Ergod. Th. & Dynam. Sys. (2005), **25**, 1357–1370
doi:10.1017/S0143385703000518
© *2005 Cambridge University Press*
Printed in the United Kingdom

A non-conventional ergodic theorem for a nilsystem

T. ZIEGLER

Department of Mathematics, The Ohio State University, 100 Math Tower, 231 West 18th Avenue, Columbus, OH 43210-1174, USA
(e-mail: tamar@math.ohio-state.edu)

(*Received* 29 *April* 2002 *and accepted in revised form* 3 *September* 2003)

Abstract. We prove a non-conventional pointwise convergence theorem for a nilsystem, and give an explicit formula for the limit.

1. *Introduction*

Let (X, μ, T) be a measure-preserving system (mps). Non-conventional averages of the form

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

where $f_j \in L^\infty(X)$, were first introduced by Furstenberg in the ergodic theoretic proof of Szemerédi's theorem on arithmetic progressions [**3**]. Recent developments imply L^2 convergence of these averages to a limit for any k [**6**]. We are interested in almost everywhere convergence. The case $k = 1$ corresponds to the classical Birkhoff ergodic theorem. Bourgain has proved almost everywhere convergence for the case $k = 2$ [**1**]. Almost everywhere convergence is not known for $k \geq 3$.

Nilsystems arise in the study of the L^2 convergence of the averages (1). A *nilspace* is a homogeneous space of a nilpotent Lie group. A *nilsystem* consists of a finite measure nilspace $(N/\Gamma, m)$ where m is the Haar measure on N/Γ, with a measure-preserving transformation given by rotation by an element of N: $Tx\Gamma = ax\Gamma$ for some $a \in N$. The system $(N/\Gamma, m, T)$ is a *k-step nilsystem* if the group N is a k-step nilpotent group (this notion is defined later). It was shown by Host and Kra [**6**], that k-step nilsystems 'characterize' the L^2 behavior of the averages in (1) (this was shown for $k = 3$ by Furstenberg and Weiss [**4**] and Conze and Lesigne [**2**], and for $k = 4$ by Host and Kra [**5**] and by the present author in her PhD thesis [**17**]). By 'characterize' we mean that the L^2 average can be calculated by projecting the characteristic factor of the average in question on to an appropriate factor. That k-step nilsystems play a role in the convergence of a $(k+1)$-length multiple recurrence average can be expected, since for a k-step nilspace

https://doi.org/10.1017/S0143385703000518 Published online by Cambridge University Press

N/Γ, $a^{k+1}x\Gamma$ is a function of $ax\Gamma, a^2x\Gamma, \ldots, a^kx\Gamma$, and this provides a constraint on the $(k+1)$ tuple $T^nx, \ldots, T^{(k+1)n}x$ in this case. For this reason it is particularly interesting to study these averages on nilsystems. Almost everywhere convergence for nilsystems was proved by Lesigne [**11**] and also follows from results of Shah [**13**].

2. *Statement of results*

We prove an almost everywhere convergence theorem for nilsystems and give an explicit formula for the limit. Let N be a k-step, simply connected nilpotent Lie group, Γ a discrete subgroup such that N/Γ is compact. Let $N_1 = N$; for $i > 1$, let $N_i = [N_{i-1}, N]$ (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 [**12**]). Let m_i be the Haar measure on N_i/Γ_i. Let $a \in N$, and for $x\Gamma \in N/\Gamma$, let $Tx\Gamma = ax\Gamma$. Then the mps $(N/\Gamma, m_1, a)$ is called a *nilsystem*. In [**10**], Lesigne proved the following theorem for two-step nilsystems.

THEOREM 2.1. (Lesigne) *Let $(N/\Gamma, m_1, a)$ be a two-step ergodic nilsystem, where N is connected, and let $f_1, f_2, f_3 \in L^\infty(N/\Gamma)$. Then for almost all $x \in N$,*

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

We explain why this integral makes sense. For a given x, consider the function F defined on $N \times N_2$ by

$$F(y_1, y_2) = f_1(xy_1\Gamma) f_2(xy_1^2y_2\Gamma) f_3(xy_1^3y_2^3\Gamma).$$

This function satisfies $F(y_1, y_2\gamma_2) = F(y_1, y_2)$ for any $\gamma_2 \in \Gamma_2$ and thus can be considered as a function on $N \times N_2/\Gamma_2$. Therefore, the integral

$$H(y_1) = \int F(y_1, y_2)\, dm_2(y_2\Gamma_2)$$

is a well-defined function on N. A direct computation gives $H(y_1) = H(y_1\gamma)$ for $\gamma \in \Gamma$, therefore H is defined on N/Γ, and the integral in (2) exists.

We generalize this result. Let

$$M = \left\{ \left(y_1, y_1^2y_2, \ldots, \prod_{j=1}^{k} y_j^{\binom{k}{j}} \right) : y_1 \in N_1, y_2 \in N_2, \ldots, y_k \in N_k \right\} \subset N^k,$$

where

$$\prod_{j=1}^{k} y_j^{\binom{k}{j}} = y_1^k y_2^{\binom{k}{2}} \ldots y_k^{\binom{k}{k}},$$

and let $\Lambda = M \cap \Gamma^k$. Then M is a subgroup of N^k [**7, 8, 16**]. As N_j are closed subgroups of N (see Theorem 3.1), M is closed and therefore a Lie subgroup of N^k. The group Λ is a discrete subgroup of M, and the nilmanifold $Y = M/\Lambda$ is embedded in $(N/\Gamma)^k$ in a natural way. As $(N/\Gamma)^k$ is compact so is M/Λ. Let ν be the Haar measure on Y.

https://doi.org/10.1017/S0143385703000518 Published online by Cambridge University Press

THEOREM 2.2. *Let $(N/\Gamma, m_1, T)$ be a k-step ergodic nilsystem, and let $f_1, \ldots, f_k \in L^\infty(N/\Gamma)$. If N is spanned by $N_{(0)}$ and a, where $N_{(0)}$ denotes the connected component of N, then for almost all $x \in N$*

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k} f_j(a^j x\Gamma) &= \int_Y \prod_{j=1}^{k} f_j(xz_j)\, d\nu(z_1, \ldots, z_k) \\ &= \int_{N/\Gamma} \left(\cdots \left(\int_{N_k/\Gamma_k} \prod_{j=1}^{k} f_j\left(x \prod_{i=1}^{j} y_i^{\binom{j}{i}} \Gamma \right) dm_k(y_k\Gamma_k) \right) \cdots \right) dm_1(y_1\Gamma_1) \quad (3)\end{aligned}$$

(we explain in §3 why the multiple integral is well defined).

Theorem 2.1 is deducible from Theorem 2.2 applied for $k = 3$. Indeed if N is a two-step nilpotent group then $N_3 = \{1\}$ and

$$M = \{(y_1, y_1^2 y_2, y_1^3 y_2^3) : y_1 \in N_1, y_2 \in N_2\}.$$

The proof is a combination of the proof of Lesigne with the work of Leibman [**8**] on polynomial sequences in groups. Let $x \in N$, and denote by A_x the closure of the orbit of $(x, \ldots, x)$ in $(N/\Gamma)^k$ under the action of $T \times \cdots \times T^k$. As this transformation is translation by $(a, a^2, \ldots, a^k)$, it is known that A_x can be given the structure of a nilmanifold for some closed subgroup M_x of N^k containing $(a, a^2, \ldots, a^k)$. As $T \times \cdots \times T^k$ acts transitively on A_x it is uniquely ergodic. Therefore, for f_i continuous, the average in (3) converges to the integral

$$\int \prod_{j=1}^{k} f_j(x_j)\, d\lambda_x(x_1, \ldots, x_k)$$

where λ_x is the Haar measure on M_x. We give an explicit formula for the measure λ_x and then for the limit.

Remark 2.3. If the nilsystem $(N/\Gamma, a)$ is ergodic then the projection of $\langle N_{(0)}, a\rangle$ on N/Γ is an open invariant set, and therefore is equal to N/Γ. Thus $N/\Gamma = \langle N_{(0)}, a\rangle/(\Gamma \cap \langle N_{(0)}, a\rangle)$. We can therefore assume that $N = \langle N_{(0)}, a\rangle$.

3. *Proof of Theorem 2.2*

We define a set $\tilde{N}$ by $\tilde{N} = N_1 \times \cdots \times N_k$. As M is a group under normal multiplication, and the map

$$(g_1, g_2, \ldots, g_k) \to \left(g_1, g_1^2 g_2, \ldots, \prod_{j=1}^{k} g_j^{\binom{k}{j}} \right)$$

is a bijection between $\tilde{N}$ and M, we can define multiplication in $\tilde{N}$ as follows. For $(x_1, \ldots, x_k), (y_1, \ldots, y_k) \in \tilde{N}$,

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

https://doi.org/10.1017/S0143385703000518 Published online by Cambridge University Press

where

$$\prod_{j=1}^{i} z_j^{\binom{i}{j}} = \prod_{j=1}^{i} x_j^{\binom{i}{j}} \prod_{j=1}^{i} y_j^{\binom{i}{j}} \quad 1 \le i \le k,$$

i.e. $(z_1, z_2, \dots, z_k)$ is the preimage of the coordinate product of

$$\left(x_1, x_1^2 x_2, \dots, \prod_{j=1}^{k} x_j^{\binom{k}{j}}\right), \quad \left(y_1, y_1^2 y_2, \dots, \prod_{j=1}^{k} y_j^{\binom{k}{j}}\right). \tag{4}$$

Let $\tilde{\Gamma} = \Gamma_1 \times \cdots \times \Gamma_k$. Then $(\tilde{\Gamma}, \star)$ is a discrete subgroup of $\tilde{N}$, and as M/Λ is compact, 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

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

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

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

The mapping I_x is one to one if

$$I_x(y_1, \dots, y_k) = I_x(z_1, \dots, z_k),$$

in which case

$$\begin{aligned} I_x((z_1, \dots, z_k)^{-1} \star (y_1, \dots, y_k)) &= I_x((z_1, \dots, z_k)^{-1} \star (z_1, \dots, z_k)) \\ &= (\Gamma, \dots, \Gamma), \end{aligned}$$

which implies that $((z_1, \dots, z_k)^{-1} \star (y_1, \dots, y_k)) \in \tilde{\Gamma}$. Let $(\gamma_1, \dots, \gamma_k) \in \tilde{\Gamma}$, then for all $(y_1, \dots, y_k) \in \tilde{N}$ we have

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

thus I_x is well defined on $\tilde{N}/\tilde{\Gamma}$.

Let $\tilde{m}$ denote the Haar measure on $\tilde{N}/\tilde{\Gamma}$. 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 $(\tilde{N}/\tilde{\Gamma}, \tilde{m}, S_x)$ and $(A_x, m_x, T \times T^2 \times \cdots \times T^k)$:

$$\begin{aligned} I_x(S_x((y_1, \dots, y_k) \star \tilde{\Gamma})) &= I_x((a[a, x], e, \dots, e) \star (y_1, \dots, y_k) \star \tilde{\Gamma}) \\ &= \left(x a[a, x] y_1 \Gamma, x (a[a, x])^2 y_1^2 y_2 \Gamma, \dots, x (a[a, x])^k \prod_{j=1}^{k} y_j^{\binom{k}{j}} \Gamma\right) \\ &= \left(a x y_1 \Gamma, a^2 x y_1^2 y_2 \Gamma, \dots, a^k x \prod_{j=1}^{k} y_j^{\binom{k}{j}} \Gamma, a^k x \prod_{j=1}^{k} y_j^{\binom{k}{j}} \Gamma\right) \\ &= (T \times T^2 \times \cdots \times T^k)(I_x((y_1, \dots, y_k) \star \tilde{\Gamma})). \end{aligned}$$

We now use a generalization of a theorem by Parry [**14**], that was proved in Leibman [**9**].

https://doi.org/10.1017/S0143385703000518 Published online by Cambridge University Press

THEOREM 3.1. (Leibman) *Let $(G/\Gamma, a)$ be a nilsystem. If G is spanned by $G_{(0)}$ and a, where $G_{(0)}$ denotes the connected component of the identity, then:*

(1) *for $i > 1$, G_i is connected, and G_i, $G_i\Gamma$ are closed subgroups of N; the Abelian group $Z = G/G_2\Gamma$ is compact;*
(2) *the nilsystem $(G/\Gamma, a)$ is ergodic if and only if the factor system (Z, a) is ergodic;*
(3) *if $(G/\Gamma, a)$ is ergodic, then it is uniquely ergodic.*

We apply this theorem to $(\tilde{N}/\tilde{\Gamma}, S_x)$. We first check that $\tilde{N}$ is spanned by the connected component of the identity $\tilde{N}_{(0)}$, and S_x. By Theorem 3.1 applied to the group N, the groups N_j are connected for $j > 1$, therefore $\tilde{N}_{(0)} = N_{(0)} \times N_2 \times \cdots \times N_k$. The action of S_x on the first coordinate is by translation by the element $a[a, x]$. Since $N_{(0)}$ is a normal subgroup of N

$$\bigcup_{n\in\mathbb{Z}} (a[a,x])^n N_{(0)} = \bigcup_{n\in\mathbb{Z}} x^{-1}a^n x N_{(0)} = x^{-1}\left(\bigcup_{n\in\mathbb{Z}} a^n N_{(0)}\right)x = N.$$

By Theorem 3.1, if the action of S_x is ergodic then it is uniquely ergodic, and thus the mps $(A_x, m_x, T \times T^2 \times \cdots \times T^k)$ is uniquely ergodic. For each x such that S_x is ergodic we have

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} F\left(a^n x y_1\Gamma, a^{2n} x y_1^2 y_2\Gamma, \ldots, a^{(k)n} x \prod_{j=1}^{k} y_j^{\binom{k}{j}}\Gamma\right)$$
$$= \int_Y F(xz_1, xz_2, \ldots, xz_k)\, d\nu(z_1, \ldots, z_k)$$

for all $(y_1, \ldots, y_k) \in \tilde{N}$ and for all continuous functions F on $(N/\Gamma)^k$. In particular, for $F = f_1 \otimes \cdots \otimes f_k$ where f_i are continuous functions, and $(y_1, \ldots, y_k) = (e, \ldots, e)$,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f_1(a^n x\Gamma)\cdots f_k(a^{(k)n}x\Gamma) = \int_Y f_1(xz_1)\cdots f_k(xz_k)\, d\nu(z_1, \ldots, z_k). \quad (5)$$

We want to show that this is equal to the multi-integral in (3). We first explain why the multi-integral in (3) is defined: this is done inductively. For a given x, consider the function F defined on $N \times N_2 \times \cdots \times N_k$ by

$$F(y_1, \ldots, y_k) = \prod_{j=1}^{k} f_j\left(x \prod_{i=1}^{j} y_i^{\binom{j}{i}}\Gamma\right).$$

Assume

$$H_l(y_1, \ldots, y_l)$$
$$= \int_{N_{l+1}/\Gamma_{l+1}} \left(\cdots\left(\int_{N_k/\Gamma_k} F(y_1, \ldots, y_k)\, dm_k(y_k\Gamma_k)\right)\cdots\right) dm_{l+1}(y_{l+1}\Gamma_{l+1})$$

is a well-defined function on $N \times N_2 \times \cdots \times N_{l-1} \times N_l/\Gamma_l$. Define a function on $N \times N_2 \times \cdots \times N_{l-1}$ by

$$H_{l-1}(y_1, \ldots, y_{l-1}) = \int_{N_l/\Gamma_l} H_l(y_1, \ldots, y_l)\, dm_l(y_l\Gamma_l).$$

https://doi.org/10.1017/S0143385703000518 Published online by Cambridge University Press

A direct computation shows that

$$H_{l-1}(y_1, \ldots, y_{l-1}) = H_{l-1}(y_1, \ldots, y_{l-1}\gamma_{l-1})$$

for $\gamma_{l-1} \in \Gamma_{l-1}$, thus H_{l-1} is well defined on $N \times N_2 \times \ldots \times N_{l-2} \times N_{l-1}/\Gamma_{l-1}$. Now let $x \in N/\Gamma$, and let $f_1, \ldots f_k$ be continuous functions on N/Γ. Define a measure ν_x on Y by

$$\int \prod_{j=1}^{k} f_j(z_j)\, d\nu_x(z_1, \ldots, z_k)$$
$$= \int_{N/\Gamma} \left(\cdots \left(\int_{N_k/\Gamma_k} \prod_{j=1}^{k} f_j\left(x \prod_{i=1}^{j} y_i^{(j_i)} \Gamma \right) dm_k(y_k\Gamma_k) \right) \cdots \right) dm_1(y_1\Gamma_1).$$

A direct computation shows that the measure ν_x is invariant under translation by elements in M, and is therefore equal to ν.

This proves the result for $f_1, \ldots, f_k$ continuous functions. For passing to bounded measurable functions use the maximal inequality in [**2**, Lemma 3, p. 173]. In order to finish the proof of the theorem, we must show that the action of S_x is ergodic for almost all (Haar) $x \in N$.

4. *Ergodicity of S_x*

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 $\tilde{N}_2$ denote the commutator subgroup of $\tilde{N}$. By Theorem 3.1 it is enough to show the ergodicity of the action of S_x on $\tilde{N}/\tilde{N}_2 \star \tilde{\Gamma}$. The transformation S_x is ergodic if and only if for all non-trivial characters σ of $\tilde{N}$ (continuous homomorphisms $\tilde{N} \to S^1$) that vanish on $\tilde{N}_2 \star \tilde{\Gamma}$ we have $\sigma(a[a, x], e, \ldots, e) \neq 1$. Let m be the Haar measure on N. As there is only a countable number of such characters ($\tilde{N}/\tilde{N}_2 \star \tilde{\Gamma}$ is a separable Abelian group and thus its dual group is countable), it is enough to show that if $\sigma \not\equiv 1$ then

$$m\{x \in N : \sigma((a[a, x], e, \ldots, e)) = 1\} = 0.$$

Let

$$\psi(x) := \sigma((a[a, x], e, \ldots, e)).$$

For $1 \leq i \leq k$ we define the functions σ_i on N_i by

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

Since

$$(y_1, \ldots, y_k) = (y_1, e, \ldots, e) \star (e, y_2, e, \ldots, e) \star \cdots \star (e, \ldots, e, y_k),$$

we have

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

As $\psi(x)$ is an analytic function on N, if $\psi(x) = 1$ on a set of positive measure in N, then $\psi(x) \equiv 1$ on some connected component of N. As $\psi(ax) = \psi(x)$ for every $x \in N$ and N

https://doi.org/10.1017/S0143385703000518 Published online by Cambridge University Press

is spanned by $N_{(0)}$ and a, it follows that $\psi(x) \equiv 1$. We will show that if $\psi(x) \equiv 1$, then $\sigma_i \equiv 1$ for $1 \le i \le k$, and thus $\sigma \equiv 1$.

Without loss of generality we assume $k \ge 2$. For $n \in \mathbb{N}$ and $y_1, \dots, y_n \in N$ we define $R(y_1, \dots, y_n)$ as follows:

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

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 $R_{k-1} = R_k$).

Example 4.1. For $k = 3, n = 2, x, y \in N$

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

LEMMA 4.2. *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 (4) 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 $i < j \le k$. □

Example 4.3. For $k = 2$ and $x, y \in N$

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

Assume $\psi \equiv 1$. Our aim is to show that σ is trivial. Using Lemma 4.2 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 $R(a, [a, x])$.

Setting $x = e$ we get $\sigma(a) = 1$, thus

$$\begin{aligned} \sigma_1([a, x]) &= \sigma_2(r_2(x))^{-1} \cdots \sigma_k^{-1}(r_k(x)) \\ &= \prod_{i=2}^{k} \prod_{m=1}^{l_i} \sigma_i(r_{i_m}(x)), \end{aligned} \tag{6}$$

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).

https://doi.org/10.1017/S0143385703000518 Published online by Cambridge University Press

Example 4.4. Before proceeding with the general case, which involves a complicated induction, we give the argument for the case $k = 3$ (the difficulties arising in the proof for large k do not show up for $k = 2$, which was treated in Lesigne [**10**]). Let N be a three-step simply connected nilpotent Lie group and let σ be a character of $\tilde{N}$ that vanishes on $\tilde{N}_2 \star \tilde{\Gamma}$. In this case, for all x in N,

$$\sigma_1([a, x]) = \sigma_2([a, [a, x]])^{-1}. \tag{7}$$

If $x \in N_2$ then $[a, [a, x]] \in N_4 = \{1\}$ thus by (7) for $x \in N_2$, $\sigma_1([a, x]) = 1$. For $x \in N_2$ compute

$$(a^{-1}, e, e) \star (x^{-1}, e, e) \star (a, e, e) \star (x, e, e) = (z_1, z_2, z_3). \tag{8}$$

By the definition of multiplication in $\tilde{N}$, (z_1, z_2, z_3) represents the vector

$$(a^{-1}x^{-1}ax, a^{-2}x^{-2}a^2x^2, a^{-3}x^{-3}a^3x^3),$$

which, since $[[a, x], x] = 1$ and $[[a, x], a] = 1$, equals

$$([a, x], [a, x]^2[a, x]^2, [a, x]^3[a, x]^6).$$

Thus $z_1 = [a, x]$, $z_2 = [a, x]^2$, $z_3 = e$. As σ is trivial on $\tilde{N}_2$ we get, for $x \in N_2$, $\sigma_1([a, x])\sigma_2([a, x]^2) = 1$ thus $\sigma_2([a, x])^2 = 1$, and as N_2 is connected $\sigma_2([a, x]) = 1$. By a similar computation to that in (8) we get from

$$(a^{-1}, e, e) \star (e, x^{-1}, e) \star (a, e, e) \star (e, x, e) = (e, [a, x]^2, [a, x]^3) \tag{9}$$

that $\sigma_3([a, x]) = 1$.

Now, if $x \in N$, then $[a, x] \in N_2$, thus by (7), $\sigma_1([a, x]) = 1$. Using the fact that $[x, yz] = [x, y][[x, y], z][x, z]$ we get

$$\sigma_1([a, xy]) = \sigma_1([a, x][[a, x], y][a, y]) = \sigma_1([a, x])\sigma_1([[a, x], y])\sigma_1([a, y])$$

(commutators involving two or more of the elements $[a, x]$, $[[a, x], y]$, $[a, y]$ are trivial as $N_4 = 1$), thus $\sigma_1([[a, x], y]) = 1$. Repeat the computation in (8) replacing a with $[a, x]$ to get that $\sigma_2([[a, x], y]) = 1$. Repeat the computation in (9) replacing a with y and x with $[a, x]$. We get $\sigma_2^2([[a, x], y])\sigma_3^3([[a, x], y]) = 1$, and as N_3 is connected $\sigma_3([[a, x], y]) = 1$. The set $\{\{a^n\Gamma\}_{n\in\mathbb{Z}}\}$ is dense in N, and σ_3 is trivial on Γ_3. Since $\sigma_3([[a, x], y]) = 1$, and $\sigma_3([a, [x, y]]) = 1$, $\sigma_3 \equiv 1$. Now repeat the computation in (8) for $x \in N$ (not N_2). As

$$\sigma_2([[a, x], x]) = \sigma_2([[a, x], a]) = \sigma_3([[a, x], x]) = \sigma_3([[a, x], a]) = 1,$$

we get $\sigma_1([a, x])\sigma_2^2([a, x]) = 1$, which implies that $\sigma_2([a, x]) = 1$. Since σ_2 is trivial on Γ_2, $\sigma_2 \equiv 1$. This means that σ_1 is a character. However, $\sigma_1(a) = 1$ and $\sigma_1(\Gamma) = 1$, thus $\sigma \equiv 1$.

We need the following lemma, which is based on the fact that σ is multiplicative and thus $\sigma \equiv 1$ on $\tilde{N}_2$.

https://doi.org/10.1017/S0143385703000518 Published online by Cambridge University Press

LEMMA 4.5. *Let $x \in N$, $y \in N_n$. We define*

$$H(x, y) = \bigcup_{i=2}^{k-1} R_i(x, y) \subset N_{n+2}.$$

If $\sigma_i(h) = 1$ for all $i > 1$, and $h \in H(x, y) \cap N_i$, 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 4.2 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 $\sigma|_{\tilde{N}_2} \equiv 1$, we have

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

We calculate the z_i:

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

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

$$z_m^{(m+1)} z_{m+1} = [x^{(m+1)}, y^{(m+1)}] = [x, y]^{(m+1)^2} \tilde{h_1'},$$

which implies

$$z_{m+1} = [x, y]^{(m+1)} \tilde{h_1}$$

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 $j = 2$,

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

which implies

$$z_{m+2} = [x, y]^{-m\binom{m+2}{2} - (m+1)(m+2) + (m+2)\binom{m+2}{2}} \tilde{h_2} = \tilde{h_2}$$

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

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

which implies

$$z_{m+j} = [x, y]^{-m\binom{m+j}{j} - (m+1)\binom{m+j}{j-1} + (m+j)\binom{m+j}{j}} \tilde{h_j} = \tilde{h_j}$$

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

$$\begin{aligned} 1 &= \prod_{i=1}^{k} \sigma_i(z_i) = \sigma_m([x, y]^m \tilde{h_0}) \sigma_{m+1}([x, y]^{m+1} \tilde{h_1}) \prod_{i=m+2}^{k} \sigma_i(\tilde{h}_{i-m}) \\ &= \sigma_m([x, y]^m) \sigma_{m+1}([x, y]^{m+1}), \end{aligned}$$

as $\tilde{H}([x, y], \tilde{h_0}), \tilde{H}([x, y], \tilde{h_1}) \subset \tilde{H}(x, y)$. □

https://doi.org/10.1017/S0143385703000518 Published online by Cambridge University Press

COROLLARY 4.6. *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 *such that* $[x,y] \in N_i$.

Proof. By Lemma 4.5, $\sigma_i([x,y])^i = 1$. By §4 N_i is connected for $i > 1$, and $\sigma_i([e,y]) = 1$. □

Notation 4.7. We use the notation $[x_1, \dots, x_n]$ for

$$[\dots[[[x_1,x_2],x_3],x_4],\dots,x_n].$$

Notation 4.8. 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

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

Example 4.9. If $A = (1,3)$, then

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

Notation 4.10.

(1) For $A = (s_1,\dots,s_j) \in \mathcal{S}$, $|A| = \sum s_j$.
(2) For $A \in \mathcal{S}$, $(i : A)$ means σ_i is trivial on $(C_A \cap N_i)$.
(3) $j^{\ge}$ $(>)$ is some $l \ge j$ $(l > j)$.
(4) Let X be j, $j^{\ge}$ or $j^{>}$, then $X+$ is a sequence of length ≥ 1 whose entries are of type j, $j^{\ge}$ or $j^{>}$, respectively.
(5) For $x \in N_j$, $|x| = j$.
(6) If $A_1 = (s_1,\dots,s_{j_1})$, $A_2 = (t_1,\dots,t_{j_2}) \in \mathcal{S}$, then

$$A_1 * A_2 := (s_1,\dots,s_{j_1},t_1,\dots,t_{j_2}).$$

PROPOSITION 4.11. *For all* $A \in \mathcal{S}$, $1 \le i \le k$, *we have* $(i : A)$.

Proof. Let $A = (s_1,\dots,s_j)$, $0 \le j \le k$, $x_n \in N_{s_n}$. We use the following induction rules.

LEMMA 4.12. (Triviality rule) *If* $|A| \ge k$, *then for all* i, $(i : A)$.

Proof. $|A| \ge k$ implies that $[a,x_1,\dots,x_j] = 1$, thus $\sigma_i([a,x_1,\dots,x_j]) = 1$. □

LEMMA 4.13. (Reducing rule) *If for all* i, $(i : (j^{>}+))$, *then* $(1 : (j))$.

Proof. Let $x \in N_j$. We use (6):

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

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 $m \ge 1$, $|y_n| \ge |b| > |x|$, and is thus of type $(j^{>}+)$. □

https://doi.org/10.1017/S0143385703000518 Published online by Cambridge University Press

LEMMA 4.14. (Expanding rule) *If the following hold:*
(1) $(1 : A * (l))$*;*
(2) $j > l$*;*
(3) *for all* i $(i : A * (l, j^{\geq}+))$*;*
(4) *for all* i $(i : A * (j-1, j^{\geq}+))$*;*
then $(1 : A * (l, j-1))$ *and* $(1 : A * (j-1, l))$.

Proof. Let $v = [a, x_1, \ldots, x_j]$, $y \in N_l$, $z \in N_{j-1}$, 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 without loss of generality

$$b_{i_m}(y, z, v) = [v, y, y_1, \ldots, y_m],$$

where $m \geq 1$ and $|y_n| \geq j$ for $n = 1, \ldots, m$. One obtains this by looking 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 without loss of generality

$$b_{i_m}(y, z, v) = [v, z, z_1, \ldots, z_m],$$

where $m \geq 1$ and $|z_n| \geq j$ for $n = 1, \ldots, m$. This is of type $A * (j-1, j^{\geq}+)$. For $A * (j-1, l)$, repeat the argument exchanging the roles of y and z. □

LEMMA 4.15. *If the following hold:*
(1) $(1 : A * (j))$*;*
(2) *for all* i $(i : A * (j^{\geq}, j^{\geq}+))$*;*
(3) *for all* i $(i : A * (j^{>}))$*;*
then for all i $(i : A * (j))$.

Proof. Let v be of type A, $x \in N_j$. By Lemma 4.5, we must show that for $h \in H(v, x) \cap N_i$, $\sigma_i(h) = 1$. However, 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 $A * (j^{\geq}, j^{\geq}+)$. □

PROPOSITION 4.16. *For* $|A| \geq 0$, $l \geq 1$ *and every* i, $(1 : A * (l))$ *implies* $(i : A * (l^{\geq}+))$.

Proof. We prove this by downward induction on $|A|$, and for $|A|$ given by downward induction on l. For $|A| = k$ and any l the result is clear (as the commutator is trivial). Suppose the statement is true for $|A| > d$, and any l. For $|A| = d$, $l = k - d$ it is clear (as the commutator is trivial). Assume that $(1 : A * (j))$ implies $(i : A * (j^{\geq}+))$ for $|A| = d$, any $j > l$. We follow the scheme shown in Figure 1.

https://doi.org/10.1017/S0143385703000518 Published online by Cambridge University Press

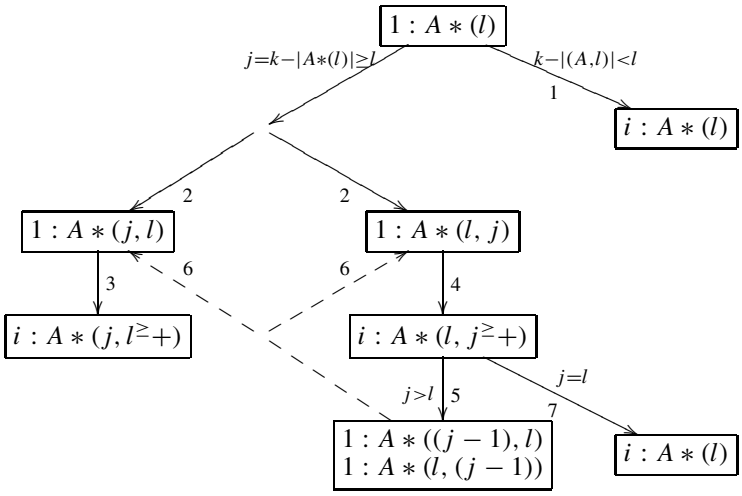


FIGURE 1.

Explanation for the numbered arrows in Figure 1.

1. If $k - |A * (l)| < l$, then $|A * (l^{\geq}, l^{\geq}+)| > k$ and therefore $(i : A * (l^{\geq}, l^{\geq}+))$ by the triviality rule (Lemma 4.12). If $x \in N_j$ and $j > l$, then $x \in N_l$. Therefore, $(1 : A*(l))$ implies $(1 : A*(j))$ for $j > l$. This implies, by the induction hypothesis, that $(i : A * (j))$ for $j > l$, and by Lemma 4.15 $(i : A * (l))$.
2. As $|A|+l+(j+1) > k$ and $|A|+j+(j+1) > k$, by the triviality rule (Lemma 4.12) we have $(i : A*(j, (j+1)^{\geq}+))$ and $(i : A*(l, (j+1)^{\geq}+))$. Then use the expanding rule (Lemma 4.14).
3. This follows from the induction hypothesis by substituting $A*(j)$ for A $(|A*(j)| > |A|)$.
4. This follows from the induction hypothesis by substituting $A * (l)$ for A $(|A * (l)| > |A|)$.
5. By the expanding rule (Lemma 4.14), the only condition we need to verify is $(i : A*((j-1), j^{\geq}+))$. If $j-1 = l$, then $(i : A*((j-1), j^{\geq}+)) = (i : A*(l, j^{\geq}+))$ and there is no problem. If $j - 1 > l$, then $(1 : A * (l))$ implies $(1 : A * (j - 1))$, which implies that $(i : A * ((j - 1)^{\geq}+))$ by the induction hypothesis.
6. Repeat steps 3, 4, 5 replacing j with $j - 1$, until $j = l$.
7. When $j = l$ we have covered all the cases $(i : A * (l^{\geq}, l^{\geq}+))$. Now repeat the argument in step 1 to get $(i : A * (l))$.

Proof of Proposition 4.11. We prove this by induction on j: we have $(1 : k)$ by the triviality rule (Lemma 4.12). $(1 : (j)) \Rightarrow \forall i$ $(i : (j^{\geq}+))$ by Proposition 4.16 $(|A| = 0)$. By Lemma 4.13 this implies $(1 : (j - 1))$. □

PROPOSITION 4.17. *For all i, $\sigma_i \equiv 1$.*

https://doi.org/10.1017/S0143385703000518 Published online by Cambridge University Press

Proof. We prove this by induction on k. By Proposition 4.11, we have

$$\sigma_k([x_1, \ldots, \overset{m}{a}, \ldots, x_{k-1}]) = 1 \quad \forall x_1, \ldots, x_{k-1} \in N, \quad 1 \le m \le k. \tag{10}$$

Since $\sigma|_{\tilde{\Gamma}} \equiv 1$,

$$\sigma_k([\gamma_1, \ldots, \gamma_k]) = 1 \quad \forall \gamma_1, \ldots, \gamma_k \in \Gamma.$$

From (10)

$$\sigma_k([a, \gamma_2, \ldots, \gamma_k]) = 1 \quad \forall \gamma_2, \ldots, \gamma_k \in \Gamma.$$

Since σ_k is a character on N_k and a acts ergodically on N/Γ, we have

$$\sigma_k([x_1, \gamma_2, \ldots, \gamma_k]) = 1 \quad \forall \gamma_2, \ldots, \gamma_k \in \Gamma, \quad x_1 \in N.$$

Using (10) and induction, we get

$$\sigma_k([x_1, \ldots, x_k]) = 1 \quad \forall x_1, \ldots, x_k \in N,$$

thus $\sigma_k \equiv 1$. Now $\sigma_k, \ldots, \sigma_{j+1} \equiv 1$ implies that σ_j is a character. For $j > 1$, by Proposition 4.11, we have

$$\sigma_j([x_1, \ldots, \overset{m}{a}, \ldots, x_{j-1}]) = 1 \quad \forall x_1, \ldots, x_{j-1} \in N, \quad 1 \le m \le j,$$

and by the same argument as above, we have $\sigma_j \equiv 1$. For $j = 1$, σ_1 is a character that satisfies

$$\sigma_1(a) = 1, \quad \sigma_1|_{\Gamma} \equiv 1,$$

thus $\sigma_1 \equiv 1$. □

Acknowledgements. The author thanks David Lehavi, Yoav Yaffe, Alexander Leibman, Vitaly Bergelson, Nikos Frantzikinakis, and the referee for valuable remarks and for helping her make the proofs more readable. This paper is a part of the author's PhD thesis. She thanks her adviser Professor Hillel Furstenberg for introducing her to ergodic theory, specifically to questions concerning non-conventional ergodic theorems, and for fruitful discussions.

References

[1] J. Bourgain. Pointwise ergodic theorems for arithmetic sets. *Publ. Math. Inst. Hautes Ètudes Sci.* **69** (1989), 5–45.

[2] J. P. Conze and E. Lesigne. Théorèmes ergodiques pour des mesures diagonales. *Bull. Soc. Math. France* **112**(2) (1984), 143–175.

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

[4] H. Furstenberg and B. Weiss. 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** (1996), 193–227.

[5] B. Host and B. Kra. Personal communication.

[6] B. Host and B. Kra. *Non-conventional ergodic averages and nilmanifolds*. Ann. Math. **161**(1) (2005), 397–488.

https://doi.org/10.1017/S0143385703000518 Published online by Cambridge University Press

[7] M. Lazard. Sur certaines suites d'éléments dans les groupes libres et leurs extensions. *C. R. Acad. Sci. Paris* **236** (1953), 36–38.
[8] A. Leibman. Polynomial sequences in groups. *J. Algebra* **201** (1998), 189–206.
[9] A. Leibman. Pointwise convergence of ergodic averages for polynomial sequences of rotations of a nilmanifold. *Ergod. Th. & Dynam. Sys.* **25**(1) (2005), 201–213.
[10] E. Lesigne. Théorèmes ergodiques pour une translation sur une nilvariété. *Ergod. Th. & Dynam. Sys.* **9**(1) (1989), 115–126.
[11] E. Lesigne. Sur une nil-variété, les parties minimales associées à une translation sont uniquement ergodiques. *Ergod. Th. & Dynam. Sys.* **11**(2) (1991), 379–391.
[12] A. I. Malcev. On a class of homogeneous spaces. *Amer. Math. Soc. Transl.* **39** (1951).
[13] N. Shah. Invariant measures and orbit closures on homogeneous spaces for actions of subgroups. *Lie groups and Ergodic Theory (Mumbai, 1996)*. Tata Institute of Fundamental Research, Bombay, 1998, 229–271.
[14] W. Parry. Ergodic properties of affine transformations and flows on nilmanifolds. *Amer. J. Math.* **91** (1969), 757–771.
[15] W. Parry. Dynamical systems on nilmanifolds. *Bull. London Math. Soc.* **2** (1970), 37–40.
[16] J. Petresco. Sur les commutateurs. *Math. Z.* **61** (1954), 348–356.
[17] T. Ziegler. Non-conventional ergodic averages. *PhD Thesis*, 2002.

https://doi.org/10.1017/S0143385703000518 Published online by Cambridge University Press