POINTWISE MULTIPLE AVERAGES FOR SUBLINEAR FUNCTIONS

SEBASTIÁN DONOSO, ANDREAS KOUTSOGIANNIS AND WENBO SUN

Abstract. For any measure preserving system $(X,\mathcal{B},\mu,T_1,\dots,T_d)$ with no commutativity assumptions on the transformations T_i, $1\le i\le d$, we study the pointwise convergence of multiple ergodic averages with iterates of different growth coming from a large class of sublinear functions. This class properly contains important subclasses of Hardy field functions of order 0 and of Fejér functions, i.e., tempered functions of order 0. We show that the convergence of the single average, via an invariant property, implies the convergence of the multiple one. We also provide examples of sublinear functions which are in general bad for convergence on arbitrary systems, but good for uniquely ergodic systems. The case where the fastest function is linear is addressed as well, and we provide, in all the cases, an explicit formula of the limit function.

1. Introduction and main results

The study of the limiting behavior, in $L^2(\mu)$ or pointwise, as $N\to\infty$, of *multiple ergodic averages* of the form

$$\frac{1}{N}\sum_{n=0}^{N-1} T_1^{a_1(n)}f_1\cdot\ldots\cdot T_d^{a_d(n)}f_d, \tag{1}$$

where $T_1,\dots,T_d: X\to X$ are invertible (usually commuting) measure preserving transformations acting on a probability space $(X,\mathcal{B},\mu)$; $f_1,\dots,f_d\in L^\infty(\mu)$ and $a_i(\cdot)$ are suitable functions taking integer values on integers for all $1\le i\le d$,[1] is a central problem in ergodic theory. With partial knowledge of the limiting behavior of (1), for the case $T_i=T$ and $a_i(n)=in$, Furstenberg provided a purely ergodic theoretical proof of Szemerédi's theorem ([22]), i.e., every subset of $\mathbb{N}$ with positive upper density contains arbitrary long arithmetic progressions.

In recent years, motivated by the work of Furstenberg, fruitful progress has been made towards the study of the existence and also of the exact value of the $L^2(\mu)$ limit of (1) for various classes of functions a_i. For the existence of the limit, we refer the readers to [5, 25, 26, 27, 28, 31, 32]. As for the explicit expression of the limit, the first result is von Neumann's mean ergodic theorem which says that for $d=1$ and $a_1(n)=n$ the limit of (1) is $\mathbb{E}(f_1|\mathcal{I}(T_1))$, where $\mathcal{I}(T)$ denotes the σ-algebra of T-invariant sets and $\mathbb{E}(f|\mathcal{I}(T))$ is the *conditional expectation* of f with respect to $\mathcal{I}(T)$. The classes of integer polynomial, integer parts of real polynomial, Hardy field (see definition in § 2) and more generally

2010 *Mathematics Subject Classification.* Primary: 37A05; Secondary: 37A30, 28A99 .

Key words and phrases. Pointwise convergence, multiple averages, sublinear functions, Fejér functions, Hardy functions.

The first author is supported by Fondecyt Iniciación en Investigación Grant 11160061.

[1]For a measurable function f and a transformation $T\colon X\to X$, Tf denotes the composition $f\circ T$.

tempered classes of functions are also studied in [7, 9, 20] for a single T and in [15, 21, 27] for commuting T_i's.

Following the work of Gowers ([23]), the averages along cubes have enjoyed a key role in establishing convergence results as illustrated by several papers by Host, Kra ([25]) and Bergelson ([8]) among others. For reasons of completeness we define this notion that lies on the background of our study. Let $(X,\mathcal{B},\mu)$ be a probability space and T_ϵ be measure preserving transformations on X for all $\epsilon\in\{0,1\}^d$. The *cubic averages* are expressions of the form

$$\frac{1}{N^d}\sum_{n_1,\dots,n_d=0}^{N-1}\prod_{\epsilon=(\epsilon_1,\dots,\epsilon_d)\in\{0,1\}^d}T_\epsilon^{n_1\epsilon_1+\dots+n_d\epsilon_d}f_\epsilon, \tag{2}$$

where $f_\epsilon\in L^\infty(\mu)$ for all $\epsilon\in\{0,1\}^d$. One of the nice properties of (2) is that the pointwise convergence for such averages holds (hence, its $L^2(\mu)$ as well) for not necessarily commuting transformations T_ϵ, $\epsilon\in\{0,1\}^d$. This property was first discovered by Assani in the papers [3, 4] and it was later extended by Chu and Frantzikinakis ([16]). We refer the interested reader to see the papers [8, 25] for $L^2(\mu)$ convergence/recurrence of cubic averages for a single transformation; the papers [5, 14] for $L^2(\mu)$ convergence for commuting transformations; the paper [29] for pointwise convergence of cubic averages for a single transformation; the papers [17, 19] for the variant of cubic averages that arise from the van der Corput trick; and finally, the papers [3, 4, 16] for pointwise convergence of cubic averages for non-commuting transformations. Hence, motivated by the fruitful results on cubic averages, it is natural to ask: Does the limit of (1), as $N\to\infty$, exists in the pointwise sense? Do we necessarily have to assume that the transformations $T_1,\dots,T_d$ commute with each other?

The results on the existence and explicit expression of the pointwise limit of (1) are not as abundant as that of the $L^2(\mu)$ limit, even for the case when $T_1,\dots,T_d$ are commuting transformations. In fact, even the $d=1$ case is not completely understood (for some results see [11, 13]). For $d=2$ and commuting transformations T_1 and T_2, Assani gave sufficient conditions so that (1) converges pointwisely a.e. for all $f,g\in L^2(\mu)$ (the argument is a nice application of the classical van der Corput trick–see [2, Proposition 4]). For $d=2$, Bourgain showed (in [12]) that the pointwise limit of (1) exists when $T_1=T_2$ and $a_1(n)=an$, $a_2(n)=bn$ for $a,b\in\mathbb{Z}$. Recently, Huang, Shao and Ye ([30]) showed the existence of the pointwise limit of (1) for $T_i=T$, $a_i(n)=in$ under the assumption that T is a distal transformation (see also [1, 24] for some particular weakly mixing systems). This result was extended in [18] for two commuting transformations generating a distal action, and for an arbitrary number of commuting transformations in [19] (also for a distal system and linear iterates).

When $T_1,\dots,T_d$ are not necessarily commuting, in order to expect (1) to have a nice behavior (in either the $L^2(\mu)$ or the pointwise sense), one needs to impose additional *distinctness* conditions on the a_i's. For the case $d=2$ and $a_1(n)=a_2(n)=n$, counterexamples where (1) does not converge in $L^2(\mu)$ were given by Berend (see [6, Example 7.1], where (1) does not even converge weakly), and by Bergelson and Leibman (see [10, Section 4], where T_1, T_2 generate a solvable group).

It should be mentioned that there are also positive results in the pointwise setting for $d = 2$ and $a_1(n) = a_2(n) = n$. Assani (in [2, Theorem 6]) provided a sufficient condition so that (1) converges pointwisely a.e. (for all $f, g \in L^2(\mu)$) when we don't assume any commutativity of T_1 and T_2.

On the other hand, arguably the only known multiple convergence result for non-commuting transformations, for a general system is due to Frantzikinakis. In [20, Theorem 2.7], under no commutativity assumption on T_i's, for integer part (denoted by $[\cdot]$) of functions a_i in $\mathcal{LE}$,[2] with $x^\varepsilon \prec a_d \prec \ldots \prec a_1 \prec x$,[3] for some $\varepsilon > 0$, he showed that the limit of (1) in $L^2(\mu)$ is equal to $\mathbb{E}(f_1|\mathcal{I}(T_1)) \cdot \ldots \cdot \mathbb{E}(f_d|\mathcal{I}(T_d))$.

In this paper, we study the pointwise convergence of (1) for integer part of sequences of functions of different growth rate which are sublinear (actually the case where the fastest function is linear is addressed as well), with no commutativity assumption on the transformations. More specifically, we obtain a pointwise version of [20, Theorem 2.7] mentioned above, to the wider class of sublinear functions $\mathcal{S}^*$ (see next subsection or the appendix for notation):

Theorem 1.1. *Let $d \in \mathbb{N}$ and (X_i, μ_i, T_i), $1 \leq i \leq d$ be measure preserving systems. Let $a_i \in \mathcal{S}^*$, $1 \leq i \leq d$ with $a_d \prec \ldots \prec a_1$ and $a'_d \prec \ldots \prec a'_1$, and ν be any coupling of the spaces (X_i, μ_i). Then for all $f_i \in L^\infty(\mu_i), 1 \leq i \leq d$, the averages*

$$\frac{1}{N}\sum_{n=0}^{N-1} T_1^{[a_1(n)]} f_1(x_1) \cdot \ldots \cdot T_d^{[a_d(n)]} f_d(x_d)$$

converge as $N \to \infty$ for ν-a.e. $(x_1, \ldots, x_d) \in X_1 \times \cdots \times X_d$ to

$$\mathbb{E}(f_1|\mathcal{I}(T_1))(x_1) \cdot \ldots \cdot \mathbb{E}(f_d|\mathcal{I}(T_d))(x_d).$$

In particular, if $(X_i, \mu_i) = (X, \mu)$, $1 \leq i \leq d$, for ν the diagonal coupling we get that

$$\frac{1}{N}\sum_{n=0}^{N-1} T_1^{[a_1(n)]} f_1(x) \cdot \ldots \cdot T_d^{[a_d(n)]} f_d(x)$$

converge as $N \to \infty$ for μ-a.e. $x \in X$ to

$$\mathbb{E}(f_1|\mathcal{I}(T_1))(x) \cdot \ldots \cdot \mathbb{E}(f_d|\mathcal{I}(T_d))(x).$$

Remark. Let $\mathcal{LE}_\varepsilon$ denote the set of logarithmico-exponential Hardy field functions a satisfying the growth condition $x^\varepsilon \prec a(x) \prec x$ for some $\varepsilon > 0$. By a variation of the argument in [20, Proposition 6.4], one can obtain a different proof of Theorem 1.1 for the special case where $a_i \in \mathcal{LE}_\varepsilon$, $1 \leq i \leq d$. The idea of [20, Proposition 6.4] is to convert the multiple averages for sublinear functions of different growth in $\mathcal{LE}_\varepsilon$, via a change of variable, to an average of the same form but with linear (equal to x) fastest function. Our method, which is applicable into a larger class of functions, has a different philosophy focusing instead on the invariance property of the averages under the transformations

[2] a is a *logarithmico-exponential Hardy field function* if it belongs to a Hardy field of real valued functions and it's defined on some $(c, +\infty)$, $c \geq 0$, by a finite combination of symbols $+, -, \times, \div, \sqrt[n]{\cdot}, \exp, \log$ acting on the real variable x and on real constants (for more on Hardy field functions and in particular for logarithmico-exponential ones check for example [20, 21]).

[3] For two functions a, b we write $a(x) \prec b(x)$, or just $a \prec b$ if $a(x)/b(x) \to 0$ as $x \to \infty$.

$T_1 \times id \times \cdots \times id$, $id \times T_2 \times \cdots \times id$, $\ldots$, $id \times \cdots \times id \times T_d$, via which we deduce the limit of the expressions of interest. Another advantage of this method is that it can also be used to show that there are certain sublinear functions for which even though the pointwise convergence might in general fail, it holds for all the uniquely ergodic systems (see § 1.2 for details).

It is worth nothing that a result similar to Theorem 1.1 holds when we replace a_1 by a linear function, i.e., a polynomial of degree 1. More specifically, we have:

Theorem 1.2. *Let $d \in \mathbb{N}$, (X_i, μ_i, T_i), $1 \leq i \leq d$ be measure preserving systems, a_1 be a linear function, $a_i \in \mathcal{S}^*$, $2 \leq i \leq d$ with $a_d \prec \ldots \prec a_1$ and $a'_d \prec \ldots \prec a'_1$, and ν be any coupling of the spaces (X_i, μ_i). Then for all $f_i \in L^\infty(\mu_i)$, $1 \leq i \leq d$ the averages*

$$\frac{1}{N}\sum_{n=0}^{N-1} T_1^{[a_1(n)]}f_1(x_1)T_2^{[a_2(n)]}f_2(x_2)\cdot\ldots\cdot T_d^{[a_d(n)]}f_d(x_d)$$

converge as $N \to \infty$ for ν-a.e. $\vec{x} = (x_1, \ldots, x_d) \in X_1 \times \cdots \times X_d$.

In particular, if $a_1(n) = kn + \ell$, $k = p/q$, $p,\ q \in \mathbb{Z} \setminus \{0\}$, then the limit is equal to

$$\frac{1}{q}\sum_{j=0}^{q-1} \mathbb{E}\Big(T_1^{[\frac{pj}{q}+\ell]}f_1|\mathcal{I}(T_1^p)\Big)(x_1)\mathbb{E}(f_2|\mathcal{I}(T_2))(x_2)\cdot\ldots\cdot\mathbb{E}(f_d|\mathcal{I}(T_d))(x_d),$$

while if $a_1(n) = \gamma n + \ell, \gamma \in \mathbb{R}\backslash\mathbb{Q}$, then the limit is equal to

$$F(x_1)\mathbb{E}(f_2|\mathcal{I}(T_2))(x_2)\cdot\ldots\cdot\mathbb{E}(f_d|\mathcal{I}(T_d))(x_d),$$

where

$$F(x) = \sum_{m\in\mathbb{Z}} \exp\Big(2\pi i\frac{m\ell}{\gamma}\Big)\cdot\frac{\exp(-2\pi i\frac{m}{\gamma})-1}{-2\pi i\frac{m}{\gamma}}\mathbb{E}(f_1|\mathcal{I}_{\gamma,m}(T))(x)$$

and $\mathcal{I}_{\gamma,m}(T)$ is the sub-σ-algebra generated by the eigenspace of T with eigenvalue $-\dfrac{m}{\gamma}$.

Via Theorems 1.1 and 1.2, applied to the diagonal coupling, we immediately get the following result on sequences of different growth rates of the form $(n^c)_n$, $0 < c \leq 1$:

Corollary 1.3. *Let $d \in \mathbb{N}$ and $(X, \mu, T_1, \ldots, T_d)$ be a measure preserving system. For all $0 < c_d < \ldots < c_1 \leq 1$ and $f_1, \ldots, f_d \in L^\infty(\mu)$, the averages*

$$\frac{1}{N}\sum_{n=0}^{N-1} T_1^{[n^{c_1}]}f_1(x)\cdot\ldots\cdot T_d^{[n^{c_d}]}f_d(x) \tag{3}$$

converge as $N \to \infty$ for μ-a.e. $x \in X$ to

$$\mathbb{E}(f_1|\mathcal{I}(T_1))(x)\cdot\ldots\cdot\mathbb{E}(f_d|\mathcal{I}(T_d))(x).$$

This answers Problem 6 of [20] for the special case where $0 < c_d < \ldots < c_1 \leq 1$.[4]

[4] Another possible approach to Corollary 1.3 is to use a variation of [20, Proposition 6.4] to convert (3), via a change of variable, to an average of the same form with $c_1 = 1$. We omit the details.

1.1. **Single convergence implies multiple convergence.** The philosophy of this article is that for a specific nice and wide class of sublinear functions, we have that "single convergence implies the multiple one". More specifically, assuming no commutativity on the transformations T_i, $1 \leq i \leq d$, we show in Theorem 1.1, that averages as in (1) with integer parts of functions of different growth rate from the aforementioned class, converge pointwisely and the limit is the expected one, i.e., the product of conditional expectations, using the fact that the single average converges. The same method also extends to the case when the fastest growing function is linear and we get Theorem 1.2. Our arguments throughout the article are elementary and have a soft touch of ergodic theory.

We introduce some additional notation. For $1 \leq i \leq d$, let (X_i, μ_i, T_i) be measure preserving (m.p.) systems (we also assume that each (X_i, T_i) is a topological dynamical system) and $\mu_i = \int \mu_{[T_i],x} d\mu_i(x)$ be the disintegration of μ_i over its factor $\mathcal{I}(T_i)$ (i.e., the *ergodic decomposition*). For $\vec{x} = (x_1, \dots, x_d)$, let $\mu_{[T_1,\dots,T_d],\vec{x}}$ be the measure on $X_1 \times \cdots \times X_d$ defined by

$$\int\limits_{X_1\times\cdots\times X_d} f_1 \otimes \cdots \otimes f_d \; d\mu_{[T_1,\dots,T_d],\vec{x}} = \mathbb{E}(f_1|\mathcal{I}(T_1))(x_1) \cdot \ldots \cdot \mathbb{E}(f_d|\mathcal{I}(T_d))(x_d)$$

for all $f_1, \dots, f_d \in L^\infty(\mu)$. It is easy to see that $\mu_{[T_1,\dots,T_d],\vec{x}} = \bigotimes_{i=1}^{d} \mu_{[T_i],x_i}$. Let also

$$\lambda_{N,\vec{x}} := \frac{1}{N}\sum_{n=0}^{N-1} \left(T_1^{[a_1(n)]} \times \cdots \times T_d^{[a_d(n)]}\right)\delta_{\vec{x}}, \tag{4}$$

where $\vec{x} \in X_1 \times \cdots \times X_d$ and $\delta_{\vec{x}}$ denotes the Dirac measure at $\vec{x}$.

Denoting with $\mathbb{R}^+$ a set of the form $(c, +\infty)$ for some $c \geq 0$, we define the class

$$S := \left\{ a \in \mathcal{C}^3(\mathbb{R}^+) \;\middle|\; a,\; \frac{1}{a'} \in \mathcal{SL} \text{ and } a^{-1} \in M_1 \cap D_0 \cap D_1 \cap (D_2 \cup M_2) \right\},$$

where for $k \in \mathbb{N} \cup \{0\}$

$$D_k := \left\{ a : \limsup_{x\to \operatorname{sgn}(a^{-1})\cdot\infty} \sup_{h\in[-1,1]} \left| \frac{a^{(k+1)}(x+h)}{a^{(k)}(x)} \right| < \infty \right\},^{5}$$

$$M_k := \left\{ a : \; a^{(k)} \text{ is eventually monotone} \right\},$$

sgn is the *sign* function and

$$\mathcal{SL} = \{a : \; a(x) \prec x\}$$

is the set of *sublinear functions* (recall that $a \prec b$ means $a(x)/b(x) \to 0$ as $x \to \infty$).

Note that the lim sup that appears in the definition of D_k can in general be any $\alpha \in [0, \infty]$. Indeed, for $k = 0$, the function $a(x) = \log x$ gives $\alpha = 0$; to get a specific $\alpha > 0$, pick β with $\beta \exp(\beta) = \alpha$ and let $a(x) = \exp(\beta x)$; while to get $\alpha = \infty$, pick $a(x) = \exp(x^2)$.

Note also that every function $a \in \mathcal{S}$ satisfies $\log x \prec a(x)$ and $a'(x) \to 0$ as $x \to \infty$. Indeed, since $a^{-1} \in M_1$, we have that a' has eventually constant sign, hence integrating

[5] The notation $\operatorname{sgn}(a^{-1})\cdot\infty$ denotes ∞ if a^{-1} is eventually positive, and $-\infty$ if a^{-1} is eventually negative.

the relation $x|a'(x)| \geq M$ (that holds eventually for $M > 0$ since $1/a' \in \mathcal{SL}$) we get $\log x \ll |a(x)|$. Using again that a, $1/a' \in \mathcal{SL}$, we have the claim since

$$\lim_{x\to\infty} \frac{\log x}{a(x)} = \lim_{x\to\infty} \frac{1/a'(x)}{x} = 0 \quad \text{and} \quad \lim_{x\to\infty} a'(x) = \lim_{x\to\infty} \frac{a(x)}{x} = 0.$$

Let $\mathcal{S}^* \subseteq \mathcal{S}$ denote the subclass of functions where $\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^{[a(n)]}x)$ exists pointwisely (a.e.) for every measure preserving system (X, μ, T) and every bounded measurable function f (we will see in Section 4 that in this case the pointwise limit is $\mathbb{E}(f|\mathcal{I}(T))$). We also stress out the fact that $\mathcal{S}^*$ is a strict subset of $\mathcal{S}$ (see § 1.2).

The following result via a density argument will lead us to the proof of Theorem 1.1.

Theorem 1.4. *Let $d \in \mathbb{N}$ and (X_i, μ_i, T_i), $1 \leq i \leq d$ be m.p. systems, $a_i \in \mathcal{S}^*$, $1 \leq i \leq d$ with $a_d \prec \ldots \prec a_1$ and $a'_d \prec \ldots \prec a'_1$, and ν be any coupling of the spaces (X_i, μ_i). Then, for ν-a.e. $\vec{x} \in X_1 \times \cdots \times X_d$, we have that $\lambda_{N,\vec{x}}$ converges to $\mu_{[T_1,\ldots,T_d],\vec{x}}$ as $N \to \infty$.*

Remark. In § 2, we show that $\mathcal{S}$ contains functions a which belong to some Hardy field and satisfy $x^\varepsilon \prec a(x) \prec x$ for some $\varepsilon > 0$. So, by [11, Theorem 3.4] we actually have that each such function a is in $\mathcal{S}^*$ (with convergence to the expected limit, i.e., the conditional expectation $\mathbb{E}(f|\mathcal{I}(T))$) while slow Hardy field functions (as $1 \prec a(x) \prec \log x \exp((\log(\log x))^m)$ for some $0 \leq m < 1$) don't belong to $\mathcal{S}^*$ (see [11, Theorem 3.6]). However, even though Theorem 1.4 in general might fail (take for example $a_i \in \mathcal{S} \backslash \mathcal{S}^*$ for some $1 \leq i \leq d$), we have its validity for uniquely ergodic systems and continuous functions on them, since in this setting the single convergence holds not only for functions in $\mathcal{S}^*$ but for all functions in $\mathcal{S}$ (see Theorem 1.5).

1.2. **Pointwise averages on uniquely ergodic systems.** A topological system (X, T) is *uniquely ergodic* if there is a unique Borel probability measure which is T-invariant. Via the following result for single convergence, which we show in § 5, under the unique ergodicity assumption of the system, we extend Theorem 1.4 (to Theorem 1.6 below).

Theorem 1.5. *Let (X, T) be a uniquely ergodic system with unique T-invariant measure μ and $a \in \mathcal{S}$. Then for any continuous function f on X, for every $x \in X$ we have that*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T^{[a(n)]} f(x) = \int_X f \, d\mu.$$

As we mentioned before, this result has been studied for general systems in [11, 13] along functions which belong to a smaller class of functions than $\mathcal{S}$ (see Theorem 4.1), and in general it might fail for $a \in \mathcal{S}$. One can show (after some elementary calculations) that

$$a(x) = \log x \log(\log x) \in S \setminus S^*$$

($a \notin S^*$ by [11, Theorem 3.6], since a satisfies the hypothesis of the slow growth rate). A similar argument as in Theorem 1.1 extends Theorem 1.5 to multiple averages:

Theorem 1.6. *Let $d \in \mathbb{N}$, (X_i, T_i) be uniquely ergodic systems with unique T_i-invariant measures μ_i, $a_i \in \mathcal{S}$, $1 \leq i \leq d$ with $a_d \prec \ldots \prec a_1$ and $a'_d \prec \ldots \prec a'_1$, and f_i be continuous* *functions on* X_i, $1 \leq i \leq d$. *Then for any coupling* ν *of these systems, the averages*

$$\frac{1}{N}\sum_{n=0}^{N-1} T_1^{[a_1(n)]} f_1(x_1) \cdot \ldots \cdot T_d^{[a_d(n)]} f_d(x_d)$$

converge as $N \to \infty$ *for* ν*-a.e.* $(x_1, \ldots, x_d) \in X_1 \times \cdots \times X_d$ *to*

$$\int f_1 \, d\mu_1 \cdot \ldots \cdot \int f_d \, d\mu_d.$$

Remark. For pointwise averages with a_1 being a linear function, a similar statement to Theorem 1.2 can be derived assuming unique ergodicity of (X_i, μ_i, T_i) for $i \geq 2$ taking averages over functions in $\mathcal{S}$ on continuous functions (the unique ergodicity of (X_1, μ_1, T_1) is not necessary since for the linear a_1 we always have single convergence).

In § 2, we introduce a specific large family $\mathcal{T}$ of sublinear functions which contains properly the Hardy field functions a with $x^{\varepsilon} \prec a(x) \prec x$ for some $\varepsilon > 0$ and is contained properly in the class of Fejér functions (see § 2 or the appendix for definitions of these two important classes of functions). We also show (in § 2 and § 4) that $\mathcal{T}$ is properly contained in $\mathcal{S}^*$ and that for functions of different growth rate in $\mathcal{T}$ we get different growth rate for their derivatives; as a result, our results hold for $a_i \in \mathcal{T}$ of different growth (see Corollary 4.2 for details).

Definitions and notations. A *measure preserving system* $(X, \mathcal{B}, \mu, T_1, \ldots, T_d)$ is a probability space $(X, \mathcal{B}, \mu)$ endowed with measure preserving transformations $T_i \colon X \to X$, meaning that $\mu(T_i^{-1}A) = \mu(A)$ for all $A \in \mathcal{B}$ and $1 \leq i \leq d$. We omit writing the σ-algebra $\mathcal{X}$ when this causes no confusion. Throughout the paper we always assume that X is a compact metric space, $\mathcal{B}$ is its Borel σ-algebra and μ is a Borel measure. We let $M(X)$ denote the convex set of probability measures on X which is compact for the weak*-topology. A *coupling* λ of two probability spaces (X_1, μ_1) and (X_2, μ_2) is a measure in $M(X \times Y)$, whose marginals are equal to μ_1 and μ_2 respectively. A *joining* of two measure preserving systems $(X_1, \mu_1, T_1, \ldots, T_d)$ and $(X_2, \mu_2, S_1, \ldots, S_d)$ is a coupling of (X_1, μ_1) and (X_2, μ_2) that is invariant under the diagonal transformations $T_1 \times S_1, \ldots, T_d \times S_d$ (these definitions extend naturally to k systems, $k \geq 2$). A *factor map* between two measure preserving systems $(X, \mu, T_1, \ldots, T_d)$ and $(Y, \nu, S_1, \ldots, S_d)$ is a measurable function $\pi \colon X \to Y$ such that the push-forward measure $\pi_* \mu$ is equal to ν and $\pi \circ T_i = S_i \circ \pi$, $1 \leq i \leq d$. Finally, for two quantities a and b, we write $a \ll b$ if there exists $c > 0$ such that $|a| \leq c \cdot |b|$, and $a \ll_{\delta_1, \ldots, \delta_r} b$ if there exists $c \equiv c(\delta_1, \ldots, \delta_r) > 0$ with $|a| \leq c \cdot |b|$.

In this paper, we study various classes of functions, whose definitions spread throughout the article. For the reader's convenience, we summarize all these definitions, and their connection, in the appendix.

Acknowledgements. We thank Nikos Frantzikinakis for bringing us to the alternative proof of Theorem 1.1 for functions of different growth from the class $\mathcal{LE}_{\varepsilon}$. We especially thank the anonymous referee whose wide insight of the topic allowed us to add important historical references and motivations to the problem we studied.

2. A nice class of sublinear functions

In this section we define a nice class of sublinear functions, $\mathcal{T}$, first appeared in [9], which we'll show that is a subclass of $\mathcal{S}^*$ defined in the previous section. More specifically, we show in this section that $\mathcal{T}$ is a proper subset of $\mathcal{S}$ and then, in § 4, we prove that $\mathcal{T}$ is a proper subset of $\mathcal{S}^*$.

Let

$$\mathcal{R} := \Big\{ a \in \mathcal{C}^3(\mathbb{R}^+) : \text{ the limits } \lim_{x\to\infty} \frac{xa'(x)}{a(x)}, \lim_{x\to\infty} \frac{xa''(x)}{a'(x)}, \text{ and } \lim_{x\to\infty} \frac{xa'''(x)}{a''(x)} \text{ exist in } \mathbb{R} \Big\}$$

and

$$\mathcal{T} := \Big\{ a \in \mathcal{R} : \lim_{x\to\infty} \frac{xa'(x)}{a(x)} \in (0,1) \text{ or } \lim_{x\to\infty} \frac{xa'(x)}{a(x)} = 1 \text{ and } \lim_{x\to\infty} a'(x) = 0 \text{ monotonicaly} \Big\}.$$

We start with the connection between $\mathcal{T}$ and some important classes of sublinear functions.

2.0.1. *Fejér functions.* A function $a \in C^1((c,\infty)), c \geq 0$, is a *Fejér* function if: (i) $a'(x)$ tends monotonically to 0 as $x \to \infty$; and (ii) $\lim_{x\to\infty} x|a'(x)| = \infty$. We denote with $\mathcal{F}$ the set of all Fejér functions. Note that every $a \in \mathcal{F}$ is eventually monotone and satisfies the growth rate conditions $\log x \prec a(x) \prec x$ (see [9] for more details on Fejér functions).

Remark. Since we can find Fejér functions which are not even $\mathcal{C}^2$, so, we have $\mathcal{S} \subsetneq \mathcal{F}$.

2.0.2. *Hardy field functions.* Let B be the collection of equivalence classes of real valued functions defined on some halfline (c,∞), $c \geq 0$, where two functions that agree eventually are identified. These equivalence classes are called *germs* of functions. A *Hardy field* is a subfield of the ring $(B, +, \cdot)$ that is closed under differentiation.[6] (See [21] for more details on Hardy field functions.)

If $\mathcal{H}$ is the union of all Hardy fields, every element of $\mathcal{H}$ has eventually constant sign, from which it follows that if $a \in \mathcal{H}$, then a is eventually monotone and the limit $\lim_{x\to\infty} a(x)$ exists (possibly infinite–as in the $\mathcal{LE}_\varepsilon$ class of functions). For functions $a,\ b \in \mathcal{H}$ with $b \neq 0$, it follows that the asymptotic growth ratio $\lim_{x\to\infty} a(x)/b(x)$ exists (possibly infinite), fact that will often justify the use of L'Hospital's rule. For some $\varepsilon > 0$ we denote with $\mathcal{H}_\varepsilon$ the set of functions a which belongs to some Hardy field $\mathcal{H}$ and satisfy $x^\varepsilon \prec a(x) \prec x$.[7]

Remark. $\mathcal{H}_\varepsilon$ is a proper subset of $\mathcal{T}$ and $\mathcal{T}$ is a proper subset of $\mathcal{F}$.

Indeed, if $a \in \mathcal{H}_\varepsilon$, then $a'(x) \to 0$ monotonically and

$$\lim_{x\to\infty} \frac{xa'(x)}{a(x)} = \lim_{x\to\infty} \frac{\log|a(x)|}{\log x} \in [\varepsilon, 1],$$

so $a \in \mathcal{T}$ (the other limits exist by the properties of a Hardy field). If $a \in \mathcal{T}$ and $\lim_{x\to\infty} \frac{xa'(x)}{a(x)} \in (0,1)$, then $a \in \mathcal{F}$ from [9, Lemma 2.1], while if $\lim_{x\to\infty} \frac{xa'(x)}{a(x)} = 1$ and $a'(x) \to 0$ monotonically, then $a \in \mathcal{F}$ by [9, Lemma 2.2].

---

[6]We use the word *function* when we refer to elements of B (understanding that all the operations defined and statements made for elements of B are considered only for sufficiently large values of $x \in \mathbb{R}^+$).

[7]We say that the Hardy field functions a which satisfy $\log x \prec a(x) \prec x$ are of *polynomial degree 0.*

Since $\log^\alpha x \in \mathcal{F}\setminus\mathcal{T}$ for all $\alpha > 1$, and $a(x) = x^{1/2}(2+\cos\sqrt{\log x}) \in \mathcal{T}\setminus\mathcal{H}_\varepsilon$ for all $\varepsilon > 0$ (the derivative of a^2 is not (eventually) monotone by [9, Section 1]), the claim follows.

The following lemma provides a sufficient condition for a function to belong to D_0 (recall the definition from § 1).

Lemma 2.1. *Let $a \in C^1(\mathbb{R})$ and α, $\beta \in \mathbb{R}$ with $\alpha \leq \frac{xa'(x)}{a(x)} \leq \beta$ eventually. Then for any $H > 0$ and large enough x, we have that*

$$\sup_{h\in[-H,H]} \left|\frac{a'(x+h)}{a(x)}\right| \ll_{\alpha,\beta} |x|^{\beta-\alpha-1}. \tag{5}$$

Proof. Since $\alpha \leq \frac{xa'(x)}{a(x)} \leq \beta$ eventually, following the proof of [9, Lemma 2.2] we have that there exist positive constants C_α and C_β such that $C_\alpha|x|^\alpha \leq |a(x)| \leq C_\beta|x|^\beta$. Using the fact that eventually we have $\left|\frac{a'(x)}{a(x)}\right| \leq \frac{\max\{|\alpha|,|\beta|\}}{|x|}$, we get

$$\sup_{h\in[-H,H]} \left|\frac{a'(x+h)}{a(x)}\right| \leq \max\{|\alpha|,|\beta|\}C_\alpha^{-1}C_\beta \cdot |x|^{\beta-\alpha-1} \max\left\{|1-H/x|^{\beta-1},\ |1+H/x|^{\beta-1}\right\}$$

and the result follows. □

Remark. If the constants α, β of Lemma 2.1 satisfy $\beta-\alpha-1 < 0$ (a special case of this is when $\lim_{x\to\pm\infty} \frac{xa'(x)}{a(x)} \in \mathbb{R}$) then the limit of (5) exists (as $x \to \infty$ or $-\infty$) and it is 0.

Since $\mathcal{T} \subseteq \mathcal{F}$, every $a \in \mathcal{T}$ is sublinear, (eventually) monotone, and has the property that $\frac{1}{a'(x)} \prec x$. Moreover, we have:

Proposition 2.2. *$\mathcal{T}$ is a proper subset of $\mathcal{S}$.*

Proof. Assuming that $a \in \mathcal{T}$ is eventually positive (the other case follows analogously), using the fact that $\lim_{x\to\infty} \frac{xa'(x)}{a(x)} = \alpha \in (0,1]$, we get $\lim_{x\to\infty} \frac{xa''(x)}{a'(x)} = \alpha-1$, and $\lim_{x\to\infty} \frac{xa'''(x)}{a''(x)} = \alpha-2$ (the properties of $\mathcal{T}$ allows us to use L'Hopital's rule), so we have that

$$\lim_{x\to\infty} \frac{x(a^{-1})'(x)}{a^{-1}(x)} = \lim_{y\to\infty} \frac{a(y)}{ya'(y)} = \frac{1}{\alpha},\quad \lim_{x\to\infty} \frac{x(a^{-1})''(x)}{(a^{-1})'(x)} = -\lim_{y\to\infty} \frac{a''(y)a(y)}{(a'(y))^2} = \frac{1}{\alpha}-1,\quad \text{and}$$

$\lim_{x\to\infty} \frac{x(a^{-1})'''(x)}{(a^{-1})''(x)} = \lim_{y\to\infty} \left(\frac{a(y)a'''(y)}{a'(y)a''(y)} - 3\frac{a(y)a''(y)}{(a'(y))^2}\right) = \frac{1}{\alpha}-2$. The previous remark implies that $a \in \mathcal{S}$. Since $a(x) = \log x \log(\log x) \in \mathcal{S}\setminus\mathcal{T}$ (a satisfies all the properties of S – we skip all the elementary calculations – and $xa'(x)/a(x) \to 0$ as $x \to \infty$), we have the claim. □

Remark. The class $\mathcal{T}$ misses not only slow functions as $a_1(x) = \log x \log(\log x)$ from S (i.e., functions a with $1 \prec a \prec x^\varepsilon$ for all $\varepsilon > 0$) but also functions as

$$a_2(x) = x^\alpha(4/\alpha + \sin\log x)^3$$

for $0 < \alpha < 1/20$ for which we have $a_2 \notin \mathcal{T}$ since the ratio $xa_2'(x)/a_2(x)$ doesn't have a limit as $x \to \infty$ (it is bounded though between positive numbers since $\alpha - \frac{3\alpha}{4-\alpha} \leq \frac{xa_2'(x)}{a_2(x)} \leq \alpha + \frac{3\alpha}{4-\alpha}$) and a_2 is not slow since $x^{\alpha/2} \prec a_2(x)$. Since $a_2(x) \prec x$, $a_2 \in M_1 \subseteq M_0$ with $a_2'(x) \to 0$ and $1/a_2'(x) \prec x$, we actually have that $a_2 \in \mathcal{F} \setminus \mathcal{T}$. We also have that $a_2^{-1} \in D_0 \cap D_1 \cap D_2$ (the lim sup that appear in these sets are limits and equal to 0 – we skip all the elementary calculations), hence $a_2(x) \in \mathcal{S} \setminus \mathcal{T}$.

We will actually show in § 4 that a_2 is a special function as $a_2 \in S^*$.

The following lemma informs us that for a function in D_0 we have that the ratios of horizontal translations are bounded.

Lemma 2.3. *Let $a \in \mathcal{C}^1(\mathbb{R})$ and $H > 0$ with $\limsup_{x\to\infty} \sup_{h\in[-H,H]} \left|\frac{a'(x+h)}{a(x)}\right| < \infty$ (resp. $x \to -\infty$). Then, the quantities $\left|\frac{a(x+\rho)}{a(x)}\right|$ are eventually bounded for all $-H \leq \rho \leq H$.*

Proof. Let $\rho \in [-H, H]$. The mean value theorem furnishes a point $h_\rho \in [-H, H]$ with

$$\left|\frac{a(x+\rho)}{a(x)} - 1\right| = \left|\frac{a'(x+h_\rho)}{a(x)}\right|$$

from which we have the result by our hypothesis. □

We close this section with a fact about the sets D_k. We show that for a function of S if the lim sup that appears in D_0 is a limit, then it has to be equal to 0.

Proposition 2.4. *If $a \in S$ with $\lim_{x\to\infty} \sup_{h\in[-1,1]} \left|\frac{(a^{-1})'(x+h)}{a^{-1}(x)}\right| = \alpha \in \mathbb{R}$, then $\alpha = 0$.*

Proof. We assume to the contrary that $\alpha > 0$ (the case where $a < 0$ is analogous) and we let $0 < \varepsilon < \alpha$. Using the hypothesis, there exists $M > 0$ such that $\frac{(a^{-1})'(x+1)}{a^{-1}(x)} > \alpha - \varepsilon$ for all $x > M$, or equivalently

$$\frac{1}{a^{-1}(x+1)a'(a^{-1}(x+1))} > (\alpha - \varepsilon)\frac{a^{-1}(x)}{a^{-1}(x+1)},$$

from which (using the fact that $1/a'(x) \prec x$ and by taking the lim sup) we have that $\lim_{x\to\infty} \frac{a^{-1}(x-1)}{a^{-1}(x)} = 0$. By the mean value theorem, there exists $\xi_x \in [x-1, x]$ with

$$\left|\frac{a^{-1}(x-1)}{a^{-1}(x)} - 1\right| = \frac{(a^{-1})'(\xi_x)}{a^{-1}(x)} \leq \frac{(a^{-1})'(x+1)}{a^{-1}(x)}.$$

Letting $x \to \infty$, we get $\alpha \geq 1$. By Lemma 2.3, there exists $c > 0$ such that

$$\frac{(a^{-1})'(x)}{a^{-1}(x)} \geq c\frac{(a^{-1})'(x+1)}{a^{-1}(x)} > c(\alpha - \varepsilon).$$

by integrating and solving for $a(x)$ this relation we get $a(x) \leq c_1 \log x + c_2$ for some c_1, c_2 constants with $c_1 > 0$. Then we have

$$0 < \frac{1}{c_1} \leq \lim_{x\to\infty} \frac{\log x}{a(x)} = \lim_{x\to\infty} \frac{1}{xa'(x)} = 0,$$

a contradiction. The claim now follows. □

3. The key invariant properties

In the following two subsections, we state and prove the invariant arguments that play a central role in our study and we be used in the proof of our main results.

3.1. **The sublinear case.** In this subsection we develop the main tool, Lemma 3.1, in order to prove Theorem 1.4, which we'll use in the proof of Theorem 1.1.

Recall from § 1 that $\lambda_{\vec{x}}$ is any weak limit in $M(X_1 \times \cdots \times X_d)$ of

$$\lambda_{N,\vec{x}} = \frac{1}{N} \sum_{n=0}^{N-1} \left(T_1^{[a_1(n)]} \times \cdots \times T_d^{[a_d(n)]}\right) \delta_{\vec{x}}.$$

Lemma 3.1. *Let $d \in \mathbb{N}$ and $a_i \in \mathcal{S}$, $1 \leq i \leq d$, with $a_d \prec \ldots \prec a_1$ and $a_d' \prec \ldots \prec a_1'$. Then $\lambda_{\vec{x}}$ is invariant under $T_1 \times id \times \cdots \times id$.*

Proof. We can assume without loss of generality that a_1 is eventually positive (the other case is analogous). For $\vec{b} = (b_1, b_2, \ldots, b_d) \in \mathbb{Z}^d$, write $\vec{b}_* = (b_2, \ldots, b_d)$ and let

$$\mathcal{U}_{b_1,\vec{b}_*} := \left|\left\{n \in \{1, \ldots, N\} : \ b_i \leq a_i(n) < b_i + 1 \ \forall\, 1 \leq i \leq d\right\}\right|, \ \delta_{b_1,\vec{b}_*} := \delta_{T_1^{b_1}x_1} \times \cdots \times \delta_{T_d^{b_d}x_d}.$$

With this notation, we have that $\lambda_{N,\vec{x}} = \frac{1}{N} \sum_{(b_1,\vec{b}_*):\, \mathcal{U}_{b_1,\vec{b}_*} \neq 0} \mathcal{U}_{b_1,\vec{b}_*} \delta_{b_1,\vec{b}_*}$, and if

$$\tilde{\lambda}_{N,\vec{x}} := (T_1 \times id \times \cdots \times id)(\lambda_{N,\vec{x}}) = \frac{1}{N} \sum_{(b_1,\vec{b}_*):\, \mathcal{U}_{b_1-1,\vec{b}_*} \neq 0} \mathcal{U}_{b_1-1,\vec{b}_*} \delta_{b_1,\vec{b}_*},$$

then,

$$\lambda_{N,\vec{x}} - \tilde{\lambda}_{N,\vec{x}} = \frac{1}{N} \sum_{(b_1,\vec{b}_*):\, \mathcal{U}_{b_1,\vec{b}_*} \neq 0,\, \mathcal{U}_{b_1-1,\vec{b}_*} \neq 0} \left(\mathcal{U}_{b_1,\vec{b}_*} - \mathcal{U}_{b_1-1,\vec{b}_*}\right) \delta_{b_1,\vec{b}_*} \tag{6}$$

$$+ \frac{1}{N} \sum_{(b_1,\vec{b}_*):\, \mathcal{U}_{b_1,\vec{b}_*} \neq 0,\, \mathcal{U}_{b_1-1,\vec{b}_*} = 0} \mathcal{U}_{b_1,\vec{b}_*} \delta_{b_1,\vec{b}_*} \tag{7}$$

$$- \frac{1}{N} \sum_{(b_1,\vec{b}_*):\, \mathcal{U}_{b_1,\vec{b}_*} = 0,\, \mathcal{U}_{b_1-1,\vec{b}_*} \neq 0} \mathcal{U}_{b_1-1,\vec{b}_*} \delta_{b_1,\vec{b}_*}. \tag{8}$$

We will study each term separately. Terms (7) and (8) are treated similarly, so we only study the Terms (6) and (7).

Term (6)**:** For $1 \leq i \leq d$, let

$$C_{a_i,b_i} := \min\{a_i^{-1}(b_i),\ a_i^{-1}(b_i+1)\} \ \text{ and } \ C'_{a_i,b_i} := \max\{a_i^{-1}(b_i),\ a_i^{-1}(b_i+1)\}.$$

The conditions $\mathcal{U}_{b_1,\vec{b}_*} \neq 0$ and $\mathcal{U}_{b_1-1,\vec{b}_*} \neq 0$ imply that

$$\max_{1\leq i\leq d}\{C_{a_i,b_i}\} < \min_{1\leq i\leq d}\{C'_{a_i,b_i}\}; \quad \text{and} \tag{9}$$

$$\max_{2\leq i\leq d}\{C_{a_1,b_1-1},\ C_{a_i,b_i}\} < \min_{2\leq i\leq d}\{C'_{a_1,b_1-1},\ C'_{a_i,b_i}\}. \tag{10}$$

Using the Relations (9) and (10) we get

$$C_{a_i,b_i} < a_1^{-1}(b_1) < C'_{a_i,b_i} \ \text{ for all } \ 2 \leq i \leq d,$$

from which we have

$$b_i < a_i \circ a_1^{-1}(b_1) < b_i + 1,$$

hence $b_i = [a_i \circ a_1^{-1}(b_1)]$ for all $2 \leq i \leq d$.

By the mean value theorem, for some $\xi_b \in (b_1 - 1, b_1 + 1)$ we have

$$\left|\mathcal{U}_{b_1,\vec{b}_*} - \mathcal{U}_{b_1-1,\vec{b}_*}\right| \leq \left|a_1^{-1}(b_1+1) - 2a_1^{-1}(b_1) + a_1^{-1}(b_1-1)\right| + 2 = \left|(a_1^{-1})''(\xi_b)\right| + 2.$$

Since $a_1^{-1} \in D_2 \cup M_2$ for large enough b_1, we have that $\left|(a_1^{-1})''(\xi_b)\right| \ll \left|(a_1^{-1})''(b_1)\right|$ (or $\left|(a_1^{-1})''(b_1 \pm 1)\right|$ which is studied similarly). So,

$$\frac{1}{N}\sum_{b_1=1}^{[a_1(N)]-1}\left|\mathcal{U}_{b_1,\vec{b}_*} - \mathcal{U}_{b_1-1,\vec{b}_*}\right| \ll \frac{1}{N}\sum_{b_1=1}^{[a_1(N)]-1}\left|(a_1^{-1})''(b_1)\right|$$

(the term $2[a_1(N)]/N$ is null by sublinearity of a_1). By the same property, for large enough b_1 and all $t \in [b_1, b_1+1]$, we have that $\left|(a_1^{-1})''(b_1)\right| \ll \left|(a_1^{-1})''(t)\right|$, so

$$\left|(a_1^{-1})''(b_1)\right| \ll \int_{b_1}^{b_1+1}\left|(a_1^{-1})''(t)\right|\, dt.$$

Hence,

$$\frac{1}{N}\sum_{b_1=1}^{[a_1(N)]-1}\left|(a_1^{-1})''(b_1)\right| \ll \frac{1}{N}\int_{*}^{[a_1(N)]}\left|(a_1^{-1})''(t)\right|\, dt \ll \frac{\left|(a_1^{-1})'([a_1(N)])\right|}{N} \tag{11}$$

(where we used the fact that $a_1^{-1} \in D_2$ so $(a_1^{-1})''$ has constant sign – in our case here is positive). Using the fact that $a_1^{-1} \in D_1$, we have that $(a_1^{-1})'([a_1(N)])/(a_1^{-1})'(a_1(N))$ is bounded for large N, so the right hand side of (11) is bounded by a constant multiple of $(Na_1'(N))^{-1}$, which goes to 0 as $N \to \infty$, and consequently Term (6) goes to 0 as $N \to \infty$.

Term (7)**:** The conditions $\mathcal{U}_{b_1,\vec{b}_*} \neq 0$ and $\mathcal{U}_{b_1-1,\vec{b}_*} = 0$ imply that

$$\max_{1\leq i\leq d}\{C_{a_i,b_i}\} < \min_{1\leq i\leq d}\{C'_{a_i,b_i}\}, \quad \text{and} \tag{12}$$

(13) as in (10) but with $C_{a_i,b_i} > a_1^{-1}(b_1) - 1$ for at least one $2 \leq i \leq d$.

If $\{i_0 = 1, i_1, \ldots, i_r\}$ is the set of indices for which $C_{a_i,b_i} > a_1^{-1}(b_1) - 1$, let

$$B_{\{i_0,i_1,\ldots,i_r\}} := \{\vec{b}_* :\ a_1^{-1}(b_1) \leq C_{a_{i_1},b_{i_1}} \leq \ldots \leq C_{a_{i_r},b_{i_r}}\}.$$

Then we have that

$$a_1^{-1}(b_1) - 1 < C_{a_{i_j},b_{i_j}} \leq C_{a_{i_r},b_{i_r}} < C'_{a_{i_j},b_{i_j}} \quad \text{for all } \ 0 < j \leq r,$$

from which we get

$$b_{i_j} \leq a_{i_j} \circ a_{i_r}^{-1}(b_{i_r} + e_{i_r}) < b_{i_j} + 1,$$

for some $e_{i_r} \in \{0, 1\}$. Hence, $b_{i_j} = \left[a_{i_j} \circ a_{i_r}^{-1}(b_{i_r} + e_{i_r})\right]$ for all $0 < j \leq r$.

For $j = 0$, we have

$$a_1^{-1}(b_1) - 1 < C_{a_{i_r},b_{i_r}} < a_1^{-1}(b_1 + 1),$$

or

$$a_1(C_{a_{i_r},b_{i_r}}) - 1 < b_1 < a_1(C_{a_{i_r},b_{i_r}} + 1).$$

So, for each b_{i_r} fixed, $b_1 = [a_1(C_{a_{i_r},b_{i_r}})]$ or $b_1 = [a_1(C_{a_{i_r},b_{i_r}} + 1)]$.

For $i \notin \{i_1, \dots, i_r\}$, we have

$$C_{a_i,b_i} \leq a_1^{-1}(b_1) < C_{a_{i_r},b_{i_r}} < C'_{a_i,b_i},$$

or

$$b_i \leq a_i \circ a_{i_r}^{-1}(b_{i_r} + e_{i_r}) < b_i + 1,$$

for some $e_{i_r} \in \{0, 1\}$, hence $b_i = \left[a_i \circ a_{i_r}^{-1}(b_{i_r} + e_{i_r})\right]$.

The average of $\left|\mathcal{U}_{b_1,\vec{b}_*}\right|$ can be split into finitely many sums over $(b_1, \vec{b}_*)$ for $\vec{b}_* \in B_{\{i_0,i_1,\dots,i_r\}}$ for some $\{i_0, i_1, \dots, i_r\}$. For $\vec{b}_* \in B_{\{i_0,i_1,\dots,i_r\}}$, $\left|\mathcal{U}_{b_1,\vec{b}_*}\right|$ is bounded by

$$\left|a_1^{-1}(b_1 + 1) - a_1^{-1}(b_1)\right| + 1 \ll \left|(a_1^{-1})'(b_1)\right| + 1 = \left|(a_1^{-1})'([a_1 \circ a_{i_r}^{-1}(b_{i_r} + e_{i_r})])\right| + 1.$$

This approximation follows by the mean value theorem and the fact that $a_1^{-1} \in D_1$. Hence, every average is estimated by a constant multiple of

$$\frac{1}{N}\left|\int_*^{[a_{i_r}(N)]} \left|(a_1^{-1})'([a_1 \circ a_{i_r}^{-1}(t)])\right| \, dt\right| \ll \frac{1}{N}\left|\int_*^{[a_{i_r}(N)]} \left|(a_1^{-1})'(a_1 \circ a_{i_r}^{-1}(t))\right| \, dt\right|, \tag{14}$$

where we used the fact that $a_{i_r} \prec a_1$ and the sublinearity of a_{i_r}. So, the right hand side of (14) is bounded by

$$\frac{1}{N}\left|\int_*^{[a_{i_r}(N)]} \frac{1}{\left|a_1'(a_{i_r}^{-1}(t))\right|} \, dt\right| = \frac{1}{N}\left|\int_*^{a_{i_r}^{-1}([a_{i_r}(N)])} \left|\frac{a_{i_r}'(t)}{a_1'(t)}\right| \, dt\right|.$$

Since $a_{i_r}^{-1} \in D_0$ we have $a_{i_r}^{-1}([a_{i_r}(N)])/N \to 1$ as $N \to \infty$, hence the last integral goes to 0, since the integrand is arbitrarily small for large N because of $a_{i_r}' \prec a_1'$.

This proves that the Terms (7) and (8) go to 0 as $N \to \infty$ and we get the conclusion. □

3.2. **Linearity of the fastest term.** In this subsection we deal with the case where the fastest function is linear. If the leading coefficient is rational, we can work as in Lemma 3.1 while if the leading coefficient is irrational, with an explicit example (see the remark after Lemma 3.2) we show that the invariant property of Lemma 3.1 can fail, but one can still deal with this case via the suspension flow and Birkhoff's ergodic theorem (see § 4.1).

3.2.1. **Rational leading coefficient**. If $a_1(n) = kn + \ell$, with $k \in \mathbb{Q} \setminus \{0\}$, $\ell \in \mathbb{R}$, it suffices to cover the case where $a_1(n) = pn$ with $p \in \mathbb{Z} \setminus \{0\}$ (see details in the proof of Theorem 1.2). As before, let $\lambda_{\vec{x}}$ be a weak limit of $\lambda_{N,\vec{x}}$.

Lemma 3.2. *For $d \in \mathbb{N}$, let $a_1(n) = pn$, $p \in \mathbb{Z} \setminus \{0\}$ and a_i, $2 \le i \le d$ be functions with $a_d \prec \ldots \prec a_1$. Then $\lambda_{\vec{x}}$ is invariant under $T_1^p \times id \times \cdots \times id$.*

Proof. We can assume that $p > 0$ (the case where $p < 0$ is analogous), hence a_1 is eventually positive. For $\vec{b} = (b_1, b_2, \ldots, b_d) \in \mathbb{Z}^d$ we write $\vec{b}_* = (b_2, \ldots, b_d)$ and we set

$\mathcal{U}_{b_1,\vec{b}_*} = \begin{cases} 1 \ , & \text{if } [a_i(b_1)] = b_i \ \forall\, 2 \le i \le d \\ 0 \ , & \text{otherwise} \end{cases}$, and $\delta_{b_1,\vec{b}_*} = \delta_{T_1^{[a_1(b_1)]}x_1} \times \delta_{T_2^{b_2}x_2} \times \cdots \times \delta_{T_d^{b_d}x_d}$

(note that the position of b_1 here is not the same as in the proof of Lemma 3.1). With this notation, we have that $\lambda_{N,\vec{x}} = \frac{1}{N} \sum_{(b_1,\vec{b}_*):\, \mathcal{U}_{b_1,\vec{b}_*} \neq 0} \mathcal{U}_{b_1,\vec{b}_*} \delta_{b_1,\vec{b}_*}$, and

$$\tilde{\lambda}_{N,\vec{x}} := (T_1^p \times id \times \cdots \times id)(\lambda_{N,x}) = \frac{1}{N} \sum_{(b_1,\vec{b}_*):\, \mathcal{U}_{b_1-1,\vec{b}_*} \neq 0} \mathcal{U}_{b_1-1,\vec{b}_*} \delta_{b_1,\vec{b}_*},$$

where we used the fact that $[a_1(b_1 - 1)] + p = [a_1(b_1)]$. Then,

$$\lambda_{N,\vec{x}} - \tilde{\lambda}_{N,\vec{x}} = \frac{1}{N} \sum_{(b_1,\vec{b}_*):\, \mathcal{U}_{b_1,\vec{b}_*} \neq 0,\, \mathcal{U}_{b_1-1,\vec{b}_*} \neq 0} \Big(\mathcal{U}_{b_1,\vec{b}_*} - \mathcal{U}_{b_1-1,\vec{b}_*}\Big) \delta_{b_1,\vec{b}_*} \tag{15}$$

$$+ \frac{1}{N} \sum_{(b_1,\vec{b}_*):\, \mathcal{U}_{b_1,\vec{b}_*} \neq 0,\, \mathcal{U}_{b_1-1,\vec{b}_*} = 0} \mathcal{U}_{b_1,\vec{b}_*} \delta_{b_1,\vec{b}_*} \tag{16}$$

$$- \frac{1}{N} \sum_{(b_1,\vec{b}_*):\, \mathcal{U}_{b_1,\vec{b}_*} = 0,\, \mathcal{U}_{b_1-1,\vec{b}_*} \neq 0} \mathcal{U}_{b_1-1,\vec{b}_*} \delta_{b_1,\vec{b}_*}. \tag{17}$$

We study the Terms (15) and (16) as in the sublinear case.

Term (15)**:** Since each $\mathcal{U}_{b_1,b_*}$ is at most 1, $\mathcal{U}_{b_1,b_*} = \mathcal{U}_{b_1-1,b_*} = 1$, so Term (15) equals to 0.

Term (16)**:** Let

$$C_{a_i,b_i} := \min\{a_i^{-1}(b_i),\ a_i^{-1}(b_i + 1)\} \ \text{ and } \ C'_{a_i,b_i} := \max\{a_i^{-1}(b_i),\ a_i^{-1}(b_i + 1)\}.$$

The conditions $\mathcal{U}_{b_1,\vec{b}_*} \neq 0$ and $\mathcal{U}_{b_1-1,\vec{b}_*} = 0$ imply that

$$\max_{1 \le i \le d} \{C_{a_i,b_i}\} \le b_1 < \min_{1 \le i \le d} \{C'_{a_i,b_i}\}, \quad \text{and} \tag{18}$$

$$C_{a_i,b_i} \ge b_1 - 1 \quad \text{for at least one } \ 2 \le i \le d. \tag{19}$$

If $\{i_1, \ldots, i_r\}$ is the set of indices for which $C_{a_i,b_i} \ge b_1 - 1$, let

$$B_{\{i_1,\ldots,i_r\}} := \{\vec{b}_* :\ b_1 - 1 \le C_{a_{i_1},b_{i_1}} \le \ldots \le C_{a_{i_r},b_{i_r}}\}.$$

As in Lemma 3.1 we get $b_1 = [C_{a_{i_r},b_{i_r}}]$ or $b_1 = [C_{a_{i_r},b_{i_r}}] + 1$ and $b_i = \left[a_i \circ a_{i_r}^{-1}(b_{i_r} + e_{i_r})\right]$, $2 \le i \le d$ for some $e_{i_r} \in \{0, 1\}$.

For fixed b_{i_r} there are at most 2^r terms $\left|\mathcal{U}_{b_1,\vec{b}_*}\right|$ in (16) whose last index equals to b_{i_r}. Since $|b_{i_r}| \leq a_{i_r}(N)$, (16) is bounded by a constant multiple of $a_{i_r}(N)/N$, which goes to 0 as $N \to \infty$ and hence the conclusion of the lemma follows. □

Remark. In the previous proof we essentially used the relation $[a_1(b_1-1)]+p=[a_1(b_1)]$, which is not in general true for expressions of the form $a(t)=\gamma n+\ell$, $\gamma \notin \mathbb{Q}$. Actually, Lemma 3.2 cannot even be extended to the $d=1$ case if $a_1(n)=\gamma n+\ell, \gamma \notin \mathbb{Q}$.

Indeed, let $d=1$ and $(\mathbb{T},\mu)$ be the 1 dimensional torus endowed with the Lebesgue measure. Let $T\colon \mathbb{T}\to\mathbb{T}$ with $Tx=\gamma^{-1}+x$. Then ($\{\cdot\}$ denotes the fractional part)

$$T^{[\gamma n+\ell]}x=\gamma^{-1}[\gamma n+\ell]+x=\gamma^{-1}\ell+x-\gamma^{-1}\{\gamma n\}.$$

Since $(\{\gamma n\})_n$ is equidistributed on $\mathbb{T}$, for all $f\in L^\infty(\mu)$, we have

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{[\gamma n+\ell]}f(x)=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}f(\gamma^{-1}\ell+x-\gamma^{-1}\{\gamma n\})=\int_{\gamma^{-1}(\ell-1)+x}^{\gamma^{-1}\ell+x}f(y)\,d\mu(y)$$

and

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{[\gamma n+\ell]+1}f(x)=\int_{\gamma^{-1}+x}^{\gamma^{-1}(\ell+1)+x}f(y)\,d\mu(y).$$

So λ_x is obviously not T-invariant. Nevertheless, we will show in the next section that there is an explicit expression for the limit $\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{[\gamma n+\ell]}f(x), \gamma\notin\mathbb{Q}$.

4. Single pointwise averages

Let (X,μ,T) be a measure preserving system. We start this section by justifying why for $a\in\mathcal{S}^*$ the average $\frac{1}{N}\sum_{n=0}^{N-1}T^{[a(n)]}f(x)$ converges to $\mathbb{E}(f|\mathcal{I}(T))(x)$ for μ-a.e $x\in X$ for every bounded measurable function f. Since we assume that the average does converge (by the definition of $\mathcal{S}^*$) it suffices to show that it converges in $L^2(\mu)$ to $\mathbb{E}(f|\mathcal{I}(T))$. By writing $f=\mathbb{E}(f|\mathcal{I}(T))+g$ where $g=f-\mathbb{E}(f|\mathcal{I}(T))$, it suffices to show that the average goes to 0 for a function g with 0 conditional expectation with respect to $\mathcal{I}(T)$. To prove that, it suffices to show it for coboundary functions, i.e., for functions of the form $g-g\circ T$. By decomposing $\mu=\int\delta_x d\mu(x)$, it is a consequence of Lemma 3.1 that the average does converge to 0 in $L^2(\mu)$ for coboundary functions.

We now establish all the pointwise limits of (1) for $d=1$ for the class $\mathcal{T}\cup\mathcal{L}_*$, where $\mathcal{L}_*$ denotes the set of linear functions with non-zero leading coefficient.

Theorem 4.1. *Let (X,μ,T) be a measure preserving system and $f\in L^\infty(\mu)$. If $a\in\mathcal{T}$, then for μ-a.e. $x\in X$ we have*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{[a(n)]}f(x)=\mathbb{E}(f|\mathcal{I}(T))(x).$$

Theorem 4.1 is proved in [11, Theorem 3.4] for $a \in \mathcal{H}_\varepsilon$ but the same result holds for $a \in \mathcal{T}$. For completeness we sketch the idea of the proof.

By [9, Theorem 3.5], if a satisfies some properties, then the average of a sequence $(x_n)_n$ along $([a(n)])_n$, i.e., $\frac{1}{N}\sum x_{[a(n)]}$, converges to 0 as long as the usual average $\frac{1}{N}\sum x_n$ converges to 0. The proof that the assumptions of [9, Theorem 3.5] are satisfied for functions from $\mathcal{T}$ can be found in [9, Lemma 2.5]. More specifically, [9, Lemma 2.5] covers the case where $a > 0$ while the case where $a < 0$ follows by the fact that $[a(n)] = -[-a(n)] - 1$ in a set of density 1 since $[a(n)] = -[-a(n)]$ only happens when $a(n)$ is an integer, which up to time N, happens at most $a(N)$ times. The claim now follows by the sublinearity of a.

Remarks. (1) Combining this result with Proposition 2.2, we get that $\mathcal{T}$ is a subset of $\mathcal{S}^*$. In fact, $\mathcal{S}^*$ properly contains $\mathcal{T}$. Indeed, recall by § 2 that for a small positive α we have that

$$a_2(x) = x^\alpha(4/\alpha + \sin\log x)^3 \in \mathcal{S} \setminus \mathcal{T}.$$

The function a_2 belongs to $\mathcal{S}^*$ as it satisfies all the conditions of [9, Theorem 3.5].

To be more precise, let $\phi(n) = |\{m \in \mathbb{N} :\ n = [a_2(m)]\}|$ and $\Phi(n) = \sum_{k=0}^{n} \phi(k)$. Then all the following conditions of [9, Theorem 3.5] hold

(i) $\lim_{n\to\infty} \phi(n) = \infty$ (obvious);

(ii) $\phi(n)$ is almost increasing, i.e., $\phi(n) = q_n + p_n$, where $(q_n)_n$ is increasing and $(p_n)_n$ is bounded (this follows by [9, Lemma 2.4] since a_2^{-1} is increasing and a_2' is decreasing);

(iii) $n\phi(n)/\Phi(n) \le c$ for some $c > 0$ (this follows by the proof of [9, Lemma 2.5] since $\frac{n\phi(n)}{\Phi(n)}$ is arbitrarily close to $\frac{n(a_2^{-1})'(n)}{a_2^{-1}(n)}$, hence it's bounded above by a positive number since, as we already mentioned in § 2, $\frac{na_2'(n)}{a_2(n)}$ is bounded below by $\alpha - \frac{3\alpha}{4-\alpha} > 0$).

Also, it is worth recalling at this point that Theorem 1.5 establishes the same result as Theorem 4.1 for a larger class of functions in the case of uniquely ergodic systems.

(2) Summarizing the info that we have for the classes of interest we have the following

$$\mathcal{H}_\varepsilon \subsetneq \mathcal{T} \subsetneq \mathcal{S}^* \subsetneq \mathcal{S} \subsetneq \mathcal{F} \subsetneq \mathcal{SL},$$

where the strictness of the inclusions are given by: $x^{1/2}(2+\cos\sqrt{\log x})$; $x^\alpha(4/\alpha+\sin\log x)^3$ for $0 < \alpha < 1/20$; $\log x \cdot \log(\log x)$; any Fejér non-$\mathcal{C}^2$ function; and $\log x$ respectively.

We can now show that Theorem 1.4 holds for the class $\mathcal{T}$.

Corollary 4.2. *Let $d \in \mathbb{N}$, (X_i, μ_i, T_i), be measure preserving systems, $a_i \in \mathcal{T}$, $1 \le i \le d$ with $a_d \prec \ldots \prec a_1$ and ν be any coupling of the spaces (X_i, μ_i). Then, for ν-a.e. $\vec{x} \in X_1 \times \cdots \times X_d$, we have that $\lambda_{N,\vec{x}}$ converges to $\mu_{[T_1,\ldots,T_d],\vec{x}}$ as $N \to \infty$.*

This corollary follows (together with the analogous results of Theorems 1.1, 1.2, 1.5 and 1.6 with the condition $a_i \in \mathcal{S}^*$, or $\mathcal{S}$, $a_d \prec \ldots \prec a_1$ and $a_d' \prec \ldots \prec a_1'$ replaced by $a_i \in \mathcal{T}$, or $\mathcal{T} \cup \mathcal{L}_*$ in Theorem 1.2, and $a_d \prec \ldots \prec a_1$) by the fact that $\mathcal{T} \subseteq \mathcal{S}^*$ and that if

$a_1, a_2 \in \mathcal{T} \cup \mathcal{L}_*$ with $a_2 \prec a_1$ (so $a_2 \in \mathcal{T}$) then we have that $a_2' \prec a_1'$ by the identity

$$\frac{a_2'(x)}{a_1'(x)} = \frac{xa_2'(x)}{a_2(x)} \cdot \frac{a_1(x)}{xa_1'(x)} \cdot \frac{a_2(x)}{a_1(x)}.$$

Of course, for functions of different growth $a_1 \prec a_2$ in S^* in general, if some of the limits $xa_i'(x)/a_i(x)$ doesn't exist as $x \to \infty$ (as in the case of the function that we saw in the remarks of this section) we don't expect to get the different growth rate of the derivatives, so we have to assume it.

If for every $\alpha \in (0,1]$ we set $\mathcal{T}(\alpha) := \left\{a \in \mathcal{T} : \lim_{x\to\infty} \frac{xa'(x)}{a(x)} = \alpha\right\}$ then the following remark gives us a relation between the growth rate of $a_i \in \mathcal{T}(\alpha_i)$, $i = 1, 2$, and α_1, α_2.

Remark. If $a_i \in \mathcal{T}(\alpha_i)$, $i = 1, 2$, with $a_2 \prec a_1$, then $\alpha_2 \le \alpha_1$.

Indeed, this follows by [9, Lemmas 2.1 and 2.6]. Let $a = a_2/a_1$ and note that $\lim_{x\to\infty} \frac{xa'(x)}{a(x)} = \alpha_2 - \alpha_1$. If $\alpha_2 > \alpha_1$, then by the argument in [9, Lemma 2.1], we have that $|a(x)| \to \infty$ as $x \to \infty$, a contradiction (note that we only used here the fact that a is bounded).
Conversely, if $\alpha_2 < \alpha_1$ then $a_2 \prec a_1$.

The same argument, for $1/a$, gives us that $a(x) \to 0$ as $x \to \infty$, hence $a_2 \prec a_1$.
Note at this point that it can happen $a_2 \prec a_1$ while both functions belong to the same $\mathcal{T}(\alpha)$, as $a_2(x) = x^{1/3}$ and $a_1(x) = x^{1/3} \log x$, where $a_2 \prec a_1$ with $a_1, a_2 \in \mathcal{T}(1/3)$.

We now deal with the irrational leading coefficient case in the class $\mathcal{L}_*$.

4.1. **Irrational leading coefficient.** In this subsection, we study the limit along linear iterates of the form $\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T^{[\gamma n+\ell]} f(x)$ when γ is irrational. To do this we have to introduce some additional notions and tools.

4.1.1. *A generalized ergodic theorem.* Let (X, μ, T) be a measure preserving system, $\mathcal{I}_{\gamma,m}(T)$ be the sub-σ-algebra generated by the eigenspace of T with eigenvalue $-\frac{m}{\gamma}$, $\mathcal{U}_{\gamma,m} \subseteq L^2(\mu)$ be the closed subalgebra generated by all functions of the form $Tg - \exp\left(2\pi i\frac{m}{\gamma}\right)g$, $\mathcal{I}_\gamma(T) = \bigvee_{m\in\mathbb{Z}} \mathcal{I}_{\gamma,m}(T)$ and $\mathcal{U}_\gamma(T) = \bigcap_{m\in\mathbb{Z}} \mathcal{U}_{\gamma,m}(T)$. We have the following structure theorem (its proof is routine and we omit it):

Theorem 4.3. *Let (X, μ, T) be a measure preserving system and $\gamma \in \mathbb{R} \setminus \mathbb{Q}$. Then*

$$L^2(\mu) = \mathcal{I}_{\gamma,m}(T) \oplus \mathcal{U}_{\gamma,m}(T)$$

for all $m \in \mathbb{Z}$. In particular,

$$L^2(\mu) = \mathcal{I}_\gamma(T) \oplus \mathcal{U}_\gamma(T).$$

We also have the following von Neumann-type mean ergodic theorem:

Proposition 4.4. *Let (X, μ, T) be a measure preserving system, $\gamma \in \mathbb{R}\setminus\mathbb{Q}$, $\ell \in \mathbb{R}$ and $f \in L^2(\mu)$. We have the following (each convergence takes place in $L^2(\mu)$):*

- *If* $\mathbb{E}(f|\mathcal{I}_\gamma(T))=0$, *then*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{[\gamma n+\ell]}f=0;$$

- *If* f *is measurable in* $\mathcal{I}_{\gamma,m}(T)$, *then*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{[\gamma n+\ell]}f=\exp\Big(2\pi i\frac{m\ell}{\gamma}\Big)\cdot\frac{\exp(-2\pi i\frac{m}{\gamma})-1}{-2\pi i\frac{m}{\gamma}}f.$$

Consequently, for all $f\in L^2(\mu)$, *we have*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{[\gamma n+\ell]}f=\sum_{m\in\mathbb{Z}}\exp\Big(2\pi i\frac{m\ell}{\gamma}\Big)\cdot\frac{\exp(-2\pi i\frac{m}{\gamma})-1}{-2\pi i\frac{m}{\gamma}}\mathbb{E}(f|I_{\gamma,m}(T)).$$

Proof. Assume that $\mathbb{E}(f|\mathcal{I}_\gamma(T))=0$. By Theorem 4.3, $f\in\mathcal{U}_\gamma(T)$. Let $\varepsilon>0$. By definition, for all $j\in\mathbb{Z}$, there exist $g_j,\varepsilon_j\in L^2(\mu)$ such that $f=Tg_j-\exp\Big(2\pi i\frac{j}{\gamma}\Big)g_j+\varepsilon_j$ and $\|\varepsilon_j\|_2\leq\varepsilon$. We may assume that $\gamma>0$ since the proof of the other case is identical. We first assume that $\gamma>1$. Note that $m=[n\gamma+\ell]$ for some $n\in\mathbb{Z}$ if and only if $\Big\{\frac{m-\ell}{\gamma}\Big\}\in\Big(1-\frac{1}{\gamma},1\Big)$. Additionally for each $m\in\mathbb{Z}$, there is at most one $n\in\mathbb{Z}$ such that $m=[n\gamma+\ell]$. So we have

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{[\gamma n+\ell]}f=\lim_{N\to\infty}\frac{\gamma}{N}\sum_{n=0}^{N-1}\mathbf{1}_{(1-\frac{1}{\gamma},1)}\Big(\Big\{\frac{n-\ell}{\gamma}\Big\}\Big)T^nf.$$

Let

$$\mathbf{1}_{(1-\frac{1}{\gamma},1)}\Big(\Big\{x-\frac{\ell}{\gamma}\Big\}\Big)=\sum_{j\in\mathbb{Z}}a_j\exp(-2\pi ijx),\quad a_j\in\mathbb{R}$$

be the Fourier expansion of the function $\mathbf{1}_{(1-\frac{1}{\gamma},1)}\Big(\Big\{\cdot-\frac{\ell}{\gamma}\Big\}\Big)$. Then

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{[\gamma n+\ell]}f=\lim_{N\to\infty}\frac{\gamma}{N}\sum_{n=0}^{N-1}\Big(\sum_{j\in\mathbb{Z}}a_j\exp\Big(-2\pi i\frac{jn}{\gamma}\Big)T^nf\Big)$$

$$=\sum_{j\in\mathbb{Z}}a_j\left(\lim_{N\to\infty}\frac{\gamma}{N}\Big(\sum_{n=1}^{N}\exp\Big(-2\pi i\frac{j(n-1)}{\gamma}\Big)T^ng_j-\sum_{n=0}^{N-1}\exp\Big(-2\pi i\frac{j(n-1)}{\gamma}\Big)T^ng_j\Big)\right)$$

$$+\sum_{j\in\mathbb{Z}}a_j\Big(\lim_{N\to\infty}\frac{\gamma}{N}\Big(\sum_{n=1}^{N}\exp\Big(-2\pi i\frac{j(n-1)}{\gamma}\Big)T^n\varepsilon_j\Big)\Big).$$

The first series is equal to 0, while the second one has L^2-norm smaller than ε. Since ε is arbitrary we conclude that $\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{[\gamma n+\ell]}f=0$.

If $0<\gamma<1$, pick $W\in\mathbb{N}$ such that $\gamma W>1$. Then by the previous computation,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{[\gamma n+\ell]}f=\frac{1}{W}\sum_{k=0}^{W-1}\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{[(\gamma W)n+(\gamma k+\ell)]}f=0.$$

This finishes the proof of the first part. Suppose now that f is measurable in $\mathcal{I}_{\gamma,m}(T)$. Then $Tf(x)=\exp\Big(2\pi i\frac{m}{\gamma}\Big)f(x)$ for μ-a.e. $x\in X$. For such $x\in X$, we have

$$\begin{aligned}\frac{1}{N}\sum_{n=0}^{N-1}T^{[\gamma n+\ell]}f(x)&=\frac{1}{N}\sum_{n=0}^{N-1}\exp\Big(2\pi i\frac{m[n\gamma+\ell]}{\gamma}\Big)f(x)\\&=\exp\Big(2\pi i\frac{m\ell}{\gamma}\Big)f(x)\cdot\frac{1}{N}\sum_{n=0}^{N-1}\exp\Big(-2\pi i\frac{m\{n\gamma+\ell\}}{\gamma}\Big).\end{aligned}$$

Since the sequence $(\{n\gamma+\ell\})_n$ is equidistributed in $\mathbb{T}$, we have that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\exp\Big(-2\pi i\frac{m\{n\gamma+\ell\}}{\gamma}\Big)=\int_{\mathbb{T}}\exp\Big(-2\pi i\frac{mx}{\gamma}\Big)\,dx=\frac{\exp(-2\pi i\frac{m}{\gamma})-1}{-2\pi i\frac{m}{\gamma}}.$$

This completes the proof. □

4.1.2. *A special extension.* Let (X,μ,T) be a measure preserving system. We build an extension system of X which will be used to prove Theorem 1.2. Let $([0,1],\mathcal{D},m)$ be the $[0,1]$ interval with the Lebesgue measure m. Let R be the equivalence relation on $[0,1]\times X$ generated by $((1,x),(0,Tx))$, $x\in X$, $Y=([0,1]\times X)/R$, $\mathcal{Y}=(\mathcal{D}\times\mathcal{B})/R$, $\nu=m\times\mu/R$, and $\tilde{T}=id\times T$. Then we have a factor map $\pi\colon(Y,\mathcal{Y},\nu,\tilde{T})\to(X,\mathcal{B},\mu,T)$, where π is the projection to the second coordinate. Let $S\colon Y\to Y$ with $S(t,x)=S(t+\gamma,x)$. Then (using the relations $[\gamma+\{\gamma'\}]+[\gamma']=[\gamma+\gamma']$ and $\{\gamma+\{\gamma'\}\}=\{\gamma+\gamma'\}$) we have

$$S^n\Big(\{\ell\},T^{[\ell]}x\Big)=(n\gamma+\ell,x)=\Big(\{n\gamma+\ell\},T^{[n\gamma+\ell]}x\Big)$$

for all $n\in\mathbb{Z}$ and $x\in X$. For a function f on X we define its extension, $\tilde{f}$ on Y, by $\tilde{f}(t,x)=f(x)$ for all $t\in[0,1),x\in X$. Then

$$T^{[\gamma n+\ell]}f(x)=\tilde{f}\Big(\{n\gamma+\ell\},T^{[n\gamma+\ell]}x\Big)=S^n\tilde{f}\Big(\{\ell\},T^{[\ell]}x\Big).$$

Corollary 4.5. *Let the quantifiers be as above. Then for μ-a.e. $x\in X$ we have*

$$\mathbb{E}\Big(\tilde{f}|\mathcal{I}(S)\Big)\Big(\{\ell\},T^{[\ell]}x\Big)=\sum_{m\in\mathbb{Z}}\exp\Big(2\pi i\frac{m\ell}{\gamma}\Big)\cdot\frac{\exp(-2\pi i\frac{m}{\gamma})-1}{-2\pi i\frac{m}{\gamma}}\mathbb{E}(f|I_{\gamma,m}(T))(x).$$

Proof. By Birkhoff's ergodic theorem, the limit $\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}S^n\tilde{f}(y,z)$ exists and equals to $\mathbb{E}\Big(\tilde{f}|\mathcal{I}(S)\Big)(y,z)$ for ν-a.e. $(y,z)\in Y$. Since $\tilde{f}(y,z)=\tilde{f}(y',z)$ for all $y,y'\in[0,1]$, $z\in X$, and μ is T-invariant, it is easy to conclude that $\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}S^n\tilde{f}\Big(\{\ell\},T^{[\ell]}x\Big)$ exists and equals to $\mathbb{E}\Big(\tilde{f}|\mathcal{I}(S)\Big)\Big(\{\ell\},T^{[\ell]}x\Big)$ for μ-a.e. $x\in X$. This implies that the limit

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{[\gamma n+\ell]}f(x)$$

exists for μ-a.e. $x\in X$, hence its pointwise limit equals to its $L^2(\mu)$ limit. So

$$\mathbb{E}\Big(\tilde{f}|\mathcal{I}(S)\Big)\Big(\{\ell\},T^{[\ell]}x\Big)=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}S^n\tilde{f}\Big(\{\ell\},T^{[\ell]}x\Big)=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{[\gamma n+\ell]}f(x)$$

for μ-a.e. $x\in X$. The result now follows by Proposition 4.4. □

5. Proof of main results

In this last section we give the proof of the statements in § 1. To lighten the notation, we omit writing the spaces along which we integrate, since they are easily deduced by the measures that we use.

Proof of Theorem 1.4. The case $d=1$ is true by the definition of $\mathcal{S}^*$. Now suppose the conclusion holds for $d-1$. Let ν be a coupling of the systems (X_i,μ_i) and let $\pi_1\colon X_1\times\cdots\times X_d\to X_1$ be the projection to the first coordinate and $\pi_2\colon X_1\times\cdots\times X_d\to X_2\times\cdots\times X_d$ be the projection to the rest coordinates. Write $\vec{x}=(x_1,x_*)$, where $x_*=(x_2,\dots,x_d)$. By induction hypothesis, $(\pi_1)_*\lambda_{\vec{x}}=\mu_{[T_1],x_1}$ ν-a.e. and $(\pi_2)_*\lambda_{\vec{x}}=\mu_{[T_2,\dots,T_d],x_*}$ ν-a.e.

For every $1\le i\le d$, fix a countable dense set of continuous functions $\mathcal{C}_i=\{g_{i,k}:k\in\mathbb{N}\}\subseteq C(X_i)$. Let $X_1'\subseteq X_1$ be a full μ_1-measure set such that $\left(\frac{1}{N}\sum_{n=0}^{N-1}T_1^n g_{1,k}(x_1)\right)_N$ converges to $\mathbb{E}(g_{1,k}|\mathcal{I}(T_1))(x_1)=\int g_{1,k}\,d\mu_{[T_1],x_1}$ as $N\to\infty$ for every $k\in\mathbb{N}$ and $x_1\in X_1'$. Since $\mu_1=\int\mu_{[T_1],x_1}\,d\mu_1(x_1)$, the same is true for $\mu_{[T_1],x_1}$-a.e. $y\in X_1$ for μ_1-a.e. $x_1\in X_1$. Let $f_i\in\mathcal{C}_i$, $1\le i\le d$ and $\vec{x}\in\pi_1^{-1}(X_1')$. By Lemma 3.1 and since $(\pi_1)_*\lambda_{\vec{x}}=\mu_{[T_1],x_1}$ we have

$$\begin{aligned}\int f_1\otimes f_2\otimes\cdots\otimes f_d\,d\lambda_{\vec{x}} &= \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\int T_1^n f_1\otimes f_2\otimes\cdots\otimes f_d\,d\lambda_{\vec{x}}\\ &= \int\Big(\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T_1^n f_1\Big)\otimes f_2\otimes\cdots\otimes f_d\,d\lambda_{\vec{x}}\\ &= \int\mathbb{E}(f_1|\mathcal{I}(T_1))(x_1)\otimes f_2\otimes\cdots\otimes f_d\,d\lambda_{\vec{x}}\\ &= \int\Big(\int f_1 d\mu_{[T_1],x_1}\Big)\otimes f_2\otimes\cdots\otimes f_d\,d\lambda_{\vec{x}}.\end{aligned}$$

Remark that the function $\left(\int f_1 \, d\mu_{[T_1],x_1}\right)$ is constant $\mu_{[T_1],x_1}$-a.e. and thus continuous $\mu_{[T_1],x_1}$-a.e. Using again the definition of $\lambda_{\vec{x}}$ and the induction hypothesis we get that

$$\begin{aligned}\int f_1 \otimes f_2 \otimes \cdots \otimes f_d \, d\lambda_{\vec{x}} &= \left(\int f_1 \, d\mu_{[T_1],x}\right)\left(\int f_2 \otimes \cdots \otimes f_d \, d\mu_{[T_2,\ldots,T_d],x_*}\right) \\ &= \mathbb{E}(f_1|\mathcal{I}(T_1))(x_1)\mathbb{E}(f_2|\mathcal{I}(T_2))(x_2) \cdot \ldots \cdot \mathbb{E}(f_d|\mathcal{I}(T_d))(x_d).\end{aligned}$$

By the density of linear combinations of functions $f_1 \otimes \cdots \otimes f_d$, $f_i \in \mathcal{C}_i$ in $C(X_1 \times \cdots \times X_d)$, we get that $\lambda_{\vec{x}} = \mu_{[T_1,\ldots,T_d],\vec{x}}$. □

We are now ready to prove Theorem 1.1 via Theorem 1.4.

Proof of Theorem 1.1. We keep the notation as above, and let $\mathcal{C}_i$ be a countable family of continuous functions which is dense in $L^2(\mu_i)$ for $1 \leq i \leq d$. For $N \in \mathbb{N}$ denote

$$A_N(x_1, \ldots, x_d) := \frac{1}{N} \sum_{n=0}^{N-1} f_1\Big(T_1^{[a_1(n)]} x_1\Big) \cdot \ldots \cdot f_d\Big(T_d^{[a_d(n)]} x_d\Big).$$

For $k \in \mathbb{N}$, pick $\widehat{f}_{i,k} \in \mathcal{C}_i$ such that $\|f_i - \widehat{f}_{i,k}\|_2 \leq \frac{1}{k}$ for $1 \leq i \leq d$ and denote

$$\widehat{A}_{N,k}(x_1, \ldots, x_d) := \frac{1}{N} \sum_{n=0}^{N-1} \widehat{f}_{1,k}\Big(T_1^{[a_1(n)]} x_1\Big) \cdot \ldots \cdot \widehat{f}_{d,k}\Big(T_d^{[a_d(n)]} x_d\Big).$$

By the definition of $\mathcal{S}^*$ and the telescoping inequality, we have that

$$\limsup_{N\to\infty} |A_N(x_1, \ldots, x_d) - \widehat{A}_{N,k}(x_1, \ldots, x_d)| \leq \sum_{i=1}^{d} \mathbb{E}(|f_i - \widehat{f}_{i,k}||\mathcal{I}(T_i))(x_i)$$

for ν-a.e. $\vec{x} = (x_1, \ldots, x_d) \in X_1 \times \cdots \times X_d$ and for all $k \in \mathbb{N}$. So for every $k \in \mathbb{N}$,

$$\limsup_{N\to\infty} \left| A_N(x_1, \ldots, x_d) - \int f_1 \otimes \cdots \otimes f_d \, d\mu_{[T_1,\ldots,T_d],\vec{x}} \right|$$

is bounded by the sum of the terms

$$\limsup_{N\to\infty} \left| A_N(x_1, \ldots, x_d) - \widehat{A}_{N,k}(x_1, \ldots, x_d) \right|, \tag{20}$$

$$\limsup_{N\to\infty} \left| \widehat{A}_{N,k}(x_1, \ldots, x_d) - \int \widehat{f}_{1,k} \otimes \cdots \otimes \widehat{f}_{d,k} \, d\mu_{[T_1,\ldots,T_d],\vec{x}} \right|; \quad \textit{and} \tag{21}$$

$$\left| \int \widehat{f}_{1,k} \otimes \cdots \otimes \widehat{f}_{d,k} \, d\mu_{[T_1,\ldots,T_d],\vec{x}} - \int f_1 \otimes \cdots \otimes f_d \, d\mu_{[T_1,\ldots,T_d],\vec{x}} \right|. \tag{22}$$

By Theorem 1.4, the Term (21) is equal to 0. Again by telescoping, the sum of the Terms (20) and (22) is bounded by $2\sum_{i=1}^{d} \mathbb{E}(|f_i - \widehat{f}_{i,k}||\mathcal{I}(T_i))(x_i)$, for ν-a.e. $(x_1, \ldots, x_d) \in X_1 \times \cdots \times X_d$. If A_m denotes the set of points $\vec{x} = (x_1, \ldots, x_d)$ such that

$$\limsup_{N\to\infty} \left| A_N(x_1, \ldots, x_d) - \int f_1 \otimes \cdots \otimes f_d \, d\mu_{[T_1,\ldots,T_d],\vec{x}} \right| \geq \frac{1}{m},$$

then Markov's inequality implies that for every $m \in \mathbb{N}$ the measure of A_m is smaller than $\frac{2dm}{k}$. Since k is arbitrary, we have that $\nu(A_m) = 0$ and $\nu(\cap_{m\in\mathbb{N}} A_m^c) = 1$. It is immediate to check that $A_N(x_1, \dots, x_d)$ converges for every $(x_1, \dots, x_d) \in \cap_{m\in\mathbb{N}} A_m^c$. □

Proof of Theorem 1.5. Let $x \in X$ and $\lambda_{N,x} = \frac{1}{N}\sum_{n=0}^{N-1} T^{[a(n)]}\delta_x$. By Lemma 3.1, we have that any weak limit of $\lambda_{N,x}$ is T-invariant and hence equal to μ by unique ergodicity. Therefore, $\lambda_{N,x}$ converges to μ as $N \to \infty$ and the conclusion follows. □

Proof of Theorem 1.6. It is the same proof as of Theorem 1.4, combined with the fact that the single average converges by Theorem 1.5. □

Proof of Theorem 1.2. **(i) Case where $a_1(n) = pn$, $p \in \mathbb{Z} \setminus \{0\}$.** The proof is very similar to the one of Theorem 1.4; we sketch it for completeness. We will give an expression for any weak limit of the average of the Dirac measure $\delta_{\vec{x}}$, $\vec{x} = (x_1, \dots, x_d)$ in a dense family of functions, and then it's routine to arrive at the conclusion (as in the proof of Theorem 1.1).

Let π_1 and π_2 be the projections as in the proof of Theorem 1.4. By Birkhoff's ergodic theorem we have that $(\pi_1)_*\lambda_{\vec{x}} = \mu_{[T_1^p],x_1}$ for ν-a.e. $\vec{x} \in X_1 \times \cdots \times X_d$. By Theorem 1.4, $(\pi_2)_*\lambda_{\vec{x}} = \mu_{[T_2,\dots,T_d],x_*}$ for ν-a.e. $\vec{x} \in X_1 \times \cdots \times X_d$.

Fix a countable dense set of continuous functions $\mathcal{C}_i = \{g_{i,k} : k \in \mathbb{N}\} \subseteq C(X_i)$. We have that there exists a set of full measure $X_1' \subseteq X_1$ such that $\left(\frac{1}{N}\sum_{n=0}^{N-1} T_1^{pn} g_{1,k}(x_1)\right)_N$ converges to $\mathbb{E}(g_{1,k}|\mathcal{I}(T_1^p))(x_1) = \int g_{1,k}\, d\mu_{[T_1^p],x_1}$ as $N \to \infty$ for every $k \in \mathbb{N}$ and $x_1 \in X_1'$. Since $\mu_1 = \int \mu_{[T_1^p],x_1}\, d\mu_1(x_1)$, the same is true for $\mu_{[T_1^p],x_1}$-a.e. $y \in X_1$ for μ_1-a.e. $x_1 \in X_1$. Let $f_i \in \mathcal{C}_i$, $1 \le i \le d$ and $\vec{x} \in \pi_1^{-1}(X_1')$. Applying Lemma 3.2 we have (as in Theorem 1.4)

$$\begin{aligned}\int f_1 \otimes f_2 \otimes \cdots \otimes f_d\, d\lambda_{\vec{x}} &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \int T_1^{pn} f_1 \otimes f_2 \otimes \cdots \otimes f_d\, d\lambda_{\vec{x}} \\ &= \int \left(\int f_1\, d\mu_{[T_1^p],x_1}\right) \otimes f_2 \otimes \cdots \otimes f_d\, d\lambda_{\vec{x}}.\end{aligned}$$

Remark that the function $\left(\int f_1\, d\mu_{[T_1^p],x_1}\right)$ is constant $\mu_{[T_1^p],x_1}$-a.e. and thus continuous $\mu_{[T_1^p],x_1}$-a.e. Using again the definition of $\lambda_{\vec{x}}$ and Theorem 1.4, we get that

$$\begin{aligned}\int f_1 \otimes f_2 \otimes \cdots \otimes f_d\, d\lambda_{\vec{x}} &= \left(\int f_1\, d\mu_{[T_1^p],x_1}\right)\left(\int f_2 \otimes \cdots \otimes f_d\, d\mu_{[T_2,\dots,T_d],x_*}\right) \\ &= \mathbb{E}(f_1|\mathcal{I}(T_1^p))(x_1)\mathbb{E}(f_2|\mathcal{I}(T_2))(x_2) \cdot \ldots \cdot \mathbb{E}(f_d|\mathcal{I}(T_d))(x_d).\end{aligned}$$

By the density of linear combinations of functions $f_1 \otimes \cdots \otimes f_d$, $f_i \in \mathcal{C}_i$ in $C(X_1 \times \cdots \times X_d)$, we get that $\lambda_{\vec{x}} = \mu_{[T_1^p,T_2,\dots,T_d],\vec{x}}$.

(ii) Case where $a_1(n) = pn + \ell$, $p \in \mathbb{Z} \setminus \{0\}$. Using Case (i), we have

$$\begin{aligned}
&\int f_1 \otimes f_2 \otimes \cdots \otimes f_d \, d\lambda_x \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T_1^{[pn+\ell]} f_1(x_1) T_2^{[a_2(n)]} f_2(x_2) \cdot \ldots \cdot T_d^{[a_d(n)]} f_d(x_d) \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T_1^{pn} \Big(T_1^{[\ell]} f_1 \Big)(x_1) T_2^{[a_2(n)]} f_2(x_2) \cdot \ldots \cdot T_d^{[a_d(n)]} f_d(x_d) \\
&= \mathbb{E}\Big(T_1^{[\ell]} f_1 | \mathcal{I}(T_1^p) \Big)(x_1) \mathbb{E}(f_2|\mathcal{I}(T_2))(x_2) \cdot \ldots \cdot \mathbb{E}(f_d|\mathcal{I}(T_d))(x_d).
\end{aligned}$$

(iii) Case where $a_1(n) = kn + \ell$, $k = p/q \in \mathbb{Q} \setminus \{0\}$. Using Case (ii), we have

$$\begin{aligned}
&\int f_1 \otimes f_2 \otimes \cdots \otimes f_d \, d\lambda_{\vec{x}} \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T_1^{[kn+\ell]} f_1(x_1) T_2^{[a_2(n)]} f_2(x_2) \cdot \ldots \cdot T_d^{[a_d(n)]} f_d(x_d) \\
&= \frac{1}{q} \sum_{j=0}^{q-1} \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T_1^{[k(qn+j)+\ell]} f_1(x_1) T_2^{[a_2(qn+j)]} f_2(x_2) \cdot \ldots \cdot T_d^{[a_d(qn+j)]} f_d(x_d) \\
&= \frac{1}{q} \sum_{j=0}^{q-1} \mathbb{E}\Big(T_1^{[\frac{pj}{q}+\ell]} f_1 | \mathcal{I}(T_1^p) \Big)(x_1) \mathbb{E}(f_2|\mathcal{I}(T_2))(x_2) \cdot \ldots \cdot \mathbb{E}(f_d|\mathcal{I}(T_d))(x_d).
\end{aligned}$$

(iv) Case where $a_1(n) = \gamma n + \ell$, $\gamma \in \mathbb{R} \setminus \mathbb{Q}$ (recall the notations of Corollary 4.5). By Case (i), and by passing to the mapping torus extension, we have

$$\begin{aligned}
&\int f_1 \otimes f_2 \otimes \cdots \otimes f_d \, d\lambda_x \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T_1^{[\gamma n+\ell]} f_1(x_1) T_2^{[a_2(n)]} f_2(x_2) \cdot \ldots \cdot T_d^{[a_d(n)]} f_d(x_d) \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} S^n \tilde{f}_1 \Big(\{\ell\}, T^{[\ell]} x_1 \Big) T_2^{[a_2(n)]} f_2(x_2) \cdot \ldots \cdot T_d^{[a_d(n)]} f_d(x_d) \\
&= \mathbb{E}\Big(\tilde{f}_1 | \mathcal{I}(S) \Big) \Big(\{\ell\}, T^{[\ell]} x_1 \Big) \mathbb{E}(f_2|\mathcal{I}(T_2))(x_2) \cdot \ldots \cdot \mathbb{E}(f_d|\mathcal{I}(T_d))(x_d).
\end{aligned}$$

By Corollary 4.5, the result follows. □

Proof of Corollary 1.3. It follows by Theorems 1.1 and 1.2. □

Appendix A. Classes of functions used in the paper

In this appendix, we summarize all the classes of functions pertaining to this paper:
$\mathcal{F}$: the set of all Fejér functions.
$\mathcal{L}_*$: the set of linear functions with non-zero leading coefficient.
$\mathcal{H}$: the union of all Hardy fields.

$\mathcal{H}_\varepsilon$: the set of functions a which belongs to some Hardy field and satisfy $x^\varepsilon \prec a(x) \prec x$.

$\mathcal{LE}$: the set of logarithmico-exponential Hardy field functions.

$\mathcal{LE}_\varepsilon$: the set of logarithmico-exponential Hardy field functions a that satisfy the growth condition $x^\varepsilon \prec a(x) \prec x$ for some $\varepsilon > 0$.

$\mathcal{SL}$: the set of sublinear functions (i.e., functions a that satisfy $a(x) \prec x$).

$$\mathcal{R} = \Big\{ a \in \mathcal{C}^3(\mathbb{R}^+) : \text{ the limits } \lim_{x\to\infty} \frac{xa'(x)}{a(x)}, \lim_{x\to\infty} \frac{xa''(x)}{a'(x)}, \text{ and } \lim_{x\to\infty} \frac{xa'''(x)}{a''(x)} \text{ exist in } \mathbb{R} \Big\}.$$

$$\mathcal{T} = \Big\{ a \in \mathcal{R} : \lim_{x\to\infty} \frac{xa'(x)}{a(x)} \in (0,1) \text{ or } \lim_{x\to\infty} \frac{xa'(x)}{a(x)} = 1 \text{ and } \lim_{x\to\infty} a'(x) = 0 \text{ monotonicaly} \Big\}.$$

$$D_k = \left\{ a : \limsup_{x\to \operatorname{sgn}(a^{-1})\cdot\infty} \sup_{h\in[-1,1]} \left| \frac{a^{(k+1)}(x+h)}{a^{(k)}(x)} \right| < \infty \right\}.$$

$$M_k = \Big\{ a : a^{(k)} \text{ is eventually monotone} \Big\}.$$

$$S = \left\{ a \in \mathcal{C}^3(\mathbb{R}^+) \;\middle|\; a, \frac{1}{a'} \in \mathcal{SL} \text{ and } a^{-1} \in M_1 \cap D_0 \cap D_1 \cap (D_2 \cup M_2) \right\}.$$

$\mathcal{S}^*$: the set of functions $a \in \mathcal{S}$ where $\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^{[a(n)]}x)$ exists pointwisely (a.e.) for every measure preserving system (X, μ, T) and every bounded measurable function f.

The relation between the classes of main interest is: $\mathcal{H}_\varepsilon \subsetneq \mathcal{T} \subsetneq \mathcal{S}^* \subsetneq \mathcal{S} \subsetneq \mathcal{F} \subsetneq \mathcal{SL}$.

References

[1] I. Assani. Multiple recurrence and almost sure convergence of weakly mixing dynamical systems. *Israel J. Math.* **103** (1998), 111–125.

[2] I. Assani. Pointwise convergence of nonconventional everages. *Coll. Math.*, vol 102, **2** (2005), 245–262.

[3] I. Assani. Averages along cubes for not necessarily commuting measure preserving transformations. Contemp. Math., vol. **430** (2007), 1–19.

[4] I. Assani. Pointwise convergence of ergodic averages along cubes. *J. Analyse Math.* **110** (2010), 241–269.

[5] T. Austin. On the norm convergence of nonconventional ergodic averages. *Ergodic Theory Dynam. Systems* **30** (2010), no. 2, 321–338.

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

[7] V. Bergelson. Weakly mixing PET. *Ergodic Theory Dynam. Systems* **7** (1987), no. 3, 337–349.

[8] V. Bergelson. The multifarious Poincaré Recurrence Theorem. Descriptive Set Theory and Dynamical Systems, Eds. M. Foreman, A.S. Kechris, A. Louveau, B. Weiss. Cambridge University Press, New York (2000), 31–57.

[9] V. Bergelson and I. Håland-Knutson. Weakly mixing implies mixing of higher orders along tempered functions. *Ergodic Theory Dynam. Systems* **29** (2009), no. 5, 1375–1416.

[10] V. Bergelson and A. Leibman. A nilpotent Roth theorem. *Invent. Math.* **147** (2002), 429–470.

[11] M. Boshernitzan, G. Kolesnik, A. Quass and M. Weirdl. Ergodic averaging sequences. *J. Anal. Math.* **95** (2005), 63–103.

[12] J. Bourgain. Double recurrence and almost sure convergence. *J. Reine Angew. Math.* **404** (1990), 140–161.

[13] J. Bourgain. Pointwise ergodic theorems for arithmetic sets. *Inst. Hautes Études Sci. Publ. Math.* **69** (1989), 5–45.
[14] Q. Chu. Convergence of multiple ergodic averages along cubes for several commuting transformations. *Studia Math.* **196**(1) (2009), 13–22.
[15] Q. Chu, N. Frantzikinakis and B. Host. Ergodic averages of commuting transformations with distinct degree polynomial iterates. *Proc. Lond. Math. Soc. (3)*, **102** (2011), 801–842.
[16] Q. Chu and N. Frantzikinakis. Pointwise convergence for cubic and polynomial ergodic averages of non-commuting transformations. *Ergodic Theory Dynam. Systems* **32**, no. 3, (2012), 877–897.
[17] S. Donoso and W. Sun. A pointwise cubic average for two commuting transformations. *Israel J. Math.* **216** (2016), no. 2, 657–678.
[18] S. Donoso and W. Sun. Pointwise multiple averages for systems with two commuting transformations. *Ergodic Theory Dynam. Systems*, 1–26. doi:10.1017/etds.2016.127.
[19] S. Donoso and W. Sun. Pointwise convergence of some multiple ergodic averages, *Adv. Math.* **330** (2018), 946–996.
[20] N. Frantzikinakis. Multiple recurrence and convergence for Hardy field sequences of polynomial growth. *J. Anal. Math.* **112** (2010), 79–135.
[21] N. Frantzikinakis. A multidimensional Szemeredi theorem for Hardy sequences of polynomial growth. *Trans. Amer. Math. Soc.* **367**, no. 8, (2015), 5653–5692.
[22] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Anal. Math.* **31** (1977), 204–256.
[23] W. T. Gowers. A new proof of Szemerédi's Theorem. *Geom. Funct. Anal.* **11** (2001), 465–588.
[24] Y. Gutman, W. Huang, S. Shao and X. Ye. Almost sure convergence of the multiple ergodic average for certain weakly mixing systems. *Acta Math. Sin. (Engl. Ser.)* **34** (2018), no. 1, 79–90.
[25] B. Host and B. Kra. Nonconventional averages and nilmanifolds. *Ann. of Math. (2)* **161** (2005), no. 1, 398–488.
[26] B. Host, Ergodic seminorms for commuting transformations and applications. *Studia Math.* **195** (2009), no. 1, 31–49.
[27] A. Koutsogiannis. Integer part polynomial correlation sequences. *Ergodic Theory Dynam. Systems* **38** (2018), no. 4, 1525–1542.
[28] A. Koutsogiannis. Closest integer polynomial multiple recurrence along shifter primes. *Ergodic Theory Dynam. Systems* **38** (2018), no. 2, 666–685.
[29] W. Huang, S. Shao and X. Ye, Strictly ergodic models under face and parallelepiped group actions,em Commun. Math. Stat. **5** (2017), no. 1, 93–122.
[30] W. Huang, S. Shao, X. Ye. Pointwise convergence of multiple ergodic averages and strictly ergodic models. To appear in *J. Anal. Math.* arXiv:1406.5930.
[31] T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. *Ergodic Theory Dynam. Systems* **28** (2008), no. 2, 657–688.
[32] M. Walsh. Norm convergence of nilpotent ergodic averages. *Ann. of Math.* **175** (2012), no. 3, 1667–1688.

(Sebastián Donoso) Universidad de O'Higgins, Instituto de Ciencias de la Ingeniería, Rancagua, Chile
Email address: sebastian.donoso@uoh.cl

(Andreas Koutsogiannis) The Ohio State University, Department of mathematics, Columbus, Ohio, USA
Email address: koutsogiannis.1@osu.edu

(Wenbo Sun) The Ohio State University, Department of mathematics, Columbus, Ohio, USA
Email address: sun.1991@osu.edu