Ergod. Th. & Dynam. Sys. (2016), **36**, 2044–2075 © *Cambridge University Press, 2015*
doi:10.1017/etds.2015.11

Joint ergodicity along generalized linear functions

V. BERGELSON†, A. LEIBMAN† and Y. SON‡

† *Department of Mathematics, The Ohio State University, Columbus, OH 43210, USA*
(e-mail: bergelson.1@osu.edu, leibman.1@osu.edu)

‡ *Faculty of Mathematics and Computer Science, The Weizmann Institute of Science, 234 Herzl Street, Rehovot 7610001, Israel*
(e-mail: younghwan.son@weizmann.ac.il)

(*Received* 8 *August* 2014 *and accepted in revised form* 21 *January* 2015)

Abstract. A criterion of joint ergodicity of several sequences of transformations of a probability measure space X of the form $T_i^{\varphi_i(n)}$ is given for the case where T_i are commuting measure-preserving transformations of X and φ_i are integer-valued generalized linear functions, that is, the functions formed from conventional linear functions by an iterated use of addition, multiplication by constants, and the greatest integer function. We also establish a similar criterion for joint ergodicity of families of transformations depending on a continuous parameter, as well as a condition of joint ergodicity of sequences $T_i^{\varphi_i(n)}$ along primes.

1. *Introduction*

Let $(X, \mathcal{B}, \mu)$ be a probability measure space. A measure-preserving transformation $T : X \longrightarrow X$ is said to be *weakly mixing* if the transformation $T \times T$, acting on the Cartesian square $X \times X$, is ergodic. The notion of weak mixing was introduced in [**vNK**] (for measure-preserving flows) and has numerous equivalent forms (see, for example, [**BeR**, **BeG**]). The following result involving weak mixing plays a critical role in Furstenberg's proof [**Fu**] of the ergodic Szemerédi theorem and forms a natural starting point for numerous further developments (see [**Be**, **BeL1**, **BeMc**, **BeH**]).

THEOREM 1.1. *If T is an invertible weakly mixing measure-preserving transformation of X, then for any $k \in \mathbb{N}$ and any $A_0, A_1, \ldots, A_k \in \mathcal{B}$ one has*

$$\lim_{N\longrightarrow\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A_0 \cap T^{-n} A_1 \cap \cdots \cap T^{-kn} A_k) = \prod_{i=0}^{k} \mu(A_i).$$

It is not hard to show that Theorem 1.1 has the following functional form. (In accordance with the well-established tradition, we write Tf for the function $f(Tx)$.)

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

THEOREM 1.2. *If T is an invertible weakly mixing measure-preserving transformation of X, then for any $k \in \mathbb{N}$, any distinct non-zero integers $a_1, \ldots, a_k$, and any $f_1, \ldots, f_k \in L^\infty(X)$ one has*

$$\lim_{N\longrightarrow\infty} \frac{1}{N} \sum_{n=1}^{N} T^{a_1 n} f_1 \cdots T^{a_k n} f_k = \prod_{i=1}^{k} \int_X f_i \, d\mu$$

in L^2 norm.

In other words, given a weakly mixing transformation T of X and distinct non-zero integers $a_1, \ldots, a_k$, the transformations $T^{a_1}, \ldots, T^{a_k}$ (or, rather, the sequences $T^{a_1 n}, \ldots, T^{a_k n}, n \in \mathbb{N}$) possess a strong joint independence property. This naturally leads to the following definition.

Definition. (cf. [**BBe1**]) Measure-preserving transformations $T_1, \ldots, T_k$ of a probability measure space X are said to be *jointly ergodic* if for any $f_1, \ldots, f_k \in L^\infty(X)$ one has

$$\lim_{N\longrightarrow\infty} \frac{1}{N} \sum_{n=1}^{N} T_1^n f_1 \cdots T_k^n f_k = \prod_{i=1}^{k} \int_X f_i \, d\mu$$

in L^2 norm.

The following theorem, proved in [**BBe1**], provides a criterion of joint ergodicity of commuting measure-preserving transformations.

THEOREM 1.3. *Let $T_1, \ldots, T_k$ be commuting invertible measure-preserving transformations of X. Then $T_1, \ldots, T_k$ are jointly ergodic if and only if the transformation $T_1 \times \cdots \times T_k$ of X^k is ergodic and the transformations $T_i^{-1} T_j$ of X are ergodic for all $i \neq j$.*

Further developments (most of which were motivated by connections with combinatorics and number theory) have revealed that the phenomenon of joint ergodicity is a rather general one. For example, as was shown in [**Be**], if T is an invertible weakly mixing measure-preserving transformation and $p_1, \ldots, p_k$ are non-constant polynomials $\mathbb{Z} \longrightarrow \mathbb{Z}$ with $p_i - p_j \not\equiv$ constant for any $i \neq j$, then for any $f_1, \ldots, f_k \in L^\infty(X)$ one has

$$\lim_{N\longrightarrow\infty} \frac{1}{N} \sum_{n=1}^{N} T^{p_1(n)} f_1 \cdots T^{p_k(n)} f_k = \prod_{i=1}^{k} \int_X f_i \, d\mu$$

in L^2 norm (see also [**FK**, **BeH**, **F**] for more results of this flavor). So, it makes sense to consider ergodicity and joint ergodicity of sequences of measure-preserving transformations of general form.

Definition. Let $\mathcal{T}(n)$, $n \in \mathbb{N}$, be a sequence of measure-preserving transformations of X; we say that $\mathcal{T}$ is *ergodic* if for any $f \in L^2(X)$

$$\lim_{N\longrightarrow\infty} \frac{1}{N} \sum_{n=1}^{N} \mathcal{T}(n) f = \int_X f \, d\mu.$$

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

Given several sequences $\mathcal{T}_1(n), \ldots, \mathcal{T}_k(n)$, $n \in \mathbb{N}$, of measure-preserving transformations of X, we say that $\mathcal{T}_1, \ldots, \mathcal{T}_k$ are *jointly ergodic* if

$$\lim_{N\longrightarrow\infty} \frac{1}{N} \sum_{n=1}^{N} \mathcal{T}_1(n) f_1 \cdots \mathcal{T}_k(n) f_k = \prod_{i=1}^{k} \int_X f_i \, d\mu$$

in L^2 norm for any $f_1, \ldots, f_k \in L^\infty(X)$.

Results obtained in [**Be**, **BeH**, **F**] lead to a natural question of what are the necessary and sufficient conditions for joint ergodicity of sequences of transformations of the form $T_1^{\varphi_1(n)}, \ldots, T_k^{\varphi_k(n)}$, where T_i are measure-preserving transformations of X and $\varphi_i(n)$ are 'sufficiently regular' sequences of integers diverging to infinity. In the case where $T_1 = \cdots = T_k = T$, where T is a weakly mixing transformation, this question has a quite satisfactory answer not only when φ_i are integer-valued polynomials, but also, more generally, are functions of the form $[\psi_i]$, where $[\cdot]$ denotes the integer part and ψ_i are either the so-called 'tempered functions' or functions of polynomial growth belonging to a Hardy field (see [**BeH**, **F**]).

Much less is known about joint ergodicity of $T_1^{\varphi_1(n)}, \ldots, T_k^{\varphi_k(n)}$ when T_i are distinct, not necessarily weakly mixing, transformations. It is our goal in this paper to extend Theorem 1.3 to the case where φ_i are integer-valued *generalized linear functions*. A generalized or *bracket* linear function (of real or integer argument) is a function constructible from conventional linear functions with the help of the operations of addition, multiplication by constants, and taking the integer part, $[\cdot]$ (or, equivalently, the fractional part, $\{\cdot\}$). (For example, $\varphi(n) = [\alpha_1 n + \alpha_2]$, $\varphi(n) = \alpha_1[\alpha_2 n + \alpha_3] + \alpha_4$, and, say, $\varphi(n) = \alpha_1[\alpha_2[\alpha_3[\alpha_4 n + \alpha_5] + \alpha_6] + \alpha_7[\alpha_8 n + \alpha_9]] + \alpha_{10} n + \alpha_{11}$, where $\alpha_i \in \mathbb{R}$, are generalized linear functions.) In complete analogy with Theorem 1.3, we have the following result.

THEOREM 1.4. (Special case of Theorem 1.7) *Let $T_1, \ldots, T_k$ be commuting invertible measure-preserving transformations of X and let $\varphi_1, \ldots, \varphi_k$ be generalized linear functions $\mathbb{Z} \longrightarrow \mathbb{Z}$. The sequences $T_1^{\varphi_1(n)}, \ldots, T_k^{\varphi_k(n)}$ are jointly ergodic if and only if the sequence $T_1^{\varphi_1(n)} \times \cdots \times T_k^{\varphi_k(n)}$ of transformations of X^k is ergodic and the sequences $T_i^{-\varphi_i(n)} T_j^{\varphi_j(n)}$ of transformations of X are ergodic for all $i \neq j$.*

Here are two corollaries of Theorem 1.4.

COROLLARY 1.5. (Cf. Corollary 6.7 below) *Let T be a weakly mixing invertible measure-preserving transformation of X and let $\varphi_1, \ldots, \varphi_k$ be unbounded generalized linear functions $\mathbb{Z} \longrightarrow \mathbb{Z}$ such that $\varphi_j - \varphi_i$ are unbounded for all $i \neq j$. Then, for any $f_1, \ldots, f_k \in L^\infty(X)$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T^{\varphi_1(n)} f_1 \cdots T^{\varphi_k(n)} f_k = \prod_{i=1}^{k} \int_X f_i \, d\mu.$$

In particular, for any distinct $\alpha_1, \ldots, \alpha_k \in \mathbb{R}\backslash\{0\}$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T^{[\alpha_1 n]} f_1 \cdots T^{[\alpha_k n]} f_k = \prod_{i=1}^{k} \int f_i \, d\mu.$$

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

For a measure-preserving transformation T of X, let $\operatorname{Eig} T$ be the set of eigenvalues of T,

$$\operatorname{Eig} T = \{\lambda \in \mathbb{C}^* : Tf = \lambda f \text{ for some } f \in L^2(X)\}.$$

For several measure-preserving transformations $T_1, \dots, T_k$ of X, we put $\operatorname{Eig}(T_1, \dots, T_k) = \prod_{i=1}^k \operatorname{Eig} T_i$.

COROLLARY 1.6. (Cf. Corollary 6.8) *Let $T_1, \dots, T_k$ be commuting invertible jointly ergodic measure-preserving transformations of X and let φ be an unbounded generalized linear function $\mathbb{Z} \longrightarrow \mathbb{Z}$. Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T_1^{\varphi(n)} f_1 \cdots T_k^{\varphi(n)} f_k = \prod_{i=1}^{k} \int f_i \, d\mu \quad \text{for any } f_1, \dots, f_k \in L^\infty(X)$$

if and only if $\lim_{N\to\infty} (1/N) \sum_{n=1}^{N} \lambda^{\varphi(n)} = 0$ *for every* $\lambda \in \operatorname{Eig}(T_1, \dots, T_k) \setminus \{1\}$. *In particular, for any irrational $\alpha \in \mathbb{R}$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T_1^{[\alpha n]} f_1 \cdots T_k^{[\alpha n]} f_k = \prod_{i=1}^{k} \int f_i \, d\mu \quad \text{for any } f_1, \dots, f_k \in L^\infty(X)$$

if and only if $e^{2\pi i \alpha^{-1}\mathbb{Q}} \cap \operatorname{Eig}(T_1, \dots, T_k) = \{1\}$.

In fact, we obtain a result more general than Theorem 1.4. Let G be a commutative group of measure-preserving transformations of X. We say that a sequence $\mathcal{T}$ of transformations of X is a *generalized linear sequence* in G if it has the form $\mathcal{T}(n) = T_1^{\varphi_1(n)} \cdots T_r^{\varphi_r(n)}$, $n \in \mathbb{Z}$, for some $T_1, \dots, T_r \in G$ and generalized linear functions $\varphi_1, \dots, \varphi_k : \mathbb{Z} \longrightarrow \mathbb{Z}$. (The sequences $T_i^{-\varphi_i(n)} T_j^{\varphi_j(n)}$ appearing in Theorem 1.4 are of this sort.) Also, we change the definitions of ergodicity and of joint ergodicity above (see Definition 6.3), replacing the averages $(1/N) \sum_{n=1}^{N}$ with the more general averages $(1/|\Phi_N|) \sum_{n\in\Phi_N}$, where (Φ_N) is an arbitrary Følner sequence in $\mathbb{Z}$. (The *uniform* ergodicity and joint ergodicity, which appear when the averages $(1/N) \sum_{n=1}^{N}$, with $N \to \infty$, are replaced by the averages $(1/(M-N)) \sum_{n=N+1}^{M}$, with $M - N \to \infty$, form a special case of it.) In this setup, we prove the following result.

THEOREM 1.7. (Cf. Theorem 6.4 below) *Generalized linear sequences $\mathcal{T}_1, \dots, \mathcal{T}_k$ in a commutative group of transformations of X are jointly ergodic if and only if the sequence $\mathcal{T}_1 \times \cdots \times \mathcal{T}_k$ of transformations of X^k is ergodic and the sequences $\mathcal{T}_i^{-1}\mathcal{T}_j$ of transformations of X are ergodic for all $i \neq j$.*

Utilizing a technique developed by Green and Tao [**GT**], and following closely an argument of Sun [**S**], we then obtain a version of Theorem 1.4 for sequences along primes. Here is a special case of Theorem 7.1 below.

THEOREM 1.8. *Let $T_1, \dots, T_k$ be commuting invertible measure-preserving transformations of X and let $\varphi_1, \dots, \varphi_k$ be generalized linear functions $\mathbb{Z} \longrightarrow \mathbb{Z}$. Assume that for any $W \in \mathbb{N}$ and any $1 \le r \le W$ relatively prime with W, the sequences*

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

$\mathcal{T}_i(n) = T_i^{\varphi_i(Wn+r)}$, $i = 1, \ldots, k$, *are jointly ergodic. Then, for any* $f_1, \ldots, f_k \in L^\infty(X)$,

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p\in\mathcal{P}(N)} T_1^{\varphi_1(p)} f_1 \cdots T_k^{\varphi_k(p)} f_k = \prod_{i=1}^{k} \int_X f_i \, d\mu$$

in L^2 *norm, where, for* $N \in \mathbb{N}$, $\mathcal{P}(N) = \{p : p \text{ is prime, } p \le N\}$, *and* $\pi(N) = |\mathcal{P}(N)|$.

The structure of the paper is as follows: §§1–4 contain technical material related to properties of generalized linear functions. In §5, we investigate ergodic properties of what we call 'a generalized linear sequence of measure-preserving transformations'—a product of several sequences of the form $T^{\varphi(n)}$, where φ is an integer-valued generalized linear function. In §6, we obtain our main result, Theorem 6.4, the criterion of joint ergodicity of several commuting generalized linear sequences. In §7, we extend Theorem 6.4 to averaging along primes. In §8, we deal with families of transformations depending on a continuous parameter, and obtain a version of Theorem 6.4 for continuous flows. By using a 'change of variable' trick we also extend this result to more general families of transformations of the form $T^{\varphi(\sigma(t))}$, where φ is a generalized linear function and σ is a monotone function of 'regular' growth. For example, we have the following version of Corollary 1.6.

PROPOSITION 1.9. *Let* $T_1^s, \ldots, T_k^s$, $s \in \mathbb{R}$, *be commuting jointly ergodic continuous flows of measure-preserving transformations of* X *and let* φ *be an unbounded generalized linear function; then, for any* $\alpha > 0$, *the families* $T_1^{\varphi(t^\alpha)}, \ldots, T_k^{\varphi(t^\alpha)}$, $t \in [0, \infty)$, *are jointly ergodic (that is,* $\lim_{b\to\infty} (1/b) \int_0^b T_1^{\varphi(t^\alpha)} f_1 \cdots T_k^{\varphi(t^\alpha)} f_k \, dt = \prod_{i=1}^k \int_X f_i \, d\mu$ *for any* $f_1, \ldots, f_k \in L^\infty(X)$*) if and only if* $\lim_{b\to\infty} (1/b) \int_0^b \lambda^{\varphi(t)} \, dt = 0$ *for every* $\lambda \in \mathrm{Eig}(T_1^1, \ldots, T_k^1) \setminus \{1\}$.

Finally, §9 contains a result pertaining to joint ergodicity of several non-commuting generalized linear sequences.

2. *Generalized linear functions*

For $x \in \mathbb{R}$, we denote by $[x]$ the integer part of x and by $\{x\}$ the fractional part $x - [x]$ of x.

The set GLF of *generalized linear functions* is the minimal set of functions $\mathbb{R} \longrightarrow \mathbb{R}$ containing all linear functions $ax + b$ and closed under addition, multiplication by constants, and the operation of taking the integer (equivalently, the fractional) part. More exactly, we define GLF inductively in the following way. We put $\mathrm{GLF}_0 = \{\varphi(x) = ax + b, \ a, b \in \mathbb{R}\}$. After GLF_k has already been defined, we define GLF_{k+1} to be the space of functions spanned by GLF_k and the set $\{[\varphi], \ \varphi \in \mathrm{GLF}_k\}$. (Equivalently, we can define GLF_{k+1} to be the space spanned by GLF_k and the set $\{\{\varphi\}, \ \varphi \in \mathrm{GLF}_k\}$.) Finally, we put $\mathrm{GLF} = \bigcup_{k=0}^\infty \mathrm{GLF}_k$. For $\varphi \in \mathrm{GLF}$, we call the minimal k for which $\varphi \in \mathrm{GLF}_k$ *the weight* of φ.

We will refer to functions from GLF as *GL-functions*.

Example. $\varphi(x) = a_1\{a_2[a_3\{a_4x + a_5\} + a_6] + a_7[a_8x + a_9]\} + a_{10}x + a_{11}$, where $a_1, \ldots, a_{11} \in \mathbb{R}$, is a GL-function.

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

Clearly, the set of GL-functions is closed under composition: if $\varphi_1, \varphi_2 \in \mathrm{GLF}$, then $\varphi_1(\varphi_2(x)) \in \mathrm{GLF}$.

We define the set BGLF inductively in the following way: $\mathrm{BGLF}_1 = \{\varphi(x) = \{ax + b\},\ a, b \in \mathbb{R}\}$; if BGLF_k has already been defined, BGLF_{k+1} is the space spanned by the set $\mathrm{BGLF}_k \cup \{\{\varphi\},\ \varphi \in \mathrm{GLF}_k\}$; and, finally, $\mathrm{BGLF} = \bigcup_{k=1}^{\infty} \mathrm{BGLF}_k$.

LEMMA 2.1. BGLF *is exactly the set of bounded GL-functions.* (Hence the abbreviation 'BGLF'.)

Proof. Clearly, all elements of BGLF are bounded GL-functions. To prove the opposite inclusion, we use induction on the weight of GL-functions. Let $\varphi \in \mathrm{GLF}_k \backslash \mathrm{GLF}_{k-1}$ be bounded. If $k = 0$, then φ must be a constant and thus belongs to BGLF. If $k \geq 1$, $\varphi = \varphi_0 + \sum_{i=1}^{m} a_i\{\varphi_i\}$, where $\varphi_0, \varphi_1, \ldots, \varphi_m \in \mathrm{GLF}_{k-1}$. Now φ_0 is bounded; thus, by induction, $\varphi_0 \in \mathrm{BGLF}$, and $\{\varphi_1\}, \ldots, \{\varphi_m\} \in \mathrm{BGLF}$ by definition, so $\varphi \in \mathrm{BGLF}$. □

We will refer to elements of BGLF as *bounded generalized linear functions* or *BGL-functions*.

LEMMA 2.2. *Any GL-function φ is uniquely representable in the form $\varphi(x) = ax + \psi(x)$, where $a \in \mathbb{R}$ and ψ is a BGL-function.*

Proof. Every $\varphi \in \mathrm{GLF}_k$ has the form $\varphi = \varphi_0 + \sum_{i=1}^{m} a_i\{\varphi_i\}$ with $\varphi_0, \varphi_1, \ldots, \varphi_m \in \mathrm{GLF}_{k-1}$. We have $\sum_{i=1}^{m} a_i\{\varphi_i\} \in \mathrm{BGLF}$, and φ_0 is representable in the form $\varphi_0(x) = ax + \psi_0(x)$ with $\psi_0 \in \mathrm{BGLF}$ by induction on k.

As for the uniqueness, if $a_1x + \psi_1(x) = a_2x + \psi_2(x)$ with $a_1, a_2 \in \mathbb{R}$ and $\psi_1, \psi_2 \in \mathrm{BGLF}$, then the function $(a_1 - a_2)x$ is bounded, and so $a_1 = a_2$. □

COROLLARY 2.3. *Any GL-function φ is uniquely representable in the form $\varphi(x) = [ax] + \xi(x)$ with $a \in \mathbb{R}$ and $\xi \in \mathrm{BGLF}$.*

For a function $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ and $\alpha \in \mathbb{R}$, 'the difference derivative' $D_\alpha\varphi$ of φ with step α is $D_\alpha\varphi(x) = \varphi(x + \alpha) - \varphi(x)$, $x \in \mathbb{R}$.

COROLLARY 2.4. *For any GL-function φ and $\alpha \in \mathbb{R}$, $D_\alpha\varphi$ is a BGL-function.*

We will refer to BGL-functions taking values in $\{0, 1\}$ as *UGL-functions*.

LEMMA 2.5. *Let φ be a BGL-function. Then, for any $a \in \mathbb{R}$, the indicator functions $1_{\{\varphi<a\}}$, $1_{\{\varphi\leq a\}}$, $1_{\{\varphi>a\}}$, and $1_{\{\varphi\geq a\}}$ of the sets $\{x : \varphi(x) < a\}$, $\{x : \varphi(x) \leq a\}$, $\{x : \varphi(x) > a\}$, and $\{x : \varphi(x) \geq a\}$ are UGL-functions.*

Proof. We start with the set $\{\varphi \geq a\}$. Let $c = \sup|\varphi| + |a| + 1$. Then the function $\xi = (\varphi - a)/c + 1$ satisfies $0 < \xi < 2$, and $\varphi \geq a$ if and only if $\xi \geq 1$. Thus, the UGL-function $[\xi]$ is just $1_{\{\varphi\geq a\}}$.

Now $1_{\{\varphi\leq a\}} = 1_{\{-\varphi\geq -a\}}$, $1_{\{\varphi<a\}} = 1 - 1_{\{\varphi\leq a\}}$, and $1_{\{\varphi>a\}} = 1 - 1_{\{\varphi\leq a\}}$. □

We will now show that the set of UGL-functions is closed under Boolean operations. For two functions φ and ψ taking values in $\{0, 1\}$, let $\varphi \vee \psi = \max\{\varphi, \psi\} = \varphi + \psi - \varphi\psi$, $\varphi \wedge \psi = \min\{\varphi, \psi\} = \varphi\psi$, and $\neg\varphi = 1 - \varphi$.

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

PROPOSITION 2.6. *If φ, ψ are UGL-functions, then $\varphi \vee \psi$, $\varphi \wedge \psi$, and $\neg\varphi$ are also UGL-functions.*

Proof. $\neg\varphi = 1 - \varphi$ is clearly a UGL-function, $\varphi \vee \psi$ is the indicator function of the set $\{\varphi + \psi > 0\}$ and thus is a UGL-function by Lemma 2.5, and $\varphi \wedge \psi = \neg(\neg\varphi \vee \neg\psi)$. □

From Proposition 2.6, we get the following generalization of Lemma 2.5.

PROPOSITION 2.7. *Let $\varphi_1, \ldots, \varphi_k$ be BGL-functions and let $\varphi = (\varphi_1, \ldots, \varphi_k)$. For any interval $I = I_1 \times \cdots \times I_k \subseteq \mathbb{R}^k$ (where I_i are intervals in $\mathbb{R}$, which may be bounded or unbounded, open, closed, half-open half-closed, or degenerate), the indicator function 1_A of the set $A = \{x : \varphi(x) \in I\}$ is a UGL-function.*

We also have the following result.

PROPOSITION 2.8. *Let φ be an unbounded GL-function $\mathbb{Z} \longrightarrow \mathbb{Z}$. Then the indicator function 1_H of the range $H = \varphi(\mathbb{Z})$ of φ is a UGL-function.*

(Notice that GL-functions $\mathbb{Z} \longrightarrow \mathbb{R}$ are restrictions of GL-functions $\mathbb{R} \longrightarrow \mathbb{R}$; thus, all the results above apply.)

Proof. By Corollary 2.3, $\varphi(n) = [an] + \psi(n)$ for some $a \in \mathbb{R}$, $\psi \in \mathrm{BGLF}$. Since φ is integer valued, ψ is integer valued, and thus the range $K = \psi(\mathbb{Z})$ of ψ is a finite set of integers. Since φ is unbounded, $a \neq 0$; let us assume that $a > 0$.

If $n, k, j \in \mathbb{Z}$ are such that $n = [ak] + j$, then $0 \leq ak - n + j < 1$, so

$$\frac{n-j}{a} \leq k < \frac{n-j+1}{a}$$

and, so,

$$k \in \left\{ \left[\frac{n-j}{a} \right] + i,\ i \in I \right\},$$

where $I = \{0, 1, \ldots, [1/a] + 1\}$. Hence, if $n \in H$, that is, if $n = \varphi(k)$ for some $k \in \mathbb{Z}$, then

$$k \in \left\{ \left[\frac{n-j}{a} \right] + i,\ i \in I,\ j \in K \right\}.$$

For each $i \in I$ and $j \in K$, define

$$\delta_{i,j}(n) = n - \varphi\left(\left[\frac{n-j}{a} \right] + i \right) = n - a\left(\left[\frac{n-j}{a} \right] + i \right) - \psi\left(\left[\frac{n-j}{a} \right] + i \right);$$

then $\delta_{i,j} \in \mathrm{BGLF}$ for all i, j, and $n \in H$ if and only if $\delta_{i,j}(n) = 0$ for some i, j. By Lemma 2.5 and Proposition 2.6, the indicator functions $1_{\{\delta_{i,j}=0\}}$ are UGL-functions for all i, j, and thus the function $1_H = \bigvee_{\substack{i \in I \\ j \in K}} 1_{\{\delta_{i,j}=0\}}$ is also a UGL-function by Proposition 2.6. □

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

3. *C-lims, D-lims, densities, and the van der Corput trick*

This is a technical section. Starting from this moment, we fix an arbitrary Følner sequence $(\Phi_N)_{N=1}^{\infty}$ in $\mathbb{Z}$ (that is, a sequence of finite subsets of $\mathbb{Z}$ with the property that for any $h \in \mathbb{Z}$, $|(\Phi_N - h)\triangle\Phi_N|/|\Phi_N| \to 0$ as $N \to \infty$).

Under 'a sequence' we will usually understand a function with domain $\mathbb{Z}$. For a sequence (u_n) of real numbers, or of elements of a normed vector space, we define $\text{C-lim}_n\, u_n = \lim_{N\to\infty} (1/|\Phi_N|) \sum_{n\in\Phi_N} u_n$, if this limit exists. When u_n are real numbers, we define $\text{C-limsup}_n\, u_n = \limsup_{N\to\infty} (1/|\Phi_N|) \sum_{n\in\Phi_N} u_n$. When u_n are elements of a normed vector space, we also define $\text{C-limsup}_{\|\cdot\|,n}\, u_n = \limsup_{N\to\infty} \|(1/|\Phi_N|) \sum_{n\in\Phi_N} u_n\|$.

For a set $E \subseteq \mathbb{Z}$, we define *the density* of E to be $\text{d}(E) = \lim_{N\to\infty} |E \cap \Phi_N|/|\Phi_N|$, if this limit exists. We also define *the upper density* and *the lower density* of E as $\text{d}^*(E) = \limsup_{N\to\infty} |E \cap \Phi_N|/|\Phi_N|$ and $\text{d}_*(E) = \liminf_{N\to\infty} |E \cap \Phi_N|/|\Phi_N|$, respectively.

We will say that a sequence (z_n) in a probability measure space (Z, λ) is *uniformly distributed* if $\text{C-lim}_n\, g(z_n) = \int_Z g\, d\lambda$ for any $g \in C(Z)$.

For a sequence (u_n) of vectors in a normed vector space, we write $\text{D-lim}_n\, u_n = u$ if, for any $\varepsilon > 0$, $\text{d}(\{n : \|u_n - u\| \geq \varepsilon\}) = 0$. Clearly, this is equivalent to $\text{C-lim}_n\, \|u_n - u\| = 0$. For a sequence (u_n) of real numbers, we also define $\text{D-limsup}_n\, u_n$ as $\inf\{u \in \mathbb{R} : \text{d}(\{n : u_n > u\}) = 0\}$.

We will be using the following version of the van der Corput trick.

LEMMA 3.1. *Let (u_n) be a bounded sequence of elements of a Hilbert space. Then, for any finite subset D of $\mathbb{Z}$,*

$$\operatorname*{C\text{-}limsup}_{n,\|\cdot\|} u_n \leq \left(\frac{1}{|D|^2} \sum_{h_1,h_2\in D} \operatorname*{C\text{-}limsup}_n \langle u_{n+h_1}, u_{n+h_2}\rangle\right)^{1/2}.$$

Thus, if for some $\varepsilon > 0$ there exists an infinite set $B \subseteq \mathbb{Z}$ such that $|\text{C-limsup}_n \langle u_{n+h_1}, u_{n+h_2}\rangle| < \varepsilon$ for all distinct $h_1, h_2 \in B$, then $\text{C-limsup}_{n,\|\cdot\|} u_n < \sqrt{\varepsilon}$.

Proof. Let $D \subseteq \mathbb{Z}$, $|D| < \infty$. For any $N \in \mathbb{N}$, we have

$$\frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} u_n = \frac{1}{|D|} \sum_{h\in D} \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} u_n = \left(\frac{1}{|D|} \sum_{h\in D} \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} u_{n+h}\right) - A_N + B_N,$$

where

$$A_N = \frac{1}{|D|} \sum_{h\in D} \frac{1}{|\Phi_N|} \sum_{\substack{n\in\Phi_N \\ n+h\notin\Phi_N}} u_{n+h} \quad \text{and} \quad B_N = \frac{1}{|D|} \sum_{h\in D} \frac{1}{|\Phi_N|} \sum_{\substack{n\notin\Phi_N \\ n+h\in\Phi_N}} u_{n+h}.$$

Since $\{\Phi_N\}_{N=1}^{\infty}$ is a Følner sequence and the sequence (u_n) is bounded, $\|A_N\|, \|B_N\| \to 0$ as $N \to \infty$. Thus,

$$\limsup_{N\to\infty} \left\| \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} u_n \right\| = \limsup_{N\to\infty} \left\| \frac{1}{|D|} \sum_{h\in D} \frac{1}{|\Phi_N|} \sum_{u\in\Phi_N} u_{n+h} \right\|.$$

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

By Schwarz's inequality,

$$\begin{aligned}\left\|\frac{1}{|D|}\sum_{h\in D}\frac{1}{|\Phi_N|}\sum_{n\in\Phi_N}u_{n+h}\right\|^2 &= \frac{1}{|D|^2}\left\|\frac{1}{|\Phi_N|}\sum_{n\in\Phi_N}\sum_{h\in D}u_{n+h}\right\|^2\\ &\le \frac{1}{|D|^2}\frac{1}{|\Phi_N|}\sum_{n\in\Phi_N}\left\|\sum_{h\in D}u_{n+h}\right\|^2\\ &= \frac{1}{|D|^2}\frac{1}{|\Phi_N|}\sum_{n\in\Phi_N}\sum_{h_1,h_2\in D}\langle u_{n+h_1},u_{n+h_2}\rangle,\end{aligned}$$

so

$$\begin{aligned}\left(\limsup_{N\to\infty}\left\|\frac{1}{|\Phi_N|}\sum_{n\in\Phi_N}u_n\right\|\right)^2 &= \left(\limsup_{N\to\infty}\left\|\frac{1}{|D|}\sum_{h\in D}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}u_{n+h}\right\|\right)^2\\ &= \limsup_{N\to\infty}\left\|\frac{1}{|D|}\sum_{h\in D}\frac{1}{|\Phi_N|}\sum_{n\in\Phi_N}u_{n+h}\right\|^2\\ &\le \limsup_{N\to\infty}\frac{1}{|D|^2}\sum_{h_1,h_2\in D}\frac{1}{|\Phi_N|}\sum_{n\in\Phi_N}\langle u_{n+h_1},u_{n+h_2}\rangle\\ &\le \frac{1}{|D|^2}\sum_{h_1,h_2\in D}\limsup_{N\to\infty}\frac{1}{|\Phi_N|}\sum_{n\in\Phi_N}\langle u_{n+h_1},u_{n+h_2}\rangle.\end{aligned}$$

To get the second assertion, for any finite set $D\subseteq B$ write

$$\begin{aligned}&\left|\frac{1}{|D|^2}\sum_{h_1,h_2\in D}\operatorname*{C\text{-}limsup}_n\langle u_{n+h_1},u_{n+h_2}\rangle\right|\\ &\quad\le\frac{1}{|D|^2}\sum_{\substack{h_1,h_2\in D\\ h_1\neq h_2}}\left|\operatorname*{C\text{-}limsup}_n\langle u_{n+h_1},u_{n+h_2}\rangle\right|+\frac{1}{|D|^2}\sum_{h\in D}\left|\operatorname*{C\text{-}limsup}_n\langle u_{n+h},u_{n+h}\rangle\right|\\ &\quad\le\varepsilon+\frac{1}{|D|}\sup_n\|u_n\|^2\end{aligned}$$

and notice that the second summand tends to zero as $|D|\to\infty$. □

We will also need a simple 'finitary version' of the van der Corput trick.

LEMMA 3.2. *Let $u_1,\dots,u_N$ be elements of a Hilbert space. Then*

$$\left\|\frac{1}{N}\sum_{n=1}^{N}u_n\right\|^2\le\frac{1}{N}\sum_{h\in\mathbb{Z}}\left|\frac{1}{N}\sum_{n=1}^{N}\langle u_n,u_{n+h}\rangle\right|,$$

where we assume that $u_m=0$ if $m\notin\{1,\dots,N\}$ (and, so, the sum on the right is finite).

Proof.

$$\begin{aligned}\left\|\frac{1}{N}\sum_{n=1}^{N}u_n\right\|^2 &= \frac{1}{N^2}\sum_{n,m=1}^{N}\langle u_n,u_m\rangle=\frac{1}{N^2}\sum_{h\in\mathbb{Z}}\sum_{n=1}^{N}\langle u_n,u_{n+h}\rangle\\ &\le\frac{1}{N}\sum_{h\in\mathbb{Z}}\left|\frac{1}{N}\sum_{n=1}^{N}\langle u_n,u_{n+h}\rangle\right|.\end{aligned}$$

□

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

4. *BGL-functions and Besicovitch almost periodicity*

We will now describe and use a 'dynamical' approach to BGL-functions. We will focus on functions $\mathbb{Z} \longrightarrow \mathbb{R}$.

Let $\mathcal{M}$ be a torus, $\mathcal{M} = V/\Gamma$, where V is a finite-dimensional $\mathbb{R}$-vector space and Γ is a cocompact lattice in V, and let π be the projection $V \longrightarrow \mathcal{M}$. We call *a polygon* any bounded subset P of V defined by a system of linear inequalities, strict or non-strict:

$$P = \{v \in V : L_1(v) < c_1, \ldots, L_k(v) < c_k, L_{k+1}(v) \leq c_{k+1}, \ldots, L_m(v) \leq c_m\},$$

where L_i are linear functions on V and $c_i \in \mathbb{R}$. Let Q be a parallelepiped in V such that $\pi|_Q : Q \longrightarrow \mathcal{M}$ is a bijection. (Q is a fundamental domain of $\mathcal{M}$ in V.) Assume that $Q = \bigcup_{j=1}^{l} \widehat{P}_j$ is a finite partition of Q into disjoint polygons. Let a function $\widetilde{F}$ on Q be the sum, $\widetilde{F} = L + E$, of a linear function L and of a function E which is constant on each of $\widehat{P}_j$. Finally, let F be the function induced by $\widetilde{F}$ on $\mathcal{M}$, $F = \widetilde{F} \circ (\pi|_Q)^{-1}$. We will call functions F obtainable in this way *polygonally broken linear* or *PGL-functions*.

Example. The function $\{2x + 1/3\}$ on $\mathbb{R}/\mathbb{Z}$ is a PGL-function.

The following result is clear.

LEMMA 4.1. *The set of PGL-functions on a torus $\mathcal{M}$ is closed under addition, multiplication by scalars, and the operation of taking the fractional part.*

The following theorem says that BGL-functions are dynamically obtainable from PGL-functions.

THEOREM 4.2. *For any BGL-function φ, there exists a torus $\mathcal{M}$, an element $u \in \mathcal{M}$, and a PGL-function F on $\mathcal{M}$ such that $\varphi(n) = F(nu)$, $n \in \mathbb{Z}$.*

Proof. For $\varphi(n) = \{an + b\}$, $a, b \in \mathbb{R}$, take $\mathcal{M} = \mathbb{R}/\mathbb{Z}$, $u = a \mod \mathbb{Z}$, and $F(x) = \{x + b\}$, $x \in \mathcal{M}$.

The set of BGL-functions satisfying the assertion of the theorem is closed under addition and multiplication by constants. Indeed, if a BGL-function φ is represented in the form $\varphi(n) = F(nu)$, $n \in \mathbb{Z}$, where F is a PGL-function on a torus $\mathcal{M}$ and $u \in \mathcal{M}$, then for $a \in \mathbb{R}$ the function aF is a PGL-function as well and $a\varphi(n) = (aF)(nu)$, $n \in \mathbb{Z}$. If BGL-functions φ_1, φ_2 are represented as $\varphi_1(n) = F_1(nu)$, $\varphi_2(n) = F_2(nv)$, $n \in \mathbb{Z}$, where F_1, F_2 are PGL-functions on tori $\mathcal{M}_1$ and $\mathcal{M}_2$, respectively, $u_1 \in \mathcal{M}_1$ and $u_2 \in \mathcal{M}_2$, then the function $F(x_1, x_2) = F_1(x_1) + F_2(x_2)$ on the torus $\mathcal{M}_1 \times \mathcal{M}_2$ is a PGL-function and $(\varphi_1 + \varphi_2)(n) = F(n(u_1, u_2))$, $n \in \mathbb{Z}$.

Also, the set of BGL-functions satisfying the assertion of the theorem is closed under the operation of taking the fractional part: if a BGL-function φ is represented as $\varphi(n) = F(nu)$, $n \in \mathbb{Z}$, where F is a PGL-function on a torus $\mathcal{M}$ and $u \in \mathcal{M}$, then the function $\{F\}$ is a PGL-function and $\{\varphi(n)\} = \{F\}(nu)$, $n \in \mathbb{Z}$.

From the inductive definition of BGL-functions, it follows that the theorem holds for all BGL-functions. □

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

Any closed subgroup Z of a torus $\mathcal{M}$ has the form $Z = \mathcal{M}' \times J$ for some subtorus $\mathcal{M}'$ of $\mathcal{M}$ and a finite abelian group J. We will say that a function F on Z is *a PGL-function* if the restriction $F_{|\mathcal{M}' \times \{i\}}$ is a PGL-function on the torus $\mathcal{M}' \times \{i\}$ for every $i \in J$.

If Z is a closed subgroup of a torus $\mathcal{M}$ and F is a PGL-function on $\mathcal{M}$, then $F_{|Z}$ is a PGL-function on Z. In the environment of Theorem 4.2, putting $Z = \overline{\mathbb{Z}u}$, we obtain the following result.

PROPOSITION 4.3. *For any BGL-function φ, there exists a compact abelian group Z, of the form $Z = \mathcal{M}' \times J$, where $\mathcal{M}'$ is a torus and J is a finite cyclic group, an element $u \in Z$, whose orbit $\mathbb{Z}u$ is dense (and, so, uniformly distributed) in Z, and a PGL-function F on Z such that $\varphi(n) = F(nu)$, $n \in \mathbb{Z}$.*

COROLLARY 4.4. *For any BGL-function φ, the limit* $\text{C-lim}_n \varphi(n)$ *exists. For any BGL-functions $\varphi_1, \ldots, \varphi_k$, for $\varphi = (\varphi_1, \ldots, \varphi_k)$, and for any polygon $P \subseteq \mathbb{R}^k$, the density of the set $\{n \in \mathbb{Z} : \varphi(n) \in P\}$ exists.*

As another corollary of Proposition 4.3, we get the following result.

PROPOSITION 4.5. *Let $\varphi : \mathbb{Z} \longrightarrow \mathbb{R}$ be a BGL-function. For any $\varepsilon > 0$, there exists $h \in \mathbb{Z}$ such that* $\text{D-limsup}_n |\varphi(n + h) - \varphi(n)| < \varepsilon$ *and there exists a trigonometric polynomial q such that* $\text{D-limsup}_n |\varphi(n) - q(n)| < \varepsilon$.

Remark. Functions with these properties are called *Besicovitch almost periodic* (at least, in the case where the Følner sequence with respect to which the densities are measured is $\Phi_N = [-N, N]$, $N \in \mathbb{N}$). Any function obtainable dynamically with the help of a rotation of a compact commutative Lie group and a Riemann integrable function thereon is such.

Proof. Represent φ in the form $\varphi(n) = F(nu)$, $n \in \mathbb{Z}$, as in Proposition 4.3. Let $Z = \bigcup_{j=1}^{l} P_j$ be the polygonal partition of Z such that F is linear on each of P_j. Let U be a δ-neighborhood of $\bigcup_{j=1}^{l} \partial P_j$ with $\delta > 0$ small enough so that $\lambda(U) < \varepsilon$, where λ is the normalized Haar measure on Z. Let $\widehat{F}$ be a continuous function on Z which coincides with F on $Z \backslash U$ and such that $\sup |\widehat{F}| \leq \sup |F| = \sup |\varphi|$. Let $\hat{\varphi}(n) = \widehat{F}(nu)$, $n \in \mathbb{Z}$. The sequence (nu) is uniformly distributed on Z; thus, $\text{d}^*(\{n \in \mathbb{Z} : nu \in U\}) < \varepsilon$, and so $\text{d}^*(\{n : \varphi(n) \neq \hat{\varphi}(n)\}) = \text{d}^*(\{n : F(nu) \neq \widehat{F}(nu)\}) < \varepsilon$. Since $\widehat{F}$ is uniformly continuous, for any $h \in \mathbb{Z}$ for which hu is close enough to 0 we have $|\widehat{F}(v + hu) - \widehat{F}(v)| < \varepsilon$ for all $v \in Z$, so $|\hat{\varphi}(n + h) - \hat{\varphi}(n)| < \varepsilon$ for all $n \in \mathbb{Z}$. This implies that $\text{D-limsup}_n |\varphi(n + h) - \varphi(n)| < \varepsilon + 2\varepsilon \sup |\varphi|$. And, if Θ is a finite linear combination of characters of Z such that $|\widehat{F} - \Theta| < \varepsilon$, then, for the trigonometric polynomial $q(n) = \Theta(nu)$, $n \in \mathbb{Z}$, we have $|\hat{\varphi}(n) - q(n)| < \varepsilon$ for all n, which implies that $\text{D-limsup}_n |\varphi(n) - q(n)| < \varepsilon + (\sup |\varphi| + \sup |q|)\varepsilon = \varepsilon + (2 \sup |\varphi| + \varepsilon)\varepsilon$. □

COROLLARY 4.6. *If φ is a BGL-function $\mathbb{Z} \longrightarrow \mathbb{Z}$, then for any $\varepsilon > 0$ there exists $h \in \mathbb{Z}$ such that* $\text{d}(\{n \in \mathbb{Z} : \varphi(n + h) = \varphi(n)\}) > 1 - \varepsilon$.

(Notice that the density of the set $\{n \in \mathbb{Z} : \varphi(n + h) = \varphi(n)\}$ exists by Corollary 4.4.) We now turn to unbounded GL-functions. From Lemma 2.2 and Theorem 4.2, we see that any GL-function φ is representable in the form $\varphi(n) = an + F(nu)$, where $a \in \mathbb{R}$, F is a

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

PGL-function on a torus $\mathcal{M}$, and $u \in \mathcal{M}$. Given several GL-functions $\varphi_1, \ldots, \varphi_k$, we can read them off a single torus: for each i, represent φ_i in the form $\varphi_i(n) = a_i n + F_i(nu_i)$, where $a_i \in \mathbb{R}$, F_i is a PGL-function on a torus $\mathcal{M}_i$, and $u_i \in \mathcal{M}_i$, put $\mathcal{M} = \prod_{i=1}^k \mathcal{M}_i$, $u = (u_1, \ldots, u_k) \in \mathcal{M}$, and lift $F_1, \ldots, F_k$ to a function on $\mathcal{M}$; then $\varphi_i(n) = a_i n + F_i(nu)$, $n \in \mathbb{Z}$, $i = 1, \ldots, k$. As a corollary, we get the following result.

PROPOSITION 4.7. *Given GL-functions $\varphi_1, \ldots, \varphi_k$, there exists a torus $\mathcal{M}$, an element $u \in \mathcal{M}$, and a polygonal partition $\mathcal{M} = \bigcup_{j=1}^l P_j$, such that for each i, j, $\varphi_i(n+h) - \varphi_i(n)$ does not depend on n if both nu, $(n+h)u \in P_j$.*

Proof. Let $\mathcal{M}$, u, and F_i be as above; let $\mathcal{M} = V/\Gamma$, where V is a vector space and Γ is a lattice in V, π be the projection $V \longrightarrow \mathcal{M}$, $Q \subset V$ be the fundamental domain of $\mathcal{M}$ in V, and $\widetilde{F}_i = F \circ \pi|_Q$, $i = 1, \ldots, k$. Choose a partition $\mathcal{M} = \bigcup_{j=1}^l P_j$ of $\mathcal{M}$ such that for each j and each i, the function F_i is linear on P_j and, additionally, for each j, $((\widehat{P}_j - \widehat{P}_j) - (\widehat{P}_j - \widehat{P}_j)) \cap \Gamma = \{0\}$, where $\widehat{P}_j = \pi^{-1}(P_j) \cap Q$. Then, for any i and j, for $v, w \in P_j$, $F_i(v) - F_i(w)$ depends on $v - w$ only. Indeed, let $v_1, w_1, v_2, w_2 \in P_j$ be such that $v_1 - w_1 = v_2 - w_2$; let $\hat{v}_t = {\pi|_Q}^{-1}(v_t)$, $\hat{w}_t = {\pi|_Q}^{-1}(w_t)$, $t = 1, 2$; then $(\hat{v}_1 - \hat{w}_1) - (\hat{v}_2 - \hat{w}_2) \in \Gamma$, so $= 0$; thus,

$$\begin{aligned} F_i(v_1) - F_i(w_1) &= \widetilde{F}_i(\hat{v}_1) - \widetilde{F}_i(\hat{w}_1) = L_i(\hat{v}_1) - L_i(\hat{w}_1) = L_i(\hat{v}_1 - \hat{w}_1) \\ &= L_i(\hat{v}_2 - \hat{w}_2) = F_i(v_2) - F_i(w_2), \end{aligned}$$

where L_i is the linear function on V that coincides with $\widetilde{F}_i$ on $\widehat{P}_j$ up to a constant. Now, if n and h are such that both nu, $(n+h)u \in P_j$ for some j, then, for any i, $\varphi_i(n+h) - \varphi_i(n) = a_i h + F_i(nu + hu) - F_i(nu)$, and $F_i(nu + hu) - F_i(nu)$ does not depend on n. □

A set $H \subseteq \mathbb{Z}$ is said to be *a Bohr set* if H contains a non-empty subset of the form $\{n \in \mathbb{Z} : nu \in W\}$, where u and W are an element and an open subset of a torus. Any Bohr set is infinite and has positive density (with respect to any Følner sequence in $\mathbb{Z}$). The following proposition says that (several) GL-functions are 'almost linear' along a Bohr set.

PROPOSITION 4.8. *For any GL-functions $\varphi_1, \ldots, \varphi_k$ and any $\varepsilon > 0$, there exist a Bohr set $H \subseteq \mathbb{Z}$ and constants $C_1, \ldots, C_k$ such that for any $h \in H$,*

$$\mathrm{d}(\{n \in \mathbb{Z} : \varphi_i(n+h) = \varphi_i(n) + \varphi_i(h) + C_i,\ i = 1, \ldots, k\}) > 1 - \varepsilon.$$

Proof. First of all, for any h, the density of the set $\{n \in \mathbb{Z} : \varphi_i(n+h) = \varphi_i(n) + \varphi_i(h) + C_i,\ i = 1, \ldots, k\}$ exists by Corollary 4.4.

Let $\mathcal{M}$ be a torus, $u \in \mathcal{M}$, and $F_1, \ldots, F_k$ be PGL-functions on $\mathcal{M}$ such that $\varphi_i(n) = F_i(nu)$, $i = 1, \ldots, k$. Let $Z = \overline{\mathbb{Z}u}$; then the sequence $(nu)_{n \in \mathbb{Z}}$ is uniformly distributed in Z. Let $Z = \bigcup_{j=1}^l P_j$ be the polygonal partition of Z such that for every i and j, $F_i|_{P_j} = L_i + C_{i,j}$, where L_i is linear and $C_{i,j}$ is a constant. Let $\delta > 0$ be small enough so that $\lambda(U) < \varepsilon$, where U is the δ-neighborhood of the set $\bigcup_{j=1}^l \partial P_j$ and λ is the normalized Haar measure on Z. Let W_0 be the δ-neighborhood of 0. Now, for any $w \in W_0$, $\mathrm{d}(\{n \in \mathbb{Z} : nu \in P_{j_1},\ nu + w \in P_{j_2},\ j_1 \neq j_2\}) < \varepsilon$. Choose j_0 for which 0 is a limit point of

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

the interior $\overset{\circ}{P}_{j_0}$ of P_{j_0}, let $W = \overset{\circ}{P}_{j_0} \cap W_0$, and let $H = \{n \in \mathbb{Z} : nu \in W\}$. Then, for any $w \in W$ and any i, whenever v, $v + w \in P_j$ for some j, we have

$$\begin{aligned} F_i(v+w) = L_i(v+w) + C_{i,j} &= L_i(v) + C_{i,j} + L_i(w) + C_{i,j_0} - C_{i,j_0} \\ &= F_i(v) + F_i(w) + C_i, \end{aligned}$$

where $C_i = -C_{i,j_0}$. For every $h \in H$, let $E_h = \{n \in \mathbb{Z} : nu, (n+h)u \in P_j \text{ for some } j\}$; then $\mathrm{d}(E_h) > 1 - \varepsilon$ and, for any $n \in E_h$ and any i, $\varphi_i(n+h) = \varphi_i(n) + \varphi_i(h) + C_i$. □

5. *Generalized linear sequences of transformations*

A generalized linear sequence (a GL-sequence) in a commutative group G is a sequence of the form $\mathcal{T}(n) = T_1^{\varphi_1(n)} \cdots T_r^{\varphi_r(n)}$, $n \in \mathbb{Z}$, where $T_1, \ldots, T_r \in G$ and $\varphi_1, \ldots, \varphi_r$ are GL-functions $\mathbb{Z} \longrightarrow \mathbb{Z}$. We say that $\mathcal{T}$ is a BGL-sequence if $\varphi_1, \ldots, \varphi_r$ are BGL-functions. Corollaries 2.3 and 4.6 and Propositions 4.7 and 4.8 imply the following properties of GL-sequences.

PROPOSITION 5.1. *Let G be a commutative group.*

(i) *If $\mathcal{T}$ is a GL-sequence in G, then, for any $h \in \mathbb{Z}$, the sequence $\mathcal{T}(n)^{-1}\mathcal{T}(n+h)$, $n \in \mathbb{Z}$, is a BGL-sequence.*

(ii) *If $\mathcal{T}$ is a BGL-sequence in G, then, for any $\varepsilon > 0$, there exists $h \in \mathbb{Z}$ such that $\mathrm{d}_*(\{n \in \mathbb{Z} : \mathcal{T}(n+h) = \mathcal{T}(n)\}) > 1 - \varepsilon$.*

(iii) *If $\mathcal{T}_1, \ldots, \mathcal{T}_k$ are GL-sequences in G (or, more generally, in distinct commutative groups $G_1, \ldots, G_k$, respectively), then there exist a torus $\mathcal{M}$, an element $u \in \mathcal{M}$, and a polygonal partition $\mathcal{M} = \bigcup_{j=1}^{l} P_j$ such that for any i, j, $\mathcal{T}_i(n)^{-1}\mathcal{T}_i(n+h)$ does not depend on n whenever nu, $(n+h)u \in P_j$.*

(iv) *If $\mathcal{T}_1, \ldots, \mathcal{T}_k$ are GL-sequences in G (or, more generally, in distinct commutative groups $G_1, \ldots, G_k$, respectively), then, for any $\varepsilon > 0$, there exist a Bohr set $H \subseteq \mathbb{Z}$ and elements $S_1, \ldots, S_k \in G$ (respectively, $S_i \in G_i$, $i = 1, \ldots, k$) such that for any $h \in H$, the set $E_h = \{n \in \mathbb{Z} : \mathcal{T}_i(n+h) = \mathcal{T}_i(n)\mathcal{T}_i(h)S_i,\ i = 1, \ldots, k\}$ satisfies $\mathrm{d}_*(E_h) > 1 - \varepsilon$.*

If $\mathcal{T}$ is a GL-sequence of unitary operators on a Hilbert space $\mathcal{H}$, then, via the spectral theorem, Corollary 4.4 implies the following result. (See Corollary 0.27 in [**BeL2**] for a more general statement.)

LEMMA 5.2. *For any $f \in \mathcal{H}$, $\text{C-lim}_n \mathcal{T}(n)f$ exists.*

We now fix a commutative group G of measure-preserving transformations of a probability measure space (X, μ), and denote by $\mathfrak{T}$ the set of GL-sequences of transformations in G.

DEFINITION 5.3. If $\mathcal{T}$ is a sequence of measure-preserving transformations of X (or just a sequence of unitary operators on a Hilbert space $\mathcal{H}$), we say that $\mathcal{T}$ is *ergodic* if $\text{C-lim}_n \mathcal{T}(n)f = \int_X f\, d\mu$ for all $f \in L^2(X)$ (respectively, $\text{C-lim}_n \mathcal{T}(n)f = 0$ for all $f \in \mathcal{H}$). We will also say that $\mathcal{T}$ is *weakly mixing* if for any $f, g \in L^2(X)$ one has $\text{D-lim}_n \int_X \mathcal{T}(n)f \cdot g\, d\mu = \int_X f\, d\mu \int_X g\, d\mu$ (respectively, $\text{D-lim}_n \langle \mathcal{T}(n)f, g\rangle = 0$ for all $f, g \in \mathcal{H}$).

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

REMARK 5.4. We have defined our C-lims and, so, ergodicity of a sequence of transformations, with respect to a fixed Følner sequence in $\mathbb{Z}$. However, since, for any GL-sequence $\mathcal{T}$ of measure-preserving transformations, or of unitary operators, and for any (function or vector) f, C-lim $\mathcal{T}(n)f$ exists with respect to any Følner sequence, this limit is the same for all Følner sequences; thus, the ergodicity of GL-sequences does not depend on the choice of the Følner sequence.

Let $\mathcal{H}_{\mathrm{c}} \oplus \mathcal{H}_{\mathrm{wm}}$ be the compact/weak mixing decomposition of $L^2(X)$ induced by G, meaning that $\mathcal{H}_{\mathrm{c}}$ is the subspace of $L^2(X)$ on which all elements of G act in a compact way and $\mathcal{H}_{\mathrm{wm}}$ is the orthocomplement of $\mathcal{H}_{\mathrm{c}}$; then for any $g \in \mathcal{H}_{\mathrm{wm}}$ there exists a transformation $T \in G$ that acts on g in a weakly mixing fashion. Notice also that if $T \in G$ is ergodic, then T is weakly mixing on $\mathcal{H}_{\mathrm{wm}}$. The following theorem says that any ergodic sequence from $\mathfrak{T}$ is weakly mixing on $\mathcal{H}_{\mathrm{wm}}$.

THEOREM 5.5. *If $\mathcal{T} \in \mathfrak{T}$ is ergodic, then for any $f \in \mathcal{H}_{\mathrm{wm}}$ and $g \in L^2(X)$ one has* D-lim$_n \int_X \mathcal{T}(n)f \cdot g \, d\mu = 0$.

We first prove that $\mathcal{T}$ has no 'eigenfunctions' in $\mathcal{H}_{\mathrm{wm}}$.

LEMMA 5.6. *If $\mathcal{T} \in \mathfrak{T}$ is ergodic, then for any $f \in \mathcal{H}_{\mathrm{wm}}$ and $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ one has* C-lim$_n \lambda^n \mathcal{T}(n)f = 0$.

Proof. We may and will assume that $|f| \le 1$. Fix $g \in \mathcal{H}_{\mathrm{wm}}$ with $|g| \le 1$ and let $T \in G$ be a transformation that acts weakly mixingly on g. We are going to apply the van der Corput trick (Lemma 3.1 above) to the sequence $f_n = \lambda^n T^n \mathcal{T}(n) f \cdot T^n g$, $n \in \mathbb{Z}$. Let $\varepsilon > 0$ and let a Bohr set $H \subseteq \mathbb{Z}$, a transformation $S \in G$, and sets $E_h \subseteq \mathbb{Z}$, $h \in H$, be as in Proposition 5.1(iv), applied to the single GL-sequence $\mathcal{T}$. Let $h_1, h_2 \in H$; for any $n \in E_{h_1} \cap E_{h_2}$, one has

$$\begin{aligned}
\langle f_{n+h_1}, f_{n+h_2} \rangle &= \int_X f_{n+h_1} \bar{f}_{n+h_2} \, d\mu \\
&= \int_X \lambda^{h_1} T^{n+h_1} \mathcal{T}(n+h_1) f \cdot T^{n+h_1} g \cdot \bar{\lambda}^{h_2} T^{n+h_2} \mathcal{T}(n+h_2) \bar{f} \cdot T^{n+h_2} \bar{g} \, d\mu \\
&= \int_X \mathcal{T}(n)(\lambda^{h_1-h_2} T^{h_1} \mathcal{T}(h_1) S f \cdot T^{h_2} \mathcal{T}(h_2) S \bar{f}) \cdot (T^{h_1} g \cdot T^{h_2} \bar{g}) \, d\mu \\
&= \int_X \mathcal{T}(n) \tilde{f}_{h_1,h_2} \cdot (T^{h_1} g \cdot T^{h_2} \bar{g}) \, d\mu,
\end{aligned}$$

where $\tilde{f}_{h_1,h_2} = \lambda^{h_1-h_2} T^{h_1} \mathcal{T}(h_1) S f \cdot T^{h_2} \mathcal{T}(h_2) S \bar{f}$. Since $\mathcal{T}$ is ergodic,

$$\text{C-lim}_n \int_X \mathcal{T}(n) \tilde{f}_{h_1,h_2} \cdot (T^{h_1} g \cdot T^{h_2} \bar{g}) \, d\mu = \left(\int_X \tilde{f}_{h_1,h_2} d\mu \right) \left(\int_X T^{h_1} g \cdot T^{h_2} \bar{g} \, d\mu \right)$$

and, since $\mathrm{d}_*(E_{h_1} \cap E_{h_2}) > 1 - 2\varepsilon$ and $|\tilde{f}_{h_1,h_2}|, |g| \le 1$,

$$\left| \text{C-limsup}_n \langle f_{n+h_1}, f_{n+h_2} \rangle \right| \le \left| \int_X T^{h_1} g \cdot T^{h_2} \bar{g} \, d\mu \right| + 2\varepsilon.$$

Since D-lim$_h \int_X T^h g \cdot g' \, d\mu = 0$ for any $g' \in L^2(X)$, and since $\mathrm{d}_*(H) > 0$, we can construct an infinite set $B \subseteq H$ such that $|\int_X T^{h_1} g \cdot T^{h_2} \bar{g} \, d\mu| < \varepsilon$ for any distinct

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

$h_1, h_2 \in B$. (Indeed, choose $b_1 \in H$ arbitrarily. The set $B_1 = \{h \in \mathbb{N} : |\int_X T^h g \cdot T^{b_1} \bar{g}\, d\mu| < \varepsilon\}$ has density one, so $B_1 \cap H \neq \emptyset$; choose $b_2 \in B_1 \cap H$. Next, the set $B_2 = \{h \in \mathbb{N} : |\int_X T^h g \cdot T^{b_2} \bar{g}\, d\mu| < \varepsilon\}$ also has density one, so there is $b_3 \in B_1 \cap B_2 \cap H$, etc. Then put $B = \{b_1, b_2, \ldots\}$.) Then, for any distinct $h_1, h_2 \in B$, we have $|\text{C-limsup}_n \langle f_{n+h_1}, f_{n+h_2}\rangle| < 3\varepsilon$, which, by Lemma 3.1, implies that $\text{C-limsup}_{\|\cdot\|,n} f_n \leq \sqrt{3\varepsilon}$. Since ε is arbitrary, $\text{C-lim}_n f_n = 0$.

Now let $\hat{f} = \text{C-lim}_n \lambda^n \mathcal{T}(n) f \in \mathcal{H}_{\text{wm}}$ (this limit exists by Lemma 5.2). Then, for any $g \in L^2(X)$,

$$\int_X \hat{f} \cdot g\, d\mu = \operatorname*{C\text{-}lim}_n \int_X \lambda^n \mathcal{T}(n) f \cdot g\, d\mu = \operatorname*{C\text{-}lim}_n \int_X f_n\, d\mu = 0.$$

Hence, $\hat{f} = 0$. □

Proof of Theorem 5.5. Let $\mathcal{T}(n) = T_1^{\varphi_1(n)} \cdots T_r^{\varphi_r(n)}$, $n \in \mathbb{Z}$, where $T_i \in G$ and φ_i are GL-functions. By Lemma 2.2, for each $j = 1, \ldots, r$ one has $\varphi_j(n) = a_j n + \psi_j(n)$, $n \in \mathbb{Z}$, where $a_j \in \mathbb{R}$ and ψ_j is a BGL-function. Considering $T_1, \ldots, T_r$ as unitary operators on $\mathcal{H}_{\text{wm}}$, embed them (by utilizing the spectral theorem) into commuting continuous unitary flows $(T_i^t)_{t\in\mathbb{R}}$, and let $T = T_1^{a_1} \cdots T_r^{a_r}$.

Based on Lemma 5.6, we are going to show that T has no eigenvectors in $\mathcal{H}_{\text{wm}}$. Assume, in the course of contradiction, that there exists $f \in \mathcal{H}_{\text{wm}}$, with $\|f\| = 1$, such that $Tf = \lambda f$. Let $\mathcal{S}(n) = T_1^{\psi_1(n)} \cdots T_r^{\psi_r(n)}$, so that $\mathcal{T}(n) = T^n \mathcal{S}(n)$, $n \in \mathbb{Z}$. Fix $\varepsilon > 0$. Let I be an interval in $\mathbb{R}^r$ that contains the range $\psi(\mathbb{Z})$ of the function $\psi = (\psi_1, \ldots, \psi_r)$. Partition I to subintervals $I_1, \ldots, I_l$ small enough so that for each $i = 1, \ldots, l$, for some $f_i \in \mathcal{H}_{\text{wm}}$, one has $(T_1^{z_1} \cdots T_r^{z_r})^{-1} f \overset{\varepsilon}{\approx} f_i$ for all $(z_1, \ldots, z_r) \in I_i$. (Here and below, for $g_1, g_2 \in L^2(X)$, '$g_1 \overset{\varepsilon}{\approx} g_2$' means that $\|g_1 - g_2\| \leq \varepsilon$.) For each $i = 1, \ldots, l$, let $A_i = \{n : \psi(n) \in I_i\}$; by Proposition 2.7, the indicator function 1_{A_i} is a UGL-function, and thus by Proposition 4.5 there exists a trigonometric polynomial q_i such that $\text{D-limsup}_n |1_{A_i}(n) - q_i(n)| < \varepsilon/l$. For any i, for any $n \in A_i$, $T^n f = \mathcal{T}(n)\mathcal{S}(n)^{-1} f \overset{\varepsilon}{\approx} \mathcal{T}(n) f_i$; thus,

$$\begin{aligned} \mathrm{d}(A_i) f = \operatorname*{C\text{-}lim}_n 1_{A_i}(n) \lambda^{-n} T^n f &\overset{\mathrm{d}(A_i)\varepsilon}{\approx} \operatorname*{C\text{-}lim}_n 1_{A_i}(n) \lambda^{-n} \mathcal{T}(n) f_i \\ &\overset{\varepsilon/l}{\approx} \operatorname*{C\text{-}lim}_n q_i(n) \lambda^{-n} \mathcal{T}(n) f_i. \end{aligned}$$

By Lemma 5.6, the last limit is equal to 0; summing this up for $i = 1, \ldots, l$ we get $f \overset{2\varepsilon}{\approx} 0$. Since ε is arbitrary, $f = 0$.

Hence, T is weakly mixing on $\mathcal{H}_{\text{wm}}$, that is, for any $f, g \in \mathcal{H}_{\text{wm}}$, $\text{D-lim}_n \langle T^n f, g\rangle = 0$. Let $f, g \in \mathcal{H}_{\text{wm}}$ and let $\varepsilon > 0$. The set $\{\mathcal{S}(n)^{-1} g,\ n \in \mathbb{Z}\}$ is totally bounded; let $\{g_1, \ldots, g_k\}$ be an ε-net in this set. Then

$$\operatorname*{D\text{-}limsup}_n |\langle \mathcal{T}(n) f, g\rangle| = \operatorname*{D\text{-}limsup}_n |\langle T^n f, \mathcal{S}(n)^{-1} g\rangle| < \operatorname*{D\text{-}lim}_n \max_i |\langle T^n f, g_i\rangle| + \varepsilon = \varepsilon.$$

Since ε is arbitrary, $\text{D-lim}_n \langle \mathcal{T}(n) f, g\rangle = 0$. □

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

6. *Joint ergodicity of several GL-sequences of transformations*

We now start dealing with several GL-sequences of measure-preserving transformations. We preserve the notation G, $\mathfrak{T}$, $\mathcal{H}_{\mathrm{c}}$, and $\mathcal{H}_{\mathrm{wm}}$ from the preceding section.

Given functions $f_1, \dots, f_k$ on X, the tensor product $\bigotimes_{i=1}^k f_i = f_1 \otimes \cdots \otimes f_k$ is the function $f(x_1, \dots, x_k) = f_1(x_1) \cdots f_k(x_k)$ on X^k (whereas the product $\prod_{i=1}^k f_i = f_1 \cdots f_k$ is the function $f_1(x) \cdots f_k(x)$ on X).

LEMMA 6.1. *If $\mathcal{T}_1, \dots, \mathcal{T}_k \in \mathfrak{T}$ are ergodic, then for any functions $f_1, \dots, f_k \in L^\infty(X)$ with $f_i \in \mathcal{H}_{\mathrm{wm}}$ for at least one i,* $\operatorname{C-lim}_n \bigotimes_{i=1}^k \mathcal{T}_i(n) f_i = 0$ *in* $L^2(X^k)$.

Proof. Assume that $f_1 \in \mathcal{H}_{\mathrm{wm}}$. Let $\hat{f} = \operatorname{C-lim}_n \bigotimes_{i=1}^k \mathcal{T}_i(n) f_i$ (the limit exists by Lemma 5.2). For any $g_1, \dots, g_k \in L^\infty(X)$, we have

$$\left\langle \hat{f}, \bigotimes_{i=1}^k g_i \right\rangle = \operatorname*{C-lim}_n \int_{X^k} \bigotimes_{i=1}^k \mathcal{T}_i(n) f_i \cdot \bigotimes_{i=1}^k \bar{g}_i \, d\mu^k = \operatorname*{C-lim}_n \prod_{i=1}^k \int_X \mathcal{T}_i(n) f_i \cdot \bar{g}_i \, d\mu = 0,$$

since $\operatorname{D-lim}_n \int_X \mathcal{T}_1(n) f_1 \cdot \bar{g}_1 \, d\mu = 0$ by Theorem 5.5. Since the functions of the form $\bigotimes_{i=1}^k g_i$ are dense in $L^2(X^k)$, $\hat{f} = 0$. □

Given transformations $T_1, \dots, T_k$ of X, $T_1 \times \cdots \times T_k$ is the transformation of X^k defined by $(T_1 \times \cdots \times T_k)(x_1, \dots, x_k) = (T_1 x_1, \dots, T_k x_k)$. Notice that if $\mathcal{T}_1, \dots, \mathcal{T}_k$ are sequences of transformations of X such that $\mathcal{T}_1 \times \cdots \times \mathcal{T}_k$ is ergodic, then $\mathcal{T}_1, \dots, \mathcal{T}_k$ are ergodic and, moreover, $\mathcal{T}_{i_1} \times \cdots \times \mathcal{T}_{i_l}$ is ergodic for any $1 \le i_1 < \cdots < i_l \le k$.

LEMMA 6.2. *Let $\mathcal{T}_1, \dots, \mathcal{T}_k \in \mathfrak{T}$ be such that the GL-sequences $\mathcal{T}_1^{-1}\mathcal{T}_2, \dots, \mathcal{T}_1^{-1}\mathcal{T}_k$ of transformations of X are ergodic and the GL-sequence $\mathcal{T}_1 \times \cdots \times \mathcal{T}_k$ of transformations of X^k is ergodic. Then the GL-sequence $(\mathcal{T}_1^{-1}\mathcal{T}_2) \times \cdots \times (\mathcal{T}_1^{-1}\mathcal{T}_k)$ of transformations of X^{k-1} is also ergodic.*

Proof. Since the span of the functions of the form $f_2 \otimes \cdots \otimes f_k$ with $f_2, \dots, f_k \in L^\infty(X)$ is dense in $L^2(X^{k-1})$, it suffices to show that for any $f_2, \dots, f_k \in L^\infty(X)$ with $\int_X f_2 \, d\mu = \cdots = \int_X f_k \, d\mu = 0$, one has

$$\operatorname*{C-lim}_n \bigotimes_{i=2}^k \mathcal{T}_1(n)^{-1} \mathcal{T}_i(n) f_i = 0$$

in $L^2(X^{k-1})$. If at least one of f_i is in $\mathcal{H}_{\mathrm{wm}}$, this is true by Lemma 6.1. If $f_2, \dots, f_k \in \mathcal{H}_{\mathrm{c}}$, we may assume that $f_2, \dots, f_k$ are non-constant eigenfunctions of the elements of G, so that $\mathcal{T}_1(n) f_i = \tau_i(n) f_i$ and $\mathcal{T}_i(n) f_i = \lambda_i(n) f_i$, $i = 2, \dots, k$, for some (multiplicative) GL-sequences τ_i, λ_i in $\{z \in \mathbb{C} : |z| = 1\}$. Put $f_1 = \overline{f_2 \cdots f_k}$. Since $\mathcal{T}_1 \times \cdots \times \mathcal{T}_k$ is ergodic, we have $\operatorname{C-lim}_n \bigotimes_{i=1}^k \mathcal{T}_i(n) f_i = 0$. But

$$\begin{aligned}\operatorname*{C-lim}_n \bigotimes_{i=1}^k \mathcal{T}_i(n) f_i &= \operatorname*{C-lim}_n \Bigg(\prod_{i=2}^k \overline{\tau_i(n)} \bar{f}_i \Bigg) \otimes \bigotimes_{i=2}^k \lambda_i f_i \\ &= \operatorname*{C-lim}_n \prod_{i=2}^k \overline{\tau_i(n)} \lambda_i(n) \Bigg(\prod_{i=2}^k \bar{f}_i \bigotimes_{i=2}^k f_i \Bigg),\end{aligned}$$

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

so $\text{C-lim}_n \prod_{i=1}^k \overline{\tau_i(n)}\lambda_i(n) = 0$. Hence, $\text{C-lim}_n \bigotimes_{i=2}^k \mathcal{T}_1(n)^{-1}\mathcal{T}_i(n) f_i = \text{C-lim}_n \prod_{i=2}^k \overline{\tau_i(n)}\lambda_i(n)(\bigotimes_{i=2}^k f_i) = 0$ as well. □

DEFINITION 6.3. We say that sequences $\mathcal{T}_1, \dots, \mathcal{T}_k$ of measure-preserving transformations of X are *jointly ergodic* if for any $f_1, \dots, f_k \in L^\infty(X)$, $\text{C-lim}_n \prod_{i=1}^k \mathcal{T}_i(n) f_i = \prod_{i=1}^k \int_X f_i \, d\mu$ in $L^2(X)$.

Notice that if $\mathcal{T}_1, \dots, \mathcal{T}_k$ are jointly ergodic, then $\mathcal{T}_1, \dots, \mathcal{T}_k$ are ergodic and, moreover, $\mathcal{T}_{i_1}, \dots, \mathcal{T}_{i_l}$ are jointly ergodic for any $1 \le i_1 < \cdots < i_l \le k$.

We are now in position to prove our main result.

THEOREM 6.4. *GL-sequences $\mathcal{T}_1, \dots, \mathcal{T}_k \in \mathfrak{T}$ are jointly ergodic if and only if the GL-sequences $\mathcal{T}_i^{-1}\mathcal{T}_j$ are ergodic for all $i \ne j$ and the GL-sequence $\mathcal{T}_1 \times \cdots \times \mathcal{T}_k$ is ergodic.*

Proof. Assume that $\mathcal{T}_1, \dots, \mathcal{T}_k \in \mathfrak{T}$ are jointly ergodic. Let $i, j \in \{1, \dots, k\}$, $i \ne j$, let $f \in L^\infty(X)$, and let $\hat{f} = \text{C-lim}_n \mathcal{T}_i^{-1}(n)\mathcal{T}_j(n) f$. Then, for any $g \in L^\infty(X)$, we have

$$\langle \hat{f}, g\rangle = \text{C-lim}_n \int_X \mathcal{T}_i^{-1}(n)\mathcal{T}_j(n) f \cdot \bar{g} \, d\mu = \int_X \text{C-lim}_n \mathcal{T}_j(n) f \cdot \mathcal{T}_i(n)\bar{g} \, d\mu$$
$$= \int_X f \, d\mu \int_X \bar{g} \, d\mu.$$

Hence, $\hat{f} = \int_X f \, d\mu$, so $\mathcal{T}_i^{-1}\mathcal{T}_j$ is ergodic.

To prove that $\mathcal{T}_1 \times \cdots \times \mathcal{T}_k$ is also ergodic, it suffices to show that for any $f_1, \dots, f_k \in L^\infty(X)$ with $\int_X f_1 \, d\mu = \cdots = \int_X f_k \, d\mu = 0$, $\text{C-lim}_n \bigotimes_{i=1}^k \mathcal{T}_i(n) f_i = 0$. If at least one of f_i is in $\mathcal{H}_{\text{wm}}$, this is true by Lemma 6.1. If $f_i \in \mathcal{H}_{\text{c}}$ for all i, we may assume that $f_1, \dots, f_k$ are non-constant eigenfunctions of the elements of G, and then $\mathcal{T}_i(n) f_i = \lambda_i(n) f_i$, $i = 1, \dots, k$, for some GL-sequences $\lambda_1, \dots, \lambda_k$ in $\{z \in \mathbb{C} : |z| = 1\}$. In this case, both $\bigotimes_{i=1}^k \mathcal{T}_i(n) f_i = \lambda(n) \bigotimes_{i=1}^k f_i$ and $\prod_{i=1}^k \mathcal{T}_i(n) f_i = \lambda(n) \prod_{i=1}^k f_i$, where $\lambda(n) = \prod_{i=1}^k \lambda_i(n)$, $n \in \mathbb{Z}$. Since $\text{C-lim}_n \prod_{i=1}^k \mathcal{T}_i(n) f_i = 0$, we have $\text{C-lim}_n \lambda(n) = 0$ and, so, $\text{C-lim}_n \bigotimes_{i=1}^k \mathcal{T}_i(n) f_i = 0$.

Conversely, assume that $\mathcal{T}_i^{-1}\mathcal{T}_j$ are ergodic for all $i \ne j$ and $\mathcal{T}_1 \times \cdots \times \mathcal{T}_k$ is ergodic, and let $f_1, \dots, f_k \in L^\infty(X)$. If all $f_i \in \mathcal{H}_{\text{c}}$, then, again, we may assume that $f_1, \dots, f_k$ are non-constant eigenfunctions of the elements of G, so $\bigotimes_{i=1}^k \mathcal{T}_i(n) f_i = \lambda(n) \bigotimes_{i=1}^k f_i$ and $\prod_{i=1}^k \mathcal{T}_i(n) f_i = \lambda(n) \prod_{i=1}^k f_i$, and since now $\text{C-lim}_n \bigotimes_{i=1}^k \mathcal{T}_i(n) f_i = 0$, we obtain that $\text{C-lim}_n \prod_{i=1}^k \mathcal{T}_i(n) f_i = 0$ as well.

It remains to show that $\text{C-lim}_n \prod_{i=1}^k \mathcal{T}_i(n) f_i = 0$ whenever $f_i \in \mathcal{H}_{\text{wm}}$ for at least one i. We will assume that $f_1 \in \mathcal{H}_{\text{wm}}$ and that $|f_i| \le 1$ for all i. We will use the van der Corput trick and induction on k. Let $\varepsilon > 0$. Let a Bohr set $H \subseteq \mathbb{Z}$, transformations $S_1, \dots, S_k \in G$, and sets $E_h \subseteq \mathbb{Z}$, $h \in H$, be as in Proposition 5.1(iv), applied to the GL-sequences

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

$\mathcal{T}_1, \dots, \mathcal{T}_k$. Let $h_1, h_2 \in H$; for any $n \in E_{h_1} \cap E_{h_2}$, we have

$$\begin{aligned}
&\left\langle \prod_{i=1}^{k} \mathcal{T}_i(n+h_1) f_i, \prod_{i=1}^{k} \mathcal{T}_i(n+h_2) f_i \right\rangle \\
&\quad = \int_X \prod_{i=1}^{k} \mathcal{T}_i(n+h_1) f_i \cdot \prod_{i=1}^{k} \mathcal{T}_i(n+h_2) \bar{f}_i \, d\mu \\
&\quad = \int_X (\mathcal{T}_1(h_1) S_1 f_1 \cdot \mathcal{T}_1(h_2) S_1 \bar{f}_1) \cdot \prod_{i=2}^{k} \mathcal{T}_1^{-1}(n) \mathcal{T}_i(n) (\mathcal{T}_i(h_1) S_i f_i \cdot \mathcal{T}_i(h_2) S_i \bar{f}_i) \, d\mu.
\end{aligned}$$

By Lemma 6.2, $\mathcal{T}_1^{-1}\mathcal{T}_2 \times \cdots \times \mathcal{T}_1^{-1}\mathcal{T}_k$ is ergodic; thus, by induction on k, $\mathcal{T}_1^{-1}\mathcal{T}_2$, $\dots$, $\mathcal{T}_1^{-1}\mathcal{T}_k$ are jointly ergodic, so

$$\begin{aligned}
\text{C-lim}_n \int_X (\mathcal{T}_1(h_1) S_1 f_1 \cdot \mathcal{T}_1(h_2) S_1 \bar{f}_1) \cdot \prod_{i=2}^{k} \mathcal{T}_1^{-1}(n) \mathcal{T}_i(n) (\mathcal{T}_i(h_1) S_i f_i \cdot \mathcal{T}_i(h_2) S_i \bar{f}_i) d\mu \\
= \prod_{i=1}^{k} \int_X \mathcal{T}_i(h_1) S_i f_i \cdot \mathcal{T}_i(h_2) S_i \bar{f}_i \, d\mu.
\end{aligned}$$

Since $\mathrm{d}_*(E_{h_1} \cap E_{h_2}) > 1 - 2\varepsilon$ and $|f_1|, \dots, |f_k| \le 1$, we get

$$\left| \text{C-limsup}_n \left\langle \prod_{i=1}^{k} \mathcal{T}_i(n+h_1) f_i, \prod_{i=1}^{k} \mathcal{T}_i(n+h_2) f_i \right\rangle \right| \le \left| \int_X \mathcal{T}_1(h_1) \tilde{f} \cdot \mathcal{T}_1(h_2) \overline{\tilde{f}} \, d\mu \right| + 2\varepsilon,$$

where $\tilde{f} = S_1 f_1$. Since $\tilde{f} \in \mathcal{H}_{\mathrm{wm}}$, by Theorem 5.5, $\text{D-lim}_h \int_X \mathcal{T}_1(h) \tilde{f} \cdot f' \, d\mu = 0$ for any $f' \in \mathcal{H}_{\mathrm{wm}}$ and, since $\mathrm{d}_*(H) > 0$, we can construct an infinite subset B of H such that $|\int_X \mathcal{T}_1(h_1) \tilde{f} \cdot \mathcal{T}_1(h_2) \overline{\tilde{f}} \, d\mu| < \varepsilon$ for any distinct $h_1, h_2 \in B$. (Indeed, choose $b_1 \in H$ arbitrarily. The set $B_1 = \{h \in \mathbb{N} : |\int_X T^h \tilde{f} \cdot T^{b_1} \overline{\tilde{f}} \, d\mu| < \varepsilon\}$ has density one, so $B_1 \cap H \neq \emptyset$; choose $b_2 \in B_1 \cap H$. Next, the set $B_2 = \{h \in \mathbb{N} : |\int_X T^h \tilde{f} \cdot T^{b_2} \overline{\tilde{f}} \, d\mu| < \varepsilon\}$ also has density one, so there is $b_3 \in B_1 \cap B_2 \cap H$, etc. Then put $B = \{b_1, b_2, \dots\}$.) Then, for any distinct $h_1, h_2 \in B$, we have $|\text{C-limsup}_n \langle \prod_{i=1}^{k} \mathcal{T}_i(n+h_1) f_i, \prod_{i=1}^{k} \mathcal{T}_i(n+h_2) f_i \rangle| < 3\varepsilon$, and so, by Lemma 3.1, $\text{C-limsup}_{\|\cdot\|,n} \prod_{i=1}^{k} \mathcal{T}_i(n) f_i \le \sqrt{3\varepsilon}$. Since ε is arbitrary, $\text{C-lim}_n \prod_{i=1}^{k} \mathcal{T}_i(n) f_i = 0$. □

REMARK 6.5. We defined our C-lims with respect to a fixed Følner sequence in $\mathbb{Z}$. However, since, for any GL-sequence $\mathcal{T}$ of measure-preserving transformations of X and any $f \in L^2(X)$, C-lim $\mathcal{T}(n) f$ exists with respect to any Følner sequence, this limit is the same for all Følner sequences (since any two such sequences can be combined to produce a new one having them as subsequences). This implies that for $\mathcal{T}_1, \dots, \mathcal{T}_k \in \mathfrak{T}$, the condition that $\mathcal{T}_i^{-1}\mathcal{T}_j$ for all $i \neq j$ and $\mathcal{T}_1 \times \cdots \times \mathcal{T}_k$ are ergodic is independent of the choice of a Følner sequence in $\mathbb{Z}$. It now follows from Theorem 6.4 that the joint ergodicity of $\mathcal{T}_1, \dots, \mathcal{T}_k$ does not depend on the choice of the Følner sequence either—which was not *a priori* evident.

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

For GL-sequences based on a single transformation, that is, of the form $\mathcal{T}(n) = T^{\varphi(n)}$, we have a simple criterion of ergodicity. Recall that for a measure-preserving transformation T of X we defined

$$\operatorname{Eig} T = \{\lambda \in \mathbb{C}^* : Tf = \lambda f \text{ for some } f \in L^2(X)\}$$

and, for several measure-preserving transformations $T_1, \dots, T_k$ of X, $\operatorname{Eig}(T_1, \dots, T_k) = (\operatorname{Eig} T_1) \cdots (\operatorname{Eig} T_k)$.

LEMMA 6.6. *Let T be an invertible measure-preserving transformation of X and let φ be an unbounded GL-function $\mathbb{Z} \longrightarrow \mathbb{Z}$. Then the GL-sequence $\mathcal{T}(n) = T^{\varphi(n)}$, $n \in \mathbb{Z}$, is ergodic if and only if T is ergodic and* $\text{C-lim}_n \lambda^{\varphi(n)} = 0$ *for every $\lambda \in \operatorname{Eig} T \setminus \{1\}$, and $\mathcal{T}$ is weakly mixing if and only if T is weakly mixing.*

Proof. The 'only if' part is clear. Let $L^2(X) = \mathcal{H}_c \oplus \mathcal{H}_{wm}$ be the compact/weak mixing decomposition induced by T. If T is ergodic and $\text{C-lim}_n \lambda^{\varphi(n)} = 0$ for every $\lambda \in \operatorname{Eig} T \setminus \{1\}$, then $\text{C-lim}_n \mathcal{T}(n) f = \int_X f \, d\mu$ for any $f \in \mathcal{H}_c$, so $\mathcal{T}$ is ergodic on $\mathcal{H}_c$. Considering T as a unitary operator on $\mathcal{H}_{wm}$ and using the spectral theorem, embed T into a continuous unitary flow. Using Lemma 2.2, write $\varphi(n) = an + \psi(n)$, $n \in \mathbb{Z}$, where $a \in \mathbb{R} \setminus \{0\}$ and ψ is a BGL-function; then T^a has no eigenfunctions in $\mathcal{H}_{wm}$, and so is weakly mixing on $\mathcal{H}_{wm}$. Since, for any $f \in \mathcal{H}_{wm}$, the set $\{T^{\psi(n)} f, \ n \in \mathbb{Z}\}$ is totally bounded, $\mathcal{T}$ is weakly mixing (and, so, ergodic) on $\mathcal{H}_{wm}$. □

We are now in position to prove Corollaries 1.5 and 1.6, formulated in the Introduction.

COROLLARY 6.7. *Let T be an invertible weakly mixing measure-preserving transformation of X and let $\varphi_1, \dots, \varphi_k$ be unbounded GL-functions such that $\varphi_j - \varphi_i$ are unbounded for all $i \neq j$; then the GL-sequences $T^{\varphi_1(n)}, \dots, T^{\varphi_k(n)}$, $n \in \mathbb{Z}$, are jointly ergodic. In particular, for any distinct $\alpha_1, \dots, \alpha_k \in \mathbb{R} \setminus \{0\}$, the GL-sequences $T^{[\alpha_1 n]}, \dots, T^{[\alpha_k n]}$ are jointly ergodic.*

Proof. By Lemma 6.6, the GL-sequences $T^{-\varphi_i(n)} T^{\varphi_j(n)} = T^{\varphi_j(n) - \varphi_i(n)}$ are ergodic for all $i \neq j$. The GL-sequence $T^{\varphi_1(n)} \times \cdots \times T^{\varphi_k(n)}$ is also ergodic by Lemma 6.1. By Theorem 6.4, $T^{\varphi_1(n)}, \dots, T^{\varphi_k(n)}$ are jointly ergodic. □

COROLLARY 6.8. *Let $T_1, \dots, T_k$ be commuting invertible jointly ergodic measure-preserving transformations of X and let φ be an unbounded GL-function $\mathbb{Z} \longrightarrow \mathbb{Z}$; then the GL-sequences $T_1^{\varphi(n)}, \dots, T_k^{\varphi(n)}$, $n \in \mathbb{Z}$, are jointly ergodic if and only if* $\text{C-lim}_n \lambda^{\varphi(n)} = 0$ *for every $\lambda \in \operatorname{Eig}(T_1, \dots, T_k) \setminus \{1\}$. In particular, for any irrational $\alpha \in \mathbb{R}$, the GL-sequences $T_1^{[\alpha n]}, \dots, T_k^{[\alpha n]}$ are jointly ergodic if and only if $e^{2\pi i \alpha^{-1} \mathbb{Q}} \cap \operatorname{Eig}(T_1, \dots, T_k) = \{1\}$.*

Proof. First of all, notice that $\operatorname{Eig}(T_1, \dots, T_k) = \operatorname{Eig}(T_1 \times \cdots \times T_k)$. Since the transformations $T_1, \dots, T_k$ are ergodic and commute, they share the set of eigenfunctions, so for any i and j we have $\operatorname{Eig}(T_i^{-1} T_j) \subseteq \operatorname{Eig} T_i \cdot \operatorname{Eig} T_j \subseteq \operatorname{Eig}(T_1, \dots, T_k)$ as well. Applying Lemma 6.6 to the transformations $T_i^{-1} T_j$ for $i \neq j$ and $T_1 \times \cdots \times T_k$, we get the first assertion.

The case $\varphi(n) = [\alpha n]$, with an irrational α, is now managed by the following lemma.

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

LEMMA 6.9. *For an irrational α and a real β one has* $\text{C-lim}_n e^{2\pi i[\alpha n]\beta} = 0$ *if and only if* $\alpha\beta \notin \mathbb{Z}\alpha + \mathbb{Q}$.

Proof. We have $[\alpha n]\beta = \alpha\beta n - \{\alpha n\}\beta$, $n \in \mathbb{Z}$. Consider the sequence $u_n = (\{\alpha\beta n\}, \{\alpha n\})$ in the torus $\mathbb{T}^2_{(x,y)} = \mathbb{R}^2/\mathbb{Z}^2$, so that the sequence $([\alpha n]\beta) \mod 1$ is its image in $\mathbb{T}$ under the map $\sigma(x, y) = (\{x\} - \beta\{y\}) \mod 1$. If $\alpha\beta$ and α are rationally independent modulo 1, (u_n) is uniformly distributed in $\mathbb{T}^2$; thus, $\sigma(u_n)$ is uniformly distributed in $\mathbb{T}$, and so $\text{C-lim}_n e^{2\pi i[\alpha n]\beta} = \text{C-lim}_n e^{2\pi i\sigma(u_n)} = 0$. Let $\alpha\beta$ and α be rationally dependent modulo 1, $k\alpha\beta = m\alpha + l$, where $k \in \mathbb{N}$ and $m, l \in \mathbb{Z}$ with $\gcd(k, m, l) = 1$. Then the sequence (u_n) is uniformly distributed in the subgroup S of $\mathbb{T}^2$ defined by the equation $kx = my$. σ maps S to k isomorphic intervals $[(m/k)j, (m/k)(j+1) - \beta)$, $j = 0, \ldots, k-1$, in $\mathbb{T}$, and the sequence $\sigma(u_n)$ is uniformly distributed in the weighted union of these intervals. It follows that $\text{C-lim}_n e^{2\pi i\sigma(u_n)} = 0$ unless all the intervals coincide, which happens if and only if $k \mid m$, that is, if and only if $\alpha\beta \in \mathbb{Z}\alpha + \mathbb{Q}$. In this situation of a single interval it could still be possible that $\text{C-lim}_n e^{2\pi i\sigma(u_n)} = 0$—if this interval covered $\mathbb{T}$ an integer number of times, that is, if and only if $(m/k) - \beta \in \mathbb{Z}\backslash\{0\}$; however, this is never the case since α is irrational. Thus, if $\alpha\beta \in \mathbb{Z}\alpha + \mathbb{Q}$, $\text{C-lim}_n e^{2\pi i\sigma(u_n)} \neq 0$. □

7. *Joint ergodicity of GL-sequences along primes*

In this section, we will adapt some results from [**GT**] and a technique from [**S**] to establish a condition for several GL-sequences to be jointly ergodic along primes.

By $\mathcal{P}$, we will denote the set of prime integers. Let us also use the following notation: for $N \in \mathbb{N}$, let $\mathcal{P}(N) = \mathcal{P} \cap \{1, \ldots, N\}$, $\pi(N) = |\mathcal{P}(N)|$, and $R(N) = \{r \in \{1, \ldots, N\} : \gcd(r, N) = 1\}$. As above, we fix a commutative group G of measure-preserving transformations of a probability measure space (X, μ) and denote by $\mathfrak{T}$ the group of GL-sequences in G. We will prove the following theorem.

THEOREM 7.1. *Let $\mathcal{T}_1, \ldots, \mathcal{T}_k \in \mathfrak{T}$ be such that for any $W \in \mathbb{N}$ and $r \in R(W)$ the GL-sequences $\mathcal{T}_{i,W,r}(n) = \mathcal{T}_i(Wn + r)$, $i = 1, \ldots, k$, are jointly ergodic. Then, for any $f_1, \ldots, f_k \in L^\infty(X)$,*

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p\in\mathcal{P}(N)} \mathcal{T}_1(p) f_1 \cdots \mathcal{T}_k(p) f_k = \prod_{i=1}^{k} \int_X f_i \, d\mu \quad \textit{(in } L^2 \textit{ norm).}$$

Remark. In general, joint ergodicity of $\mathcal{T}_i$, $i = 1, \ldots, k$, does not imply that of $\mathcal{T}_{i,W,r}$. Indeed, let $\alpha \in \mathbb{R}\backslash\mathbb{Q}$, let $X = \{0, 1\}$ with measure $\mu(\{0\}) = \mu(\{1\}) = 1/2$, let $Tx = (x + 1) \mod 2$, let $\mathcal{T}_1(n) = T^n$, and let $\mathcal{T}_2(n) = T^{[\alpha n]}$, $n \in \mathbb{N}$. Then $\mathcal{T}_1$ and $\mathcal{T}_2$ are jointly ergodic, but $\mathcal{T}_1(2n + 1)$ is not ergodic. Notice also that the assertion of the theorem does not hold for these $\mathcal{T}_1$ and $\mathcal{T}_2$: for functions f_1 and f_2 on X, $\lim_{N\to\infty} (1/\pi(N)) \sum_{p\in\mathcal{P}(N)} \mathcal{T}_1(p) f_1 \cdot \mathcal{T}_2(p) f_2 = Tf_1 \int_X f_2 \, d\mu$.

Following [**GT**], we introduce 'the modified von Mangoldt function' $\Lambda'(n) = 1_{\mathcal{P}}(n) \log n$, $n \in \mathbb{N}$. The following simple lemma allows one to rewrite the average in Theorem 7.1 in terms of Λ'.

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

LEMMA 7.2. (Cf. [**FHoK**, Lemma 1]) *For any bounded sequence* (v_n) *of vectors in a normed vector space,*

$$\lim_{N\to\infty}\left\|\frac{1}{\pi(N)}\sum_{p\in\mathcal{P}(N)}v_p-\frac{1}{N}\sum_{n=1}^{N}\Lambda'(n)v_n\right\|=0.$$

A (compact) *nilmanifold* $\mathcal{N}$ is a compact homogeneous space of a nilpotent Lie group $\mathcal{G}$; a *nilrotation* of $\mathcal{N}$ is a translation by an element of $\mathcal{G}$. Nilmanifolds are characterized by the nilpotency class and the number of generators of $\mathcal{G}$; for any $k, d\in\mathbb{N}$, there exists a universal, 'free' nilmanifold $\mathcal{N}_{k,d}$ of nilpotency class k, with d 'continuous' and d 'discrete' generators† such that any nilmanifold of class $\le k$ and with $\le d$ generators is a factor of $\mathcal{N}_{k,d}$. *A basic nilsequence* is a sequence of the form $\eta(n)=g(a^n)$, where g is a continuous function on a nilmanifold $\mathcal{N}$ and a is a nilrotation of $\mathcal{N}$. We may always assume that $\mathcal{N}=\mathcal{N}_{k,d}$ for some k and d; the minimal such k is said to be the nilpotency class of η. Given $k, d\in\mathbb{N}$ and $M>0$, we will denote by $\mathcal{L}_{k,d,M}$ the set of basic nilsequences $\eta(n)=g(a^n)$, where the function $g\in C(\mathcal{N}_{k,d})$ is Lipschitz with constant M and $|g|\le M$. (A smooth metric on each nilmanifold $\mathcal{N}_{k,d}$ is assumed to be chosen.)

Following [**GT**], for $W, r\in\mathbb{N}$ we define $\Lambda'_{W,r}(n)=(\phi(W)/W)\Lambda'(Wn+r)$, $n\in\mathbb{N}$, where ϕ is the Euler function, $\phi(W)=|R(W)|$. By $\mathcal{W}$, we will denote the set of integers of the form $W=\prod_{p\in\mathcal{P}(m)}p$, $m\in\mathbb{N}$. It is proved in [**GT**] that 'the W-tricked von Mangoldt sequences $\Lambda'_{W,r}$ are orthogonal to nilsequences'; here is a weakened version of [**GT**, Proposition 10.2].

PROPOSITION 7.3. *For any* $k, d\in\mathbb{N}$ *and* $M>0$,

$$\lim_{\substack{W\in\mathcal{W}\\W\to\infty}}\limsup_{N\to\infty}\sup_{\substack{\eta\in\mathcal{L}_{k,d,M}\\r\in R(W)}}\left|\frac{1}{N}\sum_{n=1}^{N}(\Lambda'_{W,r}(n)-1)\eta(n)\right|=0.$$

We need to extend Proposition 7.3 to sequences slightly more general than nilsequences.

LEMMA 7.4. (Cf. [**S**, Proposition 3.2]) *Let* P *be a polygonal subset of a torus* $\mathcal{M}$ *and let* $u\in\mathcal{M}$. *For any* $k, d\in\mathbb{N}$ *and* $M>0$,

$$\lim_{\substack{W\in\mathcal{W}\\W\to\infty}}\limsup_{N\to\infty}\sup_{\substack{\eta\in\mathcal{L}_{k,d,M}\\r\in R(W)}}\left|\frac{1}{N}\sum_{n=1}^{N}(\Lambda'_{W,r}(n)-1)1_P((Wn+r)u)\eta(n)\right|=0.$$

Proof. Let $Z=\overline{\mathbb{Z}u}$; Z is a finite union of subtori $\mathcal{M}_1,\dots,\mathcal{M}_l$ of $\mathcal{M}$. Let $\varepsilon>0$. Choose smooth functions g_1, g_2 on $\mathcal{M}$ such that $0\le g_1\le 1_P\le g_2\le 1$ and the set $S=\{g_1\ne g_2\}$ is polygonal with $\Lambda_{\mathcal{M}_i}(S\cap\mathcal{M}_i)\le\varepsilon\Lambda_{\mathcal{M}_i}(\mathcal{M}_i)$, $i=1,\dots,l$, where $\Lambda_{\mathcal{M}_i}$ is the normalized Haar measure on $\mathcal{M}_i$. Then, for any W and r, the sequences $\zeta_{1,W,r}(n)=g_1((Wn+r)u)$ and $\zeta_{2,W,r}(n)=g_2((Wn+r)u)$ are (one-step) basic nilsequences and, since the sequence $(Wn+r)u$ is uniformly distributed in the union of several components of Z, the set $\{n:\zeta_1(n)\ne\zeta_2(n)\}$ has density $<\varepsilon$. Notice that if M' is the sum of M and the

† 'A continuous generator' of a nilmanifold $\mathcal{N}$ is a continuous flow $(a^t)_{t\in\mathbb{R}}$ in the group $\mathcal{G}$; 'a discrete generator' is just an element of $\mathcal{G}$.

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

Lipschitz constants of g_1 and g_2 and $d' = d + \dim \mathcal{M}$, then for any $\eta \in \mathcal{L}_{k,d,M}$ one has $\zeta_{1,W,r}\eta,\ \zeta_{2,W,r}\eta \in \mathcal{L}_{k,d',M'}$.

Let $\mathcal{L}^+_{k,d,M} = \{\eta \in \mathcal{L}_{k,d,M} : \eta \geq 0\}$. For any W, r, and any $\eta \in \mathcal{L}^+_{k,d,M}$, for every $n \in \mathbb{N}$ we have

$$\begin{aligned}&(\Lambda'_{W,r}(n) - 1)1_P((Wn + r)u)\eta(n)\\ &\quad\leq (\Lambda'_{W,r}(n)\zeta_{2,W,r}(n) - \zeta_{1,W,r}(n))\eta(n)\\ &\quad= (\Lambda'_{W,r}(n) - 1)\zeta_{2,W,r}(n)\eta(n) + (\zeta_{2,W,r}(n) - \zeta_{1,W,r}(n))\eta(n).\end{aligned}$$

By Proposition 7.3,

$$\lim_{\substack{W\in\mathcal{W}\\ W\to\infty}} \limsup_{N\to\infty} \sup_{\substack{\eta\in\mathcal{L}_{k,d,M}\\ r\in R(W)}} \left|\frac{1}{N}\sum_{n=1}^{N}(\Lambda'_{W,r}(n) - 1)\zeta_{2,W,r}(n)\eta(n)\right| = 0,$$

whereas, for any W, r, and any $\eta \in \mathcal{L}^+_{k,d,M}$,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N}\left|\zeta_{2,W,r}(n) - \zeta_{1,W,r}(n)\right|\eta(n) \leq M\varepsilon;$$

thus,

$$\limsup_{\substack{W\in\mathcal{W}\\ W\to\infty}} \limsup_{N\to\infty} \sup_{\substack{\eta\in\mathcal{L}^+_{k,d,M}\\ r\in R(W)}} \frac{1}{N}\sum_{n=1}^{N}(\Lambda'_{W,r}(n) - 1)1_P((Wn + r)u)\eta(n) \leq M\varepsilon.$$

Similarly,

$$\begin{aligned}&(\Lambda'_{W,r}(n) - 1)1_P((Wn + r)u)\eta(n)\\ &\quad\geq (\Lambda'_{W,r}(n) - 1)\zeta_{1,W,r}(n)\eta(n) - (\zeta_{2,W,r}(n) - \zeta_{1,W,r}(n))\eta(n),\end{aligned}$$

so

$$\liminf_{\substack{W\in\mathcal{W}\\ W\to\infty}} \liminf_{N\to\infty} \inf_{\substack{\eta\in\mathcal{L}^+_{k,d,M}\\ r\in R(W)}} \frac{1}{N}\sum_{n=1}^{N}(\Lambda'_{W,r}(n) - 1)1_P((Wn + r)u)\eta(n) \geq -M\varepsilon.$$

Hence,

$$\lim_{\substack{W\in\mathcal{W}\\ W\to\infty}} \limsup_{N\to\infty} \sup_{\substack{\eta\in\mathcal{L}^+_{k,d,M}\\ r\in R(W)}} \left|\frac{1}{N}\sum_{n=1}^{N}(\Lambda'_{W,r}(n) - 1)1_P((Wn + r)u)\eta(n)\right| = 0;$$

since $\mathcal{L}_{k,d,M} = \mathcal{L}^+_{k,d,M} - \mathcal{L}^+_{k,d,M}$, we are done. □

Let $k,\ N \in \mathbb{N}$; for sequences $b : \{1, \ldots, N\} \longrightarrow \mathbb{R}$, *the* kth *Gowers norm* is defined by

$$\begin{aligned}\|b\|_{U^k[N]} &= \left(\frac{1}{V_k(N)}\sum_{n,h_1,\ldots,h_k\in\mathbb{Z}}\prod_{e_1,\ldots,e_k\in\{0,1\}} b(n + e_1h_1 + \cdots + e_kh_k)\right)^{1/2^k}\\ &= \left(\frac{N^2}{V_k(N)}\sum_{h_1,\ldots,h_{k-1}\in\mathbb{Z}}\left|\frac{1}{N}\sum_{n=1}^{N}\prod_{e_1,\ldots,e_{k-1}\in\{0,1\}} b(n + e_1h_1 + \cdots + e_{k-1}h_{k-1})\right|^2\right)^{1/2^k},\end{aligned}$$

where

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

$$V_k(N) = \#\{(n, h_1, \ldots, n_k) \in \mathbb{Z}^{k+1} : n + e_1h_1 + \cdots + e_{k-1}h_{k-1} \in [1, N] \text{ for any } e_1, \ldots, e_k \in \{0, 1\}\} \le N^{k+1},$$

and where we assume that $b(n) = 0$ if $n \notin \{1, \ldots, N\}$.

The next result we need is the fact that, on a certain class of sequences, 'the kth Gowers norm is continuous with respect to the system of seminorms $\|b\|_{d,M} = \sup_{\eta \in \mathcal{L}_{k,d,M}} |(1/N) \sum_{n=1}^{N} b(n)\eta(n)|$, $d \in \mathbb{N}$, $M > 0$'. We will utilize the following lemma, which is a corollary of [**GT**, Propositions 10.1 and 6.4].

LEMMA 7.5. *For any $\varepsilon > 0$ and $k \in \mathbb{N}$, there exist $d \in \mathbb{N}$, $M > 0$, and $\delta > 0$ such that for any $N \in \mathbb{N}$, if a sequence $b : \{1, \ldots, N\} \longrightarrow \mathbb{R}$ satisfies $|b| \le 1 + \Lambda'_{W,r}$ for some $W \in \mathcal{W}$ and $r \in R(W)$ and $\sup_{\eta \in \mathcal{L}_{k,d,M}} |(1/N) \sum_{n=1}^{N} b(n)\eta(n)| < \delta$, then $\|b\|_{U^k[N]} < \varepsilon$.*

Remark. Proposition 10.1 was proved in [**GT**] modulo the 'inverse Gowers-norm conjecture', which had then been confirmed in [**GTZ**].

Combining Lemmas 7.4 and 7.5, applied to the sequence

$$b(n) = (\Lambda'_{W,r}(n) - 1) \cdot 1_P((Wn + r)u)1_{[1,N]}(n),$$

we obtain the following result.

LEMMA 7.6. (Cf. [**S**, Proposition 3.2]) *Let P be a polygonal region in a torus $\mathcal{M}$ and let $u \in \mathcal{M}$. Then, for any $k \in \mathbb{N}$,*

$$\lim_{\substack{W \in \mathcal{W} \\ W \to \infty}} \limsup_{N \to \infty} \max_{r \in R(W)} \|(\Lambda'_{W,r}(n) - 1)1_P((Wn + r)u)\|_{U^k[N]} = 0.$$

From Lemma 7.6, we now deduce the following result.

PROPOSITION 7.7. (Cf. [**S**, Proposition 4.1]) *Let $\mathcal{T}_1, \ldots, \mathcal{T}_k \in \mathfrak{T}$ and $f_1, \ldots, f_k \in L^\infty(X)$. For any $f_1, \ldots, f_k \in L^\infty(X)$, we have*

$$\lim_{\substack{W \in \mathcal{W} \\ W \to \infty}} \limsup_{N \to \infty} \max_{r \in R(W)} \left\| \frac{1}{N} \sum_{n=1}^{N} (\Lambda'_{W,r}(n) - 1)\mathcal{T}_1(Wn + r)f_1 \cdots \mathcal{T}_k(Wn + r)f_k \right\|_{L^2(X)} = 0.$$

Proof. We will assume that $|f_i| \le 1$, $i = 1, \ldots, k$. Let, by Proposition 5.1(iii), $\mathcal{M}$ be a torus, $u \in \mathcal{M}$, and $\mathcal{M} = \bigcup_{j=1}^{l} P_j$ be a polygonal partition of $\mathcal{M}$ such that for every i and j, $\mathcal{T}_i(n)^{-1}\mathcal{T}_i(n + h)$ does not depend on n whenever both nu, $(n + h)u \in P_j$. We will show that for any j and any N, W, $r \in \mathbb{N}$, one has

$$\left\| \frac{1}{N} \sum_{n=1}^{N} (\Lambda'_{W,r}(n) - 1)1_{P_j}((Wn + r)u)\mathcal{T}_1(Wn + r)f_1 \cdots \mathcal{T}_k(Wn + r)f_k \right\|_{L^2(X)} \le \|(\Lambda'_{W,r}(n) - 1)1_{P_j}((Wn + r)u)\|_{U^{k+1}[N]};$$

via Lemma 7.6, this will imply that

$$\lim_{\substack{W \in \mathcal{W} \\ W \to \infty}} \limsup_{N \to \infty} \max_{r \in R(W)} \left\| \frac{1}{N} \sum_{n=1}^{N} (\Lambda'_{W,r}(n) - 1)1_{P_j}((Wn + r)u)\mathcal{T}_1(Wn + r)f_1 \cdots \mathcal{T}_k(Wn + r)f_k \right\|_{L^2(X)} = 0$$

for each $j = 1, \ldots, l$, from which the assertion of the proposition follows.

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

Fix j, N, W, and r; put $P = P_j$, define the sequence $b(n) = (\Lambda'_{W,r}(n) - 1) \cdot 1_P((Wn + r)u)1_{[1,N]}(n)$, and let $\widetilde{\mathcal{T}}_i(n) = \mathcal{T}_i(Wn + r)$ for $i = 1, \ldots, k$.

By Lemma 3.2,

$$\left\| \frac{1}{N} \sum_{n=1}^{N} b(n)\widetilde{\mathcal{T}}_1(n) f_1 \cdots \widetilde{\mathcal{T}}_k(n) f_k \right\|^2_{L^2(X)}$$
$$\leq \frac{1}{N} \sum_{h \in \mathbb{Z}} \left| \frac{1}{N} \sum_{n=1}^{N} \int_X b(n)b(n + h_1) \cdot \widetilde{\mathcal{T}}_1(n) f_1 \cdot \widetilde{\mathcal{T}}_1(n + h_1) \bar{f}_1 \right.$$
$$\left. \cdots \widetilde{\mathcal{T}}_k(n) f_k \cdot \widetilde{\mathcal{T}}_k(n + h_1) \bar{f}_k \, d\mu \right|.$$

By the definition of $P = P_j$, for each i there exists a sequence $\mathcal{S}_i(h_1)$, $h_1 \in \mathbb{Z}$, of transformations such that $\widetilde{\mathcal{T}}_i(n)^{-1}\widetilde{\mathcal{T}}_i(n + h_1) = \mathcal{S}_i(h_1)$ if $1_P((Wn + r)u) \cdot 1_P((W(n + h_1) + r)u) \neq 0$, and so, if $b(n)b(n + h_1) \neq 0$. Thus, if we put $f_{i,h_1} = f_i \cdot \mathcal{S}_i(h_1)\bar{f}_i$, $i = 1, \ldots, k$, $h_1 \in \mathbb{Z}$, we get

$$\left\| \frac{1}{N} \sum_{n=1}^{N} b(n)\widetilde{\mathcal{T}}_1(n) f_1 \cdots \widetilde{\mathcal{T}}_k(n) f_k \right\|^2_{L^2(X)}$$
$$\leq \frac{1}{N} \sum_{h_1 \in \mathbb{Z}} \left| \frac{1}{N} \sum_{n=1}^{N} \int_X b(n)b(n + h_1) \cdot \widetilde{\mathcal{T}}_1(n) f_{1,h_1} \cdots \widetilde{\mathcal{T}}_k(n) f_{k,h_1} \, d\mu \right|$$
$$= \frac{1}{N} \sum_{h_1 \in \mathbb{Z}} \left| \int_X f_{1,h_1} \frac{1}{N} \sum_{n=1}^{N} b(n)b(n + h_1) \cdot (\widetilde{\mathcal{T}}_1^{-1}\widetilde{\mathcal{T}}_2)(n) f_{2,h_1} \cdots (\widetilde{\mathcal{T}}_1^{-1}\widetilde{\mathcal{T}}_k)(n) f_{k,h_1} \, d\mu \right|$$
$$\leq \frac{1}{N} \sum_{h_1 \in \mathbb{Z}} \left\| \frac{1}{N} \sum_{n=1}^{N} b(n)b(n + h_1) \cdot (\widetilde{\mathcal{T}}_1^{-1}\widetilde{\mathcal{T}}_2)(n) f_{2,h_1} \cdots (\widetilde{\mathcal{T}}_1^{-1}\widetilde{\mathcal{T}}_k)(n) f_{k,h_1} \right\|_{L^2(X)}.$$

In the same way, for every h_1,

$$\left\| \frac{1}{N} \sum_{n=1}^{N} b(n)b(n + h_1) \cdot (\widetilde{\mathcal{T}}_1^{-1}\widetilde{\mathcal{T}}_2)(n) f_{2,h_1} \cdots (\widetilde{\mathcal{T}}_1^{-1}\widetilde{\mathcal{T}}_k)(n) f_{k,h_1} \right\|^2_{L^2(X)}$$
$$\leq \frac{1}{N} \sum_{h_2 \in \mathbb{Z}} \left\| \frac{1}{N} \sum_{n=1}^{N} b(n)b(n + h_1)b(n + h_2)b(n + h_1 + h_2) \right.$$
$$\left. \times (\widetilde{\mathcal{T}}_2^{-1}\widetilde{\mathcal{T}}_3)(n) f_{3,h_1,h_2} \cdots (\widetilde{\mathcal{T}}_2^{-1}\widetilde{\mathcal{T}}_k)(n) f_{k,h_1,h_2} \right\|_{L^2(X)}$$

for some functions f_{i,h_1,h_2} of modulus ≤ 1, and so, by Schwarz's lemma,

$$\left\| \frac{1}{N} \sum_{n=1}^{N} b(n)\widetilde{\mathcal{T}}_1(n) f_1 \cdots \widetilde{\mathcal{T}}_k(n) f_k \right\|^4_{L^2(X)}$$
$$\leq \frac{1}{N^2} \sum_{h_1,h_2 \in \mathbb{Z}} \left\| \frac{1}{N} \sum_{n=1}^{N} b(n)b(n + h_1)b(n + h_2)b(n + h_1 + h_2) \right.$$
$$\left. \times (\widetilde{\mathcal{T}}_2^{-1}\widetilde{\mathcal{T}}_3)(n) f_{3,h_1,h_2} \cdots (\widetilde{\mathcal{T}}_2^{-1}\widetilde{\mathcal{T}}_k)(n) f_{k,h_1,h_2} \right\|_{L^2(X)}.$$

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

Applying Lemma 3.2 with Schwarz's lemma $k-2$ more times, and then Schwarz's lemma once again, we arrive at

$$\begin{aligned}
&\left\| \frac{1}{N} \sum_{n=1}^{N} b(n) \widetilde{\mathcal{T}}_1(n) f_1 \cdots \widetilde{\mathcal{T}}_k(n) f_k \right\|_{L^2(X)}^{2^k} \\
&\quad \le \frac{1}{N^k} \sum_{h_1,\dots,h_k \in \mathbb{Z}} \left| \frac{1}{N} \sum_{n=1}^{N} \prod_{e_1,\dots,e_k \in \{0,1\}} b(n + e_1 h_1 + \cdots + e_k h_k) \right| \\
&\quad \le \left(\frac{1}{N^k} \sum_{h_1,\dots,h_k \in \mathbb{Z}} \left| \frac{1}{N} \sum_{n=1}^{N} \prod_{e_1,\dots,e_k \in \{0,1\}} b(n + e_1 h_1 + \cdots + e_k h_k) \right|^2 \right)^{1/2} \\
&\quad \le \|b\|_{U^{k+1}[N]}^{2^k}. \qquad \square
\end{aligned}$$

We are now in position to prove Theorem 7.1.

Proof of Theorem 7.1. For short, put $\tilde{f}(n) = \mathcal{T}_1(n) f_1 \cdots \mathcal{T}_k(n) f_k$, $n \in \mathbb{N}$. By Lemma 7.2, we have to show that $\lim_{N\to\infty} (1/N) \sum_{n=1}^{N} \Lambda'(n) \tilde{f}(n) = \prod_{i=1}^{k} \int_X f_i \, d\mu$. Let $\varepsilon > 0$. By Proposition 7.7, we can choose $W \in \mathcal{W}$ such that for any N large enough and any $r \in R(W)$ one has

$$\left\| \frac{1}{N} \sum_{n=1}^{N} (\Lambda'_{W,r}(n) - 1) \tilde{f}(Wn + r) \right\|_{L^2(X)} < \varepsilon,$$

and so

$$\left\| \frac{1}{NW} \sum_{n=1}^{N} \Lambda'(Wn + r) \tilde{f}(Wn + r) - \frac{1}{N\phi(W)} \sum_{n=1}^{N} \tilde{f}(Wn + r) \right\|_{L^2(X)} < \frac{\varepsilon}{\phi(W)}.$$

Summing this up for all $r \in R(W)$, and taking into account that $\Lambda'(Wn + r) = 0$ if $r \notin R(W)$, we obtain

$$\left\| \frac{1}{NW} \sum_{n=1}^{NW} \Lambda'(n) \tilde{f}(n) - \frac{1}{\phi(W)} \sum_{r \in R(W)} \frac{1}{N} \sum_{n=1}^{N} \tilde{f}(Wn + r) \right\|_{L^2(X)} < \varepsilon.$$

By the theorem's assumption, for any $r \in R(W)$,

$$\left\| \frac{1}{N} \sum_{n=1}^{N} \tilde{f}(Wn + r) - \prod_{i=1}^{k} \int_X f_i \, d\mu \right\|_{L^2(X)} < \varepsilon$$

for all N large enough. Hence, $\|(1/NW) \sum_{n=1}^{NW} \Lambda'(n) \tilde{f}(n)\|_{L^2(X)} < 2\varepsilon$ for such N, and so

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \Lambda'(n) \tilde{f}(n) = \lim_{N\to\infty} \frac{1}{NW} \sum_{n=1}^{NW} \Lambda'(n) \tilde{f}(n) = \prod_{i=1}^{k} \int_X f_i \, d\mu. \qquad \square$$

We will now collect some special cases of Theorem 7.1. It was shown in [**B**] that if $T_1, \dots, T_k$, with $k \ge 2$, are commuting, invertible, jointly ergodic measure-preserving

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

transformations, then they are actually totally jointly ergodic, that is, for any $W \in \mathbb{N}$ and $r \in \mathbb{Z}$, $T_1^{Wn+r}, \ldots, T_k^{Wn+r}$ are jointly ergodic. Hence, by Theorem 7.1, we obtain the following result.

THEOREM 7.8. *Let $T_1, \ldots, T_k$, where $k \geq 2$, be commuting, invertible, jointly ergodic measure-preserving transformations of X. Then, for any $f_1, \ldots, f_k \in L^\infty(X)$, in the L^2 norm,*

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p\in\mathcal{P}(N)} T_1^p f_1 \cdots T_k^p f_k = \prod_{i=1}^k \int_X f_i \, d\mu.$$

The following is a corollary of Theorem 7.1 and Corollary 1.5.

COROLLARY 7.9. *Let T be a weakly mixing invertible measure-preserving transformation of X and let $\varphi_1, \ldots, \varphi_k$ be unbounded GL-functions $\mathbb{Z} \longrightarrow \mathbb{Z}$ such that $\varphi_j - \varphi_i$ are unbounded for all $i \neq j$. Then, for any $f_1, \ldots, f_k \in L^\infty(X)$,*

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p\in\mathcal{P}(N)} T^{\varphi_1(p)} f_1 \cdots T^{\varphi_k(p)} f_k = \prod_{i=1}^k \int_X f_i \, d\mu.$$

In particular, for any distinct $\alpha_1, \ldots, \alpha_k \in \mathbb{R}\backslash\{0\}$,

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p\in\mathcal{P}(N)} T^{[\alpha_1 p]} f_1 \cdots T^{[\alpha_k p]} f_k = \prod_{i=1}^k \int f_i \, d\mu.$$

From Theorem 7.1 and Corollary 1.6, we obtain the following result.

COROLLARY 7.10. *Let $T_1, \ldots, T_k$ be commuting invertible jointly ergodic measure-preserving transformations of X and let φ be an unbounded GL-function $\mathbb{Z} \longrightarrow \mathbb{Z}$ such that $\lim_{N\to\infty} (1/N) \sum_{n=1}^N \lambda^{\varphi(Wn+r)} = 0$ for every $\lambda \in \mathrm{Eig}(T_1, \ldots, T_k)$, $W \in \mathcal{W}$, and $r \in R(W)$. Then, for any $f_1, \ldots, f_k \in L^\infty(X)$,*

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p\in\mathcal{P}(N)} T_1^{\varphi(p)} f_1 \cdots T_k^{\varphi(p)} f_k = \prod_{i=1}^k \int f_i \, d\mu.$$

In particular, if $\alpha \in \mathbb{R}$ is irrational and such that $e^{2\pi i\alpha^{-1}\mathbb{Q}} \cap \mathrm{Eig}(T_1, \ldots, T_k) = \{1\}$, then

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p\in\mathcal{P}(N)} T_1^{[\alpha p]} f_1 \cdots T_k^{[\alpha p]} f_k = \prod_{i=1}^k \int f_i \, d\mu \quad \text{for any } f_1, \ldots, f_k \in L^\infty(X).$$

8. *GL-families of a continuous parameter*

Let $\mathcal{T}(t)$, $t \in \mathbb{R}$, be a family of measure-preserving transformations of X. We say that $\mathcal{T}$ is *ergodic* if, for any $f \in L^2(X)$, $\lim_{b\to\infty} (1/b) \int_0^b \mathcal{T}(t) f \, dt = \int_X f \, d\mu$ in L^2 norm, and *uniformly ergodic* if, for any $f \in L^2(X)$, $\lim_{b\to\infty} (1/(b-a)) \int_a^b \mathcal{T}(t) f \, dt = \int_X f \, d\mu$. Given several families $\mathcal{T}_1(t), \ldots, \mathcal{T}_k(t)$, $t \in \mathbb{R}$, of measure-preserving transformations of X, we say that $\mathcal{T}_1, \ldots, \mathcal{T}_k$ are *jointly ergodic* if

$$\lim_{b\to\infty} \frac{1}{b} \int_0^b \mathcal{T}_1(t) f_1 \cdots \mathcal{T}_k(t) f_k \, dt = \prod_{i=1}^k \int_X f_i \, d\mu$$

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

in L^2 norm for any $f_1, \dots, f_k \in L^\infty(X)$, and *uniformly jointly ergodic* if

$$\lim_{b-a\to\infty} \frac{1}{b-a} \int_a^b \mathcal{T}_1(t) f_1 \cdots \mathcal{T}_k(t) f_k \, dt = \prod_{i=1}^k \int_X f_i \, d\mu$$

for any $f_1, \dots, f_k \in L^\infty(X)$.

Let G be a commutative group of measure-preserving transformations of X. In analogy with the terminology adopted in previous sections, we will call a family $\mathcal{T}(t)$, $t \in \mathbb{R}$, of elements of G *a GL-family* if it is of the form $\mathcal{T}(t) = T_1^{\varphi_1(t)} \cdots T_r^{\varphi_r(t)}$, $t \in \mathbb{R}$, where $T_1, \dots, T_r$ are continuous homomorphisms $\mathbb{R} \longrightarrow G$ and $\varphi_1, \dots, \varphi_r$ are GL-functions $\mathbb{R} \longrightarrow \mathbb{R}$. Let $\mathfrak{T}_\mathbb{R}$ denote the set of GL-families of transformations from G. We have the following analogue of Theorem 6.4.

THEOREM 8.1. *Let $\mathcal{T}_1, \dots, \mathcal{T}_k \in \mathfrak{T}_\mathbb{R}$. Then the following are equivalent:*

(i) *$\mathcal{T}_1, \dots, \mathcal{T}_k$ are jointly ergodic;*

(ii) *$\mathcal{T}_1, \dots, \mathcal{T}_k$ are uniformly jointly ergodic;*

(iii) *the GL-families $\mathcal{T}_i^{-1}\mathcal{T}_j$ are ergodic for all $i \neq j$ and the GL-family $\mathcal{T}_1 \times \cdots \times \mathcal{T}_k$ is ergodic.*

One can verify that a (properly modified) proof of Theorem 6.4 works in the situation at hand as well. An alternative and simpler approach is to derive Theorem 8.1 from Theorem 6.4 with the help of the techniques developed in [**BeLM**]. Namely, we can use the following fact.

THEOREM 8.2. [**BeLM**] *Let $\tau : \mathbb{R} \longrightarrow V$ be a bounded measurable mapping to a Banach space V such that, for every $t \in \mathbb{R}$, the limit $L_t = \lim_{N\to\infty} (1/N) \sum_{n=1}^N \tau(nt)$ (respectively, $L_t = \lim_{N-M\to\infty} (1/(N-M)) \sum_{n=M+1}^N \tau(nt)$) exists for almost every $t \in \mathbb{R}$. Then the limit $L = \lim_{b\to\infty} (1/b) \int_0^b \tau(t)\, dt$ (respectively, $L = \lim_{b-a\to\infty} (1/(b-a)) \int_a^b \tau(t)\, dt$) also exists, and $L_t = L$ for almost every $t \in \mathbb{R}$.*

To apply this result, we need to verify that for any GL-family $\mathcal{T}$ and any $t \in \mathbb{R}$ the sequence $\mathcal{T}(nt)$, $n \in \mathbb{Z}$, is a GL-sequence. This is indeed so: any GL-function φ can be written in the form $\varphi(t) = \sum_{j=1}^l [\varphi_j(t)] a_j + ct + d$, where φ_j are GL-functions and $a_j, c, d \in \mathbb{R}$; thus, for any flow T and any $t \in \mathbb{R}$, $T^{\varphi(nt)} = (\prod_{j=1}^l (T^{a_j})^{[\varphi_j(nt)]})(T^{ct})^n T^d$, in which expression all the factors are GL-sequences in the group generated by the transformations $T^{a_1}, \dots, T^{a_l}, T^{ct}, T^d$. We may now apply Theorem 8.2 in conjunction with Lemma 5.2 to the family $\tau(t) = \mathcal{T}(t) f$, where $\mathcal{T}$ is a GL-family and $f \in L^2(X)$, and see that the limits $\lim_{b\to\infty} (1/b) \int_0^b \mathcal{T}(t) f \, dt$ and $\lim_{b-a\to\infty} (1/(b-a)) \int_a^b \mathcal{T}(t) f \, dt$ exist, and $\mathcal{T}$ is ergodic and is uniformly ergodic if and only if the GL-sequences $\mathcal{T}(nt)$, $n \in \mathbb{Z}$, are ergodic (= uniformly ergodic) for almost all $t \in \mathbb{R}$.

Proof of Theorem 8.1. Assume that the GL-families $\mathcal{T}_1, \dots, \mathcal{T}_k$ are jointly ergodic. For any distinct i and j, $\mathcal{T}_i$ and $\mathcal{T}_j$ are jointly ergodic, which implies that $\mathcal{T}_i^{-1}\mathcal{T}_j$ is ergodic. It remains to show that $\mathcal{T}_1 \times \cdots \times \mathcal{T}_k$ is ergodic. $\mathcal{T}_1, \dots, \mathcal{T}_k$ are ergodic, so the GL-sequences $\mathcal{T}_1(nt), \dots, \mathcal{T}_k(nt)$, $n \in \mathbb{Z}$, are ergodic for almost every $t \in \mathbb{R}$. Thus, by Lemma 6.1, C-$\lim_n \bigotimes_{i=1}^k \mathcal{T}_i(nt) f_i = 0$ for almost every $t \in \mathbb{R}$ whenever $f_i \in \mathcal{H}_{\mathrm{wm}}$ for

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

at least one of i; by Theorem 8.2, this implies that $\lim_{b\to\infty}(1/b)\int_0^b \prod_{i=1}^k \mathcal{T}_i(t)f_i\,dt = 0$ whenever $f_i \in \mathcal{H}_{\mathrm{wm}}$ for at least one of i. If $f_i \in \mathcal{H}_{\mathrm{c}}$ for all i, we may assume that all f_i are non-constant eigenfunctions of elements of G, so $\mathcal{T}_i(t)f_i = \lambda_i(t)f_i$, $i = 1, \dots, k$, where λ_i are functions $\mathbb{R} \longrightarrow \mathbb{C}$. In this case,

$$0 = \lim_{b\to\infty}\frac{1}{b}\int_0^b \prod_{i=1}^k \mathcal{T}_i(t)f_i\,dt = \left(\lim_{b\to\infty}\frac{1}{b}\int_0^b \prod_{i=1}^k \lambda_i(t)\,dt\right)\prod_{i=1}^k f_i,$$

so $\lim_{b\to\infty}(1/b)\int_0^b \prod_{i=1}^k \Lambda_i(t)\,dt = 0$, so

$$\lim_{b\to\infty}\frac{1}{b}\int_0^b \bigotimes_{i=1}^k \mathcal{T}_i(t)f_i\,dt = \left(\lim_{b\to\infty}\frac{1}{b}\int_0^b \prod_{i=1}^k \lambda_i(t)\,dt\right)\bigotimes_{i=1}^k f_i = 0.$$

Hence, $\mathcal{T}_1 \times \cdots \times \mathcal{T}_k$ is ergodic.

Conversely, if $\mathcal{T}_i^{-1}\mathcal{T}_j$ for all $i \neq j$ and $\mathcal{T}_1 \times \cdots \times \mathcal{T}_k$ are ergodic, then, by Lemma 5.2 and Theorem 8.2, the GL-sequences $(\mathcal{T}_i^{-1}\mathcal{T}_j)(nt)$ and $(\mathcal{T}_1 \times \cdots \times \mathcal{T}_k)(nt)$ are ergodic for almost every $t \in \mathbb{R}$, so, by Theorem 6.4, the GL-sequences $\mathcal{T}_1(nt), \dots, \mathcal{T}_k(nt)$ are jointly ergodic (= uniformly jointly ergodic, see remark after the proof of Theorem 6.4) for almost every $t \in \mathbb{R}$, so $\mathcal{T}_1, \dots, \mathcal{T}_k$ are jointly ergodic and uniformly jointly ergodic by Theorem 8.2. □

For a continuous parameter, Corollaries 1.5 and 1.6 take the following form.

COROLLARY 8.3. *Let T^s, $s \in \mathbb{R}$, be a weakly mixing continuous flow of measure-preserving transformations of X and let $\varphi_1, \dots, \varphi_k$ be unbounded GL-functions $\mathbb{R} \longrightarrow \mathbb{R}$ such that $\varphi_j - \varphi_i$ are unbounded for all $i \neq j$; then the GL-families $T^{\varphi_1(t)}, \dots, T^{\varphi_k(t)}$, $t \in \mathbb{R}$, are jointly ergodic. In particular, for any distinct $\alpha_1, \dots, \alpha_k \in \mathbb{R}\backslash\{0\}$, the GL-families $T^{[\alpha_1 t]}, \dots, T^{[\alpha_k t]}$, $t \in \mathbb{R}$, are jointly ergodic.*

COROLLARY 8.4. *Let $T_1^s, \dots, T_k^s$, $s \in \mathbb{R}$, be commuting jointly ergodic continuous flows of measure-preserving transformations of X and let φ be an unbounded GL-function; then the GL-families $T_1^{\varphi(t)}, \dots, T_k^{\varphi(t)}$, $t \in \mathbb{R}$, are jointly ergodic if and only if $\lim_{b\to\infty}(1/b)\int_0^b \lambda^{\varphi(t)}\,dt = 0$ for every $\lambda \in \mathrm{Eig}(T_1^1, \dots, T_k^1)\backslash\{1\}$.*

We would also like to remark that, in the case of a continuous parameter, by using a 'change of variable' trick one can easily extend the results above, proved for GL-families, to more general families of transformations of the form $\mathcal{T}(\sigma(t))$, where $\mathcal{T}$ is a GL-family and σ is a monotone function of 'regular' growth. What we mean is stated in the following proposition.

PROPOSITION 8.5. *Let $\mathcal{T}_1(t), \dots, \mathcal{T}_k(t)$, $t \in \mathbb{R}$, be jointly ergodic families of measure-preserving transformations of X and let $\sigma : \mathbb{R} \longrightarrow \mathbb{R}$ be a strictly increasing C^1-function such that σ' is monotone and*

$$\lim_{b-a\to\infty}\frac{(\sigma^{-1})'(a)}{\sigma^{-1}(b)-\sigma^{-1}(a)} = \lim_{b-a\to\infty}\frac{(\sigma^{-1})'(b)}{\sigma^{-1}(b)-\sigma^{-1}(a)} = 0.$$

Then the families $\mathcal{T}_1(\sigma(t)), \dots, \mathcal{T}_k(\sigma(t))$ are also jointly ergodic.

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

Remark. Of course, Proposition 8.5 remains true when the families $\mathcal{T}_i$ are defined only on a ray $[r, \infty)$. In this form, it applies when σ is of the form $\sigma(t) = \sum_{i=1}^{d} c_i t^{\alpha_i}$, where α_i are non-negative reals, on the ray $[r, \infty)$ where σ' becomes monotone. Moreover, for σ of this sort, σ^{-1} also satisfies the assumptions of the proposition; thus, $\mathcal{T}_1, \ldots, \mathcal{T}_k$ are jointly ergodic if and only if $\mathcal{T}_1(\sigma(t)), \ldots, \mathcal{T}_k(\sigma(t))$ are.

Let us say that a function $\tau : \mathbb{R} \longrightarrow \mathbb{R}$ *has uniform Cesàro limit* L if $\lim_{b-a\to\infty} (1/(b-a)) \int_a^b \tau(t)\, dt = L$. Proposition 8.5 is simply a special case of the following general fact.

PROPOSITION 8.6. *Let σ be as in Proposition 8.5. Then, if a bounded function $\tau : \mathbb{R} \longrightarrow \mathbb{R}$ has uniform Cesàro limit L, the function $\tau(\sigma(t))$ also does.*

This proposition must be well known to aficionados, but, since we have not been able to find any references, we will sketch its proof. Making the substitution $s = \sigma(t)$, we get

$$\lim_{b-a\to\infty} \frac{1}{b-a} \int_a^b \tau(\sigma(t))\, dt = \lim_{q-p\to\infty} \frac{1}{\sigma^{-1}(q) - \sigma^{-1}(p)} \int_p^q \tau(s)(\sigma^{-1})'(s)\, ds.$$

What we have in the right-hand part of this formula, $\lim_{b-a\to\infty} (1/(\int_a^b \omega)) \int_a^b \tau(t)\omega(t)\, dt$, is *the weighted uniform Cesàro limit of τ with weight $\omega = (\sigma^{-1})'$*. Rewriting Proposition 8.6 in terms of ω, we reduce it to the following lemma.

LEMMA 8.7. *Let $\omega : \mathbb{R} \longrightarrow \mathbb{R}$ be a positive monotone function with the property that $\lim_{b-a\to\infty} \omega(a)/\int_a^b \omega = \lim_{b-a\to\infty} \omega(b)/\int_a^b \omega = 0$. Then, if a bounded function $\tau : [0, \infty) \longrightarrow \mathbb{R}$ has uniform Cesàro limit L, then the weighted uniform Cesàro limit of τ with weight ω is equal to L (and, in particular, exists).*

Proof. We will assume that ω is increasing; the case of decreasing ω is similar. Let $M = \sup |\tau|$. Let $\varepsilon > 0$. Find $c > 0$ such that $(1/c) \int_x^{x+c} \tau(t)\, dt \overset{\varepsilon}{\approx} L$ for every $x > 0$. Averaging this equation with weight ω over an interval $[a, b]$ and changing the order of integration, we get

$$L \overset{\varepsilon}{\approx} \frac{1}{\int_a^b \omega} \int_a^b \omega(x) \left(\frac{1}{c} \int_x^{x+c} \tau(t)\, dt \right) dx = \frac{1}{c \int_a^b \omega} \int_a^b \left(\int_{t-c}^{t} \omega(x)\, dx \right) \tau(t)\, dt$$
$$- \frac{1}{c \int_a^b \omega} \int_a^{a+c} \left(\int_{t-c}^{a} \omega(x)\, dx \right) \tau(t)\, dt + \frac{1}{c \int_a^b \omega} \int_b^{b+c} \left(\int_{t-c}^{b} \omega(x)\, dx \right) \tau(t)\, dt.$$

The moduli of the second and of the third summands in the right-hand part of this equality are majorized by $(M\omega(b)c^2/2)/(c \int_a^b \omega)$ and tend to 0 as $b - a \longrightarrow \infty$. We now claim that, for large $b - a$, the first summand is close to $(1/(\int_a^b \omega)) \int_a^b \tau(t)\omega(t)\, dt$. Indeed,

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

taking into account the monotonicity of ω, we have

$$\begin{aligned}
&\left|\frac{1}{c\int_a^b \omega}\int_a^b\left(\int_{t-c}^{t}\omega(x)\,dx\right)\tau(t)\,dt-\frac{1}{\int_a^b\omega}\int_a^b\tau(t)\omega(t)\,dt\right|\\
&\quad=\frac{1}{c\int_a^b\omega}\left|\int_a^b\left(\int_{t-c}^{t}(\omega(x)-\omega(t))\,dx\right)\tau(t)\,dt\right|\\
&\quad\le\frac{M}{c\int_a^b\omega}\int_a^b\left(\int_{t-c}^{t}\left|\omega(x)-\omega(t)\right|dx\right)dt\\
&\quad\le\frac{M}{\int_a^b\omega}\int_a^b(\omega(t)-\omega(t-c))\,dt\\
&\quad=\frac{M}{\int_a^b\omega}\int_{b-c}^{b}\omega(t)\,dt-\frac{M}{\int_a^b\omega}\int_{a-c}^{a}\omega(t)\,dt\le Mc\frac{\omega(b)}{\int_a^b\omega},
\end{aligned}$$

which tends to 0 as $b-a\to\infty$. □

9. *Non-commuting GL-sequences*

If GL-sequences $\mathcal{T}_1,\dots,\mathcal{T}_k$ do not commute, the situation becomes much more complicated. Recall that we introduced the notions of ergodicty and joint ergodicity of sequences of transformations (Definitions 5.3 and 6.3 above) with respect to an arbitrary fixed Følner sequence in $\mathbb{Z}$. However, for commuting GL-sequences the property of being ergodic or jointly ergodic has turned out to be Følner sequence independent (see Remarks 5.4 and 6.5). An example in [**BBe2**] shows that this is no longer the case if $\mathcal{T}_i$ do not commute, even in the conventional case $\mathcal{T}_i(n)=T_i^n$. It follows that one cannot expect to have a criterion of joint ergodicity in terms of ergodicity of a certain collection of sequences of transformations, unless the ergodicity of these sequences is itself Følner sequence dependent.

One nevertheless has the following generalization of [**BBe2**, Theorem 2.1].

THEOREM 9.1. *Let $G_1,\dots,G_k$ be several commutative groups of measure-preserving transformations of X and, for each $i=1,\dots,k$, let $\mathcal{T}_i$ be a GL-sequence in G_i. Then $\mathcal{T}_1,\dots,\mathcal{T}_k$ are jointly ergodic if and only if $\mathcal{T}_1,\dots,\mathcal{T}_k$ are ergodic and* $\mathrm{C\text{-}lim}_n\int_X\prod_{i=1}^k\mathcal{T}_i(n)f_i\,d\mu=\prod_{i=1}^k\int_X f_i\,d\mu$ *for any $f_1,\dots,f_k\in L^\infty(X)$.*

Proof. The 'only if' direction is clear; we will prove the 'if' statement. Let $f_1,\dots,f_k\in L^\infty(X)$, with $|f_i|\le 1$ for all i. First, assume that for some i, f_i is in the $\mathcal{H}_{\mathrm{wm}}$ space corresponding to the group G_i. We will assume that $i=1$; then $\mathcal{T}_1$ is weakly mixing on f_1 by Theorem 5.5. Let $\varepsilon>0$, and let a Bohr set $H\subseteq Z$, transformations $S_i\in G_i$ for $i=1,\dots,k$, and sets $E_h\subseteq\mathbb{Z}$ for $h\in H$ be as in Proposition 5.1(iv). Then, for any $h_1,h_2\in H$,

$$\begin{aligned}
&\left|\mathrm{C\text{-}lim}_n\left\langle\prod_{i=1}^k\mathcal{T}_i(n+h_1)f_i,\prod_{i=1}^k\mathcal{T}_i(n+h_2)f_i\right\rangle\right|\\
&\quad\le\left|\mathrm{C\text{-}lim}_n\int_X\prod_{i=1}^k\mathcal{T}_i(n)(\mathcal{T}_i(h_1)S_if_i\cdot\mathcal{T}_i(h_2)S_i\bar{f}_i)\,d\mu\right|+2\varepsilon
\end{aligned}$$

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

$$= \left| \prod_{i=1}^{k} \int_X \mathcal{T}_i(h_1) S_i f_i \cdot \mathcal{T}_i(h_2) S_i \bar{f}_i \, d\mu \right| + 2\varepsilon$$
$$\leq \left| \int_X \mathcal{T}_1(h_1) S_1 f_1 \cdot \mathcal{T}_1(h_2) S_1 \bar{f}_1 \, d\mu \right| + 2\varepsilon.$$

Since $\text{D-lim}_h \langle \mathcal{T}_1(h) S_1 f, f' \rangle = 0$ for any $f' \in L^2(X)$, we can construct an infinite set $B \subseteq H$ such that $|\int_X \mathcal{T}_1(h_1) S_1 f_1 \cdot \mathcal{T}_1(h_2) S_1 \bar{f}_1 \, d\mu| < \varepsilon$ for any distinct $h_1, h_2 \in B$. By Lemma 3.1, $\text{C-limsup}_{\|\cdot\|,n} \prod_{i=1}^{k} \mathcal{T}_i(n) f_i < \sqrt{3\varepsilon}$. Since ε is arbitrary, $\text{C-lim}_n \prod_{i=1}^{k} \mathcal{T}_i(n) f_i = 0$.

Now assume that for each i, $\mathcal{T}_i$ acts on f_i in a compact way. We then may assume that, for each i, f_i is a non-constant eigenfunction of G_i, and so $\mathcal{T}_i(n) f_i = \lambda_i(n) f_i$, $n \in \mathbb{Z}$, for some GL-sequence λ_i in $\{z \in \mathbb{C} : |z| = 1\}$. In this case, $\prod_{i=1}^{k} \mathcal{T}_i(n) f_i = \lambda(n) \prod_{i=1}^{k} f_i$, where $\lambda(n) = \prod_{i=1}^{k} \lambda_i(n)$. Since $\text{C-lim}_n \int_X \prod_{i=1}^{k} \mathcal{T}_i(n) f_i \, d\mu = 0$, we have $\text{C-lim}_n \lambda(n) = 0$, and so $\text{C-lim}_n \prod_{i=1}^{k} \mathcal{T}_i(n) f_i = 0$. □

Acknowledgements. The first two authors were partially supported by NSF grant DMS-1162073.

References

[**B**] D. Berend. Joint ergodicity and mixing. *J. Anal. Math.* **45** (1985), 255–284.

[**BBe1**] D. Berend and V. Bergelson. Jointly ergodic measure preserving transformations. *Israel J. Math.* **49**(4) (1984), 307–314.

[**BBe2**] D. Berend and V. Bergelson. Characterization of joint ergodicity for non-commuting transformations. *Israel J. Math.* **56**(1) (1986), 123–128.

[**Be**] V. Bergelson. Weakly mixing PET. *Ergod. Th. & Dynam. Sys.* **7**(3) (1987), 337–349.

[**BeG**] V. Bergelson and A. Gorodnik. Weakly mixing group actions: a brief survey and an example. *Modern Dynamical Systems and Applications*. Cambridge University Press, New York, 2004, pp. 3–25.

[**BeH**] V. Bergelson and I. J. Håland Knutson. Weak mixing implies weak mixing of higher orders along tempered functions. *Ergod. Th. & Dynam. Sys.* **29**(5) (2009), 1375–1416.

[**BeL1**] V. Bergelson and A. Leibman. Polynomial extensions of van der Waerden's and Szemeredi's theorems. *J. Amer. Math. Soc.* **9** (1996), 725–753.

[**BeL2**] V. Bergelson and A. Leibman. Distribution of values of bounded generalized polynomials. *Acta Math.* **198** (2007), 155–230.

[**BeLM**] V. Bergelson, A. Leibman and C. G. Moreira. From discrete- to continuous-time ergodic theorems. *Ergod. Th. & Dynam. Sys.* **32**(2) (2012), 383–426.

[**BeMc**] V. Bergelson and R. McCutcheon. Uniformity in the polynomial Szemeredi theorem. *Ergodic Theory of $\mathbb{Z}^d$ Actions (London Mathematical Society, Lecture Note Series, 228)*. Cambridge University Press, Cambridge, 1996, pp. 273–296.

[**BeR**] V. Bergelson and J. Rosenblatt. Mixing actions of groups. *Illinois J. Math.* **32**(1) (1988), 65–80.

[**F**] N. Frantzikinakis. Multiple recurrence and convergence for Hardy sequences of polynomial growth. *J. Anal. Math.* **112** (2010), 79–135.

[**FHoK**] N. Frantzikinakis, B. Host and B. Kra. Multiple recurrence and convergence for sequences related to prime numbers. *J. Reine Angew. Math.* **611** (2007), 131–144.

[**FK**] N. Frantzikinakis and B. Kra. Ergodic averages for independent polynomials and applications. *J. Lond. Math. Soc.* (2) **74** (2006), 131–142.

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

[**GT**] B. Green and T. Tao. Linear equations in primes. *Ann. of Math. (2)* **171**(3) (2010), 1753–1850.

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press

[**GTZ**] B. Green, T. Tao and T. Ziegler. An inverse theorem for the Gowers U^{s+1}-norm. *Ann. of Math.* **176**(2) (2012), 1231–1372.

[**S**] W. Sun. Multiple recurrence and convergence for certain averages along shifted primes. *Ergod. Th. & Dynam. Sys.*, available on CJO2014, doi:10.1017/etds.2014.6.

[**vNK**] J. von Neumann and B. O. Koopman. Dynamical systems of continuous spectra. *Proc. Natl. Acad. Sci. USA* **18** (1932), 255–263.

https://doi.org/10.1017/etds.2015.11 Published online by Cambridge University Press