Non-Conventional Ergodic Averages

Thesis Submitted for the Degree
Doctor of Philosophy
by
Tamar Ziegler

Submitted to the Senate of the Hebrew University of Jerusalem.
December 2002

This work was done under the supervision of professor Hillel Furstenberg

I thank Benjamin Weiss and Shahar Moses for sharing their insights with me. I thank Vitaly Bergelson and Alexander Leibman for valuable remarks on the second part of the thesis. I thank Yair Glasner for his patience with my frustrations. I thank my husband David, my children Yoni and Tzur, and my parents Ilana and Zvi for their love. Most of all I thank Hillel Furstenberg for showing me the beauty of ergodic theory and for his constant encouragement and support.

CONTENTS

1. INTRODUCTION

1.1. Non-conventional ergodic averages of the form

$$\frac{1}{N}\sum_{n=1}^{N}\prod_{j=1}^{k} f(T^{jn}x) \tag{1}$$

were first introduced by Furstenberg [8] in his ergodic theoretic proof of Szemèredi's theorem on arithmetic sequences in sets of positive density in $\mathbb{Z}$. He proved the following theorem:

1.2. Theorem (Furstenberg)**.** *Let $(X,\mathcal{B},\mu,T)$ be a measure preserving system. Let A be a set of positive measure, and $f = 1_A$. Then*

$$\liminf_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\int\prod_{j=1}^{k} f(T^{jn}x) > 0.$$

The theorem above ensures that there exists an integer n, such that the points $x, T^n x, \dots, T^{kn}x$ are in A. Questions related to the limiting behavior of the averages (1) (pointwise or in L^2), and related averages are an active field of research. For $k = 1$ the average in (1) is the classical ergodic average. Pointwise convergence is the classical Birkhoff ergodic theorem, and L^2 convergence is given by the mean ergodic theorem. Convergence of the averages (1) for the case $k = 2$ is relatively easy in the L^2 mode. Pointwise convergence in this case is difficult and was shown by Bourgain [4]. Convergence in L^2 for $k = 3$ is much more difficult and was shown by Furstenberg and Weiss [10], Conze and Lesigne [5],[6],[7], Host and Kra [11],[12]. Pointwise convergence for $k \geq 3$ is not known.

1.3. We concentrate on the behavior of the averages in (1) for the case $k = 4$. The idea is to show that the limiting behavior of the ergodic averages in question depends on a certain factor which will be described later, and to take advantage of the special nature of these factors for which a pointwise theorem is, in fact, available. Most of my work revolves about the study of this factor, but I have also devoted a chapter to the ergodic pointwise theorem (theorem 1.19).

1.4. By ergodic decomposition it will suffice to study the limit of (1) with the additional hypothesis of ergodicity, but the nature of the limit will depend on mixing properties of the system. The maximal degree of mixing relevant is *weak mixing*; indeed in this case Furstenberg has shown:

1.5. Theorem (Furstenberg)**.** *If $(X,\mathcal{B},\mu,T)$ is weak mixing then*

$$\frac{1}{N}\sum_{n=1}^{N}\prod_{j=1}^{k} f_j(T^{jn}x) \xrightarrow{L^2(X)} \prod_{j=1}^{k}\int f_j(x)d\mu.$$

The interpretation we can give to this is that under the hypothesis of weak mixing the points $x, T^n x, \dots, T^{kn}x$ move almost independently in X.

1.6. **Notation.** We write Tf for the function $Tf(x) = f(Tx)$.

1.7. For a general ergodic system $(X, \mathcal{B}, \mu, T)$ the averages in equation (1) need not converge to a constant function. Indeed, if the system $(X, \mathcal{B}, \mu, T)$ is not weakly mixing there exists a nontrivial eigenfunction ψ. If $T\psi(x) = \lambda\psi(x)$ then

$$T^n\psi^2(x)T^{2n}\psi^{-1}(x) = \psi(x)$$

for all n, thus

$$\frac{1}{N}\sum_{n=1}^{N} T^n\psi^2(x)T^{2n}\psi^{-1}(x) = \psi(x).$$

Therefore the set of limiting functions spans an algebra which contains the algebra spanned by the eigenfunctions - the *Kronecker algebra.* The Kronecker algebra is realized by the "Kronecker factor" $(Z, \mathrm{Borel}, \mathrm{Haar}, \alpha)$, where Z is a compact Abelian group and the action is by rotation by α. It is not surprising that an Abelian group factor should come up when studying the relations between $x, T^nx, T^{2n}x$, as $x, x+n\alpha, x+2n\alpha$ form an arithmetic sequence: $x = 2(x+n\alpha)-(x+2n\alpha)$. It turns out that this constraint imposed by the Kronecker factor is the only constraint on the triple $x, T^nx, T^{2n}x$, and in a manner to be made precise, the Kronecker factor is "characteristic" for the limit of the averages $\frac{1}{N}\sum_{n=1}^{N} f(T^nx)g(T^{2n}x)$.

1.8. **Definition.** Let $(Y, \mathcal{D}, \nu, T)$ be a factor of the system $(X, \mathcal{B}, \mu, T)$ so that there is a measurable map $\alpha : X \to Y$. The space $L^2(Y, \mathcal{D}, \nu, T)$ can be thought of as a subspace of $L^2(X, \mathcal{B}, \mu, T)$. For $f \in L^2(X, \mathcal{B}, \mu, T)$ let $E(f|Y)$ be the projection of f on $L^2(Y, \mathcal{D}, \nu, T)$. The system $(Y, \mathcal{D}, \nu, T)$ is *characteristic for the averaging scheme* $(a_1, a_2, \ldots, a_k)$ if for any bounded functions $f_1, \ldots, f_k$,

$$\frac{1}{N}\sum_{n=1}^{N}\prod_{j=1}^{k} T^{a_jn}f_j - \frac{1}{N}\sum_{n=1}^{N}\prod_{j=1}^{k} T^{a_jn}E(f_j|Y) \xrightarrow{L^2(X)} 0.$$

It is in this sense that the Kronecker factor is characteristic for $(1,2)$; i.e. for calculating the limit of the averages $\frac{1}{N}\sum_{n=1}^{N} f(T^nx)g(T^{2n}x)$. The system $(Y, \mathcal{D}, \nu, T)$ is *characteristic for schemes of length k* if it characteristic for all schemes of length k. The system $(Y, \mathcal{D}, \nu, T)$ is a *minimal characteristic factor for schemes of length k* if it is a factor of any characteristic factor for schemes of length k.

1.9. For averaging schemes of length ≥ 3 the Kronecker algebra does not suffice. Let φ be a *second order eigenfunction,* i.e.: $T\varphi = \psi\varphi$, where ψ is itself an ordinary eigenfunction. Then one can check that

$$T^n\varphi^3(x)T^{2n}\varphi^{-3}(x)T^{3n}\varphi(x) = \varphi(x)$$

for all n; thus

$$\frac{1}{N}\sum_{n=1}^{N} T^n\varphi^3(x)T^{2n}\varphi^{-3}(x)T^{3n}\varphi = \varphi(x).$$

Note that if $(Y, \mathcal{D}, \nu, T)$ is a characteristic factor of $(X, \mathcal{B}, \mu, T)$ for a certain scheme $(a_1, \ldots a_k)$ then the limit of the averages $\frac{1}{N}\sum_{n=1}^{N}\prod_{j=1}^{k} f_j(T^{a_j n}x)$ is always measurable with respect to the subalgebra of $\mathcal{B}$ lifted from $\mathcal{D}$. From the foregoing we see that the characteristic factor for $(1,2,3)$ must be large enough so that all second order eigenfunctions are measurable with respect to it. It is natural to conjecture that the algebra spanned by the second order eigenfunction is characteristic for $(1,2,3)$, but Furstenberg and Weiss presented a counterexample: Let

$$X = \begin{pmatrix} 1 & \mathbb{R} & \mathbb{R} \\ & 1 & \mathbb{R} \\ & & 1 \end{pmatrix} \Big/ \begin{pmatrix} 1 & \mathbb{Z} & \mathbb{Z} \\ & 1 & \mathbb{Z} \\ & & 1 \end{pmatrix} = N/\Gamma.$$

This space carries a measure m invariant under translations by any group element. Consider the system $(N/\Gamma, \mathrm{Borel}, m, T)$ with $Tx\Gamma = ax\Gamma$ for $a \in N$. This system has no second order eigenfunctions, but there are relations between $x\Gamma, a^n x\Gamma, a^{2n}x\Gamma, a^{3n}x\Gamma$ not coming from Kronecker factor: In N, $x = (a^n x)^3(a^{2n}x)^{-3}a^{3n}x$, and this relation carries over to N/Γ. The system N/Γ is a circle extension of a 2 dimensional torus (which is the Kronecker factor). The foregoing relation implies that the points $x, T^n x, T^{2n}x, T^{3n}x$ do not move about freely in the fibers above the Kronecker factor. In particular the characteristic factor for the averages $\frac{1}{N}\sum_{n=1}^{N} T^n f T^{2n} g T^{3n} h$ will contain functions other than first and second order eigenfunctions.

1.10. In general, if N is a k-step nilpotent group $(N_{k+1} = 1)$, then there exists a function F such that for $g \in N$

$$g = F(a^n g, a^{2n} g, \ldots, a^{(k+1)n} g),$$

and this relation carries on to a homogeneous space: If $\Gamma < N$, $x \in G/\Gamma$,

$$x = \tilde{F}(a^n x, a^{2n} x, \ldots, a^{(k+1)n} x).$$

1.11. **Conjecture.** *The* only *constraints on the points* $x, T^n x, \ldots, T^{kn} x$ *come from a k-step nilpotent factor. In particular any persistent relations between the points* $x, T^n x, \ldots, T^{kn} x$ *come from an* algebraic *structure in* X.

1.12. **Definition.** A *nilsystem* consists of a space X on which a nilpotent group N acts transitively preserving a measure μ, and a transformation T which acts by translation by a group element a: $Tx = ax$. If N is a k-step nilpotent Lie group, Γ a lattice, then N/Γ is a *nilmanifold.* The system $(N/\Gamma, \mathrm{Borel}, \mathrm{Haar}, T)$ is a k-step *nilflow* if the action of T is given by $Tx\Gamma = ax\Gamma$ for some $a \in N$. An inverse limit of nilflows is a *pro-nilsystem.*

Another formulation of the above conjecture:

1.13. **Conjecture.** *The minimal characteristic factor for schemes of length k is a $k-1$-step pro-nilsystem.*

1.14. The case $k = 3$ was analyzed by Conze and Lesigne ([5],[6],[7],[15][16]), Furstenberg and Weiss [10], Host and Kra ([11],[12]), Rudolph [27]. The following theorem is proved in [12]:

1.15. **Theorem.** *A characteristic factor for schemes of length* 3 *is a* 2*-step pro-nilsystem.*

1.16. We prove two related results: In the first result we determine the minimal characteristic factor for $k = 4$:

1.17. **Theorem.** *The minimal characteristic factor for averaging schemes of length* 4 *is a* 3*-step pro-nilsystem.*

confirming the foregoing conjecture for the case $k = 4$.

1.18. The second result is a general pointwise convergence theorem for nilflows, which also gives an explicit description of limit.

1.19. **Theorem.** *Let* $(N/\Gamma, a)$ *be a* k*-step nilflow, where* N *is connected, and let* $f_j \in L^\infty(N/\Gamma)$*, then*

$$\frac{1}{N}\sum_{n=1}^{N}\prod_{j=1}^{k+1} f_j(a^{jn}x\Gamma) \quad \textit{converges for almost all (Haar) } x \in N$$

(to $\int_{N/\Gamma}\cdots\int_{N_k/\Gamma_k}\prod_{j=1}^{k+1} f_j(x\prod_{i=1}^{j} y_i^{\binom{j}{i}}\Gamma) \quad \prod_{j=1}^{k+1} dm_j(y_j\Gamma^j)$*).*

1.20. By theorem 1.17 in order to have L^2 convergence of the averages in (1) for the case $k = 4$, it is enough to prove an L^2 convergence theorem for 3-step pro-nilsystems. Convergence in L^2 for pro-nilsystems follows from convergence for nilflows which is given by theorem 1.19 in the case where N is connected, and in general follows from a general result by Shah [28].

1.21. **Remark.** The methods used in proving theorem 1.17 seem to work for general k; the main problem is notational.

1.22. **Remark.** In the course of this work it was brought to my attention that Kra and Host proved by different methods convergence of the averages in (1) for all k [13].

Part 1. Preliminaries

In the following section we give some of the measure-theoretic and ergodic-theoretic preliminaries. The definitions and theorems in sections $1-4$ can be found in [9], [23].

1.23. A *measure preserving system* (m.p.s) is a system $(X,\mathcal{B},\mu,T)$ where X is an arbitrary space, $\mathcal{B}$ is a σ-algebra of subsets of X, μ is a σ-additive probability measure on the sets of $\mathcal{B}$, and T a measure preserving transformation. When there is no confusion we sometimes write (X,T) in short for $(X,\mathcal{B},\mu,T)$.

1.24. The action of T is *ergodic*, if $T^{-1}A=A$, for all $A\in\mathcal{B}$, implies $\mu(A)=0$ or $\mu(A)=1$. In this case we also say that μ is ergodic with respect to the action of T. The transformation T induces a natural operator on $L^2(X,\mathcal{B},\mu)$ defined by $Tf=f\circ T$, and the ergodicity of the action of T is equivalent to the assertion that there are no non-constant T-invariant functions.

1.25. Theorem (Mean Ergodic Theorem). *Let $(X,\mathcal{B},\mu,T)$ be a m.p.s., and let $f\in L^2(X)$. Let Pf denote the orthogonal projection of f on the subspace of the T invariant functions, then*

$$\frac{1}{N}\sum_{n=1}^{N} f\circ T^n \overset{L^2(X)}{\longrightarrow} Pf.$$

2. Factors and disintegration of measure

2.1. Let $(X,\mathcal{B},\mu,T)$, $(Y,\mathcal{D},\nu,T)$ be 2 m.p.s.. A measurable, measure preserving map $\alpha:X\to Y$ defines a *homomorphism* of $(X,\mathcal{B},\mu,T)$ to $(Y,\mathcal{D},\nu,T)$ if for almost every $x\in X$, $\alpha(Tx)=T\alpha(x)$. In this case we say that $(Y,\mathcal{D},\nu,T)$ is a *factor* of $(X,\mathcal{B},\mu,T)$, and $(X,\mathcal{B},\mu,T)$ is an *extension* of $(Y,\mathcal{D},\nu,T)$. The two measure preserving systems are *equivalent* if the homomorphism of one to the other is invertible.

2.2. A m.p.s. $(X,\mathcal{B},\mu,T)$ is *regular* if X is a compact metric space, $\mathcal{B}$ the Borel algebra of X, μ a Borel measure. A m.p.s. is *separable* if $\mathcal{B}$ is generated by a countable subset. As every separable m.p.s. is equivalent to a regular m.p.s., we will confine our attention to regular m.p.s.

2.3. Let $\alpha:(X,\mathcal{B},\mu,T)\to(Y,\mathcal{D},\nu,T)$ be a homomorphism of m.p.s. The map $f\to f^\alpha=f\circ\alpha$ identifies $L^2(Y,\mathcal{D},\nu)$ with a closed subspace $L^2(Y,\mathcal{D},\nu)^\alpha\subset L^2(X,\mathcal{B},\mu)$. If P denotes the orthogonal projection of $L^2(X,\mathcal{B},\mu)$ onto $L^2(Y,\mathcal{D},\nu)^\alpha$, then we define $E(f|Y)$ for $f\in L^2(X,\mathcal{B},\mu)$ by

$$E(f|Y)\in L^2(Y,\mathcal{D},\nu),\quad E(f|Y)^\alpha=Pf.$$

Since T leaves $L^2(Y,\mathcal{D},\nu)^\alpha$ invariant, the operator $E(\cdot|Y)$ commutes with the action of T, i.e. for each $f\in L^1(X,\mathcal{B},\mu)$, $E(Tf|Y)=TE(f|Y)$.

2.4. Let $(X, \mathcal{B}, \mu)$ be a regular measure space, and let $\alpha : (X, \mathcal{B}, \mu) \to (Y, \mathcal{D}, \nu)$ be a homomorphism to another measure space. Suppose α is induced by a map $\varphi : X \to Y$. In this case the measure μ has a disintegration in terms of fiber measures μ_y, where μ_y is concentrated on the fiber $\varphi^{-1}(y) = X_y$. We denote by $\mathcal{M}(X)$ the compact metric space of probability measures on X.

2.5. **Theorem.** *There exists a measurable map from Y to $\mathcal{M}(X)$, $y \to \mu_y$ which satisfies:*

(1) *For every $f \in L^1(X, \mathcal{B}, \mu)$, $f \in L^1(X, \mathcal{B}, \mu_y)$ for a.e. $y \in Y$, and $E(f|Y)(y) = \int f d\mu_y$ for a.e. $y \in Y$*
(2) *$\int \{\int f d\mu_y\} d\nu(y) = \int f d\mu$ for every $f \in L^1(X, \mathcal{B}, \mu)$.*

The map $y \to \mu_y$ is characterized by condition (1). We shall write $\mu = \int \mu_y d\nu$ and refer to this as the disintegration of the measure μ with respect to the factor Y.

If $(X, \mathcal{B}, \mu, G)$ is a m.p.s., $\mathcal{D}$ the algebra of all G-invariant sets, $\mu = \int \mu_x d\mu(x)$ the disintegration of μ with respect to $\mathcal{D}$, then μ_x is G-invariant and ergodic, for a.e. x.

3. THE KRONECKER FACTOR

3.1. An action of a measure preserving transformation T on a measure space $(X, \mathcal{B}, \mu)$ is a *Kronecker action* if X is a compact Abelian group, μ the Haar measure on X, and T acts by rotation: $Tx = x + \alpha$. The system $(X, \mathcal{B}, \mu, T)$ is called a *Kronecker system* (or an *almost periodic system*). Equivalently, $(X, \mathcal{B}, \mu, T)$ is Kronecker if the eigenfunctions of T span $L^2(X)$. Every ergodic system has a maximal almost periodic factor:

3.2. **Theorem.** *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic m.p.s, then there is a map $\pi : X \to Z$ where Z is a compact Abelian group, and a Kronecker action on Z such that $T\pi(x) = \pi(T(x))$ for a.e. $x \in X$. For every character χ on Z the function $\chi'(x) = \chi(\pi(x))$ satisfies*

$$\chi'(Tx) = \chi(\pi(x) + \alpha) = \chi(\alpha)\chi'(x)$$

and so is an eigenvector of T. Moreover, every eigenfunction of the T-action comes about this way.

The factor system $(Z, \mathcal{D}, m, T)$, where $\mathcal{D}$ is the algebra of Borel sets, and m the Haar measure, is unique up to isomorphism and is called the Kronecker factor of $(X, \mathcal{B}, \mu, T)$. For the proof see [8].

4. ISOMETRIC EXTENSIONS

4.1. Let $(X, \mathcal{B}, \mu, T)$ be an ergodic m.p.s., and let $(Y, \mathcal{D}, \nu, T)$ be a factor. Consider the ring $L^\infty(Y)$ as a subring of functions on X. A subspace $V \subset L^2(X)$ is a *finite rank module* over $L^\infty(Y)$ if there exist finitely many functions $\varphi_1, \ldots, \varphi_k$, such that any function $f \in V$ can be expressed as

$f = \sum_{i=1}^{k} a_i(y)\varphi_i(x)$. We say that $(X, \mathcal{B}, \mu, T)$ is an *isometric extension* of $(Y, \mathcal{D}, \nu, T)$ if $L^2(X)$ is spanned by finite rank T invariant modules over $L^\infty(Y)$. It can be shown that in this case $(X, \mathcal{B}, \mu, T)$ is isomorphic to a *skew product* $(Y \times M, \tilde{\mathcal{B}}, \nu \times m, T_\rho)$ where M is a homogeneous compact metric space on which there is a unique probability measure, m, invariant under the transitive group of isometries. The action of T_ρ is given by $T_\rho(y, m) = (Ty, \rho(y)m)$, where $\rho : Y \to Isom(M)$. For example, a Kronecker system is an isometric extension of a point. Define $\rho_n : Y \to G$ by $T_\rho^n(y, m) = (T^n y, \rho_n(y)m)$; then ρ_n satisfies a 1-cocycle equation for the action of $\mathbb{Z}$ on functions from Y to G

$$\rho_{n+m}(y) = \rho_n(T^m y)\rho_m(y).$$

Since $\rho_n(y)$ is determined bi $\rho_1(y)$we shall focus on $\rho(y) = \rho_1(y)$ and refer to it as the *extension cocycle* (or just *cocycle*). We shall denote the extension given by a cocycle ρ by $Y \times_\rho G/H$.

4.2. Let $(X, \mathcal{B}, \mu, T)$ be an ergodic m.p.s., and let $(Y, \mathcal{D}, \nu, T)$ be a factor. Consider the subspace of $L^2(X)$, spanned by all finite rank T-invariant modules over $L^\infty(Y)$. This subspace will be defined by some factor $(\hat{Y}, \hat{\mathcal{D}}, \hat{\nu}, T)$ between X and Y. The system $(\hat{Y}, \hat{\mathcal{D}}, \hat{\nu}, T)$ is called the *maximal isometric extension of* Y *in* X.

4.3. Let $(X_i, \mathcal{B}_i, \mu_i, T_i)$, $i = 1, \ldots, k$, be measure preserving systems, and let $(Y_i, \mathcal{D}_i, \nu_i, T_i)$ be corresponding factors. A measure μ on ΠX_i defines a *joining* of the systems X_i if it is invariant under $T_1 \times \ldots \times T_k$ and maps onto μ_i under the natural map $\Pi X_i \to X_i$. Let μ_{i,y_i} represent the disintegration of μ_i with respect to (Y_i, ν_i), and ν the projection of μ on ΠY_i. We say that μ is the *conditional product joining relative* to ΠY_i if

$$\mu = \int \mu_{1,y_1} \ldots \mu_{k,y_k} d\nu(y_1, \ldots, y_k)$$

4.4. **Theorem.** *Let* X_i, Y_i *be as above. Assume* $(X_i, \mathcal{B}_i, \mu_i, T_i)$ *has finitely many ergodic components. Let* $(\hat{Y}_i, \hat{\mathcal{D}}_i, \hat{\nu}_i, T_i)$ *be the maximal isometric extension of* $(Y_i, \mathcal{D}_i, \nu_i, T_i)$ *in* $(X_i, \mathcal{B}_i, \mu_i, T_i)$. *Let* μ *be a measure on* ΠX_i *defining a joining which is a conditional product relative to* ΠY_i, *then almost every ergodic component of* μ *is a conditional product relative to* $\Pi \hat{Y}_i$; *i.e. If* $F \in L^2(\Pi X_i, \Pi \mathcal{B}_i, \mu)$ *is invariant under* $T_1 \times \ldots \times T_k$ *then there exists a function* $\Phi \in L^2(\Pi \hat{Y}_i, \Pi \hat{\mathcal{D}}_i, \Pi \hat{\nu}_i)$ *for* $\hat{\nu}_i$ *the image of* μ *on* $\Pi \hat{Y}_i$ *so that*

$$F(x_1, \ldots, x_k) = \Phi(\hat{\pi}_1(x_1), \ldots, \hat{\pi}_k(x_k)).$$

Proof. see [8] theorem 9.5. □

4.5. $(X, \mathcal{B}, \mu, T)$ is a *group extension* of $(Y, \mathcal{D}, \nu, T)$ if it is isomorphic to a skew product $(Y \times G, \tilde{\mathcal{B}}, \nu \times Haar, \tilde{T})$ where G is a compact group, and $\tilde{T}(y, g) = (Ty, \rho(y)g)$, where $\rho : Y \to G$. (This is a special case of an isometric extension where $H = \{1\}$) .

4.6. **Lemma.** *Suppose* $(X, \mathcal{B}, \mu, T)$ *is an ergodic isometric extension of* $(Y, \mathcal{D}, \nu, T)$ *so that we can express* $X = Y \times_\rho G/H$. *Using the function* ρ, *we can define a group extension* $Y \times_\rho G$. *Then* G *and* H *can be chosen so that the extension* $Y \times_\rho G$ *is an ergodic group extension.*

Proof. [10] lemma 7.2. □

4.7. **Lemma.** *Let* $X = Y \times_\rho G$ *be an ergodic group extension of* Y, *and let* W *be an intermediate factor between* X *and* Y, *then* X *is a group extension of* W.

Proof. [10] lemma 7.3. □

5. EQUIVALENT COCYCLES AND THE MACKEY GROUP

5.1. Let (Y, ν, T) be an ergodic m.p.s., G a compact metrizable group. Let $(Y \times_\rho G, \mu \times m_G)$ be a group extension: $T_\rho(y, g) = (Ty, \rho(y)g)$. We can reparameterize $Y \times G$ replacing (y, g) with $F(y, g) = (y, f(y)g)$ for some measurable function $f : Y \to G$. Let $\rho'(y) := f(Ty)\rho(y)f(y)^{-1}$, then the systems $(T_\rho, Y \times_\rho G)$, $(T_{\rho'}, Y \times_{\rho'} G)$ are isomorphic, and ρ, ρ' are called *equivalent cocyles*. If ρ is equivalent to the identity cocycle then ρ is a *coboundary.*

5.2. If ρ takes values in a closed subgroup H of G, the extension $Y \times_\rho G$ will not be ergodic (any function on H/G will be invariant). By the foregoing discussion if ρ is equivalent to a cocycle taking values in a closed subgroup H, then the extension $Y \times_\rho G$ will not be ergodic. Mackey's theory says that all ergodic measures μ on $Y \times_\rho G$ that map onto ν under the natural projection come about by replacing the cocycle by an equivalent one, taking its values in an appropriately chosen subgroup.. More precisely:

5.3. **Theorem.** *Let* $Y \times_\rho G, \mu$ *be an ergodic group extension. Suppose* μ *is mapped onto* ν *under the natural projection. There is a closed subgroup* $H < G$, *unique up to conjugacy, so that:*

(1) ρ *is equivalent to a cocycle* ρ' *taking values in* H.
i.e. $\rho'(y) = f(Ty)\rho(y)f(y)^{-1} \in H$.
(2) *The transformation* $T_{\rho'}$ *has ergodic invariant measures* $\nu \times m_{H\gamma}$ *where* $\gamma \in G$, *and* $m_{H\gamma}$ *is the translated Haar measure on* H *(supported on the coset* $H\gamma$*).*
(3) *Any ergodic* $T_{\rho'}$*-invariant measure on* $Y \times G$ *extending* ν *has the form* $\nu \times m_{H\gamma}$ *for some coset* $H\gamma$, *and the ergodic* T_ρ *invariant measures are obtained by applying* F^{-1} *to the ergodic* T_ρ*-invariant measures. The group* H *is called the* Mackey group *of the extension* $Y \times_\rho G$.

6. ABELIAN EXTENSIONS

6.1. **Notation.** We use additive notation for Abelian groups with the exception of the group S^1 which will play a special role in the future. In particular, if ρ, ρ' are equivalent cocycles (defined in the foregoing section) taking values in an Abelian group G, then there exists a function $f : Y \to G$ such that

$$\rho(y) = f(Ty) + \rho'(y) - f(y).$$

6.2. If $Y \times_\rho G$ is an Abelian group extension, the Mackey group defined in the foregoing section H is unique. Let

$$\hat{H} = \{\chi \in \hat{G} : \chi(g) = 1 \text{ for all } g \in H\}$$

be the annihilator of H. If ρ is equivalent to a cocycle taking values in H then $\chi \circ \rho$ is a coboundary for all $\chi \in \hat{H}$. The converse is also true: by Pontryagin duality $G \hookrightarrow (S^1)^{\hat{G}}$, and $\rho = (\rho_\chi)_{\chi \in \hat{G}}$ where $\rho_\chi(y) = \chi \circ \rho(y)$. If $\chi \circ \rho(y) = \frac{f_\chi(Ty)}{f_\chi(y)}$ is a coboundary for all $\chi \in \hat{H}$, then define $\rho'_\chi(y) = \rho_\chi(y)\frac{f_\chi^{-1}(Ty)}{f_\chi^{-1}(y)}$ for all $\chi \in \hat{G}$. Thus

$$\hat{H} = \{\chi \in \hat{G} : \chi \circ \rho \text{ is a coboundary}\}.$$

6.3. **Proposition** (Furstenberg Weiss). *Let $Y \times_\rho G$ be an Abelian extension, and let M be the Mackey group of this extension. Let $f \in L^2(Y \times_\rho G)$ be s.t. for all $\chi \in M^\perp$,*

$$\int f(y,g)\chi(g)dm_G(g) = 0$$

for a.e $y \in Y$. Then f is orthogonal to the space of T_ρ invariant functions.

6.4. **Notation.** Denote $U_d = d$ dimensional unitary matrices, $C(U_d)$ the center of U_d, and $P : U_d \to \mathbb{P}U_d = U_d/C(U_d)$ the natural projection.

We need the following lemma:

6.5. **Lemma.** *Let H be a compact Abelian connected group, and $A : H \to U_d$ a measurable function. If $P \circ A$ is a homomorphism, then $A(H)$ is a commuting set of matrices.*

Proof. Let $g, h \in H$. Suppose $[A(h), A(g)] = \delta I$. If v is an eigenvector of $A(h)$ with eigenvalue γ, then

$$A(h)A(g)v = \delta A(g)A(h)v = \gamma\delta A(g)v$$

thus $A(g)v$ is an eigenvector of $A(h)$ with eigenvalue $\gamma\delta$. This implies that $A^k(g)v$ is an eigenvector of $A(h)$ with eigenvalue $\gamma\delta^k$, thus δ is a root of unity of order $\leq d$. Denote $C_{d!}$ the group of order $d!$ roots of unity. Then the commutator set

$$\{[A(h), A(g)]\}_{h,g\in H} \subset C_d,$$

Fix g. Denote

$$E(g,\delta) = \{h : [A(g), A(h)] = \delta\}$$

The set $E(g,1)$ is a closed subgroup of H since for $h_1, h_2 \in H$

$$\begin{aligned}[A(g), A(h_1 + h_2)] &= [A(g), cA(h_1)A(h_2)]\\ &= [A(g), A(h_1)][A(g), A(h_2)]\\ &= [A(g), A(h_1)][A(g), A(h_2)].\end{aligned}$$

Let $h \in E(g,\delta)$, then by the above computation

$$E(g,\delta) = h + E(g,1),$$

thus $E(g,1)$ is a closed subgroup of index $\leq d!$. As H is connected $E(g,1) = H$. $\square$

6.6. Theorem. *Let (Y,T) be an ergodic m.p.s, $Y \times_\rho H$ an ergodic extension by a connected Abelian group. Let $F : Y \times H \times H \to S^1$ be a measurable function. Let $\sigma_1(y,\varphi_1)$,$\sigma_2(y,\varphi_2) : Y \times H \to S^1$ be measurable functions. Suppose*

$$\sigma_1(y,\varphi_1)\sigma_2(y,\varphi_2) = \frac{F(Ty,\varphi_1+\rho(y),\varphi_2+\rho(y))}{F(y,\varphi_1,\varphi_2)}.$$

Then there exists measurable functions $h, H : Y \to S^1$ such that

$$\sigma_1(y,\varphi_1) = h(y)\frac{H(Ty,\varphi_1+\rho(y))}{H(y,\varphi_1)}$$

Proof. Construct the system $(X,S) = (Y \times H \times H \times S^1 \times S^1, S)$, where

$$S(y,\varphi_1,\varphi_2,\zeta_1,\zeta_2) = (Ty,\varphi_1+\rho(y),\varphi_2+\rho(y),\sigma_1(y,\varphi_1)\zeta_1,\sigma_2(y,\varphi_2)\zeta_2)$$

Denote

$$T_1(y,\varphi_1,\zeta_1) = (Ty,\varphi_1+\rho(y),\sigma_1(y,\varphi_1)\zeta_1)$$
$$T_2(y,\varphi_2,\zeta_2) = (Ty,\varphi_2+\rho(y),\sigma_2(y,\varphi_2)\zeta_2),$$

and

$$X_1 = (Y \times H \times S^1, T_1); \quad X_2 = (Y \times H \times S^1, T_2).$$

Then

$$(X,S) = (X_1,T_1) \times_Y (X_2,T_2)$$

is a conditional product joining relative to Y. The function

$$(2) \qquad \tilde{F}(y,\varphi_1,\varphi_2,\zeta_1,\zeta_2) = F(y,\varphi_1,\varphi_2)\zeta_1^{-1}\zeta_2^{-1}$$

is invariant under S, and therefore by theorem 4.4 it is measurable with respect to $\hat{Y_1} \times \hat{Y_2}$, where $\hat{Y_i}$ is the max isometric extension of Y in X_i for $i = 1,2$. Isometric extensions are spanned by finite rank modules (see 4.1) Thus

$$\tilde{F}(y,\varphi_1,\varphi_2,\zeta_1,\zeta_2) = \sum \left\langle \vec{\psi}_j^1\,(y,\varphi_1,\zeta_1), \vec{\psi}_j^2\,(y,\varphi_2,\zeta_2) \right\rangle$$

where

$$T_1\,\vec{\psi}_j^1\,(y,\varphi_1,\zeta_1) = H_j^1(y)\,\vec{\psi}_j^1\,(y,\varphi_1,\zeta_1)$$
$$T_2\,\vec{\psi}_j^2\,(y,\varphi_2,\zeta_i) = H_j^2(y)\,\vec{\psi}_j^2\,(y,\varphi_2,\zeta_2),$$

and $H_j^1(y), H_j^2(y)$ are $d_j \times d_j$ unitary matrices. Substituting the Fourier expansions:

$$\vec{\psi}_j^1\,(y,\varphi_1,\zeta_1) = \sum \vec{\psi}_{j,k}^1\,(y,\varphi_1)\zeta_1^k$$
$$\vec{\psi}_j^2\,(y,\varphi_2,\zeta_2) = \sum \vec{\psi}_{j,k}^2\,(y,\varphi_2)\zeta_2^k$$

in equation(2) we get that for $k = -1$ there exists j such that $\overrightarrow{\psi^1_{j,-1}} \neq 0$. Apply T_1 to get

$$\sigma_1^{-1}(y,\varphi_1)\ \overrightarrow{\psi^1_{j,-1}}\ (Ty,\varphi_1+\rho(y)) = H^1_j(y)\ \overrightarrow{\psi^1_{j,-1}}\ (y,\varphi_1)$$

For simplicity we drop the indices:

$$\sigma^{-1}(y,\varphi)\ \overrightarrow{\psi}\ (Ty,\varphi+\rho(y)) = H(y)\ \overrightarrow{\psi}\ (y,\varphi) \tag{3}$$

For each y consider the distribution of $\overrightarrow{\psi}\ (y,\varphi)$ in the fiber over y, and look at the vector space spanned by the support of this distribution. Call this $V(y)$, so that $V(y) \subset \mathbb{C}^d$, and $V(Ty) = H(y)V(y)$. Since $H(y)$ is unitary, $dimV(Ty) = dimV(y)$, thus by ergodicity $dimV(y) = \hat{d}$ for a.s. y. For each y choose a basis for $\mathbb{C}^d$ s.t. $V(y)$ is spanned by the first $\hat{d}$ elements. As the transformation matrix is a function of y, we may assume $d = \hat{d}$.

Denote $\overrightarrow{\tilde{\psi}}$ the projection on $\mathbb{P}V$, and $\tilde{H}$ the projection on $\mathbb{P}U_d$. Thus:

$$\overrightarrow{\tilde{\psi}}\ (Ty,\varphi+\rho(y)) = \tilde{H}(y)\ \overrightarrow{\tilde{\psi}}\ (y,\varphi).$$

Consider the group extension $(Y \times H) \times_{\tilde{H}} \mathbb{P}U_d$. Then

$$\overrightarrow{\tilde{\psi}}\ (T^n y,\varphi+\rho_n(y)) = \tilde{H}_n(y)\ \overrightarrow{\tilde{\psi}}\ (y,\varphi)$$

For fixed y, $\overrightarrow{\tilde{\psi}}\ (y,\varphi)$ spans the space, so whenever $(T^n y,\varphi+\rho_n(y))$ is close to (y,φ) (by ergodicity this happens for a generic y), $\tilde{H}_n(y)$ is close to the identity. This implies that the foregoing group extension is not ergodic, and furthermore - the Mackey group is trivial. Thus for some projective unitary matrix function $\tilde{M}$:

$$\tilde{M}(Ty,\varphi+\rho(y)) = \tilde{H}(y)\tilde{M}(y,\varphi) \tag{4}$$

Also for any φ'

$$\tilde{M}(Ty,\varphi+\varphi'+\rho(y)) = \tilde{H}(y)\tilde{M}(y,\varphi+\varphi').$$

Thus

$$\tilde{M}^{-1}(Ty,\varphi+\varphi'+\rho(y))\tilde{M}(Ty,\varphi+\rho(y)) = \tilde{M}^{-1}(y,\varphi+\varphi')\tilde{M}(y,\varphi)$$

By ergodicity

$$\tilde{M}^{-1}(y,\varphi+\varphi')\tilde{M}(y,\varphi) = \tilde{A}(\varphi'),$$

for all φ', a.e.(y,φ). By Fubini's theorem there exists φ_0 such that

$$\tilde{M}(y,\varphi) = \tilde{M}(y,\varphi_0)\tilde{A}^{-1}(\varphi-\varphi_0). \tag{5}$$

a.e.(y,φ). The function $\tilde{A}(\varphi')$ is a homomorphism of H:

$$\begin{aligned}
\tilde{A}(\varphi'+\varphi'') &= \tilde{M}^{-1}(y,\varphi+\varphi'+\varphi'')\tilde{M}(y,\varphi) \\
&= \tilde{M}^{-1}(y,\varphi+\varphi'+\varphi'')\tilde{M}(y,\varphi+\varphi')\tilde{M}^{-1}(y,\varphi+\varphi')\tilde{M}(y,\varphi) \\
&= \tilde{A}(\varphi'')\tilde{A}(\varphi')
\end{aligned}$$

Recall $P : U_d \to \mathbb{P}U_d$ is the natural projection. We can find a measurable function $A : H \to U_d$ so that $P \circ A = \tilde{A}$.

$$A(H) = P^{-1}\tilde{A}(H).$$

Then by lemma 6.5 $A(H)$ is a commuting set. Substituting equation (5) in equation (4) we get

$$\begin{aligned}\tilde{M}(Ty, \varphi + \rho(y)) =& \tilde{M}(Ty, \varphi_0)\tilde{A}^{-1}(\varphi + \rho(y) - \varphi_0)\\ =& \tilde{H}(y)\tilde{M}(y, \varphi)\\ =& \tilde{H}(y)\tilde{M}(y, \varphi_0)\tilde{A}^{-1}(\varphi - \varphi_0)\end{aligned}$$

Thus

$$\tilde{H}(y) = \tilde{M}(Ty, \varphi_0)\tilde{A}^{-1}(\rho(y))\tilde{M}^{-1}(y, \varphi_0)$$

or

$$H(y) = M(Ty, \varphi_0)A(-\rho(y))M^{-1}(y, \varphi_0)d(y) \tag{6}$$

where $d(y)$ is a scalar matrix. As $A(H)$ is a commuting set, it is simultaneously diagonalizable:

$$A(\varphi) = N^{-1}D(\varphi)N \tag{7}$$

therefore

$$H(y) = M(Ty, \varphi_0)N^{-1}D(-\rho(y))NM^{-1}(y, \varphi_0)d(y)$$

Denote $M'(y) = M(y, \varphi_0)$. Substitute $H(y)$ in equation (3):

$$\sigma(y, \varphi)NM'^{-1}(Ty)\overrightarrow{\psi}(Ty, \varphi + \rho(y)) = D(-\rho(y))d(y)NM'^{-1}(y)\overrightarrow{\psi}(y, \varphi)$$

Now each coordinate gives us the desired result. □

6.7. **Remark.** If H in theorem 6.6 is not necessarily connected, but the cocycle ρ is cohomologous to a constant: $\rho(y) = c\frac{f(Ty)}{f(y)}$, then the result holds as for some scalar matrix $d(y) : Y \to S^1$

$$A(\rho(y)) = A(cf(Ty)f^{-1}(y)) = A(f(Ty))A(c)A^{-1}(f(y))d(y)$$

Now diagonalize $A(c)$: $A(c) = HDH^{-1}$ and substitute in equation (6).

6.8. **Corollary.** *Let (Y,T) be an ergodic m.p.s, $Y \times_\rho H$ an ergodic Abelian extension where either H is connected or the cocycle ρ is cohomologous to a constant. Suppose there exists a family of functions $\{f_h\}_{h\in H}$, $f_h : Y \times H \to S^1$ such that*

$$\frac{\sigma(y, \varphi + h)}{\sigma(y, \varphi)} = \frac{f_h(Ty, \varphi + \rho(y))}{f_h(y, \varphi)},$$

then

$$\sigma(y, \varphi) = g(y)\frac{F(Ty, \varphi + \rho(y))}{F(y, \varphi)}.$$

Proof. Make the coordinate change: $\varphi_1 = \varphi$; $\varphi_2 = \varphi + h$. Then

$$f_h(y,\varphi) = f(y,h,\varphi) = f'(y,\varphi+h,\varphi) = f'(y,\varphi_1,\varphi_2)$$

and

$$f_h(Ty,\varphi+\rho(y)) = f'(Ty,\varphi_1+\rho(y),\varphi_2+\rho(y)).$$

Now apply theorem 6.6. □

6.9. **Notation.** We write $f \sim g$ if f/g =const.

6.10. **Lemma.** *Let $(X = Y \times_\rho H, T_\rho)$ be an Abelian extension of (T,Y). Let $\sigma : Y \times H \to S^1$ be such that for all $h \in H$ there exists a function $f_h : Y \times H \to S^1$ such that*

$$\frac{\sigma(y,\varphi+h)}{\sigma(y,\varphi)} = \lambda_h \frac{f_h(Ty,\varphi+\rho(y))}{f_h(y,\varphi)}, \tag{8}$$

and the functions $\{f_h\}_{h\in H}$ depend measurably on h. Then there exists a measurable family of functions $\{f_h\}_{h\in H}$ satisfying the above equation, and a a neighborhood U of zero in H such that

$$f_{h_1+h_2}(y,\varphi) \sim f_{h_2}(y,\varphi+h_1) f_{h_1}(y,\varphi)$$
$$\lambda_{h_1+h_2} = \lambda_{h_1}\lambda_{h_2}$$

for $h_1, h_2 \in U$.

Proof. For measurability see remark 6.17. Using equation (8) we get

$$\begin{aligned}\frac{\sigma(y,\varphi+h_1+h_2)}{\sigma(y,\varphi)} &= \lambda_{h_1+h_2}\frac{Tf_{h_1+h_2}(y,\varphi)}{f_{h_1+h_2}(y,\varphi)}\\ &= \lambda_{h_1}\lambda_{h_2}\frac{Tf_{h_1}(y,\varphi+h_2)}{f_{h_1}(y,\varphi+h_2)}\frac{Tf_{h_2}(y,\varphi)}{f_{h_2}(y,\varphi)}\end{aligned}$$

this implies that

$$\frac{f_{h_1+h_2}(y,\varphi)}{f_{h_1}(y,\varphi+h_2)f_{h_2}(y,\varphi)}$$

is an eigenfunction of T and that

$$\frac{\lambda_{h_1}\lambda_{h_2}}{\lambda_{h_1+h_2}}$$

is an eigenvalue. Let (Z,α) be the Kronecker factor of (X,T_ρ), $\pi : X \to Z$ the projection map, let N parameterize $\hat{Z}$, and let $\psi_{N(h_1,h_2)}(z)$ be a character of Z s.t.:

$$\frac{f_{h_1+h_2}(y,\varphi)}{f_{h_2}(y,\varphi+h_1)f_{h_1}(y,\varphi)} \sim \psi_{N(h_1,h_2)} \circ \pi(y,\varphi) \tag{9}$$

and

$$\frac{\lambda_{h_1}\lambda_{h_2}}{\lambda_{h_1+h_2}} = \psi_{N(h_1,h_2)}(\alpha) \tag{10}$$

We now show that $\psi_{N(h_1,h_2)}$ satisfies a 2-cocycle equation:

$$\psi_{N(h_1+h_2,h_3)}\circ\pi(y,\varphi)\sim\frac{f_{h_1+h_2+h_3}(y,\varphi)}{f_{h_3}(y,\varphi+h_1+h_2)f_{h_1+h_2}(y,\varphi)}$$
$$\psi_{N(h_1,h_2+h_3)}\circ\pi(y,\varphi)\sim\frac{f_{h_1+h_2+h_3}(y,\varphi)}{f_{h_2+h_3}(y,\varphi+h_1)f_{h_1}(y,\varphi)}$$

Thus

$$\psi_{N(h_1+h_2,h_3)}\circ\pi(y,\varphi)f_{h_3}(y,\varphi+h_1+h_2)f_{h_1+h_2}(y,\varphi)$$
$$\sim\psi_{N(h_1,h_2+h_3)}\circ\pi(y,\varphi)f_{h_2+h_3}(y,\varphi+h_1)f_{h_1}(y,\varphi)$$

Dividing both sides by

$$f_{h_1}(y,\varphi)f_{h_3}(y,\varphi+h_1+h_2)f_{h_2}(y,\varphi+h_1)$$

we get

$$\psi_{N(h_1+h_2,h_3)}\circ\pi(y,\varphi)\frac{f_{h_1+h_2}(y,\varphi)}{f_{h_1}(y,\varphi)f_{h_2}(y,\varphi+h_1)}$$
$$\sim\psi_{N(h_1,h_2+h_3)}\circ\pi(y.\varphi)\frac{f_{h_2+h_3}(y,\varphi+h_1)}{f_{h_2}(y,\varphi+h_1)f_{h_3}(y,\varphi+h_1+h_2)}.$$

Combining the above equation with equation (9),

$$\psi_{N(h_1+h_2,h_3)}(z)\psi_{N(h_1,h_2)}(z)=\psi_{N(h_1,h_2+h_3)}(z)\psi_{N(h_2,h_3)}(z). \tag{11}$$

As $h\to f_h$ is a measurable function, $f_{h_2}(y), f_{h_2+h_1}(y)$ are close in measure for small h_1, most h_2, and the same goes for $f_{h_2}(y,\varphi), f_{h_2}(y,\varphi+h)$. Therefore the expression in equation (9) is close (in measure) to $\bar{f}_{h_1}(y,\varphi)$. But $N_1\neq N_2$ implies

$$\|\psi_{N_1}-\psi_{N_2}\|_2=\sqrt{2},$$

thus by equation (9), $\psi_{N(h_1,h_2)}=\psi_{\tilde{N}(h_1)}$ for $h_1\in U'$ a neighborhood of zero in H, $h_2\in A$ a set of positive measure. The set $A-A$ contains a neighborhood of zero U''. Let $U=U'\cap U''$. Take any $h_1,h_2,h_1+h_2\in U$, and find an element $h_3\in A$ such that $h_3+h_2\in A$, then by (11)

$$\psi_{N(h_1,h_2)}=\psi_{\tilde{N}(h_1)}\psi^{-1}_{\tilde{N}(h_1+h_2)}\psi_{\tilde{N}(h_2)}$$

For $h_1,h_2\in U$, denote

$$\tilde{f}_h(y,\varphi)=\psi_{\tilde{N}(h)}\circ\pi(y,\varphi)f_h(y,\varphi)$$

, and

$$\tilde{\lambda}_h=\lambda_h\psi^{-1}_{\tilde{N}(h)}(\alpha).$$

By equations (9), if $h_1,h_2,h_1+h_2\in U$, then :

$$\tilde{f}_{h_1+h_2}(y,\varphi)\sim\tilde{f}_{h_2}(y,\varphi+h_1)\tilde{f}_{h_1}(y,\varphi). \tag{12}$$

By equation(10), if $h_1,h_2,h_1+h_2\in U$ then

$$\tilde{\lambda}_{h_1+h_2}=\tilde{\lambda}_{h_1}\tilde{\lambda}_{h_2}. \tag{13}$$

□

6.11. Lemma. *Let H be a torus (possibly infinite dimensional) and let $Y \times_\rho H$ be an Abelian extension of Y. Suppose*

$$\frac{\sigma(y,\varphi+h)}{\sigma(y,\varphi)} = \lambda_h \frac{f_h(Ty,\varphi+\rho(y))}{f_h(y,\varphi)}. \tag{14}$$

Then there is a subgroup $H_2 < H$ such that $H/H_2 = \mathbb{T}^n$ such that if $\pi : H \to H/H_2$ is the natural projection then there exists a function $\tilde{\sigma} : Y \times H/H_2 \to S^1$ such that

$$\sigma(y,\varphi) = \tilde{\sigma}(y,\pi(\varphi)) \frac{T_\rho F(y,\varphi)}{F(y,\varphi)}.$$

Proof. By lemma 6.10 the functions f_h can be chosen such that λ_h is multiplicative in U a zero neighborhood in H. The neighborhood U contains H_1 - a closed connected subgroup of H, such that $H/H_1 = \mathbb{T}^l$, thus λ_h is a character of H_1. Thinking of H as $H/H_1 \times H_1$ with coordinates (h_0, h_1) the above equation becomes

$$\frac{\sigma(y,h_0,h_1+h)}{\sigma(y,\varphi)} = \lambda_h \frac{f_h(Ty,\varphi+\rho(y))}{f_h(y,\varphi)}. \tag{15}$$

where $h \in H_1$. This is the same as

$$\frac{\lambda_{h_1+h}^{-1}\sigma(y,h_0,h_1+h)}{\lambda_{h_1}^{-1}\sigma(y,\varphi)} = \frac{f_h(Ty,\varphi+\rho(y))}{f_h(y,\varphi)}. \tag{16}$$

Applying corollary 6.8 replacing Y with $(Y \times H/H_1)$ and H with H_1 we get

$$\lambda_{h_1}^{-1}\sigma(y,h_0,h_1) = \tilde{\sigma}(y,h_0) \frac{T_\rho F(y,\varphi)}{F(y,\varphi)}.$$

or

$$\sigma(y,h_0,h_1) = \lambda_{h_1}\tilde{\sigma}(y,h_0) \frac{T_\rho F(y,\varphi)}{F(y,\varphi)}.$$

Now for h_1 in the kernel of λ we have $\lambda_{h_1} = 1$. The image of λ is S^1 thus if $ker\lambda$ is H_2 then $H/H_2 = \mathbb{T}^{l+1}$. ☐

6.12. Remark. If H is any connected compact Abelian group (not necessarily a torus) then H_1 in the foregoing proof is not necessarily connected. By the same proof we will get that σ is cohomologous to a cocycle lifted from a product of a finite torus and a totally disconnected compact Abelian group.

6.13. Lemma. *Let and $Y \times_\rho H$ be an Abelian extension of Y with $\rho(y)$ cohomologous to a constant function (now H is any compact Abelian group). Let σ be as in lemma 6.11. Then there is a subgroup $H_2 < H$, and a finite group C_k, such that $H/H_2 = \mathbb{T}^n \times C_k$ and if $\pi : H \to H/H_2$ is the natural projection then there exists a function $\tilde{\sigma} : Y \times H/H_2 \to S^1$ such that*

$$\sigma(y,\varphi) = \tilde{\sigma}(y,\pi(\varphi)) \frac{T_\rho F(y,\varphi)}{F(y,\varphi)}.$$

Proof. By lemma 6.10 the functions f_h can be chosen such that λ_h is multiplicative in U a zero neighborhood in H. The neighborhood U contains H_1 - a closed subgroup of H, such that $H/H_1 = \mathbb{T}^l \times C_j$ where C_j is a finite group, thus λ_h is a character of H_1. Now proceed as in lemma 6.11 (the image of λ is either S^1 or a finite group).. ☐

6.14. **Lemma.** *Let* $Y = Z \times_\rho H$ *be an ergodic Abelian extension, and* $F : Z \times H \to S^1$, $g : Z \to S^1$ *measurable functions such that*

$$g(\theta) = \frac{TF(y)}{F(y)}.$$

Then there exists $\chi \in \hat{H}$, *and* $k : Z \to S^1$ *such that*

$$F(\theta, \varphi) = k(\theta)\chi(\varphi).$$

Proof. Take the Fourier expansion of F:

$$F(\theta, \varphi) = \sum k_i(\theta)\chi_i(\varphi).$$

Then for all i

$$k_i(\theta + \alpha)\chi_i(\varphi)\chi_i(\rho(\theta)) = g(\theta)k_i(\theta)\chi_i(\varphi).$$

Ergodicity of T implies $|k_i(\theta)|$ is constant a.e.. The fact that $|F| = 1$ implies that there exist an i for which $|k_i(\theta)| \neq 0$. If there are two such indices i, j, then

$$\frac{\chi_i}{\chi_j}(\rho(\theta))$$

is a coboundary. As T is ergodic $\chi_i = \chi_j$. ☐

6.15. **Notation.** Let $(X_1, \mathcal{B}_1)$, $(X_2, \mathcal{B}_2)$ be measure spaces. Denote

$$B(X_1, X_2) = \{f : X_1 \to X_2, f \text{ measurable}\}.$$

6.16. **Proposition.** *Let* $Y = Z \times_\rho H$ *be an ergodic Abelian extension* (X, μ) *a measure space, and* $x \to f_x(y)$ *be a Borel measurable function from* X *to* $B(Y, S^1)$. *Suppose for all* x *there are functions* $g_x(\theta), F_x(y) \in B(Y, S^1)$ *such that*

$$f_x(y) = g_x(\theta)\frac{TF_x(y)}{F_x(y)}. \tag{17}$$

Then there is an x *measurable choice of* $g_x(\theta), F_x(y)$.

Proof. Endowed with the L^2 topology, $B(Y, S^1)$ is a polish group. Let $B(Z, S^1)$ be the closed subgroup of $B(Y, S^1)$ of functions that depend only on the θ coordinate, and let $f \to \bar{f}$ be the natural projection onto $\bar{B} = B(Y, S^1)/B(Z, S^1)$, with the induced topology. By Dixmier ([3] theorem 1.2.4) there is a measurable section $\bar{B} \to B$. Equation 17 implies

$$\bar{f}_x(y) = \frac{T\bar{F}_x(y)}{\bar{F}_x(y)}.$$

Define $\varphi : \bar{B} \to \bar{B}$

$$\varphi(\bar{f}) = \frac{T\bar{f}}{\bar{f}}.$$

If $\varphi(\bar{f}) = \varphi(\bar{g})$, then

$$\frac{T\frac{f}{g}(y)}{\frac{f}{g}(y)} = h(\theta).$$

By 6.14 this implies that up to multiplication by a function of θ, $\frac{f}{g}$ belongs to a countable set, Thus φ is countable to one. By Lusin [19] $\varphi(\bar{B})$ is a measurable set and there is a measurable function $\psi : \varphi(\bar{B}) \to \bar{B}$ s.t.

$$\varphi \circ \psi = Id|_{\varphi(\bar{B})}$$

Now if

$$\psi(\bar{f}_x) = \bar{F}_x,$$

then

$$\bar{f}_x = \varphi \circ \psi(\bar{f}_x) = \frac{T\bar{F}_x}{\bar{F}_x}$$

the combination

$$x \to f_x \to \bar{f}_x \to \bar{F}_x \to F_x$$

gives a measurable choice of F_x, and g_x is measurable as a quotient of measurable functions. □

6.17. **Remark.** If $g_x(\theta) \in B(Y, *)$ (g_x is constant) then the same proof works to give a measurable choice of g_x, F_x.

7. NILSYSTEMS

7.1. **Definition.** Let N be a group. Denote $N^0 = N$, $N^k = [N^{k-1}, N]$. N is k -step nilpotent if $N^{k+1} = 1$. Let N be a k step nilpotent group acting transitively on a measure space X preserving a measure μ. For $a \in N$ let T be the transformation $Tx = ax$. Then (X, T) is a k-step *nilsystem.* If N is a Lie group, Γ a lattice then N acts transitively on N/Γ preserving a unique measure μ. We call N/Γ a *nilmanifold* and $(N/\Gamma, T)$ a *nilflow.*

7.2. **Definition.** Let $(X, \mathcal{B}, \mu, T)$ be a m.p.s. Let $\mathcal{A} \subset \mathcal{B}$ be a T invariant sub σ-algebra. If

$$\mathcal{F} = \{f : f \text{ measurable}, \ |f| = 1, \ Tf/f \text{ is } \mathcal{A} \text{ measurable}\}.$$

We define $D(\mathcal{A})$ as the smallest σ-algebra with respect to which the functions of $\mathcal{F}$ are measurable, and define $D_n(\mathcal{A}) = D(D_{n-1}(\mathcal{A}))$ where $D_0(\mathcal{A}) = D(\mathcal{A})$. T is said to have *generalized discrete spectrum [mod $\mathcal{A}$] of finite type* if for some $n \in \mathbb{N}$, $D_n(\mathcal{N})$ $[D_n(\mathcal{A})]$ is $\mathcal{B}$, where $\mathcal{N}$ is the trivial σ -algebra of null sets and their complements. Since $D_n(\mathcal{N}) \subset D_n(\mathcal{A})$ generalized discrete spectrum of finite type implies generalized discrete spectrum mod $\mathcal{A}$ of finite type. The qualification "generalized" is dropped when $n = 1$.

7.3. **Example.** If $(Z \times H, T_\rho)$ is an Abelian extension of the Kronecker system (Z, T), then T_ρ has discrete spectrum of finite type mod the Kronecker algebra, and generalized discrete spectrum of finite type (mod the trivial algebra). Another example - if $(N/\Gamma, T)$ is a nilflow then T has generalized discrete spectrum of finite type.

7.4. **Proposition** (Parry [25]). *If T is ergodic with discrete spectrum mod $\mathcal{A}$ then there exists a compact Abelian group G of measure preserving transformations such that $T(gx) = gTx$ for $g \in G$ and $\mathcal{A} = \{B \in \mathcal{B} : gB = B \ \forall g \in G\}$.*

8. CHARACTERISTIC FACTORS

8.1. **Definition.** We say that (Y, D, ν, T) is a *characteristic factor for schemes of length k* if (Y, D, ν, T) is a factor of (X, B, μ, T), and for any distinct $a_1, \dots, a_k \in \mathbb{Z}$.

$$\frac{1}{N}\sum_{n=1}^{N} T^{a_1 n} f_1 \dots T^{a_k n} f_k - \frac{1}{N}\sum_{n=1}^{N} T^{a_n} E(f_1|Y) \dots T^{a_k n} E(f_k|Y) \longrightarrow 0.$$

We say that (Y, D, ν, T) is a *minimal characteristic factor (m.c.f) for schemes of length k* if it is a factor of any other characteristic factor for schemes of length k.

8.2. **Lemma.** *Let Y_1, Y_2 be characteristic factors of X for schemes of length k*

Then there exists a characteristic factor of X, which is a factor of both Y_1, Y_2.

Proof. Denote P, Q the orthogonal projections onto $L^2(Y_1), L^2(Y_2)$ (seen as subspaces of $L^2(X)$) respectfully. Then $P^2 = P^* = P$ (same for Q). We show that $(PQP)^n$ strongly converges to a self adjoint operator projection W: P is a projection thus $P \leq I$.

$$\left\langle (PQP)^2 x, x \right\rangle = \langle PQPx, QPx \rangle \leq \langle QPx, QPx \rangle = \langle PQPx, x \rangle,$$

Inductively, the sequence $(PQP)^n$ is a decreasing sequence of operators, thus $\langle (PQP)^n x, x \rangle$ converges for all x. The sequence $(PQP)^n x$ is a Cauchy sequence as

$$\begin{aligned} ||(PQP)^n x - (PQP)^m x||^2 = &\left\langle (PQP)^{2n} x, x \right\rangle + \left\langle (PQP)^{2m} x, x \right\rangle \\ &- 2\left\langle (PQP)^{(n+m)} x, x \right\rangle \to 0. \end{aligned}$$

Let $W = \lim_{n\to\infty} (PQP)^n$, then $W^2 = W = W^*$. If $Wx = x$ then $Px = PWx = Wx = x$, and

$$PQx = PQPx = PQPWx = Wx = x \Rightarrow Qx = x.$$

$W(L^2(X))$ is a characteristic factor as for all m:

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}T^{a_1n}f_1\dots T^{a_kn}f_k$$
$$=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}T^{a_1n}((PQP)^m f_1)\dots T^{a_kn}((PQP)^m f_k).$$

□

8.3. **Corollary.** *There exists a unique minimal characteristic factor for all schemes of length k.*

Proof. By Zorn's lemma. □

The advantage of the definition of a characteristic factor for all schemes of length k is that it is natural in the sense that any morphism of measure preserving systems induces a morphism between their minimal characteristic factors for schemes of length k - as will be shown in corollary 8.5 (this may also be true for characteristic factors of a specific scheme).

8.4. **Lemma.** *Let V be the σ-algebra spanned by the partial L^2 limits of the sequences $\{\frac{1}{N}\sum_{n=1}^{N}T^{a_1n}f_1\dots T^{a_kn}f_k\}$, where f_i are bounded functions, $a_1,\dots,a_k\in\mathbb{Z}$. Then V is the minimal characteristic factor for schemes of length k.*

Proof. Obviously V is a subspace of the minimal characteristic factor. We must show that V is a characteristic factor: Let $g\perp V$, then for any f_1

$$\left\langle g,\frac{1}{N}\sum_{n=1}^{N}T^{a_1n}f_1\dots T^{a_kn}f_k\right\rangle=\frac{1}{N}\sum_{n=1}^{N}\int gT^{a_1n}f_1T^{a_2n}f_2\dots T^{a_kn}f_kd\mu$$
$$=\frac{1}{N}\sum_{n=1}^{N}\int f_1T^{-a_1n}g\dots T^{(a_k-a_1)n}f_kd\mu$$
$$\xrightarrow{N\to\infty}0.$$

□

8.5. **Corollary.** *If $\pi:(X,\mathcal{B},\mu)\to(Y,\mathcal{D},\nu)$ is a factor map, and W_X,W_Y are minimal for averaging schemes of length k for X ,Y respectfully, then π induces a map between W_X and W_Y.*

9. THE CONZE-LESIGNE FACTOR

9.1. **Definition.** A *CL extension* of a Kronecker system (Z,α) (or a *CL system*) is an ergodic system (Y,T) where $Y=Z\times_\rho H$ for H a compact metrizable Abelian group, $T(\theta,\varphi)=(\theta+\alpha,\varphi+\rho(\theta))$, and the extension cocycle ρ satisfies the (additive) CL - equation: there exists a family of functions $\{\rho_\beta\}$ such that for all $\beta\in Z$

$$\rho(\theta+\beta)-\rho(\theta)=h(\alpha,\beta)+\rho_\beta(\theta+\alpha)-\rho_\beta(\theta) \tag{18}$$

for a.e. $\theta \in Z$. We call ρ a *CL character.* If Z is the Kronecker factor of $Z \times_\rho H$ then the extension is called *weakly mixing.*

9.2. **Remark.** The multiplicative form of the above equation 18 is known as the Conze-Lesigne equation: Let χ be a character of H, the CL equation implies:

$$\frac{\chi \circ \rho(\theta+\beta)}{\chi \circ \rho(\theta)} = \chi(h(\alpha,\beta)) \frac{\chi \circ \rho_\beta(\theta+\alpha)}{\chi \circ \rho_\beta(\theta)}.$$

9.3. **Lemma.** *If $Z \times_\rho H$ is a CL extension then either the extension is weakly mixing, or for some $\chi \neq 1 \in \hat{H}$, $\chi \circ \rho$ is equivalent to a constant.*

Proof. [12] lemma 4. ☐

9.4. **Corollary.** *If $Z \times_\rho H$ is a weakly mixing CL extension then $\hat{H}$ is torsion free so $kH = H$ for every non zero integer.*

Proof. [12] lemma 5. ☐

9.5. **Corollary.** *If $Z \times_\rho H$ is a weakly mixing CL extension then H is a compact connected Abelian group.*

9.6. **Proposition.** *Let $Z \times_\rho H$ be a weakly mixing CL extension. If χ is a character of H, then $\chi \circ \rho$ is lifted from a function on a finite dimensional toral factor $\mathbb{T}^n$ of Z. Denote $p : Z \to \mathbb{T}^n$ the projection map. There exists a unique matrix $M \in M_n(\mathbb{Z})$, $N \in \mathbb{Z}^n$ (at least one of them not zero), such that for β in a neighborhood of zero in Z*

$$\frac{\chi \circ \rho(\theta+\beta)}{\chi \circ \rho(\theta)} = e^{2\pi i (Mp(\alpha)\cdot p(\beta) + N \cdot p(\beta))} \frac{\chi \circ \rho_\beta(p(\theta+\alpha))}{\chi \circ \rho_\beta(p(\theta))}.$$

Proof. The fact that $\chi \circ \rho$ is lifted from a function on a finite dimensional toral factor $\mathbb{T}^n$ of Z is shown in [27] corollary 5.5. The existence of M, N are shown in [27], and in [18] lemma 4. ☐

9.7. **Lemma.** *If for all $\gamma \in Z$,*

$$\frac{\chi \circ \rho(\theta+\gamma)}{\chi \circ \rho(\theta)}$$

is a CL cocycle then $\chi \circ \rho$ is a CL cocycle.

Proof. [12] proposition 7. ☐

9.8. **Proposition.** *Let l be a positive integer and let $a_1, \ldots, a_k$ be distinct non zero integers. Let $\rho^{(1)}, \ldots, \rho^{(l)}$ be multiplicative cocycles on Z. Suppose*

$$\prod_{k=1}^{l} \rho^{(k)}_{a_k}(\theta + a_k\beta) = \frac{F(z + a_1(\beta+\alpha), \ldots, z + a_l(\beta+\alpha))}{F(z + a_1\beta, \ldots, z + a_l\beta)},$$

(recall that $\rho^{(k)}_j(\theta) = \rho^{(k)}(\theta + (j-1)\alpha) \ldots \rho^{(k)}(\theta+\alpha)\rho^{(k)}(\theta)$) then $\rho^{(k)}$ is a CL cocycle for $1 \leq k \leq l$.

Proof. [12] proposition 10 (the notation there is reversed). ☐

9.9. **Definition.** Let $Z \times_\rho H$ be a CL extension and denote by $B(Z,H)$ the measurable functions $Z \to H$.

$$\mathcal{G} := \{(\beta, f) : \beta \in Z; f \in B(Z,H)\}$$

Then $\mathcal{G}$ is a group under the multiplication:

$$(\beta, f)(\beta', f') = (\beta + \beta', f^{\beta'} + f'),$$

where $f^{\beta'}(z) = f(z + \beta')$. Endow $B(Z,H)$ with the topology:

$$f_n \to f \iff \forall \chi \in \hat{H} : \chi \circ f_n \overset{L^2(Z)}{\longrightarrow} \chi \circ f.$$

Then $\mathcal{G} \subset Z \times B(Z,H)$ is a polish group, and acts on $Y = Z \times H$ by

$$(\beta, f)(\theta, \varphi) = (\theta + \beta, \varphi + f(\theta)).$$

Denote $[u,v] = u^{-1}v^{-1}uv$. Let $u_\alpha = (\alpha, \rho)$ (the element that represents the action of T). Define

$$\mathcal{U} := \{u \in \mathcal{G} : [u, u_\alpha] = (0, const)\}. \tag{19}$$

Let $C(\mathcal{U})$ be the center of $\mathcal{U}$, then

$$\{(0,h);\ h \in H\} = C(\mathcal{U}).$$

As a subgroup of $\mathcal{G}, \mathcal{U}$ acts on $Z \times H$, and we have:

9.10. **Lemma.** *The group $\mathcal{U}$ is a 2-step nilpotent group and acts transitively on $Y = Z \times H$.*

Proof. [27] theorem 3.8, or [21]. □

9.11. **Notation.**
We use T_u for the action of the element u.
We denote V_h (vertical rotation) the action of the element $(0,h)$.

9.12. **Lemma.** *Let $u = (\beta, \rho_\beta)$, $v = (\gamma, \rho_\gamma)$. there exists a constant $h(\beta,\gamma)$, such that*

$$[T_u, T_v](\theta, \varphi) = (\theta, \varphi + h(\beta,\gamma)) = V_{h(\beta,\gamma)}(\theta, \varphi).$$

9.13. **Lemma.** *Let $Y = Z \times_\rho H$ be a CL system. Then (Y,T) is isomorphic to a 2-step pro-nilsystem $lim_\leftarrow N_i/\Gamma_i$. Denote $N := lim_\leftarrow N_i$. Let G be a compact Abelian group of measure preserving transformations acting on Y and commuting with the action of T. Then G commutes with the action of N.*

Proof. Let $Y = Z \times_\rho H$ be a CL system. We may assume that Z is the Kronecker factor of the system. Let $\pi : Y \to Y'$ be a factor map. Let Z' be the Kronecker factor of Y'. Then π induces a map between Z and Z'. In particular, any transformation $T_g : Y \to Y$ that commutes with T induces a map from $Z \to Z$: $\theta \to \theta + \alpha_g$. We first show that T_g respects the skew product structure, i.e. that $T_g(\theta, \varphi) = (\theta + \alpha_g, \varphi + f_g(\theta)$ (a priori f_g is a function of θ and φ). For $g \in G$:

$$TT_g(\theta, \varphi) = T(\theta + \alpha_g, f_g(\theta, \varphi)) = (\theta + \alpha_g + \alpha, \rho(\theta + \alpha_g) + f_g(\theta, \varphi)).$$

On the other hand,

$$T_gT(\theta,\varphi) = T_g(\theta+\alpha,\varphi+\rho(\theta))) = (\theta+\alpha+\alpha_g, f_g(T(\theta,\varphi))).$$

Thus using the CL equation

$$\begin{aligned} f_g(T(\theta,\varphi)) =& \rho(\theta+\alpha_g) + f_g(\theta,\varphi) \\ =& \rho(\theta) + h(\alpha,\alpha_g) + \rho_{\alpha_g}(\theta+\alpha) - \rho_{\alpha_g}(\theta) + f_g(\theta,\varphi). \end{aligned}$$

Denote

$$\psi_g(\theta,\varphi) = f_g(\theta,\varphi) - \rho_{\alpha_g}(\theta) - \varphi,$$

then ψ_g is an eigenfunction of T and is thus a function of θ. Therefore

$$f_g(\theta,\varphi) = \varphi + \rho_{\alpha_g}(\theta) + \psi_g(\theta).$$

Denote

$$f_g(\theta) = \rho_{\alpha_g}(\theta) + \psi_g(\theta),$$

then

$$T_g(\theta,\varphi) = (\theta+\alpha_g, \varphi + f_g(\theta)).$$

Denote $g = (\alpha_g, f_g)$. Then the fact that T and T_g commute implies that $g \in \mathcal{U}$ (by the definition of $\mathcal{U}$). Since $\mathcal{U}$ is 2-step nilpotent, for any $u \in \mathcal{U}$, there is an element $\lambda(u,g) \in H$

$$[u,g] = (0, \lambda(u,g)).$$

This is a multiplicative function from $U \to \hat{G}$ thus trivial on U_0 - the connected component of the identity. But $\lambda(u_\alpha,\alpha_g) = 1$, and u_α, U_0 generate a dense subgroup of $\mathcal{U}$ (as the action of T is ergodic) thus $\lambda(u,\alpha_g) \equiv 1$. □

9.14. **Proposition.** *Any factor of a CL system is a CL system.*

Proof. Let $(Y,\mathcal{D},\mu,T)$ be a CL system, $(Y',\mathcal{D}',\mu',T)$ a factor. The system $(Y,\mathcal{D},\mu,T)$ has generalized discrete spectrum mod $\mathcal{D}'$ of finite type (see 7.2). Now apply the foregoing lemma, proposition 7.4 and induction. □

9.15. **Theorem.** *[HK] Let $(Y,\mathcal{D},\mu,T)$ be an ergodic m.p.s. then the maximal CL extension is a characteristic factor for length 3 schemes. Furthermore for any scheme (a_1,a_2,a_3) ($a_i \in \mathbb{Z}$) the limit*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f_1(T^{a_1n}y)f_2(T^{a_2n}y)f_3(T^{a_3n}y)$$

exists in $L^2(Y)$.

9.16. **Corollary.** *Let $(Y,\mathcal{D},\mu,T)$ be an ergodic m.p.s. then the minimal characteristic factor for schemes of length 3 is a 2-step pro-nilsystem.*

Part 2. The minimal characteristic factor for linear schemes of length 4

Let (X,B,μ,T) be an ergodic m.p.s., $f_1,\dots,f_k \in L^\infty(X)$, $a_1,\dots,a_k \in \mathbb{Z}$. We wish to study the limiting behavior of non-conventional ergodic averages of the kind

$$\frac{1}{N}\sum_{n=1}^{N} T^{a_1 n} f_1 \dots T^{a_k n} f_k.$$

9.17. By the ergodic theorem, the minimal characteristic factor for schemes of length 1 is an inverse limit of cyclic groups (if T is totally ergodic it's just a point). It is not difficult to show that the minimal characteristic factor for the schemes of length 2 is the Kronecker factor - an Abelian group rotation factor ([10]).

9.18. Schemes of length 3 are already much more difficult to analyze. Existence of the limits in the totally ergodic case was shown by Conze and Lesigne ([5],[6],[7],[15],[16],[18]), and in the general case by Furstenberg and Weiss (in the process of analyzing related averages) [10], and by Host and Kra ([11],[12]).

We prove the following theorem:

9.19. Theorem. *The minimal characteristic factor for schemes of length* 4 *is a* 3*-step pro-nilsystem.*

Our analysis seems to carry over for any k; the main problem is notational. We intend to work it out in the near future.

10. REDUCTION TO AN ABELIAN EXTENSION OF THE CL FACTOR

10.1. To avoid cumbersome notation we perform the analysis for the scheme $(1,2,3,4)$. We start with a series of reductions that will eventually leave us with handling an Abelian extension of the CL factor of the system. We follow the lines of the work of Furstenberg and Weiss [10].

10.2. *Reduction to an isometric extension of the CL factor* Our first step is to reduce the question to the case where X is the maximal isometric extension of the CL factor. The following lemma is known as the van der Corput lemma. The formulation below is due to Bergelson [2].:

10.3. Lemma (van der Corput). *Let $\{u_n\}$ be a bounded sequence of vectors in a Hilbert space $\mathcal{H}$. Assume that for each m the limit*

$$\gamma_m := \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \langle u_n, u_{n+m}\rangle$$

exists, and

$$\lim_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} \gamma_m = 0. \tag{20}$$

Then

$$\frac{1}{N}\sum_{n=1}^{N} u_n \xrightarrow{\mathcal{H}} 0.$$

Proof. Let M be large enough so that the expression in (20) is small. Let N be large enough with respect to M so that the two expressions

$$\frac{1}{NM}\sum_{n=1}^{N}\sum_{m=1}^{M} u_{n+m}, \qquad \frac{1}{N}\sum_{n=1}^{N} u_n$$

are close. We have:

$$\begin{aligned}\|\frac{1}{NM}\sum_{n=1}^{N}\sum_{m=1}^{M} u_{n+m}\|^2 &\leq \frac{1}{N}\sum_{n=1}^{N}\|\frac{1}{M}\sum_{m=1}^{M} u_{n+m}\|^2\\ &= \frac{1}{NM^2}\sum_{n=1}^{N}\sum_{m_1,m_2=1}^{M}\langle u_{n+m_1}, u_{n+m_2}\rangle\\ &\xrightarrow{N\to\infty} \frac{1}{M^2}\sum_{m_1,m_2=1}^{M}\gamma_{m_2-m_1}\end{aligned}$$

which is small. □

10.4. **Definition.** Let Y be the CL factor of the system. Define the measure μ_Y^* on Y^4 by:

$$\int\prod_{k=1}^{4} g_k(y_k)d\mu_Y^*(y_1,y_2,y_3,y_4) := \lim\frac{1}{N}\sum_{n=1}^{N}\int\prod_{k=1}^{4} T^{kn}g_k(y)d\nu(y)$$

(the limit exists by Theorem 9.15).

Define a measure μ^* on X^4 by:

$$\int\prod_{k=1}^{4} f_k(x_k)d\mu^* := \int\prod_{k=1}^{4} E(f_k|Y)(y_k)d\mu_Y^*.$$

Then μ^* defines a joining of $(X,T),\ldots,(X,T^4)$ which is the conditional product joining (4.3) relative to (Y^4,μ_Y^*).

10.5. **Remark.** Both measures μ_Y^* and μ^* are invariant under $T\times T\times T\times T$ and $T\times T^2\times T^3\times T^4$.

10.6. **Theorem.** *Let* $\hat{Y}$ *be the maximal isometric extension of* Y *in* X*. Then* $\hat{Y}$ *is a characteristic factor for the scheme* $(n,2n,3n,4n)$*.*

Proof. We apply the van der Corput lemma 10.3 with

$$u_n = \prod_{k=1}^{4} T^{kn}f_k(x).$$

We calculate γ_m:

$$\begin{aligned}\gamma_m &= \lim \frac{1}{N}\sum_{n=1}^{N}\langle u_n, u_{n+m}\rangle \\ &= \lim \frac{1}{N}\sum_{n=1}^{N}\int\prod_{k=1}^{4} T^{kn} f_k(x) T^{kn+km} f_k(x) d\mu(x) \\ &= \lim \frac{1}{N}\sum_{n=1}^{N}\int\prod_{k=1}^{4} T^{kn}(f_k T^{km} f_k(x)) d\mu(y) \\ &= \lim \frac{1}{N}\sum_{n=1}^{N}\int\prod_{k=1}^{4} T^{kn} E(f_k T^{km} f_k|Y)(y) d\nu(y) \\ &= \int\prod_{k=1}^{4} E(f_k T^{km} f_k|Y)(y_k) d\mu_Y^*(y_1,y_2,y_3,y_4) \\ &= \int\prod_{k=1}^{4} f_k T^{km} f_k(x_k) d\mu^*(x_1,x_2,x_3,x_4).\end{aligned}$$

By the ergodic theorem, there exists a $(L^2(\mu^*), T\times\ldots\times T^4)$ invariant function D such that

$$\lim \frac{1}{M}\sum_{m=1}^{M}\gamma_m = \int\prod_{k=1}^{4} f_k(x_k) D(x_1,\ldots,x_4) d\mu^*(x_1,x_2,x_3,x_4). \tag{21}$$

By 4.3 this function is measurable w.r.t $\hat{Y}^4$. If f_k is orthogonal to $\hat{Y}$ then the average (21) is zero, and by VDC so is the original average. □

10.7. **Remark.** If Y_k is a characteristic factor for the scheme $(n, 2n, \ldots, kn)$, then by the same proof, $\hat{Y}_k$ is a characteristic factor for $(n, 2n, \ldots, (k+1)n)$.

11. THE ERGODIC COMPONENTS OF μ_Y^*

11.1. **Notation.**

1. For $y \in Y$, $\bar{y} = (y,y,y,y)$, $\tilde{y} = (y_1,y_2,y_3,y_4)$. For $\theta \in Z$, $\varphi \in H$, $\bar{\theta}, \bar{\varphi}, \tilde{\theta}, \tilde{\varphi}$ defined similarly.
2. T stands for $T \times T \ldots \times T$ with length depending on context.
3. $\tau = T \times T^2 \times T^3 \times T^4$.
4. $\tau_u = T_u \times T_u^2 \times T_u^3 \times T_u^4$ for $u \in \mathcal{U}$ ($\mathcal{U}$ is defined in 9.9).
5. For $\tilde{h} \in H^4$, $V_{\tilde{h}}(\tilde{\theta}, \tilde{\varphi}) = (\tilde{\theta}, \tilde{\varphi} + \tilde{h})$.

11.2. **Lemma.** *Let*

$$M := \{(h_1,h_2,h_3,h_4) : 3h_1 - 3h_2 + h_3 = 0;\ 4h_1 - 6h_2 + 4h_3 - h_4 = 0\}.$$

Then M is the Mackey group characterizing the ergodic components of the Abelian extension $\{\bar{z} + \tilde{\beta}\}_{\beta\in Z} \times_{\tilde{\rho}} H^4$ with respect to τ.

Proof. See [12]. We give another way to compute the Mackey group implemented for $k = 4$ in corollary 13.4. ☐

11.3. Notation. For $\tilde{m} \in M$, $V_{\tilde{m}}(\tilde{\theta}, \tilde{\varphi}) = (\tilde{\theta}, \tilde{\varphi} + \tilde{m})$.

11.4. Lemma. *Let $\{\rho_\beta\}_{\beta \in Z}$ be a measurable family of of functions satisfying the CL equation with ρ, then the ergodic components of μ_Y^* are supported on*

$$L_y := \{\tau_{(\beta,\rho_\beta)} V_{\tilde{m}}(\bar{y})\}_{\beta \in Z,\ \tilde{m} \in M}$$

where y ranges over Y.

Proof. We need to show invariance under τ (ergodicity follows from the properties of the Mackey group). The functions ρ_β are defined up to an eigenfunction (taking values in H), thus there exists an eigenfunction ψ such that

$$(\alpha, \rho)(\beta, \rho_\beta) = (\alpha + \beta, \rho^\beta + \rho_\beta) = (\alpha + \beta, \psi + \rho_{(\alpha+\beta)})$$

Thus for $m \leq 4$:

$$\begin{aligned}
&(\alpha, \rho)^m (\beta, \rho_\beta)^m \\
&\quad = (\alpha + \beta, \psi + \rho_{(\alpha+\beta)})^m [(\alpha, \rho), (\beta, \rho_\beta)]^{\binom{m}{2}} \\
&\quad = \left(m(\alpha + \beta), \binom{m}{2} \psi(\alpha + \beta) + m\psi + \rho_{m(\alpha+\beta)} \right) [(\alpha, \rho), (\beta, \rho_\beta)]^{\binom{m}{2}}.
\end{aligned}$$

But

$$(\psi(\theta), 2\psi(\theta), 3\psi(\theta), 4\psi(\theta))\,;\ (0, h(\alpha,\beta), 3h(\alpha,\beta), 6h(\alpha,\beta)) \in M.$$

☐

11.5. Corollary. *The ergodic components of τ are invariant under $\tau_u, V_{\tilde{m}}$, for $u \in \mathcal{U}$, $\tilde{m} \in M$.*

Proof. Invariance under $\tau_{\tilde{m}}$ is obvious. Invariance under τ_u follows as by the same computation in lemma 11.4, there exists $\tilde{m} \in M$ such that $\tau_u \tau_v = \tau_{uv} V_{\tilde{m}}$. ☐

11.6. Remark. One can see directly that different y lead to disjoint sets; thus the ergodic of μ_Y^* components are parameterized by $y \in Y$.

11.7. *Reduction to a group extension of the CL factor* By Theorem 10.6 we may assume that X is an isometric extension of the CL factor Y. Thus by 4.1 it is of the form $Y \times_\sigma G/H$, where G/H is a homogeneous space of a compact metric group G. By 4.6 we may assume that $Y \times_\sigma G$ is an ergodic group extension. We will prove the following theorem:

11.8. Theorem. *If X is an ergodic group extension of its CL factor, then the minimal characteristic factor of X for schemes of length 4 is a 3-step pro-nilsystem.*

11.9. **Corollary.** *If X is an isometric extension of its CL factor $X = Y \times_\sigma G/H$, then the minimal characteristic factor for schemes of length* 4 *is a* 3*-step pro-nilsystem.*

Proof. Choose G,H such that $X' = Y \times_\sigma G$ is an ergodic group extension of Y. Let Y' be the m.c.f. of X' for schemes of length 3. By corollary 8.5, the map $X' \to X$ induces a map from Y' onto Y. By lemma 4.7, X' is a group extension of Y'. By theorem 11.8 its minimal characteristic factor for schemes of length 4 is a 3 step nilsystem. By proposition 13.47 (to be proved later) a factor of a 3 step pro-nilsystem is also a 3 step pro-nilsystem. ☐

,

11.10. *Reduction to an Abelian extension of the CL factor* By corollary 11.9 it is enough to analyze the case where the system (X, T) is a group extension of the CL factor Y, i.e. $X = Y \times_\sigma G$, where G is a compact metrizable group. Denote $G' = [G, G]$. We show that $Y \times_\sigma G/G'$ is a characteristic factor of X. The proofs in this section carry over verbatim replacing 4 with any k, and Y with the characteristic factor for the $k - 1$ scheme.

11.11. **Lemma.** *Let $G_1, \dots, G_4$ be* 4 *groups, $\tilde{G} = \prod_{j=1}^4 G_j$. Denote by π_i the natural projection:*

$$\pi_i : \tilde{G} \to \prod_{j=1, j \neq i}^{4} G_j$$

Let N be a subgroup of $\tilde{G}$ satisfying $\pi_i(N) = \pi_i(\tilde{G})$. Then

- *There exists an Abelian group K and homomorphisms $\psi_i : G_i \to K$ so that*
$$N = \{(g_1, \dots, g_4) | \psi_1(g_1) \dots \psi_4(g_4) = 1\}$$
- *If G_i' denotes the commutator subgroup of G_i then $G_1' \times \dots \times G_4' \subset N$.*

Proof. We first show that N is a normal subgroup. Let

$$N_1 = 1 \times \prod_{j=2}^{4} G_j \cap N,$$

and define N_i similarly. We claim that N_i is a normal subgroup of $\tilde{G}$. Let $\tilde{g} = (1, g_2, g_3, g_4) \in N_1$, and take $\tilde{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4) \in \tilde{G}$. There exists $\delta \in G_1$, such that $\alpha = (\delta, g_2, g_3, g_4) \in N$, thus

$$\tilde{\beta}\tilde{g}\tilde{\beta}^{-1} = (1, \beta_2 g_2 \beta_2^{-1}, \beta_3 g_3 \beta_3^{-1}, \beta_4 g_4 \beta_4^{-1}) = \alpha \tilde{g} \alpha^{-1} \in N_1.$$

Next we claim $N = N_2 N_3$. Let $(g_1, \dots, g_4) \in M$. we can find $h \in G_1$ s.t. $(h, 1, g_3, g_4) \in N_2$. The element

$$(h^{-1} g_1, g_2, 1, 1) = (h, 1, g_3, g_4)^{-1} (g_1, \dots, g_4) \in N_3,$$

we can therefore express

$$(g_1, \dots, g_4) = (h, 1, g_3, g_4)(h^{-1} g_1, g_2, 1, 1) \in N_2 N_3.$$

Thus M is normal in $\tilde{G}$. Let $K = \tilde{G}/N$. Define

$$\psi_1(g_1) = (g_1, 1, 1, 1)N \in K,$$

and ψ_i similarly. Clearly

$$(g_1, \dots, g_4) \in N \iff \psi_1(g_1) \dots \psi_4(g_4) = 1.$$

Each map $\psi_i : G_i \to K$ is onto and for $i \neq j$, $\psi_i(g_i)\psi_j(g_j) = \psi_j(g_j)\psi_i(g_i)$. It follows that K is Abelian. □

11.12. **Lemma.** *Let $F(y_1, g_1, \dots, y_4, g_4)$ be a function in $L^2(\mu^*)$ invariant under τ. Let G' be the commutator of G and let $\varphi : (Y \times G)^4 \to (Y \times G/G')^4$ be the natural map. Then there exists a function $\tilde{F}$ on $(Y \times G/G')^4$ so that $F = \tilde{F} \circ \varphi$*

Proof. The statement of the lemma is equivalent to the assertion that almost all ergodic components of μ^* with respect to the transformation τ are invariant under the action of $(G')^4$. For $y \in Y$ let L_y be an ergodic component of τ on Y^4 (see lemma 11.4) . The ergodic components of τ on $(Y \times G)^4$ are the ergodic components of the G^4 extension of the L_y. For each such y we obtain a Mackey group $N_y \subset G^4$ defined up to conjugacy. Denote by $[N_y]$ the conjugacy class of the subgroup N_y. The set of conjugacy classes is a compact metric space, and the map $y \to [N_y]$ is measurable. Since $T \times \ldots \times T$ commutes with τ it follows that $[N_y] = [N_{Ty}]$. By ergodicity $[N_y] = [N]$ for almost all y. We now claim that $\pi_i(N) = G^3$. For $i = 4$ the statement is equivalent to showing that the Mackey group of $T \times T^2 \times T^3$ on $\pi_4(L_y) \times G^3$ relative to $\pi_4(L_y)$ is G^3. This is the same as saying that the $T \times T^2 \times T^3$ invariant functions with respect to $\pi_4(\mu^*)$ are lifted from invariant functions on the base space $\pi_i(L_y)$ but this is satisfied by the CL -factor (this is equivalent to saying the the CL -factor is a characteristic factor for the scheme $(1, 2, 3)$). The argument for $i < 4$ is the same. □

11.13. We return to the average in 10.6. If f_1 is orthogonal to $Y \times G/G'$ then the integral in equation (21) will be zero.

11.14. **Lemma.** *Let $J_0 = \cap_{k=1}^4$ ker ψ_k, Then we may replace $J = G/G'$ with J/J_0.*

Proof. $J_0 \times J_0 \times J_0 \times J_0 \subset N$, and μ^* invariant functions are invariant under rotation by elements of N. □

11.15. **Corollary.** *We may assume that $J_0 = \{1\}$.*

11.16. By Pontryagin duality $J \hookrightarrow (S^1)^{\hat{J}}$ so we replace $Y \times J$ with a join of systems of the form $Y \times_{\omega \circ \sigma} S^1$ where $\omega \in \hat{J}$. We now wish to study S^1 extensions of the CL factor Y. To make the notation more simple we replace $\omega \circ \sigma$ with σ.

12. THE MACKEY GROUP FOR μ^*

12.1. The Mackey Group For μ^*. In accordance with 11.16 we can reduce the study of the characteristic factor of a general system for schemes of length 4 to that of $X = Y \times_\sigma S^1$ where Y is a CL system. We decompose the system

$$X^4 = (Y \times S^1)^4, \quad \mu^* = \mu_Y^* \times Haar^4$$

into its ergodic components. For $y \in Y$, let L_y be an ergodic component of τ (see lemma 11.4). Consider the system

$$(L_y \times (S^1)^4, \tau).$$

This is a group extension of an ergodic system. Let N_y be the associated Mackey group in $(S^1)^4$ (5.1). As $T \times T \times T \times T$ commutes with τ, and the action of the group generated by $\tau, T \times T \times T \times T$ on (X^4, μ^*) is ergodic, it follows that N_y is a.e. a constant function of y , thus $N = N_y$. If $N = (S^1)^4$ then then μ^* invariant functions are already measurable with respect to Y^4 (the characteristiv factoor in this case would be Y). Otherwise as the projection of N on any 3 coordinates is full (see lemma 11.12) the group $N^\perp$ is generated by one element (m_1, m_2, m_3, m_4) , and by lemma 11.14, $gcd\{m_k\} = 1$. By theorem 5.3

$$\prod_{k=1}^{4} \sigma_k^{m_k}(y_k) = \frac{\tau F_y(\tilde{y})}{F_y(\tilde{y})} \tag{22}$$

for a.e. $y \in Y$ $\tilde{y} \in L_y$. Define

$$L := \bigcup_y L_y \tag{23}$$

The set L is the support of the measure μ_Y^*. By proposition 6.17 there is a measurable choice of F_y, thus we can write

$$\prod_{k=1}^{4} \sigma_k^{m_k}(y_k) = \frac{\tau F(\tilde{y})}{F(\tilde{y})} \tag{24}$$

a.e. $\tilde{y} \in L$.

12.2. Remark. Let $y = (\theta, \varphi)$ then $V_{(h,h,h,h)} L_(\theta, \varphi) = L_{\theta, \varphi + h}$, Thus L is invariant under $V_{(h,h,h,h)}$ for all $h \in H$.

13. SOLUTION OF THE FUNCTIONAL EQUATION

13.1. We are trying to extract information on σ from the functional equation (24) given by the Mackey group. We shall arrive at the following CL type equation:

$$\frac{\sigma(\theta, \varphi + h)}{\sigma(\theta, \varphi)} = \lambda_h \frac{f_h(Ty)}{f_h(y)}. \tag{25}$$

We start by showing that the homogeneous equation (when $\lambda_h \equiv 1$) implies that $\sigma(\theta,\varphi) = g(\theta)\frac{F(Ty)}{F(y)}$, and by reusing the original equation (24), we deduce that g must satisfy the CL equation. This means that if σ and σ' are solutions to equation (24) then their quotient satisfies the CL equation. Given the set $\{\lambda_h\}$ arising from equation (25), we construct a 3-step pro-nilsystem with a cocycle satisfying equations (24), and (25). We do that by first reducing the problem to the case where the CL factor Y is a torus $\mathbb{T}^m \times_\rho \mathbb{T}^n$. Then the group $\mathcal{U}$ defined in 9.9 is a locally connected 2 step nilpotent Lie group. Next we construct a collection of functions $\{F_u\}_{u\in\mathcal{U}}$ (that depend measurably on u) that satisfy 'nice' commutation relations with σ. This family will produce a set of constants $\lambda(u,h)$, that satisfy a 'Jacobi' condition. We use these constants to construct a 3-step nilsystem. We have thus constructed a 3-step nilsystem with the same λ_h as the original one .Using the information gathered regarding the homogeneous equation, we find that the original cocycle differs from the one constructed by a CL cocycle factor. Combining the solutions to the CL case with what we have obtained then produces a system isomorphic to the original characteristic system, and shows that it is a (pro)nilsystem. We remark that most of the following analysis applies for general k (using induction).

13.2. **Construction of f_h.** Let (X,T) be an ergodic m.p.s. which is a circle extension of its CL factor; i.e $X = Y \times_\sigma S^1$, where Y is the CL factor of the system. Then $Y = (Z \times_\rho H, T)$, where (Z,α) is the Kronecker factor of (X,T), H a connected compact Abelian group, and ρ satisfies the CL equation (9.1). Suppose the extension (multiplicative) cocycle $\sigma(\theta,\varphi)$: $Z \times H \to S^1$ satisfies the functional equation (24).

13.3. **Proposition.** *There exists a measurable family of functions $\{f_h\}_{h\in H}$, such that*

$$\frac{\sigma_1(\theta,\varphi+h)}{\sigma_1(\theta,\varphi)} = \lambda_h \frac{Tf_h(\theta,\varphi)}{f_h(\theta,\varphi)}, \tag{26}$$

for all $h \in H$, a.e. θ,φ.

Proof. Recall the definitions of $L \subset Y^4$, $M \subset H^4$ from lemma 11.2, and equation (23). By lemma 11.4 the set L is invariant under V_m for $m \in M$. By remark 12.2, L is also invariant under $V_{(h,h,h,h)}$ for any $h \in H$. Thus it is invariant under

$$V_{(6h,6h,6h,6h)-(3h,5h,6h,6h)} = V_{(3h,h,0,0)}.$$

Substituting $(\tilde\theta, \tilde\varphi + (3h,h,0,0))$ in equation (24),

$$(\bigotimes_{k=1}^{4} \sigma_k^{m_k})(\tilde\theta,\tilde\varphi+(3h,h,0,0)) = \frac{\tau F(\tilde\theta,\tilde\varphi+(3h,h,0,0))}{F(\tilde\theta,\tilde\varphi+(3h,h,0,0))}. \tag{27}$$

Denote

$$F_h(\tilde\theta,\tilde\varphi) = \frac{F(\tilde\theta,\tilde\varphi+(3h,h,0,0))}{F(\tilde\theta,\tilde\varphi)}.$$

Dividing equations (27),(24) , the functions

$$\sigma_3^{m_3}(\theta_3,\varphi_3),\sigma_4^{m_4}(\theta_4,\varphi_4)$$

appear in both equations, and thus disappear in the quotient:

$$\frac{\sigma_1^{m_1}(\theta_1,\varphi_1+3h)}{\sigma_1^{m_1}(\theta_1,\varphi_1)}\frac{\sigma_2^{m_2}(\theta_2,\varphi_2+h)}{\sigma_2^{m_2}(\theta_2,\varphi_2)}=\frac{\tau F_h(\tilde{\theta},\tilde{\varphi})}{F_h(\tilde{\theta},\tilde{\varphi})} \tag{28}$$

We fix a measurable family of functions $\{\rho_\beta\}_{\beta\in Z}$ satisfying the CL equation (9.1) with ρ. Recall that

$$(\theta_1,\theta_2,\theta_3,\theta_4)=(\theta+\beta,\theta+2\beta,\theta+3\beta,\theta+4\beta).$$

This implies that θ_1,θ_2 determine β and θ ($\theta_2-\theta_1=\beta$, $2\theta_1-\theta_2=\theta$), and thus also determine θ_3,θ_4 and the value $\rho_\beta(\theta)$. Using lemma 11.4, the elements in L satisfy:

$$\begin{aligned}&\theta_4=\theta_1-3\theta_2+3\theta_3\\&\varphi_4=\varphi_1-3\varphi_2+3\varphi_3+\rho_\beta(\theta)-3\rho_{2\beta}(\theta)+3\rho_{3\beta}(\theta)-\rho_{4\beta}(\theta)\end{aligned} \tag{29}$$

In other words there is some function $G:(Z\times H)^3\to Z\times H$ such that for $(\tilde{\theta},\tilde{\varphi})\in L$:

$$(\theta_4,\varphi_4)=G(\theta_1,\theta_2,\theta_3,\varphi_1,\varphi_2,\varphi_3).$$

As L is τ invariant

$$T^4(\theta_4,\varphi_4)=T\times T^2\times T^3G(\theta_1,\theta_2,\theta_3,\varphi_1,\varphi_2,\varphi_3)$$

Furthermore, if we take $y=(\theta,\varphi_0)$ and set

$$L_{\varphi_0}=\bigcup_\theta L_{\theta,\varphi_0},$$

then by lemma 11.4, in L_{φ_0}

$$(\theta_3,\varphi_3)=H(\theta_1,\varphi_1,\theta_2,\varphi_2),$$

and as L_{φ_0} is invariant under τ,

$$T^3(\theta_3,\varphi_3)=T\times T^2H(\theta_1,\varphi_1,\theta_2,\varphi_2).$$

Thus restricting to L_{φ_0}, equation (28) can be expressed as

$$\frac{\sigma_1^{m_1}(\theta_1,\varphi_1+3h)}{\sigma_1^{m_1}(\theta_1,\varphi_1)}\frac{\sigma_2^{m_2}(\theta_2,\varphi_2+h)}{\sigma_2^{m_2}(\theta_2,\varphi_2)}=\frac{T\times T^2K_{\varphi_0,h}(\theta_1,\varphi_1,\theta_2,\varphi_2)}{K_{\varphi_0,h}(\theta_1,\varphi_1,\theta_2,\varphi_2)}. \tag{30}$$

In L_{φ_0} there are no restrictions on φ_1,φ_2, thus the above equation is true for a.e. $\theta_1,\theta_2\in Z$, $\varphi_1,\varphi_2\in H$ for fixed $h\in H$. the solution to equation (28) is (see [21])

$$\frac{\sigma_1^{m_1}(\theta_1,\varphi_1+3h)}{\sigma_1^{m_1}(\theta_1,\varphi_1)}=\lambda_{3h}\frac{Tf_{3h}(\theta_1,\varphi_1)}{f_{3h}(\theta_1,\varphi_1)} \tag{31}$$

$$\frac{\sigma_2^{m_2}(\theta_2,\varphi_2+h)}{\sigma_2^{m_2}(\theta_2,\varphi_2)}=\lambda_{3h}^{-1}\frac{T^2g_h(\theta_2,\varphi_2)}{g_h(\theta_2,\varphi_2)}. \tag{32}$$

As $h \to 3h$ is onto (H is connected),

$$\frac{\sigma_1^{m_1}(\theta, \varphi + h)}{\sigma_1^{m_1}(\theta, \varphi)} = \lambda_h \frac{T f_h(\theta, \varphi)}{f_h(\theta, \varphi)} \tag{33}$$

By iteration it follows that for any l:

$$\frac{\sigma_l^{m_1}(\theta, \varphi + h)}{\sigma_l^{m_1}(\theta, \varphi)} = \lambda_h^l \frac{T^l f_h(\theta, \varphi)}{f_h(\theta, \varphi)}. \tag{34}$$

Specifically for $l = 2$ we have

$$\frac{\sigma_2^{m_1}(\theta, \varphi + h)}{\sigma_2^{m_1}(\theta, \varphi)} = \lambda_h^2 \frac{T^2 f_h(\theta, \varphi)}{f_h(\theta, \varphi)}. \tag{35}$$

combine equations (32) (35) with $m_1 l_1 = m_2 l_2$:

$$\lambda_{3h}^{2l_2} \lambda_h^{4l_1} = \text{ eigenvalue of } T^2 \text{ for all } h.$$

The constants λ_h, f_h can be chosen such that λ_h is multiplicative in a zero neighborhood (lemma 6.10). As H is connected and T^2 has only a countable number of eigenvalues, we get $2l_1 + 3l_2 = 0$ and thus:

$$(m_1 : m_2) = k(-2, 3).$$

The $\tilde{\varphi}$ coordinate is invariant under addition of $(h, 2h, 3h, 4h)$, $(h, 4h, 9h, 16h)$ thus substituting equation (34) in equation (24)

$$\lambda_h^{m_1 + 4m_2 + 9m_3 + 16m_4}, \ \lambda_h^{m_1 + 8m_2 + 27m_3 + 64m_4}$$

are eigenvalues of τ for all h. As λ_h, f_h can be chosen such that λ_h is multiplicative in a zero neighborhood (lemma 6.10)

$$m_1 + 4m_2 + 9m_3 + 16m_4 = 0; \ m_1 + 8m_2 + 27m_3 + 64m_4 = 0.$$

Altogether we get:

$$(m_1, m_2, m_3, m_4) = k(-4, 6, -4, 1).$$

As $\gcd(m_k) = 1$, $k = 1$. By applying the foregoing computation for the pair σ_1, σ_4 (the element $(-2h, -3h, -3h, -2h) \in M$ thus L is invariant under

$$V_{(h,0,0,h)} = V_{(-2h,-3h,-3h,-2h)+(3h,3h,3h,3h)}$$

which enables us to eliminate σ_2 and σ_3), we get (as $m_4 = 1$):

$$\frac{\sigma_4(\theta, \varphi + h)}{\sigma_4(\theta, \varphi)} = \delta_h \frac{T^4 g_h(\theta, \varphi)}{g_h(\theta, \varphi)}. \tag{36}$$

As T, T^4 commute we have:

$$\sigma_1(T^4(\theta, \varphi))\sigma_4(\theta, \varphi) = \sigma_4(T(\theta, \varphi)\sigma_1(\theta, \varphi).$$

By equation (36):

$$\frac{T^4 \frac{\sigma_1(\theta,\varphi+h)}{\sigma_1(\theta,\varphi)}}{\frac{\sigma_1(\theta,\varphi+h)}{\sigma_1(\theta,\varphi)}} = \frac{T \frac{\sigma_4(\theta,\varphi+h)}{\sigma_4(\theta,\varphi)}}{\frac{\sigma_4(\theta,\varphi+h)}{\sigma_4(\theta,\varphi)}} = \frac{T \delta_h \frac{T^4 g_h(\theta,\varphi)}{g_h(\theta,\varphi)}}{\delta_h \frac{T^4 g_h(\theta,\varphi)}{g_h(\theta,\varphi)}} = \frac{T \frac{T^4 g_h(\theta,\varphi)}{g_h(\theta,\varphi)}}{\frac{T^4 g_h(\theta,\varphi)}{g_h(\theta,\varphi)}} = \frac{T^4 \frac{T g_h(\theta,\varphi)}{g_h(\theta,\varphi)}}{\frac{T g_h(\theta,\varphi)}{g_h(\theta,\varphi)}}.$$

Thus the function

$$\left(\frac{\sigma_1(\theta,\varphi+h)}{\sigma_1(\theta,\varphi)}\right)^{-1}\frac{Tg_h(\theta,\varphi)}{g_h(\theta,\varphi)}$$

is invariant under T^4.If T^4 is ergodic we're done; otherwise

$$\frac{\sigma_1(\theta,\varphi+h)}{\sigma_1(\theta,\varphi)}=\Lambda_h(\theta+4Z)\frac{Tg_h(\theta,\varphi)}{g_h(\theta,\varphi)}.$$

By proposition 13.10 (to be proved later), this would imply that $\Lambda_h(\theta+4Z)$ satisfies the CL equation. By Rudolph ([27] corollary 5.5) it depends only on the connected part of Z - and is thus constant. Measurability follows from remark 6.17. □

13.4. **Remark.** The above computation gives a description of the Mackey group:

$$\{(\zeta_1,\zeta_2,\zeta_3,\zeta_4):\zeta_1^{-4}\zeta_2^{6}\zeta_3^{-4}\zeta_4=1\}.$$

13.5. **Corollary.** *By lemma 6.10 there exists a measurable family of functions* $\{f_h\}_{h\in H}$ *satisfying equation (26), such that*

$$f_{h_1+h_2}(\theta,\varphi)\sim f_{h_1}(\theta,\varphi+h_2)f_{h_2}(\theta,\varphi)$$
$$\lambda_{h_1+h_2}=\lambda_{h_1}\lambda_{h_1}$$

for h_1,h_2 *in* U *- a neighborhood of zero in* H*.*

13.6. **Definition.** Let $B(Y,S^1)$ be the set of measurable functions from $Y\to S^1$ with the L^2 topology. Let

$$\mathcal{H}=\{(u,f):u\in\mathcal{U};\ f\in B(Y,S^1)\}.$$

$\mathcal{H}$ is a group under the multiplication

$$(u,f)(v,g)=(uv,f^vg),\qquad (f^vg)(y)=f(T_vy)g(y).$$

$\mathcal{H}$ is a polish group which acts on $X=Y\times S^1$ by:

$$(u,f)(y,\zeta)=(T_uy,f(y)\zeta).$$

We will also denote the elements of $\mathcal{H}$ as pairs (T_u,f) or (V_h,f) where we are identifying $\mathcal{U}$ with its action on Y. For (T_u,f), (T_v,g) in $\mathcal{H}$ we have

$$[(T_u,f),(T_v,g)]=(T_{[u,v]},k),\tag{37}$$

where $k(y)$ satisfies

$$\frac{T_vf(y)}{f(y)}=k(T_{vu}y)\frac{T_ug(y)}{g(y)}.$$

For $u=(0,h)$ equation (26) implies

$$[(V_h,f_h),(T,\sigma)]=(0,\lambda_h).$$

Denote $C(\mathcal{H})$ the center of $\mathcal{H}$.

13.7. **Proposition.** *If* $f_h(\theta,\varphi)$ *satisfies equation (26) for all* h*, then*

$$[(V_h,f_h),(V_g,f_g)]=0$$

Proof. The function element in $[(V_h, f_h), (V_g, f_g)]$ is invariant under T, therefore constant,

$$[(V_h, f_h), (V_g, f_g)] = (0, \gamma(g, h))$$

As $f_{g_1}(\theta, \varphi + g_2) f_{g_2}(\theta, \varphi)$ also satisfies equation (26) for $h = g_1 + g_2$,

$$f_{g_1+g_2}(\theta, \varphi)\chi_{g_1,g_2}(\theta) = f_{g_1}(\theta, \varphi + g_2) f_{g_2}(\theta, \varphi).$$

where $\chi_{g_1,g_2}(\theta)$ is an eigenfunction of T (and thus invariant under V_h).

$$\begin{aligned}[(V_h, f_h), (V_{g_1}, f_{g_1})][(V_h, f_h), (V_{g_2}, f_{g_2})]&\\ =&[(V_h, f_h), (V_{g_1}, f_{g_1})(V_{g_2}, f_{g_2})]\\ =&[(V_h, f_h), (V_{g_1+g_2}, \chi_{g_1,g_2} f_{g_1+g_2})]\\ =&[(V_h, f_h), (V_{g_1+g_2}, f_{g_1+g_2})].\end{aligned}$$

Therefore

$$\gamma(g_1 + g_2, h) = \gamma(g_1, h)\gamma(g_2, h)$$

for all g_1, g_2, h, and from symmetry it is also true for the second coordinate. This implies that $\gamma_h(g) = \gamma(h, g)$ is a character of H (we assume that $f_0 \equiv 1$,thus $\gamma(g_1, h) \equiv 1$). But any homomorphism from H (which is connected) to $\hat{H}$ is constant. □

13.8. **Lemma.** *If the cocycle σ satisfies for all $h \in H$:*

$$\frac{V_h\sigma(\theta, \varphi)}{\sigma(\theta, \varphi)} = \frac{T f_h(\theta, \varphi)}{f_h(\theta, \varphi)}, \tag{38}$$

Then

$$\sigma(\theta, \varphi) = h(\theta)\frac{TH(\theta, \varphi)}{H(\theta, \varphi)}$$

Proof. By corollary 6.8 for the case where H is connected. □

13.9. We need to get more restrictions on $g(\theta)$, so we return to the original functional equation (24). Iterating we get

$$\sigma_k(\theta, \varphi) = g_k(\theta)\frac{T^k F(\theta, \varphi)}{F(\theta, \varphi)}$$

where $g_k(\theta)$ is a $\mathbb{Z}$ cocycle for the action of T. By equation (24):

$$\prod_{k=1}^{4} g_k^{m_k}(\theta_k) = \frac{\tau F(\tilde{y})}{F(\tilde{y})}. \tag{39}$$

13.10. **Proposition.** *If $g(\theta)$ satisfies the functional equation (39) then $g(\theta)$ satisfies the CL equation.*

Proof. We already showed in the proof of proposition 13.3 that this equation reduces to

$$\prod_{k=1}^{4} g_k^{m_k}(\theta_k) = h(\theta_1, \theta_2) = \frac{T \times T^2 F(y_1, y_2)}{F(y_1, y_2)}. \tag{40}$$

As (y_1, y_2) are free coordinates, F can be expanded in the following form:

$$F(y_1, y_2) = \sum_{n_1, n_2} F_{n_1,n_2}(\theta_1, \theta_2)\psi_{n_1}(\varphi_1)\psi_{n_2}(\varphi_2) \tag{41}$$

Where ψ_{n_i} are characters of H. combining equations (40), (41) we get: for all n_1, n_2

$$T \times T^2 F_{n_1,n_2}(\theta_1, \theta_2)\psi_{n_1}(\rho_1(\theta_1))\psi_{n_2}(\rho_2(\theta_2)) = h(\theta_1, \theta_2)F_{n_1,n_2}(\theta_1, \theta_2) \tag{42}$$

The function $|F_{n_1,n_2}(\theta_1, \theta_2)|$ is constant on ergodic components of $T \times T^2$. As

$$|F(y_1, y_2)| = 1 \qquad \text{a.e} \quad (y_1, y_2),$$

almost all ergodic components have (n_1, n_2) s.t.

$$|F_{n_1,n_2}(\theta_1, \theta_2)| \neq 0.$$

The ergodic components are invariant under $T_u \times T_u^2$ for all $u = (\beta, f) \in \mathcal{U}$, by corollary 11.5, so applying $T_u \times T_u^2$ to equation (42) and dividing the two equations:

$$\frac{\frac{(T_u \times T_u^2)(T \times T^2)F_{n_1,n_2}(\theta_1,\theta_2)}{(T_u \times T_u^2)F_{n_1,n_2}(\theta_1,\theta_2)}}{\frac{(T \times T^2)F_{n_1,n_2}(\theta_1,\theta_2)}{F_{n_1,n_2}(\theta_1,\theta_2)}} = \frac{(T_u \times T_u^2)h(\theta_1, \theta_2)\psi_{n_1}^{-1}(\rho_1(\theta_1))\psi_{n_2}^{-1}(\rho_2(\theta_2))}{h(\theta_1, \theta_2)\psi_{n_1}^{-1}(\rho_1(\theta_1))\psi_{n_2}^{-1}(\rho_2(\theta_2))}$$

We now use the fact that the transformations $T \times T^2$, $T_u \times T_u^2$ commute on the Kronecker factor, and the fact that $\psi_{n^i}^{-1}(\rho_i(\theta_i))$ satisfies CL to get:

$$\frac{(T \times T^2)F_{u,n_1,n_2}(\theta_1, \theta_2)}{F_{u,n_1,n_2}(\theta_1, \theta_2)} = \delta(n_1, n_2)\frac{(T_u \times T_u^2)h(\theta_1, \theta_2)}{h(\theta_1, \theta_2)}$$

We repeat the procedure with $T_v \times T_v^2$ for $v = (\gamma, g)$ and get

$$\frac{(T \times T^2)F_{u,v,n_1,n_2}(\theta_1, \theta_2)}{F_{u,v,n_1,n_2}(\theta_1, \theta_2)} = \frac{\frac{(T_u \times T_u^2)(T_v \times T_v^2)h(\theta_1,\theta_2)}{(T_v \times T_v^2)h(\theta_1,\theta_2)}}{\frac{(T_u \times T_u^2)h(\theta_1,\theta_2)}{h(\theta_1,\theta_2)}}.$$

Now substitute

$$h(\theta_1, \theta_2) = \prod_{k=1}^{4} g_k^{m_k}(\theta_k)$$

in the foregoing equation to get that

$$\frac{(T \times T^2)F_{u,v,n_1,n_2}(\theta_1, \theta_2)}{F_{u,v,n_1,n_2}(\theta_1, \theta_2)} = \prod_{k=1}^{4} \frac{\frac{g_k^{m_k}(\theta_k + k\beta + k\gamma)}{g_k^{m_k}(\theta_k + k\gamma)}}{\frac{g_k^{m_k}(\theta_k + k\beta)}{g_k^{m_k}(\theta_k)}}.$$

By proposition 9.8

$$G(\theta) = \frac{\frac{g(\theta+\beta+\gamma)}{g(\theta+\beta)}}{\frac{g(\theta+\gamma)}{g(\theta)}}$$

is a CL cocycle (take $\rho^{(k)} = G^{m_k}$ and recall that $m_4 = 1$), and by proposition 9.8, g itself is a CL cocycle. $\square$

13.11. **Corollary.** *If* $\sigma(\theta,\varphi)$ *and* $\tau(\theta,\varphi)$ *satisfy both the functional equation (24), and equation (26) with the same* λ_h *for* h *in a neighborhood of zero in* H*, then*

$$\frac{\sigma(\theta,\varphi)}{\tau(\theta,\varphi)} = g(\theta)\frac{TG(\theta,\varphi)}{G(\theta,\varphi)}$$

where $g(\theta)$ *satisfies the CL equation.*

Proof. By iteration, as H is connected,

$$\frac{\frac{\sigma(\theta,\varphi+h)}{\tau(\theta,\varphi+h)}}{\frac{\sigma(\theta,\varphi)}{\tau(\theta,\varphi)}} = \frac{TG_h(\theta,\varphi)}{G_h(\theta,\varphi)}$$

now apply lemma 13.8 and proposition 13.10. ☐

13.12. **Lemma.** (reduction to an extension of Kronecker by a finite torus) *There exists a subgroup* $H_2 < H$ *such that* $H/H_2 = \mathbb{T}^n$ *and if* $\pi : H \to H/H_2$ *is the natural projection, then there exists a function* $\tilde{\sigma} : Z \times H/H_2 \to S^1$ *such that*

$$\sigma(\theta,\varphi) = \tilde{\sigma}(\theta,\pi(\varphi)))\frac{TF(\theta,\varphi)}{F(\theta,\varphi)}.$$

Proof. By lemma 6.11, and remark 6.12 we can find H_2 so that $H/H_2 = \mathbb{T}^n \times H_0$ where H_0 is a compact totally disconnected Abelian group, but this would be a contradiction to Z being the CL factor of the system (see 9.4). ☐

13.13. Now $Z \times_\rho H$ is an extension of Kronecker by a finite torus $H = \mathbb{T}^n$. By lemma 9.6 there exists a subgroup $Z_1 < Z$, such that $Z/Z_1 = \mathbb{T}^m$, and ρ is lifted from a function on $\mathbb{T}^m$; i.e. if $\pi : Z \to Z/Z_1$ is the natural projection, then $\rho = \rho \circ \pi$. Let C be a measurable section $Z/Z_1 \to Z$: $C(z+Z_1)-z \in Z_1$. The system (Z,T) is isomorphic to the system $Z/Z_1 \times Z_1$ under the map

$$z \to (z+Z_1, C(z+Z_1)-z).$$

We compute the action of T:

$$\begin{aligned} z+\alpha &\to (z+\alpha+Z_1, C(z+\alpha+Z_1)-z-\alpha) \\ &= (z+\alpha+Z_1, C(z+Z_1)-z+(C(z+\alpha+Z_1)-C(z+Z_1)-\alpha)), \end{aligned}$$

thus

$$T(z+Z_1,z_1) = (z+\alpha+Z_1, z_1+(C(z+\alpha+Z_1)-C(z+Z_1)-\alpha)).$$

If $\beta \in Z_1$ then

$$z+\beta \to (z+\beta+Z_1, C(z+\beta+Z_1)-z-\beta) = (z+Z_1, C(z+Z_1)-z-\beta)$$

i.e translation in $-\beta$ in the second coordinate. Therefore the system (Y,T) is isomorphic to the system $(\mathbb{T}^m \times Z_1) \times \mathbb{T}^n$ with the action of T given by:

$$T(\theta_1,z_1,\varphi) = (\theta_1+\alpha_1, z_1+f(\theta_1), \varphi+\rho(\theta_1)).$$

where f is equivalent to a constant, and if $\gamma_1 \in Z_1$, the element $(0,\gamma_1,0)$ is in $\mathcal{U}$ and commutes with $u_\alpha = (\alpha_1, f, \rho)$, and $(0,0,h)$ for all $h \in H$.

13.14. Notation. When there is no confusion we denote $T_{(0,\beta_1,0)} = T_{\beta_1}$, and $\theta = (\theta_1, z_1)$

13.15. Lemma. *There exist a measurable set of functions* $\{g_{\beta_1}, F_{\beta_1}\}_{\{\beta_1 \in Z_1\}}$ *such that*

$$\frac{T_{\beta_1}\sigma(y)}{\sigma(y)} = g_{\beta_1}(\theta)\frac{TF_{\beta_1}(y)}{F_{\beta_1}(y)}.$$

Proof. For all $h \in \mathbb{T}^n$:

$$\frac{V_h \frac{T_{\beta_1}\sigma(y)}{\sigma(y)}}{\frac{T_{\beta_1}\sigma(y)}{\sigma(y)}} = \frac{T_{\beta_1}\frac{V_h\sigma(y)}{\sigma(y)}}{\frac{V_h\sigma(y)}{\sigma(y)}} = \frac{T_{\beta_1}\frac{Tf_h(y)}{f_h(y)}}{\frac{Tf_h(y)}{f_h(y)}} = \frac{T\frac{T_{\beta_1}f_h(y)}{f_h(y)}}{\frac{T_{\beta_1}f_h(y)}{f_h(y)}}$$

Now use corollary 6.8. Measurability follows from proposition 6.16. □

13.16. Lemma. *The functions* $g_{\beta_1}(\theta)$ *satisfy the CL equation.*

Proof. As the transformations T_{β_1}, T commute, and σ satisfies equation 24, the proof is the same as in proposition 13.10. □

13.17. Lemma. *Let* $\mathcal{V}$ *be a closed subgroup of* $\mathcal{U}$, $\{g_v\}_{v\in\mathcal{V}}$ *a measurable set of CL cocycles such that* $g_0 \equiv 1$ *and in* U *- a neighborhood of zero in* $\mathcal{V}$

$$\frac{g_u(T_v\theta)g_v(\theta)}{g_{uv}(\theta)} \sim \frac{TG_{u,v}(\theta)}{G_{u,v}(\theta)},$$

then there exists a neighborhood of zero U_0 *in* $\mathcal{V}$ *such that in* U_0

$$g_u(\theta) = C_u\frac{TK_u(\theta)}{K_u(\theta)}.$$

Proof. The cocycles $g_u(\theta)$ satisfy the CL equation, thus for each u there is an $m \in \mathbb{Z}$, $n(u) \in \mathbb{Z}^m$, and $M(u) \in M_m(\mathbb{Z})$, such that $g_u(\theta)$ is lifted from a function defined on $\mathbb{T}^m$ and for δ in a zero neighborhood of $\mathbb{T}^m$:

$$\frac{g_u(\theta+\delta)}{g_u(\theta)} = e^{2\pi i(M(u)\alpha\delta + n(u)\delta)}\frac{K_{u,\delta}(\theta+\alpha)}{K_{u,\delta}(\theta)},$$

for $u, v, uv \in U$

$$\begin{aligned}&(M(u)\alpha_m\delta + M(v)\alpha_m\delta - M(uv)\alpha_m\delta) + (n(u)+n(v)-n(uv))\,\delta\\ &= \text{eigenvalue of } T.\end{aligned}$$

Linearity in δ (in a neighborhood of zero) implies

$$n(u) + n(v) - n(uv) = 0$$
$$M(u) + M(v) - M(uv) = 0.$$

Let

$$U_{n,M} = \{u \in U : n(u) = n;\ M(u) = M\},$$

then for some n, M, $m(U_{n,M}) > 0$, thus $U_{n,M}U_{n,M}^{-1}$ contains a neighborhood of zero U_0. For $u \in U_0$

$$\frac{g_u(\theta+\delta)}{g_u(\theta)} = \frac{K_{u,\delta}(\theta+\alpha)}{K_{u,\delta}(\theta)},$$

By corollary 6.8

$$g_u(\theta) = C_u \frac{TK_u(\theta)}{K_u(\theta)}.$$

□

13.18. **Proposition.** *Let $\mathcal{V}$ be a subgroup of $\mathcal{U}$ acting continuously on Y. Let $\{h_u, g_u, F_u\}$ be a family of measurable function $Y \to S^1$ such that*

1. *g_u are CL characters.*
2. *$h_u(y) = g_u(\theta)\frac{TF_u(y)}{F_u(y)}$.*
3. *$h_{uv} \sim h_u^v h_v$.*

Then there exists a neighborhood of zero V_0 in $\mathcal{V}$, such that for $u \in V_0$, h_u is cohomologous to a constant, i.e.

$$h_u(y) = C_u \frac{TG_u(y)}{G_u(y)}. \tag{43}$$

Proof. The functions g_u, F_u satisfy

$$\frac{g_u^v g_v}{g_{uv}}(\theta) \sim \frac{\frac{F_{uv}}{F_u^v F_v}(Ty)}{\frac{F_{uv}}{F_u^v F_v}(y)}. \tag{44}$$

thus by lemma 6.14

$$\frac{F_{uv}}{F_u^v F_v}(y) = k_{u,v}(\theta)\chi_{u,v}(h) \tag{45}$$

where $\chi_{u,v}$ is a character of H,

$$\begin{aligned} F_{uvw}(y) =& k_{uv,w}(\theta)\chi_{uv,w}(h)F_{uv}(T_w y)F_w(y) \\ =& k_{u,vw}(\theta)\chi_{u,vw}(h)F_u(T_v T_w y)F_{vw}(y) \end{aligned} \tag{46}$$

Let $w = (\delta, \rho_\delta)$. Dividing the two equalities in equation (46) by

$$F_u(T_v T_w y)F_v(T_w y)F_w(y),$$

we get

$$k_{uv,w}(\theta)\chi_{uv,w}(h)k_{u,v}(\theta+\delta)\chi_{u,v}(h+\rho_\delta(\theta)) = k_{u,vw}(\theta)\chi_{u,vw}(h)k_{v,w}(\theta)\chi_{v,w}(h).$$

As

$$\chi_{u,v}(h + \rho_\delta(\theta)) = \chi_{u,v}(h)\chi_{u,v}(\rho_\delta(\theta)),$$

we get

$$\chi_{u,v}\chi_{uv,w} = \chi_{u,vw}\chi_{v,w}. \tag{47}$$

As $v \to F_v$ is a measurable function, there exists a neighborhood of zero $U \subset \mathcal{U}$ such that for $v \in U$,

$$m\{u : ||F_{uv} - F_u||_2 > \epsilon\} < \delta.$$

As the action of T_v is continuous, there exists a neighborhood of zero $U' \subset \mathcal{U}$ such that for $v \in U'$,

$$m\{u : ||F_u(T_v y) - F_u(y)||_2 > \epsilon\} < \delta.$$

Therefore, by equation (45), for $v \in U \cap U'$, $u \in A$ - a set of positive measure, the character $\chi_{u,v}$ is independent of u:

$$\chi_{u,v} = \tilde{\chi}_v \qquad \text{for } u \in A,\ v \in U \cap U'.$$

Let $v, w, vw \in U \cap U'$, and let $u \in A$ be such that $uv \in A$. Using equation (47):

$$\chi_{v,w} = \chi_{u,v}\chi_{uv,w}\chi_{u,vw}^{-1} = \tilde{\chi}_v\tilde{\chi}_w\tilde{\chi}_{vw}^{-1}.$$

Define

$$\tilde{F}_u(y) = F_u(y)\tilde{\chi}_u(h); \quad \tilde{g}_u(\theta) = g_u(\theta)\tilde{\chi}_u^{-1}(\rho(\theta)).$$

The functions $\tilde{F}_u(y), \tilde{g}_u(\theta)$ satisfy:

$$\tilde{g}_u(\theta)\frac{T\tilde{F}_u(y)}{\tilde{F}_u(y)} = g_u(\theta)\tilde{\chi}_u^{-1}(\rho(\theta))\frac{F_u(Ty)\tilde{\chi}_u(\rho(\theta))}{F_u(y)} = g_u(\theta)\frac{TF_u(y)}{F_u(y)}$$

and if $v = (\gamma, \rho_\gamma)$,

$$\begin{aligned}\tilde{F}_{uv}(y) =& F_{uv}(y)\tilde{\chi}_{uv}(h)\\ =& F_u(T_v y)F_v(y)k_{u,v}(\theta)\chi_{u,v}(h)\tilde{\chi}_{uv}(h)\\ =& F_u(T_v y)F_v(y)k_{u,v}(\theta)\tilde{\chi}_v(h)\tilde{\chi}_v(h)\\ =& \rho_\gamma(\theta)k_{u,v}(\theta)\tilde{F}_u(T_v y)\tilde{F}_v(y).\end{aligned}$$

Thus if $G_{u,v}(\theta) = \rho_\gamma(\theta)k_{u,v}(\theta)$ then by equation equation (44) (with $\tilde{F}_u, \tilde{g}_u$):

$$\frac{\tilde{g}_u(T_v\theta)\tilde{g}_v(\theta)}{\tilde{g}_{uv}(\theta)} \sim \frac{TG_{u,v}(\theta)}{G_{u,v}(\theta)}.$$

Now use lemma 13.17. □

13.19. Corollary. *There exists a neighborhood of zero W in Z_1 such that for $\beta_1 \in W$,*

$$\frac{T_{\beta_1}\sigma(y)}{\sigma(y)} = C_{\beta_1}\frac{TF_{\beta_1}(y)}{F_{\beta_1}(y)}.$$

Proof. The subgroup $\{((0, \beta_1), 0)\}$ of $\mathcal{U}$, and the function

$$h_{\beta_1}(y) = g_{\beta_1}(\theta)\frac{F_{\beta_1}(Ty)}{F_{\beta_1}(y)},$$

satisfy the conditions of proposition 13.18. □

13.20. Corollary. *There exists a subgroup Z_2 of Z_1, and a cyclic group C_k, such that $Z_1/Z_2 = \mathbb{T}^{m_1} \times C_k$, and if $\pi_1 : Z_1 \to Z_1/Z_2$ is the natural projection, then there exists a function $\tilde{\sigma} : (\mathbb{T}^m \times Z_1) \times \mathbb{T}^n \to S^1$ such that*

$$\sigma(\theta_1, z_1, \varphi) = \tilde{\sigma}(\theta_1, \pi_1(z_1), \varphi)\frac{TF(\theta_1, z_1, \varphi)}{F(\theta_1, z_1, \varphi)}.$$

Proof. Recall the neighborhood W from corollary 13.19. Let $W_1 \subset W$ be a subgroup such that $Z_1/W_1 = \mathbb{T}^{m_1} \times C_k$, and for $\beta_1 \in W_1$

$$\frac{T_{\beta_1}\sigma(y)}{\sigma(y)} = C_{\beta_1}\frac{TF_{\beta_1}(y)}{F_{\beta_1}(y)}.$$

Now apply lemma 6.13 with $Y = \mathbb{T}^{m+m_1} \times C_k \times \mathbb{T}^n$, $H = W_1$. □

13.21. We are left with analyzing the system $Y = (C_k \times \mathbb{T}^m) \times_\rho \mathbb{T}^n$, $X = Y \times_\sigma S^1$. The system $(T^k, \{1\} \times \mathbb{T}^m \times \mathbb{T}^n \times_\sigma S^1)$ is ergodic, and the system (X, T) consists of k copies of this system, where the transformation T moves us from one copy to the next. Thus we must understand ergodic systems of the form $(T, \mathbb{T}^m \times \mathbb{T}^n \times S^1)$, where the cocycle σ satisfies the functional equation 24. In this case the group $\mathcal{U}$ can be identified with $(\mathbb{Z} \times \mathbb{R}^m) \times \mathbb{R}^n$ with multiplication defined by $B = (B_1, \ldots, B_n)$, where $B_i \in M_{m+1,m+1}(\mathbb{Z})$ are bilinear forms with coefficients in $\mathbb{Z}$: for $r \in \mathbb{Z} \times \mathbb{T}^m$, $s \in \mathbb{R}^n$:

$$(r, s)(r', s') = (r'', s'')$$

where

$$\begin{aligned} r'' =& r + r' \\ s'' =& (s''_1, \ldots s''_n),\ s''_i = s_i + s'_i + B_i(r, r'). \end{aligned}$$

(see Rudolph [27]), and the action of T is given by the element $(\alpha, 0)$ ($\alpha = (1, \tilde{\alpha}) \in \mathbb{Z} \times \mathbb{R}^m$). We now have the vertical rotations V_s for $s \in \mathbb{R}^n$ (for $n \in \mathbb{Z}^n$, V_n acts trivially)

13.22. **Lemma.** *For all $s \in \mathbb{R}^n$ we can choose f_s, λ_s satisfying equation (26) such that*

$$f_{s_1}(\theta, \varphi + s_2) f_{s_2}(\theta, \varphi) \sim f_{s_1+s_2}(\theta, \varphi).$$

Proof. We now use lemma 6.10 to define f_s, λ_s for all $s \in \mathbb{R}^n$. Notice that multiplying f_h by a constant does not affect equation (26). If $s \in \mathbb{R}$, $s = s_1 + \ldots + s_k$, and $s_1, \ldots, s_k \in U$, define

$$f_s(\theta, \varphi) = f_{s_1}(\theta, \varphi) f_{s_2}(\theta, \varphi + s_1) \ldots f_{s_k}(\theta, \varphi + s_1 + \ldots + s_{k-1}).$$

We claim this is well defined (up to a constant multiple) on $\mathbb{R}^n$: Given two sequences $s_1, \ldots s_k$ and $s'_1, \ldots s'_l$ with equal sum, we can break up the "steps" s_j into an equal number of small steps and we can interpolate a sequence of such paths where two consecutive paths differ only within a small cube which can be translated to be inside U. Since the resulting λ's and f's will be the same for consecutive paths, they will be the same for the initial and final one.

□

13.23. **Lemma.** *(V_s, f_s), (V_t, f_t) commute.*

Proof. Iterate proposition 13.7. □

13.24. **Lemma.** *There exists a matrix $N \in M_{m+1,n}(\mathbb{Z})$, and a vector $j \in \mathbb{Z}^n$ such that*

$$\lambda_s = e^{2\pi i(N\alpha + j)\cdot s}$$

Proof. Now f_s is defined on $\mathbb{R}^n$,

$$\frac{\sigma(\theta,\varphi+s)}{\sigma(\theta,\varphi)} = \lambda_s \frac{Tf_s(\theta,\varphi)}{f_s(\theta,\varphi)}, \tag{48}$$

and satisfies the equations of lemma 6.10 for all $s,t \in \mathbb{R}^n$. λ_s is continuous and is thus of the form $e^{2\pi i rs}$. Let e_i denote the standard basis for $\mathbb{R}^n$. Each f_{e_i} is an eigenfunction (as the left side of equation (48) is 1); thus there exists $j(e_i) \in \mathbb{Z}^n$ such that $f_{e_i}(\theta,\varphi) = Ce^{2\pi i j(e_i)\theta}$, with eigenvalue $e^{2\pi i j(e_i)\alpha}$. Finally for each i there is $k_i \in \mathbb{Z}$ such that

$$re_i = j(e_i)\alpha + k_i.$$

$\square$

13.25. **Notation.** For $u = (r,s)$, $v = (r',s')$

$$B(u,v) := B(r,r') - B(r',r).$$

13.26. **Proposition.** *There exists a measurable family of pairs* $\{g_u, F_u\}$ *such that*

$$[(T,\sigma),(T_u,F_u)] = ([T,T_u], g_u(\theta) f_{B(u_\alpha,u)}(y)) \tag{49}$$

Proof. First note that if $u = (0,h)$ then equation (49) holds by equation (48) (using the formula in equation (37)). As $[T_u,V_s] = id_{\mathcal{U}}$, we have

$$\frac{V_s\frac{\sigma(T_u(\theta,\varphi))}{\sigma(\theta,\varphi)}}{\frac{\sigma(T_u(\theta,\varphi))}{\sigma(\theta,\varphi)}} = \frac{T_u\frac{\sigma(V_s(\theta,\varphi))}{\sigma(\theta,\varphi)}}{\frac{\sigma(V_s(\theta,\varphi))}{\sigma(\theta,\varphi)}} = \frac{T_u\frac{f_s(T(\theta,\varphi))}{f_s(\theta,\varphi)}}{\frac{f_sT(\theta,\varphi))}{f_s(\theta,\varphi)}}$$

$$= \frac{f_s(V_{B(u_\alpha,u)}T_uT(\theta,\varphi))}{f_s(T_uT(\theta,\varphi))} \frac{T\frac{f_s(T_u(\theta,\varphi))}{f_s(\theta,\varphi)}}{\frac{f_sT_u(\theta,\varphi)}{f_s(\theta,\varphi)}}$$

$$= \frac{f_{B(u_\alpha,u)}(V_sT_uT(\theta,\varphi))}{f_{B(u_\alpha,u)}(T_uT(\theta,\varphi))} \frac{T\frac{f_s(T_u(\theta,\varphi))}{f_s(\theta,\varphi)}}{\frac{f_sT_u(\theta,\varphi)}{f_s(\theta,\varphi)}}.$$

Thus

$$\frac{V_s(\frac{\sigma(T_u(\theta,\varphi))}{\sigma(\theta,\varphi)}\bar{f}_{B(u_\alpha,u)}(T_uT(\theta,\varphi)))}{(\frac{\sigma(T_u(\theta,\varphi))}{\sigma(\theta,\varphi)}\bar{f}_{B(u_\alpha,u)}(T_uT(\theta,\varphi)))} = \frac{T\frac{f_s(T_u(\theta,\varphi))}{f_s(\theta,\varphi)}}{\frac{f_sT_u(\theta,\varphi)}{f_s(\theta,\varphi)}}.$$

By corollary 6.8:

$$\frac{T_u\sigma(\theta,\varphi)}{\sigma(\theta,\varphi)}\bar{f}_{B(u_\alpha,u)}(T_uT(\theta,\varphi)) = \tilde{g}_u(\theta)\frac{TF_u(\theta,\varphi)}{F_u(\theta,\varphi)}. \tag{50}$$

Now define

$$g_u(\theta) := \tilde{g}_u((T_uT)^{-1}\theta),$$

and use equation (37). By proposition 6.16 there is a measurable choice of F_u, g_u. $\square$

13.27. We first show that $[(T_u,F_u),(V_s,f_s)]$ is in the center:

13.28. **Proposition.**

$$[(T_u, F_u), (V_s, f_s)] = (0, \lambda(u, s))$$

Proof. We show that

$$\frac{\frac{F_u(V_s y)}{F_u(y)}}{\frac{f_s(T_u y)}{f_s(y)}}$$

is a T-invariant function. As $g_u(\theta)$ is invariant under V_s:

$$\frac{T\frac{F_u(V_s y)}{F_u(y)}}{\frac{F_u(V_s y)}{F_u(y)}} = \frac{\bar{f}_{B(u_\alpha,u)}(T_u T V_s y)}{\bar{f}_{B(u_\alpha,u)}(T_u T y)} \frac{\frac{\sigma(T_u V_s y)}{\sigma(V_s y)}}{\frac{\sigma(T_u y)}{\sigma(y)}}$$

$$\frac{T\frac{f_s(T_u y)}{f_s(y)}}{\frac{f_s(T_u y)}{f_s(y)}} = \frac{f_s(T_u T y)}{f_s(T T_u y)} \frac{\frac{\sigma(T_u V_s y)}{\sigma(V_s y)}}{\frac{\sigma(T_u y)}{\sigma(y)}}.$$

By 13.23 (V_s, f_s),$(V_{B(u_\alpha,u)}, f_{B(u_\alpha,u)})$ commute, thus:

$$\frac{f_s(T_u T y)}{f_s(T T_u y)} = \frac{f_s(T_u T y)}{f_s(V_{B(u_\alpha,u)} T_u T y)} = \frac{f_{B(u_\alpha,u)}(T_u T y)}{f_{B(u_\alpha,u)}(T_u T V_s y)}$$

□

13.29. We must get some restrictions on $g_u(\theta)$, so we return to the original functional equation (exactly as we did in 13.10), and show that $g_u(\theta)$ satisfies a CL equation.

13.30. **Notation.**
$F_u^{\otimes} = \otimes_{j=1}^4 F_{ju}^{m_j}$
$f_{B(u_\alpha,u)}^{\otimes} = \otimes_{j=1}^4 f_{j^2 B(u_\alpha,u)}^{m_j}$
$g_u^{\otimes} = \otimes_{j=1}^4 g_{ju}^{m_j}$
$\sigma^{\otimes} = \otimes_{j=1}^4 \sigma_j^{m_j}$

We will need the following lemma:

13.31. **Lemma.** *Let F satisfy the functional equation (24). Then*

$$k(\tilde{y}) = \frac{F([\tau, \tau_u]\tilde{y})}{F(\tilde{y}) f_{B(u_\alpha,u)}^{\otimes}(\tilde{y})}$$

is a τ-invariant function.

Proof. As $\sum j^3 m_j = 0$,

$$\frac{f_{B(u_\alpha,u)}^{\otimes}(\tau\tilde{y})}{f_{B(u_\alpha,u)}^{\otimes}(\tilde{y})} = \frac{\sigma^{\otimes}([\tau, \tau_u]\tilde{y})}{\sigma^{\otimes}(\tilde{y})} = \frac{\frac{F(\tau[\tau,\tau_u]\tilde{y})}{F([\tau,\tau_u]\tilde{y})}}{\frac{F(\tau\tilde{y})}{F(\tilde{y})}} = \frac{\frac{F([\tau,\tau_u]\tau\tilde{y})}{F(\tau\tilde{y})}}{\frac{F([\tau,\tau_u]\tilde{y})}{F(\tilde{y})}}.$$

□

13.32. **Corollary.** *for any $v \in \mathcal{U}$: $k(\tau_v\tilde{y}) = k(\tilde{y})$.*

Proof. Use lemmas 13.31, 11.5. □

13.33. Proposition. *$g_u(\theta)$ satisfies CL.*

Proof. We apply τ_u to the functional equation (24), and divide the two. Using equation (50):

$$\frac{\frac{F(\tau\tau_u\tilde{y})}{F(\tau_u\tilde{y})}}{\frac{F(\tau\tilde{y})}{F(\tilde{y})}} = \frac{\sigma^{\otimes}(\tau_u\tilde{y})}{\sigma^{\otimes}(\tilde{y})} = g_u^{\otimes}(\tilde{\theta})f_{B(u_\alpha,u)}^{\otimes}(\tau_u\tau\tilde{y})\frac{F_u^{\otimes}(\tau\tilde{y})}{F_u^{\otimes}(\tilde{y})}$$

Thus,

$$g_u^{\otimes}(\tilde{\theta}) = \frac{F([\tau,\tau_u]\tau_u\tau\tilde{y})}{F(\tau_u\tau\tilde{y})f_{B(u_\alpha,u)}^{\otimes}(\tau_u\tau\tilde{y})}\frac{\frac{F(\tau_u\tau\tilde{y})}{F(\tau_u\tilde{y})}F_u^{\otimes}(\tau\tilde{y})}{\frac{F(\tau\tilde{y})}{F(\tilde{y})}F_u^{\otimes}(\tilde{y})}.$$

by lemma 13.31

$$\frac{F([\tau,\tau_u]\tau_u\tau\tilde{y})}{F(\tau_u\tau\tilde{y})f_{B(u_\alpha,u)}^{\otimes}(\tau_u\tau\tilde{y})}$$

is a τ invariant function, therefore it is constant on the ergodic components of τ. By lemma 11.5 it is invariant under τ_v for all $v \in \mathcal{U}$. We now have

$$\text{(51)}\quad \frac{g_u^{\otimes}(\tau_v\tilde{\theta})}{g_u^{\otimes}(\tilde{\theta})} = \frac{\frac{\frac{F(\tau_u\tau\tau_v\tilde{y})}{F(\tau_u\tau_v\tilde{y})}F_u^{\otimes}(\tau\tau_v\tilde{y})}{\frac{F(\tau\tau_v\tilde{y})}{F(\tau_v\tilde{y})}F_u^{\otimes}(\tau_v\tilde{y})}}{\frac{\frac{F(\tau_u\tau\tilde{y})}{F(\tau_u\tilde{y})}(F_u^{\otimes})(\tau\tilde{y})}{\frac{F(\tau\tilde{y})}{F(\tilde{y})}F_u^{\otimes}(\tilde{y})}} = \frac{\frac{\frac{F(\tau_u\tau_v\tau\tilde{y})}{F(\tau_u\tau_v\tilde{y})}F_u^{\otimes}(\tau_v\tau\tilde{y})}{\frac{F(\tau_v\tau\tilde{y})}{F(\tau_v\tilde{y})}F_u^{\otimes}(\tau_v\tilde{y})}}{\frac{\frac{F(\tau_u\tau\tilde{y})}{F(\tau_u\tilde{y})}F_u^{\otimes}(\tau\tilde{y})}{\frac{F(\tau\tilde{y})}{F(\tilde{y})}F_u^{\otimes}(\tilde{y})}}\frac{\frac{F(\tau_u\tau\tau_v\tilde{y})}{F(\tau\tau_v\tilde{y})}F_u^{\otimes}(\tau\tau_v\tilde{y})}{\frac{F(\tau_u\tau_v\tau\tilde{y})}{F(\tau_v\tau\tilde{y})}F_u^{\otimes}(\tau_v\tau\tilde{y})}.$$

Let's study the right part of the right expression.

$$\begin{aligned}\frac{\frac{F(\tau_u\tau\tau_v\tilde{y})}{F(\tau\tau_v\tilde{y})}F_u^{\otimes}(\tau\tau_v\tilde{y})}{\frac{F(\tau_u\tau_v\tau\tilde{y})}{F(\tau_v\tau\tilde{y})}F_u^{\otimes}(\tau_v\tau\tilde{y})} &= \frac{\frac{F(\tau_u\tau\tau_v\tilde{y})}{F(\tau\tau_v\tilde{y})}F_u^{\otimes}([\tau,\tau_v]\tau_v\tau\tilde{y})}{\frac{F(\tau_u\tau_v\tau\tilde{y})}{F(\tau_v\tau\tilde{y})}F_u^{\otimes}(\tau_v\tau\tilde{y})} \\ &= \frac{\frac{F(\tau_u\tau\tau_v\tilde{y})}{F(\tau\tau_v\tilde{y})}f_{B(u_\alpha,v)}^{\otimes}(\tau_u\tau_v\tau\tilde{y})}{\frac{F(\tau_u\tau_v\tau\tilde{y})}{F(\tau_v\tau\tilde{y})}f_{B(u_\alpha,v)}^{\otimes}(\tau_v\tau\tilde{y})}.\end{aligned}$$

But

$$\frac{F(\tau\tau_v\tilde{y})}{F(\tau_v\tau\tilde{y})}f_{B(u_\alpha,v)}^{\otimes}(\tau_v\tau\tilde{y})$$

is invariant under τ_u,$\tau_{[u,vu_\alpha]}$, thus the quotient is 1!. Return to equation (51) to get

$$\frac{g_u^{\otimes}(\tau_v\tilde{\theta})}{g_u^{\otimes}(\tilde{\theta})} = \frac{\tau G_v(\tilde{y})}{G_v(\tilde{y})}.$$

Now apply proposition 13.10. ☐

13.34. Proposition. *There exists a family of functions $\{F_u, g_u\}_{u\in\mathcal{U}}$ satisfying equation (49), and a family of functions functions $\{f_{u,v}(\theta)\}_{u,v\in\mathcal{U}}$ such that*

$$F_u^vF_v = f_{u,v}(\theta)F_{uv}; \quad g_u(\theta) = C_u$$

for u, v in a zero neighborhood of $\mathcal{U}$.

Proof. By equation (50)

$$\frac{\sigma(T_{uv}y)}{\sigma(y)} = g_{uv}(\theta) f_{B(u_\alpha,uv)}(T_{uv}Ty)\frac{F_{uv}(Ty)}{F_{uv}(y)}.$$

On the other hand:

$$\begin{aligned}\frac{\sigma(T_{uv}y)}{\sigma(y)} =&\frac{\sigma(T_uT_vy)}{\sigma(T_vy)}\frac{\sigma(T_vy)}{\sigma(y)}\\ =&g_u(T_v\theta)g_v(\theta)f_{B(u_\alpha,u)}(T_uTT_vy)f_{B(u_\alpha,v)}(T_vTy)\\ &\frac{F_u(TT_vy)}{F_u(T_vTy)}\frac{T(F_u^vF_v(y))}{(F_u^vF_v(y))}.\end{aligned} \tag{52}$$

$$\frac{F_u(TT_vy)}{F_u(T_vTy)} = \frac{F_u([T,T_v]T_vTy)}{F_u(T_vTy)} = \lambda(u, B(u_\alpha, v))\frac{f_{B(u_\alpha,v)}(T_uT_vTy)}{f_{B(u_\alpha,v)}(T_vTy)}.$$

By lemma 13.22

$$f_{B(u_\alpha,u)}(T_uTT_vy)f_{B(u_\alpha,v)}(T_uT_vTy) \sim f_{B(u_\alpha,uv)}(T_uT_vTy). \tag{53}$$

Therefore:

$$\frac{\sigma(T_{uv}y)}{\sigma(y)} \sim g_u(T_v\theta)g_v(\theta)f_{B(u_\alpha,uv)}(T_{uv}Ty)\frac{F_u^vF_v(Ty)}{F_u^vF_v(y)}. \tag{54}$$

Define

$$h_u(y) := \frac{\sigma(T_uy)}{\sigma(y)}\bar{f}_{B(u_\alpha,u)}(T_uTy).$$

Then

$$h_u(y) = g_u(\theta)\frac{F_u(Ty)}{F_u(y)},$$

and

$$h_{uv}(y) \sim h_u^v(y)h_v(y), \tag{55}$$

The functions h_u satisfy the conditions of proposition 13.18. Take $\tilde{g}_u(\theta) = C_u$ and $\tilde{F}_u(y) = G_u(y)$ from equation (43), and by equation (55), $\tilde{F}_{uv}/\tilde{F}_u^v\tilde{F}_v$ is an eigenfunction. □

13.35. Corollary. *The functions F_u can be chosen such that for all $u \in \mathcal{U}$, g_u is constant.*

Proof. The group $\mathcal{U}$ is generated by U, u_α. Now iterate using equation (54) with $g_u(\theta) = C(u)$ for u in a neighborhood of zero and choose $F_{u_\alpha} = \sigma$, $C(u_\alpha) = 1$. Using the argument in lemma 13.22 we get g_u constant for u in the connected component of the identity, then using u_α we get it for all $u \in \mathcal{U}$. □

13.36. Corollary. *The functions F_u are defined up to an eigenfunction .*

13.37. Remark. We can identify f_s with $F_{(0,s)}$.

13.38. Corollary.

$$\lambda(u, s + s') = \lambda(u, s)\lambda(u, s').$$

Proof. By multiplicativity of $F_{(0,s)}$ (up to constants). ☐

13.39. **Corollary.**

$$\lambda(uv,s)=\lambda(u,s)\lambda(v,s).$$

Proof. Eigenfunctions are invariant under V_s ☐

13.40. **Notation.** Denote

$$J(u,v,w)=\lambda(u,B(v,w))\lambda(w,B(u,v))\lambda(v,B(w,u)).$$

13.41. **Corollary.** *For $v,w\in\mathcal{U}$: $J(u_\alpha,w,v)=1$.*

Proof. By the Jacobi identity

$$\begin{aligned}&[[(T,\sigma),(T_v,F_v)],(T_w,F_w)][[(T_w,F_w),(T,\sigma)],(T_v,F_v)]\\&[[(T_v,F_v),(T_w,F_w)],(T,\sigma)]=1\end{aligned}$$

By corollary 13.36 there exists an eigenfunction of T - $\psi_{v,w}$ such that

$$F_v^w F_w=\psi_{v,w}F_{vw},$$

thus there exists an eigenfunction of T - $\tilde{\psi}_{v,w}$ such that

$$[(T_v,F_v),(T_w,F_w)]=(T_{B(v,w)},\tilde{\psi}_{v,w}F_{B(v,w)}),$$

therefore

$$\begin{aligned}[[(T_v,F_v),(T_w,F_w)],(T,\sigma)]=&[(T_{B(v,w)},\tilde{\psi}_{v,w}F_{B(v,w)}),(T,\sigma)]\\=&\frac{\tilde{\psi}_{v,w}(T\theta)}{\tilde{\psi}_{v,w}(\theta)}[(T_{B(v,w)},F_{B(v,w)}),(T,\sigma)]\end{aligned}$$

As $g_w(\theta)$ is constant,

$$[(T_w,F_w),(T,\sigma)]\sim(T_{B(w,u_\alpha)},F_{B(w,u_\alpha)}).$$

Thus $J(u_\alpha,w,v)$ is an eigenvalue of T. Linearity in v,w, gives the result in the connected component of the identity. As $B(u,v)$ is antisymmetric $J(u,u,v)=1$. Now

$$J(u_\alpha,u_\alpha w,v)=J(u_\alpha,u_\alpha,v)J(u_\alpha,w,v)=J(u_\alpha,w,v).$$

and u_α along with the connected component of the identity generate $\mathcal{U}$. ☐

13.42. **An Explicit Solution.** The situation is as follows. We are trying to construct a system (X,T) which is a circle extension of $\mathbb{T}^m\times\mathbb{T}^n$ by a cocycle σ which satisfies the equation:

$$\frac{V_h\sigma(\theta,\varphi)}{\sigma(\theta,\varphi)}=\lambda(u_\alpha,h)\frac{TF_{(0,h)}(\theta,\varphi)}{F_{(0,h)}(\theta,\varphi)}.$$

for all $h\in\mathbb{T}^n$. We will obtain a cocycle cohomologous to the original one "up to a CL character". This will then be modified to obtain a system isomorphic to the original one. We defined a function $\lambda(u,s)$ on $\mathcal{U}\times\mathbb{R}^n$ which coincides with $\lambda(u_\alpha,h)$ in a neighborhood of zero in $\mathbb{R}^n$, is multiplicative in both coordinates, and satisfies a 'Jacobi' equation (13.41).

13.43. **Proposition.** *There exists* $\tilde{N} \in M_{m+1,n}(\mathbb{Z})$*, such that if* $u = (r, s)$

$$\lambda(u, s') = e^{2\pi i \tilde{N}(r,s')}.$$

Proof. If $u = ((l, x), s)$, $l \in \mathbb{Z}$, $x \in \mathbb{R}^m$, then as in lemma 13.24 the fact that of f_{e_i} are eigenfunctions of T implies:

$$\lambda(u, s') = e^{2\pi i(\tilde{N}(u)(x,s')+j(u)s')}.$$

where $\tilde{N}(u) \in M_{n,m}(\mathbb{Z})$, $j(u) \in \mathbb{Z}^n$. For $u = ((l,0),0)$, $\tilde{N}(u) = 0$ thus multiplicativity of $\lambda(*, s')$ implies $\tilde{N}(u) = \tilde{N}$, and j is a linear function which depends only on l - thus determined by $j(1) = j_0$. Thus for $N = \binom{j_0}{\tilde{N}}$

$$\lambda(u, s') = e^{2\pi i N(r,s')}$$

☐

13.44. **Corollary.** *for all* $u, v, w \in \mathcal{U}$, $J(u, v, w) = 1$.

Proof. Let $\Gamma = \mathbb{Z}^{m+1+n}$ then the group generated by u_α, Γ is dense in $\mathcal{U}$. By the foregoing proposition, if $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{Z}^{m+1+n}$ then $J(\gamma_1, \gamma_2, \gamma_3) = 1$. The corollary follows from corollary 13.41. ☐

13.45. **Theorem.** *There exists a* 3*-step nilpotent group* $\mathcal{M}$*, a discrete subgroup* $\mathcal{L}$ *and an element* $a \in \mathcal{M}$*, such that the system* $(\mathcal{M}/\mathcal{L}, a)$ *is isomorphic to* (X, T)*.*

Proof. We construct the following 3-step nilpotent group:

$$\mathcal{M} = (\mathbb{Z} \times \mathbb{R}^n) \times \mathbb{R}^m \times S^1$$

with multiplication:

$$(r, s, \zeta)(r', s', \zeta') = ((r, s)(r', s'), \zeta\zeta' e^{2\pi i(\xi(r,s,r',s'))})$$

where

$$\xi(r, s, r', s') = N(r, s') + C(r, r', r) + D(r, r', r') + E(r, r'), \tag{56}$$

Where $C(r, r', r''), D(r, r', r'')$ trilinear forms, and E a bilinear form. In order for this multiplication to be associative C, D must satisfy:

$$N(r, B(r', r'')) = -D(r, r', r'') + C(r', r'', r) + (C - D)(r, r'', r').$$

For fixed i, j, k we must solve

$$-d_{ijk} - d_{ikj} + c_{jki} + c_{ikj} = l_{ijk}$$

where $l_{ijk} \in \mathbb{Z}$. The conditions of integrality on C and D are that $\chi(r, s, r', s')$ be an integer when the arguments are integer vectors. Denote

$$c'_{ijk} = c_{jki} + c_{ikj}$$

and similarly

$$d'_{ijk} = d_{ijk} + d_{ikj}.$$

By the Jacoby identity (see 13.41), the l_{ijk} satisfy

$$l_{ijk} + l_{jki} + l_{kij} - l_{kji} - l_{jik} - l_{ikj} = 0.$$

The system

$$c'_{ijk} + d'ijk = l_{ijk}$$

has a solution in integers with the right symmetry properties:

$$c'_{ijk} = c'_{jik}; \quad d'_{ijk} = d'_{ikj},$$

If i, j, k are different then we are done. If $i = j$, then c_{iji} may be a half integer. set

$$E_1(r, r') = -\sum c_{iji} r_i r'_j,$$

then for $r, r' \in \mathbb{Z}^{m+1}$,

$$\sum c_{iji} r_i r'_j r_i - \sum c_{iji} r_i r'_j \in \mathbb{Z}$$

thus

$$C(r, r', r) + E_1(r, r') \in \mathbb{Z}.$$

If $j = k$ do the same for D using a matrix E_2 and set $E = E_1 + E_2$. Now $\mathcal{M}$ is a 3-step nilpotent group, and $\mathcal{K} = \mathbb{Z}^{m+1} \times \mathbb{Z}^n \times \{1\}$ is a discrete subgroup of $\mathcal{M}$. Let $\mathcal{N} = \mathbb{Z} \times \mathbb{R}^m \times \mathbb{R}^n$ with the induced multiplication, $\mathcal{L} = \mathbb{Z}^{m+1} \times \mathbb{Z}^n$. The projection $\mathcal{M}/\mathcal{K} \to \mathcal{N}/\mathcal{L}$ is onto, and let

$$\tilde{L} : \mathcal{N} \to \mathcal{N}/\mathcal{L} \to \mathcal{M}/\mathcal{K}$$
$$(r, s) \to (r, s, L(r, s))\mathcal{K}$$

be a measurable section. We have

$$\begin{aligned}
&\tilde{L}((r, s)(m_1, m_2)) \\
&\quad =((r, s)(m_1, m_2), L((r, s)(m_1, m_2)))\mathcal{K} \\
&\quad =((r, s), L(r, s))\mathcal{K} \\
&\quad =((r, s), L(r, s))(m_1, m_2, 1)\mathcal{K} \\
&\quad =((r, s)(m_1, m_2), L(r, s)e^{2\pi i(\xi(r,s,m_1,m_2))})\mathcal{K}
\end{aligned}$$

Thus

$$L((r, s)(m_1, m_2)) = L(r, s)e^{2\pi i(\xi(r,s,m_1,m_2))}$$

The action of the element $(\alpha, 0, 1)$ is given by

$$(\alpha, 0, 1)(r, s, \zeta) = (r + \alpha, s + B(\alpha, r), \zeta e^{2\pi i(\xi(\alpha,0,r,s))})$$

Set

$$f(r, s) = (\frac{L(r, s)}{L(r + \alpha, s + B(\alpha, r))})e^{2\pi i(\xi(\alpha,0,r,s))},$$

then f is defined on $\mathcal{N}/\mathcal{L}$ and represents the action of $(\alpha, 0, 1)$:

$$\begin{aligned}
&(\alpha, 0, 1)(r, s, L(r, s)\zeta) \\
&\quad =(r + \alpha, s + B(\alpha, r), L(r, s)e^{2\pi i(\xi(\alpha,0,r,s))}\zeta) \\
&\quad =(r + \alpha, s + B(\alpha, r), f(r, s)L(r + \alpha, s + B(\alpha, r))\zeta).
\end{aligned}$$

Define

$$L_t(r, s) = \frac{L(r, s + t)}{L(r, s)},$$

then $L_t(r,s)$ is defined on $\mathcal{N}/\mathcal{L}$:

$$\frac{L_t((r,s)(m_1,m_2))}{L_t(r,s)} = \frac{e^{2\pi i \xi(r,s+t,m_1,m_2)}}{e^{2\pi i \xi(r,s,m_1,m_2)}} = 1$$

finally

$$\begin{aligned}\frac{f(r,s+t)}{f(r,s)} =& \frac{\frac{L(r,s+t)}{L(r+\alpha,s+B(\alpha,r)+t)}e^{2\pi i(\xi(\alpha,0,r,s+t))}}{\frac{L(r,s)}{L(r+\alpha,s+B(\alpha,r))}e^{2\pi i(\xi(\alpha,0,r,s))}}\\ =&\frac{L_t(r,s)}{L_t(r+\alpha,s+B(\alpha,r))}e^{2\pi i N(\alpha,t)}.\end{aligned}$$

The system we constructed is isomorphic to the original system up to a CL character. Let $\tilde{B}(r,r')$ denote the bilinear form representing this CL character. If we replace χ in equation (56) by

$$\xi(r,s,r',s') = N(r,s') + C(r,r',r) + D(r,r',r') + E(r,r') + \tilde{B}(r,r'),$$

we get the desired system. □

To complete the picture we must prove that a factor of a 3 step pronilsystem is also of this type.

13.46. Lemma. *Let $X = Y \times_\sigma J$ where Y is the m.c.f for schemes of length 3, J a compact Abelian group, and σ satisfies equation (24). Then (X,T) is isomorphic to a 3 step pro-nilsystem $\mathcal{M}/\mathcal{L} = \lim\leftarrow M_i/L_i$. Let K be a compact Abelian group of measure preserving transformations acting on X and commuting with the action of T. Then K commutes with the action of $\mathcal{M}$.*

Proof. We use additive notation for J. Denote T_k the action of the element $k \in K$. By corollary 8.5 the transformation $T_k : X \to X$ induces a map from $Y \to Y$. Denote t_k the action of k on Y, and by lemma 9.13 the action of K on $Y = \mathcal{N}/\mathcal{D}$ is by rotation by a central element of $\mathcal{N}$: For $k \in K$:

$$TT_k(y,j) = T(t_k y, f_k(y,j)) = (Tt_k y, \sigma(t_k y) + f_k(y,j)).$$

On the other hand,

$$T_k T(y,j) = T_k(Ty, j+\sigma(y))) = (t_k Ty, f_k(T(y,j))).$$

Thus

$$f_k(T(y,j)) = \sigma(t_k y) + f_k(y,j).$$

The transformation t_k is in the center of $\mathcal{U}$; thus $t_k y = V_{h_k} y$ for some $h_k \in H$ (see 9.9). Thus by equation (25)

$$T(f_k(y,j) - j - f_{h_k}(y)) = \lambda_{h_k} + f_k(y,j) - j - f_{h_k}(y),$$

which implies that for some eigenfunction $\psi_k(\theta)$

$$f_k(y,j) = j + f_{h_k}(y) + \psi_k(\theta) = j + f'_{h_k}(y).$$

The transformation T_k is thus of the form $T_k(y,j) = (t_k y, j + f'_{h_k}(y))$. For each character $\chi \in \hat{J}$, $\chi \circ \sigma$ is isomorphic to a cocycle which is lifted from a finite torus factor of Y, and as the action of K commutes with the action of $\mathcal{M}$ on Y we are reduced to the case where (X,T) is a nilflow. This case was treated in [26] theorem 4.3 (one can also show this directly as in 9.13). □

13.47. **Proposition.** *Any factor of a* 3*-step pro-nilsystem is a* 3*-step pro-nilsystem.*

Proof. Let $(X,\mathcal{B},T)$ be a 3-step pro-nilsystem, $(X',\mathcal{B}',T)$ a factor. The system $(X,\mathcal{B},T)$ has generalized discrete spectrum mod $\mathcal{B}'$ of finite type (see 7.4). Now apply the foregoing lemma, proposition 7.4 and induction. □

Part 3. A non-conventional ergodic theorem for a nilflow

14. INTRODUCTION

We prove an almost everywhere convergence theorem for nilflows. Let N be a k-step connected, simply connected nilpotent Lie group, Γ a discrete subgroup s.t. N/Γ is compact. Let $N_1 = N$, and for $i > 1$: $N_i = [N_{i-1}, N]$ (N is a k-step nilpotent group if $N_{k+1} = \{1\}$), and for $i \geq 1$ let $\Gamma_i = \Gamma \cap N_i$. Then Γ_i is a discrete subgroup of N_i, and N_i/Γ_i is compact (cf. Malcev [20]). Let m_i be the probability measure on N_i/Γ_i, invariant under translation by elements of N_i. Let $a \in N$, and for $x\Gamma \in N/\Gamma$, let $Tx\Gamma = ax\Gamma$. Then the measure preserving system $(N/\Gamma, m_1, T)$, is called a *nilflow.* In [17], E. Lesigne proved the following theorem for 2-step nilflows:

14.1. **Theorem.** *(Lesigne) Let $(N/\Gamma, m_1, T)$ be a 2-step nilflow. Suppose T acts ergodically on N/Γ, $f_1, f_2, f_3 \in L^\infty(N/\Gamma)$, then for almost all $x \in N$*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T^n f_1(x\Gamma) T^{2n} f_2(x\Gamma) T^{3n} f_3(x\Gamma)$$
$$= \int_{N/\Gamma} \int_{N_2/\Gamma_2} f_1(xy_1\Gamma) f_2(xy_1^2 y_2\Gamma) f_3(xy_1^3 y_2^3\Gamma) dm_1(y_1\Gamma_1) dm_2(y_2\Gamma_2)$$

We generalize this result and prove the following ergodic theorem:

14.2. **Theorem.** *Let $(N/\Gamma, m_1, T)$ be a k-step nilflow. Suppose T acts ergodically on N/Γ. If $f_1, \ldots, f_{k+1} \in L^\infty(N/\Gamma)$, then for almost all $x \in N$*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k+1} T^{jn} f_j(x\Gamma)$$
$$= \int_{N/\Gamma} \cdots \int_{N_k/\Gamma_k} \prod_{j=1}^{k+1} f_j(x \prod_{i=1}^{j} y_i^{\binom{j}{i}} \Gamma) \quad \prod_{j=1}^{k+1} dm_j(y_j\Gamma_j)$$

(note that as $N_{k+1} = \{1\}$, $y_{k+1} = 1$ and the measure m_{k+1} is trivial).

The proof is a combination of the proof of Lesigne with the work of Leibman [14] on geometric sequences in groups. Let $x \in N$, and denote A_x the closure of the orbit of $(x, \ldots, x)$ in $(N/\Gamma)^{k+1}$ under $T \times \ldots \times T^{k+1}$. We show that the system $(A_x, T \times \ldots \times T^{k+1})$ is isomorphic to some nilflow $(\tilde{N}/\tilde{\Gamma}, S_x)$. We then use the fact that for nilflowss ergodicity implies unique ergodicity. The main technical difficulty will be with proving that the transformation S_x is ergodic on $\tilde{N}/\tilde{\Gamma}$ for almost all (Haar) $x \in N$.

15. Proof of Theorem 14.2

We define a set $\tilde{N}$ by $\tilde{N} = N_1 \times \ldots \times N_k$. For $(x_1, \ldots, x_k), (y_1, \ldots, y_k) \in \tilde{N}$ we define

$$z_1 = x_1 y_1,$$

$$\prod_{j=1}^{i} z_j^{\binom{i}{j}} = \prod_{j=1}^{i} x_j^{\binom{i}{j}} \prod_{j=1}^{i} y_j^{\binom{i}{j}} \qquad 2 \le i \le k \; .$$

$$(x_1, \ldots, x_k) \star (y_1, \ldots, y_k) = (z_1, \ldots, z_k),$$

One should think of $(x_1, \ldots, x_k), (y_1, \ldots, y_k)$ as representing the vectors

$$(x_1, x_1^2 x_2, \ldots, \prod_{j=1}^{k} x_j^{\binom{k}{j}}), \quad (y_1, y_1^2 y_2, \ldots, \prod_{j=1}^{k} y_j^{\binom{k}{j}}) \tag{57}$$

respectively. Then $(z_1, \ldots, z_k)$ represents the coordinate product of these vectors. By Leibman [14] (corollary 1.3), $\tilde{N}$ is a group under the multiplication defined above. Now $\tilde{N}$ is a nilpotent connected simply connected Lie group. Let $\tilde{\Gamma} = \Gamma_1 \times \ldots \times \Gamma_k$, then $(\tilde{\Gamma}, \star)$ is a discrete subgroup of $\tilde{N}$, and the quotient $\tilde{N}/\tilde{\Gamma}$ is compact. For $x \in N$ we define a transformation $S_x : \tilde{N}/\tilde{\Gamma} \to \tilde{N}/\tilde{\Gamma}$ by:

$$S_x((y_1, \ldots, y_k) \star \tilde{\Gamma}) = ((a[a, x], e, \ldots, e) \star (y_1, \ldots, y_k)) \star \tilde{\Gamma},$$

and a mapping $I_x : \tilde{N} \to (N/\Gamma)^{k+1}$ by:

$$I_x(y_1, \ldots, y_k) = (x y_1 \Gamma, x y_1^2 y_2 \Gamma, \ldots, x \prod_{j=1}^{k} y_j^{\binom{k+1}{j}} \Gamma).$$

Let $(\gamma_1, \ldots, \gamma_k) \in \tilde{\Gamma}$ then we have for all $(y_1, \ldots, y_k) \in \tilde{N}$:

$$\begin{aligned} I_x((y_1, \ldots, y_k) \star (\gamma_1, \ldots, \gamma_k)) & \\ &= (x y_1 \gamma_1 \Gamma, x y_1^2 y_2 \gamma_1^2 \gamma_2 \Gamma, \ldots, x \prod_{j=1}^{k} y_j^{\binom{k+1}{j}} \prod_{j=1}^{k} \gamma_j^{\binom{k+1}{j}} \Gamma) \\ &= (x y_1 \Gamma, x y_1^2 y_2 \Gamma, \ldots, x \prod_{j=1}^{k} y_j^{\binom{k+1}{j}} \Gamma) \\ &= I_x(y_1, \ldots, y_k), \end{aligned}$$

thus I_x can be defined on $\tilde{N}/\tilde{\Gamma}$. Let $\tilde{m}$ denote the probability measure on $\tilde{N}/\tilde{\Gamma}$, invariant under the action of $\tilde{N}$. By uniqueness of the Haar measure $\tilde{m} = m_1 \times \ldots \times m_k$. Denote by m_x, A_x the images of $\tilde{m}$, $\tilde{N}/\tilde{\Gamma}$ under I_x. The mapping I_x is an isomorphism of the systems $(A_x, m_x, T \times T^2 \times \ldots \times T^{k+1})$

and $(\tilde{N}/\tilde{\Gamma}, \tilde{m}, S_x)$:

$$\begin{aligned}
&I_x(S_x((y_1, \dots, y_k) \star \tilde{\Gamma})) \\
&\quad = I((a[a,x], e, \dots, e) \star (y_1, \dots, y_k) \star \tilde{\Gamma}) \\
&\quad = (xa[a,x]y_1\Gamma, x(a[a,x])^2 y_1^2 y_2\Gamma, \dots, x(a[a,x])^{k+1} \prod_{j=1}^{k+1} y_j^{\binom{k+1}{j}} \Gamma) \\
&\quad = (axy_1\Gamma, a^2xy_1^2y_2\Gamma, \dots, a^k x \prod_{j=1}^{k} y_j^{\binom{k}{j}} \Gamma, a^{k+1}x \prod_{j=1}^{k+1} y_j^{\binom{k+1}{j}} \Gamma) \\
&\quad = (T \times T^2 \times \dots \times T^{k+1})(I_x((y_1, \dots, y_k) \star \tilde{\Gamma}))
\end{aligned}$$

(note that $y_{k+1} \in N_{k+1} = \{1\}$ thus $y_{k+1} = 1$; we use it just to simplify the notation). If the action of S_x is ergodic, then by Parry [24] (theorem 4), the m.p.s $(\tilde{N}/\tilde{\Gamma}, \tilde{m}, S_x)$ is uniquely ergodic (Parry shows that unique ergodicity is equivalent to minimality of the action on $\tilde{N}/[\tilde{N}, \tilde{N}] \star \tilde{\Gamma}$ which is a torus), and thus the m.p.s $(A_x, m_x, T \times T^2 \times \dots \times T^{k+1})$ is uniquely ergodic. For each x such that S_x is ergodic we have: for all $(y_1, \dots, y_k) \in \tilde{N}$ and for all continuous functions F on $(N/\Gamma)^{k+1}$,

$$\begin{aligned}
&\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} F(a^n xy_1\Gamma, a^{2n}xy_1^2y_2\Gamma, \dots, a^{(k+1)n} x \prod_{j=1}^{k+1} y_j^{\binom{k+1}{j}} \Gamma) \\
&= \int_{N/\Gamma} \cdots \int_{N_k/\Gamma_k} F(xy_1\Gamma, xy_1^2y_2\Gamma, \dots, x \prod_{j=1}^{k+1} y_j^{\binom{k+1}{j}} \Gamma) \prod_{j=1}^{k+1} dm_j(y_j\Gamma_j).
\end{aligned}$$

In particular, for all such functions F

(58)

$$\begin{aligned}
&\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} F(a^n x\Gamma, a^{2n}x\Gamma, \dots, a^{(k+1)n}x\Gamma) \\
&= \int_{N/\Gamma} \cdots \int_{N_k/\Gamma_k} F(xy_1\Gamma, xy_1^2y_2\Gamma, \dots, x \prod_{j=1}^{k+1} y_j^{\binom{k+1}{j}} \Gamma) \prod_{j=1}^{k+1} dm_j(y_j\Gamma_j)
\end{aligned}$$

(once again note that $y_{k+1} \in N_{k+1} = \{1\}$, and μ_{k+1} is trivial). If S_x is ergodic then we proved the theorem for $f_1, \dots, f_{k+1}$ continuous (take $F = f_1 \otimes \dots \otimes f_{k+1}$). As the subspace of continuous functions is dense in $L^2(N/\Gamma)$ equation (58) holds for all bounded measurable functions.

In order to finish the proof of the theorem, we must show that the action of S_x is ergodic for almost all (Haar) $x \in N$.

16. ERGODICITY OF S_x

16.1. We wish to show that the action of S_x on $\tilde{N}/\tilde{\Gamma}$ is ergodic for almost all (Haar) $x \in N$. Let $\tilde{N}'$ denote the commutator subgroup of $\tilde{N}$. Let σ be a character of $\tilde{N}$ s.t. σ vanishes on $\tilde{N}' \star \tilde{\Gamma}$. By Green [1] we must show that if $\sigma \not\equiv 1$ then

$$m_h\{x \in N : \sigma((a[a,x],e,\dots,e)) = 1\} = 0$$

(as there is only a countable number of characters on $\tilde{N}/\tilde{N}'\tilde{\Gamma}$). Suppose this set is of positive measure. Let

$$\psi(x) = \sigma((a[a,x],e,\dots,e)),$$

then ψ is an analytic function on N, therefore $\psi(x) = 1$ on a set of positive measure implies $\psi \equiv 1$. For $1 \leq i \leq k$ we define the functions σ_i on N_i by

$$\sigma_i(y) = \sigma((e,\dots,e,y,e,\dots,e)) \quad \text{where } y \text{ is in the } i\text{'th place.}$$

We have

$$\sigma(y_1,\dots,y_k) = \prod_{i=1}^{k} \sigma_i(y_i).$$

We will show that if $\psi(x) \equiv 1$ then $\sigma_i \equiv 1$ for $1 \leq i \leq k$, and thus $\sigma \equiv 1$.

16.2. W.l.o.g $k \geq 2$. For $n \in \mathcal{N}$ and $y_1,\dots,y_n \in N$ we define $R(y_1,\dots,y_n)$ as follows:

$$\begin{aligned} R_0(y_1,\dots,y_n) &= \{y_1,\dots,y_n\} \\ R_j(y_1,\dots,y_n) &= \{[R_i,R_l] : 0 \leq i,l \leq j-1\} \\ R(y_1,\dots,y_n) &= \bigcup_{i=1}^{k-1} R_i(y_1,\dots,y_n). \end{aligned}$$

In other words, $R(y_1,\dots,y_n)$ is the set of commutators involving $y_1,\dots,y_n$ (note that as N is k-step nilpotent $R_{k-1} = R_k$).

16.3. Example. For $k = 3$, $n = 2$, $x,y \in N$

$$R(x,y) = \{[x,y]^{\pm 1}, [[x,y]^{\pm 1},y]^{\pm 1}, [[x,y]^{\pm 1},x]^{\pm 1}\}.$$

16.4. Lemma. *If $x,y \in N_i$ then*

$$\sigma_i(xy) = \sigma_i(x)\sigma_i(y)\theta_i(x,y)$$

where $\theta_i(x,y) = \prod_{j>i} \sigma_j(r_j(x,y))$, where $r_j(x,y)$ is a finite product of elements from $R(x,y) \cap N_j$.

Proof. Recall from equation (57) that $(e,\dots,e,x,e,\dots,e)$ represents the vector $(e,\dots,e,x,x^{\binom{i+1}{i}},\dots,x^{\binom{k}{i}})$. Then

$$\begin{aligned} \sigma_i(x)\sigma_i(y) =& \sigma(e,\dots,e,x,e,\dots,e)\sigma(e,\dots,e,y,e,\dots,e) \\ =& \sigma((e,\dots,e,x,e,\dots,e) \star (e,\dots,e,y,e,\dots,e)) \\ =& \sigma(e,\dots,e,xy,r_{i+1}(x,y),\dots,r_k(x,y)) \end{aligned}$$

where $r_j(x,y)$ is a product of elements from $R(x,y) \cap N_j$ for $i < j \leq k$. ☐

16.5. Example. For $k = 2$, $x,y \in N$

$$\sigma_1(xy) = \sigma_1(x)\sigma_1(y)\sigma_2^{-1}([x,y]).$$

16.6. Assume $\psi \equiv 1$. Our aim is to show that σ is trivial. Using lemma 16.4 we have: for all x

$$\begin{aligned} 1 = \psi(x) &= \sigma((a[a,x], e, \dots, e)) = \sigma_1(a[a,x]) \\ &= \sigma_1(a)\sigma_1([a,x])\sigma_2(r_2(x)) \dots \sigma_k(r_k(x)), \end{aligned}$$

where $r_i(x)$ are products of elements of $R(a, [a,x])$.

16.7. Setting $x = e$ we get $\sigma_1(a) = 1$, thus

$$\begin{aligned} \sigma_1([a,x]) &= \sigma_2(r_2(x))^{-1} \dots \sigma_k^{-1}(r_k(x)) \\ &= \prod_{i=2}^{k} \prod_{m=1}^{l_i} \sigma_i(r_{i_m}(x)), \end{aligned} \tag{59}$$

for all $x \in N$, where $r_{i_m}(x) \in R(a, [a,x])$ (as $R(a, [a,x])$ is closed under the operation of taking commutators).

16.8. We will need the following lemma, which is based on the fact that $\sigma \equiv 1$ on $\tilde{N}'$:

16.9. Lemma. *Let $x \in N$, $y \in N_n$. We define*

$$H(x,y) = \bigcup_{i=2}^{k-1} R_i(x,y)$$

If $\sigma_i(h) = 1$ $\forall i > 1$, $h \in H(x,y) \cap N_i$, then $\sigma_m([x,y])^m \sigma_{m+1}([x,y])^{m+1} = 1$, for all $m \leq n$.

Proof: We denote by $\tilde{H}(x,y)$ the set of finite products of elements of $H(x,y)$. By lemma (16.4) and since $H(x,y)$ is closed under commutators, $\sigma_i(\tilde{h}) = 1$ $(i > 1)$ for all $\tilde{h} \in \tilde{H}(x,y) \cap N_i$. As $\sigma|_{\tilde{N}'} \equiv 1$, we have

$$\begin{aligned} 1 &= \sigma((x^{-1}, e, \dots, e) \star (e, \dots, \overset{m}{y^{-1}}, \dots, e) \star (x, e, \dots, e) \star (e, \dots, \overset{m}{y}, \dots, e)) \\ &= \sigma(z_1, \dots, z_k) = \prod_{i=1}^{k} \sigma_i(z_i), \end{aligned}$$

we calculate the z_i:

$$\begin{aligned} z_1 = \dots = z_{m-1} &= e, \\ z_m = [x^m, y] &= [x,y]^m \tilde{h_0} \end{aligned}$$

for some $\tilde{h_0} \in \tilde{H}(x,y)$.

$$\begin{aligned} &z_m^{(m+1)} z_{m+1} = [x^{(m+1)}, y^{(m+1)}] = [x,y]^{(m+1)^2} \tilde{h_1'} \qquad \Rightarrow \\ &z_{m+1} = [x,y]^{(m+1)} \tilde{h_1} \end{aligned}$$

for some $\tilde{h_1}, \tilde{h'_1} \in \tilde{H}(x,y)$. We show by induction that $z_{m+j} = \tilde{h_j}$ for $j > 1$, $\tilde{h_j} \in \tilde{H}(x,y)$. For $j = 2$:

$$z_m^{\binom{m+2}{2}} z_{m+1}^{(m+2)} z_{m+2} = [x^{(m+2)}, y^{\binom{m+2}{2}}] = [x,y]^{(m+2)\binom{m+2}{2}} \tilde{h'_2} \qquad \Rightarrow$$
$$z_{m+2} = [x,y]^{-m\binom{m+2}{2}-(m+1)(m+2)+(m+2)\binom{m+2}{2}} \tilde{h_2} = \tilde{h_2}$$

for some $\tilde{h_2}, \tilde{h'_2} \in \tilde{H}(x,y)$. Suppose it is true for $2 \le i < j$, we have

$$z_m^{\binom{m+j}{j}} z_{m+1}^{\binom{m+j}{j-1}} \dots z_{m+j} = [x^{(m+j)}, y^{\binom{m+j}{j}}] = [x,y]^{(m+j)\binom{m+j}{j}} \tilde{h'_j} \qquad \Rightarrow$$
$$z_{m+j} = [x,y]^{-m\binom{m+j}{j}-(m+1)\binom{m+j}{j-1}+(m+j)\binom{m+j}{j}} \tilde{h_j} = \tilde{h_j}$$

for some $\tilde{h_j}, \tilde{h'_j} \in \tilde{H}(x,y)$. Thus

$$1 = \prod_{i=1}^{k} \sigma_i(z_i) = \sigma_m([x,y]^m \tilde{h}_0)\sigma_{m+1}([x,y]^{m+1}\tilde{h}_1) \prod_{i=m+2}^{k} \sigma_i(\tilde{h}_{i-m})$$
$$= \sigma_m([x,y]^m)\sigma_{m+1}([x,y]^{m+1}),$$

as $\tilde{H}([x,y],\tilde{h}_0)$, $\tilde{H}([x,y],\tilde{h}_1) \subset \tilde{H}(x,y)$. □

16.10. **Corollary.** *If* $\sigma_1([x,y]) = 1$, *and* $\sigma_i(h) = 1$ *for all* $h \in H(x,y) \cap N_i$ *then* $\sigma_i([x,y]) = 1$ *for all* i *s.t.* $[x,y] \in N_i$.

Proof: By lemma 16.9, $\sigma_i([x,y])^i = 1$. N is connected, and $\sigma_i([e,y]) = 1$. □

16.11. **Notation.** We use $[x_1,\dots,x_n]$ for $[\dots[[[x_1,x_2],x_3],x_4],\dots,x_n]$.

16.12. **Notation.** Let $\mathcal{S} = \{(s_1,\dots,s_j) : 0 \le j \le k, 1 \le s_n \le k\}$. For $A = (s_1,\dots,s_n) \in \mathcal{S}$ define C_A by:

$$C_A^0 = a$$
$$C_A^j = \{[C_A^{j-1}, x_j]^{\pm 1} : x_j \in N_{s_j}\}$$
$$C_A = C_A^n$$

16.13. **Example.** If $A = (1,3)$ then

$$C_A = \{[[a,x]^{\pm 1}, y]^{\pm 1} : x \in N, y \in N_3\}.$$

16.14. **Notation.** (1) For $A = (s_1,\dots,s_j) \in \mathcal{S}$, $|A| = \sum s_j$

(2) For $A \in \mathcal{S}$, $(i : A)$ means: $\forall v \in (C_A \cap N_i) : \sigma_i(v) = 1$ (if $|A| < i - 1$ the intersection is empty).

(3) $j^{\geq}$ $(>)$ is some $l \ge j$ $(l > j)$.

(4) Let X be either j, $j^{\geq}$, $j^{>}$, then $X+$ is a sequence of length ≥ 1, whose entries are of type j, $j^{\geq}$, $j^{>}$ respectively.

(5) For $x \in N_j$, $|x| = j$.

16.15. **Proposition.** *For all* $A \in \mathcal{S}$, $1 \leq i \leq k$ *we have* $(i : A)$.

Proof: We will be using the following induction rules (for example rule (2) in the next lemma means that if for all A of the form $A = (s_1, \dots, s_n)$ where $n \geq 1$, $s_l > j$ for $1 \leq l \leq n$, we have for all $i \geq 1$, $v \in C_A \cap N_i$, $\sigma_i(v) = 1$ then for all $w \in C_{(j)}$ we have $\sigma_1(w) = 1$).

16.16. **Lemma.** *We prove the following rules:*

(1) $|A| \geq k \Longrightarrow \forall i\ (i : A)$. *(triviality rule)*
(2) $\forall i\ (i : j^{>}+) \Longrightarrow (1 : j)$. *(reducing rule)*
(3) $(1 : (A, l))$ *&* $j > l$ *&* $\forall i\ (i : (A, l, j^{\geq}+))$ *&* $\forall i\ (i : (A, j-1, j^{\geq}+)) \Longrightarrow (1 : (A, l, j-1))$ *&* $(1 : (A, j-1, l))$. *(expanding rule)*
(4) $(1 : (A, j))$ *&* $\forall i\ (i : (A, j^{\geq}, j^{\geq}+))$ *&* $\forall i\ (i : (A, j^{>})) \Longrightarrow \forall i\ (i : (A, j))$.

Proof: Let $A = (s_1, \dots, s_j)$, $0 \leq j \leq k$, $x_n \in N_{s_n}$.

(1) $|A| \geq k$ implies that $[a, x_1, \dots, x_j] = 1$ thus $\sigma_i([a, x_1, \dots, x_j]) = 1$.
(2) Let $x \in N_j$.
We use equation 59:

$$\sigma_1([a, x]) = \prod_{i=2}^{k} \prod_{m=1}^{l_i} \sigma_i(r_{i_m}(x)),$$

where $r_{i_m}(x) \in R(a, [a, x])$. Denote $b = [a, x]$, then $r_{i_m}(x) \in R(a, b)$. Looking at the a appearing in a minimal number of brackets, $r_{i_m}(x)$ is of the form: $[[a, y_1]^{\pm 1}, \dots, y_m]$ where $m \geq 1$, $|y_n| \geq |b| > |x|$, and is thus of type $(j^{>}+)$.
(3) Let $v = [a, x_1, \dots, x_j]$, $y \in N_l, z \in N_{j-1}$ then $yz \in N_l$.

$$\begin{aligned} 1 &= \sigma_1([v, yz]) = \sigma_1([v, y][v, y, z][v, z]) \\ &= \sigma_1([v, y])\sigma_1([v, y, z])\sigma_1([v, z]) \prod_{i=2}^{k} \prod_{m=1}^{l_i} \sigma_i(b_{i_m}(y, z, v)) \\ &= \sigma_1([v, y, z]) \prod_{i=2}^{k} \prod_{m=1}^{l_i} \sigma_i(b_{i_m}(y, z, v)), \end{aligned}$$

where $b_{i_m} \in R([v, y], [v, z], [v, y, z])$. If $[v, y]$ appears in $b_{i_m}(y, z, v)$ then w.l.o.g.

$$b_{i_m}(y, z, v) = [v, y, y_1, \dots, y_m],$$

where $m \geq 1$, $|y_n| \geq j$ (look at the $[v, y]$ appearing in the minimal number of brackets) This is of type $(A, l, j^{\geq}+)$. Otherwise $[v, y]$ does not appear in $b_{i_m}(y, z, v)$, thus w.l.o.g.

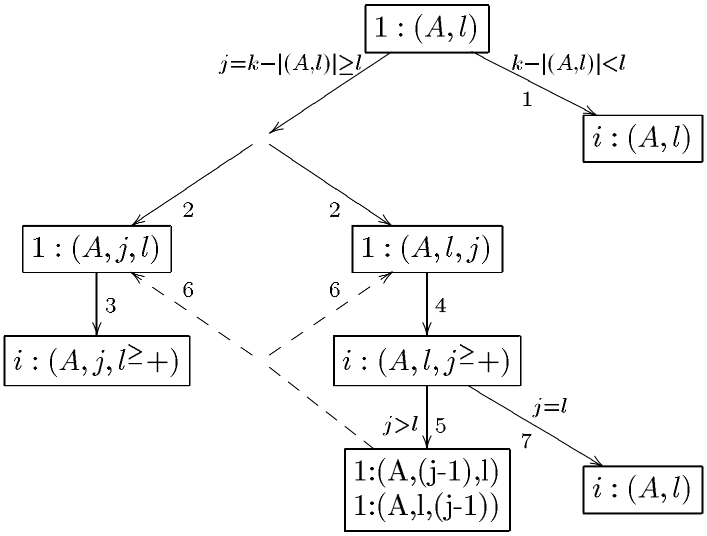
$$b_{i_m}(y, z, v) = [v, z, z_1, \dots, z_m],$$

where $m \geq 1$, $|z_n| \geq j$. This is of type $(A, j-1, j^{\geq}+)$.
For $(A, j-1, l)$, do the same with $[v, zy]$.

(4) Let v be of type A, $x \in N_j$. By lemma 16.9, we must show that for $h \in H(v,x) \cap N_i$, $\sigma_i(h) = 1$. But any such h is either of the form $[v, r(v,x)]$ where $r(v,x) \in R(v,x)$ and thus of type $(A, j^{>})$, or $[v, z_1, \ldots, z_m]$, $m \geq 1$, $|z_n| \geq j$, and thus of type $(A, j^{\geq}, j^{\geq}+)$.

16.17. **Proposition.** $(1 : (A,l)) \Longrightarrow (i : (A, l^{\geq}+))$, *where* $|A| \geq 0$, $l \geq 1$.

Proof: We prove this by induction on $|A|$, l. For $|A| = k$, and for any l, it is clear (as the commutator is trivial). Suppose the statement is true for $|A| > d$, any l. For $|A| = d$, $l = k - d$ it is clear (as the commutator is trivial). Suppose the statement is true for $|A| = d$, any $j > l$. We follow the following scheme:



Explanation for numbered arrows::

(1) $k - |(A,l)| < l$ implies $(i : (A, l^{\geq}, l^{\geq}+))$ by the triviality rule 1. $(1 : (A,l))$ implies $(1 : (A,j))$ for $j > l$ (as $x \in N_j \Rightarrow x \in N_l$) By the induction hypothesis $(i : (A,j))$ for $j > l$, and by rule 4 $(i : (A,l))$.
(2) By expanding rule 3 and triviality rule 1.
(3) By the induction hypothesis (as $|A,j| > |A|$).
(4) By the induction hypothesis (as $|A,l| > |A|$).
(5) By expanding rule: $(j-1) \geq l$, $(1 : (A,l))$, $(i : (A, l, j^{\geq}+))$, and $(i : (A, (j-1), j^{\geq}+))$ as if $j - 1 = l$ it is the previous condition, and if $j - 1 > l$ then $(1 : (A,l))$ implies $(1 : (A, j-1))$ which implies by the induction $(i : (A, j - 1^{\geq}+))$.
(6) Repeat the procedure until $j = l$.
(7) By this procedure we now know $(i : A, l^{\geq}, l^{\geq}+)$. Repeat the argument in 1.

Proof of proposition 16.15 We prove this by induction on j: $(1 : k)$ by triviality rule 1. $(1:j) \Rightarrow \forall i\ (i : j^{\geq}+)$ by proposition 16.17 ($|A| = 0$). By rule 2 this implies $(1 : j-1)$. ☐

16.18. **Proposition.** *For all i, $\sigma_i \equiv 1$.*

Proof: We prove this by induction on k. By proposition 16.15, we have

$$\sigma_k([x_1, \dots, \overset{m}{a}, \dots, x_{k-1}]) = 1 \qquad \forall x_1, \dots, x_{k-1} \in N, \quad 1 \leq m \leq k \tag{60}$$

Since $\sigma|_{\tilde{\Gamma}} \equiv 1$,

$$\sigma_k([\gamma_1, \dots, \gamma_k]) = 1 \qquad \forall \gamma_1, \dots, \gamma_k \in \Gamma.$$

From (60)

$$\sigma_k([a, \gamma_2, \dots, \gamma_k]) = 1 \qquad \forall \gamma_2, \dots, \gamma_k \in \Gamma.$$

Since σ_k is a character on N_k, and a acts ergodically on N/Γ, we have

$$\sigma_k([x_1, \gamma_2, \dots, \gamma_k]) = 1 \qquad \forall \gamma_2, \dots, \gamma_k \in \Gamma, \quad x_1 \in N.$$

Using (60) and induction, we get

$$\sigma_k([x_1, \dots, x_k]) = 1 \qquad \forall x_1, \dots, x_k \in N,$$

thus $\sigma_k \equiv 1$. Now $\sigma_k, \dots, \sigma_{j+1} \equiv 1$ implies that σ_j is a character. For $j > 1$, by Proposition 16.15, we have

$$\sigma_j([x_1, \dots, \overset{m}{a}, \dots, x_{j-1}]) = 1 \qquad \forall x_1, \dots, x_{j-1} \in N, \quad 1 \leq m \leq j,$$

and by the same argument as above we have $\sigma_j \equiv 1$. For $j = 1$, σ_1 is a character which satisfies

$$\sigma_1(a) = 1, \qquad \sigma_1|_\Gamma \equiv 1,$$

thus $\sigma_1 \equiv 1$. ☐

References

[1] Auslander, L., Green, L., Hahn, F. *Flows on homogeneous spaces.* Annals of Mathematics Studies, No. 53 (1963).

[2] Bergelson, V. *Weakly mixing PET.* Erg. Th. & Dyn. Sys. 7: 337-349 (1987).

[3] Becker, H. Kechris, S. *The Descriptive Set Theory of Polish Group Actions.* London Math. Society Lecture Notes Series 232 (1996).

[4] Bourgain, J. *Pointwise ergodic theorems for arithmetic sets.* Inst. Hautes Études Sci. Publ. Math. No. 69: 5-45 (1989).

[5] Conze, J.P., Lesigne, E. *Théorèmes ergodiques pour des mesures diagonales.* Bull. Soc. Math. France 112 no. 2: 143-175 (1984).

[6] Conze, J.P., Lesigne, E. *Sur un théorème ergodique pour des mesures diagonales.* Probabilités, Publ. Inst. Rech. Math. Rennes: 1-31 (1987).

[7] Conze, J.P., Lesigne, E. *Sur un théorème ergodique pour des mesures diagonales.* C. R. Acad. Sci. Paris Sér. I Math. 306, no. 12: 491-493 (1988).

[8] Furstenberg, H. *Ergodic behavior of diagonal measures and a theorem of Szemèredi on arithmetic progressions.* J. Analyse Math. 31: 204-256 (1977).

[9] Furstenberg, H. *Recurrence in Ergodic Theory and Combinatorial Number Theory.* Princeton Univ. Press (1981).

[10] Furstenberg, H., Weiss B. *A mean ergodic theorem for* $(1/N)\sum_{n=1}^{N} f(T^n x)g(T^{n^2} x)$. Convergence in ergodic theory and probability Ohio State Univ. Math. Res. Inst. Publ., 5, de Gruyter, Berlin: 193-227 (1996).

[11] Host,B.; Kra, B. *Convergence of Conze-Lesigne averages.* Erg. Th. & Dyn. Sys. 21, no. 2: 493–509 (2001).

[12] Host,B.; Kra, B. *An odd Furstenberg-Szemerdi theorem and quasi-affine systems.* J. Anal. Math. 86: 183–220 (2002).

[13] Host, B. , Kra, B : personal communication.

[14] Leibman, A *Polynomial sequences in groups.* Journal of Algebra, 201: 189-206 (1998).

[15] Lesigne, E. *Résolution d'une équation fonctionnelle.* Bull. Soc. Math. France 112, no. 2: 177–196 (1984).

[16] Lesigne, E. *Théorèmes ergodiques ponctuels pour des mesures diagonales. Cas des systmes distaux.* Ann. Inst. H. Poincaré Probab. Statist. 23, no. 4: 593–612 (1987).

[17] Lesigne, E. *Théorèmes ergodiques pour une translation sur une nilvariete.* Erg. Th. & Dyn. Sys. 9(1): 115-126 (1989).

[18] Lesigne, E. *Equations fonctionelles, couplages de produits guaches et théorèmes ergodiques ponctuels pour mesures diagonales.* Bull. Soc. Math. France 121: 315-351 (1993).

[19] Lusin, N. *Leçons sur les Ensembles Analytiques.* (1930).

[20] Malcev, A.I. *On a class of homogeneous spaces.* Amer. Math. Soc. Translation no. 39 (1951).

[21] Meiri D. *Generalized corrolation series and nilpotent systems.* MSc thesis (1990).

[22] Montgomery D., Zippin L. *Topological Tranformation Groups.* Interscience Publishers, New York-London, (1955).

[23] Peterson, C. *Ergodic Theory* (1983).

[24] Parry, W. *Ergodic properties of affine transformations and flows on nilmanifolds.* Amer. J. Math. 91: 757-771 (1969).

[25] Parry, W. *Dynamical systems on nilmanifolds.* Bull. London Math. Soc. 2: 37-40 (1970).

[26] Parry, W. *Dynamical representations in nilmanifolds.* Compositio Math. 26: 159-174 (1973).

[27] Rudolph, D. *Eigenfunctions of* $T \times S$ *and the Conze-Lesigne algebra.* Ergodic theory and its connections with harmonic analysis (Alexandria, 1993), 369-432, London Math. Soc. Lecture Note Ser., 205, Cambridge Univ. Press, Cambridge (1995).

[28] Shah, N. *Invariant measures and orbit closures on homogeneous spaces for actions of subgroups.* Lie groups and ergodic theory (Mumbai, 1996) Tata Inst. Fund. Res. Bomb ay: 229-271 (1998).

[29] Ziegler, T. *A non-conventional ergodic theorem for a nilsystem.* availiable at http://www.math.ohio-state.edu/ tamar/.