Topological **M**ethods in **N**onlinear **A**nalysis
Journal of the Juliusz Schauder Center
Volume 30, 2007, 193–210

LIFTING ERGODICITY IN (G, σ)-EXTENSIONS

Mahesh Nerurkar

Abstract. Given a compact dynamical system (X, T, m) and a pair (G, σ) consisting of a compact group G and a continuous group automorphism σ of G, we consider the twisted skew-product transformation on $G \times X$ given by

$$T_\varphi(g, x) = (\sigma[(\varphi(x)g], Tx),$$

where $\varphi: X \to G$ is a continuous map. If (X, T, m) is ergodic and aperiodic, we develop a new technique to show that for a large class of groups G, the set of φ's for which the map T_φ is ergodic (with respect to the product measure $\nu \times m$, where ν is the normalized Haar measure on G) is residual in the space of continuous maps from X to G. The class of groups for which the result holds contains the class of all connected abelian and the class of all connected Lie groups. For the class of non-abelian fiber groups, this result is the only one of its kind.

1. Introduction

By a *dynamical system* we mean a triple (X, T, m) where X is a compact metric space, m is a Borel probability measure on X, and T is an m-preserving homeomorphism of X. Let G be a compact metric topological group and let σ be a continuous group automorphism of G. Given a continuous map $\varphi: X \to G$, define a σ-*skew-product* transformation $T_\varphi: G \times X \to G \times X$ by the rule

$$T_\varphi(g, x) = (\sigma[\varphi(x)g], Ty), \quad (g, x) \in G \times X.$$

2000 *Mathematics Subject Classification.* 37A99, 37A05.
Key words and phrases. Affine extensions, Dynamical Systems, ergodicity, cocycles..

Let ν denote the Haar measure on G, normalized so that $\nu(G) = 1$. Then $\nu \times m$ is invariant under T_φ. The dynamical system $(G \times X, T_\varphi, \nu \times m)$ is called a (G, σ)-extension or an *affine extension* of (X, T, m). The more commonly studied "group extension" is a special case when $\sigma = I$ — the identity automorphism of G.

Such skew-product extensions have been extensively studied in ergodic theory and topological dynamics. However, inspite of the vast literature on the subject, several basic questions remain unanswered, particularly when the fiber group G is *non-abelian.* One such question is: Is the set

$$C_{erg}(X, G) \equiv \{\varphi \in C(X, G) : (G \times X, T_\varphi, \nu \times m) \text{ is ergodic}\}$$

non-empty? and if so, how "big" is it? In this paper we shall show that for large class of compact groups G, this set residual in $C(X, G)$ — the metric space of continuous maps from X to G with the supremum metric.

Before starting with the proof, we briefly describe related results and techniques known so far. In the class of group extensions, there are basically two main techniques to prove such generic lifting of minimality and ergodicity. The first one is due to R. Ellis which proves generic lifting of minimality in compact, connected Lie group extensions (cf. [1]). Later H. Keynes and D. Newton (cf. [3]) extended this technique to lift minimality in affine extensions as well. The second technique is what is now called the "conjugation approximation technique". This method first appeared in the works of D. Anosov and A. Katok, then used by A. Fathi and M. Herman and the subsequently refined by S. Glasner and B. Weiss. This technique allows one to lift minimality as well as ergodicity but in a rather special class of "closures of coboundaries" (and not in the class of all cocycles). Using a version of this technique we proved a generic ergodicity lifting result in the class of "closures of σ-coboundaries" when the fiber group is the n-torus (cf. [2]). Still a proper "ergodic analog" of Ellis's technique to lift minimality was not yet known even for group extensions. For group extensions, this was developed in a joint work with H. Sussmann (cf. [7]) and in the present paper we extend that method to lift ergodicity in affine extensions. We remark that a "smooth version" of Ellis's technique and (ours as well) is yet to be developed and it seems there are some fundamental obstructions to doing this in general.

Now we shall begin by describing the conditions we need to impose on the pair (G, σ).

Definition 1.1. Consider a pair (G, σ), where G is a compact metric group with metric d and σ is a continuous group automorphism of G.

(a) The pair (G, σ) has *Property A* if given a neighbourhood W of the identity e_G, there exist a positive integer κ such that, if $W_0, \dots, W_{\kappa-1}$

are any κ right translates of W, then

$$\prod_{i=0}^{\kappa-1} \sigma^i(W_i) \stackrel{\text{def}}{=} \sigma(W_{\kappa-1}) \cdot \sigma^2(W_{\kappa-2}) \cdot \ldots \cdot \sigma^\kappa(W_0) = G.$$

(b) The pair (G,σ) has *Property B* if given a compact metric space X and a $\varphi \in C(X,G)$ and $\varepsilon > 0$, there exists a $\delta > 0$ such that if F is a finite subset of X and $\psi\colon F \to G$ is a function that satisfies $d(\varphi(x),\psi(x)) < \delta$ for all $x \in F$, then ψ can be extended to a continuous map on X such that $d(\varphi(x),\psi(x)) < \varepsilon$ for all $x \in X$.

Following [5], we shall call the pair (G,σ) *admissible* if it satisfies both Properties A and B.

We refer to [3] and [5] for the proof of the following.

Proposition 1.2. *If G is either*

(a) *a compact, metric connected abelian group or*
(b) *a compact, connected Lie group or*
(c) *a compact connected metric group with finite center.*

Then (G,σ) is admissible for any continuous automorphism σ.

Before stating the main theorem, we recall that (X,T,m) *aperiodic* if

$$m(\{x \in X : T^k x = x \text{ for some } k \in \mathbb{N}\}) = 0$$

and *ergodic* if every T-invariant Borel subset of X has m measure either zero or one.

Theorem 1.3. *Let (X,T,m) be a dynamical system and (G,σ) be a pair where G is a compact group and σ its continuous automorphism. Suppose that*

(a) *(X,T,m) is a ergodic and aperiodic, and*
(b) *(G,σ) is admissible.*

Then the set

$$C_{\text{erg}}(X,G) = \{\varphi \in C(X,G) : (G \times X, T_\varphi, \nu_G \times m) \textit{ is ergodic}\},$$

is a residual subset of $C(X,G)$.

2. Reduction of the proof

First we describe a crucial result (Proposition 2.2, which is Corollary 2.4 in [4]) that reduces the problem of lifting ergodicity from the case of a general σ to the case when σ is equicontinuous. (An automorphism σ is *equicontinuous* if the family of map $\{\sigma^i : i \in \mathbb{N}\}$ is an equicontinuous family of maps from G to G.) First we begin with some notation.

Definition 2.1. Consider a given pair (G,σ) with G compact.

(a) Let $\widehat{G}$ denote the "dual object" i.e. the set of equivalence classes of finite dimensional irreducible representations of G. A typical element of $\widehat{G}$ will be denoted by $[\gamma]$ where γ denotes a representative of the class $[\gamma]$.

(b) Let $\widehat{\sigma}\colon \widehat{G}\to\widehat{G}$ be the map induced by σ, i.e. $\widehat{\sigma}[\gamma]=[\gamma\circ\sigma]$.

(c) Given a $[\gamma]\in\widehat{G}$, let $\mathbb{Z}_{[\gamma]}$ denote the *stabilizer* subgroup of $[\gamma]$. More precisely,

$$\mathbb{Z}_{[\gamma]} \stackrel{\text{def}}{=} \{n\in\mathbb{Z} : (\widehat{\sigma})^n[\gamma]=[\gamma]\}.$$

Let

$$\widehat{G}_{\text{eq}} = \{[\gamma]\in\widehat{G} : \mathbb{Z}_{[\gamma]} \text{ has finite index in } \mathbb{Z}\},$$

and let G_{eq} be the *annihilator* of $\widehat{G}_{\text{eq}}$, i.e.

$$G_{\text{eq}} \stackrel{\text{def}}{=} \operatorname{Ann}\widehat{G}_{\text{eq}} \stackrel{\text{def}}{=} \bigcap_{[\gamma]\in\widehat{G}} \ker\gamma.$$

Then the set G_{eq} is in fact a closed normal σ invariant subgroup of G. The subgroup G_{eq} is the obstruction to σ being equicontinuous, (hence the suffix "eq").

(d) Set $G^*=G/G_{\text{eq}}$ and σ^* denote the automorphism induced by σ on G^*.

(e) Let $\pi\colon G\to G^*\equiv G/G_{\text{eq}}$ be the canonical homomorphism (which is a continuous and open map). A typical element of G^* will be denoted by $g^* \stackrel{\text{def}}{=} \pi(g)$. In the language of topological dynamics, (G^*,σ^*) is a *maximal equicontinuous factor* of the dynamical system (G,σ) under the factor map π.

(f) Let ν^* denote the normalized Haar measure on G^*. Observe that $\pi^*(\nu)=\nu^*$.

(g) Given a $\varphi\in C(X,G)$, set $\varphi^*=\pi\circ\varphi$. Thus $\varphi\to\varphi^*$ defines a continuous map from $C(X,G)\to C(X,G^*)$.

(h) For a given $\varphi\in C(X,G)$, define transformation T_{φ^*} on $G^*\times X$ by setting

$$T_{\varphi^*}(g^*,x) = (\sigma^*[\varphi^*(x)g],Tx)\,.$$

Metrics on all the underlying spaces (such as X, G and G^*) will be denoted by d.

The following theorem essentially allows one to reduce the proof to the case when σ is equicontinuous (cf. [4]).

Proposition 2.2. *With the notation as above, given $\varphi\in C(X,G)$, the dynamical system $(G\times X,T_\varphi,\nu\times m)$ is ergodic if and only if the system $(G^*\times X,T_{\varphi^*},\nu^*\times m)$ is ergodic.*

This proposition implies that

$$C_{\mathrm{erg}}(X,G)=\{\varphi\in C(X,G):(G^*\times X,T_{\varphi^*},\nu^*\times m)\text{ is ergodic}\}.$$

We will in fact prove a slightly stronger result, of which Theorem 1.3 is a direct consequence. We begin by introducing some notation. If $f\in L^2(G^*,\nu^*)$, we use $\langle f\rangle$ to denote the average of f over G^* with respect to ν^*, so $\langle f\rangle=\int_{G^*}f(g)\,d\nu^*(g)$. If f belongs to $L^2(G^*\times X,\nu^*\times m)$ and $x\in X$, we use f^x to denote the function $G^*\ni g^*\mapsto f(g^*,x)\in\mathbb{C}$. Then $f^x\in L^2(G^*,\nu^*)$ for almost all $x\in X$ and $\int_X\|f^x\|_{L^2}^2\,dm(x)\le\|f\|_{L^2}^2$. For $f\in L^2(G^*\times X,\nu^*\times m)$ set

$$\widehat{f}(g^*,x)\stackrel{\mathrm{def}}{=}\langle f^x\rangle\stackrel{\mathrm{def}}{=}\int_{G^*}f(g^*,x)\,d\nu^*(g^*).$$

Then $\widehat{f}$ is in fact the orthogonal projection of f on the space of square-integrable functions on $G^*\times X$ that are functions of x only. We let

$$\begin{aligned}\mathcal{H}&=L^2(G^*\times X,\nu^*\times m),\\ \mathcal{H}_0&=\left\{f\in\mathcal{H}:\int_{G^*\times X}f\,d(\nu^*\times m)=0\right\},\\ \mathcal{H}_{0,\mathrm{av}}&=\{f\in\mathcal{H}:\widehat{f}\equiv 0\}.\end{aligned}$$

Given $\varphi\in C(X,G)$, define a unitary operator U_{φ^*} on $\mathcal{H}$ by setting

$$U_{\varphi^*}f=f\circ T_{\varphi^*}=f\circ T_{\varphi\circ\pi}.$$

Furthermore, let

$$W_{n,\varphi^*}=\sum_{i=0}^{n-1}U_{\varphi^*}^i.$$

We will then prove the following.

Theorem 2.3. *Let (X,T,m) be a discrete dynamical system and G be a compact metric group. Suppose that*

(a) *(X,T,m) is ergodic and aperiodic, and*
(b) *(G,σ) is admissible.*

Then the set

$$C_{\mathrm{erg,av}}(X,G)=\left\{\varphi\in C(X,G):\lim_{n\to\infty}\frac{1}{n}W_{n,\varphi^*}f=0\text{ for all }f\in\mathcal{H}_{0,\mathrm{av}}\right\},$$

is a residual subset of $C(X,G)$.

Proof of Theorem 1.3. The assumptions of Theorem 1.3 allow us to apply Theorem 2.3. Let $\varphi\in C_{\mathrm{erg,av}}(X,G)$. If $f\in\mathcal{H}_0$, then $\widetilde{f}\stackrel{\mathrm{def}}{=}f-\widehat{f}\in\mathcal{H}_{0,\mathrm{av}}$.

Hence $(1/n)W_{n,\varphi^*}\widetilde{f} \to 0$ as $n \to \infty$. Let $h(x) \stackrel{\text{def}}{=} \langle f^x \rangle$, then $\int_X h(x)\,dm(x) = 0$. Since T is ergodic,

$$\frac{1}{n}W_n h \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=0}^{n-1} h \circ T^i \to 0.$$

Observe that $W_n h = W_{n,\varphi^*}\widehat{f}$. Therefore

$$\frac{1}{n}W_{n,\varphi^*}f = \frac{1}{n}(W_{n,\varphi^*}\widetilde{f} + W_{n,\varphi^*}\widehat{f}) \to 0 \quad \text{as } n \to \infty.$$

Since this is true for all $f \in \mathcal{H}_0$, it follows that $(G^* \times X, T_{\varphi^*}, \nu^* \times m)$ is ergodic. Thus Proposition 2.2 shows that $C_{\text{erg,av}}(X,G) \subseteq C_{\text{erg}}(X,G)$ and consequently $C_{\text{erg}}(X,G)$ is residual. □

3. Proof of Theorem 2.3

Given $f \in \mathcal{H}_{0,\text{av}}$, $\varepsilon > 0$ and $\overline{n} \in \mathbb{N}$, define a set $E(f,\varepsilon,\overline{n})$ as follows:

$$E(f,\varepsilon,\overline{n}) = \Big\{\varphi \in C(X,G) : \frac{1}{n}||W_{n,\varphi^*}f||_2 < \varepsilon \quad \text{for some } n \in \mathbb{N} \text{ such that } n > \overline{n}\Big\}.$$

LEMMA 3.1. *Let $\varphi \in C(X,G)$, and let $\mathcal{F}$ be a dense subset of $\mathcal{H}_{0,\text{av}}$. If $\varphi \in E(f,(1/n),\overline{n})$ for all $f \in \mathcal{F}$, $n \in \mathbb{N}$, $\overline{n} \in \mathbb{N}$, then $\varphi \in C_{\text{erg,av}}(X,G)$.*

PROOF. Fix an $f \in \mathcal{F}$. By the L^2 ergodic theorem, the sequence $(1/n)W_{n,\varphi^*}f$ converges in $\mathcal{H}$ as $n \to \infty$ to some $f^* \in \mathcal{H}$. Since $\varphi \in E(f,(1/n),\overline{n})$ for all $n, \overline{n} \in \mathbb{N}$, f^* must be the zero function. So $(1/n)W_{n,\varphi^*}f \to 0$ for every $f \in \mathcal{F}$. Since $\mathcal{F}$ is dense in $\mathcal{H}_{0,\text{av}}$, and $\|(1/n)W_{n,\varphi^*}\| \le 1$ for all n, it follows that $(1/n)W_{n,\varphi^*}f \to 0$ as $n \to \infty$ for all $f \in \mathcal{H}_{0,\text{av}}$, so $\varphi \in C_{\text{erg,av}}(X,G)$. □

In view of Lemma 3.1, Theorem 1.3 will follow from the Baire category theorem — together with the facts that

(a) $C(X,G)$ is a complete metric space, because G is a compact metric group,
(b) $C(G^* \times X) \cap \mathcal{H}_{0,\text{av}}$ is dense $\mathcal{H}_{0,\text{av}}$, and
(c) $C(G^* \times X)$ is separable,

if we prove following lemma.

LEMMA 3.2. *Let $f \in C(G^* \times X) \cap \mathcal{H}_{0,\text{av}}$. Then $E(f,(1/n),\overline{n})$ is open and dense in $C(X,G)$.*

The openness of the set $E(f,\varepsilon,\overline{n})$ is easily checked and we shall establish its density. The proof of density involves a series of constructions begining with an application of the following general result about approximating the integral of a function over a compact group by average of its translates. For the proof of next lemma we refer to Proposition 6.1 of [7].

Lemma 3.3. *Let G be a compact metric group. If $\mathcal{K} \subseteq C(X,G)$ is compact and $\eta > 0$, then there exists a positive integer P, and members $\gamma_1, \dots, \gamma_P$ of G, such that*

$$\left\| \langle h \rangle - \frac{1}{P} \sum_{r=1}^{P} \mathcal{T}_{\gamma_r} h \right\|_{\sup} \le \eta \quad \text{for all } h \in \mathcal{K},$$

where $(\mathcal{T}_g h)(x) = h(gx)$ is the isometry on $C(G)$ generated by the left translation on the group.

Density of $E(f,\varepsilon,\overline{n})$. So we turn to the proof of the density of $E(f,\varepsilon,\overline{n})$. First, a word about notation. For this purpose, we fix the following objects:

(a) an $f \in C(G^* \times X) \cap \mathcal{H}_{0,\mathrm{av}}$;
(b) a $\varphi \in C(X,G)$;
(c) a positive number ε;
(d) U-a symmetric neighbourhood of e-the identity of G;
(e) a positive integer $\overline{n}$.

In a series of steps we shall construct a function $\psi \in E(f,\varepsilon,\overline{n})$ such that $\psi(x)\varphi(x)^{-1} \in U$ for all $x \in X$. While going through these steps, for clarity, the reader may also refer to the outline of the procedure summarized at the end of the paper.

Step 1. For the given φ and ε, let δ be as in Property B and without loss of generality we shall assume that $B_\delta \overset{\mathrm{def}}{=} \{g \in G : d(g, e_G) < \delta\} \subseteq U$. Set

$$\mathcal{V}(\varphi, U) = \{\psi \in C(X,G) : \psi(x)\varphi(x)^{-1} \in U \text{ for all } x \in X\}.$$

We shall construct a map $\psi \in \mathcal{V}(\varphi,U) \cap E(f,\varepsilon,\overline{n})$ which shows that the later set is non-empty. It follows that $E(f,\varepsilon,\overline{n})$ is dense in $C(X,G)$, concluding our proof. Before starting with the construction of ψ, we introduce some more notation and make a few observations.

For $\varphi \in C(X,G)$, the iterates of T_φ are given by

$$T_\varphi^j f(g,x) = f(\varphi(x,j)\sigma^j(g), T^j x),$$

where by the abuse of notation we use the same letter φ to denote the map (i.e. the cocycle): $X \times \mathbb{Z} \to G$ generated by φ. Note that,

$$\varphi(x,j) = \sigma(\varphi(T^{j-1}x)) \cdot \sigma^2(\varphi(T^{j-2}x)) \cdot \ldots \cdot \sigma^j(\varphi(x)) \quad \text{for } j > 0, \tag{3.1}$$

and we let $\varphi(x,0) = e$. Note that φ satisfies the following "σ-cocycle identity"

$$\varphi(x, m+n) = \varphi(T^m x, n)\sigma^n[\varphi(x,m)], \quad m,n \in \mathbb{N}. \tag{3.2}$$

Note that identities (3.1) and (3.2) remain valid when φ and σ are replaced by φ^* and σ^*, respectively.

For $h \in C(G^*)$ and $g^* \in G^*$ define functions $\mathcal{T}_{g^*}(h)$ and $\mathcal{T}_{\sigma^*}(h)$ by setting

$$(\mathcal{T}_{g^*}h)y^* = h(g^*y^*),$$
$$(\mathcal{T}_{\sigma^*}h)(y^*) = h(\sigma^*(y^*)) \quad y^* \in G^*.$$

As before $\mathcal{T}_{g^*}$ is an isometry on $C(G^*)$ and since σ^* is equicontinuous on G^*, without loss of generality we can assume that $\mathcal{T}_{\sigma^*}$ is an isometry as well. Now we continue with the construction of a map $\psi \in \mathcal{V}(\varphi, U) \cap E(f, \varepsilon, \overline{n})$ in a series of steps.

Step 2. Let $\mathcal{K}$ to be the closed convex hull (in $C(G^*)$) of the set of all functions given by

$$\{\mathcal{T}^i_{\sigma^*}(\mathcal{T}_{g^*}f^x) : g^* \in G^*,\ i \in \mathbb{N} \cup \{0\},\ x \in X\}.$$

(We recall that $f^x(h^*) = f(h^*, x)$.) Since σ^* is equicontinuous on G^*, $\mathcal{K}$ is a compact, convex subset of $C(G^*)$. Moreover, $\mathcal{K}$ is invariant under $\mathcal{T}_{\sigma^*}$ and $\mathcal{T}_{g^*}$ for all $g^* \in G^*$, furthermore $\langle h \rangle = \int_{G^*} h \, d\nu^* = 0$ for every $h \in \mathcal{K}$.

Next, we fix a finite subset $\mathcal{K}_0$ of $\mathcal{K}$ such that every $h \in \mathcal{K}$ satisfies the following inequality

$$\|h - h_0\|_{\sup} < \frac{\varepsilon}{16}$$

for some $h_0 \in \mathcal{K}_0$. We let $\widehat{\kappa}$ be the cardinality of $\mathcal{K}_0$. We let

$$\beta = \frac{\varepsilon}{8\widehat{\kappa}}. \tag{3.3}$$

Next, let $\delta_1 > 0$ be such that

(3.4) if ${g_1}^*, {g_2}^* \in G^*$ and $d({g_1}^*, {g_2}^*) < \delta_1$,

$$\text{then } |H({g_1}^*) - H({g_2}^*)| < \frac{\beta}{2}, \quad \text{for all } H \in \mathcal{K}.$$

Step 3. Now applying Lemma 3.3 to the group G^*, with $\mathcal{K}$ as above and with $\eta = \beta$, we get a positive integer P and ${\gamma_1}^*, \dots, {\gamma_P}^* \in G^*$ such that

$$\left| \frac{1}{P} \sum_{r=1}^{P} \mathcal{T}_{{\gamma_r}^*} h \right|_{\sup} \leq \beta \quad \text{for all } h \in \mathcal{K}. \tag{3.5}$$

Step 4. For the given neighbourhood U of e, using Property A, we get a positive integer κ such that for any κ right translates $U_0, \dots, U_{\kappa-1}$ of U we have,

$$\sigma(U_{\kappa-1})\sigma^2(U_{\kappa-2}) \dots \sigma^\kappa(U_0) = G. \tag{3.6}$$

Step 5. Next, we want to apply Rokhlin's lemma to get a Rokhlin stack of height N. We want to have this N large and of the form $N = (\kappa + \mu)\lambda$, where κ is as in the previous step and μ and λ are chosen as follows. Recall that we need to produce a positive integer n, $(n > \overline{n})$ so that the function ψ that we are going to construct satisfies $||W_{n,\psi^*} f||_2/n < \varepsilon$, (where $\psi^* = \pi \circ \psi$). This n

will be of the form $n = (\kappa+\mu)\rho$, where ρ is another large positive integer (but much smaller than λ) to be chosen. Now we describe how (and in what order) the choice of these integers is made.

First select ρ so that

$$\frac{||f||_{\sup}}{\rho} \le \frac{\varepsilon}{16} \quad \text{and} \quad \frac{P\widehat{\kappa}||f||_{\sup}}{\rho} \le \frac{\varepsilon}{8}. \tag{3.7}$$

Next, select λ so that

$$\frac{\rho}{\lambda}||f||^2_{\sup} \le \frac{\varepsilon^2}{64}, \tag{3.8}$$

Finally choose μ so that

$$\frac{\kappa||f||_{\sup}}{\mu} \le \frac{\varepsilon}{8}, \quad n \stackrel{\text{def}}{=} (\kappa+\mu)\rho > \overline{n}, \tag{3.9}$$

and

$$d({\sigma^*}^{i(\kappa+\mu)}(g^*), g^*) < \delta_1, \quad \text{for all } g^* \in G^* \text{ and all } i \in \{1, \ldots, \lambda\}, \tag{3.10}$$

where δ_1 is as in Step 2 (see equation (3.4)). By the equicontinuity of σ^* such choice of μ is possible. Let

$$N = (\kappa+\mu)\lambda.$$

Step 6. For this N, Using Rokhlin's lemma, pick a Borel subset $\widehat{E}$ of X with the property that the sets $\widehat{E}, T\widehat{E}, \ldots, T^{N-1}\widehat{E}$ are pairwise disjoint, and

$$m(\widehat{E} \cup T\widehat{E} \cup \ldots \cup T^{N-1}\widehat{E}) \ge 1 - c_1,$$

where c_1 is a positive constant such that

$$c_1||f||^2_{\sup} \le \frac{\varepsilon^2}{64}. \tag{3.11}$$

Next, we need to partition each $\widehat{E}$ into "small pieces" (see Step 10). The next three steps describe how fine this partition must be.

Step 7. Using uniformly continuity of f on $G^* \times X$, choose

(a) a positive number c_2 such that $|f(g^*, y) - f(g^*, z)| \le \varepsilon/8$ whenever $g^* \in G^*$, $y, z \in X$, and $d(y,z) \le c_2$,

(b) a neighbourhood V of the identity e^* of G^* such that $|f(g^*, x) - f(h^*, x)| \le \varepsilon/8$ whenever $g^*, h^* \in G^*$, $x \in X$, and $g^*(h^*)^{-1} \in V$.

Step 8. Let Φ^N be the subset of $C(X,G)$ consisting of all functions $\eta\colon X \to G$ of the form

$$\eta(x) = g_r \cdot \sigma\big(\varphi(T^r x)\big) \cdot g_{r-1} \cdot \sigma^2\big(\varphi(T^{r-1}x)\big) \cdot \;\cdots\; \cdot \sigma^{r+1}\big(\varphi(x)\big) \cdot g_0, \tag{3.12}$$

as r ranges over $\{0, \ldots, N-1\}$ and $(g_0, \ldots, g_r)$ ranges over the set of all $r+1$-tuples of members of G. Then Φ^N is compact, because the right-hand side of (3.12), regarded as a function of x, depends continuously on the $r+1$-tuple

$(g_0, \dots, g_r)$, which takes values in the compact space G^{r+1}. It follows that the family Φ^N is equicontinuous, so corresponding to the neighborhood V of e^* chosen in Step 7, there exists a constant $\omega(N, V)$ with the property that $\pi(\eta(y)\eta(z)^{-1}) \in V$ whenever $\eta \in \Phi^N$, $y, z \in X$, and $d(y, z) \le \omega(N, V)$.

Step 9. Write $J = \{0, \dots, N-1\}$. Using the continuity of maps T^j, $(j \in J)$, pick a positive c_3 such that

(a) $d(T^j y, T^j z) \le c_2$ whenever $d(y, z) \le c_3$ and $j \in J$,
(b) $c_3 \le \omega(N, V)$.

Step 10. We partition X into finitely many Borel-measurable sets $X_1, \dots, X_s$ of diameter less than c_3, and let $\widehat{E}_i = \widehat{E} \cap X_i$ for $i = 1, \dots, s$. Write $I = \{1, \dots, s\}$. Using the fact that m is regular, pick compact sets E_i such that $E_i \subseteq \widehat{E}_i$ and $m(\widehat{E}_i \setminus E_i) < (c_1/sN)$. Write $E_{ij} = T^j E_i$ for $i \in I$, $j \in J$ (so $E_{i0} = E_i$). Then the sets E_{ij} are pairwise disjoint, and

$$m\Big(\bigcup_{i\in I}\bigcup_{j\in J} E_{ij}\Big) \ge 1 - 2c_1.$$

We discard any E_i's that have measure zero, so $m(E_{ij}) > 0$ for all $i \in I$, $j \in J$.

Step 11. We pick points $x_i \in E_i$ for each $i \in I$, and then let $x_{i,j} = T^j x_i$, so $x_{i,j} \in E_{ij}$.

Step 12. Now we are going to define the map ψ belonging to $\mathcal{V}(\varphi, U)$. For this purpose, we begin by dividing the integer interval $J = [0, 2, \dots, N-1]$ into λ blocks $J_1, \dots, J_\lambda$ of length $\kappa + \mu$, given by

$$J_\ell = \{(\ell-1)(\kappa+\mu), (\ell-1)(\kappa+\mu)+1, \dots, \ell(\kappa+\mu)-1\} \quad \text{for } \ell \in L, \tag{3.13}$$

where $L = \{1, \dots, \lambda\}$. We then split each block J_ℓ into two blocks J_ℓ^1, J_ℓ^2, of lengths κ and μ, by letting

$$\begin{aligned} J_\ell^1 &= \{(\ell-1)(\kappa+\mu), (\ell-1)(\kappa+\mu)+1, \dots, \ell\kappa + (\ell-1)\mu - 1\}, \\ J_\ell^2 &= \{\ell\kappa + (\ell-1)\mu, \ell\kappa + (\ell-1)\mu + 1, \dots, \ell(\kappa+\mu)-1\}. \end{aligned}$$

The map ψ will satisfy

$$\psi(x_{i,j}) = \varphi(x_{i,j}) \qquad \text{if } j \in J_\ell^2, \tag{3.14}$$

$$\psi(x_{i,j}) = g_{ij}\varphi(x_{i,j}) \quad \text{if } j \in J_\ell^1, \tag{3.15}$$

where the g_{ij} are members of U whose choice will be described later. By Property B there exists a continuous map $\theta\colon X \to U$ such that $\theta(x_{i,j}) = g_{ij}$ for $j \in J_\ell^1$ and $\theta(x_{i,j}) = e$ for $j \in J_\ell^2$. If we then define $\psi(x) = \theta(x)\varphi(x)$, then ψ satisfies (3.14) and (3.15) and furthermore $\psi \in \mathcal{V}(\varphi, U)$.

Estimating $(1/n)\|W_{n,\psi^*}f\|_{L_2}$. Now we are now going to estimate the norm $\|W_{n,\psi^*}f\|_{L_2}$where n is chosen as in (3.9). We let

$$A=\bigcup_{j=0}^{N-n-1}\bigcup_{i=1}^{s}E_{ij}.$$

Since $m\big(\bigcup_{j=N-n}^{N-1}\bigcup_{i=1}^{s}E_{ij}\big)\le(n/N)=(\rho/\lambda)$, (because $m\big(\bigcup_{i=1}^{s}E_{ij}\big)\le(1/N)$ for each j), the set A satisfies

$$m(A)\ge 1-2c_1-\frac{\rho}{\lambda}.$$

Outside $G^*\times A$, the function $(1/n)W_{n,\psi^*}f$ is bounded pointwise by $\|f\|_{\sup}$, so using (3.11) and the estimate on $m(A)$ obtained above, we have

$$\int_{G^*\times(X\setminus A)}\left|\frac{1}{n}W_{n,\psi^*}f(g^*)\right|^2d\nu^*(g^*)\le\|f\|_{\sup}^2\left(2c_1+\frac{\rho}{\lambda}\right)\le\frac{3\varepsilon^2}{64}. \tag{3.16}$$

If $(g^*,x)\in G^*\times A$, then there exists unique $i\in I$, $j\in J$, such that $j\le N-n-1$ and $x\in E_{ij}$. Write

$$W_{n,\psi^*;i,j}f(g^*)\stackrel{\text{def}}{=}W_{n,\psi^*}f(g^*,x_{i,j})=\sum_{k=0}^{n-1}f(\psi^*(x_{i,j},k){\sigma^*}^k(g^*),T^kx_{i,j}). \tag{3.17}$$

Since both x and $x_{i,j}$ belong to E_{ij}, the distance $d(x,x_{i,j})$ is not larger than c_3. Then $d(T^kx,T^kx_{i,j})\le c_2$ for all $0\le k\le n-1$. Therefore (by (b) of Steps 7 and 9),

$$|f(\psi^*(x,k)h^*,T^kx)-f(\psi^*(x,k)h^*,T^kx_{i,j})|\le\frac{\varepsilon}{8}, \tag{3.18}$$

for all $0\le k\le n-1$ and $h^*\in G^*$. Now notice that for each $0\le k\le n$, the function $x\to\psi(x,k)$ restricted to set A "belongs to Φ^N" (i.e. equals the restriction to A of some member of Φ^N). Since $c_3\le\omega(N,V)$, we conclude that

$$\pi(\psi(x,k)h(\psi(x_{i,j},k)h)^{-1})=\psi^*(x,k)\psi^*(x_{i,j},k)\in V,$$

for all $x\in A$, $h\in G$ and $0\le k\le n-1$. Therefore

$$|f(\psi^*(x,k)h^*,T^kx_{i,j})-f(\psi^*(x_{i,j},k)h^*,T^kx_{i,j})|\le\frac{\varepsilon}{8}. \tag{3.19}$$

By taking $h^*={\sigma^*}^k(g^*)$, and using (3.18) and (3.19) we have:

$$|W_{n,\psi^*;i,j}f(g^*)-W_{n,\psi^*}f(g^*,x)|\le\frac{\varepsilon}{4}\quad\text{whenever } g^*\in G^*,x\in E_{ij}. \tag{3.20}$$

Now fix an $i\in I$ and $j\in J$ such that $j\le N-n-1$. Then for each $k\in\{0,\dots,n-1\}$, $T^kx_{i,j}=x_{i,j+k}\in E_{i,j+k}$. As k varies from 0 to $n-1$, the index $j+k$ takes all values in the integer interval $\Gamma_n(j)=\{j,\dots,j+n-1\}$, which is a subset of $J\equiv\{0,\dots,N-1\}$ because $j\le N-n-1$. We can then write $\Gamma_n(j)=\Gamma_n^c(j)\cup\Gamma_n^b(j)$, where the "central" part $\Gamma_n^c(j)$ is a disjoint union of intervals J_ℓ (see (3.13)) with ℓ such that $J_\ell \subseteq \Gamma_n(j)$, and the "bad" (or the "end") part $\Gamma_n^b(j)$ which has at most $2(\kappa+\mu)$ members. Thus,

$$W_{n,\psi^*;i,j}f(g^*) = W^c_{n,\psi^*;i,j}f(g^*) + W^b_{n,\psi^*;i,j}f(g^*),$$

where $W^c_{n,\psi^*;i,j}f(g^*)$ and $W^b_{n,\psi^*;i,j}f(g^*)$ are the contributions to the right-hand side of (3.17) by the terms for which $k+j \in \Gamma^c_n(j)$ and $k+j \in \Gamma^b_n(j)$, respectively. Thus,

$$W^c_{n,\psi^*;i,j}f(g^*) = \sum_{k+j\in\Gamma^c_n(j)} f(\psi^*(x_{i,j},k){\sigma^*}^k(g^*), x_{i,j+k}).$$
$$W^b_{n,\psi^*;i,j}f(g^*) = \sum_{k+j\in\Gamma^b_n(j)} f(\psi^*(x_{i,j},k){\sigma^*}^k(g^*), x_{i,j+k}).$$

Since $\Gamma^b_n(j)$ has at most $2(\kappa+\mu)$ indices, we get the bound

$$\left|\frac{1}{n}W^b_{n,\psi^*;i,j}f(g^*)\right| \le \frac{2(\kappa+\mu)\|f\|_{\sup}}{n} = \frac{2\|f\|_{\sup}}{\rho}. \tag{3.21}$$

Now let us analyze the term $W^c_{n,\psi^*;i,j}f(g^*)$. Let $\Delta_n(j) = \{\ell \in J_\ell \subseteq \Gamma_n(j)\}$, then

$$W^c_{n,\psi^*;i,j}f(g^*) = \sum_{\ell\in\Delta_n(j)} \sum_{k+j\in J_\ell} f(\psi^*(x_{i,j},k){\sigma^*}^k(g^*), x_{i,j+k}).$$

This expression allows us to write $W^c_{n,\psi^*;i,j}f(g^*)$ as a sum of contributions from the individual terms $W^c_{n,\psi^*;i,j;\ell}f(g^*)$, $(\ell \in \Delta_n(j))$, defined by

$$\begin{aligned} W^c_{n,\psi^*;i,j;\ell}f(g^*) &= \sum_{k+j\in J_\ell} f(\psi^*(x_{i,j},k){\sigma^*}^k(g^*), x_{i,j+k}) \\ &= \sum_{u\in J_\ell} f(\psi^*(x_{i,j},u-j){\sigma^*}^{(u-j)}(g^*), x_{i,u}). \end{aligned}$$

Let $J_\ell = \{a_\ell, a_\ell+1, \dots, a_\ell+\kappa+\mu-1\}$. Rewriting the summation with $u = a_\ell + v$ and using the cocycle identity (3.2) we get

$$\begin{aligned} W^c_{n,\psi^*;i,j;\ell}f(g^*) &= \sum_{v=0}^{\kappa+\mu-1} f(\psi^*(x_{i,j}, v+a_\ell-j){\sigma^*}^{(v+a_\ell-j)}(g^*), x_{i,a_\ell+v}) \\ &= \sum_{v=0}^{\kappa+\mu-1} f(\psi^*(x_{i,a_\ell},v){\sigma^*}^v[\psi^*(x_{i,j},a_\ell-j)]{\sigma^*}^{(v+a_\ell-j)}(g^*), x_{i,a_\ell+v}) \\ &= \sum_{v=0}^{\kappa+\mu-1} f(\psi^*(x_{i,a_\ell},v){\sigma^*}^v[\psi^*(x_{i,j},a_\ell-j){\sigma^*}^{(a_\ell-j)}(g^*)], x_{i,a_\ell+v}). \end{aligned}$$

We now split the above sum once more, by separating out the first κ terms from the remaining μ ones. We get

$$W^c_{n,\psi^*;i,j;\ell}f(g^*) = W^{c,-}_{n,\psi^*;i,j;\ell}f(g^*) + W^{c,+}_{n,\psi^*;i,j;\ell}f(g^*),$$

where

$$W^{c,-}_{n,\psi^*;i,j;\ell} f(g^*) = \sum_{v=0}^{\kappa-1} f(\psi^*(x_{i,a_\ell}, v)\sigma^{*v}[\psi^*(x_{i,j}, a_\ell - j)\sigma^{*(a_\ell - j)}(g^*)], x_{i,a_\ell+v}),$$

$$W^{c,+}_{n,\psi^*;i,j;\ell} f(g^*) = \sum_{w=0}^{\mu-1} f(\psi^*(x_{i,a_\ell}, \kappa+w)\sigma^{*(\kappa+w)}[\psi^*(x_{i,j}, a_\ell - j)\sigma^{*(a_\ell - j)}(g^*)], x_{i,a_\ell+\kappa+w}).$$

Clearly, each function $W^{c,-}_{n,\psi^*;i,j;\ell} f$ is pointwise bounded by $\kappa\|f\|_{\sup}$. Moreover, it is clear that the number of members of $\Delta_n(j)$ is at most ρ. Therefore

$$\frac{1}{n}\Big|\sum_{\ell \in \Delta_n(j)} W^{c,-}_{n,\psi^*;i,j;\ell} f(g^*)\Big| \le \frac{\rho\kappa\|f\|_{\sup}}{n} = \frac{\kappa\|f\|_{\sup}}{\kappa+\mu}. \tag{3.22}$$

We now turn to the crucial estimate, namely, the bound for the functions $W^{c,+}_{n,\psi^*;i,j;\ell} f$. Recall that $x_{i,j} = T^j x_i$ then the cocycle identity (3.2) yields

$$\psi^*(x_{i,j}, a_\ell - j) = \psi^*(x_i, a_\ell)\sigma^{*(a_\ell - j)}[\psi^*(x_i, j)^{-1}].$$

Now, consider the expression for $W^{c,+}_{n,\psi^*;i,j;\ell} f$, using the cocycle identity we obtain

$$\begin{aligned}
&\psi^*(x_{i,a_\ell}, \kappa+w)\sigma^{*(\kappa+w)}[\psi^*(x_{i,j}, a_\ell - j)\sigma^{*(a_\ell - j)}(g^*)] \\
&= \psi^*(x_{i,a_\ell+\kappa}, w)\sigma^{*w}[\psi^*(x_{i,a_\ell}, \kappa)]\sigma^{*(\kappa+w)}[\psi^*(x_{i,j}, a_\ell - j)\sigma^{*(a_\ell - j)}(g^*)] \\
&= \psi^*(x_{i,a_\ell+\kappa}, w)\sigma^{*w}(\psi^*(x_{i,a_\ell}, \kappa)\sigma^{*\kappa}[\psi^*(x_{i,j}, a_\ell - j)]\sigma^{*(\kappa+a_\ell - j)}(g^*)) \\
&= \psi^*(x_{i,a_\ell+\kappa}, w)\sigma^{*w}(\psi^*(x_{i,a_\ell}, \kappa)\sigma^{*\kappa} \\
&\quad \cdot [\psi^*(x_i, a_\ell)\sigma^{*(a_\ell - j)}(\psi^*(x_i, j))^{-1}]\sigma^{*(\kappa+a_\ell - j)}(g^*)) \\
&= \psi^*(x_{i,a_\ell+\kappa}, w)\sigma^{*w}(\xi_{\ell,i}\sigma^{*(\kappa+a_\ell - j)}(\psi^*(x_i, j))^{-1}g^*)),
\end{aligned}$$

where

$$\xi_{\ell,i} = \psi^*(x_{i,a_\ell}, \kappa)\sigma^{*\kappa}[\psi^*(x_i, a_\ell)].$$

It is at this point that we will describe the choice of g_{ij}'s. For this purpose, we will want to assign to each index $\ell \in L$ and each $i \in I$ an integer $r(\ell, i) \in \{1, \dots, P\}$. We let $F_{i,\ell}$ be the function on G^* defined by

$$F_{i,\ell}(g^*) = \frac{1}{\mu}\sum_{w=0}^{\mu-1} f(\psi^*(x_{i,a_\ell+\kappa}, w)\sigma^{*w}(g^*), x_{i,a_\ell+\kappa+w}).$$

Our first observation is that the functions $F_{i,\ell}$ do not depend on ψ at all, i.e. on the choice of g_{ij}'s, in fact

$$F_{i,\ell}(g^*) = \frac{1}{\mu}\sum_{w=0}^{\mu-1} f(\varphi^*(x_{i,a_\ell+\kappa}, w)\sigma^{*w}(g^*), x_{i,a_\ell+\kappa+w}).$$

Next, observe that

$$W^{c,+}_{n,\psi^*;i,j;\ell} f(g^*) = \mu F_{i,\ell}(\xi_{\ell,i}\sigma^{*(\kappa+a_\ell-j)}[\psi^*(x_i,j)^{-1}g^*)]).$$

Clearly, each $F_{i,\ell}$ is a convex combination of elements of the set

$$\{(\mathcal{T}^i_{\sigma^*}\circ\mathcal{T}_{g^*})f^x : g^*\in G^*,\ i\in\mathbb{N}\cup\{0\},\ x\in X\}.$$

Hence $F_{i,\ell}$ belong to $\mathcal{K}$. So for each i, ℓ we can choose $H_{i,\ell}\in\mathcal{K}_0$ such that $\|F_{i,\ell}-H_{i,\ell}\|_{\sup}\le(\varepsilon/8)$. We fix one such choice from now on.

For each $i\in I$, $H\in\mathcal{K}_0$, let $A(i,H)$ be the set of those indices $\ell\in L$ such that $H_{i,\ell}=H$. Enumerate the members of $A(i,H)$, from left to right, so

$$A(i,H)=\{\alpha_{i,H,1},\dots,\alpha_{i,H,\Lambda(i,H)}\},$$

with $\alpha_{i,H,1}<\alpha_{i,H,2}<\dots<\alpha_{i,H,\Lambda(i,H)}$. Then for each $i\in I$, $\ell\in L$, there exists a unique $k=k(i,\ell)\in\{1,\dots,\Lambda(i,H_{i,\ell})\}$ such that $\ell=\alpha_{i,H_{i,\ell},k}$. We then define $r(i,\ell)$ to be the remainder of $k(i,\ell)$ modulo P, so $k(i,\ell)=\widehat{k}(i,\ell)P+r(i,\ell)$, where $\widehat{k}(i,\ell)$ and $r(i,\ell)$ are nonnegative integers and $0\le r(i,\ell)<P$.

We now define the g_{ij} for $i\in I$, $j\in J$. As explained above, we let $g_{ij}=e$ if $j\in\bigcup_\ell J^2_\ell$. We then define the family $\{g_{ij}\}_{i\in I,j\in J^1_\ell}$ for each $j\in J^1_\ell$, inductively with respect to ℓ.

Fix an ℓ, and assume that the g_{ij} have already been chosen for $j\in J^1_{\ell'}$ for all ℓ' such that $\ell'<\ell$, but not yet for $j\in J^1_\ell$. Then the g_{ij} are in fact determined for all j such that $j<a_\ell$. For any $i\in I$, we have

$$\begin{aligned}\psi^*(x_{i,a_\ell},\kappa)\cdot\sigma^{*\kappa}[\psi^*(x_i,a_\ell)]\\ &=\sigma^*[\psi^*(x_{i,a_\ell+\kappa-1})]\sigma^{*2}[\psi^*(x_{i,a_\ell+\kappa-2})]\cdots\sigma^{*\kappa}[\psi^*(x_{i,a_\ell})]\sigma^{*\kappa}[\psi^*(x_i,a_\ell)]\\ &=\sigma^*[g^*_{i,a_\ell+\kappa-1}\cdot z^*_{\kappa-1}]\cdot\sigma^{*2}[g^*_{i,a_\ell+\kappa-2}\cdot z^*_{\kappa-2}]\cdot\ldots\cdot\sigma^{*k}[g^*_{i,a_\ell}\cdot z_0{}^*],\end{aligned}$$

where $z_k=\varphi(x_{i,a_\ell+k})$ for $0<k<\kappa$, $z_0=\varphi(x_{i,a_\ell})\cdot\psi(x_i,a_\ell)$, (and recall that $z^*=\pi(z)$, $\psi^*=\psi\circ\pi$). Note that z_0 is determined, because the g_{ij} have already been selected for $j<a_\ell$. Now (3.6) guarantees that $\sigma[Uz_{\kappa-1}]\cdot\sigma^2[Uz_{\kappa-2}]\cdot\ldots\cdot\sigma^\kappa[Uz_0]=G$. Since the map $\pi\colon G\to G^*\equiv G/G_{\mathrm{eq}}$ is onto, we can choose $g_{i0},\dots,g_{i,\kappa-1}$ such that

$$\psi^*(x_{i,a_\ell},\kappa)\cdot\sigma^{*\kappa}[\psi^*(x_i,a_\ell)]=\gamma^*_{r(\ell,i)}.$$

Then, for any i, j,

$$\begin{aligned}\sum_{\ell\in\Delta_n(j)}W^{c,+}_{n,\psi^*;i,j;\ell}f(g^*)&=\mu\sum_{\ell\in\Delta_n(j)}F_{i,\ell}(\gamma^*_{r(\ell,i)}\sigma^{*(\kappa+a_\ell-j)}[\psi^*(x_i,j)^{-1}g^*])\\ &=\mu\sum_{H\in\mathcal{K}_0}\ \sum_{\ell\in A(i,H)\cap\Delta_n(j)}\mathcal{T}_{\gamma^*_{r(\ell,i)}}F_{i,\ell}(\sigma^{*(\kappa+a_\ell-j)}[\psi^*(x_i,j)^{-1}g^*]).\end{aligned}$$

Now let,

$$W^{\sigma}_{n,\psi^*;i,j}(g^*) \stackrel{\text{def}}{=} \mu \sum_{H\in\mathcal{K}_0} \sum_{\ell\in A(i,H)\cap\Delta_n(j)} \mathcal{T}_{\gamma^*_{r(\ell,i)}} H(\sigma^{*a_\ell}(g^*)).$$

Now fix a $g^* \in G^*$ and set $t^*_{ij} = \sigma^{*(\kappa-j)}[\psi^*(x_i, j)^{-1} g^*]$. Since $\|H_{i,\ell} - F_{i,\ell}\|_{\sup} \leq \varepsilon/16$ whenever $\ell \in A(i, H)$, we have the bound

$$\left| W^{\sigma}_{n,\psi^*;i,j}(t^*_{ij}) - \sum_{\ell\in\Delta_n(j)} W^{c,+}_{n,\psi^*;i,j;\ell} f(g^*) \right| \leq \frac{\mu\rho\varepsilon}{16}\,,$$

using the fact that $\Delta_n(j)$ has at most ρ members. Observing that $(\mu\rho/n) \leq 1$, we get

$$\frac{1}{n}\left| W^{\sigma}_{n,\psi^*;i,j}(t^*_{ij}) - \sum_{\ell\in\Delta_n(j)} W^{c,+}_{n,\psi^*;i,j;\ell} f(g^*) \right| \leq \frac{\varepsilon}{16}, \tag{3.23}$$

because $n = (\kappa + \mu)\rho$.

We now turn to the final task of estimating $W^{\sigma}_{n,\psi^*;i,j}$. First, notice that when $\sigma = I$-the identity automorphism, we have

$$W^{I}_{n,\psi^*;i,j}(g^*) = \mu \sum_{H\in\mathcal{K}_0} \sum_{\ell\in A(i,H)\cap\Delta_n(j)} \mathcal{T}_{\gamma^*_{r(\ell,i)}} H(g^*).$$

Let us also write

(3.24)

$$W^{\sigma}_{n,\psi^*;i,j;H}(g^*) \stackrel{\text{def}}{=} \sum_{\ell\in A(i,H)\cap\Delta_n(j)} \mathcal{T}_{\gamma^*_{r(\ell,i)}} H(\sigma^{*a_\ell} g^*) \quad \text{and}$$

(3.25)

$$W^{I}_{n,\psi^*;i,j;H}(g^*) \stackrel{\text{def}}{=} \sum_{\ell\in A(i,H)\cap\Delta_n(j)} \mathcal{T}_{\gamma^*_{r(\ell,i)}} H(g^*).$$

Then we have

$$W^{\sigma}_{n,\psi^*;i,j}(g^*) = \mu \sum_{H\in\mathcal{K}_0} W^{\sigma}_{n,\psi^*;i,j;H}(g^*), \tag{3.26}$$

$$W^{I}_{n,\psi^*;i,j}(g^*) = \mu \sum_{H\in\mathcal{K}_0} W_{n,\psi^*;i,j;H}(g^*). \tag{3.27}$$

Since $\Delta_n(j)$ has at most ρ members and $a_\ell = \ell(\kappa + \mu)$, our choice of $\kappa + \mu$ in Step 5 (see (3.10)) gives

$$|W^{\sigma}_{n,\psi^*;i,j;H}(g^*) - W^{I}_{n,\psi^*;i,j;H}(g^*)| < \frac{\rho\beta}{2} \quad \text{for all } g^* \in G^* \text{ and } H \in \mathcal{K}.$$

Since the cardinality of $\mathcal{K}_0$ is $\widehat{\kappa}$, using (3.3) we have

$$|W^{\sigma}_{n,\psi^*;i,j}(g^*) - W^{I}_{n,\psi^*;i,j}(g^*)| < \frac{\mu\kappa\rho\beta}{2} = \frac{\varepsilon\mu\rho}{8} \quad \text{for all } g^* \in G^*.$$

Combining this with (3.23) and recalling that $(\mu\rho/n) \le 1$ yields,

$$\frac{1}{n}\Big|\sum_{\ell\in\Delta_n(j)} W^{c,+}_{n,\psi^*;i,j;\ell} f(g^*) - W^I_{n,\psi^*;i,j}(t^*_{ij})\Big| \le \frac{\varepsilon}{8}, \tag{3.28}$$

(recall that $t^*_{ij} = \sigma^{*(\kappa-j)}[\psi^*(x_i,j)^{-1}g^*]$). Now we estimate $W^I_{n,\psi^*;i,j}(g^*)$. Recall that

$$W^I_{n,\psi^*;i,j}(g^*) = \mu \sum_{H\in\mathcal{K}_0}\ \sum_{\ell\in A(i,H)\cap\Delta_n(j)} (\mathcal{T}_{\gamma^*_{r(\ell,i)}} H)(g^*).$$

Now suppose that $A(i,H)\cap\Delta_n(j)$ has $q_{i,j,H}P + p_{i,j,H}$ members, where $q_{i,j,H}$ and $p_{i,j,H}$ are integers such that $q_{i,j,H}\ge 0$ and $0\le p_{i,j,H} < P$. Then the sum of the first $q_{i,j,H}P$ terms of the above expression is equal to $q_{i,j,H}\sum_{r=1}^P \mathcal{T}_{\gamma_r}H(g^*)$, whose absolute value is bounded by $P\beta q_{i,j,H}$, because $\int_{G^*} H d\nu^* = 0$.The sum of the remaining $p_{i,j,H}$ terms is bounded by $p_{i,j,H}\|f\|_{\sup}$. Thus

$$|W^I_{n,\psi^*;i,j;H}(g^*)| \le P\beta q_{i,j,H} + p_{i,j,H}\|f\|_{\sup}.$$

Each number $q_{i,j,H}$ is bounded by ρ/P, and $p_{i,j,H}\le P$. So

$$|W^I_{n,\psi^*;i,j;H}(g^*)| \le \rho\beta + P\|f\|_{\sup}.$$

Therefore

$$\|W^I_{n,\psi^*;i,j}\|_{\sup} \le \mu\widehat{\kappa}\rho\beta + \mu\widehat{\kappa}P\|f\|_{\sup}. \tag{3.29}$$

If we combine (3.29) and (3.28) we find

$$\frac{1}{n}\Big\|\sum_{\ell\in\Delta_n(j)} W^{c,+}_{n,\psi^*;i,j;\ell} f(g^*)\Big\|_{\sup} \le \frac{\varepsilon}{8} + \frac{\mu\widehat{\kappa}\rho\beta}{n} + \frac{\mu\widehat{\kappa}P}{n}\|f\|_{\sup},$$

We now use (3.22) and get

$$\frac{1}{n}|W^c_{n,\psi^*;i,j}|_{\sup} = \frac{1}{n}\Big|\sum_{\ell\in\Delta_n(j)} W^c_{n,\psi^*;i,j;\ell}\Big|_{\sup} \le \frac{\kappa\|f\|_{\sup}}{\kappa+\mu} + \frac{\varepsilon}{8} + \frac{\mu\widehat{\kappa}\rho\beta}{n} + \frac{\mu\widehat{\kappa}P}{n}\|f\|_{\sup}.$$

Then (3.21) implies

$$\frac{1}{n}|W_{n,\psi^*;i,j}(g^*)| \le \frac{2\|f\|_{\sup}}{\rho} + \frac{\kappa\|f\|_{\sup}}{\kappa+\mu} + \frac{\varepsilon}{8} + \frac{\mu\widehat{\kappa}\rho\beta}{n} + \frac{\mu\widehat{\kappa}P}{n}\|f\|_{\sup}.$$

Finally, we use (3.20) and get the pointwise estimate

$$\frac{1}{n}|W_{n,\psi^*}f(g^*,x)| \le \frac{\varepsilon}{4} + \frac{2\|f\|_{\sup}}{\rho} + \frac{\kappa\|f\|_{\sup}}{\kappa+\mu} + \frac{\varepsilon}{8} + \frac{\mu\widehat{\kappa}\rho\beta}{n} + \frac{\mu\widehat{\kappa}P}{n}\|f\|_{\sup},$$

valid whenever $g^*\in G^*$ and $x\in A$. Note that each term on the right hand side of above inequality except the first one is less than or equal to $\varepsilon/8$ (this follows from equations (3.3) (3.7), (3.8) and (3.10)). Futhermore $n > \overline{n}$. Thus we have shown that depending on the initial given data (i.e. f, ε, φ, U and $\overline{n}$ we have found n such that

$$|W_{n,\psi^*}f(g^*,x)| \leq \frac{7\varepsilon}{8} \quad \text{whenever } g^* \in G^* \text{ and } x \in A.$$

This, together with (3.16), implies that

$$\int_{G^*\times X} |W_{n,\psi^*}f|^2 d\nu^* \times m \leq \frac{49\varepsilon^2}{64} + \left(2c_1 + \frac{3\varepsilon^2}{64}\right) < \varepsilon^2,$$

and we conclude that $\|W_{n,\psi^*}f\|_{L^2} < \varepsilon$. Since $n > \overline{n}$, we have proved that $\psi \in E(f,\varepsilon,\overline{n})$. We have already checked that $\varphi \in E(f,\varepsilon,\overline{n}) \cap \mathcal{V}(\varphi, U)$, concluding the proof of Lemma 3.2 and subsequently the proof of Theorem 2.3 as well. $\square$

Finally, for the readers convinience we briefly outline the steps taken in getting a pointwise estimate for $W_{n,\psi^*}f$. First, we remark that the levels of our Rokhlin tower (of height N and with base $\bigcup_{i\in I} E_i$) are grouped into λ blocks J_ℓ, $(1 \leq \ell \leq \Lambda)$. Levels in each λ block J_ℓ are further grouped into first κ steps and later μ steps, (where $N = (\kappa + \mu)\lambda$). After fixing a $g^* \in G^*$ and $x \in E_{ij}$ we begin by:

(a) approximating (pointwise) $W_{n,\psi^*}f$ by the quantity $W_{n,\sigma^*;i,j}f$ (see (3.17) and (3.20)).

(b) Contribution to each $W_{n,\psi^*;i,j}$ comes from the "central terms" $W^c_{n,\psi^*;i,j}f$ and the "end terms" or the "bad terms" $W^b_{n,\psi^*;i,j}f$. Since n is much larger than $\kappa + \mu$, the term $|W^b_{n,\psi^*;i,j}f|/n$ is small (see (3.21)).

(c) Next, we write $W^c_{n,\psi^*;i,j}$ as a sum of contributions from the individual terms $W^c_{n,\psi^*;i,j;\ell}f$, $(\ell \in \Delta_n(j))$ and each sum $W^c_{n,\psi^*;i,j;\ell}f(g^*)$ is then further split as a sum of terms $W^{c,-}_{n,\psi^*;i,j;\ell}f(g^*)$ and $W^{c,+}_{n,\psi^*;i,j;\ell}f(g^*)$ corresponding to the first κ terms and the remaining μ ones, respectively. Since κ is small compared to n, $|W^{c,-}_{n,\psi^*;i,j;\ell}f(g^*)|/n$ is small, (see (3.22)).

(d) For each i,ℓ we think of $W^{c,+}_{n,\psi^*;i,j;\ell}f$ as a function on G^*, namely $\mu F_{i\ell}$. For each i,ℓ, we approximate each $F_{i\ell}$ by $H_{i\ell} \in \mathcal{K}_0$. Then the sum of $W^{c,+}_{n,\psi^*;i,j;\ell}f(g^*)$, as ℓ varies over $\Delta_n(j)$, can be approximated by the quantity $W^\sigma_{n,\psi^*;i,j}$ (see (3.23)). Next, $W^\sigma_{n,\psi^*;i,j}$ itself is a sum of terms $W^\sigma_{n,\psi^*;i,j;H}$ as H varies over $\mathcal{K}_0$ (see (3.26)–(3.27)). Now the term $W^\sigma_{n,\psi^*;i,j;H}$ involves sum of the translates $\mathcal{T}_{g^*}H$ of $H \equiv H_{i,\ell}$, for each H in $\mathcal{K}_0$, (see (3.24)). (Here we remark that the functions $F_{i\ell}$'s and hence $H_{i\ell}$'s themselves do not depend on the choice of g_{ij}'s, their translates involve g_{ij}'s.) Our choice of the g_{ij}'s was such that the average (over P terms) of these translates of H's in $\mathcal{K}_o$ has small supremum norm, uniformly for $H \in \mathcal{K}_0$ (see (3.5)). Since P is much smaller than n we can make $||W_{n,\psi^*}f||_{L^2}/n$ as small as we desire.

(e) The only problem left now is that the sum $W^\sigma_{n,\psi^*;i,j;H}$ involves translates of $H \in \mathcal{K}_0$ at $\sigma^{*a_\ell}(g^*)$ rather an at g^*. But again the equicontinuity of σ^* has allowed us to choose a_ℓ's in such a way (see (3.10)) that $W^{\sigma}_{n,\psi^*;i,j;H}(g^*)$ is close to $W^{I}_{n,\psi^*;i,j;H}(g^*)$.

References

[1] R. Ellis, *The construction of minimal discrete flows*, Amer. J. Math. **LXXXVII** (1965), no. 3, 564–574.

[2] H. Keynes and M. Nerurkar, *Ergodicity in affine skew product toral extensions*, Pacific J. Math. **123** (1986), no. 1, 115–126.

[3] H. Keynes and D. Newton, *Minimal (G, σ) extensions*, Pacific J. Math. **77** (1978), 145–163.

[4] ______, *Ergodicity in (G, σ) extensions*, Lecture Notes in Math. **819**, Springer–Verlag, 265–290.

[5] ______, *Minimality for non-abelian (G, σ)-extensions*, Lecture Notes in Math. **668** (1977), Springer–Verlag, 173–178.

[6] M. Nerurkar and H. Sussmann, *Construction of minimal cocycles arising from specific differential equations*, Israel J. Math. **100** (1997), 309–326.

[7] ______, *Construction of ergodic cocycles arising from specific differential equations*, J. Modern Dynamics **1** (2007), no. 2 (to appear).

Manuscript received February 8, 2007

Mahesh Nerurkar
Rutgers University
Camden, NJ 08102, USA

E-mail address: nmahesh@crab.rutgers.edu