \text{ and } \mathbf{u}\in U^{[k]}\ .$$

J is an isometry and its range $\mathcal{H}_\chi$ is a closed subspace of $L^2(\mu^{[k]})$. Furthermore, for $f\in L^2(\mu_{k-1}^{[k]})$,

$$J\big(\chi(\Delta^k\rho)\cdot f\circ T^{[k]}\big)=\big(Jf\big)\circ T^{[k]}$$

and so the space $\mathcal{H}_\chi$ is invariant under $T^{[k]}$. Since the function Ψ belongs to $\mathcal{H}_\chi$, the function $\mathbb{E}(\Psi \mid \mathcal{I}^{[k]})$ also belongs to this space. We get that there exists a non-identically zero function f on $Z_{k-1}^{[k]}$ with

$$\chi(\Delta^k\rho)\cdot f\circ T^{[k]} = f \qquad \mu_{k-1}^{[k]}\text{a.e.} \tag{22}$$

Let $A=\{\mathbf{y}\in Z_{k-1}^{[k]}\colon f(\mathbf{y})\neq 0\}$. Then $\mu_{k-1}(A)\neq 0$ and A is $T^{[k]}$-invariant by Equation (22). We use the ergodic decomposition given by Formula (4), but for the measure $\mu_{k-1}^{[k]}$. Since A is invariant, it corresponds to a subset B of Ω_k, with $P_k(B)\neq 0$.

Define

$$C=\bigl\{\omega\in\Omega_k\colon \chi\circ\Delta^k\rho \text{ is a coboundary of } (Z_{k-1}^{[k]},\mu_{k-1,\omega}^{[k]},T^{[k]})\bigr\}\ .$$

Then C is measurable in Ω_k and it contains B by Equation (22) and the definition of B. Thus $P_k(C)>0$. We show now that C is invariant under the group $\mathcal{T}_{k-1}^{[k]}$ of side transformations. Let $\omega\in\Omega_k$ and let α be a side of V_k not containing $\mathbf{0}$ so that $T_\alpha^{[k]}\omega\in C$. Let $\phi\colon Z_{k-1}^{[k]}\to\mathbb{T}$ be chosen with its coboundary for $T^{[k]}$ equal to $\chi\circ\Delta^k\rho$ almost everywhere for the measure $\mu^{[k]}_{k-1,T_\alpha^{[k]}\omega}$. The coboundary of $\phi\circ T_\alpha^{[k]}$ for $T^{[k]}$ is equal to $\chi\circ(\Delta^k\rho)\circ T_\alpha^{[k]}$ almost everywhere for the measure $\mu_{k-1,\omega}^{[k]}$. But the map $(\Delta^k\rho)\circ T_\alpha^{[k]}-\Delta^k\rho$ from $Y^{[k]}$ to U is the coboundary for $T^{[k]}$ of the map

$$\mathbf{y}\mapsto\sum_{\epsilon\in\alpha}s(\epsilon)\rho(y_\epsilon)\ .$$

Therefore $\chi\circ\Delta^k\rho$ is a coboundary of the system $(Z_{k-1}^{[k]},\mu_{k-1,\omega}^{[k]},T^{[k]})$ and $\omega\in C$. Thus the set C is invariant under $T_\alpha^{[k]}$. By Corollary 3.6, the action of the group $\mathcal{T}_*^{[k]}$ on Ω_k is ergodic. As $P(C)>0$, we have $P(C)=1$.

Therefore, for P_k-almost every $\omega\in\Omega_k$, $\chi\circ\Delta^k\rho$ is a coboundary of the system $(Z_{k-1}^{[k]},\mu_{k-1,\omega}^{[k]},T^{[k]})$. By Corollary C.4, $\chi\circ\Delta^k\rho$ is a coboundary of $(Z_{k-1}^{[k]},\mu_{k-1}^{[k]},T^{[k]})$. As this holds for every $\chi\in\widehat{U}$, $\Delta^k\rho$ is a coboundary of this system by Lemma C.1 and the first part of the Proposition is proven.

(2) We identify the dual group of $U^{[k]}$ with $\widehat{U}^{[k]}$. For $\boldsymbol{\theta}=(\theta_\epsilon:\epsilon\in V_k)\in\widehat{U}^{[k]}$ and $\mathbf{u}=(u_\epsilon:\epsilon\in V_k)\in U^{[k]}$,

$$\boldsymbol{\theta}(\mathbf{u})=\prod_{\epsilon\in V_k}\theta_\epsilon(u_\epsilon)\ .$$

Let $\mathcal{H}$ be the subspace of $L^2(\mu^{[k]})$ consisting in functions invariant under $T^{[k]}$. For $\boldsymbol{\theta}\in\widehat{U}^{[k]}$, we write $\mathcal{L}_{\boldsymbol{\theta}}$ for the subspace of $L^2(\mu^{[k]})$ consisting in functions of the form

$$(\mathbf{y},\mathbf{u})\mapsto f(\mathbf{y})\boldsymbol{\theta}(\mathbf{u}) \tag{23}$$

for some $f\in L^2(\mu_{k-1}^{[k]})$. As above, $\mathcal{L}_{\boldsymbol{\theta}}$ is a closed subspace of $L^2(\mu^{[k]})$, invariant under $T^{[k]}$. Since the measure $\mu^{[k]}$ is relatively independent over $\mu_{k-1}^{[k]}$, using the Fourier Transform it is immediate that $L^2(\mu^{[k]})$ is the Hilbert sum of the spaces $\mathcal{L}_{\boldsymbol{\theta}}$ for $\boldsymbol{\theta}\in\widehat{U}^{[k]}$. Therefore, the invariant subspace $\mathcal{H}$ of $L^2(\mu^{[k]})$ is the Hilbert sum of the invariant subspaces $\mathcal{H}\cap\mathcal{L}_{\boldsymbol{\theta}}$ of $\mathcal{L}_{\boldsymbol{\theta}}$.

Let $\boldsymbol{\theta} \in \widehat{U}^{[k]}$ and assume that $\mathcal{H} \cap \mathcal{L}_{\boldsymbol{\theta}}$ contains a non-identically zero function ϕ. Let $\alpha = (\epsilon, \eta)$ be an edge of V_k and let $u \in U$. By Equation (23) we have $\phi \circ u_\alpha^{[k]} = \phi \cdot \theta_\epsilon(u)\theta_\eta(u)$. But by Part (2) of Proposition 6.3, $\phi \circ u_\alpha^{[k]} = \phi$ and we get that $\theta_\epsilon(u)\theta_\eta(u) = 1$. Since this holds for every $u \in U$, $\theta_\epsilon\theta_\eta = 1$. As it holds for every edge $\alpha = (\epsilon, \eta)$, there exists $\chi \in \widehat{U}$ with $\theta_\epsilon = \chi^{s(\epsilon)}$ for every $\epsilon \in V_k$. Finally, ϕ is a function of the form

$$\phi(\mathbf{y}, \mathbf{u}) = f(\mathbf{y}) \cdot \chi\big(\sum_{\epsilon \in V_k} s(\epsilon) u_\epsilon\big)$$

for some $f \in L^2(\mu_{k-1}^{[k]})$, and

$$\phi(\mathbf{y}, \mathbf{u}) = \chi\big(-\Phi(\mathbf{y}, \mathbf{u})\big) \cdot \chi(F(\mathbf{y})) f(\mathbf{y})$$

where Φ is the map defined by Equation (21). Since Φ and ϕ are invariant under $T^{[k]}$, the function $\chi \circ F \cdot f$ is also invariant under this transformation and is measurable with respect to $\mathcal{I}^{[k]}(Z_{k-1})$. We conclude that ϕ is measurable with respect to the σ-algebra spanned by Φ and $\mathcal{I}^{[k]}(Z_{k-1})$.

Since the invariant space $\mathcal{H}$ of $L^2(\mu^{[k]})$ is the Hilbert sum of the spaces $\mathcal{H} \cap \mathcal{L}_{\boldsymbol{\theta}}$, every function in $\mathcal{H}$ is measurable with respect to this σ-algebra and the second part of the Proposition in proven. $\square$

6.4. **More terms.** The next Proposition is used only in the proof of Corollary 6.6, which in turn is only used in the proof of Lemma 10.6.

Proposition 6.5. *Let (X, μ, T) be a system of order k. Then for $\ell > k$ the invariant σ-algebra $\mathcal{I}^{[\ell]}$ is spanned by the σ-algebras ${\xi_\alpha^{[\ell]}}^{-1}\big(\mathcal{I}^{[k]}\big)$, where α is a k-face of V_ℓ.*

Proof. First Step. Let (X, μ, T) be a system of order k. We use the notations of Proposition 6.4 and the maps F and Φ defined in Equations. (20) and (21). Let $\ell > k$.

We identify $X^{[\ell]}$ with $Z_{k-1}^{[\ell]} \times U^{[\ell]}$. As the projection of $\mu^{[\ell]}$ on $Z_{k-1}^{[\ell]}$ is $\mu_{k-1}^{[\ell]}$, for $\mu_{k-1}^{[\ell]}$-almost every $\mathbf{y} \in Z_{k-1}^{[\ell]}$ there exists a measure $\lambda_{\mathbf{y}}$ on $U^{[\ell]}$ such that

$$\mu^{[\ell]} = \int_{Z_{k-1}^{[\ell]}} \delta_{\mathbf{y}} \times \lambda_{\mathbf{y}} \, d\mu_{k-1}^{[\ell]}(\mathbf{y}) \ .$$

For every $u \in U$, the corresponding vertical rotation (see the definition in Subsection C.1) is an automorphism of X and acts trivially on Z_{k-1}. By Lemma 5.5, for every $(\ell - k)$-face β of V_ℓ the measure $\mu^{[\ell]}$ is invariant under $u_\beta^{[\ell]}$. It follows that the measure $\lambda_{\mathbf{y}}$ is invariant under this transformation for $\mu_{k-1}^{[\ell]}$-almost every $\mathbf{y}$. By separability, for almost every $\mathbf{y}$ the measure $\lambda_{\mathbf{y}}$ is invariant under the translation by any element of the group $U_{\ell-k}^{[\ell]}$.

We identify $U^{[\ell]}$ with $U^{[\ell-1]} \times U^{[\ell-1]}$ and we write $\mathbf{u} = (\mathbf{u}', \mathbf{u}'')$ for an element of $U^{[\ell]}$; we write also $\mathbf{y} = (\mathbf{y}', \mathbf{y}'')$ for a point of $Z_{k-1}^{[\ell]} = Z_{k-1}^{[\ell-1]} \times Z_{k-1}^{[\ell-1]}$; and $\mathbf{x} = (\mathbf{y}', \mathbf{u}', \mathbf{y}'', \mathbf{u}'')$ for a point of $X^{[\ell]}$, with $\mathbf{y} = (\mathbf{y}', \mathbf{y}'') \in Z_{k-1}^{[\ell]}$ and $\mathbf{u} = (\mathbf{u}', \mathbf{u}'') \in U^{[\ell]}$.

Let γ be a k-face of $V_{\ell-1}$. As the map $\Phi_k \circ \xi_\gamma^{[\ell-1]} \colon X^{[\ell-1]} \to U$ is invariant, it follows from the construction of $\mu^{[\ell]}$ that $\Phi_k \circ \xi_\gamma^{[\ell-1]}(\mathbf{x}') = \Phi_k \circ \xi_\gamma^{[\ell-1]}(\mathbf{x}'')$ for

$\mu_k^{[\ell]}$-almost every $\mathbf{x}$, that is

$$\sum_{\epsilon\in\gamma} s(\epsilon)u'_\epsilon - \sum_{\epsilon\in\gamma} s(\epsilon)u''_\epsilon = F(\xi_\gamma^{[\ell-1]}\mathbf{y}') - F(\xi_\gamma^{[\ell-1]}\mathbf{y}'')\ \mu^{[\ell]}\text{-a.e.}$$

For $\mu_{k-1}^{[\ell]}$-almost every $\mathbf{y} = (\mathbf{y}', \mathbf{y}'') \in Z_{k-1}^{[\ell]}$, this identity is true for $\lambda_{\mathbf{y}}$-almost every $\mathbf{u} = (\mathbf{u}', \mathbf{u}'') \in U^{[\ell]}$ and the measure $\lambda_{\mathbf{y}}$ is concentrated on a coset of the group

$$\big\{(\mathbf{u}', \mathbf{u}'') \in U^{[\ell]} : \sum_{\epsilon\in\gamma} s(\epsilon)u'_\epsilon - \sum_{\epsilon\in\gamma} s(\epsilon)u''_\epsilon = 0\big\}\ .$$

We write δ for the $(k+1)$-face $\gamma\times\{0,1\}$ of V_ℓ, and we notice that this group is equal to

$$\big\{\mathbf{u} \in U^{[\ell]} : \sum_{\epsilon\in\delta} s(\epsilon)u_\epsilon = 0\big\} = {\xi_\delta^{[\ell]}}^{-1}(U_1^{[k+1]})\ .$$

By permutation of coordinates, the same property holds for any $k+1$-face δ of V_ℓ, and $\lambda_{\mathbf{y}}$ is concentrated on a coset of the intersection

$$\big\{\mathbf{u} \in U^{[\ell]} : \sum_{\epsilon\in\delta} s(\epsilon)u_\epsilon = 0 \text{ for every } (k+1)\text{-face } \delta \text{ of } V_\ell\big\}$$

of the corresponding subgroups of $U^{[\ell]}$. By an elementary algebraic computation, we see that this group is equal to $U_{\ell-k}^{[\ell]}$.

Finally, $\lambda_{\mathbf{y}}$ is invariant under translation by $U_{\ell-k}^{[\ell]}$ and is concentrated on a coset of this group. Thus this measure is the image of the Haar measure of this group by some translation. Moreover, for almost every $\mathbf{y} \in Z_{k-1}^{[\ell]}$, the measure $\lambda_{T^{[\ell]}\mathbf{y}}$ is the image of the measure $\lambda_{\mathbf{y}}$ by the translation by $\rho^{[\ell]}(\mathbf{y})$. We conclude that:

The system $(X^{[\ell]}, \mu^{[\ell]}, T^{[\ell]})$ is an extension of $(Z_{k-1}^{[\ell]}, \mu_{k-1}^{[\ell]}, T^{[\ell]})$ by the compact abelian group $U_{\ell-k}^{[\ell]}$.

Step 2. We keep the notation and hypotheses of the first step. It follows from the description of $\mu^{[\ell]}$ just above that the Hilbert space $L^2(\mu^{[\ell]})$ can be decomposed as in the proof of Proposition 6.4: $L^2(\mu^{[\ell]})$ is the Hilbert sum for $\boldsymbol{\theta} \in \widehat{U_{\ell-k}^{[\ell]}}$ of the subspaces

$$\mathcal{L}_{\boldsymbol{\theta}} = \big\{f(\mathbf{u}\cdot\mathbf{x}) = \boldsymbol{\theta}(\mathbf{u})f(\mathbf{x})\ \mu^{[\ell]}\text{-a.e. for every } \mathbf{u} \in U_{\ell-k}^{[\ell]}\big\}\ .$$

(Here we see characters as taking values in the circle group.) Each space $\mathcal{L}_{\boldsymbol{\theta}}$ is invariant under $T^{[\ell]}$ and thus the $T^{[\ell]}$-invariant subspace $\mathcal{H}$ of $L^2(\mu^{[\ell]})$ is the Hilbert sum of the spaces $\mathcal{H}_{\boldsymbol{\theta}} = \mathcal{H}\cap\mathcal{L}_{\boldsymbol{\theta}}$.

On the other hand, by Lemmas 5.5 and 5.3, each function in $\mathcal{H}$ is invariant under the map $\mathbf{x}\mapsto\mathbf{u}\cdot\mathbf{x}$ for any $\mathbf{u} \in U_{\ell-k+1}^{[\ell]}$. Therefore $\mathcal{H}_{\boldsymbol{\theta}}$ is trivial except if $\boldsymbol{\theta}$ belongs to the annihilator of $U_{\ell-k+1}^{[\ell]}$ in the dual group of $U_{\ell-k}^{[\ell]}$. By the same algebraic computation as above, we get that

$$U_{\ell-k+1}^{[\ell]} = \big\{\mathbf{u} \in U^{[\ell]} : \sum_{\epsilon\in\alpha} s(\epsilon)u_\epsilon = 0 \text{ for every } k\text{-face } \alpha \text{ of } V_\ell\big\}\ .$$

It follows that the annihilator of $U_{\ell-k+1}^{[\ell]}$ in $\widehat{U^{[\ell]}}$ is $(\widehat{U})_k^{[\ell]}$. Therefore, the subspace $\mathcal{H}$ of $L^2(\mu_k^{[\ell]})$ is the closed linear span of the family of invariant functions of the type

$$\phi(\mathbf{y},\mathbf{u}) = \psi(\mathbf{y})\boldsymbol{\theta}(\mathbf{u}) \text{ where } \psi \in L^2(\mu_{k-1}^{[\ell]}) \text{ and } \boldsymbol{\theta} \in \widehat{U}_k^{[\ell]}\ .$$

We consider an invariant function ϕ of this type. As $\widehat{U}_k^{[\ell]}$ is spanned by the elements of the form $\chi_\alpha^{[\ell]}$, where $\chi \in \widehat{U}$ and α is a k-face of V_ℓ, there exist k-faces $\alpha_1, \dots, \alpha_m$ of V_ℓ and characters $\chi_1, \dots, \chi_m \in \widehat{U}$ with

$$\boldsymbol{\theta}(\mathbf{u}) = \prod_{j=1}^{m} \prod_{\epsilon \in \alpha_j} \chi_j(u_\epsilon)$$

for $\mathbf{u} \in U^{[\ell]}$. For each j the function $\chi_j \circ \Phi_k \circ \xi_{\alpha_j}^{[\ell]}$ is invariant, and thus so is the function

$$\phi \cdot \prod_{j=1}^{m} \chi_j \circ \Phi_k \circ \xi_{\alpha_j}^{[\ell]}.$$

But this function factors clearly through $Z_{k-1}^{[\ell]}$ and is measurable with respect to $\mathcal{I}^{[\ell]}(Z_{k-1})$. Therefore, the function ϕ is measurable with respect to the σ-algebra spanned by $\mathcal{I}^{[\ell]}(Z_{k-1})$ and ${\xi_{\alpha_m}^{[\ell]}}^{-1}(\mathcal{I}^{[k]}(X))$, $1 \le j \le m$. We get:

The σ-algebra $\mathcal{I}^{[\ell]}(X)$ is spanned by the σ-algebras $\mathcal{I}^{[\ell]}(Z_{k-1})$ and the σ-algebras ${\xi_\alpha^{[\ell]}}^{-1}(\mathcal{I}^{[k]}(X))$, for α a k-face of V_ℓ.

Last step. We now prove the assertion of Proposition 6.5 by induction on $k \ge 0$. For $k = 0$ the system X is trivial and there is nothing to prove. We take $k > 0$ and assume that the assertion holds for every system of order $k-1$. Let X be a system of order k and let $\ell > k$. We use the notation of the first two steps. By the inductive hypothesis $\mathcal{I}^{[\ell]}(Z_{k-1})$ is spanned by the σ-algebras ${\xi_\alpha^{[\ell]}}^{-1}(\mathcal{I}^{[k]}(Z_{k-1}))$ for α a k-face of V_ℓ. But, for each α, ${\xi_\alpha^{[\ell]}}^{-1}(\mathcal{I}^{[k]}(Z_{k-1})) \subset {\xi_\alpha^{[\ell]}}^{-1}(\mathcal{I}^{[k]}(X))$ and the result follows from the conclusion of the second step. $\square$

Corollary 6.6. *Let (X, μ, T) be a system of order k and let $x \mapsto g \cdot x$ be a measure preserving transformation of X satisfying the property $(\mathcal{P}_k)$ of Definition 5.1. Then $g \in \mathcal{G}(X)$.*

Proof. We have to show that the property $(\mathcal{P}_\ell)$ holds for every ℓ. For $\ell = k$ there is nothing to prove. For $\ell < k$, $(\mathcal{P}_\ell)$ follows immediately from $(\mathcal{P}_k)$ by projection (see the fourth remark after Definition 5.1). For $\ell > k$ we proceed by induction. Let $\ell > k$ and assume that $\mathcal{P}_{\ell-1}$ holds. By Lemma 5.3, the measure $\mu^{[\ell]}$ is invariant under $g_\beta^{[\ell]}$ for any $(\ell-1)$-face β of V_ℓ and it follows immediately that it is invariant under $g^{[\ell]}$. By hypothesis, $g^{[k]}$ acts trivially on $\mathcal{I}^{[k]}$ and it follows that for every k-face α of V_ℓ the transformation $g^{[\ell]}$ acts trivially on the σ-algebra ${\xi_\alpha^{[\ell]}}^{-1}(\mathcal{I}^{[k]})$. By Proposition 6.5, $g^{[\ell]}$ acts trivially on $\mathcal{I}^{[\ell]}$. $\square$

7. Cocycles of type k and systems of order k

Notation. Let (X, μ) be a probability space and U a compact abelian group. We write $\mathcal{C}(X, U)$ for the group of measurable maps from X to U. We also write $\mathcal{C}(X)$ instead of $\mathcal{C}(X, \mathbb{T})$.

$\mathcal{C}(X, U)$ is endowed with the topology of convergence in probability. It is a Polish group.

When (X, μ, T) is a system, an element of $\mathcal{C}(X, U)$ is called an U-valued cocycle. (see Appendix C.) For the notation $\Delta^k \rho$ see Subsection 6.3.

Definition 7.1. Let $k \geq 1$ be an integer, (X, μ, T) an ergodic system, U a compact abelian group (written additively) and $\rho\colon X \to U$ a cocycle. We say that ρ is a *cocycle of type* k if the cocycle $\Delta^k \rho\colon X^{[k]} \to U$ is a coboundary of $(X^{[k]}, \mu^{[k]}, T^{[k]})$.

7.1. **First properties.** We have shown in the preceding section that for every ergodic system X and integer $k \geq 1$, $Z_k(X)$ is an extension of $Z_{k-1}(X)$ associated to a cocycle of type k.

Remark 7.2. A cocycle cohomologous to a cocycle of type k is also of type k.

By Lemma C.1 we get:

Remark 7.3. $\rho\colon X \to U$ is of type k if and only if $\chi \circ \rho\colon X \to \mathbb{T}$ is of type k for every character χ of U. It follows that for any closed subgroup V of U, a V-valued cocycle is of type k if and only if it is of type k as a U-valued cocycle.

A cocycle $\rho\colon X \to U$ is of type 1 if and only if $\rho(x) - \rho(y)$ is a coboundary on X^2. Equivalently, $\chi \circ \rho$ is a quasi-coboundary for every $\chi \in \widehat{U}$. (See Appendix C.4 for the definition and properties.) When U is a torus, this property means simply that ρ itself is a quasi-coboundary (see Lemma C.5).

Cocycles $\rho\colon X \to U$ so that $\Delta^k \rho = 0$ are obviously of type k. In the sequel we use some properties of these cocycles.

Notation. Let (X, μ, T) be an ergodic system, $k \geq 1$ be an integer, and U a compact abelian group. Let $\mathcal{D}_k(X, U)$ denote the family of cocycles $\rho\colon X \to U$ with $\Delta^k \rho = 0$.

Lemma 7.4. *Let (X, μ, T) be an ergodic system, $k \geq 1$ be an integer, and U a compact abelian group. Then $\mathcal{D}_k(X, U)$ is a closed subgroup of $\mathcal{C}(X, U)$. Moreover, it admits the group U of constant cocycles as an open subgroup.*

Proof. The first assertion is obvious. We prove the second statement by induction on k. By definition, a cocycle in $\mathcal{D}_1(X)$ is constant. Assume that the assertion holds for some $k \geq 1$. We use the formula (5) for $\mu^{[k+1]}$. ρ belongs to $\mathcal{D}_{k+1}(X, U)$ if and only if $\Delta(\Delta^k \rho) = 0$, $\mu_\omega^{[k]} \times \mu_\omega^{[k]}$-almost everywhere for P_k-almost $\omega \in \Omega_k$. This condition means that for P_k-almost $\omega \in \Omega_k$, $\Delta^k \rho$ is equal to some constant, $\mu_\omega^{[k]}$-almost everywhere. Thus $\Delta^k \rho$ is an invariant map on $X^{[k]}$. As $(\Delta^k \rho) \circ T^{[k]} = \Delta^k(\rho \circ T)$, this condition is equivalent to $\Delta^k(\rho \circ T - \rho) = 0$. Thus $\rho \circ T - \rho \in \mathcal{D}_k(X, U)$.

The coboundary map $\partial\colon \rho \mapsto \rho \circ T - \rho$ is a continuous group homomorphism from $\mathcal{D}_{k+1}(X, U)$ to $\mathcal{D}_k(X, U)$ and the kernel of this homomorphism is the group U of constant cocycles. There exist only countably many constants in U which are coboundaries of some cocycle on X and thus $\partial(\mathcal{D}_{k+1}(X, U)) \cap U$ is countable. By the induction hypothesis, $\partial(\mathcal{D}_{k+1}(X, U))$ is countable and so the compact group U has countable index in the Polish group $\mathcal{D}_{k+1}(X, U)$ and the result is proven. $\square$

In fact, the proof shows that $\mathcal{D}_k(X, U)$ consists of those cocycles ρ for which the k-iterated coboundary $\partial^k \rho$ is equal to 0.

7.2. **Cocycles of type k and automorphisms.**

Corollary 7.5. *Let (X, μ, T) be an ergodic system, $\rho\colon X \to U$ a cocycle and k an integer.*

1. *If ρ is of type $k \geq 1$, then for any automorphism S of X the cocycle $\rho \circ S - \rho$ is of type $k - 1$.*

(2) *If X is of order $k \geq 2$ and ρ is of type k, then for any vertical rotation $x \mapsto u \cdot x$ of X over Z_{k-1} the cocycle $\rho \circ u - \rho$ is a coboundary.*
(3) *If X is of order $k \geq 1$ and ρ is of type $k+1$, then for any vertical rotation $x \mapsto u \cdot x$ of X over Z_{k-1} the cocycle $\rho \circ u - \rho$ is of type 1.*

For the definition of a vertical rotation, see Appendix C.1.

Proof. (1) Let $F \colon X^{[k]} \to U$ be a map with $F \circ T^{[k]} - F = \Delta^k \rho$. Let α be the first side of V_k. By Lemma 5.5, the measure $\mu^{[k]}$ is invariant under $S_\alpha^{[k]}$. As this transformation commutes with $T^{[k]}$, by the definition of F we have

$$\big(\Delta^{k-1}(\rho \circ S - \rho)\big) \circ \xi_\alpha^{[k]} = (F \circ S_\alpha^{[k]} - F) \circ T^{[k]} - (F \circ S_\alpha^{[k]} - F) \ .$$

By Lemma C.7, $\Delta^{k-1}(\rho \circ S - \rho)$ is a coboundary on $X^{[k-1]}$ and $\rho \circ S - \rho$ is of type $k-1$.

(2) By Proposition 6.3, $X = Z_{k-1} \times W$ for some compact abelian group W. The measure $\mu^{[k]}$ is conditionally independent over $Z_{k-1}^{[k]}$ and thus invariant under the vertical rotation by $w_\epsilon^{[k]}$ for every $\epsilon \in V_k$ and every $w \in W$. The same computation as above shows that $(\rho \circ w - \rho) \circ \xi_\epsilon^{[k]}$ is a coboundary on $X^{[k]}$ and so $\rho \circ w - \rho$ is a coboundary on X.

(3) Let W be as in Part (2). Let $w \in W$. For any $\epsilon \in V_k$, the measure $\mu^{[k]}$ is invariant under $w_\epsilon^{[k]}$. This transformation commutes with $T^{[k]}$ and thus maps the σ-algebra $\mathcal{I}^{[k]}$ to itself. By Lemma 5.5, for any edge α of V_{k+1} the measure $\mu^{[k+1]}$ is invariant under $w_\alpha^{[k+1]}$. We conclude as in Part (2) . □

7.3. **Cocycles of type k and group extensions.** Let Y be an ergodic extension of a system X by a compact abelian group U. Then for every $u \in U$ the associated vertical rotation of Y above X is an automorphism of Y and belongs to $\mathcal{G}(Y)$ by Lemma 5.5. By Lemma 5.2, for every k this transformation induces a measure preserving transformation $p_k u$ of $Z_k(Y)$, which belongs to $\mathcal{G}(Z_k(Y))$ and is actually an automorphism of $Z_k(Y)$. (This follows also from Proposition 4.6.)

Proposition 7.6. *Let (X, μ, T) be an ergodic system, U a compact abelian group, $\rho \colon X \to U$ an ergodic cocycle and $(Y, \nu, S) = (X \times U, \mu \times m_U, T_\rho)$ the extension it defines. (See Appendix C.2 for the definition.) Let $k \geq 1$ be an integer. For $u \in U$, let $p_k u$ be the automorphism of $Z_k(Y)$ defined just above. Let $W = \{u \in U : p_k u = \mathrm{Id}\}$. Then*

(1) *W is a closed subgroup of U.*
(2) *The annihilator $W^\perp$ of W in $\widehat{U}$ is the subgroup $\Gamma = \{\chi \in \widehat{U} : \chi \circ \rho$ is of type $k\}$.*
(3) *The cocycle $\rho \mod W \colon X \to U/W$ is of type k.*
(4) *$Z_k(Y)$ is an extension of $Z_k(X)$ by the compact abelian group U/W, given by a cocycle $\rho' \colon Z_k(X) \to U/W$ of type k. Moreover, the cocycle $\rho' \circ \pi_{X,k}$ is cohomologous to $\rho \mod W \colon X \to U/W$.*

Proof. (1) is obvious. For every $u \in U$, let $\overline{u}$ denote its image in U/W.

We view factors as invariant sub-σ-algebras. Then $\mathcal{X}$ consists in the sets in $\mathcal{Y}$ which are invariant under the vertical rotation associated to any $u \in U$. By Proposition 4.6 we have $\mathcal{Z}_k(X) = \mathcal{Z}_k(Y) \cap \mathcal{X}$. Thus $\mathcal{Z}_k(X)$ consists in those sets in

$\mathcal{Z}_k(Y)$ which are invariant under $p_k u$ for every $u \in U$. Therefore, as an extension of $Z_k(X)$, $Z_k(Y)$ is isomorphic to an extension by the compact abelian group U/W.

We identify $Z_k(Y)$ with $Z_k(X) \times U/W$ and Y with $X \times U$ and study the factor map $\pi_{Y,k}\colon X \times U \to Z_k(X) \times U/W$. By construction, for $(x,u) \in X \times U$, the first coordinate of $\pi_{Y,k}(x,u)$ is equal to $\pi_{X,k}(x)$. Moreover, for every $v \in U$, the transformation $p_k v$ is given by $p_k v(z,\overline{u}) = (z, \overline{v} + \overline{u})$. That is, it is the vertical rotation by $\overline{v}$ of $Z_k(Y)$ over X. Since $\pi_{Y,k} \circ v = p_k v \circ \pi_{Y,k}$, it follows that there exists $\phi\colon X \to U/W$ such that $\pi_{Y,k}(x,u) = \big(\pi_{X,k}(x), \overline{u} + \phi(x)\big)$.

Let $\rho'\colon Z_k(X) \to U/W$ be a cocycle defining the extension $Z_k(X) \times U/W$ of $Z_k(X)$. Since $\pi_{Y,k}\colon X \times U \to Z_k(X) \times U/W$ is a factor map, we get $\rho' \circ \pi_{X,k}(x) = \overline{\rho(x)} + \phi(Tx) - \phi(x)$ and $\rho' \circ \pi_{X,k}$ is cohomologous to $\overline{\rho} = \rho \mod W$.

Let $\chi \in \widehat{U/W} = W^{\perp}$. Here we consider χ as taking values in the circle group $\mathcal{S}^1$. We define a map ψ on $Z_k(Y) = Z_k(X) \times U/W$ by $\psi(x,u) = \chi(u)$ and define a function Ψ on $Z_k(Y)^{[k]} = Z_k(X)^{[k]} \times (U/W)^{[k]}$ by

$$\Psi(\mathbf{x}, \overline{\mathbf{u}}) = \chi\big(\sum_{\epsilon \in V_k} s(\epsilon)\overline{u}_\epsilon\big) \text{ for } \mathbf{x} \in Z_k(X)^{[k]} \text{ and } \overline{\mathbf{u}} \in (U/W)^{[k]}$$

and continue exactly as in the proof of the first part of Proposition 6.4. Then $\chi \circ \rho'$ is of type k.

As this holds for every $\chi \in \widehat{U/W}$, the cocycle ρ' is of type k and Part (4) of the Proposition is proven. Part (3) follows immediately, as does the inclusion $W^{\perp} \subset \Gamma$. We now prove the opposite inclusion.

Let $\chi \in \Gamma$. Then $\chi \circ \rho$ is a cocycle of type k. We consider χ as taking values in $\mathbb{T}$. Let $F\colon X^{[k]} \to \mathbb{T}$ be a map with $F \circ T^{[k]} - F = \Delta^k(\chi \circ \rho)$ $\mu^{[k]}$-almost everywhere. We define a map Φ from $Y^{[k]} = X^{[k]} \times U^{[k]}$ to $\mathbb{T}$ by

$$\Phi(\mathbf{x}, \mathbf{u}) = F(\mathbf{x}) - \sum_{\epsilon \in V_k} s(\epsilon)\chi(u_\epsilon) \text{ for } \mathbf{x} \in X^{[k]} \text{ and } \mathbf{u} \in U^{[k]}\ .$$

The projection of $\nu^{[k]}$ on $X^{[k]}$ is $\mu^{[k]}$ and each of the one-dimensional marginals of $\nu^{[k]}$ is ν. From these remarks and the definition of F we get that $\Phi \circ S^{[k]} = \Phi$ $\nu^{[k]}$-almost everywhere. The map Φ is measurable with respect to $\mathcal{I}(Y)^{[k]}$.

Let $w \in W$ and $\epsilon \in V_k$. The measure $\nu^{[k]}$ is relatively independent with respect to $Z_{k-1}(Y)$ and thus with respect to $Z_k(Y)$. Since the vertical rotation w acts trivially on $Z_k(Y)$, the measure $\nu^{[k]}$ is invariant under $w_\epsilon^{[k]}$. Moreover this transformation acts trivially on $\mathcal{Z}_k^{[k]}(Y)$, thus also on $\mathcal{I}^{[k]}(Y)$, and $\Phi \circ w_\epsilon^{[k]} = \Phi$ $\nu^{[k]}$-almost everywhere. But $\Phi \circ w_\epsilon^{[k]} - \Phi$ is equal to the constant $s(\epsilon)\chi(w)$ and we get that $\chi(w) = 1$. As this holds for every $w \in W$, we have $\chi \in W^{\perp}$ and so $\Gamma \subset W^{\perp}$. Combining the two inclusions, we have the statement of Part (2). □

Corollary 7.7. *Let $k \geq 1$ be an integer, (X, μ, T) a system of order k, U a compact abelian group and $\rho\colon X \to U$ an ergodic cocycle. Then the extension of X associated to ρ is of order k if and only if ρ is of type k.*

Proof. We use the notation of Proposition 7.6. If Y is of order k then $Z_k(Y) = Y$, W is the trivial subgroup of U and ρ is of type k. If ρ is of type k, then $\Gamma = \widehat{U}$, thus W is trivial, and $Z_k(Y) = Y$. □

Corollary 7.8. *Assume that (X, μ, T) and (Y, ν, S) are ergodic systems and that X is of order k for some integer $k \geq 1$. Assume that $\pi \colon X \to Y$ is a factor map and $\rho \colon Y \to U$ is a cocycle. Then ρ is of type k on Y if and only if $\rho \circ \pi$ is if type k on X.*

Proof. If ρ is of type k, it follows immediately from the definition that $\rho \circ \pi$ is of type k.

Assume that $\rho \circ \pi$ is of type k. It suffices to show that $\chi \circ \pi$ is of type k for every $\chi \in \widehat{U}$. Since $\chi \circ (\rho \circ \pi)$ is of type k, without loss of generality we can assume that $U = \mathbb{T}$.

The set $\{c \in \mathbb{T} : c + \rho \text{ is not ergodic }\}$ is either empty or is a coset of the countable subgroup $\{c \in \mathbb{T} : nc \text{ is an eigenvalue for some } n \neq 0\}$. Therefore, there exists $c \in \mathbb{T}$ so that $\rho + c$ is ergodic. Substituting $\rho + c$ for ρ, we can assume that ρ is ergodic.

By Proposition 7.6, the extension of X associated to $\rho \circ \pi$ is of order k because ρ is of type k. Furthermore, the extension of Y associated to ρ is a factor of this and so is of order k as well. Therefore ρ is of type k. □

Corollary 7.9. *Let (X, μ, T) be an ergodic system, U a compact abelian group, and $\rho \colon X \to U$ a cocycle of type k for some integer $k \geq 1$. Then there exists a cocycle $\rho' \colon Z_k(X) \to U$ of type k so that ρ is cohomologous to $\rho' \circ \pi_k$.*

Proof. If ρ is ergodic, the result follows immediately from the preceding Proposition, since by Part (2), the subgroup W is trivial.

Assume that ρ is not ergodic. There exist a closed subgroup V of U and an ergodic cocycle $\sigma \colon X \to V$ so that ρ and σ are cohomologous as U-valued cocycles (see [Zim76]). σ is of type k as a U-valued cocycle, thus also as a V-valued cocycle. There exists a cocycle $\rho' \colon Z_k(X) \to V$ of type k so that σ is cohomologous to $\rho' \circ \pi_k$, as V-valued cocycles on $Z_k(X)$. Thus, as a U-valued cocycle, ρ' is of type k and $\rho' \circ \pi_k$ is cohomologous to ρ. □

Corollary 7.10. *Let $k \geq 2$ be an integer, (X, μ, T) be a system of order k and $\rho \colon X \to U$ a cocycle of type k. Assume that X is an extension of Z_{k-1} by a compact connected abelian group. Then there exists a cocycle $\rho' \colon Z_{k-1} \to U$ of type k so that ρ is cohomologous to $\rho' \circ \pi_{k-1}$.*

Proof. Write $X = Z_{k-1} \times V$ and assume that V is connected. By Corollary 7.5, for every $v \in V$ the cocycle $\rho \circ v - \rho$ is a coboundary. By Lemma C.9, there exists a cocycle ρ' on Z_{k-1} so that $\rho' \circ \pi_{k-1}$ is cohomologous to ρ. By Corollary 7.8, ρ' is of type k. □

8. Initializing the induction: Systems of order 2

In this Section we study the systems of order 2. These systems appeared earlier in the literature (see [CL88], [CL87] and [Ru95]) as 'Conze-Lesigne algebras' and were studied with a different point of view (in [HK01] and [HK02]) under the name of 'quasi-affine systems'. Our purpose here is twofold. In the following sections we establish properties of systems of order k for arbitrary k. As the proofs are a bit intricate, we hope that the proofs in the easier case $k = 2$ aid in understanding the overall plan. Moreover, we prove some technical results which are useful as the starting points of the inductive proofs for higher k.

8.1. **Systems of order 1.** We have shown that for any ergodic system X, $Z_1(X)$ is it Kronecker factor. Thus an ergodic system is of order 1 if and only if it is an ergodic rotation.

Let (Z, t) be an ergodic rotation. For every $s \in Z$, the rotation $z \mapsto sz$ is an automorphism of Z and thus belongs to $\mathcal{G}(Z)$. Conversely, by Corollary 5.9 $\mathcal{G}(Z)$ is abelian. As the rotation $T\colon z \mapsto tz$ lies in $\mathcal{G}(Z)$, every element of $\mathcal{G}(Z)$ is a measure preserving transformation of Z commuting with T and thus is itself a rotation $z \mapsto sz$ for some s. Therefore, the group $\mathcal{G}(Z)$ is equal to Z, acting on itself by translations.

A compact abelian group is a Lie group if and only if its dual group is finitely generated. Thus every compact abelian group is the inverse limit of a sequence of compact abelian Lie groups. Therefore, a system of order 1 is the inverse limit of a sequence of ergodic rotations (Z, t) where each group Z is a compact abelian Lie group.

In the rest of this section, we study the systems of order 2. By Proposition 6.3 and Corollary 7.7, an ergodic system is of order 2 if and only if it is an extension of an ergodic rotation (Z, t) by a compact abelian group U, given by an ergodic cocycle $\sigma\colon Z \to U$ of type 2. By the remark after Definition 7.1, $\sigma\colon Z \to U$ is of type 2 if and only if $\chi \circ \sigma\colon Z \to \mathbb{T}$ is of type 2 for every $\chi \in \widehat{U}$.

8.2. **The Conze-Lesigne Equation and applications.** Throughout this section, (Z, t) denotes an ergodic rotation: Z is a compact abelian group, endowed with the Haar measure $m = m_Z$ and with the ergodic transformation $T\colon z \mapsto tz$, where t is a fixed element of Z.

Lemma 8.1. *Let (Z, t) be an ergodic rotation, U be a torus and $\rho\colon Z \to U$ a cocycle of type 2. For every $s \in Z$, there exist $f\colon Z \to U$ and $c \in U$ so that*

$$\rho(sx) - \rho(x) = f(tx) - f(x) + c\ . \tag{CL}$$

This functional equation was originally introduced by Conze and Lesigne in[CL84], and we call it the *Conze-Lesigne Equation*.

Proof. For every $s \in Z$, the map $z \mapsto sz$ is an automorphism of Z. By Corollary 7.5 the cocycle $z \mapsto \rho(sz) - \rho(z)$ is of type 1. Since U is a torus, the cocycle is a quasi-coboundary by Lemma C.5 and we obtain the functional equation. □

Lemma 8.2. *Let (Z, t) be an ergodic rotation and $\rho\colon Z \to \mathbb{T}$ be a cocycle of type 2 and assume that there exists an integer $n \neq 0$ so that $n\rho$ is a quasi-coboundary. Then ρ is a quasi-coboundary.*

Proof. Let s, f and c be as in Equation (CL). Since $n\rho$ is a quasi-coboundary, the map $z \mapsto n(\rho(sz) - \rho(z))$ is a coboundary. Substituting into Equation (CL), we have that the constant nc is a coboundary, i.e. an eigenvalue of (Z, t). So for all s, f and c satisfying Equation (CL), c belongs to the countable subgroup Γ of $\mathbb{T}$, where

$$\Gamma = \{c \in \mathbb{T} : nc \text{ is an eigenvalue of } (X, \mu, T)\}\ .$$

Define

$$Z_0 = \{s \in Z : \text{the cocycle } x \mapsto \rho(sx) - \rho(x) \text{ is a coboundary}\}\ .$$

Clearly, Z_0 is a Borel subgroup of X. Let (s, f, c) and (s', f', c') satisfy Equation (CL). If $c = c'$, the map $x \mapsto \rho(s'x) - \rho(sx)$ is a coboundary. Thus so is the map $x \mapsto \rho(s's^{-1}x) - \rho(x)$ and $s's^{-1} \in Z_0$. As Γ is countable, Z_0 has countable index in Z. As Z_0 is Borel, Z_0 is an open subgroup of Z. But Z_0 obviously contains t. By density, $Z_0 = Z$ and the cocycle $x \mapsto \rho(sx) - \rho(x)$ is a coboundary for every $s \in Z$.

In other words, the map $(z_0, z_1) \mapsto \rho(z_1) - \rho(z_0)$ is a coboundary of the system $(Z \times Z, m \times m, T \times T)$. By Lemma C.5, ρ is a quasi-coboundary. $\square$

Lemma 8.3. *Let (Z, t) be an ergodic rotation, U a torus and $\rho\colon Z \to U$ a cocycle of type 2. Then there exist a closed subgroup Z_0 of Z so that Z/Z_0 is a compact abelian Lie group and a cocycle $\rho\colon Z/Z_0 \to U$ of type 2 so that ρ is cohomologous to $\rho' \circ \pi$, where $\pi\colon Z \to Z/Z_0$ is the natural projection.*

In this statement, we mean that Z/Z_0 is endowed with the rotation by $\pi(t)$. $(Z/Z_0, \pi(t))$ is an ergodic rotation and π is a factor map.

Proof. By Equation (CL), for every $s \in Z$ the cocycle $z \mapsto \rho(sz) - \rho(z)$ is a quasi-coboundary. Applying Lemma C.10 with the action of Z on itself by translations and Corollary 7.8, we get the result. $\square$

8.3. **Systems of order 2.** .

Corollary 8.4. *For every ergodic system (X, μ, T), $Z_2(X)$ is an extension of $Z_1(X)$ by a compact* connected *abelian group.*

Proof. By Proposition 6.3, Z_2 is an extension of Z_1 by a compact abelian group U given by an ergodic cocycle $\sigma\colon Z_1 \to U$ of type 2.

Assume that U is not connected. Then it admits an open subgroup U_0 so that U/U_0 is isomorphic to $\mathbb{Z}/n\mathbb{Z}$ for some integer $n > 1$. Write $\overline{\sigma}\colon Z_1 \to U/U_0$ for the reduction of σ modulo U_0, meaning the composition of σ with the quotient map $U \mapsto U/U_0$. It is an ergodic cocycle of type 2. Using the isomorphism from U/U_0 to $\mathbb{Z}/n\mathbb{Z}$ and an embedding of $\mathbb{Z}/n\mathbb{Z}$ as a finite closed subgroup of $\mathbb{T}$, we get a (non-ergodic) cocycle $\rho\colon Z_1 \to \mathbb{T}$ of type 2 with $n\rho = 0$. By Lemma 8.2, ρ is a quasi-coboundary and thus of type 1. Viewed as a cocycle with values in $\mathbb{Z}/n\mathbb{Z}$, ρ is also of type 1 (even if it is not a quasi-coboundary) and $\overline{\sigma}$ is of type 1.

By Corollary 7.7 the extension $T_{\overline{\sigma}}$ associated to $\overline{\sigma}$ is system of order 1, meaning it is an ergodic rotation. But this extension is obviously a factor of Z_2, which is the extension of Z_1 associated to σ and thus also a factor of X. The maximal property (Proposition 4.11) of Z_1 provides a contradiction. $\square$

Definition 8.5. A system X of order 2 is *toral* if its Kronecker factor Z_1 is a compact abelian Lie group and X is an extension of Z_1 by a torus.

Proposition 8.6. *Every system of order 2 is the inverse limit of a sequence of toral systems of order 2.*

Proof. Let X be a system of order 2. By Corollary 8.4, X is an extension of its Kronecker factor Z_1 by a compact connected abelian group U, given by a cocycle $\rho\colon Z_1 \to U$. Therefore, U is an inverse limit of a sequence of tori. This means that there exists a decreasing sequence $\{V_n\}$ of closed subgroups of U, with $\bigcap_n V_n = \{0\}$ so that $U_n = U/V_n$ is a torus for each n. For each n, let $\rho_n\colon Z_1 \to U_n$ be the reduction of ρ modulo V_n and let X_n be the extension of Z_1 by U_n, associated to the cocycle ρ_n. Then X is clearly the inverse limit of the sequence $\{X_n\}$.

By Lemma 8.3, for each n there exists a subgroup K_n of Z_1 such that Z_1/K_n is a compact abelian Lie group, and a cocycle $\rho'_n\colon Z_1/K_n \to U_n$ so that ρ_n is cohomologous to $\rho'_n \circ \pi_n$, where $\pi_n = Z_1 \to Z_1/K_n$ is the natural projection. We can clearly modify the groups K_n, by induction, so that these properties remain valid and so that the sequence $\{K_n\}_n$ of subgroups is decreasing and has trivial intersection. For each n, let Y_n be the extension of Z_1/K_n by U_n associated to the cocycle ρ'_n. Each of these systems is a factor of X and is toral. This sequence of factors of X is increasing and its inverse limit is clearly X. □

8.4. **The group of a system of order 2.** In this section, we study the group $\mathcal{G} = \mathcal{G}(X)$ associated to a system (X, μ, T) of order 2. We restrict to the case that X is an extension of its Kronecker factor (Z_1, t) by a torus U and write $\rho\colon Z_1 \to U$ for the cocycle defining this extension. As usual, we identify X with $Z_1 \times U$.

We use the notation of Appendices A and C. $\mathcal{C}(Z_1, U)$ denotes the group of measurable maps from Z_1 to U, endowed with the topology of convergence in probability. A map $f\colon Z_1 \to U$ is said to be *affine* if it is the sum of a constant and a continuous group homomorphism from Z_1 to U and we write $\mathcal{A}(Z_1, U)$ for the group of affine maps. It is a closed group of $\mathcal{C}(Z_1, U)$ and is the direct sum of the compact group U of constants and the discrete group of continuous group homomorphisms from Z_1 to U.

As in Section A.1, for each $s \in Z_1$ and $f \in \mathcal{C}(Z_1, U)$, let $S_{s,f}$ denote the measure preserving transformation of $Z_1 \times U$ given by

$$S_{s,f}(z, u) = (sz, u + f(z)) \ . \tag{24}$$

These transformations form the skew product of Z_1 and $\mathcal{C}(Z_1, U)$. Endowed with the topology of convergence in probability, it is a Polish group.

Lemma 8.7. *The group $\mathcal{G}$ consists in the transformations of X of the type given by Equation* (24)*, for $s \in Z_1$ and $f\colon Z_1 \to U$ satisfying Equation* (CL) *for some constant c.*

Proof. Let $g \in \mathcal{G}$. By Lemma 5.2, g induces a measure preserving transformation of Z_1 belonging to $\mathcal{G}(Z_1)$ and thus of the form $z \mapsto sz$ for some $s \in Z_1$. Moreover, by Corollary 5.10, the transformation g commutes with all vertical rotations of X over Z_1 and thus is of the form given by Equation (24) for some map $f\colon Z_1 \to U$. We notice that the the commutator $[g; T]$ induces the trivial transformation of Z_1. As $\mathcal{G}$ is 2-step nilpotent, $[g, T]$ belongs to the center of $\mathcal{G}$ and thus commutes with T. It follows that $[g, T]$ is a vertical rotation of X over Z_1, given by some $c \in U$ (see the definition of a vertical rotation in Subsection C.1). By definition of the commutator, s, f and c satisfy Equation (CL).

Conversely, let $s \in Z_1$ and $f\colon Z_1 \to U$ be such that Equation (CL) is satisfied for some $c \in U$. We show that the transformation $g = S_{s,f}$ belongs to $\mathcal{G}$. Let α be an edge of V_2. The transformation $s\colon z \mapsto sz$ of Z_1 induced on Z_1 by g belongs to $\mathcal{G}(Z_1)$ and thus the transformation $s_\alpha^{[2]}$ leaves the measure $\mu_1^{[2]}$ invariant and maps the σ-algebra $\mathcal{I}(Z_1)^{[2]}$ to itself. We define a map $F\colon Z_1^{[2]} \to U$ and a map $\Phi\colon X^{[2]} \to U$ as in Proposition 6.4. An immediate computation shows that the map $\Phi \circ g_\alpha^{[2]} - \Phi$ is invariant under $T^{[2]}$ and so $\Phi \circ g_\alpha^{[2]}$ is also invariant under this transformation. By Proposition 6.4, the transformation $g_\alpha^{[2]}$ maps the σ-algebra $\mathcal{I}(X)^{[2]}$ to itself. By Lemma 5.3 and Corollary 6.6, $g \in \mathcal{G}$. □

We recall that $\mathcal{G}$ is endowed with the topology of convergence in probability. The map $p\colon S_{s,f} \mapsto s$ is a continuous group homomorphism from $\mathcal{G}$ to Z_1 and is onto by Lemma 8.1. The kernel of this homomorphism is the group of transformations of the kind $S_{1,f}$, where $f(tz) - f(z)$ is constant. By ergodicity of the rotation (Z_1, t), a map $f \in \mathcal{C}(Z_1, U)$ satisfies this condition if and only if it is affine. The map $f \mapsto S_{1,f}$ is then an algebraic and topological embedding of $\mathcal{A}(Z_1, U)$ in $\mathcal{G}$ with range $\ker(p)$. In the sequel we identify $\mathcal{A}(Z_1, U)$ with $\ker(p)$. This identification generalizes the preceding identification of U with the group of vertical rotations. $\mathcal{G}$ is a group of the type which is studied in Appendix A. By Corollary A.2, $\mathcal{G}$ is locally compact.

Lemma 8.8. *Every toral system of order* 2 *is isomorphic to a nilsystem.*

(See Section B for the meaning of a nilsystem.)

Proof. We keep the same notation as above and assume furthermore that Z_1 is a compact abelian Lie group. The kernel $\mathcal{A}(Z_1, U)$ of p is the direct sum of the torus U and a discrete group and thus it is a Lie group also. By Lemma A.3, $\mathcal{G}$ is a Lie group. We recall that $\mathcal{G}$ is 2-step nilpotent.

Let Γ be the stabilizer of $(1, 0) \in X$ for the action of $\mathcal{G}$ on this space. Then Γ consists in the transformations associated to $(1, f)$, where f is a continuous group homomorphism from Z_1 to U. Thus Γ is discrete. The map $g \mapsto g \cdot (1, 0)$ induces a bijection j from the nilmanifold $\mathcal{G}/\Gamma$ onto X. For any $g \in \mathcal{G}$, the transformation $j^{-1} \circ g \circ j$ of $\mathcal{G}/\Gamma$ is the (left) translation by g on the nilmanifold $\mathcal{G}/\Gamma$. In particular, $j^{-1} \circ T \circ j$ is the (left) translation $x \mapsto T \cdot x$ by $T \in \mathcal{G}$. Moreover, since every $g \in G$ is a measure preserving transformation of X, the image of μ under j^{-1} is invariant under the (left) action of $\mathcal{G}$ on $\mathcal{G}/\Gamma$ and thus is the Haar measure on this space. The map j is the announced isomorphism. $\square$

8.5. **Countable number of cocycles.** We show that the number of $\mathbb{T}$-valued cocycles of type 2 on an ergodic rotation Z, up to quasi-boundary, is countable.

Proposition 8.9. *Let (Z, t) be an ergodic rotation. Up to the addition of a quasi-coboundary, there are only countably many $\mathbb{T}$-valued cocycles of type* 2 *on Z.*

Proof. We make use of explicit distances on some groups of functions. For $u \in \mathbb{T}$, write

$$\|u\| = |\exp(2\pi i u) - 1| \ .$$

For $f \in \mathcal{C}(Z) = \mathcal{C}(Z, \mathbb{T})$, write

$$\|f\| = \Bigl(\int \|f(z)\|^2 \, dm(z)\Bigr)^{1/2} \ .$$

The distance between two cocycles $f, g \in \mathcal{C}(Z)$ is defined to be $\|f - g\|$. As above, $\mathcal{A}(Z) = \mathcal{A}(Z, \mathbb{T})$ denotes the closed group of affine cocycles. For $c, c' \in \mathbb{T}$ and $\gamma, \gamma' \in \widehat{Z}$, we have $\|(c + \gamma) - (c' + \gamma')\| \geq \sqrt{2}$ whenever $\gamma \neq \gamma'$.

Let $\mathcal{Q}(Z)$ denote the quotient group $\mathcal{Q}(Z) = \mathcal{C}(Z)/\mathcal{A}(Z)$ and write $q\colon \mathcal{C}(Z) \to \mathcal{Q}(Z)$ for the quotient map. The quotient distance between $\Phi \in \mathcal{Q}$ and $0 \in \mathcal{Q}$ is written $\|\Phi\|_{\mathcal{Q}}$ and the quotient distance between two elements Φ, Ψ of this group is $\|\Phi - \Psi\|_{\mathcal{Q}}$. Endowed with this distance, $\mathcal{Q}(Z)$ is a Polish group.

We also use the group $\mathcal{F}$ of continuous maps from Z to $\mathcal{Q}$, endowed with the distance of uniform convergence: If $s \mapsto \Phi(s)$ is an element of $\mathcal{F}$, write

$$|\!|\!|\Phi|\!|\!|_\infty = \sup_{s\in Z} |\!|\!|\Phi(s)|\!|\!|_{\mathcal{Q}} \ .$$

The distance between two elements Φ and $\Psi \in \mathcal{F}$ is $|\!|\!|\Phi - \Psi|\!|\!|_\infty$. As Z is compact and $\mathcal{Q}$ is a Polish group, $\mathcal{F}$ is also a Polish group.

First Step. Let $\rho \in \mathcal{C}(Z)$ be a weakly mixing cocycle of type 2. Let X be the extension of Z associated to this cocycle. X is of order 2 and $Z_1(X) = Z$. We use the notation of Section 8.4.

Let $s \mapsto S_{s,f_s}$ be an arbitrary cross section of the map $p\colon \mathcal{G} \to Z$. For every $s \in Z$, f_s belongs to $\mathcal{C}(Z)$ and satisfies Equation (CL) for some $c \in \mathbb{T}$. Define $\Phi_\rho(s) \in \mathcal{Q}(Z)$ to be the image of f_s under q. Since the kernel of $p\colon \mathcal{G} \to Z$ is $\mathcal{A}(Z)$, $\Phi_\rho(s)$ does not depend on the choice of f_s. In fact, the map $s \mapsto \Phi_\rho(s)$ from Z to $\mathcal{Q}(Z)$ is the reciprocal of the isomorphism $\mathcal{G}/\ker(p) \to Z$ and thus it is continuous. In other words, this map is an element of $\mathcal{F}$.

Second Step. We continue assuming that ρ is a weakly mixing cocycle of type 2. Φ_ρ is defined as above.

Lemma 8.10. *If* $|\!|\!|\Phi_\rho|\!|\!|_\infty < 1/20$*, then* ρ *is cohomologous to an affine map.*

Proof of lemma 8.10. Define a subset $\mathcal{K}$ of $\mathcal{G}$ by

$$\mathcal{K} = \{S_{s,f} \in \mathcal{G} : \text{ There exists } c \in \mathbb{T} \text{ with } \|c + f\| \le 1/10\} \ .$$

Let $s \in Z$. By hypothesis $|\!|\!|\Phi_\rho(s)|\!|\!|_{\mathcal{Q}} < 1/20$ and there exists $f \in \mathcal{C}(Z)$ with $S_{s,f} \in \mathcal{G}$ and $\|f\| < 1/20$, thus $S_{s,f} \in \mathcal{K}$. The restriction $p|_{\mathcal{K}}$ of $p\colon \mathcal{G} \to Z$ to $\mathcal{K}$ is therefore onto.

Claim: $\mathcal{K}$ is a subgroup of $\mathcal{G}$.

Let $S_{s,f}$ and $S_{s',f'} \in \mathcal{K}$. We have $S_{s',f'} \circ S_{s,f} = S_{s's,f''}$ where $f''(z) = f(s'z) + f'(z)$. Choose $c, c' \in \mathbb{T}$ with $\|c + f\| \le 1/10$ and $\|c' + f'\| \le 1/10$. Then $\|f'' + c + c'\| \le \|f + c\| + \|f' + c'\| \le 1/5$. On the other hand, there exists an element of $\mathcal{K}$ with projection on Z equal to ss'. This means that there exists $g \in \mathcal{C}(Z)$ with $\|g\| < 1/20$ and $S_{s's,g} \in \mathcal{G}$. We get that $S_{1,f''-g} \in \mathcal{G}$ and thus $f'' - g \in \mathcal{A}(Z)$ and $f'' - g + c + c' \in \mathcal{A}(Z)$. But $\|f'' - g + c + c'\| \le \|f'' + c + c'\| + \|g\| \le 1/4$ and so $f'' - g + c + c'$ is equal to a constant $d \in \mathbb{T}$. Finally, $\|f'' + c + c' - d\| = \|g\| < 1/20$ and $S_{s's,f''} \in \mathcal{K}$. Clearly, the identity transformation $S_{1,0}$ belongs to $\mathcal{K}$ and the inverse of an element of $\mathcal{K}$ belongs to $\mathcal{K}$. The claim is proven.

$\mathcal{K}$ clearly contains the group $\mathbb{T}$ of vertical rotations. If f is an affine map and $\|c + f\| \le 1/10$ for some constant c, then f is constant. It follows that the kernel of the group homomorphism $p|_{\mathcal{K}}\colon \mathcal{K} \to Z$ is the group $\mathbb{T}$ of vertical rotations. Moreover, $\mathcal{K}$ is clearly closed in $\mathcal{G}$ and is locally compact. Since the kernel $\mathbb{T}$ and the range Z of $p|_{\mathcal{K}}$ are compact, $\mathcal{K}$ is a compact group.

Claim: $\mathcal{K}$ is abelian.

We consider the commutator map $(g, h) \mapsto [g; h]$. It is continuous and bilinear because $\mathcal{K}$ is 2-step nilpotent. But the commutator group $\mathcal{K}'$ is included in $\mathbb{T}$ because $\mathcal{K}'$ is the kernel of the group homomorphism $p_{\mathcal{K}}$ ranging in the abelian group Z. Thus the commutator map has range in $\mathbb{T}$. Moreover, $\mathbb{T}$ is included in the center of $\mathcal{K}$. (This can be seen either by applying Proposition 6.3 or by checking directly.) Thus the commutator map is trivial on $\mathbb{T} \times \mathcal{K}$ and $\mathcal{K} \times \mathbb{T}$. Therefore, it induces a continuous bilinear map from $\mathcal{K}/\mathbb{T} \times \mathcal{K}/\mathbb{T} \to \mathbb{T}$ and finally a continuous bilinear map $b\colon Z \times Z \to \mathbb{T}$. Choose $f \in \mathcal{C}(Z)$ with $S_{t,f} \in \mathcal{K}$. For all integers m, n the transformations $S_{t,f}^m$ and $S_{t,f}^n$ commute and by definition of b, $b(t^m, t^n) = 0$. Since (Z, t) is an ergodic rotation, $\{t^n : n \in \mathbb{Z}\}$ is dense in Z and so the bilinear map b is trivial. Returning to the definition, the commutator map $\mathcal{K} \times \mathcal{K} \to \mathcal{K}'$ is trivial and the second claim is proven.

The compact abelian group $\mathcal{K}$ admits $\mathbb{T}$ as a closed subgroup, with quotient Z. Thus it is isomorphic to $\mathbb{T} \oplus Z$. This means that the group homomorphism $p|_{\mathcal{K}}\colon \mathcal{K} \to Z$ admits a cross section $Z \to \mathcal{K}$, which is a group homomorphism and is continuous. This cross section has the form $s \mapsto S_{s,f_s}$ and the map $s \mapsto f_s$ is continuous from Z to $\mathcal{C}(Z)$ satisfies for all $s, s' \in Z$

$$f_{ss'}(z) = f_{s'}(sz) + f_s(z) \text{ for almost every } z \in Z.$$

By Lemma C.8, there exists $f \in \mathcal{C}(Z)$ so that $f_s(z) = f(sz) - f(z)$ for every $s \in Z$.

Define $\rho'(z) = \rho(z) - f(tz) + f(z)$. The cocycle ρ' is cohomologous to ρ. Moreover, for every s we have $S_{s,f_s} \in \mathcal{K} \subset \mathcal{G}$ and this means that s and f_s satisfy Equation (CL) for some constant c. Substituting in the definition of ρ' we have $\rho'(sz) - \rho(z) = c$. As this holds for every $s \in Z$, ρ' is an affine cocycle. This completes the proof of Lemma 8.10. □

End of the proof of Proposition 8.9.

Let $\mathcal{W}$ be the family of weakly mixing cocycles of type 2 on Z. To every cocycle $\rho \in \mathcal{W}$, we have associated an element Φ_ρ of $\mathcal{F}$. Since $\mathcal{F}$ is separable, there exists a countable family $\{\rho_i : i \in I\}$ in $\mathcal{W}$ so that for every $\rho \in \mathcal{W}$, there exists $i \in I$ with $|\!|\!|\Phi_\rho - \Phi_{\rho_i}|\!|\!| < 1/20$.

Let $\rho\colon Z \to \mathbb{T}$ be a cocycle of type 2.

Assume first that ρ is not weakly mixing. There exists an integer $n \neq 0$ so that $n\rho$ is a quasi-coboundary and by Lemma 8.2 ρ itself is a quasi-coboundary.

Assume now that ρ is weakly mixing. Choose $i \in I$ so that $|\!|\!|\Phi_\rho - \Phi_{\rho_i}|\!|\!| < 1/20$. If $\rho - \rho_i$ is not weakly mixing, by the same argument as above this cocycle is a quasi-coboundary and ρ is the sum of ρ_i and a quasi-coboundary. If $\rho - \rho_i$ is weakly mixing, then $\Phi_{\rho-\rho_i} = \Phi_\rho - \Phi_{\rho_i}$. Thus $|\!|\!|\Phi_{\rho-\rho_i}|\!|\!| < 1/20$ and by Lemma 8.10 the cocycle $\rho - \rho_i$ is cohomologous to some affine map. In this case, ρ is the sum of ρ_i, a character $\gamma \in \widehat{Z}$ and a quasi-coboundary.

The proof of Proposition 8.9 is complete. □

9. The main induction

We now generalize the results for systems of order 2 of Section 8 to higher orders. We start with a more detailed study of the ergodic decomposition of $\mu \times \mu$.

9.1. **The systems X_s.**

In this section, we use the following notation. Let (X, μ, T) be an ergodic system. For every integer $k \geq 2$, $Z_k = Z_k(X)$ is an extension of Z_{k-1} by a compact abelian group U_k, given by a cocycle $\rho_k\colon Z_{k-1} \to U_k$ of type k.

We recall the ergodic decomposition of formula (7)

$$\mu \times \mu = \int_{Z_1} \mu_s \, d\mu_1(s)$$

of $\mu \times \mu$ for $T \times T$.

Notation. For every $s \in Z_1$, let X_s denote the system $(X \times X, \mu_s, T \times T)$.

We recall that X_s is ergodic for μ_1-almost every $s \in Z_1$ (see Subsection 3.2).

Lemma 9.1. *Let (X, μ, T) be an ergodic system, U a compact abelian group, $\rho\colon X \to U$ a cocycle and $k \geq 0$ an integer. Then the subset*

$$A = \{s \in Z_1 : \Delta\rho \text{ is a cocycle of type } k \text{ of } X_s\}$$

of Z_1 is measurable and $\mu_1(A) = 0$ or 1. Furthermore, the cocycle ρ is of type $k+1$ if and only if $\mu_1(A) = 1$.

Proof. We recall that $\Delta\rho$ is defined on $X \times X$ by $\Delta\rho(x', x'') = \rho(x') - \rho(x'')$. Under the identification of $X^{[k+1]}$ with $(X \times X)^{[k]}$, we can write $\Delta^{k+1}\rho = \Delta^k(\Delta\rho)$. Moreover, by Equation (8) we have $\int (\mu_s)^{[k]}\, d\mu_1(s) = \mu^{[k+1]}$. By using the definition of a cocycle of type $k+1$ on X, the definition of a cocycle of type k on X_s and Corollary C.4, we get immediately that A is a measurable subset of Z_1 and that ρ is of type $k+1$ if and only if $\mu_1(A) = 1$. It only remains to show that $\mu_1(A) = 0$ or 1.

Let $s \in Z_1$ with $Ts \in A$. The map $\mathrm{Id} \times T$ is an isomorphism of X_s onto X_{Ts}. Thus $\Delta\rho \circ (\mathrm{Id} \times T)$ is a cocycle of type k on X_s. But $\Delta\rho \circ (\mathrm{Id} \times T) - \Delta\rho$ is the coboundary of the map $(x', x'') \mapsto -\rho(x'')$. Thus, $\Delta\rho$ is of type k on X_s and $s \in A$. Therefore, the subset A of Z_1 is measurable and invariant under T and so has measure 0 or measure 1. □

Before stating the next property we need some notation. Let $p\colon (X, \mu, T) \to (Y, \nu, S)$ be a factor map. p induces a factor map p_1 from the Kronecker factor $Z_1(X)$ of X to the Kronecker factor $Z_1(Y)$ of Y. By an abuse of notation, for $s \in Z_1(X)$ we often write ν_s instead of $\nu_{p_1(s)}$ and Y_s instead of $Y_{p_1(s)}$. By the ergodic decomposition, for μ_1-almost every $s \in Z_1$ the measure ν_s is the image of μ_s under $p \times p$. In other words, $p \times p$ is a factor map from X_s to Y_s.

Lemma 9.2. *Let (X, μ, T) be an inverse limit of a sequence $\{X_n\}_n$ of ergodic systems. Then for μ_1-almost every $s \in Z_1$, $X_s = \varprojlim X_{n,s}$, where $X_{n,s}$ is the system associated to X_n in the same way that X_s is associated to X.*

Proof. There exists a countable family $\{f_i : i \in I\}$ of bounded functions defined everywhere on X, dense in $L^2(\mu)$ and so that the linear span of the family $\{f_i \otimes f_j : i, j \in I\}$ is dense in $L^2(\nu)$ for every probability measure ν on $X \times X$. For every i and every n, we consider $\mathbb{E}(f_i \mid \mathcal{X}_n)$ as a function defined everywhere on X.

For every $i \in I$, $\mathbb{E}(f_i \mid \mathcal{X}_n)$ converges to f_i μ-almost everywhere. There exists a subset X_0 of X, with $\mu(X_0) = 1$, so that $\mathbb{E}(f_i \mid \mathcal{X}_n)(x) \to f_i(x)$ for all $i \in I$ and all $x \in X_0$. For μ_1-almost every $s \in Z_1$, we have $\mu_s(X_0 \times X_0) = 1$.

Fix such an s, and consider $X \times X$ as endowed with μ_s. For every $i, j \in I$, $\mathbb{E}(f_i \mid \mathcal{X}_n) \otimes \mathbb{E}(f_j \mid \mathcal{X}_n)$ converges to $f_i \otimes f_j$ on $X_0 \times X_0$, thus μ_s-almost everywhere. For every n, $\mathbb{E}(f_i \mid \mathcal{X}_n) \otimes \mathbb{E}(f_j \mid \mathcal{X}_n)$ is measurable with respect to $\mathcal{X}_n \otimes \mathcal{X}_n$ and it follows that $f_i \otimes f_j$ is measurable with respect to the inverse limit $\varprojlim X_{n,s}$ of the factors $X_{n,s}$ of X_s. By density, every function in $L^2(\mu_s)$ is measurable with respect to $\varprojlim X_{n,s}$. □

9.2. **The factors $Z_k(X_s)$.** We compute the factors $Z_k(X_s)$ of X_s.

As above, for every integer $k \geq 2$, Z_k is an extension of Z_{k-1} by a compact abelian group U_k, given by a cocycle $\rho_k\colon Z_{k-1} \to U_k$ of type k. We recall that for every integer k, the system Z_k has the same Kronecker factor Z_1 as X.

For every k and μ_1-almost every $s \in Z_1$, we associate to the system (Z_k, μ_k, T) a measure $\mu_{k,s}$ on $Z_k \times Z_k$ in the same way that μ_s is associated to (X, μ, T). Let $Z_{k,s}$ denote the system $(Z_k \times Z_k, \mu_{k,s}, T \times T)$.

The measure μ_s is a relatively independent joining of μ over the joining $\mu_{1,s}$ of μ_1. Thus, for every k, $\mu_{k,s}$ is a relatively independent joining of μ_k over $\mu_{1,s}$ and thus over the joining $\mu_{k-1,s}$ of μ_{k-1}. Therefore, the system $(Z_{k,s}, \mu_{k,s}, T \times T)$ is an extension of $(Z_{k-1,s}, \mu_{k-1,s}, T \times T)$ by the group $U_k \times U_k$, given by the cocycle $\rho_k \times \rho_k \colon (x', x'') \mapsto (\rho_k(x'), \rho_k(x''))$.

Lemma 9.3. *Let $k \geq 1$ be an integer. Then:*

(1) *For μ_1-almost every $s \in Z_1$, $\rho_k \times \rho_k$ is a cocycle of type k on $Z_{k-1,s}$.*
(2) *For μ_1-almost every $s \in Z_1$, $Z_{k,s}$ is a system of order k. In particular, if X is of order k then X_s is of order k for μ_1-almost every $s \in Z_1$.*

Proof. (1) We identify $Z_{k-1}^{[k]} \times Z_{k-1}^{[k]}$ with $(Z_{k-1}^2)^{[k]}$ and with $Z_{k-1}^{[k+1]}$. We recall that there exists $F_k \colon Z_{k-1}^{[k]} \to U_k$ with $\Delta^k \rho_k = F_k \circ T^{[k]} - F_k$, $\mu_{k-1}^{[k]}$-almost everywhere. Define $G\colon Z_{k-1}^{[k]} \times Z_{k-1}^{[k]} \to U_k \times U_k$ by $G(\mathbf{x}', \mathbf{x}'') = (F_k(\mathbf{x}'), F_k(\mathbf{x}''))$. As each of the two projections of $\mu_{k-1}^{[k+1]}$ on $Z_{k-1}^{[k]}$ is equal to $\mu_{k-1}^{[k]}$, we get that the equality $\Delta^k(\rho_k \times \rho_k) = G \circ T^{[k+1]} - G$ holds $\mu_{k-1}^{[k+1]}$-almost everywhere. As $\mu_{k-1}^{[k+1]} = \int_{Z_1} (\mu_{k-1,s})^{[k]}\, d\mu_1(s)$, for μ_1-almost every s, the same relation holds $(\mu_{k-1,s})^{[k]}$-almost everywhere and $\rho_k \times \rho_k$ is a cocycle of type k of $Z_{k-1,s}$.
(2) This follows by induction on k, using Proposition 7.7 at each step. □

Proposition 9.4. *For every integer $k \geq 1$ and μ_1-almost every $s \in Z_1$, $Z_k(X_s)$ is a factor of $Z_{k+1,s}$; it is an extension of $Z_{k,s}$ by U_{k+1}, given by the cocycle $\Delta\rho_{k+1} \colon (x', x'') \mapsto \rho_{k+1}(x') - \rho_{k+1}(x'')$, when viewed as a cocycle on $Z_k(X_s)$. Furthermore, $Z_{k+1,s}$ is an extension of $Z_k(X_s)$ by U_{k+1}, given by the cocycle $(x', x'') \mapsto \rho_{k+1}(x'')$.*

Proof. By Proposition 4.7, the invariant σ-algebra $\mathcal{I}^{[k+1]}(X)$ of the system $(X^{[k+1]}$, $\mu^{[k+1]}$, $T^{[k+1]})$ is measurable with respect to $\mathcal{Z}_{k+1}^{[k+1]}$. As $\mu^{[k+1]} = \int \mu_s^{[k]}\, d\mu_1(s)$, by classical arguments for μ_1-almost every $s \in Z_1$, the invariant σ-algebra of $X_s^{[k]} = (X^{[k+1]}, \mu_s^{[k]}, T^{[k+1]})$ is measurable with respect to the same σ-algebra, that is, with respect to $(\mathcal{Z}_{k+1} \times \mathcal{Z}_{k+1})^{[k]}$. By the minimality property of the factor $Z_k(X_s)$ (Proposition 4.7 again), the σ-algebra $\mathcal{Z}_k(X_s)$ is measurable with respect to $\mathcal{Z}_{k+1} \times \mathcal{Z}_{k+1}$. In other words, $Z_k(X_s)$ is a factor of $Z_{k+1,s}$.

Let $\chi', \chi'' \in \widehat{U_{k+1}}$ and consider here these characters as taking values in $\mathbb{T}$. Write $\chi = (\chi', \chi'') \in \widehat{U_{k+1}} \times \widehat{U_{k+1}}$, which we identify with the dual group of $U_{k+1} \times U_{k+1}$. Let $\sigma \colon Z_k \times Z_k \to U_{k+1}$ be the map given by

$$\sigma(x', x'') = \chi'(\rho_{k+1}(x')) + \chi''(\rho_{k+1}(x''))\ .$$

Define

$$A = \bigl\{ s \in Z_1 : \sigma \text{ is a cocycle of type } k \text{ of } Z_{k,s} \bigr\}\ .$$

By the same method as in the proof of Lemma 9.1, we get that A is invariant under T and $\mu_1(A) = 0$ or 1.

Let us assume that $\mu_1(A) = 1$. For μ_1-almost every $s \in Z_1$, $\Delta^k \sigma$ is a coboundary of the system $(Z_k^{[k+1]}, \mu_{k,s}^{[k+1]}, T^{[k+1]})$. Thus $\Delta^k \sigma$ is a coboundary of the system

$(Z_k^{[k+1]}, \mu_k^{[k+1]}, T^{[k+1]})$ and there exists a map $F\colon Z_k^{[k+1]} \to U_{k+1}$, with

$$F(T^{[k+1]}\mathbf{x}) - F(\mathbf{x}) = \sum_{\epsilon \in V_{k+1}} s(\epsilon)\chi_\epsilon(\rho_{k+1}(x_\epsilon))$$

where

$$\chi_\epsilon = \begin{cases} \chi' & \text{if } \epsilon_1 = 0 \\ -\chi'' & \text{if } \epsilon_1 = 1. \end{cases}$$

The function Φ, defined on $Z_{k+1}^{[k+1]} = Z_k^{[k]} \times U_{k+1}^{[k+1]}$ by

$$\Phi(\mathbf{x}, \mathbf{u}) = F(\mathbf{x}) - \sum_{\epsilon \in V_{k+1}} s(\epsilon)\chi_\epsilon(u_\epsilon)\ ,$$

is invariant under $T^{[k+1]}$. By Proposition 6.3, it is invariant under $u_\alpha^{[k+1]}$ for every edge $\alpha = (\epsilon, \eta)$ of V_{k+1} and every $u \in U_{k+1}$. This means that $s(\epsilon)\chi_\epsilon(u) + s(\eta)\chi_\eta(u) = 1$ and thus $\chi_\epsilon(u) = \chi_\eta(u)$. As this holds for every $u \in U_{k+1}$, $\chi_\epsilon = \chi_\eta$. This holds for every edge α and so $\chi'' = -\chi'$.

Summarizing, if $\chi'' \neq -\chi'$, then $\mu_1(A) \neq 1$ and so $\mu_1(A) = 0$. Then for μ_1-almost every s, the cocycle σ of $Z_{k,s}$ is not of type k. If $\chi'' = -\chi'$, then $\sigma = \chi' \circ \Delta\rho_{k+1}$, which is a cocycle of type k on $Z_{k,s}$ for μ_1-almost every $s \in Z_1$ by Lemma 9.1.

We recall that $Z_{k+1,s}$ is the extension of $Z_{k,s}$ associated to the cocycle $\rho_{k+1} \times \rho_{k+1}$ with values in $U_{k+1} \times U_{k+1}$ and apply Proposition 7.6. The annihilator of the group W appearing in this proposition is $\{(\chi', -\chi') : \chi \in \widehat{U_{k+1}}\}$. Thus $W = \{(u', u') : u' \in U_{k+1}\}$. The map $(u', u'') \mapsto (u' - u'', u'')$ is an isomorphism of $U_{k+1} \times U_{k+1}$ on itself. It maps W to $\{0\} \times U_{k+1}$ and we can identify $(U_{k+1} \times U_{k+1})/W$ with U_{k+1}. Under this identification, the cocycle $\rho_{k+1} \times \rho_{k+1} \mod W$ is simply $\Delta\rho_{k+1}$. We get that $Z_k(X_s)$ is the extension of $Z_{k,s}$ associated to the cocycle $\Delta\rho_{k+1}$. Using the identification of the subgroup W with U_{k+1} explained above, we have the last statement of the Proposition. □

9.3. **Connectivity.** We generalize the connectivity result established for systems of order 2 in Section 8 to higher orders. We show that for an ergodic system (X, μ, T) and integer $k \geq 1$, $Z_{k+1}(X)$ is an extension of $Z_k(X)$ by a *connected* compact abelian group. In fact, we prove simultaneously two results by induction:

Theorem 9.5. *Let $k \geq 1$ be an integer.*

(1) *Let (X, μ, T) be a system of order k, $\rho\colon X \to \mathbb{T}$ a cocycle of type $k+1$ and $n \neq 0$ an integer. If $n\rho$ is of type k, then ρ itself is of type k.*
(2) *For every ergodic system (X, μ, T), $Z_{k+1}(X)$ is an extension of $Z_k(X)$ by a compact connected abelian group.*

Proof. For $k = 1$ these results have been proven in Section 8 (Lemma 8.2 and Corollary 8.4).

Let $k > 1$ and assume that the two properties hold for $k - 1$. Let X, ρ and n be as in the first statement of the Theorem.

X is an extension of $Z_{k-1} = Z_{k-1}(X)$ by a compact abelian group U, which is connected by the inductive hypothesis. As usual, for $u \in U$ we also use u to denote the corresponding vertical rotation of X over Z_{k-1}.

Since $n\rho$ is of type k, by Corollary 7.10 there exists a cocycle $\sigma\colon Z_{k-1}\to\mathbb{T}$ and a map $f\colon X\to\mathbb{T}$ so that

$$n\rho=\sigma\circ\pi_{k-1}+f\circ T-f\ .$$

Let $u\in U$. By Part (3) of Corollary 7.5, the cocycle $\rho\circ u-\rho$ is a quasi-coboundary and so there exist $\phi\colon X\to\mathbb{T}$ and $c\in\mathbb{T}$ with

$$\rho\circ u-\rho=\phi\circ T-\phi+c\ .$$

Plugging into the preceding equation, we get that the constant nc is a coboundary of X. That is, nc is an eigenvalue of this system and c belongs to the countable subgroup

$$\Gamma=\{c\in\mathbb{T}:nc\text{ is an eigenvalue of }X\}$$

of $\mathbb{T}$. For every $c\in\Gamma$, define

$$U_c=\{u\in U:\rho\circ u-\rho-c\text{ is a coboundary of }X\}\ .$$

Each of these sets is a Borel subset of U and their union is U. Thus there exists $c\in\Gamma$ such that $m_U(U_c)>0$, where m_U is the Haar measure of U. But U_0 is clearly a subgroup of U and U_c a coset of this subgroup. It follows that $m_U(U_0)>0$ and that U_0 is an open subgroup of U. Since U is connected, $U_0=U$. Thus for every $u\in U$ the cocycle $\rho\circ u-\rho$ is a coboundary. By Lemma C.9, there exists $\tau\colon Z_{k-1}\to\mathbb{T}$ and $g\colon X\to\mathbb{T}$ with

$$\rho=\tau\circ\pi_{k-1}+g\circ T-g\ .$$

By considering X as a system of order $k+1$, τ is a cocycle of type $k+1$ on Z_{k-1} by Corollary 7.8 and $n\tau$ is a cocycle of type k.

We use the notation and results of Section 9.1, applied to the system Z_{k-1}. By Lemma 9.3, $Z_{k-1,s}$ is a system of order $k-1$ for almost every $s\in Z_1$. By Lemma 9.1, for almost every s, the cocycle $\Delta\tau$ of the system $Z_{k-1,s}$ is of type k and the cocycle $n\Delta\tau$ of this system is of type $k-1$. By the inductive assumption, $\Delta\tau$ is a cocycle of type $k-1$ of this system. Using Lemma 9.1 again, τ is a cocycle of type k of the system Z_{k-1} and by Corollary 7.8 ρ is a cocycle of type k on X.

The first assertion of Theorem 9.5 is proven for k. It remains to show the second assertion for k.

We deduce it from the first part exactly as in the proof of Corollary 8.4. We reproduce it here for completeness. Z_{k+1} is an extension of Z_k by a compact abelian group U, given by a cocycle ρ of type $k+1$. Assume that U is not connected. This group admits an open subgroup U_0 such that U/U_0 is isomorphic to $\mathbb{Z}/n\mathbb{Z}$ for some integer $n>1$. We write $\overline{\rho}\colon Z_k\to U/U_0$ for the reduction of ρ modulo U_0; it is a cocycle of order $k+1$. Using the isomorphism from U/U_0 onto $\mathbb{Z}/n\mathbb{Z}$ and the natural embedding of $\mathbb{Z}/n\mathbb{Z}$ as a subgroup of $\mathbb{T}$, we get a cocycle $\tau\colon Z_k\to\mathbb{T}$, of type $k+1$, so that $n\tau=0$. Thus $n\tau$ is of type k and by the first part of the Theorem, τ is of type k.

Therefore $\overline{\rho}$ is of type k. The extension of Z_k associated to this cocycle is a factor of X and is of type k by Corollary 7.7. Proposition 4.11 provides a contradiction. □

9.4. **Countability.** The countability result that we have shown for the cocycles of order 2 (Proposition 8.9) cannot be generalized to higher orders. However, the weaker result proved in this section suffices for our purposes.

Notation. We let $\mathcal{C}_k(X)$ denote the subgroup of $\mathcal{C}(X)$ consisting in cocycles of type k.

Theorem 9.6. *Let $k \geq 2$ an integer, (X, μ, T) be an ergodic system, (Ω, P) a (standard) probability space and $\omega \mapsto \rho_\omega$ a measurable map from Ω to $\mathcal{C}_k(X)$. Then there exists a subset Ω_0 of Ω, with $P(\Omega_0) > 0$, so that $\rho_\omega - \rho_{\omega'} \in \mathcal{C}_1(X)$ for every $(\omega, \omega') \in \Omega_0 \times \Omega_0$.*

Proof. We proceed by induction on k.

By Corollary 7.9, Theorem 9.5 and Corollary 7.10, for every cocycle ρ of type 2 on X there exists a cocycle ρ' of type 2 on Z_1 so that ρ is cohomologous to $\rho' \circ \pi_1$. By Proposition 8.9, $\mathcal{C}_1(Z_1)$ has countable index in $\mathcal{C}_2(Z_1)$ and so $\mathcal{C}_1(X)$ has countable index in $\mathcal{C}_2(X)$. The statement of the Theorem follows immediately for $k = 2$.

Fix an integer $k \geq 2$ and assume that the Theorem holds for k. Let (X, μ, T), (Ω, P) be as in the statement of the Theorem and let $\omega \mapsto \rho_\omega$ be a measurable map from Ω to $\mathcal{C}_{k+1}(X)$.

We use the usual ergodic decomposition (formula (7)) of $\mu \times \mu$ for $T \times T$ and formula (8) for $\mu^{[k+1]}$. The map $\omega \mapsto \Delta\rho_\omega$ from Ω to $\mathcal{C}(X \times X)$ is measurable. By Lemma C.3 the subset

$$A = \big\{(\omega, s) \in \Omega \times Z_1 : \Delta\rho_\omega \in \mathcal{C}_k(X_s)\big\}$$

of $\Omega \times Z_1$ is measurable. In the same way, the subset

$$B = \big\{(\omega, \omega', s) \in \Omega \times \Omega \times Z_1 : \Delta\rho_\omega - \Delta\rho_{\omega'} \in \mathcal{C}_1(X_s)\big\}$$

of $\Omega \times \Omega \times Z_1$ is measurable. By Lemma 9.1, for all $\omega, \omega' \in \Omega$ the subset

$$B_{\omega,\omega'} = \{s \in Z_1 : (\omega, \omega', s) \in B\}$$

of Z_1 has measure 0 or 1. Moreover, for every $\omega \in \Omega$ the cocycle ρ_ω is of type $k+1$ by hypothesis and so by Lemma 9.1, the cocycle $\Delta\rho$ is of type k on X_s for μ_1-almost every $s \in Z_1$. Thus $(P \times \mu_1)(A) = 1$. Therefore, for μ_1-almost every $s \in Z_1$, using the inductive hypothesis applied to the system X_s and the map $\omega \mapsto \Delta\rho_\omega$, we get that

$$(P \times P)\big\{(\omega, \omega') \in \Omega \times \Omega : (\omega, \omega', s) \in B\big\} > 0 \ .$$

Therefore $(P \times P \times \mu_1)(B) > 0$ and the subset

$$C = \Big\{(\omega, \omega') \in \Omega \times \Omega : \mu_1(B_{\omega,\omega'}) > 0\Big\} = \Big\{(\omega, \omega') \in \Omega \times \Omega : \mu_1(B_{\omega,\omega'}) = 1\Big\}$$

of $\Omega \times \Omega$ has positive measure under $P \times P$. By applying Lemma 9.1 again, for $(\omega, \omega') \in C$, the cocycle $\rho_\omega - \rho_{\omega'}$ belongs to $\mathcal{C}_2(X)$. By the base step of the induction, $\mathcal{C}_1(X)$ has countable index in $\mathcal{C}_2(X)$ and so there exists $\rho \in \mathcal{C}_2(X)$ so that the set

$$D = \big\{(\omega, \omega') \in C : \rho_\omega - \rho_{\omega'} - \rho \in \mathcal{C}_1(X)\big\}$$

satisfies $(P \times P)(D) > 0$. Choose $\omega_0 \in \Omega$ so that the set

$$\Omega_0 = \{\omega \in \Omega : (\omega_0, \omega) \in D\}$$

has positive measure. Then for $\omega, \omega' \in \Omega_0$, $\rho_\omega - \rho_{\omega'} \in \mathcal{C}_1(X)$. $\square$

Corollary 9.7. *Let (X, μ, T) be an ergodic system and $\{S_u : u \in U\}$ a free action of a compact abelian group U on X by automorphisms. Let $\rho \colon X \to \mathbb{T}$ be a cocycle of type k for some integer $k \geq 2$. Then there exist a closed subgroup U_1 of U such* *that U/U_1 is a toral group and a cocycle ρ' cohomologous to ρ with $\rho' \circ S_u = \rho'$ for every $u \in U_1$.*

Proof. Define

$$U_0 = \{u \in U : \rho \circ \rho - \rho \text{ is a quasi-coboundary}\} .$$

Clearly, U_0 is a measurable subgroup of U.

The map $u \mapsto \rho \circ S_u - \rho$ is a measurable map from U to $\mathcal{C}_k(X)$ (and even to $\mathcal{C}_{k-1}(X)$ by Corollary 7.5). By Theorem 9.6 there exists a subset U_2 of U, with $m_U(U_2) > 0$, so that $\rho \circ S_u - \rho \circ S_v$ is a quasi-coboundary for every $u, v \in U_2$. We get immediately that $U_2 - U_2 \subset U_0$ and so $m_U(U_0) > 0$. Thus U_0 is an open subgroup of U.

By Lemma C.10 applied to the action $\{S_u : u \in U_0\}$, there exist a subgroup U_1 of U_0 and a cocycle ρ' on X with the required properties. (Note that U/U_1 is toral because U_0/U_1 is toral and U/U_0 is finite). □

10. Systems of order k and nilmanifolds

By using the tools developed in the preceding sections, we can now describe the structure of systems of order k. We show:

Theorem 10.1 (Structure Theorem)**.** *Any system of order $k \geq 1$ can be expressed as an inverse limit of a sequence of k-step nilsystems.*

The definition of nilsystems and the properties we use are summarized in Appendix B.

The proof splits into two parts. First we show show that every system of order k can be expressed as an inverse limit of simpler ones, called *toral systems* (Theorem 10.3). Then we show that each toral system of order k is actually a k-step nilsystem (Theorem 10.5).

10.1. **Reduction to toral systems.**

Definition 10.2. An ergodic system (X, μ, T) of order $k \geq 1$ is *toral* if $Z_1(X)$ is a compact abelian Lie group and for $1 \leq j < k$, $Z_{j+1}(X)$ is an extension of $Z_j(X)$ by a torus.

Theorem 10.3. *Any system of order $k \geq 1$ is an inverse limit of a sequence of toral systems of order k.*

We begin with a Lemma.

Lemma 10.4. *Let (X, μ, T) be an ergodic system, U a torus and $\rho\colon X \to U$ a cocycle of type $k+1$ for an integer $k \geq 0$. Assume that X is an inverse limit of a sequence $\{X_i : i \in \mathbb{N}\}$ of systems. Then ρ is cohomologous to a cocycle $\rho'\colon X \to U$, which is measurable with respect to $\mathcal{X}_i$ for some i.*

Proof of Lemma 10.4. We show by induction on ℓ that:

(*) *For integers $0 \leq \ell \leq k$, there exist $i_\ell \in \mathbb{N}$ and a cocycle ρ_ℓ cohomologous to ρ that is measurable with respect to $\mathcal{Z}_{k-\ell}(X) \vee \mathcal{X}_{i_\ell}$.*

By Corollary 7.9, ρ is cohomologous to a cocycle which factorizes through $Z_{k+1}(X)$. By Theorem 9.5, $Z_{k+1}(X)$ is an extension of $Z_k(X)$ by a connected compact abelian group. Using Corollary 7.10, there exists a cocycle ρ_0, cohomologous to ρ and measurable with respect to $\mathcal{Z}_k(X)$, and a fortiori with respect to $\mathcal{Z}_k(X) \vee \mathcal{X}_1$. The claim (*) holds for $\ell = 0$.

Let $0 \le \ell < k$ and assume that (*) holds for ℓ. Let i_ℓ and ρ_ℓ be as in the statement of the claim. By Corollary 7.8, ρ_ℓ is of type $k+1$.

Let Y be the factor of X corresponding to the σ-algebra $\mathcal{Y} = \mathcal{Z}_{k-\ell}(X) \vee \mathcal{X}_{i_\ell}$ and let W be the factor of X corresponding to $\mathcal{W} = \mathcal{Z}_{k-\ell-1}(X) \vee \mathcal{X}_{i_\ell}$. As $Z_{k-\ell}(X)$ is an extension of $Z_{k-\ell-1}(X)$ by a compact abelian group, by the first part of Lemma C.2, Y is an extension of W by a compact abelian group V. We identify Y with $W \times V$. As usual, for $v \in V$ we also let $v\colon Y \to Y$ denote the associated vertical rotation of Y above W.

By Corollary 9.7, there exist a closed subgroup V_1 of V so that V/V_1 is a compact abelian Lie group and a cocycle ρ', cohomologous to ρ_ℓ and thus to ρ, so that $\rho'(v \cdot y) = \rho(y)$ for every $v \in V_1$. We consider ρ' as a cocycle defined on the factor $W \times V/V_1$ of Y.

Since V/V_1 is a compact abelian Lie group, its dual group $\widehat{V/V_1} = V_1^\perp$ is finitely generated. Choose a finite generating set $\{\gamma_1, \dots, \gamma_m\}$ for $V_1^\perp$. For $1 \le j \le m$, consider γ_j as taking values in the circle group $\mathcal{S}^1$ and define the function f_j on $Y = W \times V$ by $f_j(w,v) = \gamma_j(v)$. Since X is the inverse limit of the sequence $\{X_i\}$, there exists $i \ge i_\ell$ so that for $1 \le j \le m$, $\mathbb{E}(f_j \mid \mathcal{X}_i) \neq 0$. Thus, $\mathbb{E}(f_j \mid \mathcal{W} \vee \mathcal{X}_i) \neq 0$. By Lemma C.2 the functions f_j are measurable with respect to $\mathcal{W} \vee \mathcal{X}_i$. But the functions f_j, $1 \le j \le m$, together with the σ-algebra $\mathcal{W}$, span the σ-algebra of the system $W \times V/V_1$. As ρ' is measurable with respect to this system, it is measurable with respect to $\mathcal{W} \vee \mathcal{X}_i = \mathcal{Z}_{k-\ell-1} \vee \mathcal{X}_i$. Therefore, (*) holds for $\ell+1$ with $i_{\ell+1} = i$. Property (*) with $\ell = k$ is the announced result. □

Proof of Theorem 10.3. We proceed by induction. For $k = 1$ the result is proven in Section 8.1.

Let $k \ge 1$ be an integer and assume that the result holds for k. Let Y be a system of order $k+1$. Write $X = Z_k(Y)$. Then Y is an extension of X by a compact abelian group U and let $\rho\colon X \to U$ be the cocycle defining this extension. By Theorem 9.5, U is connected and can be written as $\varprojlim U_j$, where each U_j is a torus. Let $\rho_j\colon X \to U_j$ be the projection of ρ on the quotient U_j of U.

By the inductive hypothesis, X can be written as an inverse limit $\varprojlim X_i$, where each X_i is toral. By Lemma 10.4, for every j there exist i_j and a U_j-valued cocycle ρ'_j, measurable with respect to $\mathcal{X}_{i_j}$, and cohomologous to ρ_j. We can clearly assume that the sequence $\{i_j\}$ is increasing. Each system $X_{i_j} \times_{\rho'_j} U_j$ is toral and $Y = X \times_\rho U$ is clearly the inverse limit of these systems. □

10.2. **Building nilmanifolds.** Here we show that every toral system can be given the structure of a k-step nilsystem. This is obtained by showing that the group $\mathcal{G}$ associated to this system as in Section 5 is a Lie group and acts transitively.

Theorem 10.5. *Let (X, μ, T) be a toral system of order $k \ge 1$. Then:*

(1) *$\mathcal{G} = \mathcal{G}(X)$ is a Lie group and is k-step nilpotent.*
(2) *Let G be the subgroup of $\mathcal{G}$ spanned by the connected component of the identity and T. Then G admits a discrete co-compact subgroup Λ so that the system X is isomorphic to the nilmanifold G/Λ, endowed with Haar measure and left translation by T.*

(See Appendix B for more on nilmanifolds.)

The proof is by induction on the order k of the system. When $k=1$, the system is a rotation on a compact abelian Lie group Z. We have $\mathcal{G}(X)=Z$, acting on itself by translations and the first statement is obvious. By ergodicity $G=Z$ and the second statement holds with $\Lambda=\{1\}$.

Let $k\geq 1$ be an integer and assume that both statements of the Theorem hold for every toral system of order k.

10.2.1. *Conditions for lifting.* Throughout this section, $k\geq 1$ is an integer and (Y,ν,S) is a *toral* system of order $k+1$. We write (X,μ,T) for $Z_k(Y)$, where Y is an extension of X by a torus U, given by a cocycle $\rho\colon X\to U$ of type $k+1$. By the inductive hypothesis, $\mathcal{G}(X)$ is a Lie group.

By Lemma 5.2, every element $\overline{g}$ of $\mathcal{G}(Y)$ induces a transformation $p_k\overline{g}$ of X, which belongs to $\mathcal{G}(X)$. We now study the inverse problem. We say that an element g of $\mathcal{G}(X)$ *can be lifted* to an element of $\mathcal{G}(Y)$ if there exists $\overline{g}\in\mathcal{G}(Y)$ with $p_k\overline{g}=g$. We now establish conditions for lifting.

We use the maps $F\colon X^{[k+1]}\to U$ and $\Phi\colon Y^{[k+1]}\to U$ introduced in Proposition 6.4. We have

$$\Delta^{k+1}\rho=F\circ T^{[k+1]}-F \text{ and } \Phi(\mathbf{x},\mathbf{u})=F(\mathbf{x})-\sum_{\epsilon\in V_{k+1}}s(\epsilon)u_\epsilon \tag{25}$$

under the identification of $Y^{[k+1]}$ with $X^{[k+1]}\times U^{[k+1]}$. By Proposition 6.4, the σ-algebra $\mathcal{I}^{[k+1]}(Y)$ is spanned by the σ-algebra $\mathcal{I}^{[k+1]}(X)$ and the map Φ.

Lemma 10.6. *Let $g\in\mathcal{C}(X)$. If $\overline{g}\in\mathcal{G}(Y)$ is a lift of g, then $\overline{g}$ is given by*

$$\overline{g}\cdot(x,u)=(g\cdot x,u+\phi(x)) \tag{26}$$

where $\phi\colon X\to U$ is a map satisfying

$$F\circ g^{[k+1]}-F=\Delta^{k+1}\phi\ . \tag{27}$$

Conversely, if $\phi\colon X\to U$ satisfies Equation (27), then the transformation $\overline{g}$ of Y given by Equation (26) is a lift of g to $\mathcal{G}(Y)$.

Proof. Let $g\in\mathcal{G}(X)$ and assume that g admits a lift $\overline{g}\in\mathcal{G}(Y)$. By Corollary 5.10, the vertical rotations of Y over X belong to the center of $\mathcal{G}(Y)$ and thus commute with $\overline{g}$. It follows that $\overline{g}$ has the form given by Equation (26) for some $\phi\colon X\to U$. As $\overline{g}\in\mathcal{G}(Y)$, the transformation $\overline{g}^{[k+1]}$ of $Y^{[k+1]}$ acts trivially on $\mathcal{I}^{[k+1]}(Y)$ and thus leaves the map Φ invariant. This implies immediately that ϕ satisfies Equation (27).

Conversely, let $g\in\mathcal{G}(X)$, $\phi\colon X\to U$ be a map satisfying Equation (27) and let $\overline{g}$ be the measure preserving transformation of Y given by Equation (26). Since $\nu^{[k+1]}$ is conditionally independent over $\mu^{[k+1]}$ and $g^{[k+1]}$ leaves the measure $\mu^{[k+1]}$ invariant, $\overline{g}^{[k+1]}$ leaves the measure $\mu^{[k+1]}$ invariant. Moreover, Equation (27) means exactly that the map Φ is invariant under $\overline{g}^{[k+1]}$. Since $g\in\mathcal{G}(X)$, $g^{[k+1]}$ acts trivially on $\mathcal{I}^{[k+1]}(X)$ By Proposition 6.4, $\overline{g}^{[k+1]}$ acts trivially on $\mathcal{I}^{[k+1]}(Y)$. By Corollary 6.6, $\overline{g}\in\mathcal{G}(Y)$. $\square$

Corollary 10.7. *The kernel of the group homomorphism $p_k\colon\mathcal{G}(Y)\to\mathcal{G}(X)$ consists in the transformations of the form $(x,u)\mapsto(x,u+\phi(x))$, where $\phi\in\mathcal{D}_{k+1}(X,U)$ (see Section 7.1).*

In order to build lifts of elements of $\mathcal{G}(X)$, we progress from $\mathcal{G}^{(k-1)}(X)$ to $\mathcal{G}(X)$ along the lower central series of $\mathcal{G}(X)$. For $1 \leq j < k$, we show that 'many' elements of $\mathcal{G}^{(j)}(X)$ satisfy a property stronger than the lifting condition of Lemma 10.6. We need some notation.

Notation. Let β be a ℓ-face of V_{k+1} and $\phi\colon X \to U$ a map. We write $\Delta_\beta^{k+1}: X^{[k+1]} \to U$ for the map given by

$$\Delta_\beta^{k+1}\phi(\mathbf{x}) = \sum_{\epsilon \in \beta} s(\epsilon)\phi(x_\epsilon)\ .$$

The projection $\xi_\beta^{[k+1]}\colon X^{[k+1]} \to X^{[\ell]}$ is defined in Section 2.1. We have that

$$\Delta_\beta^{k+1}\phi(\mathbf{x}) = \pm\Delta^\ell\phi\big(\xi_\beta^{[k+1]}(\mathbf{x})\big)\ ,$$

where the sign depends on the face β.

Lemma 10.8. *Let j be an integer with $0 \leq j < k$. For $g \in \mathcal{G}^{(j)}(X)$ and $\phi\colon X \to U$, the following are equivalent:*

1. *For every $(k+1-j)$-face β of V_{k+1}, $F \circ g_\beta^{[k+1]} - F = \Delta_\beta^{k+1}\phi$.*
2. *For every $(k-j)$-face α of V_{k+1}, $F \circ g_\alpha^{[k+1]} - F - \Delta_\alpha^{k+1}\phi$ is invariant on $X^{[k+1]}$.*

Notation. We write $\mathcal{G}_0^{(j)}$ for the set of $g \in \mathcal{G}^{(j)}(X)$ so that there exists $\phi\colon X \to U$ satisfying the properties of Lemma 10.8.

Proof. The proof is similar to the proof of Lemma 10.6. Let $g \in \mathcal{G}^{(j)}(X)$. Let $\phi\colon X \to U$ and let $\overline{g}$ be the measure preserving transformation of $Y = X \times U$ given by Equation (26). As $g \in \mathcal{G}^{(j)}(X)$, the measure $\mu^{[k+1]}$ is invariant under $g_\alpha^{[k+1]}$ whenever α is a $(k-j)$-face of V_{k+1}. Also, $\nu^{[k+1]}$ is invariant under $\overline{g}_\alpha^{[k+1]}$ because this measure is conditionally independent over $\mu^{[k+1]}$. So for a $(k-j+1)$-face β, $\nu^{[k+1]}$ is invariant under $\overline{g}_\beta^{[k+1]}$.

The first property means that the function Φ (see Proposition 6.4) defined above is invariant under $\overline{g}_\beta^{[k+1]}$ for every $(k+1-j)$-face β of V_{k+1}. Moreover, by Lemma 5.8, $g_\beta^{[k+1]}$ acts trivially on $\mathcal{I}^{[k+1]}(X)$ because $g \in \mathcal{G}^{(j)}(X)$. Therefore, the first property means that $\overline{g}_\beta^{[k+1]}$ acts trivially on $\mathcal{I}^{[k+1]}(Y)$ for any $(k+1-j)$-face β of V_{k+1}.

Similarly, the second property means that for every $(k-j)$-face α of V_{k+1}, $\overline{g}_\alpha^{[k+1]}$ maps the σ-algebra $\mathcal{I}^{[k+1]}(Y)$ to itself.

The equivalence of these properties follows from Lemma 5.3. $\square$

Note that for $j = 0$ the first property of Lemma 10.8 coincides with the condition given in Lemma 10.6. Therefore, $\mathcal{G}_0^{(0)}$ consists in the elements of $\mathcal{G}(X)$ which can be lifted to an element of $\mathcal{G}(Y)$ and $\mathcal{G}_0^{(0)} = p_k\big(\mathcal{G}(Y)\big)$.

More generally, let $g \in \mathcal{G}_0^{(j)}$ for some j and ϕ satisfying the first property of Lemma 10.8. Then ϕ obviously satisfies Equation (27), and the transformation $\overline{g}$ of Y given by Equation (26) is a lift of g in $\mathcal{G}(Y)$. Therefore, p_k maps $p_k^{-1}\big(\mathcal{G}_0^{(j)}\big)$ onto $\mathcal{G}_0^{(j)}$. Each element $\overline{g}$ of $\mathcal{G}(Y)$ is given by Equation (26) for $g = p_k(\overline{g})$ and some ϕ, and $p_k^{-1}\big(\mathcal{G}_0^{(j)}\big)$ consists in those $\overline{g}$ for which the map ϕ satisfies the conditions of Lemma 10.8. Therefore, $p_k^{-1}\big(\mathcal{G}_0^{(j)}\big)$ is a closed subgroup of $\mathcal{G}(Y)$.

10.2.2. *Lifting results.* We maintain the same notations as in Section 10.2.1.

Lemma 10.9. *Each element of $\mathcal{G}^{(k-1)}(X)$ can be lifted to an element of $\mathcal{G}(Y)$. More precisely, $\mathcal{G}_0^{(k-1)} = \mathcal{G}^{(k-1)}(X)$.*

Proof. Let $g \in \mathcal{G}^{(k-1)}(X)$. We use the results of Section 5. Since $\mathcal{G}(X)$ is k-step nilpotent, g belongs to the center of $\mathcal{G}(X)$ and thus commutes with T and is an automorphism of X. Since $\mathcal{G}(Z_{k-1})$ is $(k-1)$-step nilpotent, g induces the trivial transformation on Z_{k-1}. Thus g is a vertical rotation of X over Z_{k-1}. For every edge α of V_{k+1}, the transformation $g_\alpha^{[k+1]}$ leaves the measure $\mu^{[k+1]}$ invariant and commutes with $T^{[k+1]}$ by Corollary 5.4. By Equation (25),

$$\begin{aligned}(28)\quad \partial\big(F \circ g_\alpha^{[k+1]} - F\big) = \big(\Delta^{k+1}\rho\big) \circ g_\alpha^{[k+1]} - \Delta^{k+1}\rho &= \Delta_\alpha^{k+1}(\rho \circ g - \rho) \\ &= \pm\Delta(\rho \circ g - \rho) \circ \xi_\alpha^{[k+1]} \ .\end{aligned}$$

By Lemma C.7, $\Delta(\rho \circ g - \rho)\colon X^2 \to U$ is a coboundary. As U is a torus, by Lemma C.5, $\rho \circ g - \rho$ is a quasi-coboundary. Thus there exists $\phi\colon X \to U$ and $c \in U$ with

$$\rho \circ g - \rho = \phi \circ T - \phi + c \ . \tag{29}$$

Using this in Equation (28), we get that for every edge α there exists an invariant map $i\colon X^{[k+1]} \to U$, with

$$F \circ g_\alpha^{[k+1]} - F = \Delta_\alpha^{[k+1]}\phi + i \ .$$

By Lemma 10.8, $g \in \mathcal{G}_0^{(k-1)}$. □

The next Proposition is the crucial step in the proof. We recall that $\mathcal{G}(X)$ is a Lie group.

Proposition 10.10. *For an integer j with $0 \le j < k$, $\mathcal{G}_0^{(j)}$ is open in $\mathcal{G}^{(j)}(X)$.*

Proof. We proceed by induction downwards on j. For $j = k-1$, $\mathcal{G}_0^{(j)} = \mathcal{G}^{(j)}(X)$ by Lemma 10.9. Take j with $0 < j \le k-1$ and assume that $\mathcal{G}_0^{(j)}$ is open in $\mathcal{G}^{(j)}(X)$. We prove now that $\mathcal{G}_0^{(j-1)}$ is open in $\mathcal{G}^{(j-1)}(X)$.

Since $\mathcal{G}_0^{(j)}$ is an open subgroup of $\mathcal{G}^{(j)}(X)$, it is also closed and it is locally compact and Polish (actually it is a Lie group). We have noted that the continuous group homomorphism $p_k\colon p_k^{-1}\big(\mathcal{G}_0^{(j)}\big) \to \mathcal{G}_0^{(j)}$ is onto. By Theorem A.1, this homomorphism admits a Borel cross section.

Let $\mathcal{H} = \{g \in \mathcal{G}^{(j-1)}(X) : [g^{-1}; T^{-1}] \in \mathcal{G}_0^{(j)}\}$. By the inductive hypothesis, $\mathcal{H}$ is open in $\mathcal{G}^{[j-1]}(X)$, and is locally compact. Consider the Borel map $\kappa\colon \mathcal{H} \to \mathcal{G}(Y)$ obtained by composing the continuous map $g \mapsto [g^{-1}; T^{-1}]$ from $\mathcal{H}$ to $\mathcal{G}_0^{(j)}$ with a Borel cross section $\mathcal{G}_0^{(j)} \to \mathcal{G}(Y)$. For $g \in \mathcal{H}$, $\kappa(g)$ is given by Equation (26) for some map $\psi_g\colon X \to U$ so that the properties of Lemma 10.8 are satisfied with $[g^{-1}; T^{-1}]$. That is, for every $(k+1-j)$-face β of V_{k+1},

$$F \circ [g^{-1}; T^{-1}]_\beta^{[k+1]} - F = \Delta_\beta^{k+1}\psi_g \ .$$

Define $\theta_g = \psi_g \circ Tg + \rho \circ g - \rho$.

Let β be a $(k+1-j)$-face of V_{k+1}. Then

$$\begin{aligned}
\bigl(F\circ g_\beta^{[k+1]}-F\bigr)\circ T^{[k+1]}-\bigl(F\circ g_\beta^{[k+1]}-F\bigr)&\\
=&\bigl(F\circ[g^{-1};T^{-1}]_\beta^{[k+1]}T^{[k+1]}g_\beta^{[k+1]}-F\circ T^{[k+1]}g_\beta^{[k+1]}\bigr)\\
&+\bigl(F\circ T^{[k+1]}g_\beta^{[k+1]}-F\circ g_\beta^{[k+1]}\bigr)-\bigl(F\circ T^{[k+1]}-F\bigr)\\
=&\Delta_\beta^{k+1}\psi_g\circ T^{[k+1]}g_\beta^{[k+1]}+(\Delta^{k+1}\rho)\circ g_\beta^{[k+1]}-\Delta^{k+1}\rho\\
=&\Delta_\beta^{k+1}\theta_g\\
=&\pm\Delta^{k+1-j}\theta_g\circ\xi_\beta^{[k+1]}\ .
\end{aligned}$$

Thus the cocycle $\Delta_\beta^{k+1}\theta_g$ is a coboundary of the system $X^{[k+1]}$. As already noted, the cocycle $(\Delta^{k+1-j}\theta_g)\circ\xi_\beta^{[k+1]}$ is equal to this coboundary or to its opposite and thus is a coboundary. By Lemma C.7, $\Delta^{k+1-j}\theta_g$ is a coboundary of the system $X^{[k+1-j]}$ and θ_g is a cocycle of type $k+1-j\leq k$ on X.

Since the map κ defined above is Borel, the map $g\mapsto\psi_g$ from $\mathcal{H}$ to $\mathcal{C}(X,U)$ is Borel, and the map $g\mapsto\theta_g$ is a Borel map from $\mathcal{H}$ to the group $\mathcal{C}_{k+1-j}(X,U)$ of U-valued cocycles of type $k+1-j$ on X. Choose a probability measure λ on $\mathcal{H}$, equivalent to the Haar measure of $\mathcal{H}$ and apply Theorem 9.6. Then there exists a measurable subset A of $\mathcal{H}$, with $\lambda(A)>0$, so that $\theta_g-\theta_h$ is a quasi-coboundary for every $(g,h)\in A\times A$.

Let $g,h\in A$. Let $\theta\colon X\to U$ and $c\in U$ be such that $\theta_g-\theta_h=\partial\theta+c$. For any $(k+1-j)$-face β of V_{k+1}, by the last equation we get that

$$\partial\bigl(F\circ g_\beta^{[k+1]}-F\circ h_\beta^{[k+1]}\bigr)=\partial\Delta_\beta^{k+1}\theta\ .$$

Thus $F\circ g_\beta^{[k+1]}-F\circ h_\beta^{[k+1]}-\Delta_\beta^{k+1}\theta$ is an invariant function on $X^{[k+1]}$. As $h\in\mathcal{G}^{(j-1)}(X)$, the transformation $h_\beta^{[k+1]}$ maps the σ-algebra $\mathcal{I}^{[k+1]}(X)$ to itself. Therefore, the function $F\circ(gh^{-1})_\beta^{[k+1]}-F-\Delta_\beta^{k+1}(\theta\circ h^{-1})$ is invariant. The second property of Lemma 10.8 is satisfied and $gh^{-1}\in\mathcal{G}_0^{(j-1)}$.

Therefore $A\cdot A^{-1}\subset\mathcal{G}_0^{(j-1)}$. Since $\mathcal{H}$ is open in $\mathcal{G}^{(j-1)}$, A has positive Haar measure in $\mathcal{G}^{(j-1)}$ and it follows that $\mathcal{G}_0^{(j-1)}$ also has positive Haar measure in $\mathcal{G}^{(j-1)}$. Since $\mathcal{G}_0^{(j-1)}$ is a Borel subgroup of $\mathcal{G}^{(j-1)}(X)$, it is an open subgroup. □

10.2.3. *End of the proof of Theorem 10.5.*

Proof. Recall that $k\geq 1$ is an integer and that we assume that the properties of Theorem 10.5 hold for every toral system of order k. Let (Y,ν,S) be a toral system of order $k+1$. We write $(X,\mu,T)=Z_k(Y)$. By the inductive hypothesis, the conclusions of Theorem 10.5 hold for this system. Let G and Λ be as in this Theorem and let $\mathcal{G}_0^{(0)}$ be as in the preceding subsection.

(1) By Proposition 10.10 used with $j=0$ the group $\mathcal{G}_0^{(0)}$ is open in $\mathcal{G}^{(0)}(X)=\mathcal{G}(X)$ and thus is a Lie group. The restriction map $p_k\colon\mathcal{G}(Y)\to\mathcal{G}(X)$ is a continuous group homomorphism and maps $\mathcal{G}(Y)$ onto $\mathcal{G}_0^{(0)}$. Its kernel is $\mathcal{D}_{k+1}(X,U)$ by Corollary 10.7 and thus is a Lie group. Since $\mathcal{G}_0^{(0)}$ and $\mathcal{D}_{k+1}(X,U)$ are both Lie groups, $\mathcal{G}(Y)$ is a Lie group by Corollary A.2 and Lemma A.3 (see Appendix A).

(2) Let H be the subgroup of $\mathcal{G}(Y)$ spanned by the connected component of the identity and S. The image under p_k of the connected component of the identity of $\mathcal{G}(Y)$ is included in the connected component of the identity of $\mathcal{G}(X)$; moreover $p_k(S) = T$ and thus $p_k(H) \subset G$. Since p_k maps $\mathcal{G}(Y)$ onto $\mathcal{G}_0^{(0)}$, it is an open map and $p_k(H)$ is an open subgroup of $\mathcal{G}_0^{(0)}$ and thus also of $\mathcal{G}(X)$. Therefore $p_k(H)$ contains the connected component of the identity in $\mathcal{G}(X)$ and so it contains G. We get that $p_k(H) = G$.

On the other hand, for every $u \in U$, the corresponding vertical rotation belongs to $\mathcal{G}(Y)$ and it defines an embedding of U in $\mathcal{G}(Y)$. $H \cap U$ is an open subgroup of U and since U is connected, $U \subset H$.

By the inductive assumption, $X = G/\Lambda$. This means that G acts transitively on X and that Λ is is the stabilizer of the point x_1 of X, image of the identity element of G under the natural projection $G \to G/\Lambda = X$. Choose a lift y_1 of x_1 in Y and consider the map $f\colon H \to Y$ given by $f(h) = h \cdot x_1$. Since $U \subset H$, the range of this map is invariant under all vertical rotations. The projection of this range on X is onto. Therefore f is onto.

This defines a bijection of H/Γ onto Y, where Γ is the stabilizer of y_1 in H. This bijection commutes with the actions of H on Y and H/Γ. The measure on H/Γ corresponding to ν through this bijection is invariant under the action of H and thus is the Haar measure of H/Γ.

Thus we are left only with checking that Γ is discrete and cocompact in H. Clearly, $\Gamma \cdot U = p_k^{-1}(\Lambda)$. Since $\Gamma \cap U$ is trivial, Γ is discrete. This also implies that $H/\Gamma U$ is homeomorphic to G/Λ and thus is compact. Since U is compact, Γ is cocompact in H. □

11. The measures $\mu^{[k]}$

We can prove a converse to Theorem 10.5, showing that every k-step ergodic nilsystem is a system of order k. Therefore the expressions "toral system of order k" and "k-step ergodic nilsystem" are actually synonymous. However, as we have no need for this result, we do not prove it and we keep using the term "toral system of order k".

When (X, μ, T) is a toral system of order ℓ for some integer ℓ, the measures $\mu^{[k]}$, $k \geq 1$, have a simple description, which is used in the proof of Theorem 1.2 (convergence for "cubic averages").

11.1. **Algebraic preliminaries.** In this Section G is a nilpotent Lie group. We study the sequence of groups $G_{k-1}^{[k]}$ for $k \geq 1$ and the relations between two consecutive groups of this form.

Temporarily, we slightly modify the definition of $G_{k-1}^{[k]}$ given by Definition 18: $G_{k-1}^{[k]}$ is the subgroup of $G^{[k]}$ spanned by

$$\{g_\alpha^{[k]} : g \in G \text{ and } \alpha \text{ is an } \ell\text{-face of } V_k.\}$$

Therefore the group $G_{k-1}^{[k]}$ with the preceding definition is the closure of the present group $G_{k-1}^{[k]}$. Below we show that this group is actually closed and thus the two definitions coincide. Recall that the groups $\mathcal{G}^{(j)}$ are equal to the algebraic iterated groups of commutators (see Lemma B.1).

Let $k \geq 1$ be an integer. As usual, we write $\mathbf{g} = (\mathbf{g}', \mathbf{g}'')$ for a point of $G^{[k+1]}$, where $\mathbf{g}', \mathbf{g}'' \in G^{[k]}$ are given by

$$\mathbf{g}'\epsilon : \mathbf{g}_{\epsilon 0} \text{ and } \mathbf{g}''_\epsilon = \mathbf{g}_{\epsilon 1} \text{ for } \epsilon \in V_k \ .$$

We also identify the element $\mathbf{g} = (\mathbf{g}', \mathbf{g}'')$ of $G^{[k+1]}$ with the element $(g'_\epsilon, g''_\epsilon : \epsilon \in V_k)$ of $(G \times G)^{[k]}$ and thus we have $G^{[k+1]} = (G \times G)^{[k]}$.

Lemma 11.1. *Let*

$$G_\bullet^{[k]} = \left\{ \mathbf{g} \in G^{[k]} : (\mathbf{g}, 1^{[k]}) \in G_k^{[k+1]} \right\} \ .$$

Then $G_\bullet^{[k]}$ is a normal subgroup of $G_{k-1}^{[k]}$ and

$$G_k^{[k+1]} = \left\{ (\mathbf{g}', \mathbf{g}'') \in G_{k-1}^{[k]} \times G_{k-1}^{[k]} : \mathbf{g}'{\mathbf{g}''}^{-1} \in G_\bullet^{[k]} \right\} \ . \tag{30}$$

Proof. For $\mathbf{g}' \in G_{k-1}^{[k]}$, we have $(\mathbf{g}', \mathbf{g}') \in G_k^{[k+1]}$. For $\mathbf{h} = (\mathbf{h}', \mathbf{h}'') \in G_k^{[k+1]}$, we have $\mathbf{h}', \mathbf{h}'' \in G_{k-1}^{[k]}$. The result follows. □

We also note that $g^{[k]} \in G_\bullet^{[k]}$ for every $g \in G$.

Lemma 11.2. *Define*

$$\widetilde{G} = \left\{ (g', g'') \in G \times G : g''{g'}^{-1} \in G^{(1)} \right\} \ .$$

Then $\widetilde{G}_{k-1}^{[k]}$ is a normal subgroup of $G_k^{[k+1]}$.

Moreover, writing ζ for the side $\{\epsilon \in V_{k+1} : \epsilon_{k+1} = 0\}$, we have

$$G_k^{[k+1]} = \left\{ h_\zeta^{[k+1]} \mathbf{g} : h \in G,\ \mathbf{g} \in \widetilde{G}_{k-1}^{[k]} \right\} \ .$$

If $h_\zeta^{[k+1]}\mathbf{g} = {h'}_\zeta^{[k+1]}\mathbf{g}'$ for some $h, h' \in G$ and $\mathbf{g}, \mathbf{g}' \in \widetilde{G}_{k-1}^{[k]}$, then $h' = hu^{-1}$ and $\mathbf{g}' = u_\alpha^{[k+1]}\mathbf{g}$ for some $u \in G^{(1)}$.

(Here we consider $\widetilde{G}_{k-1}^{[k]}$ as a subgroup of $G^{[k+1]}$.)

Proof. We claim that, for every $g \in G$ and every $\mathbf{h} \in \tilde{G}_{k-1}^{[k]}$ we have

$$\left(g_\zeta^{[k+1]}\right)^{-1} \mathbf{h}\, g_\zeta^{[k+1]} \in \tilde{G}_{k-1}^{[k]} \ . \tag{31}$$

First we consider the case that $\mathbf{h} = (h, h)_\alpha^{[k]}$ for some $h \in G$ and some side α of V_k. Then, under the identification of $(X \times X)^{[k]}$ with $X^{[k+1]}$, $\mathbf{h} = h_\beta^{[k+1]}$ where β is the side $\alpha \times \{0, 1\}$ of V_{k+1}. We notice that $\beta \cap \zeta = \alpha \times \{0\}$. By Equation (19) we have

$$\left[\mathbf{h}; g_\zeta^{[k+1]}\right] = [h; g]_{\beta\cap\zeta}^{[k+1]} = \left([h; g], 1\right)_\alpha^{[k]} \in \tilde{G}_{k-1}^{[k]}$$

because $[h; g] \in G^{(1)}$ and thus $([h; g], 1) \in \tilde{G}$. The relation (31) holds in this case.

We consider now the case that $\mathbf{h} = (1, u)_\alpha^{[k]}$ for some $u \in G^{(1)}$ and some side α of V_k. We have $\mathbf{h} = u_\gamma^{[k+1]}$ where γ is the $(k-1)$-face $\alpha \times \{1\}$ of V_{k+1}. We notice that $\gamma \cap \zeta = \emptyset$. It follows that $[\mathbf{h}; g_\zeta^{[k+1]}] = 1$ and the relation (31) holds in this case also.

Therefore, when α is a side of V_k, this relation holds whenever $\mathbf{h} = (g', g'')_\alpha^{[k+1]}$ for any $(g', g'') \in \tilde{G}$. This relation holds for every $\mathbf{h} \in \tilde{G}_{k-1}^{[k]}$ by definition of this group. The claim is proven.

By definition, every element of $G_k^{[k+1]}$ can be expressed as a product of elements of one of the following three types.

(1) $g^{[k+1]}$ for some $g\in G$.
(2) $g_\beta^{[k]}$ for some $g\in G$ and some side β of V_{k+1} defined by fixing a coordinate $j<k+1$.
(3) $g_\zeta^{[k+1]}$ for some $g\in G$.

Let $g\in G$. $g^{[k+1]}=(g,g)^{[k]}\in\tilde{G}^{[k]}_{k-1}$ because $(g,g)\in\tilde{G}$. Let β be a side of V_{k+1} defined by fixing a coordinate $j<k+1$. Then $\beta=\alpha\times\{0,1\}$ where α is a side of V_k and $g_\beta^{[k+1]}=(g,g)_\alpha^{[k]}\in\tilde{G}^{[k]}_{k-1}$.

Therefore, every element of the types (1) or (2) above belongs to $\tilde{G}^{[k]}_{k-1}$. The first two assertions of Lemma 11.2 follows immediately from the relation (31).

If we have $h_\zeta^{[k+1]}\mathbf{g}={h'}_\zeta^{[k+1]}\mathbf{g'}$ as in the third statement of the Lemma, then $(hh'^{-1})_\zeta^{[k+1]}\in\widetilde{G}^{[k]}_{k-1}$. Thus $(hh'^{-1},1)\in\widetilde{G}$ and $hh'^{-1}\in G^{(1)}$. $\square$

By induction, the commutator subgroups $\widetilde{G}^{(j)}$, $j\geq 0$, of $\widetilde{G}$ are given by

$$\widetilde{G}^{(j)}=\bigl\{(g',g'')\in G^{(j)}\times G^{(j)}:g''{g'}^{-1}\in G^{(j+1)}\bigr\}\ .$$

Lemma 11.3. *Let*

$$G_*^{[k]}=\{\mathbf{g}\in G^{[k]}:(\mathbf{g},1^{[k]})\in\widetilde{G}^{[k]}_{k-1}\}\ .$$

Then $G_*^{[k]}$ is a normal subgroup of $G_\bullet^{[k]}$ and

$$G_\bullet^{[k]}=\{h^{[k]}\mathbf{g}:h\in G,\ \mathbf{g}\in G_*^{[k]}\}\ .$$

Proof. We claim that

$$(G^{(1)})^{[k]}_{k-1}\subset G_*^{[k]}\subset(G^{(1)})^{[k]}\ .$$

When $u\in G^{(1)}$ and α is a side of V_k we have $(u_\alpha^{[k]},1^{[k]})=(u,1)_\alpha^{[k]}\in\tilde{G}^{[k]}_{k-1}$ because $(1,u)\in\tilde{G}$ and thus $u_\alpha^{[k]}\in G_*^{[k]}$. The first inclusion follows. Moreover, when $\mathbf{g}\in G_*^{[k]}$, we have $(\mathbf{g},1^{[k]})\in\tilde{G}^{[k]}_{k-1}$ thus for every $\epsilon\in V_k$ we have $(g_\epsilon,1)\in\tilde{G}$ and thus $g_\epsilon\in G^{(1)}$. The second inclusion follows and the claim is proven.

Since $\widetilde{G}^{[k]}_{k-1}$ is a normal subgroup of $G_k^{[k+1]}$, it follows from the definition of $G_*^{[k]}$ and $G_\bullet^{[k]}$ that $G_*^{[k]}$ is a normal subgroup of $G_\bullet^{[k]}$.

Let $\mathbf{q}\in G_\bullet^{[k]}$. We have $(\mathbf{q},1^{[k]})\in G_k^{[k+1]}$. By Lemma 11.2, there exists $h\in G$ and $\mathbf{g}\in\widetilde{G}^{[k]}_{k-1}$ with $(\mathbf{q},1^{[k]})=h_\zeta^{[k+1]}\mathbf{g}$. The element $\mathbf{g}$ has the form $\mathbf{g}=(\mathbf{g'},1^{[k]})$, $\mathbf{g'}\in G_*^{[k]}$ by definition and $\mathbf{q}=h^{[k]}\mathbf{g'}$. $\square$

If for some $\mathbf{q}\in G_\bullet^{[k]}$ and some $\epsilon\in V_k$ we have $q_\epsilon\in G^{(1)}$, then, writing $\mathbf{q}=h^{[k]}\mathbf{g}$ as in Lemma 11.3, we have that $h\in G^{(1)}$. Thus $h^{[k]}\in G_*^{[k]}$ and $\mathbf{q}\in G_*^{[k]}$. This proves:

Lemma 11.4. *For every $\epsilon\in V_k$,*

$$G_*^{[k]}=\{\mathbf{q}\in G_\bullet^{[k]}:q_\epsilon\in G^{(1)}\}\ .$$

In particular, $G_*^{[k]}=G_\bullet^{[k]}\cap(G^{(1)})^{[k]}$.

11.2. **Topological results.**

Lemma 11.5. *Let G be a nilpotent Lie group. For any integer $k \geq 1$, the group $G_{k-1}^{[k]}$ is closed in $G^{[k]}$.*

Proof. By induction on k. For $k = 1$, $G_0^{[1]} = G^{[1]} = G \times G$ and there is nothing to prove. Take $k \geq 1$ and assume that the result holds for k and any nilpotent Lie group. We use the notation of the preceding subsection.

Since $\widetilde{G}$ is a nilpotent Lie group, by the inductive hypothesis $\widetilde{G}_{k-1}^{[k]}$ is closed in $\widetilde{G}^{[k]}$. Thus it is complete and thus it is closed in $G^{[k+1]}$. Therefore $G_*^{[k]}$ is closed in $G^{[k]}$.

Let $\{\mathbf{g}_n\}$ be a sequence in $G_\bullet^{[k]}$, converging in $G^{[k]}$ to some element $\mathbf{g}$. For every integer n, let θ_n be the image of the first coordinate $(g_n)_\mathbf{0}$ of $\mathbf{g}_n$ in $G/G^{(1)}$. Then θ_n converges to the projection of $g_\mathbf{0}$ in $G/G^{(1)}$. As $G/G^{(1)}$ is endowed with the quotient topology, the sequence $\{\theta_n\}$ can be lifted in a sequence $\{h_n\}$ in G, convergent to some $h \in G$. The sequence $\{({h_n}^{[k]})^{-1}\mathbf{g}_n\}$ converges in $G^{[k]}$ to $(h^{[k]})^{-1}\mathbf{g}$. For every n, we have that $({h_n}^{[k]})^{-1}\mathbf{g}_n \in G_\bullet^{[k]}$ and its $\mathbf{0}$ coordinate is equal to 1. Thus by Lemma 11.4, this element belongs to $G_*^{[k]}$. Since this group is closed, $(h^{[k]})^{-1}\mathbf{g} \in G_*^{[k]}$ and it follows that $\mathbf{g} \in G_\bullet^{[k]}$. Therefore $G_\bullet^{[k]}$ is closed in $G^{[k]}$.

The announced result follows now immediately from Lemma 11.1. □

Along the way, we have shown that

$$G_*^{[k]} \textit{ and } G_\bullet^{[k]} \textit{ are closed subgroups of } G^{[k]}.$$

Recall that if Λ is a discrete cocompact subgroup of a nilpotent Lie group G, then for every j the group $G^{(j)}\Lambda$ is closed in G (see Lemma B.1). It follows that for every j, the group $\Lambda \cap G^{(j)}$ is cocompact in $G^{(j)}$.

Lemma 11.6. *Let G be a nilpotent Lie group and Λ a discrete cocompact subgroup of G. For every integer $k \geq 1$, the group $\Lambda^{[k]} \cap G_{k-1}^{[k]}$ is cocompact in $G_{k-1}^{[k]}$.*

Proof. By induction on k. For $k = 1$ there is nothing to prove. We take $k \geq 1$ and assume that the result holds for k and for any nilpotent Lie group G and any discrete cocompact subgroup Λ.

We use the notation of the preceding sections. The group $\widetilde{G}$ is a nilpotent Lie group. We define

$$\widetilde{\Lambda} := \widetilde{G} \cap (\Lambda \times \Lambda) = \left\{ (\lambda', \lambda'') \in \Lambda \times \Lambda : \lambda'' {\lambda'}^{-1} \in \Lambda \cap G^{(1)} \right\}$$

and we note that $\widetilde{\Lambda}$ is cocompact in $\widetilde{G}$.

Claim. $G_*^{[k]} \cap \Lambda^{[k]}$ is cocompact in $G_*^{[k]}$.

Proof. Let $\{\mathbf{g}_n\}$ be a sequence in $G_*^{[k]}$. Consider the sequence $\{(\mathbf{g}_n, \mathbf{1})\}$ in $\widetilde{G}_{k-1}^{[k]}$. By the inductive hypothesis, $\widetilde{\Lambda}^{[k]} \cap \widetilde{G}_{k-1}^{[k]}$ is cocompact in $\widetilde{G}_{k-1}^{[k]}$. Therefore, for each integer n, there exists $(\boldsymbol{\lambda}'_n, \boldsymbol{\lambda}''_n) \in \widetilde{\Lambda}^{[k]} \cap \widetilde{G}_{k-1}^{[k]}$ and $(\mathbf{h}'_n, \mathbf{h}''_n) \in \widetilde{G}_{k-1}^{[k]}$ so that the sequence $\{(\mathbf{h}'_n, \mathbf{h}''_n)\}$ is bounded and for every n,

$$\mathbf{g}_n = \mathbf{h}'_n \boldsymbol{\lambda}'_n \text{ and } 1^{[k]} = \mathbf{h}''_n \boldsymbol{\lambda}''_n \ .$$

The sequence $\{\boldsymbol{\lambda}''_n\}$ is bounded; since Λ is discrete, this sequence takes only finitely many values. Let $\boldsymbol{\lambda} \in \Lambda^{[k]} \cap G_{k-1}^{[k]}$ be one of these values and let $E = \{n : \boldsymbol{\lambda}''_n = \boldsymbol{\lambda}\}$.

For $n \in E$, we have $(\boldsymbol{\lambda}, \boldsymbol{\lambda}) \in \widetilde{G}^{[k]}_{k-1}$. Thus $(\mathbf{h}'_n\boldsymbol{\lambda}, \mathbf{1}) \in \widetilde{G}^{[k]}_{k-1}$ and $\mathbf{h}'_n\boldsymbol{\lambda} \in G^{[k]}_*$. We have written the sequence $\{\mathbf{g}_n : n \in E\}$ as the product of the bounded sequence $\{\mathbf{h}'_n\boldsymbol{\lambda} : n \in E\}$ in $G^{[k]}_*$ with the sequence $\{\boldsymbol{\lambda}^{-1}\boldsymbol{\lambda}'_n : n \in E\}$ in $G^{[k]}_* \cap \Lambda^{[k]}$.

Since $\mathbb{N}$ is a finite union of sets E with this property, it follows that $G^{[k]}_* \cap \Lambda^{[k]}$ is cocompact in $G^{[k]}_*$. □

Claim. $G^{[k]}_\bullet \cap \Lambda^{[k]}$ is cocompact in $G^{[k]}_\bullet$.

Proof. Let $\{\mathbf{q}_n\}$ be a sequence in $G^{[k]}_\bullet$. By using Lemma 11.3 and the fact that λ is cocompact in G, for every n we can write $\mathbf{q}_n = h_n^{[k]}\lambda_n^{[k]}\mathbf{g}_n$, where $\{h_n\}$ is a bounded sequence in G, $\lambda_n \in \Lambda$ for every n, and $\mathbf{g}_n \in G^{[k]}_*$ for every n. We have that $\lambda_n^{[k]}\mathbf{g}_n{\lambda_n^{[k]}}^{-1} \in G^{[k]}_*$. Using the first claim, we write $\lambda_n^{[k]}\mathbf{g}_n{\lambda_n^{[k]}}^{-1} = \mathbf{v}_n\boldsymbol{\mu}_n$, where $\{\mathbf{v}_n\}$ is a bounded sequence in $G^{[k]}_*$ and $\boldsymbol{\mu}_n \in G^{[k]}_* \cap \Lambda^{[k]}$ for every n. The claim follows. □

The Lemma follows immediately from Equation 30 and the inductive hypothesis. □

As a corollary of the two claims we have:

Corollary 11.7. *$G^{[k]}_*(\Lambda^{[k]} \cap G^{[k]}_{k-1})$ and $G^{[k]}_\bullet(\Lambda^{[k]} \cap G^{[k]}_{k-1})$ are closed subgroups of $G^{[k]}_{k-1}$.*

11.3. **The measures $\mu^{[k]}$.** Here (X, μ, T) is a toral system of order ℓ for some integer ℓ. By Theorem 10.5, this system can be represented as an ℓ-step nilsystem $X = G/\Lambda$, where G is nilpotent Lie group, Λ is a cocompact subgroup, μ is the Haar measure of X and the transformation T is left translation by some fixed element of G which we also write as T. Recall that G is the subgroup of $\mathcal{G}(X)$ spanned by the connected component of the identity and T.

For every integer k, the group $\Lambda^{[k]} \cap G^{[k]}_{k-1}$ is cocompact in $G^{[k]}_{k-1}$ by Lemma 11.6 and we can define the nilmanifold

$$X_k := G^{[k]}_{k-1}/(\Lambda^{[k]} \cap G^{[k]}_{k-1}) \tag{32}$$

and let ν_k denote its Haar measure. The nilmanifold X_k is included in $X^{[k]} = G^{[k]}/\Lambda^{[k]}$ in the natural way.

For every $g \in G$ we have $g^{[k]} \in G^{[k]}_{k-1}$. It follows that, for every $x \in X$, X_k contains the *diagonal point* $(x, x, \dots, x)$ of $X^{[k]}$.

Lemma 11.8. *For every $k \geq 1$, the measure $\mu^{[k]}$ is the Haar measure of the nilmanifold X_k.*

Proof. The proof is by induction. The assertion is obvious for $k = 1$, because $X_1 = X \times X$ and $G^{[1]}_0 = G^{[1]} = G \times G$. We assume that it holds for some $k \geq 1$.

By Corollary 11.7 $G^{[k]}_\bullet(\Lambda^{[k]} \cap G^{[k]}_{k-1})$ is a closed subgroup of $G^{[k]}_{k-1}$ and we can define the space

$$Y_k := G^{[k]}_{k-1} \,/\, G^{[k]}_\bullet(\Lambda^{[k]} \cap G^{[k]}_{k-1}) \ .$$

Write $\phi_k \colon X_k \to Y_k$ for the the natural continuous surjection.

For $\mathbf{x} \in X_k$, the subset $G^{[k]}_\bullet \cdot \mathbf{x} := \{\mathbf{g} \cdot \mathbf{x} : \mathbf{g} \in G^{[k]}_\bullet\}$ of X_k is the inverse image of the point $\phi_k(\mathbf{x}) \in Y_k$ under ϕ_k and thus it is closed. So the action of $G^{[k]}_\bullet$ on X_k by left translations has closed orbits and we can identify Y_k with the quotient of X_k under this action.

Claim. The invariant σ-algebra $\mathcal{I}^{[k]}$ of $(X^{[k]}, \mu^{[k]}, T^{[k]})$ is equal up to $\mu^{[k]}$ null sets to the inverse image under ϕ_k of the Borel σ-algebra of Y_k.

Proof of the Claim. Let $\mathcal{B}$ be this inverse image. This σ-algebra consists in the Borel subsets of X_k which are invariant under translation by any element of $G_\bullet^{[k]}$. Since $T \in G$, $T^{[k]} \in G_\bullet^{[k]}$ and every set belonging to $\mathcal{B}$ is invariant under $T^{[k]}$ and thus belongs to $\mathcal{I}^{[k]}$.

On the other hand, as $G \subset \mathcal{G}(X)$, the measure $\mu^{[k+1]}$ is invariant under $\mathbf{g}$ for any $\mathbf{g} \in G_k^{[k+1]}$ by Corollary 5.4. In particular $\mu^{[k+1]}$ is invariant under $(1_k^{[k]}, \mathbf{h})$ for any $\mathbf{h} \in G_\bullet^{[k]}$. Proceeding exactly as for the implication (2) $\Longrightarrow$ (3) in the proof of Lemma 5.3, we have that every $\mathbf{h} \in G_\bullet^{[k]}$ acts trivially on $\mathcal{I}^{[k]}$ and we conclude that $\mathcal{I}^{[k]}$ is measurable with respect to $\phi^{-1}(\mathcal{B})$. The claim is proven. □

From Equation (30), it follows immediately that X_{k+1} consists in the pairs $(\mathbf{x}', \mathbf{x}'') \in X_k \times X_k$, with $\phi_k(\mathbf{x}') = \phi_k(\mathbf{x}'')$. Using the inductive hypothesis and the definition of the measure $\mu^{[k+1]}$, we get that this measure is concentrated on the nilmanifold X_{k+1}. By Lemma 5.2, this measure is invariant under the translation by any of the generators of $G_k^{[k+1]}$ and thus by translation by every element of this group. It is therefore the Haar measure of the nilmanifold X_{k+1} and the statement of the Lemma is proven for $k+1$. □

12. Arithmetic Progressions

We now use the tools assembled to study convergence along arithmetic progressions in order to obtain Theorem 1.1.

12.1. **Characteristic Factor for Arithmetic Progressions.** We first show that we can modify the original system and replace it by some factor so that convergence of the factor system implies convergence in the original system. This is based on the notion of a *characteristic factor* used by Furstenberg and Weiss in [FW96].

We can always assume that the system is ergodic by using, if necessary, ergodic decomposition.

Theorem 12.1. *Let (X, μ, T) be an ergodic system. Assume that $f_1, \dots, f_k$ are bounded functions on X with $\|f_j\|_\infty \leq 1$ for $j = 1, \dots, k$. Then*

$$\limsup_{N\to+\infty} \Big\| \frac{1}{N} \sum_{n=0}^{N-1} \Big(\prod_{j=1}^{k} f_j \circ T^{jn} \Big) \Big\|_{L^2(\mu)} \;\leq\; \min_{1\leq \ell \leq k} \Big(\ell \cdot |\!|\!| f_\ell |\!|\!|_k \Big) \ . \tag{33}$$

Proof. We proceed by induction. For $k = 1$ by the Ergodic Theorem,

$$\Big\| \frac{1}{N} \sum_{n=0}^{N-1} f_1 \circ T^n \Big\|_{L^2(\mu)} \to \Big| \int f_1 \, d\mu \Big| = |\!|\!| f_1 |\!|\!|_1 \ .$$

Let $k \geq 1$ and assume that the majorization (33) holds for k. Let $f_1, \dots, f_{k+1} \in L^\infty(\mu)$ with $\|f_j\|_\infty \leq 1$ for $j = 1, \dots, k+1$. Choose $\ell \in \{2, \dots, k+1\}$. (The case $\ell = 1$ is similar.) Write

$$\xi_n = \prod_{j=1}^{k+1} f_j \circ T^{jn} \ .$$

By the van der Corput Lemma (Lemma D.2),

$$\limsup_{N\to\infty}\Big\|\frac{1}{N}\sum_{n=1}^{N}\xi_n\Big\|^2_{L^2(\mu)} \leq \limsup_{H\to\infty}\frac{1}{H}\sum_{h=1}^{H}\Big(\limsup_{N\to\infty}\Big|\frac{1}{N}\sum_{n=1}^{N}\int \xi_{n+h}\cdot\xi_n\,d\mu\Big|\Big)\ .$$

Letting M denote the last lim sup, we need to show that $M \leq \ell^2 |\!|\!| f_\ell |\!|\!|_{k+1}^2$.

For any integer $h \geq 1$,

$$\begin{aligned}
&\Big|\frac{1}{N}\sum_{n=1}^{N}\int \xi_{n+h}\cdot\xi_n\,d\mu\Big| \\
&\qquad\qquad = \Big|\int (f_1\cdot f_1\circ T^h)\cdot\frac{1}{N}\sum_{n=1}^{N}\Big(\prod_{j=2}^{k+1}(f_j\cdot f_j\circ T^{jh})\circ T^{(j-1)n}\Big)\,d\mu\Big| \\
&\qquad\qquad \leq \|f_1\cdot f_1\circ T^h\|_{L^2(\mu)}\cdot\Big\|\frac{1}{N}\sum_{n=1}^{N}\Big(\prod_{j=2}^{k+1}(f_j\cdot f_j\circ T^{jh})\circ T^{(j-1)n}\Big)\Big\|_{L^2(\mu)}
\end{aligned}$$

and by the inductive assumption,

$$\limsup_{N\to\infty}\Big|\frac{1}{N}\sum_{n=1}^{N}\int \xi_{n+h}\cdot\xi_n\,d\mu\Big| \leq \ell\cdot|\!|\!| f_\ell\cdot f_\ell\circ T^{\ell h}|\!|\!|_k\ .$$

We get

$$\begin{aligned}
M &\leq \ell\cdot\limsup_{H\to\infty}\frac{1}{H}\sum_{h=1}^{H}|\!|\!| f_\ell\cdot f_\ell\circ T^{\ell h}|\!|\!|_k \leq \ell^2\cdot\limsup_{H\to\infty}\frac{1}{H}\sum_{h=1}^{H}|\!|\!| f_\ell\cdot f_\ell\circ T^{h}|\!|\!|_k \\
&\leq \ell^2\cdot\limsup_{H\to\infty}\Big(\frac{1}{H}\sum_{h=1}^{H}|\!|\!| f_\ell\cdot f_\ell\circ T^{h}|\!|\!|_k^{2^k}\Big)^{1/2^k}\ .
\end{aligned}$$

Define $F(\mathbf{x}) = \prod_{\epsilon\in V_k} f_\ell(x_\epsilon)$. The last average becomes

$$\frac{1}{H}\sum_{h=1}^{H}\int F\circ(T^{[k]})^h\cdot F\,d\mu^{[k]}$$

by definition of the seminorm $|\!|\!|\cdot|\!|\!|_k$. When $H\to+\infty$, this average converges to

$$\int \mathbb{E}(F\mid\mathcal{I}^{[k]})^2\,d\mu^{[k]} = \int F\otimes F\,d\mu^{[k+1]} = |\!|\!| f_\ell|\!|\!|_{k+1}^{2^{k+1}}$$

by definition of the seminorm $|\!|\!|\cdot|\!|\!|_{k+1}$, and the proof is complete. □

12.2. **Convergence for Arithmetic Progressions.** We prove Theorem 1.1.

Let f_j, $1\leq j\leq k$, be k bounded functions on X. By Theorem 12.1, the difference between the average (1) and the same average with f_j replaced by $\mathbb{E}(f_j|\mathcal{Z}_k)$ for $1\leq j\leq k$ tends to 0 in $L^2(X)$. Thus it suffices to prove Theorem 1.1 when all functions are measurable with respect to $\mathcal{Z}_k$. In particular, we can assume that the system $X = Z_k(X)$, that is, that X is a system of type k. Such a system is an inverse limit of translations on nilmanifolds by Theorem 10.3 and so it suffices to prove Theorem 1.1 for a translation $x\mapsto t\cdot x$ on a nilmanifold $X=\mathcal{G}/\Lambda$ endowed with its Haar measure. By density, it is also sufficient to prove the convergence when the functions $f_1,\ldots,f_k$ are continuous.

Several independent proofs already exist for the convergence of the averages (1) in this case (see Appendix A). Leibman [Lb02] uses Theorem B.3 applied to the

the translation by $s = (t, t^2, \dots, t^k)$ on the nilmanifold $X^k = \mathcal{G}^k/\Lambda^k$, and obtain the convergence everywhere. Ziegler ([Zie02a]) builds an explicit partition of X^k into invariant nilmanifolds and shows that almost every nilmanifold is ergodic and thus uniquely ergodic for the translation by s; the convergence almost everywhere follows.

13. Cubes

We are now ready to complete the proof of Theorem 1.2. As for the arithmetic progressions, we can assume that the system is ergodic. We first describe an appropriate characteristic factor.

Let (X, μ, T) be an ergodic system. Given an integer $k \geq 1$ and 2^k bounded functions f_ϵ, $\epsilon \in V_k$, on X, we study the convergence of the sequence of numerical averages:

$$(\mathcal{A}_k) \qquad \prod_{i=1}^{k} \frac{1}{N_i - M_i} \sum_{n \in [M_1, N_1) \times \cdots \times [M_k, N_k)} \int \prod_{\epsilon \in V_k} f_\epsilon \circ T^{\epsilon_1 n_1 + \cdots + \epsilon_k n_k} \, d\mu$$

and the convergence in $L^2(\mu)$ of the averages

$$(\mathcal{B}_k) \qquad \prod_{i=1}^{k} \frac{1}{N_i - M_i} \sum_{n \in [M_1, N_1) \times \cdots \times [M_k, N_k)} \prod_{\epsilon \in V_k^*} f_\epsilon \circ T^{\epsilon_1 n_1 + \cdots + \epsilon_k n_k}$$

when $N_1 - M_1, \dots, N_k - M_k$ tend to $+\infty$. We show:

Theorem 13.1. (1) *The averages $(\mathcal{A}_k)$ converge to*

$$\int_{X^{[k]}} \prod_{\epsilon \in V_k} f_\epsilon(x_\epsilon) \, d\mu^{[k]}(\mathbf{x}) \; . \tag{34}$$

(2) *The averages $(\mathcal{B}_k)$ converge in $L^2(\mu)$. The limit is the function*

$$x \mapsto \mathbb{E}\Big(\bigotimes_{\epsilon \in V_k^*} f_\epsilon \big| \mathcal{J}^{[k]^*} \Big)(x) \tag{35}$$

where we have identified the σ-algebra $\mathcal{J}^{[k]^*}$ with the factor $\mathcal{Z}_{k-1}(X)$ (see Section 4.2).

13.1. The case of a toral system.

Lemma 13.2. *The results of Theorem 13.1 hold when X is a toral system of order ℓ for some integer $\ell \geq 1$.*

Proof. Let $k \geq 1$ be an integer. For this proof we let T_i, $1 \leq i \leq k$, denote the transformation $T_{\alpha_i}^{[k]}$ of $X^{[k]}$, where $\alpha_1, \dots, \alpha_k$ are the sides of V_k not containing $\mathbf{0}$. We recall that the group of transformations $\mathcal{T}_*^{[k]}$ of $X^{[k]}$ is spanned by $\{T_i : 1 \leq i \leq k\}$ and that the group $\mathcal{T}_{k-1}^{[k]}$ is spanned by $\mathcal{T}_*^{[k]}$ and $T^{[k]}$.

We assume that X is a toral system of order ℓ. By Lemma 11.8, $\mu^{[k]}$ is the Haar measure of the nilmanifold $X_k = G_{k-1}^{[k]}/(\Lambda^{[k]} \cap G_{k-1}^{[k]})$ introduced in Subsection 11.3. By Corollary 3.5, $\mu^{[k]}$ is ergodic under the group $\mathcal{T}_{k-1}^{[k]}$. As the transformations T_i, $1 \leq i \leq k$, and $T^{[k]}$ of X_k are translations by commuting elements of $G_{k-1}^{[k]}$, it follows from Theorem B.2 that X_k is uniquely ergodic for the action of $\mathcal{T}_{k-1}^{[k]}$.

Let f_ϵ, $\epsilon \in V_k$, be 2^k continuous functions on X. For integers $n, n_1, \dots, n_k$ the transformation $T^nT_1^{n_1}\dots T_k^{n_k}$ is given by

$$\big(T^nT_1^{n_1}\dots T_k^{n_k}\mathbf{x}\big)_\epsilon = T^{n+\epsilon_1 n_1+\cdots+\epsilon_k n_k}x_\epsilon \text{ for every } \epsilon \in V_k\ .$$

Therefore, by unique ergodicity, when $N_1 - M_1, \dots, N_k - M_k$ and N tend to $+\infty$, the functions

$$\mathbf{x} \qquad \mapsto \qquad \prod_{i=1}^k \frac{1}{N_i - M_i} \sum_{\substack{M_1\leq n_1<N_1,\\ \dots,\\ M_k\leq n_k<N_k}} \frac{1}{N}\sum_{n=0}^{N-1}\prod_{\epsilon\in V_k} f_\epsilon(T^{n+\epsilon_1 n_1+\cdots+\epsilon_k n_k}x_\epsilon)$$

converge uniformly on X_k to the constant given by the integral

$$\int_{X_k}\prod_{\epsilon\in V_k} f_\epsilon(x_\epsilon)\,d\mu^{[k]}(\mathbf{x})\ . \tag{36}$$

Thus, they converge uniformly to this constant on the 'diagonal' subset of X_k (the subset consisting in points $\mathbf{x} = (x, x, \dots, x)$). This means that the averages

$$x \mapsto \prod_{i=1}^k \frac{1}{N_i - M_i} \sum_{\substack{M_1\leq n_1<N_1,\\ \dots,\\ M_k\leq n_k<N_k}} \frac{1}{N}\sum_{n=0}^{N-1}\prod_{\epsilon\in V_k} f_\epsilon(T^{n+\epsilon_1 n_1+\cdots+\epsilon_k n_k}x)$$

converge uniformly on X to this constant. Taking the integral we get that the averages $(\mathcal{A}_k)$ converge to this constant. Part (1) of Theorem 13.1 holds for a toral system when the functions f_ϵ are continuous. The case of arbitrary bounded functions follows by density.

Let f_ϵ, $\epsilon \in V_k$, be $2^k - 1$ continuous functions on X. By Theorem B.3 the averages

$$\prod_{i=1}^k \frac{1}{N_i - M_i} \sum_{\substack{M_1\leq n_1<N_1,\\ \dots,\\ M_k\leq n_k<N_k}} \prod_{\epsilon\in V_k^*} f_\epsilon(T^{\epsilon_1 n_1+\cdots+\epsilon_k n_k}x_\epsilon)$$

converge for every $\mathbf{x} \in X_k$ and in particular for every diagonal point $\mathbf{x} = (x, x, \dots, x)$. Therefore the averages $(\mathcal{B}_k)$ converge for every $x \in X$. Let $\phi(x)$ be the limit. By Part (1), for every bounded function $f_{\mathbf{0}}$ on X,

$$\begin{aligned}\int_X f_{\mathbf{0}}(x)\phi(x)\,d\mu(x) &= \int_{X^{[k]}} f_{\mathbf{0}}(x_{\mathbf{0}})\prod_{\epsilon\in V_k^*} f_\epsilon(x_\epsilon)\,d\mu^{[k]}(\mathbf{x})\\ &= \int_X f_{\mathbf{0}}(x)\mathbb{E}\Big(\bigotimes_{\epsilon\in V_k^*} f_\epsilon \mid \mathcal{J}^{[k]^*}\Big)(x)\,d\mu(x)\end{aligned}$$

by Lemma 4.2, under the identification of the σ-algebras $\mathcal{J}^{[k]^*}$ and $\mathcal{Z}_{k-1}(X)$. It follows that the function ϕ is equal to the conditional expectation (35). By density, the same result holds for arbitrary bounded functions on X. Part (2) of Theorem 10.3 is proven for a toral system. □

Corollary 13.3. *The results of Theorem 13.1 hold when X is a system of level ℓ for some $\ell \geq 1$.*

Proof. Let X be a system of order ℓ. By Theorem 10.3, X can be represented as an inverse limit of a sequence of toral systems of order ℓ. Let Y be one of these systems and $p\colon X \to Y$ the corresponding factor map.

Let g_ϵ, $\epsilon \in V_k$, be bounded functions on Y. $p^{[k]}\colon X^{[k]} \to Y^{[k]}$ is a factor map by Lemma 4.5 and thus it follows from Lemma 13.2 that Part (1) of Theorem 13.1 holds for X and the functions $f_\epsilon = g_\epsilon \circ p$.

By Proposition 4.6, $p^{-1}(\mathcal{Z}_k(Y)) = Z_k(X) \cap p^{-1}(\mathcal{Y})$ and Part (2) of Theorem 13.1 also follows from Lemma 13.2 for the functions $f_\epsilon = g_\epsilon \circ p$.

By density the same results hold for every bounded functions on X. □

13.2. **The general case.** In the proof, we consider the averages $(\mathcal{A}_k)$ with $f_{\mathbf{0}} = 1$ separately. That is, the averages

$$(\mathcal{C}_k) \qquad \prod_{i=1}^{k} \frac{1}{N_i - M_i} \sum_{n \in [M_1,N_1)\times\cdots\times[M_k,N_k)} \int \prod_{\epsilon \in V_k^*} f_\epsilon \circ T^{\epsilon_1 n_1 + \cdots + \epsilon_k n_k} \, d\mu \ .$$

We prove Theorem 13.1 by induction. For $k = 1$, the averages are

$$\frac{1}{N-M} \sum_{n=M}^{N-1} \int f_0 \cdot f_1 \circ T^n \, d\mu$$

and

$$\frac{1}{N-M} \sum_{n=M}^{N-1} f_1 \circ T^n \, d\mu$$

where f_0 and f_1 are bounded functions on X. Since $\mu^{[1]} = \mu \times \mu$, the results are obvious.

Henceforth, fix an integer $k > 1$ and assume that the two statements of Theorem 13.1 hold for $k-1$.

13.2.1. *The averages $(\mathcal{C}_k)$.*

Lemma 13.4. *Let g_η, $\eta \in V_{k-1}$, be bounded functions on X. Then the* $\limsup$ *for $N_1 - M_1, \dots, N_{k-1} - M_{k-1} \to +\infty$ and $N - M \to +\infty$ of*

$$(37) \qquad \prod_{i=1}^{k-1} \frac{1}{(N_i - M_i)} \sum_{\substack{M_1 \le n_1 < N_1, \\ \dots, \\ M_{k-1} \le n_{k-1} < N_{k-1}}} \int \Big| \frac{1}{N-M} \sum_{p=M}^{N-1} \prod_{\eta \in V_{k-1}^*} g_\eta \circ T^{\eta \cdot n - p} \Big|^2 \, d\mu$$

is less than or equal to

$$(38) \qquad \int \Big| \mathbb{E}\Big(\bigotimes_{\eta \in V_{k-1}} g_\eta \mid \mathcal{I}^{[k-1]} \Big) \Big|^2 \, d\mu^{[k-1]} \ .$$

Proof. Without loss of generality, we can assume that $|g_\eta| \le 1$ for each $\eta \in V_{k-1}$. Fix an integer $H > 0$. By the finite van der Corput Lemma (Lemma D.2), for each $n = (n_1, \dots, n_{k-1})$ the integral in (37) is bounded by

$$\frac{N-M+H-1}{(N-M)H} + \frac{N-M+H-1}{N-M} \sum_{h=1}^{H-1} 2 \frac{H-h}{H^2} \, \frac{N-h}{N} \int \prod_{\eta \in V_{k-1}} (g_\eta \cdot g_\eta \circ T^h) \circ T^{\eta \cdot n} \, d\mu \ .$$

Thus the lim sup of expression (37) is bounded by

$$\frac{1}{H}+\sum_{h=1}^{H-1}2\frac{H-h}{H^2}\limsup_{\substack{N_1-M_1\to\infty,\\ \dots,\\ N_{k-1}-M_{k-1}\to\infty}}\prod_{i=1}^{k}\frac{1}{N_i-M_i}\sum_{\substack{M_1\leq n_1<N_1,\\ \dots,\\ M_k\leq n_k<N_k}}\int\prod_{\eta\in V_{k-1}}(g_\eta\cdot g_\eta\circ T^h)\circ T^{\eta\cdot n}\,d\mu\ .$$

By the inductive hypothesis Theorem 13.1 holds for $k-1$ and this expression is equal to

$$\frac{1}{H}+\sum_{h=1}^{H-1}2\frac{H-h}{H^2}\int_{X^{[k-1]}}\bigotimes_{\eta\in V_{k-1}}(g_\eta\cdot g_\eta\circ T^h)\,d\mu^{[k-1]}\ .$$

Taking the limit when $H\to\infty$, we get the result. □

Lemma 13.5. *The factor Z_{k-2} is characteristic for the convergence of the averages $(\mathcal{C}_k)$. In other words, if for some $\epsilon\in V_k^*$ we have $\mathbb{E}(f_\epsilon\mid\mathcal{Z}_{k-2})=0$, then these averages converge to* 0.

Proof. Without loss of generality, we can assume that $|f_\epsilon|\leq 1$ for every $\epsilon\in V_k^*$.

First assume that $\mathbb{E}(f_\epsilon\mid\mathcal{Z}_{k-2})=0$ for some $\epsilon\in V_k^*$ with $\epsilon_1=0$. Define g_η, $\eta\in V_k^*$ by $g_\eta=f_{0\eta}$ and $g_{\mathbf{0}}=1$ and h_η, $\eta\in V_k$, by $h_\eta=f_{1\eta}$. Then the average $(\mathcal{C}_k)$ can be written

$$\prod_{i=1}^{k-1}\frac{1}{N_i-M_i}\sum_{M_1\leq n_1<N_1,\dots,M_{k-1}\leq n_{k-1}<M_{k-1}}$$
$$\int\Big(\prod_{\eta\in V_{k-1}}h_\eta\circ T^{\eta\cdot n+p}\Big)\cdot\Big(\frac{1}{N_k-M_k}\sum_{p=M_k}^{N_k-1}\prod_{\eta\in V_{k-1}}g_\eta\circ T^{\eta\cdot n-p}\Big)\,d\mu\ .$$

By the Cauchy-Schwartz inequality, the square of this average is bounded by (37). By Lemma 13.4, the lim sup of the square of this average is bounded by (38).

The measure $\mu^{[k-1]^*}$ is relatively independent with respect to $\mathcal{Z}_{k-2}^{[k-1]^*}$ and at least one of the functions g_η, $\eta\in V_k^*$, has zero conditional expectation with respect to $\mathcal{Z}_{k-2}$. Therefore $\mathbb{E}(\otimes_{\eta\in V_k^*}g_\eta\mid\mathcal{Z}_{k-2}^{[k]^*})=0$. But by Part (2) of Proposition 4.9, the σ-algebra $\mathcal{I}^{[k-1]^*}$ is measurable with respect to $\mathcal{Z}_{k-2}^{[k-1]^*}$. Thus $\mathbb{E}(\bigotimes_{\eta\in V_k^*}g_\eta\mid\mathcal{I}^{[k-1]^*})=0$ and also $\mathbb{E}(\bigotimes_{\eta\in V_k}g_\eta\mid\mathcal{I}^{[k-1]})=0$. The bound (38) is equal to 0 and the averages $(\mathcal{C}_k)$ converge to 0.

By permuting the coordinates, we get the same result when $\mathbb{E}(f_\epsilon\mid\mathcal{Z}_{k-2})=0$ for some ϵ with $\epsilon_j=0$ for some j, that is, for some $\epsilon\neq 11\dots1$.

Finally assume that $\mathbb{E}(f_{11\dots1}\mid\mathcal{Z}_{k-2})=0$. By the preceding proof, the lim sup of the absolute value of the averages $(\mathcal{C}_k)$ remains unchanged when we substitute $\mathbb{E}(f_\epsilon\mid\mathcal{Z}_{k-2})$ for f_ϵ, for every $\epsilon\neq 11\dots1$. Without loss of generality, we can therefore assume that for each $11\dots1\neq\epsilon\in V_k^*$, the function f_ϵ is measurable with respect to $\mathcal{Z}_{k-2}$. But in this case the integral in the average $(\mathcal{C}_k)$ is equal to 0 and the result is obvious. □

Corollary 13.6. *The averages $(\mathcal{C}_k)$ converge to*

$$\int_{X^{[k]^*}}\prod_{\epsilon\in V_k^*}f_\epsilon(x_\epsilon)\,d\mu^{[k]^*}(\tilde{x})\ . \tag{39}$$

Proof. By Lemma 13.5 the difference between the averages $(\mathcal{C}_k)$ and the same averages with the functions $\mathbb{E}(f_\epsilon \mid \mathcal{Z}_{k-2})$ substituted for f_ϵ converges to zero. As the natural projection $X^{[k]^*} \to Z_{k-2}^{[k]^*}$ is a factor map, the announced result follows immediately from Corollary 13.3. □

13.2.2. *The averages* $(\mathcal{A}_k)$ *and* $(\mathcal{B}_k)$.

Lemma 13.7. *The factor Z_{k-1} of X is characteristic for the convergence in $L^2(\mu)$ of the averages $(\mathcal{B}_k)$*

Proof. Assume that for some $\epsilon \in V_k^*$ we have $\mathbb{E}(f_\epsilon \mid \mathcal{Z}_{k-1}) = 0$. By Proposition 4.9 the measure $\mu^{[k]^*}$ is conditionally independent with respect to $\mathcal{Z}_{k-1}$ and thus $\mathbb{E}(\bigotimes_{\epsilon\in V_k^*} f_\epsilon \mid \mathcal{Z}_{k-1}^{[k]^*}) = 0$. Moreover by Proposition 4.9 the σ-algebra $\mathcal{J}^{[k]^*}$ is measurable with respect to $\mathcal{Z}_{k-1}^{[k]^*}$ and thus

$$\mathbb{E}(\bigotimes_{\epsilon\in V_k^*} f_\epsilon \mid \mathcal{J}^{[k]^*}) = 0 \ .$$

For $n = (n_1, \dots, n_k) \in \mathbb{Z}^k$, set

$$g_n = \prod_{\epsilon\in V_k^*} f_\epsilon \circ T^{\epsilon\cdot n}$$

and we have to show that

$$\prod_{i=1}^k \frac{1}{N_i - M_i} \sum_{\substack{M_1\le n_1<N_1,\\ \dots,\\ M_k\le n_k<N_k}} g_n \to 0 \text{ in } L^2(\mu)$$

as $N_1 - M_1, \dots, N_k - M_k \to +\infty$. For $h = (h_1, \dots, h_k) \in \mathbb{Z}^k$, by Corollary 13.6

$$\prod_{i=1}^k \frac{1}{N_i - M_i} \sum_{\substack{M_1\le n_1<N_1,\\ \dots,\\ M_k\le n_k<N_k}} \int g_{n+h} \cdot g_n \, d\mu \to \gamma_h \ ,$$

when $N_1 - M_1$, $\dots$, $N_k - M_k$ tend to $+\infty$, where

$$\gamma_h = \int_{X^{[k]^*}} \bigotimes_{\epsilon\in V_k^*} (f_\epsilon \cdot f_\epsilon \circ T^{\epsilon\cdot h}) \, d\mu^{[k]^*} \ .$$

When $H \to \infty$, we have

$$\sum_{\substack{-H\le h_1\le H,\\ \vdots\\ -H\le h_k\le H}} \prod_{i=1}^k \frac{H-|h_i|}{H^2} \gamma_h \ \to \ \Big\|\mathbb{E}\Big(\bigotimes_{\epsilon\in V_k^*} f_\epsilon \big| \mathcal{J}^{[k]^*}\Big)\Big\|^2_{L^2(\mu^{[k]^*})} = 0$$

and the statement of the Lemma follows from the multidimensional van der Corput Lemma (Lemma D.3). □

As for arithmetic progressions, we combine the fact that the factors Z_k are characteristic with the proof of convergence for nilsystems to prove Theorem 13.1:

Proof of Theorem 13.1. We study the convergence of the averages $(\mathcal{A}_k)$ and $(\mathcal{B}_k)$ for an arbitrary ergodic system.

Recall that the natural projections $X^{[k]} \to Z_{k-1}^{[k]}$ and $X^{[k]^*} \to Z_{k-1}^{[k]^*}$ are factor maps and that the σ-algebra $\mathcal{J}^{[k]^*}$ is measurable with respect to $\mathcal{Z}_{k-1}^{[k]^*}$ (Proposition 4.9). Then Theorem 13.1 follows immediately from Corollary 13.3 and Lemma 13.7. □

13.3. **Proof of Theorem 1.3.** Using ergodic decomposition, we restrict to the case where the system X is ergodic. By part (1) of Theorem 13.1, applied to $f_\epsilon = \mathbf{1}_A$ for every $\epsilon \in V_k$, the averages appearing in the statement of Theorem 1.3 converge to

$$\int_{X^{[k]}} \prod_{\epsilon \in V_k} \mathbf{1}_A(x_\epsilon)\, d\mu^{[k]}(\mathbf{x}) = |\!|\!| \mathbf{1}_A |\!|\!|_k^{2^k} \ .$$

By part (3) of Lemma 3.9 we have $|\!|\!| \mathbf{1}_A |\!|\!|_k \geq |\!|\!| \mathbf{1}_A |\!|\!|_1 = \mu(A)$ and the result follows. □

13.4. **Proof of Theorem 1.5.** Theorem 1.3 has the following corollary:

Corollary 13.8. *Let $(X, \mathcal{B}, \mu, T)$ be an invertible measure preserving probability system, let $A \in \mathcal{B}$ and let $k \geq 1$ be an integer. Then for any $c > 0$, the set of $n \in \mathbb{Z}^k$ so that*

$$\mu\Big(\bigcap_{\epsilon \in V_k} T^{\epsilon \cdot n} A\Big) \geq \mu(A)^{2^k} - c$$

is syndetic.

Proof. Let E be the subset of $\mathbb{Z}^k$ appearing in Theorem 1.3. If E is not syndetic, there exist intervals $[M_{1_i}, N_{1_i}), [M_{2_i}, N_{2_i}), \dots, [M_{k_i}, N_{k_i})$ with the lengths of the intervals tending to $+\infty$ so that

$$E \cap \big([M_{1_i}, N_{1_i}) \times [M_{2_i}, N_{2_i}) \times \ldots \times [M_{k_i}, N_{k_i})\big) = \emptyset \ .$$

Taking averages along these k dimensional cubes in Theorem 1.3, we get a contradiction. □

Theorem 1.5 follows by combining Furstenberg's correspondence principle and Corollary 13.8.

Appendix A. Groups

A.1. **Polish groups.** We summarize the main results we need (see Chapter 1 of [BK96]):

Theorem A.1. *Let G and H be Polish groups and let $p\colon G \to H$ be a group homomorphism that is continuous and onto. Then p is an open map. Moreover, p admits a Borel cross section, that is, a Borel map $s\colon H \to G$ with $p \circ s = \mathrm{Id}$.*

Let G, H and p be as above and let the quotient $G/\ker(p)$ be endowed with the quotient distance. It follows from the Theorem the natural group isomorphism $G/\ker(p) \to H$ is a homeomorphism.

Corollary A.2. *Let H be a closed normal subgroup of the Polish group G. If H and G/H are locally compact, then G is locally compact. If H and G/H are compact, then G is compact.*

We often build groups by a skew product construction and so present it here. Let G be a Polish group and let (X, μ) be a probability space. A measure preserving action on G on X is a measurable map $(g, x) \mapsto g \cdot x$ of $G \times X$ to X so that

(1) For every $g \in G$, the map $x \mapsto g \cdot x$ is a measure preserving bijection from X onto itself.
(2) For every $g, h \in G$, $gh \cdot x = g \cdot (h \cdot x)$ almost everywhere.

Let U be a compact abelian group, written additively. We recall that $\mathcal{C}(X, U)$ denotes the additive group of measurable maps from X to U. Endowed with the topology of convergence in probability, it is an abelian Polish group. For $g \in G$ and $f \in \mathcal{C}(X, U)$ we write $S_{g,f}$ for the measure preserving transformation of $(X \times U, \mu \times m_U)$ given by

$$S_{g,f}(x, u) = \big(g \cdot x, u + f(x)\big) \ .$$

These transformations form a group, called the *skew product* of G and written $G \ltimes \mathcal{C}(X, U)$. Endowed with the topology of convergence in probability, it is a Polish group. A sequence $\{S_{g_n, f_n}\}$ converges to $S_{g,f}$ in $G \ltimes \mathcal{C}(X, U)$ if and only if g_n converges to g in G and f_n converges to f in $\mathcal{C}(X, U)$.

The map $p \colon S_{g,f} \mapsto g$ is a continuous group homomorphism from $G \ltimes \mathcal{C}(X, U)$ onto G and thus is an open map.

A.2. **Lie groups.** We call a locally compact group a *Lie group* when it can be given the analytic structure of a Lie Group, although we never use the analytic structure. From the characterization of Lie groups in [MZ55] it can be deduced:

Lemma A.3. *Let G be a locally compact group and H a closed normal subgroup. If H and G/H are Lie groups then G is a Lie group.*

A.3. **Nilpotent Lie groups.** Let G be a Polish or locally compact group. For $g, h \in G$, we write $[g; h]$ for the commutator $g^{-1}h^{-1}gh$ of g and h. If A, B are subsets of G, we write $[A; B]$ for the *closed* subgroup of G spanned by $\{[a; b] : a \in A, b \in B\}$. The subgroups $G^{(j)}$, $j \geq 0$, of G are defined by $G^{(0)} = G$ and $G^{(j+1)} = [G; G^{(j)}]$ for $j \geq 0$. We say that G is *k-step nilpotent* if $G^{(k)}$ is the trivial subgroup $\{1\}$ of G.

(This definition of nilpotency is stronger than the purely algebraic definition, but the two definitions coincide for Lie groups.)

Appendix B. Nilmanifolds

Let G be a k-step nilpotent Lie group and Λ a discrete cocompact subgroup. The compact manifold $X = G/\Lambda$ is called a *k-step nilmanifold.* The group G acts on X by left translations and we write $(g, x) \mapsto g \cdot x$ for this action. There exists a unique probability measure μ on X invariant under this action; it is called *the Haar measure of X*. The fundamental properties of nilmanifolds were established by Malcev [Ma51]. We use the following properties of the commutator:

Lemma B.1. *Let G be a nilpotent Lie group and Λ a discrete cocompact subgroup. Then:*

(1) *The groups $G^{(j)}$, $j \geq 1$, are equal to the algebraic subgroups of iterated commutators. This means that for $j \geq 1$ the group $G^{(j)}$ is algebraically spanned by $\{[g; h] : g \in G,\ h \in G^{(j-1)}\}$.*
(2) *For every $j \geq 1$, the subgroup $G^{(j)}\Lambda$ of G is closed in G.*

Let $X = G/\Lambda$ be a k-step nilmanifold with Haar measure μ, let $t \in G$ and $T\colon X \to X$ the transformation $x \mapsto t \cdot x$. Then the system (X, μ, T) is called *a k-step nilsystem.*

The dynamical properties of nilsystems were studied by Auslander, Green and Hahn [AGH63], Parry ([P69], [P70]), Lesigne [L91] and Leibman [Lb02], between others.

Theorem B.2. *Let $X = G/\Lambda$ be a nilmanifold with Haar measure μ and let $t_1, \dots, t_\ell$ be commuting elements of G. If the group spanned by the translations by $t_1, \dots, t_\ell$ acts ergodically on (X, μ), then X is uniquely ergodic for this group.*

This result was shown by Parry [P69] in the case of a single translation, by using methods of [F61]. A similar proof for the general case can be found in [Lb02].

Theorem B.3. *Let $X = G/\Lambda$ be a nilmanifold and let $t_1, \dots, t_\ell$ be commuting elements of G. Then for any continuous function f on X the averages*

$$\prod_{i=1}^{k} \frac{1}{N_i - M_i} \sum_{\substack{M_1 \le n_1 < N_1 \\ \dots \\ M_k \le n_k < N_k}} f\bigl(t_1^{n_1} \dots t_k^{n_k} x\bigr)$$

converge everywhere on X when $N_1 - M_1, \dots, N_k - M_k$ tend to infinity.

This theorem can be viewed as a special case of the general results of M. Ratner and N. Shah (see [Ra91] and [Sh96]). A proof of this result is given in [L91] for a single transformation, under the additional hypothesis that the group G is connected. The preprint [Lb02] contains a similar proof for the general case. We do not reproduce it here, but indicate the different steps. By distality, for every $x \in X$, its closed orbit

$$Y_x = \overline{\{t_1^{n_1} \dots t_k^{n_k} x : (n_1, \dots, n_k) \in \mathbb{Z}^k\}}$$

is minimal for the the action spanned by the translations by $t_1, \dots, t_k$. The crucial point is that Y can be given the structure of a nilmanifold. By [P69], a minimal nilmanifold is uniquely ergodic, and the result follows.

We notice that in Theorem B.3 the "cubes" $[M_1, N_1) \times \dots \times [M_k, N_k)$ can be replaced by an arbitrary Følner sequence of subsets of $\mathbb{Z}^k$.

Appendix C. Cocycles

C.1. **Cocycles and extensions.** Let (X, μ, T) be a system and U a compact abelian group. We generally assume here that U is written with additive notation. (The changes needed when multiplicative notation is used are obvious.) A *cocycle with values in U* is a measurable map $\rho\colon X \to U$. We let $\mathcal{C}(X, U)$ denote the family of U-valued cocycles on X and we write $\mathcal{C}(X)$ instead of $\mathcal{C}(X, \mathbb{T})$. $\mathcal{C}(X, U)$ is endowed with pointwise addition and the topology of convergence in probability. It is a Polish group.

The *extension of (X, μ, T) by U associated to the cocycle $\rho \in \mathcal{C}(X)$* is the system $(X \times U, \mu \times m_U, T_\rho)$, where $T_\rho\colon X \times U \to X \times U$ is given by

$$T_\rho(x, u) = \bigl(Tx, u + \rho(x)\bigr)\ .$$

If $(X \times U, \mu \times m_U, T_\rho)$ is ergodic, we say that the cocycle ρ is *ergodic*. If moreover $(X \times U, \mu \times m_U, T_\rho)$ has the same Kronecker factor as X, we say that ρ is *weakly mixing.*

The factor map $(x, u) \mapsto x$ is called the *natural projection.* For $v \in U$, we also let v denote for the measure preserving transformation of $X \times U$ given by

$$v \cdot (x, u) = (x, v + u) .$$

A transformation of this type is called a *vertical rotation* or in case of ambiguity, a vertical rotation above X. We continuously identify the group U with the group of vertical rotations. The vertical rotations commute with T_ρ and preserve the natural projection on X. When ρ is ergodic, they are exactly characterized by these properties.

C.2. **Cocycles and coboundaries.** For $\rho \in \mathcal{C}(X, U)$, the *coboundary* of ρ is the cocycle $\rho \circ T - \rho$ and when there is no ambiguity, we write it $\partial\rho$. Let $\partial\mathcal{C}(X)$ denote the subgroup of $\mathcal{C}(X)$ consisting of coboundaries.

Assume that X is ergodic. Then a cocycle $\rho \in \mathcal{C}(X, U)$ is ergodic if and only if there exists no nontrivial character $\chi \in \widehat{U}$ so that the cocycle $\chi \circ \rho \in \mathcal{C}(X)$ is a coboundary.

The following result is found in Moore and Schmidt [MS80]:

Lemma C.1. *Let (X, μ, T) be a system, U a compact abelian group and $\rho \in \mathcal{C}(X, U)$. Then ρ is a coboundary if and only if for every $\chi \in \widehat{U}$, the cocycle $\chi \circ \rho \colon X \to \mathbb{T}$ is a coboundary.*

Two cocycles $\rho, \rho' \in \mathcal{C}(X, U)$ are said to be *cohomologous* if $\rho - \rho'$ is a coboundary. In this case, the extensions they define are isomorphic (i.e., there is an isomorphism between these two systems which preserves the natural projections).

Lemma C.2. *Let (X, μ, T) and (Y, ν, S) be ergodic systems, U a compact abelian group, $\rho \colon X \to U$ an ergodic cocycle and W the extension of X by U associated to ρ. Assume that W and Y are factors of the same ergodic system K and let L and M be the factors of K associated to the invariant sub σ-algebras $\mathcal{L} = \mathcal{X} \vee \mathcal{Y}$ and $\mathcal{M} = \mathcal{W} \vee \mathcal{Y}$, respectively. Then M is an extension of L by a closed subgroup V of U.*

Let $\gamma \in \widehat{U}$ and consider γ as taking values in $\mathcal{S}^1$. Define a function f_γ on W by $f_\gamma(x, u) = \gamma(u)$. If $\mathbb{E}(f_\gamma \mid \mathcal{L}) \neq 0$, then f_γ is measurable with respect to $\mathcal{L}$ and $\gamma \in V^\perp$.

This Lemma is essentially a reformulation of more or less classical results and similar Lemmas can be found, in particular, in Furstenberg and Weiss [FW96]. We only give an outline of the proof.

Proof. The system L can be represented as an ergodic joining λ of (X, μ, T) and (Y, ν, S). In the same way, M can be represented as an ergodic joining τ of W and Y. τ is a measure on $W \times Y = X \times Y \times U$ and the projection of τ on $X \times Y$ is λ. Moreover, τ is invariant under the transformation of $(X \times Y) \times U$ associated to the cocycle $\sigma \colon (x, y) \mapsto \rho(x)$ of the ergodic system $(X \times Y, \lambda, T \times S)$.

Therefore τ is an ergodic component of the extension of this system by U, defined by the cocycle σ. Thus it is an extension of this system by a closed subgroup V of U, the *Mackey group* of σ in the terminology of Furstenberg and Weiss [FW96]. For $\gamma \in \widehat{U}$, we have $\gamma \in V^\perp$ if and only if $\gamma \circ \sigma$ is a coboundary of the system $(X \times Y, \lambda, T \times S)$. That is, if and only if $\gamma \circ \rho$ is a coboundary of L.

Let $\gamma \in \widehat{U}$ and assume that $\mathbb{E}(f_\gamma \mid \mathcal{L}) \neq 0$. We have

$$f_\gamma(T_\rho(x,u)) = \gamma(\rho(x)) \cdot f_\gamma(x,u)$$

and moreover the map $(x,y,u) \mapsto \gamma(\rho(x))$ is measurable with respect to $\mathcal{L}$. Thus

$$\mathbb{E}(f_\gamma \mid \mathcal{L}) \circ T = \mathbb{E}(f_\gamma \circ T_\rho \mid \mathcal{L}) = \gamma \circ \rho \cdot \mathbb{E}(f_\gamma \mid \mathcal{L}) \ .$$

The function $\mathbb{E}(f_\gamma \mid \mathcal{L}) \cdot \overline{f_\gamma}$ is invariant on M and thus is constant by ergodicity. Therefore f_γ is measurable with respect to $\mathcal{L}$ and $\gamma \circ \rho$ is a coboundary on $\mathcal{L}$. By the first part, $\gamma \in V^\perp$.

□

C.3. **Measurability properties.** Let X be a system and U a compact abelian group. Then the coboundaries form a subgroup of $\mathcal{C}(X,U)$, which is Borel because it it the range of the continuous group homomorphism $\partial\colon \rho \mapsto \rho \circ T - \rho$ from the Polish group $\mathcal{C}(X,U)$ to itself ([BK96]).

Lemma C.3. *Let (X,μ,T) be (non-ergodic) system, (Y,ν) a (standard) probability space, and $y \mapsto \mu_y$ a weakly measurable map from Y to the space of probability measures on X. Assume that*

- *For every $y \in Y$, the measure μ_y is invariant under T.*
- *$\mu = \int_Y \mu_y \, d\nu(y)$.*

Let (Ω,P) be a (standard) probability space and let $\omega \mapsto \rho_\omega$ be a measurable map from Ω to $\mathcal{C}(X,\mathcal{S}^1)$. Then:

(1) *The set*

$$A = \big\{(\omega,y) \in \Omega \times Y : \rho_\omega \textit{ is a coboundary of } (X,\mu_y,T)\big\}$$

is a measurable subset of $\Omega \times Y$.

(2) *For $\omega \in \Omega$, ρ_ω is a coboundary of (X,μ,T) if and only if the set*

$$A_\omega = \{y \in Y : (\omega,y) \in A\}$$

satisfies $\nu(A_\omega) = 1$.

A cocycle $\rho \in \mathcal{C}(X,\mathcal{S}^1)$ is a map from X to $\mathcal{S}^1$ which is defined only μ-almost everywhere. This makes the definition of the set A in the Lemma appear 'problematic' and so we begin with an explanation.

We recall that $\mathcal{C}(X,\mathcal{S}^1)$ is endowed with the topology of convergence in probability and this topology coincides with the topology of L^1. By a classical result (see for example [Va70], p. 65) there exists a map $R\colon \Omega \times X \to \mathcal{S}^1$, defined everywhere and measurable, such that for every $\omega \in \Omega$, $\rho_\omega(x) = R(\omega,x)$ for μ-almost every x. In the statement above and in the proof below we write $\rho_\omega(x)$ instead of the more precise but heavier notation $R(\omega,x)$.

Proof. (1) For $\omega \in \Omega$ and an integer $n \geq 0$, write

$$\rho_\omega^{(n)}(x) = \rho_\omega(x)\rho_\omega(Tx)\dots\rho_\omega(T^{n-1}x) \ .$$

For a bounded function (defined everywhere) on X, we write $B_{\omega,f}$ for the set of points $x \in X$ where the averages

$$\frac{1}{N}\sum_{n=0}^{N-1} \rho_\omega^{(n)}(x) f(T^n x) \tag{40}$$

converge as $N \to +\infty$. Define the function $\psi_{\omega,f}$ on $B_{\omega,f}$ to be the limit of these averages. The set $B_{\omega,f}$ is clearly invariant under T and the function $\psi_{\omega,f}$ satisfies

$$\psi_{\omega,f}(Tx) = \psi_{\omega,f}(x)\overline{\rho_\omega(x)} \quad \text{for } x \in B_{\omega,f} \ . \tag{41}$$

Define

$$C_{\omega,f} = \{x \in B_{\omega,f} : \psi_{\omega,f}(x) \neq 0\} \ .$$

Then $C_{\omega,f}$ is invariant under T. For every bounded function f on X, the subset

$$C_f = \{(\omega, x) \in \Omega \times X : x \in C_{\omega,f}\}$$

is measurable in $\Omega \times X$.

We show now that $\mu(B_{\omega,f}) = 1$. Let $X \times \mathcal{S}^1$ be endowed with the transformation associated to the cocycle ρ_ω and let ϕ be the function defined on $X \times \mathcal{S}^1$ by $\phi(x, u) = f(x)u$. By applying the ergodic theorem on the system $X \times \mathcal{S}^1$ and the function ϕ, we get that the averages (40) converge almost everywhere. That is, $\mu(B_{\omega,f}) = 1$. Therefore, the function $\psi_{\omega,f}$ is defined μ-almost everywhere, and satisfies (41) μ-almost everywhere. By the same argument, for every $y \in Y$, the same properties hold with μ_y substituted for μ.

Choose a countable family $\{f_j : j \in J\}$ of bounded functions on X that is dense in $L^2(\mu)$ and dense in $L^2(\mu_y)$ for every $y \in Y$. Define

$$C_\omega = \bigcup_{j\in J} C_{\omega,f_j} \text{ and } C = \bigcup_{j\in J} C_{f_j} \ .$$

We claim that

$$A = \{(\omega, y) \in \Omega \times Y : \mu_y(C_\omega) = 1\} \ . \tag{42}$$

Let $\omega \in \Omega$ and $y \in Y$ so that $(\omega, y) \in A$. There exists $f\colon X \to \mathcal{S}^1$ so that $\rho_\omega(x) = \overline{f(Tx)}f(x)$ for μ_y-almost every x and by construction, $\psi_{\omega,f} = f$ μ_y-a.e. Choose a sequence $\{j_k\}$ in J so that $f_{j_k} \to f$ in $L^2(\mu_y)$. The sequence of functions $\{\psi_{\omega,f_{j_k}}\}$ converges in $L^2(\mu_y)$ to $\psi_{\omega,f} = f$, which is of modulus 1. By definition of these sets, $\mu_y\bigl(\bigcup_{k=1}^\infty C_{\omega,f_{j_k}}\bigr) = 1$ and thus finally $\mu_y(C_\omega) = 1$.

Conversely, assume that $\mu_y(C_\omega) = 1$. This set is the union for $j \in J$ of the invariant sets C_{ω,f_j}. Thus we can find a sequence $\{D_j\}$ of measurable subsets of X, invariant and pairwise disjoint, with

$$D_j \subset C_{\omega,f_j} \text{ for every } j \text{ and } \bigcup_{j\in J} D_j = C_\omega \ .$$

Define a function f on C_ω by $f(x) = f_j(x)$ for $x \in D_j$. As the sets D_j are invariant, it follows from the construction that for every j and every $x \in D_j$ we have $\psi_{\omega,f}(x) = \psi_{\omega,f_j}(x) \neq 0$. Then $\psi_{\omega,f} \neq 0$ on C_ω and so μ_ω-almost everywhere. By dividing the two sides of Equation (41) by $|\psi_{\omega,f}|$, we get that ρ_ω is a coboundary of (X, μ_ω, T) and that $(\omega, y) \in A$.

Our claim (42) is proven and the first part of Lemma C.3 follows.

(2) If ρ_ω is a coboundary of (X, μ, T), there exists $f \in \mathcal{C}(X, \mathcal{S}^1)$ with $\rho_\omega = f \circ T \cdot \overline{f}$, μ-almost everywhere. As $\mu = \int \mu_y \, d\nu(y)$, for ν-almost every y the same relation holds μ_y-almost everywhere and ρ_ω is a coboundary of (X, μ_y, T).

Conversely, assume that for ν-almost every y the cocycle ρ_ω is a coboundary of (X, μ, T). Define the sets C_{ω,f_j} and C_ω as above. For ν-almost every y we have $(\omega, y) \in A$ and thus $\mu_y(C_\omega) = 1$. It follows that $\mu(C_\omega) = 1$. Use the sets D_j and the function f defined above, with the measure μ substituted for μ_y. The function $\psi_{\omega,f}$

is defined and non-zero μ-almost everywhere and satisfies Equation (41) μ-almost everywhere. Therefore, ρ is a coboundary of (X, μ, T). $\square$

For simplicity, we stated and proved the preceding Lemma only for cocycles with values in the circle group $\mathcal{S}^1$. But it follows immediately from Lemma C.1 that a similar result holds for cocycles with values in any compact abelian group. (We recall that implicitly we assume that all compacts abelian groups are metrizable.)

On the other hand, the full form of Lemma C.3 is used only in the proof of Theorem 9.6. Several times we use a weaker form with a single cocycle, corresponding to a constant map $\omega \mapsto \rho_\omega$:

Corollary C.4. *Let (X, μ, T), (Y, ν) and μ_ω be as in Lemma C.3. Let U be a compact abelian group and $\rho\colon X \to U$ a cocycle. Then the subset*

$$A_\rho = \{y \in Y : \rho \text{ is a coboundary of } (X, \mu_y, T)\}$$

of Y is measurable. The cocycle ρ is a coboundary of (X, μ, T) if and only if $\nu(A_\rho) = 1$.

C.4. **Quasi-coboundaries and cocycles on squares.** Let (X, μ, T) be an ergodic system, U a torus and $\rho\colon X \to U$ a cocycle. ρ is a *quasi-coboundary* if it is the sum of a coboundary and a constant.

We recall that ρ is weakly mixing if and only if there exists no nontrivial character γ of U so that $\gamma \circ \rho\colon X \to \mathbb{T}$ is a quasi-coboundary.

A proof of the following result can be found in Moore and Schmidt [MS80]:

Lemma C.5. *Let (X, μ, T) be an ergodic system, U a torus and $\rho\colon X \to U$ a cocycle. If the map $(x, x') \mapsto \rho(x) - \rho(x')\colon X \times X \to U$ is a coboundary of $(X \times X, \mu \times \mu, T \times T)$, then ρ is a quasi-coboundary.*

We note that the analogous result does not hold for a cocycle with values in an arbitrary compact abelian group.

Lemma C.6. *Let (X, μ, T) be an ergodic system, U a compact abelian group and $\rho \in \mathcal{C}(X, U)$ a cocycle. Assume that the map $(x, x') \mapsto \rho(x)\colon X \times X \to \mathbb{T}$ is a coboundary on $(X \times X, \mu \times \mu, T \times T)$. Then ρ is a coboundary.*

Proof. By Lemma C.1, we can reduce to the case that U is the circle group $\mathcal{S}^1$. Write (Z, t) for the Kronecker factor of X and $\pi\colon X \to Z$ for the natural projection.

By hypothesis, there exists a function $f\colon X \times X \to \mathcal{S}^1$ with

$$f(Tx, Tx')\overline{f(x, x')} = \rho(x) \ .$$

The function defined on $X \times X \times X$ by $(x, x', x'') \mapsto f(x, x')\overline{f(x, x'')}$ is invariant under $T \times T \times T$ and thus is measurable with respect to $\mathcal{Z} \times \mathcal{Z} \times \mathcal{Z}$. It follows that the function f is measurable with respect to $\mathcal{X} \times \mathcal{Z}$. Taking the Fourier transform of f with respect to the second variable, we can write

$$f(x, x') = \sum_{\gamma \in \widehat{Z}} g_\gamma(x)\gamma(\pi(x')) \ . \tag{43}$$

Then

$$f(x, x')\,\overline{f(x, x'')} = \sum_{\gamma, \theta \in \widehat{Z}} g_\gamma(x)\,\overline{g_\theta(x)}\,\gamma(\pi(x'))\,\overline{\theta(\pi(x''))} \ .$$

As this function is invariant under $T\times T\times T$, by unicity of the Fourier transform we get that for every $\gamma,\theta\in\widehat{Z}$,

$$g_\gamma(Tx)\,\overline{g_\theta(Tx)}\,\overline{g_\gamma(x)}\,g_\theta(x)=\overline{\gamma(t)}\theta(t)\ .$$

The function $x\mapsto g_\gamma(x)\,\overline{g_\theta(x)}$ is an eigenfunction of X for the eigenvalue $\overline{\gamma(t)}\theta(t)$ and so there exists a constant $c_{\gamma,\theta}$ with

$$g_\gamma(x)\,\overline{g_\theta(x)}=c_{\gamma,\theta}\,\overline{\gamma(\pi(x))}\,\theta(\pi(x))\ .$$

Finally, there exists a function ϕ on X and for every $\gamma\in\widehat{U}$ there exists a constant c_γ so that

$$g_\gamma(x)=c_\gamma\,\phi(x)\,\overline{\gamma(\pi(x))}\ .$$

Using the values of the functions g_γ in Equation (43), there exists a function g on Z with $f(x,x')=\phi(x)g(\pi(x)-\pi(x'))$. As f is of modulus 1, the functions g and ϕ have constant modulus and so we can assume that $|\phi|=1$. We have $\rho(x)=\phi(Tx)\overline{\phi(x))}$. □

The next Lemma uses the definition and properties of the measures $\mu^{[k]}$ introduced in Section 3. The notation $\xi^{[k]}_\alpha$ was introduced in Section 2.1.

Lemma C.7. *Let (X,μ,T) be an ergodic system, $1\leq\ell\leq k$ integers and let α be an ℓ-face of V_k. Let U be a compact abelian group and $\rho\colon X^{[\ell]}\to U$ a cocycle. If the cocycle $\rho\circ\xi^{[k]}_\alpha=X^{[k]}\to U$ is a coboundary of $(X^{[k]},\mu^{[k]},T^{[k]})$, then ρ is a coboundary of $(X^{[\ell]},\mu^{[\ell]},T^{[\ell]})$.*

Proof. We begin by the case $\ell=0$. Here ρ is a cocycle on X. Assuming that for some vertex ϵ of V_k the cocycle $\mathbf{x}\mapsto\rho(x_\epsilon)$ is a coboundary of $X^{[k]}$, we have to show that ρ is a coboundary on X. By permuting coordinates, we can restrict to the case that ϵ is the vertex $\mathbf{0}$.

We proceed by induction on k. For $k=1$, the result is exactly Lemma C.6. Take $k\geq 1$ and assume that the result holds for k. Assume that the cocycle $\mathbf{x}\mapsto\rho(x_{\mathbf{0}})$ is a coboundary of $X^{[k+1]}$. We use the ergodic decomposition (4) of $\mu^{[k]}$ and the formula (5) for $\mu^{[k+1]}$. By Corollary C.4, for almost every ω the cocycle $\mathbf{x}\mapsto\rho(x_{\mathbf{0}})$ is a coboundary on the Cartesian square of $(X^{[k]},\mu^{[k]}_\omega,T^{[k]})$. This cocycle depends only on the first coordinate of this square and by Lemma C.6 we get that the map $\mathbf{x}'\mapsto\rho(x'_{\mathbf{0}})$ is a coboundary of the system $(X^{[k]},\mu^{[k]}_\omega,T^{[k]})$. As this holds for almost every ω, the map $\mathbf{x}'\mapsto\rho(x'_{\mathbf{0}})$ is a coboundary of the system $(X^{[k]},\mu^{[k]},T^{[k]})$ by Corollary C.4. By the induction hypothesis, ρ is a coboundary of X. This completes the proof when $\ell=0$.

Consider the case that $\ell>0$. We use the ergodic decomposition given by Formula (5) for $\mu^{[\ell]}$ and by Lemma 3.1 we get

$$\mu^{[k]}=\int_{\Omega_\ell}\big(\mu^{[\ell]}_\omega\big)^{[k-l]}\,dP_\ell(\omega)\ .$$

We use Corollary C.4 and the first part of the proof with $k-\ell$ substituted for k and $(X^{[\ell]},\mu^{[\ell]}_\omega,T^{[\ell]})$ substituted for (X,μ,T). The result follows. □

C.5. **Cocycles and group of automorphisms.** Let (X, μ) be a probability space, G a compact abelian group and $(g, x) \mapsto g \cdot x$ an action of G on X by measure preserving transformations. This action is said to be *free* if there exists a probability space (Y, ν) and a measurable bijection $j\colon Y \times G \to X$, mapping $\nu \times m_G$ to μ, with $j(y, gh) = g \cdot j(h)$ for $y \in Y$ and $g, h \in G$.

The vertical rotations introduced in Section C.1 are free actions. The action of a compact abelian group on itself by translations is free. The restriction of a free action to a closed subgroup is free.

The next Lemma says that a free action of a compact abelian group G is 'cohomologically' free. It is a classical result, but we give a proof for completeness.

Lemma C.8. *Let $\{S_g : g \in G\}$ be a free action of the compact abelian group G on the probability space (X, μ) and let $g \mapsto \phi_g$ be a measurable map from G to $\mathcal{C}(X, \mathcal{S}^1)$ so that*

$$\phi_{gh} = \phi_g \cdot (\phi_h \circ g) \text{ for every } g, h \in G \ . \tag{44}$$

Then there exists $\phi \in \mathcal{C}(X, \mathcal{S}^1)$ so that $\phi_g = (\phi \circ S_g) \cdot \overline{\phi}$ for every $g \in G$.

Proof. For $g \in G$, let S_g be the unitary operator on $L^2(\mu)$ given by $S_g f(x) = \phi_g(x) f(g \cdot x)$. The hypothesis (44) means that $\{S_g : g \in G\}$ is a unitary representation of the compact abelian group G in $L^2(\mu)$. Therefore, $L^2(\mu)$ is the Hilbert sum of the spaces H_γ, $\gamma \in \widehat{G}$, where

$$H_\gamma = \{f \in L^2(\mu) : S_g f = \gamma(g)\, f \text{ for every } g \in G\} \ .$$

If $f \in H_\gamma$, the function $|f|$ is invariant under the action of G and thus so is the set $\{x \in X : f(x) \neq 0\}$. Therefore, there exists a partition $X = \bigcup_n X_n$ of X into invariant sets and there exists for every n a character $\gamma_n \in \widehat{G}$ and a function $f_n \in H_{\gamma_n}$ with $f_n(x) \neq 0$ for $x \in X_n$. As the action of G is free, for every n there exists a function $h_n\colon X \to \mathcal{S}^1$ with $h_n \circ g = \gamma_n(g) h_n$ for every $g \in G$. The function ϕ defined on X by

$$\phi(x) = h_n(x) \frac{f_n(x)}{|f_n(x)|} \text{ for } x \in X_n$$

satisfies the announced property. □

Lemma C.9. *Let (X, μ, T) be an ergodic system, U a compact abelian group and let $(u, x) \mapsto u \cdot x$ be a free action of U on X by automorphisms. Let $\rho \in \mathcal{C}(X)$ be a cocycle so that $\rho \circ S_u - \rho$ is a coboundary for every $u \in U$. Then there exists an open subgroup U_0 of U and a cocycle ρ', cohomologous to ρ, with $\rho' \circ S_u = \rho'$ for every $u \in U_0$.*

Proof. By hypothesis, for every $u \in U$ there exists $f \in \mathcal{C}(X)$ with

$$\rho \circ S_u - \rho = f \circ T - f \ . \tag{45}$$

As in Appendix A, for $f \in \mathcal{C}(X)$ and $u \in U$ we write $S_{u,f}$ for the measure preserving transformation of $X \times \mathbb{T}$ given by $S_{u,f}(x, t) = (S_u x, t + f(x))$. The skew product group $U \ltimes \mathcal{C}(X)$ consists in all transformations of this kind. Let $\mathcal{K}$ be the subset of $U \ltimes \mathcal{C}(X)$ consisting in the transformations $S_{u,f}$, where u, f satisfy Equation (45). Clearly, $\mathcal{K}$ is a closed subgroup of $U \ltimes \mathcal{C}(X)$. By hypothesis, the natural projection $p\colon \mathcal{K} \to U$ is onto and its kernel is $\{S_{1,c} : c \in \mathbb{T}\}$, which is a group homeomorphically isomorphic to $\mathbb{T}$. By Corollary A.2, $\mathcal{K}$ is compact. We identify $\ker(p)$ with $\mathbb{T}$.

As p is a homomorphism to an abelian group, its kernel $\mathbb{T}$ contains the commutator subgroup $\mathcal{K}'$ of $\mathcal{K}$. But $\mathbb{T}$ is obviously included in the center of $\mathcal{K}$. Thus $\mathcal{K}$ is a ≤ 2-step nilpotent group, and the commutator map $\mathcal{K}\times\mathcal{K}\to\mathbb{T}$ is bilinear. This map is also continuous and is trivial on $\mathcal{K}\times\mathbb{T}$ and on $\mathbb{T}\times\mathcal{K}$. Thus it induces a continuous bilinear map $\mathcal{K}/\mathbb{T}\times\mathcal{K}/\mathbb{T}\to\mathbb{T}$. As $\mathcal{K}/\mathbb{T}$ can be identified with U, this map can be viewed as a bilinear map from $U\times U$ to $\mathbb{T}$ and by duality we see it as a continuous group homomorphism from U to $\widehat{U}$. As $\widehat{U}$ is discrete, the kernel of this last homomorphism is an open subgroup U_0 of U. Following these identifications back, we get that $p^{-1}(U_0)$ is abelian.

The compact abelian group $p^{-1}(U_0)$ admits $\mathbb{T}$ as a closed subgroup, with quotient equal to U_0. Thus it is isomorphic to $U_0\oplus\mathbb{T}$. This means that the restriction of p to $p^{-1}(U_0)$ admits a cross section which is a continuous group homomorphism. This cross section has the form $u\mapsto S_{u,f_u}$ and $u\mapsto f_u$ is a continuous map from $U_0\to\mathcal{C}(X)$, with

$$\text{for all } u\in U_0,\ \rho\circ u-\rho=f_u\circ T-f_u\ ; \tag{46}$$

$$\text{for all } u,v\in U_0,\ f_{uv}(x)=f_u(x)+f_v(S_u x)\ . \tag{47}$$

Since the action of U on X is free, by Equation (47) and Lemma C.8, there exists $f\in\mathcal{C}(X)$ so that $f_u=f\circ u-f$ for every $u\in U_0$. Write $\rho'=\rho-f\circ T+f$. This cocycle is cohomologous to ρ and by Equation (46), $\rho'\circ u=\rho'$ for $u\in U_0$. □

Lemma C.10. *Let (X,μ,T) be an ergodic system, U a compact abelian group and $(u,x)\mapsto u\cdot x$ a free action of U on X by automorphisms. Let $\rho\in\mathcal{C}(X)$ be a cocycle, so that $\rho\circ u-\rho$ is a quasi-coboundary for every $u\in U$. Then there exists a closed subgroup U_1 of U so that U/U_1 is toral and there exists a cocycle ρ', cohomologous to ρ, with $\rho'\circ S_u=\rho'$ for every $u\in U_1$.*

Proof. The beginning of the proof is similar to the proof of Lemma C.9. For every $u\in U$, there exists $f\in\mathcal{C}(X)$ and a constant $c\in\mathbb{T}$ so that

$$\rho\circ u-\rho=f\circ T-f+c\ . \tag{48}$$

Let $\mathcal{H}$ be the subset of $U\ltimes\mathcal{C}(X)$ consisting in transformations $S_{u,f}$ so that u and f satisfy Equation (48) for some c. Clearly, $\mathcal{H}$ is a closed subgroup of $U\ltimes\mathcal{C}(X)$. By hypothesis, the projection $p\colon\mathcal{H}\to U$ is onto and its kernel is $\{S_{1,f} : f \text{ is an eigenfunction of } X\}$. Thus $\ker(p)$ is homeomorphically isomorphic to the group $\mathcal{A}(Z)$ of affine functions on the Kronecker factor Z of X (for this notation see Section 8.4). This group can be identified with $\mathbb{T}\oplus\widehat{Z}$ and in particular, it is locally compact. By Corollary A.2, $\mathcal{H}$ is locally compact.

A direct computation shows that the commutator subgroup $\mathcal{K}'$ of $\mathcal{K}$ is included in the subgroup $\mathbb{T}$ of $\mathcal{H}$. Thus $\mathcal{K}=\mathcal{H}/\mathbb{T}$ is a locally compact abelian group. We write $q\colon\mathcal{K}\to U$ for the continuous group homomorphism induced by p.

For $S_{u,f}\in\mathcal{H}$, the constant c appearing in Equation (48) is well defined and the map $\psi\colon S_{u,f}\mapsto c$ induces a continuous group homomorphism from $\mathcal{H}$ to $\mathbb{T}$. This homomorphism is trivial on $\mathbb{T}$ and it induces a character ϕ of $\mathcal{K}=\mathcal{H}/\mathbb{T}$.

By the Structure Theorem of Locally compact Abelian Groups, $\mathcal{K}$ admits an open subgroup $\mathcal{L}$ isomorphic to $K\oplus\mathbb{R}^d$, where K is a compact abelian group and $d\geq 0$ is an integer. We identify $\mathcal{L}$ with $K\oplus\mathbb{R}^d$ and write $K_0=K\cap\ker(\phi)$ and U_0 for the closed subgroup $q(K_0)$ of U.

For $u\in U_0$, there exists by definition $f\in\mathcal{C}(X)$ so that $S_{u,f}\in\mathcal{H}$ and $\psi(S_{u,f})=0$. In other words, u and f satisfy Equation (48) with $c=0$, meaning, they satisfy

Equation (45). By Lemma C.9, there exist an open subgroup U_1 of U_0 and a cocycle ρ', cohomologous to ρ, with $\rho' \circ u = \rho'$ for every $u \in U_1$.

It remains to show that U/U_1 is a toral group. As $\mathcal{L}$ is open in $\mathcal{K}$ and q is an open map, $q(\mathcal{L})$ is an open subgroup of U and thus $U/q(\mathcal{L})$ is finite. $q(\mathcal{L})/q(K)$ is a quotient of $\mathcal{L}/K = \mathbb{R}^d$ and is compact and thus it is a torus. K/K_0 is isomorphic to $\phi(K)$, which is a closed subgroup of $\mathbb{T}$ and so is equal to $\mathbb{T}$ or is finite. $q(K)/U_0$ is a quotient of K/K_0 and so it is either finite or isomorphic to $\mathbb{T}$. Finally, U_0/U_1 is open and the proof is complete. □

Appendix D. the van der Corput Lemma

We use several extensions of the classical van der Corput inequality, as found for example in [KN74]. Both deal with sequences in a Hilbert space. Here $\mathcal{H}$ is a Hilbert space, with norm $\|\cdot\|$ and inner product $\langle \cdot \mid \cdot \rangle$. Let $\mathrm{Re}(z)$ denote the real part of the complex number z.

Lemma D.1 ([Be87])**.** *Let $\{x_n\}$ be a sequence in $\mathcal{H}$. For integers N and H with $1 \leq H \leq N$ we have*

$$H^2\Big\|\sum_{n=1}^{N} x_n\Big\|^2 \leq H(N+H-1)\sum_{n=1}^{N}\|x_n\|^2 + 2(N+H-1)\sum_{h=1}^{H-1}(H-h)\sum_{n=1}^{N-h}\mathrm{Re}\langle x_n \mid x_{n+h}\rangle \ .$$

Taking limits in this inequality, we get:

Lemma D.2. *Let $\{x_n\}$ be a bounded sequence in $\mathcal{H}$. We have*

$$\limsup_{N\to\infty}\Big\|\frac{1}{N}\sum_{n=1}^{N} x_n\Big\|^2 \leq \limsup_{H\to\infty}\frac{1}{H}\sum_{h=1}^{H}\limsup_{N\to\infty}\Big|\frac{1}{N}\sum_{n=1}^{N}\langle x_n \mid x_{n+h}\rangle\Big| \ .$$

We need also a similar result for multidimensional sequences. The following Lemma can be found in the proof of Lemma A6 of [BMC00]. Here we write $n = (n_1, \ldots, n_k)$ for a point in $\mathbb{Z}^k$.

Lemma D.3. *Let $\{x_n : n \in \mathbb{Z}^k\}$ be a bounded sequence in $\mathcal{H}$. Assume that for every $h = (h_1, \ldots, h_k) \in \mathbb{Z}^k$*

$$\prod_{i=1}^{k}\frac{1}{N_i-M_i}\sum_{\substack{M_1\leq n_1<N_1,\\ \ldots,\\ M_k\leq n_k<N_k}}\mathrm{Re}\langle x_{n+h} \mid x_n\rangle \to \gamma_h$$

as $M_1 - N_1, \ldots, N_k - M_k \to +\infty$, and that

$$\sum_{\substack{-H\leq h_1\leq H,\\ \ldots,\\ -H\leq h_k\leq H}}\prod_{i=1}^{k}\frac{H-|h_i|}{H^2}\cdot\gamma_h \longrightarrow 0$$

as $H \to +\infty$. Then

$$\Big\|\prod_{i=1}^{k}\frac{1}{N_i-M_i}\sum_{\substack{M_1\leq n_1<N_1,\\ \ldots,\\ M_k\leq n_k<N_k}} x_n\Big\| \longrightarrow 0$$

as $N_1 - M_1, \ldots, N_k - M_k \to +\infty$.

References

[AGH63] L. Auslander, L. Green and F. Hahn. Flows on homogeneous spaces. *Ann. Math. Studies* **53**, Princeton Univ. Press (1963).

[BK96] H. Becker and A. S. Kechris. The descriptive theory of Polish groups actions. *London Math. Soc. Series* 232, Cambridge Univ. Press (1996).

[Be87] V. Bergelson. Weakly mixing PET. *Erg. Th. & Dyn. Sys.*, **7** (1987), 337–349.

[Be00] V. Bergelson. The multifarious Poincaré Recurrence Theorem. *Descriptive Set Theory and Dynamical Systems*, Eds. M. Foreman, A.S. Kechris, A. Louveau, B. Weiss. Cambridge University Press, New York (2000), 31-57.

[BMC00] V. Bergelson and R. McCutcheon. An Ergodic IP Polynomial Szemerédi Theorem. *Mem. Amer. Math. Soc.* **146** (2000), #146.

[Bo89] J. Bourgain. Pointwise ergodic theorems for arithmetic sets. *Inst. Hautes Études Sci. Publ. Math.*, **69** (1989), 5–45.

[CL84] J.-P. Conze and E. Lesigne. Théorèmes ergodiques pour des mesures diagonales. *Bull. Soc. Math. France*, **112** (1984), 143–175.

[CL87] J.-P. Conze and E. Lesigne. Sur un théorème ergodique pour des mesures diagonales. *Publications de l'Institut de Recherche de Mathématiques de Rennes, Probabilités*, 1987.

[CL88] J.-P. Conze and E. Lesigne. Sur un théorème ergodique pour des mesures diagonales. *C. R. Acad. Sci. Paris, Série I*, **306** (1988), 491–493.

[F61] H. Furstenberg. Strict ergodicity and transformations of the torus. *Amer. J. of Mathematics*, **83** (1961), 573–601.

[F77] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szeméredi on arithmetic progressions. *J. d'Analyse Math.*, **31** (1977), 204–256.

[F81] H. Furstenberg. Recurrence in Ergodic Theory and Combinatorial Number Theory. Princeton Univ. Press (1981).

[FW96] H. Furstenberg and B. Weiss. A mean ergodic theorem for $\frac{1}{N}\sum_{n=1}^{n} f(T^n x)g(T^{n^2}x)$. *Convergence in Ergodic Theory and Probability*, Eds.:Bergelson, March, Rosenblatt. Walter de Gruyter & Co, Berlin, New York (1996), 193–227

[G01] T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, **11** (2001), 465-588.

[HK01] B. Host and B. Kra. Convergence of Conze-Lesigne Averages. *Erg. Th. & Dyn. Sys.*, **21** (2001), 493–509.

[HK02] B. Host and B. Kra. An Odd Furstenberg-Szemerédi Theorem and quasi-affine systems. *J. d'Analyse Math.*, **86** (2002), 183–220.

[HK02] B. Host and B. Kra. Convergence of polynomial ergodic averages, submitted. Available at: http://www.math.psu.edu/kra/.

[HK04] B. Host and B. Kra. Averaging along cubes. *Dynamical Systems and Related Tpopics*, Eds. Brin, Hasselblatt, Pesin. Cambridge University Press, Cambridge (2004).

[K34] A. Y. Khintchine. Eine Verschärfung des Poincaréschen "Wiederkehrsatzes" *Comp. Math.*, **1** (1934), 177–179.

[KN74] L. Kuipers and H. Niederreiter. Uniform distribution of sequences. John Wiley and Sons (1974).

[Lb02] A. Leibman. Pointwise convergence of ergodic averages for polynomial sequences of rotations of a nilmanifold. Preprint (2002). Available at http://www.math.ohio-state.edu/~leibman/preprints .

[L84] E. Lesigne. Résolution d'une équation fonctionelle. *Bull. Soc. Math. France*, **112** (1984), 177–19.

[L87] E. Lesigne. Théorèmes ergodiques ponctuels pour des mesures diagonales. Cas des systeèmes distaux. *Ann. Inst. Henri Poincaré*, **23** (1987), 593–612.

[L89] E. Lesigne. Théorèmes ergodiques pour une translation sur une nilvariété. *Erg. Th. & Dyn. Sys.*, **9-1** (1989), 115–126.

[L91] E. Lesigne. Sur une nil-variété les parties minimales associèes á une translation sont uniquement ergodiques. *Erg. Th. & Dyn. Sys.*, **11-2** (1991), 379–391.

[L93] E. Lesigne. Équations fonctionelles, couplages de produits gauches et théorèmes ergodiques pour mesures diagonales. *Bull. Soc. Math. France*, **121** (1993), 315–351.

[Mo48] D. Montgomery. Dimensions of factor spaces. *Ann. Math.*, **49** 2 (1948), 373–378.

[Ma51] A. Malcev. On a class of homogeneous spaces. *Amer. Math. Soc. Transl.* **39** (1951)

[MS80] C. C. Moore and K. Schmidt. Coboundaries and homomorphisms for non-singular actions and a problem of H. Helson. *Proc. London. Math. Soc.*, **3** 40 (1980), 443–475.

[MZ55] D. Montgomery and L. Zippin. Topological Transformation Groups. Interscience Publishers (1955)

[P69] W. Parry. Ergodic properties of affine transformations and flows on nilmanifolds. *Amer. J. Math.*, **91**, 3 (1969), 757-771.

[P70] W. Parry. Dynamical systems on nilmanifolds. *Bull. London Math. Soc.* **2** (1970), 37–40.

[Ra91] M. Ratner. On Raghunathan's measure conjecture. *Ann. Math.*, **134** (1991), 545-607.

[Ru95] D.J. Rudolph. Eigenfunctions of $T \times S$ and the Conze-Lesigne algebra. *Ergodic Theory and its Connections with Harmonic Analysis*, Eds.:Petersen/Salama, Cambridge University Press, New York (1995), 369–432.

[Sh96] N. Shah. Invariant measures and orbit closures on homogeneous spaces for actions of subgroups. Lie groups and ergodic theory (Mumbai, 1996) Tata Inst. Fund. Res., Bombay (1998), 229–271.

[Va70] V.S. Varadarajan. Geometry of Quantum Theory, Vol. II. Van Nostrand (1970).

[Zie02a] T. Ziegler. A nonconventional ergodic theorem for a nilsystem. Available at http://www.arxiv.org, math.DS/0204058 v1 (2002).

[Zie02b] T. Ziegler. *Personal communication.*

[Zim76] R. Zimmer. Extensions of ergodic group actions. *Illinois J. Math.*, **20** (1976), 373-409.

Equipe d'analyse et de mathématiques appliquées, Université de Marne la Vallée, 77454 Marne la Vallée Cedex, France

E-mail address: `host@math.univ-mlv.fr`

Department of Mathematics, McAllister Building, The Pennsylvania State University, University Park, PA 16802

E-mail address: `kra@math.psu.edu`