. . **Lifting Ergodicity in (G, σ)-Extensions**

Mahesh G. Nerurkar[1] and Hector J. Sussmann[2]

Departiment of Mathematics, Rutgers University, Camden, NJ 08102.
Department of Mathematics, Rutgers University, New Brunswick, NJ 08903.

§1. Introduction

By a *dynamical system* we mean a triple (Y, S, m) where Y is a compact metric space, m is a Borel probability measure on Y, and S is an m-preserving homeomorphism of Y. Let G be a compact metric topological group and let σ be a continuous group automorphism of G. Letter d will denote metric on both, Y and G.

Given a continuous map $\varphi : Y \to G$, define a *σ-skew-product* transformation $S_\varphi : G \times Y \to Y \times G$ by the rule

$$S_\varphi(g, y) = (\sigma[\varphi(y)g], Sy) \qquad (g, y) \in G \times Y.$$

Let ν denote the Haar measure on G, normalized so that $\nu(G) = 1$. Then $\nu \times m$ is invariant under S_φ. The dynamical system $(G \times Y, S_\varphi, \nu \times m)$ is called a (G, σ)-extension or an *affine extension* of (Y, S, 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

$$\mathcal{C}_{erg} \equiv \{\varphi \in C(Y, G) : (G \times Y, S_\varphi, \nu \times m) \text{ is ergodic}\}$$

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 $C(Y, G)$-the metric space of continuous maps from Y to G with the supremum metric. The only other generic ergodicity lifting result for (G, σ) extensions was proved in [KNe] when G is an -n-torus and the genericity was in the-'not so natural'-class of 'closures of coboundaries'.

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

[1]The research was supported by the NSF grant DMS-9972132
[2]The research was supported by the NSF grant DMS-?????

1.1 Definition. *Consider a pair (G,σ) consisting of a compact metric group G and its continuous automorphism σ.*

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

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

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

Following [KN3], we shall call the pair (G,σ) admissible if it satisfies both, Property A and B.

We refer to [KN1,3] for the proof of the following.

1.2 Proposition. *If G is either*

(1) *a compact, metric connected abelian group or*
(2) *a compact, connected Lie group or*
(3) *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 (Y,S,m) *aperiodic* if $m(\{y\in Y: S^ky=y$ for some $k\in\mathbb{N}\}=0$ and *ergodic* if every S-invariant Borel subset E of Y has m measure either zero or one.

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

(1) *(Y,S,m) is a ergodic and aperiodic, and*
(2) *(G,σ) is admissible.*

Then the set

$$C_{erg}(Y,G)=\{\varphi\in C(Y,G):(G\times Y,S_\varphi,\nu_G\times m)\ \textit{is ergodic}\},\tag{1.1}$$

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

§2. Reduction of the Proof.

First we describe a crucial result (Proposition 2.2, which is Corollary 2.4 in [KN2]) that reduces the problem of lifting ergodicity from the case of a general σ to the case when σ is equicontinuous. (The 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.

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

(1) *Let $\hat{G}$ denote the 'dual object' i.e. the set of equivalence classes of finite dimensional irreducible representations of G. A typical element of $\hat{G}$ will be denoted by $[\gamma]$ where γ denotes a representative of the class $[\gamma]$.*
(2) *Let $\hat{\sigma} : \hat{G} \to \hat{G}$ be the map induced by σ, i.e. $\hat{\sigma}[\gamma] = [\gamma \circ \sigma]$.*
(3) *Given a $[\gamma] \in \hat{G}$, let $\mathbb{Z}_{[\gamma]}$ denote the stabilizer subgroup of $[\gamma]$. More precisely,*

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

(4) *Let*

$$\hat{G}_q = \{[\gamma] \in \hat{G} : \mathbb{Z}_{[\gamma]} \textit{ has finite index in } \mathbb{Z}\}.$$

(5) *Let G_q be the annihilator of $\hat{G}_q$, i.e.*

$$G_q \overset{def}{=} Ann\ \hat{G}_q \overset{def}{=} \cap_{[\gamma]\in\hat{G}}\ ker\ \gamma.$$

Then G_q is in fact a closed normal σ invariant subgroup of G. Let
(6) *$\overline{G} = G/G_q$,*
(7) *$\bar{\sigma}$ denote the automorphism induced by σ on $\bar{G}$.*
(8) *Let $\pi_q : G \to G/G_q$ be the canonical homomorphism (which is a continuous and open map). A typical element of $\overline{G}$ will be denoted by*

$$\bar{g} \overset{def}{=} \pi_q(g).$$

(9) *Let $\bar{\nu}$ denote the normalized Haar measure on $\overline{G}$. Observe that $\pi_q^*(\nu) = \bar{\nu}$.*
(10) *Given a $\varphi \in C(Y,G)$ we shall set*

$$\bar{\varphi} = \pi_q \circ \varphi.$$

Thus $\varphi \to \bar{\varphi}$ defines a continuous map from $C(Y,G) \to C(Y,\bar{G})$.
(11) *Let $\bar{S}_{\bar{\varphi}}(\bar{g},y) = (\bar{\sigma}[\bar{\varphi}(y)g], Ty)$, where $\varphi \in C(Y,G)$.*

The subgroup G_q is the *obstruction* to σ being equicontinuous. More precisely, in the language of topological dynamics $(\bar{G},\bar{\sigma})$ is a *maximal equicontinuous factor* of the dynamical system (G,σ). The following theorem essentially allows one to reduce the proof to the case when σ is equicontinuous.

2.2 Proposition. *With the notation as above, given $\varphi \in C(Y,G)$, the dynamical system $(G \times Y, S_\varphi, \nu \times m)$ is ergodic if and only if the system $(\bar{G} \times Y, \bar{S}_{\bar{\varphi}}, \bar{\nu} \times m)$ is ergodic.*

This Proposition implies that

$$C_{erg}(Y,G) = \{\varphi \in C(Y,G) : (\bar{G} \times Y, \bar{S}_{\bar{\varphi}}, \bar{\nu} \times m) \text{ is ergodic}\}.$$

We want to show that this set is residual. We will in fact prove a slightly stronger result, of which Theorem 1.3 is a direct consequence.

To state this result, we first introduce some notation. If $f \in L^2(\bar{G}, \bar{\nu})$, we use $\langle f \rangle$ to denote the average of f over $\overline{G}$ with respect to $\bar{\nu}$, so $\langle f \rangle = \int_{\overline{G}} f(g)\, d\bar{\nu}(g)$. If f belongs to $L^2(\bar{G} \times Y, \bar{\nu} \times m)$ and $y \in Y$, we use f^y to denote the function $\overline{G} \ni \bar{g} \mapsto f(\bar{g}, y) \in \mathbb{C}$. Then $f^y \in L^2(\bar{G}, \bar{\nu})$ for almost all $y \in Y$, and $\int_Y \|f^y\|_{L^2}^2 dm(y) \leq \|f\|_{L^2}^2$.

For $f \in L^2(\bar{G} \times Y, \bar{\nu} \times m)$ set

$$\hat{f}(\bar{g}, y) = \langle f \rangle \stackrel{def}{=} \int_{\overline{G}} f(\bar{g}, y) d\bar{\nu}(\bar{g}).$$

Then $\hat{f}$ is in fact the orthogonal projection of f on the space of square-integrable functions on $\overline{G} \times Y$ that are functions of y only.

We use

$$\begin{aligned} \mathcal{H} &= L^2(\bar{G} \times Y, \bar{\nu} \times m)\,, \\ \mathcal{H}_0 &= \{f \in \mathcal{H} : \int_{\overline{G} \times Y} f\, d(\bar{\nu} \times m) = 0\}\,, \\ \mathcal{H}_{0,av} &= \{f \in \mathcal{H} : \hat{f} \equiv 0\}\,. \end{aligned}$$

Given $\varphi \in C(Y, G)$, define a unitary operator $\hat{\bar{S}}_{\bar{\varphi}}$ on $\mathcal{H}$ by setting

$$\hat{\bar{S}}_{\bar{\varphi}} f = f \circ \bar{S}_{\bar{\varphi}} = f \circ \bar{S}_{\varphi \circ \pi_q}\,. \tag{2.1}$$

Furthermore, let

$$W_{n,\bar{\varphi}} = \frac{1}{n} \sum_{i=0}^{n-1} \hat{\bar{S}}_{\bar{\varphi}}^{\,i}\,. \tag{2.2}$$

We will then prove the following

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

(1) *(Y, S, m) is ergodic and aperiodic and*
(2) *(G, σ) is admissible.*

Then the set

$$C_{erg,av}(Y, G) = \{\varphi \in C(Y, G) : \lim_{n \to \infty} W_{n,\bar{\varphi}} f = 0 \textit{ for all } f \in \mathcal{H}_{0,av}\}, \tag{2.3}$$

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

PROOF THAT THEOREM 2.3 IMPLIES THEOREM 1.3. The assumptions of Theorem 1.3 allow us to apply Theorem 2.3. Let $\varphi \in C_{erg,av}(Y, G)$. If $f \in \mathcal{H}_0$, then $\tilde{f} \overset{def}{=} f - \hat{f} \in \mathcal{H}_{0,av}$. Hence $W_{n,\bar{\varphi}}\tilde{f} \to 0$ as $n \to \infty$. Since S is ergodic, and $\int_Y h(y)\, dm(y) = 0$, $W_n h = \frac{1}{n}\sum_{i=0}^{n-1} h \circ S^i$ converge to 0, where $h(y) \overset{def}{=} \langle f^y \rangle$. Observe that

$$W_n h = W_{n,\bar{\varphi}}\hat{f}.$$

Therefore $W_{n,\bar{\varphi}}f = W_{n,\bar{\varphi}}\tilde{f} + W_{n,\bar{\varphi}}\hat{f} \to 0$ as $n \to \infty$. Since this is true for all $f \in \mathcal{H}_0$, it follows that $(\overline{G} \times Y, \bar{S}_{\bar{\varphi}}, \bar{\nu} \times m)$ is ergodic. Thus Proposition 2.2 shows that $C_{erg,av}(Y,G) \subseteq C_{erg}(Y,G)$ and consequently $C_{erg}(Y,G)$ is residual. $\square$

PROOF OF THEOREM 2.3. Given $f \in \mathcal{H}_{0,av}$, $\varepsilon > 0$ and $\bar{n} \in \mathbb{N}$, define a set $E(f, \varepsilon, \bar{n})$ as follows:

$$E(f,\varepsilon,\bar{n}) = \{\varphi \in C(Y,G) : ||W_{n,\bar{\varphi}}f||_2 < \varepsilon \text{ for some } n \in \mathbb{N} \text{ such that } n > \bar{n}\}. \tag{2.4}$$

Lemma 2.4. *Let $\varphi \in C(Y,G)$, and let $\mathcal{F}$ be a dense subset of $\mathcal{H}_{0,av}$. If $\varphi \in E(f, \frac{1}{n}, \overline{n})$ for all $f \in \mathcal{F}$, $n \in \mathbb{N}$, $\overline{n} \in \mathbb{N}$, then $\varphi \in C_{erg,av}(Y,G)$.*

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

In view of Lemma 2.4, Theorem 1.3 will follow from the Baire category theorem— together with the facts that (a) $C(Y,G)$ is a complete metric space, because G is a compact metric group, (b) $C(G \times Y) \cap \mathcal{H}_{0,av}$ is dense $\mathcal{H}_{0,av}$, and (c) $C(Y \times G)$ is separable— if we prove

Lemma 2.5. *Let $f \in C(\bar{G} \times Y) \cap \mathcal{H}_{0,av}$. Then $E(f, \frac{1}{n}, \overline{n})$ is open and dense in $C(Y,G)$.*

The openness of the set $E(f,\varepsilon,\bar{n})$ is easily checked and we shall establish its density in the next section. 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. First we introduce the necessary notation.

If $h \in C(G)$ and $g \in G$, we define functions $\tau_g h$ by setting

$$(\tau_g h)(x) = h(g^{-1}x), \qquad \text{for } x \in G.$$

Clearly, each τ_g is an isometric linear maps from $C(G)$ onto $C(G)$. Furthermore, $\tau_{g_1}\tau_{g_2} = \tau_{g_1 g_2}$ whenever $g_1, g_2 \in G$.

Lemma 2.6. *Let G be a compact metric group. If $\mathcal{K} \subseteq C(Y,G)$ is compact and $\eta > 0$, then there exists a positive integer P, and members $\gamma_1, \ldots, \gamma_P$ of G, such that*

$$\Big\| \langle h \rangle - \frac{1}{P} \sum_{r=1}^{P} \tau_{\gamma_r^{-1}} h \Big\|_{sup} \leq \eta \text{ for all } h \in \mathcal{K}\,. \tag{2.5}$$

Proof. Since $\mathcal{K}$ is a compact subset of $C(G)$, it is bounded and equicontinuous. Hence there exist (a) a constant C such that $|h(x)| \leq C$ whenever $x \in G$ and $h \in \mathcal{K}$, and (b) a neighborhood W of e_G such that $|h(x) - h(y)| < \frac{\eta}{2}$ whenever $xy^{-1} \in W$ and $h \in \mathcal{K}$. Partition G into measurable subsets $G_1, \ldots, G_R$ such that each G_r is contained in a left translate $\xi_r W$ of W. Let $\tilde{p}_r = \nu(G_r)$.

If $h \in \mathcal{K}$ and $x \in G$, then

$$\begin{aligned}\langle h \rangle &= \int_G h(gx)\, d\nu(g) = \sum_{r=1}^{R} \int_{G_r} h(gx)\, d\nu(g) \\ &= \sum_{r=1}^{R} \int_{G_r} (h(gx) - h(\xi_r x))\, d\nu(g) + \sum_{r=1}^{R} \tilde{p}_r (\tau_{\xi_r^{-1}} h)(x)\,.\end{aligned}$$

If $g \in G_r$, then $(gx)(\xi_r x)^{-1} = g\xi_r^{-1} \in W$, so $|h(gx) - h(\xi_r x)| \leq \frac{\eta}{2}$. Therefore

$$\Big| \langle h \rangle - \sum_{r=1}^{R} \tilde{p}_r \tau_{\xi_r^{-1}} h(x) \Big| \leq \frac{\eta}{2}$$

for all $h \in \mathcal{K}$, $x \in G$. Now, let M be a positive integer, and write

$$\tilde{p}_r = \frac{p_r}{M} + \hat{p}_r\,,$$

where $0 \leq \hat{p}_r < \frac{1}{M}$. Let $P = p_1 + \ldots + p_R$. Then $1 - \frac{R}{M} \leq \frac{P}{M} \leq 1$. Moreover,

$$\Big| \sum_{r=1}^{R} \hat{p}_r \tau_{\xi_r^{-1}} h(x) \Big| \leq \frac{CR}{M}\,.$$

So

$$\Big| \langle h \rangle - \sum_{r=1}^{R} \frac{p_r}{M} \tau_{\xi_r^{-1}} h(x) \Big| \leq \frac{\eta}{2} + \frac{CR}{M}\,.$$

Therefore

$$\Big| \langle h \rangle - \sum_{r=1}^{R} \frac{p_r}{P} \tau_{\xi_r^{-1}} h(x) \Big| \leq \frac{\eta}{2} + \frac{2CR}{M}\,.$$

Now choose M such that $4CR < M\eta$. Then

$$\Big\| \langle h\rangle - \sum_{r=1}^{R} \frac{p_r}{P}\tau_{\xi_r^{-1}} h \Big\|_{sup} \leq \eta \text{ for all } h \in \mathcal{K}\,.$$

If we now define

$$\begin{gathered} \gamma_1 = \gamma_2 = \cdots = \gamma_{p_1} = \xi_1\,, \\ \gamma_{p_1+1} = \gamma_{p_1+2} = \cdots = \gamma_{p_1+p_2} = \xi_2\,, \\ \cdots \\ \gamma_{p_1+p_2+\cdots+p_{R-1}+1} = \gamma_{p_1+p_2+\cdot+p_{R-1}+2} = \cdots = \gamma_P = \xi_R\,, \end{gathered}$$

then

$$\Big\| \langle h\rangle - \frac{1}{P}\sum_{r=1}^{P} \tau_{\gamma_r^{-1}} h \Big\|_{sup} \leq \eta \text{ for all } h \in \mathcal{K}\,. \tag{2.6}$$

□

§3. Density of $E(f,\varepsilon,\bar{n})$.

So we turn to the proof of the density of $E(f,\varepsilon,\bar{n})$. For this purpose, we fix the following objects:

(1) an $f \in C(\bar{G}\times Y)\cap \mathcal{H}_{0,av}$;
(2) a $\varphi \in C(Y,G)$;
(3) a positive number ε;
(4) $\tilde{U}$-a neighbourhood of e_G-the identity of G.
(5) a positive integer $\bar{n}$.

For the given φ and ε let δ be as in Property B. Let $U = \{g \in \tilde{U} : d(g,e_G) < \delta\}$ and set

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

We will then show that $\mathcal{V}(\varphi,U)\cap E(f,\varepsilon,\bar{n}) \neq \emptyset$. If this is true for arbitrary $\tilde{U}$ and φ, it will follow that $E(f,\varepsilon,\bar{n})$ is dense in $C(Y,G)$, concluding our proof.

For $\varphi \in C(Y,G)$, the iterates $\hat{S}^j_\varphi$ of $\hat{S}_\psi$ are given by

$$\hat{S}^j_\varphi f(g,y) = f(\varphi(y,j)\sigma^j(g), S^j y)\,, \tag{3.1}$$

where

$$\varphi(y,j) = \sigma\big(\varphi(S^{j-1}y)\big)\cdot\sigma^2\big(\varphi(S^{j-2}y)\big)\cdot\cdots\cdot\sigma^j\big(\varphi(y)\big) \text{ for } j>0, \tag{3.2}$$

and we let $\varphi(y,0)=e_G$. Note that φ satisfies the following 'cocycle identity'

$$\varphi(y,m+n)=\varphi(S^m y,n)\sigma^n[\varphi(y,m)], \qquad m,n\in\mathbb{N}. \tag{3.3}$$

Identities (3.1,3.2) remain valid when φ, S and σ are replaced by $\bar\varphi, \bar S$ and $\bar\sigma$ respectively.

We shall denote by $\tau_{\bar g}$ the linear isometry on $C(\overline{G})$ induced by the translation by $\bar g\in\overline{G}$. Furthermore, if $h\in C(\overline{G})$ define

$$\hat{\bar\sigma}^i(h)(x)=h(\bar\sigma^i(x)) \qquad \text{for } x\in\overline{G} \text{ and } i\in\mathbb{N}\cup\{0\}.$$

Then $\hat{\bar\sigma}^i$ also defines a linear isometry from $C(\overline{G})$ onto $C(\overline{G})$ for each $i\in\mathbb{N}\cup\{0\}$. With this notation at hand, we begin describing the construction of a map $\psi\in\mathcal{V}(\varphi,U)\cap E(f,\varepsilon,\bar n)$ in a series of steps.

Step 1 : Let $\mathcal{K}$ to be the closed convex hull of the set of all functions given by

$$\{\tau_{\bar g}\big(\hat{\bar\sigma}^i f^y\big) : \bar g\in\overline{G},\ i\in\mathbb{N}\cup\{0\},\ y\in Y\}.$$

Since $\bar\sigma$ is equicontinuous on $\overline{G}$, $\mathcal{K}$ is a compact convex subset of $C(\overline{G})$. moreover it is invariant under $\hat{\bar\sigma}$ and $\tau_{\bar g}$ for all $\bar g\in\overline{G}$ and $\langle h\rangle=0$ 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 inequality

$$\|h-h_0\|_{sup}<\frac{\varepsilon}{16} \tag{3.4}$$

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

$$\beta=\frac{\varepsilon}{8\hat\kappa}\ . \tag{3.5}$$

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

$$\text{if } d(\bar g_1,\bar g_2)<\delta_1,\ \text{ then } |H(\bar g_1)-H(\bar g_2)|<\frac{\beta}{2},\ \bar g_1,\bar g_2\in\overline{G}, \tag{3.6}$$

for all $H\in\mathcal{K}$.

Step 2 : Now applying Lemma 2.6 to the above set $\mathcal{K}$ and with $\eta=\beta$, we get a positive integer P and $\gamma_1,\cdots,\gamma_P\in G$ such that

$$\Big\|\frac{1}{P}\sum_{r=1}^{P}\tau_{\gamma_r^{-1}}h\Big\|_{sup}\le\beta \ \text{ for all } h\in\mathcal{K}\,. \tag{3.7}$$

Step 3 : For the given neighbourhood U of e_G, using Property A, we get a positive integer κ such that for any κ right translates $U_0,\cdots,U_{\kappa-1}$ of U we have,

$$\sigma\big(U_{\kappa-1}\big)\sigma^2\big(U_{\kappa-2}\big)\cdot\cdots\cdot\sigma^{\kappa}\big(U_0\big)=G.$$

Step 4 : Next, we choose $N = (\kappa + \mu)\lambda$ where μ and λ are large positive integers, the precise choice of which will be described later, but for now, we make sure that

$$d(\bar{\sigma}^{i(\kappa+\mu)}(\bar{g}), \bar{g}) < \delta_1, \qquad \text{for all } \bar{g} \in \overline{G} \text{ and for all } i \in \{1, 2, \cdots, \lambda\},$$

where δ_1 is as in Step 1. By the equicontinuity of $\bar{\sigma}$ such choice of μ and λ is possible.

Step 5 : Using Rokhlin's lemma, pick a Borel subset $\hat{E}$ of Y with the property that the sets $\hat{E}$, $S\hat{E}, \ldots, S^{N-1}\hat{E}$ are pairwise disjoint, and

$$m(\hat{E} \cup S\hat{E} \cup \ldots \cup S^{N-1}\hat{E}) \geq 1 - c_1,$$

where c_1 is a positive constant whose choice will be described later.

Step 6 : Using uniformly continuity of f on $\overline{G} \times Y$, choose

(1) a positive number c_2 such that $|f(\bar{g}, y) - f(\bar{g}, z)| \leq \frac{\varepsilon}{8}$ whenever $\bar{g} \in \overline{G}$, $y, z \in Y$, and $d(y, z) \leq c_2$,

(2) a neighborhood V of the identity $\bar{e}_G$ in $\overline{G}$ such that $|f(\bar{g}, y) - f(\bar{h}, y)| \leq \frac{\varepsilon}{8}$ whenever $\bar{g}, \bar{h} \in \overline{G}$, $y \in Y$, and $\bar{g}\bar{h}^{-1} \in V$.

Step 7 : Let Φ^N be the subset of $C(Y, G)$ consisting of all functions $\eta : Y \to G$ of the form

$$\eta(y) = g_r \cdot \sigma\big(\varphi(S^r y)\big) \cdot g_{r-1} \cdot \sigma^2(\varphi(S^{r-1}y)\big) \cdot \cdots \cdot \sigma^{r+1}\big(\varphi(y)\big) \cdot g_0 \tag{3.8}$$

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 (3.8), regarded as a function of y, depends continuously on the $r+1$-tuple $(g_0, \ldots, g_r)$, 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 $\bar{e}_G$ chosen in Step 6, there exists a constant $\omega(N, V)$ with the property that $\bar{\pi}_q\big(\eta(y)\eta(z)^{-1}\big) \in V$ whenever $\eta \in \Phi^N$, $y, z \in Y$, and $d(y, z) \leq \omega(N, V)$.

Step 8 : Write $J = \{0, 1, \ldots, N-1\}$. Using the continuity of maps S^j, $(j \in J)$, pick a positive c_3 such that

(1) $d(S^j y, S^j z) \leq c_2$ whenever $d(y, z) \leq c_3$ and $j \in J$,

(2) $c_3 \leq \omega(N, V)$.

Step 9 : We partition Y into finitely many Borel-measurable sets $Y_1, \ldots, Y_s$ of diameter less than c_3, and let $\hat{E}_i = \hat{E} \cap Y_i$ for $i = 1, \ldots, s$. Write $I = \{1, \ldots, s\}$. Using the fact that m is regular, pick compact sets E_i such that $E_i \subseteq \hat{E}_i$ and $m(\hat{E}_i \backslash E_i) < \frac{c_1}{sN}$. Write $E_{ij} = S^j E_i$ for $i \in I$, $j \in J$ (so $E_{i0} = E_i$). Then the sets E_{ij} are pairwise disjoint, and

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

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

Step 10 : We pick points $\bar{y}_i \in E_i$ for each $i \in I$, and then let $\bar{y}_{i,j} = S^j \bar{y}_i$, so $\bar{y}_{i,j} \in E_{ij}$.

Step 11 : 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 $J = [0, 2, \cdots, N-1]$ into λ blocks $J_1, \dots, J_\lambda$ of length $\kappa + \mu$, given by

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

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

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

The map ψ will satisfy

$$\psi(\bar{y}_{i,j}) = \varphi(\bar{y}_{i,j}) \text{ if } j \in J^2_\ell\,, \tag{3.9}$$

and

$$\psi(\bar{y}_{i,j}) = g_{ij}\varphi(\bar{y}_{i,j}) \text{ if } j \in J^1_\ell\,, \tag{3.10}$$

where the g_{ij} are members of U whose choice will be described later. By Property B there exists a continuous map $\theta : Y \to U$ such that $\theta(\bar{y}_{i,j}) = g_{ij}$ for $j \in J^1_\ell$ and $\theta(\bar{y}_{i,j}) = e_G$ for $j \in J^2_\ell$. If we then define $\psi(y) = \theta(y)\varphi(y)$, then ψ satisfies (3.9) and (3.10) and furthermore $\psi \in \mathcal{V}(\varphi, U)$.

Step 12 : Now we are now going to estimate the norm $\|W_{n,\bar{\psi}} f\|_{L_2}$ for large n. We will choose $n = (\kappa+\mu)\rho$, where ρ is a large positive integer whose choice will be described later, but it is such that $\frac{\lambda}{\rho}$ is still a large positive integer.

Now we begin estimating $\|W_{n,\bar{\psi}} f\|_{L_2}$. We let

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

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

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

Outside $\overline{G} \times A$, the function $W_{n,\bar{\psi}} f$ is bounded pointwise by $\|f\|_{sup}$, so

$$\int_{\overline{G}\times(Y\setminus A)} |W_{n,\bar{\psi}} f(\bar{g})|^2 d\bar{\nu}(\bar{g}) \le \|f\|^2_{sup} \left(2c_1 + \frac{\rho}{\lambda}\right). \tag{3.11}$$

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

$$\overline{W}_{n,\bar{\psi};i,j} f(\bar{g}) \overset{def}{=} W_{n,\bar{\psi}} f(\bar{g}, \bar{y}_{i,j}) = \frac{1}{n} \Big(\sum_{k=0}^{n-1} f(\bar{\psi}(\bar{y}_{i,j}, k) \bar{\sigma}^k(\bar{g}), S^k \bar{y}_{i,j}) \Big) . \tag{3.12}$$

Since both y and $\bar{y}_{i,j}$ belong to E_{ij}, the distance $d(y, \bar{y}_{i,j})$ is not larger than c_3. Then $d(S^k y, S^k \bar{y}_{i,j}) \leq c_2$ for all k. Therefore

$$|f(\bar{\psi}(y,k)\bar{h}, S^k y) - f(\bar{\psi}(y,k)\bar{h}, S^k \bar{y}_{i,j})| \leq \frac{\varepsilon}{8} \tag{3.13}$$

for all k and $\bar{h} \in \overline{G}$. Now notice that for each $0 \leq k \leq n$, the restriction of the function $y \to \psi(y,k)$ to the set A 'belongs to Φ^N' (i.e. equals restriction to A of some member of Φ^N). Since $c_3 \leq \omega(N, V)$, we conclude that

$$\pi_q\big(\psi(y,k)\bar{h}\big(\psi(\bar{y}_{i,j},k)\bar{h}\big)^{-1}\big) = \bar{\psi}(y,k)\bar{\psi}(\bar{y}_{i,j},k) \in V,$$

for all $y \in A$, $\bar{h} \in \overline{G}$ and $0 \leq k \leq n$. Therefore

$$|f(\bar{\psi}(y,k)\bar{h}, S^k \bar{y}_{i,j}) - f(\bar{\psi}(\bar{y}_{i,j},k)\bar{h}, S^k \bar{y}_{i,j})| \leq \frac{\varepsilon}{8} . \tag{3.14}$$

By taking $\bar{h} = \bar{\sigma}^k(\bar{g})$, and using (3.13) and (3.14) we have

$$|\overline{W}_{n,\bar{\psi};i,j} f(\bar{g}) - W_{n,\bar{\psi}} f(\bar{g}, y)| \leq \frac{\varepsilon}{4} \quad \text{whenever } \bar{g} \in \overline{G}, y \in E_{ij} . \tag{3.15}$$

Now fix an $i \in I$ and $j \in J$ such that $j \leq N - n - 1$. Then for each $k \in \{0, \ldots, n-1\}$ the point $S^k \bar{y}_{i,j}$ is equal to $\bar{y}_{i,j+k}$, and belongs to $E_{i,j+k}$. As k varies from 0 to $n-1$, the index $j+k$ takes all values in the integer interval $\Gamma_n(j) = \{j, j+1, \ldots, j+n-1\}$, which is a subset of J because $j \leq N-n-1$. 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, and the "bad" part $\Gamma_n^b(j)$ has at most $2(\kappa+\mu)$ members. Let $\Delta_n(j)$ be the set of indices ℓ such that $J_\ell \subseteq \Gamma_n(j)$. Then

$$\overline{W}_{n,\bar{\psi};i,j} f(\bar{g}) = \frac{1}{n} \Big(\overline{W}^c_{n,\bar{\psi};i,j} f(\bar{g}) + \overline{W}^b_{n,\bar{\psi};i,j} f(\bar{g}) \Big) , \tag{3.16}$$

where $\overline{W}^c_{n,\bar{\psi};i,j} f(\bar{g})$, $\overline{W}^b_{n,\bar{\psi};i,j} f(\bar{g})$ are, respectively, the contributions to the right-hand side of (3.12) of the terms for which $k+j \in \Gamma_n^c(j)$ and those for which $k+j \in \Gamma_n^b(j)$, so that, in particular,

$$\overline{W}^c_{n,\bar{\psi};i,j} f(\bar{g}) = \sum_{\ell \in \Delta_n(j)} \sum_{k+j \in J_\ell} f(\bar{\psi}(\bar{y}_{i,j}, k) \bar{\sigma}^k(\bar{g}), \bar{y}_{i,j+k}) , \tag{3.17}$$

and

$$\overline{W}^{b}_{n,\bar\psi;i,j}f(\bar g)=\sum_{k+j\in\Gamma^b_n(j)} f(\bar\psi(\bar y_{i,j},k)\bar\sigma^k(\bar g),\bar y_{i,j+k})\,. \tag{3.18}$$

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

$$\frac{1}{n}\Big|\overline{W}^{b}_{n,\bar\psi;i,j}f(\bar g)\Big|\le\frac{2(\kappa+\mu)\|f\|_{sup}}{n}=\frac{2\|f\|_{sup}}{\rho}\,. \tag{3.19}$$

Now let us analyze the contribution to $\overline{W}^{c}_{n,\bar\psi;i,j}f(\bar g)$ of the terms of an individual sum

$$\overline{W}^{c}_{n,\bar\psi;i,j;\ell}f(\bar g)=\sum_{k+j\in J_\ell} f(\bar\psi(\bar y_{i,j},k)\bar\sigma^k(\bar g),\bar y_{i,j+k})=\sum_{u\in J_\ell} f(\bar\psi(\bar y_{i,j},u-j)\bar\sigma^{u-j}(\bar g),\bar y_{i,u})\,.$$

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 (3.3) we get

$$\begin{aligned}\overline{W}^{c}_{n,\bar\psi;i,j;\ell}f(\bar g)&=\sum_{v=0}^{\kappa+\mu-1} f\big(\bar\psi(\bar y_{i,j},v+a_\ell-j)\bar\sigma^{v+a_\ell-j}(\bar g),\bar y_{i,a_\ell+v}\big)\\&=\sum_{v=0}^{\kappa+\mu-1} f\big(\bar\psi(\bar y_{i,a_\ell},v)\bar\sigma^{v}[\bar\psi(\bar y_{i,j},a_\ell-j)]\bar\sigma^{v+a_\ell-j}(\bar g),\bar y_{i,a_\ell+v}\big)\\&=\sum_{v=0}^{\kappa+\mu-1} f\big(\bar\psi(\bar y_{i,a_\ell},v)\bar\sigma^{v}\big[\bar\psi(\bar y_{i,j},a_\ell-j)\bar\sigma^{a_\ell-j}(\bar g)\big],\bar y_{i,a_\ell+v}\big).\end{aligned}$$

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

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

where

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

and

$$\begin{aligned}\overline{W}^{c,+}_{n,\bar\psi;i,j;\ell}f(\bar g)&=\sum_{w=0}^{\mu-1} f\big(\bar\psi(\bar y_{i,a_\ell},\kappa+w)\bar\sigma^{\kappa+w}\big[\bar\psi(\bar y_{i,j},a_\ell-j)\bar\sigma^{a_\ell-j}(\bar g)\big],\bar y_{i,a_\ell+\kappa+w}\big)\\&=\sum_{w=0}^{\mu-1} f\big(\bar\psi(\bar y_{i,a_\ell+\kappa},w)\bar\sigma^{w}\big[\bar\psi(\bar y_{i,a_\ell},\kappa)\big]\bar\sigma^{\kappa+w}\big[\bar\psi(\bar y_{i,j},a_\ell-j)\bar\sigma^{a_\ell-j}(\bar g)\big],\bar y_{i,a_\ell+\kappa+w}\big)\\&=\sum_{w=0}^{\mu-1} f\big(\bar\psi(\bar y_{i,a_\ell+\kappa},w)\bar\sigma^{w}\big[\bar\psi(\bar y_{i,a_\ell},\kappa)\bar\sigma^{\kappa}[\bar\psi(\bar y_{i,j},a_\ell-j)\bar\sigma^{a_\ell-j}(\bar g)]\big],\bar y_{i,a_\ell+\kappa+w}\big).\end{aligned}$$

In the above computation we have used the cocycle identity (3.3) and $\bar{y}_{i,j} = S^j\bar{y}$. Once again the cocycle identity allows us to write

$$\bar{\psi}(\bar{y}_{i,j}, a_\ell - j) = \bar{\psi}(\bar{y}_i, a_\ell)\bar{\sigma}^{(a_\ell - j)}\left[\bar{\psi}(\bar{y}_i, j)^{-1}\right].$$

Hence,

$$\overline{W}^{c,+}_{n,\bar{\psi};i,j;\ell} f(\bar{g})$$
$$= \textstyle\sum_{w=0}^{\mu-1} f\big(\bar{\psi}(\bar{y}_{i,a_\ell+\kappa}, w)\bar{\sigma}^w\left[\bar{\psi}(\bar{y}_{i,a_\ell}, \kappa)\bar{\sigma}^\kappa[\bar{\psi}(\bar{y}_i, a_\ell)\bar{\sigma}^{(a_\ell - j)}[\bar{\psi}(\bar{y}_i, j))^{-1}\bar{g}]]\right], \bar{y}_{i,a_\ell+\kappa+w}\big).$$

Clearly, each function $\overline{W}^{c,-}_{n,\bar{\psi};i,j;\ell} f$ 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

$$\frac{1}{n}\Big|\sum_{\ell\in\Delta_n(j)} \overline{W}^{c,-}_{n,\bar{\psi};i,j;\ell} f(\bar{g})\Big| \le \frac{\rho\kappa\|f\|_{sup}}{n} = \frac{\kappa\|f\|_{sup}}{\kappa+\mu}\,. \tag{3.20}$$

We now turn to the crucial estimate, namely, the bound for the functions $\overline{W}^{c,+}_{n,\bar{\psi};i,j;\ell} f$. It is at this point that we will choose the g_{ij}. 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

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

Notice that

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

so that in fact the functions $F_{i,\ell}$ do not depend at all on ψ, i.e., on the choice of the g_{ij}. Next, observe that

$$\overline{W}^{c,+}_{n,\bar{\psi};i,j;\ell} f(\bar{g}) = \mu F_{i,\ell}\big(\bar{\psi}(\bar{y}_{i,a_\ell}, \kappa)\bar{\sigma}^\kappa\left[\bar{\psi}(\bar{y}_i, a_\ell)\bar{\sigma}^{a_\ell - j}[\bar{\psi}(\bar{y}_i, j)^{-1}\bar{g})]\right]\big).$$

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

$$\{\tau_{\bar{g}}\big(\hat{\bar{\sigma}}^i f^y\big) : \bar{g} \in \overline{G},\ i \in \mathbb{N}\cup\{0\},\ y \in Y\}.$$

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 $\|F_{i,\ell} - H_{i,\ell}\|_{sup} \le \frac{\varepsilon}{8}$. 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 $\alpha_{i,H,1} < \alpha_{i,H,2} < \ldots < \alpha_{i,H,\Lambda(i,H)}$. Then for each $i \in I$, $\ell \in L$, there exists a unique $k = k(i,\ell) \in \{1,\ldots,\Lambda(i,H_{i,\ell})\}$ 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 $k(i,\ell) = \hat{k}(i,\ell)P + r(i,\ell)$, where $\hat{k}(i,\ell)$ 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 $g_{ij} = e_G$ if $j \in \cup_\ell J^2_\ell$. 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

$$\begin{aligned}
&\bar{\psi}(\bar{y}_{i,a_\ell},\kappa)\cdot\bar{\sigma}^\kappa[\bar{\psi}(\bar{y}_i,a_\ell)] \\
&\qquad = \bar{\sigma}[\bar{\psi}(\bar{y}_{i,a_\ell+\kappa-1})]\bar{\sigma}^2[\bar{\psi}(\bar{y}_{i,a_\ell+\kappa-2})]\cdots\bar{\sigma}^\kappa[\bar{\psi}(\bar{y}_{i,a_\ell})]\bar{\sigma}^\kappa[\bar{\psi}(\bar{y}_i,a_\ell)] \\
&\qquad = \bar{\sigma}[\bar{g}_{i,a_\ell+\kappa-1}\cdot\bar{z}_{\kappa-1}]\cdot\bar{\sigma}^2[\bar{g}_{i,a_\ell+\kappa-2}\cdot\bar{z}_{\kappa-2}]\cdot\cdots\bar{\sigma}^k[\bar{g}_{i,a_\ell}\cdot\bar{z}_0]\,,
\end{aligned} \tag{3.21}$$

where $z_k = \varphi(\bar{y}_{i,a_\ell+k})$ for $0 < k < \kappa$, $z_0 = \varphi(\bar{y}_{i,a_\ell})\cdot\psi(\bar{y}_i,a_\ell)$, (and recall that $\bar{z} = \pi_q(z)$, $\bar{\psi} = \psi\circ\pi_q$). Note that z_0 is determined, because the g_{ij} have already been selected for $j < a_\ell$. Now, our choice of κ in Step 3 guarantees that $\sigma[Uz_{\kappa-1}]\cdot\sigma^2[Uz_{\kappa-2}]\cdot\cdots\sigma^\kappa[Uz_0] = G$. Since the map $\pi_q : G \to G/G_q$ is onto, we can choose $g_{i0}, g_{i1},\cdots, g_{i,\kappa-1}$ such that

$$\bar{\psi}(\bar{y}_{i,a_\ell},\kappa)\cdot\bar{\sigma}^\kappa[\bar{\psi}(\bar{y}_i,a_\ell)] = \gamma_{r(\ell,i)}\,. \tag{3.22}$$

Then, for any i, j,

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

Let

$$\mathcal{W}^\sigma_{n,\bar{\psi};i,j}(\bar{g}) = \mu\sum_{H\in\mathcal{K}_0}\ \sum_{\ell\in A(i,H)\cap\Delta_n(j)} \tau_{\gamma^{-1}_{r(\ell,i)}} H(\bar{\sigma}^{a_\ell}(\bar{g})).$$

Now fix a $\bar{g} \in \overline{G}$ and set $\bar{g}^*_{i,j} = \bar{\sigma}^{\kappa-j}[\bar{\psi}(\bar{y}_i,j)^{-1}\bar{g}]$. Since $\|H - F_{i,\ell}\|_{sup} \le \frac{\varepsilon}{16}$ whenever $\ell \in A(i,H)$, we have the bound

$$\Big|\mathcal{W}^\sigma_{n,\bar{\psi};i,j}(\bar{g}^*_{i,j}) - \sum_{\ell\in\Delta_n(j)} \overline{W}^{c,+}_{n,\bar{\psi};i,j;\ell} f(\bar{g})\Big| \le \frac{\mu\rho\varepsilon}{16}\,,$$

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

$$\frac{1}{n}\Big|\mathcal{W}^\sigma_{n,\bar{\psi};i,j}(\bar{g}^*_{i,j}) - \sum_{\ell\in\Delta_n(j)} \overline{W}^{c,+}_{n,\bar{\psi};i,j;\ell} f(\bar{g})\Big| \le \frac{\varepsilon}{16}\,, \tag{3.23}$$

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

We now turn to the task of estimating $\mathcal{W}^{\sigma}_{n,\bar{\psi};i,j}$. First, let

$$\mathcal{W}_{n,\bar{\psi};i,j}(\bar{g}) = \mu \sum_{H\in\mathcal{K}_0} \sum_{\ell\in A(i,H)\cap\Delta_n(j)} \tau_{\gamma^{-1}_{r(\ell,i)}} H(\bar{g}).$$

We write

$$\mathcal{W}^{\sigma}_{n,\bar{\psi};i,j}(\bar{g}) = \mu \sum_{H\in\mathcal{K}_0} \mathcal{W}^{\sigma}_{n,\psi;i,j;H}(\bar{g}) \qquad \text{and} \qquad \mathcal{W}_{n,\bar{\psi};i,j}(\bar{g}) = \mu \sum_{H\in\mathcal{K}_0} \mathcal{W}_{n,\bar{\psi};i,j;H}(\bar{g}),$$

where

$$\mathcal{W}^{\sigma}_{n,\bar{\psi};i,j;H}(\bar{g}) = \sum_{\ell\in A(i,H)\cap\Delta_n(j)} \tau_{\gamma^{-1}_{r(\ell,i)}} H(\bar{\sigma}^{a_\ell}\bar{g}) \text{ and}$$

$$\mathcal{W}_{n,\bar{\psi};i,j;H}(\bar{g}) = \sum_{\ell\in A(i,H)\cap\Delta_n(j)} \tau_{\gamma^{-1}_{r(\ell,i)}} H(\bar{g}). \tag{3.24}$$

Since $\Delta_n(j)$ has at most ρ members and $a_\ell = \ell(\kappa+\mu)$, our choice of $\kappa+\mu$ in Step 4 gives

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

Again, since the cardinality of $\mathcal{K}_0$ is $\hat{\kappa}$ and $\frac{\mu\rho}{n} \le 1$, we get

$$\left|\mathcal{W}^{\sigma}_{n,\bar{\psi};i,j}(\bar{g}) - \mathcal{W}_{n,\bar{\psi};i,j}(\bar{g})\right| < \frac{\varepsilon}{16} \quad \text{for all } \bar{g}\in\overline{G}.$$

Combining this with (3.23) yields

$$\frac{1}{n}\Big| \sum_{\ell\in\Delta_n(j)} \overline{W}^{c,+}_{n,\bar{\psi};i,j;\ell} f(\bar{g}) - \mathcal{W}_{n,\bar{\psi};i,j}(\bar{g})\Big| \le \frac{\varepsilon}{8}. \tag{3.25}$$

Now we estimate $\mathcal{W}_{n,\bar{\psi};i,j}(\bar{g})$. Recall that

$$\mathcal{W}_{n,\bar{\psi};i,j}(\bar{g}) = \mu \sum_{H\in\mathcal{K}_0} \sum_{\ell\in A(i,H)\cap\Delta_n(j)} \tau_{\gamma^{-1}_{r(\ell,i)}} H(\bar{g}).$$

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 (3.24) is equal to $q_{i,j,H}\sum_{r=1}^{P} \tau_{\gamma_r^{-1}} H(\bar{g})$, whose absolute value is bounded by $P\beta q_{i,j,H}$, because $\langle H\rangle = 0$. The sum of the remaining $p_{i,j,H}$ terms is bounded by $p_{i,j,H}\|f\|_{sup}$. Thus

$$\left|\mathcal{W}_{n,\bar{\psi};i,j;H}(\bar{g})\right| \le P\beta q_{i,j,H} + p_{i,j,H}\|f\|_{sup}\,.$$

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

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

Therefore

$$\|\mathcal{W}_{n,\psi;i,j}\|_{sup} \le \mu\hat\kappa\rho\beta + \mu\hat\kappa P\|f\|_{sup}\,. \tag{3.26}$$

If we combine (3.25)and (3.26), we find

$$\frac{1}{n}\Big\|\sum_{\ell\in\Delta_n(j)}\overline{W}^{c,+}_{n,\bar\psi;i,j;\ell}f(\bar g)\Big\|_{sup} \le \frac{\varepsilon}{8} + \frac{\mu\hat\kappa\rho\beta}{n} + \frac{\mu\hat\kappa P}{n}\|f\|_{sup}\,, \tag{3.27}$$

We now use (3.20) and get

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

Then (3.19) implies

$$|\overline{W}_{n,\bar\psi;i,j}(\bar g)| \le \frac{2\|f\|_{sup}}{\rho} + \frac{\kappa\|f\|_{sup}}{\kappa+\mu} + \frac{\varepsilon}{8} + \frac{\mu\hat\kappa\rho\beta}{n} + \frac{\mu\hat\kappa P}{n}\|f\|_{sup}\,. \tag{3.28}$$

Finally, we use (3.15) and get the pointwise estimate

$$|W_{n,\bar\psi}f(\bar g,y)| \le \frac{\varepsilon}{4} + \frac{2\|f\|_{sup}}{\rho} + \frac{\kappa\|f\|_{sup}}{\kappa+\mu} + \frac{\varepsilon}{8} + \frac{\mu\hat\kappa\rho\beta}{n} + \frac{\mu\hat\kappa P}{n}\|f\|_{sup}\,, \tag{3.29}$$

valid whenever $\bar g \in G$ and $y \in A$. Now note that depending on the initial given data (i.e. $f, \varepsilon, \varphi, \tilde U$ and $\bar n$ we have so far chosen $\kappa, \hat\kappa$ and P. Next, we choose ρ and μ such that

$$n = (\kappa+\mu)\rho > \bar n\,, \quad \frac{\|f\|_{sup}}{\rho} \le \frac{\varepsilon}{16}\,, \quad \frac{\kappa\|f\|_{sup}}{\mu} \le \frac{\varepsilon}{8}\,, \quad \text{and} \quad \frac{P\hat\kappa\|f\|_{sup}}{\rho} \le \frac{\varepsilon}{8}\,.$$

Then (3.29) implies that

$$|W_{n,\bar\psi}f(\bar g,y)| \le \frac{7\varepsilon}{8} \quad \text{whenever} \quad g\in G \text{ and } y \in A\,. \tag{3.30}$$

This, together with (3.11), implies that

$$\int_{\overline{G}\times Y}|W_{n,\bar\psi}f|^2 \le \frac{49\varepsilon^2}{64} + \Big(2c_1 + \frac{\rho}{\lambda}\Big)\|f\|^2_{sup}\,.$$

Finally we choose λ and c_1 so that $(2c_1+\frac{\rho}{\lambda})\|f\|^2_{sup} \le \frac{14\varepsilon^2}{64}$, and conclude that $\|W_{n,\bar\psi}f\|_{L^2} < \varepsilon$.

Since $n > \bar n$, we have proved that $\psi \in E(f,\varepsilon,\bar n)$. We have already checked that $\psi \in \mathcal{V}(\varphi, U)$. Hence

$$\varphi \in E(f,\varepsilon,\bar n) \cap \mathcal{V}(\varphi, U)\,,$$

concluding the proof of Lemma 2.5 and subsequently the proof of Theorem 2.3 as well. ∎

References

[E] R. Ellis : The construction of minimal discrete flows, Amer. J. Math., vol. LXXXVII, no. 3 (1965), 564-574.

[KNe] H. Keynes and M. Nerurkar : Ergodicity in affine skew product toral extensions, Pacific Journal of Mathematics, Vol. 123, No. 1, (1986) 115-126.

[KN1] H. Keynes and D. Newton : Minimal (G, σ) extensions, Pacific J. Math. 77 (1978), 145-163.

[KN2] H. Keynes and D. Newton : Ergodicity in (G, σ) extensions, Proceedings of a conference on Dynamical Systems, Springer Verlag Lecture Notes in Mathematics V. 819, 265-290

[KN3] H. Keynes and D. Newton : Minimality for non-abelian (G, σ)-extensions, Springer Verlag Lecture Notes no. 668, NDSU Cnference on Dynamical Systems (1977) 173-178.

[NS] M. Nerurkar, H. Sussmann : Construction of minimal cocycles arising from specific differential equations, Israel Journal of Mathematics, Vol.100 (1997) 309-326.