Can. J. Math. Vol. **47** (4), 1995 pp. 852–876

# ERGODIC AVERAGES FOR WEIGHT FUNCTIONS MOVED BY NON-LINEAR TRANSFORMATIONS ON $\mathbb{R}^n$

DAVID I. McINTOSH

ABSTRACT. Let $\mathbb{R}^+$ denote the non-negative half of the real line, and let $\lambda$ denote Lebesgue measure on the Borel sets of $\mathbb{R}^n$. A function $\varphi\colon \mathbb{R}^n \longrightarrow \mathbb{R}^+$ is called a *weight function* if $\int_{\mathbb{R}^n} \varphi\, d\lambda = 1$. Let $(X, \mathcal{F}, \mu)$ be a non-atomic, finite measure space, let $f\colon X \longrightarrow \mathbb{R}^+$, and suppose $\{T_v\}_{v\in\mathbb{R}^n}$ is an ergodic, aperiodic $\mathbb{R}^n$-flow on $X$. We consider the weighted ergodic averages

$$\mathcal{A}^{\varphi_k} f(x) = \int_{\mathbb{R}^n} f(T_v x)\varphi_k(v)\lambda(dv)$$

where $\{\varphi_k\}_{k=1}^{\infty}$ is a sequence of weight functions. Sufficient as well as necessary and sufficient conditions for the pointwise, almost-everywhere convergence of $\mathcal{A}^{\varphi_k} f(x)$ are developed for a particular class of weight functions $\varphi_k$. Specifically, let $\{\tau_k\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n\}$ be a sequence of measurable, non-singular maps with measurable, non-singular inverses such that the Radon-Nikodym derivatives $d\lambda\circ\tau_k/d\lambda$ and $d\lambda\circ\tau_k^{-1}/d\lambda$ are $L_\infty(\mathbb{R}^n)$, and such that $\tau_k$ and $\tau_k^{-1}$ map bounded sets to bounded sets. We examine convergence for the sequence

$$\varphi_k = \frac{d\lambda\circ\tau_k}{d\lambda}\cdot\theta_k\circ\tau_k$$

where $\theta_k$ is an a.e.-convergent sequence of weight functions which are dominated by a fixed $L_1(\mathbb{R}^n)$ function with bounded support.

## 1. **Introduction.**

*Background.* In recent years a great deal of work has been done with Hardy-Littlewood types of maximal inequalities and related convergence results. In 1984, Nagel and Stein [17] examined the maximal operator $\mathcal{M}_Q$ defined by

$$\mathcal{M}_Q f(x_0) = \sup_{(x,y)\in Q} \frac{1}{\lambda\big(B_0(y)\big)} \int_{B_0(y)} |f(x_0 + x + v)|\lambda(dv)$$

where $Q \subset \mathbb{R}^n \times \mathbb{R}^+$ is open. They obtain a characterization of the regions $Q$ for which $\mathcal{M}_Q$ is weak type $(1,1)$ and strong type $(p,p)$ for $p > 1$. In 1987, Sueiro [21] provided a short, elegant proof of Nagel's and Stein's result.

In 1990, using techniques similar to those of Nagel, Stein and Sueiro, and using Calderón's transference principle, Bellow, Jones and Rosenblatt [4] derive similar maximal estimates and convergence results for the sequence

$$\mathcal{A}_k = \frac{1}{l_k}\sum_{i=1}^{l_k} f(T^{a_k+i}x)$$

Research done at the University of Toronto and supported in part by an NSERC scholarship and by the University of Toronto.

Received by the editors June 24, 1993; revised March 2, 1994.

AMS subject classification: Primary: 28D99; secondary: 60F99.

Key words and phrases: maximal ergodic theorems.

of ergodic averages where $\{(a_k, l_k)\} \subset \mathbb{Z} \times \mathbb{Z}^+$. Jones and Olsen [12] generalize these results from $\mathbb{Z}$ actions of $T^k$ to $\mathbb{Z}^n$ actions $\{T_v\}_{v \in \mathbb{Z}^n}$.

Turning to the continuous ergodic case, Broise, Déniel and Derriennic [8] have recently examined the averages

$$\mathcal{A}^{\varphi} f(x) = \int_{\mathbb{R}^n} f(T_v x) \varphi(v) \lambda(dv) \tag{1}$$

where now $\{T_v\}$ is an $\mathbb{R}^n$-flow on $X$ and $\varphi$ is a weight function. Specifically, they examine maximal estimates and convergence for the sequence of weight functions $\{\varphi_k\}$, obtained from a fixed weight function $\varphi$ by a sequence $\{\tau_k\}$ of linear transformations on $\mathbb{R}^n$, via the formula

$$\varphi_k = \frac{1}{\det \tau_k} \varphi \circ \tau_k.$$

Akçoğlu and Déniel examined similar problems for affine transformations $\tau_k$ on $\mathbb{R}$. Under various conditions on $f$ and $\varphi$, they obtain maximal estimates, and necessary-and-sufficient conditions for the a.e. convergence, of the averages $\mathcal{A}^{\varphi_k} f(x)$ for sequences of affine transformations $\tau_k(t) = r_k^{-1}(t - a_k)$.

Many others have examined similar and related ergodic averages and harmonic averages, of which we can only mention a few. See, for example, [13, 14, 19].

In the present treatment, we examine further problems, similar to those discussed above. Specifically, we examine the case of $\mathbb{R}^n$-flows $\{T_v\}_{v \in \mathbb{R}^n}$ on a finite-measure space $X$, and we examine a very large class of transformations $\tau_k$ on $\mathbb{R}^n$ which are used to generate a sequence of weight functions $\varphi_k$ via the formula

$$\varphi_k(v) = \frac{d\lambda \circ \tau_k}{d\lambda}(v) \cdot \varphi \circ \tau_k(v).$$

We obtain a maximal theorem for the maximal operator corresponding to the ergodic averages (1). In very rough terms, Theorem 5.1 states that, under suitable regularity conditions, the following inequality holds:

$$\mu\Big(\{x \mid \exists k, \mathcal{A}^{\varphi_k} f(x) > \alpha\}\Big) \leq \frac{C_\varphi}{\alpha} \int_{\mathbb{R}} f^*(t) \varphi^*(t)\, d\ell \quad \forall \alpha > 0$$

where $f^*$ and $\varphi^*$ are the decreasing re-arrangements. From this we obtain pointwise convergence results for the sequence $\{\mathcal{A}^{\varphi_k} f\}$.

The organization of the paper is broadly as follows. In the remainder of Section 1, we present the required notation and definitions, followed by a precise statement of the main results. The reader may wish to go immediately to the main results, for motivation, prior to examining all the definitions. Section 2 contains a discussion of the meaning of some of the regularity conditions, and a few related lemmas to be used in the sequel. In particular Lemma 2.1 is non-trivial. In Section 3 we obtain divergence of the ergodic averages using techniques similar to those used in [3] and elsewhere. For the reader's convenience Section 4 contains various technical lemmas to be used later. Their proofs can be found in [16]. Section 5 contains the statement and proof of the maximal theorem.

This is the key theorem of the paper. Section 6 then uses the maximal theorem to prove the convergence results.

*Notations and definitions.* The cardinality of a set $A$ will be denoted by $\#(A)$. The characteristic function of a set $E$ in any space will be denoted by $\mathbf{1}_E$. The support of a non-negative function $f$ on a set $X$ will be denoted by $S_f = \{x \in X \mid f(x) > 0\}$. If $A$ and $B$ are two sets in a topological space, $\bar{A}$ denotes the closure of $A$, $B^c$ denotes the complement of $B$, $A \setminus B$ denotes $B^c \cap A$, and $A \triangle B$ denotes the symmetric difference $A \setminus B \bigcup B \setminus A$.

Let $\mathbb{N}$ be the set of natural numbers, $\mathbb{Z}$ the integers, and $\mathbb{R}$ the real numbers. $k$ will *always* denote a positive integer, so when we speak of a sequence $\{\bullet_k\}_{k\in\mathbb{N}}$, we will usually omit the specification $k \in \mathbb{N}$. Denote the non-negative half of the reals by $\mathbb{R}^+ = \{t \in \mathbb{R} \mid t \geq 0\}$. In general, we will use a superscript $^+$ to indicate objects which take on only non-negative values, the meaning should be clear from the context. If $t \in R$, then $\lfloor t \rfloor$ will denote the integer part of $t$ and $[[t]]$ will denote the fractional part of $t$; thus $t = \lfloor t \rfloor + [[t]]$. We will be dealing extensively with $\mathbb{R}^n$ for fixed $n \in \mathbb{N}$. For subsets of $\mathbb{R}$ and $\mathbb{R}^n$, "measurable" will always mean Borel measurable. $\ell$ will denote Lebesgue measure on $\mathbb{R}$, and $\lambda$ will denote Lebesgue measure on $\mathbb{R}^n$.

The results in this paper hold for any norm $|\cdot|$ on $\mathbb{R}^n$. For any such norm, the open ball of radius $r > 0$ and centre $v \in \mathbb{R}^n$ will be denoted by $B_v(r) = \{u \in \mathbb{R}^n \mid |v-u| < r\}$. When dealing with a collection of balls such as $\{B_{v_k}(r_k)\}$, we will often abbreviate $B_{v_k}(r_k)$ by $B_k$. If $a > 0$ and $B$ is any ball, then $aB$ denotes the ball with *the same* centre, but radius multiplied by the factor $a$; thus $aB_v(r) = B_v(ar)$. (This is not standard notation, but it will prove very useful.) For $v \in \mathbb{R}^n$, $E \subset \mathbb{R}^n$, we retain the standard definition of $v + E = \{v + u \mid u \in E\}$. We shall reserve the notation $d(\cdot,\cdot)$ to denote the *Euclidean* (*i.e.* $\ell_2$) distance between two points (or a point and a set, or two sets). The diameter of a set $E \in \mathbb{R}^n$ will be denoted by $\Omega(E) = \sup\{d(u,v) \mid u, v \in E\}$. Open intervals, half-open intervals and closed intervals in $\mathbb{R}$ are denoted by $(a,b)$, $(a,b]$ or $[a,b)$, and $[a,b]$ respectively.

Throughout the paper, we let $(X, \mathcal{F}, \mu)$ be a non-atomic, finite measure space. By this we mean that we have a $\sigma$-algebra $\mathcal{F}$ of subsets of $X$, and a measure $\mu$ on $\mathcal{F}$ such that $\mu$ is non-atomic and $0 < \mu(X) < \infty$. Also, to avoid technicalities, we will assume $X$ is a Lebesgue space.

We let $\{T_v\}_{v\in\mathbb{R}^n}$ be a measurable $\mathbb{R}^n$-flow on $X$. By this we mean that:

i) For each $v \in \mathbb{R}^n$, $T_v: X \longrightarrow X$ is a measurable, measure-preserving transformation on $X$.

ii) $T_u \circ T_v = T_{u+v}$ for all $u, v \in \mathbb{R}^n$.

iii) $T_0(X)$ is a measurable set.

iv) If $E \subset X$ is measurable, then $\{(v,x) \mid T_v(x) \in E\}$ is a measurable subset of $\mathbb{R}^n \times X$ with the product $\sigma$-algebra and measure $\lambda \times \mu$.

From these properties, $\{T_v\}_{v\in\mathbb{R}^n}$ is a commutative group under composition, and there is a set $\tilde{X} \subset X$ of full measure on which all $T_v$ are invertible. $\{T_v\}$ is called *ergodic* if whenever $B$ satisfies $\mu(B \triangle T_v B) = 0$ for all $v \in \mathbb{R}^n$, then either $\mu(B) = 0$ or $\mu(B^c) = 0$.

$\{T_v\}$ is called *aperiodic* if there is a set $\tilde{X} \subset X$ of full measure so that if $x \in \tilde{X}$ and $v \in \mathbb{R}^n$, $v \neq 0$, then $T_v(x) \neq x$.

A function $\varphi: \mathbb{R}^n \longrightarrow \mathbb{R}^+$ is called a *weight function* if $\int_{\mathbb{R}^n} \varphi\, d\lambda = 1$; it is called a *compact weight function* if the closure of its support is compact. Given $f: X \longrightarrow \mathbb{R}^+$, we wish to consider the pointwise convergence of the ergodic averages defined by

$$\mathcal{A}^{\varphi_k} f(x) = \int_{\mathbb{R}^n} f(T_v x)\varphi_k(v)\lambda(dv) \tag{2}$$

where $\{\varphi_k\}_{k\in\mathbb{N}}$ is a sequence of weight functions defined as follows. Let $\{\tau_k\}_{k\in\mathbb{N}}$ be a sequence of measurable, non-singular transformations $\tau_k: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ with measurable, non-singular inverses. (By non-singular $\tau$, we mean that $\tau^{-1}$ takes sets of $\lambda$-measure zero to sets of $\lambda$-measure zero.) We consider the sequence

$$\varphi_k(v) = \frac{d\lambda \circ \tau_k}{d\lambda}(v) \cdot \varphi \circ \tau_k(v)$$

where $\varphi$ is a fixed weight function and $d\lambda \circ \tau_k / d\lambda$ is the Radon-Nikodym derivative. Essentially, $\tau_k$ is the device used to "spread out" or "compress" the weight function $\varphi$, and $d\lambda \circ \tau_k / d\lambda$ is the natural correction factor used to keep the total weight at 1. In Theorem 1.4 below, we also consider more general sequences.

When investigating the convergence of the averages (2), we make use of various regularity conditions, specified below, and prove a Hardy-Littlewood-type maximal theorem. We need some definitions in order to formulate these regularity conditions. Let $\{E_k\}$ be a sequence of bounded, measurable, non-null sets in $\mathbb{R}^n$. (A measurable set is called *non-null* if its Lebesgue measure is non-zero.)

SUPER-REGULARITY. For any bounded, measurable set $E \subset \mathbb{R}^n$, define $\mathcal{R}(E) = \inf\{r \mid \exists v \in \mathbb{R}^n, E \subset B_v(r)\}$. Thus $\mathcal{R}(E)$ is the radius of the smallest closed ball which is a super-set of (*i.e.* contains) $E$. The sequence $\{E_k\}$ will be called *super-regular* if $\sup_k \mathcal{R}^n(E_k)/\lambda(E_k) < \infty$.

$\beta$-SEQUENCES. The sequence $\{E_k\}$ is called a *$\beta$-sequence* if there is a constant $\beta$ such that, for any open ball $B \subset \mathbb{R}^n$,

$$\lambda(\{v \in \mathbb{R}^n \mid \exists k, v + E_k \subset B\}) \leq \beta\lambda(B).$$

ASYMPTOTIC CONVEXITY. We will say that $\{E_k\}$ is *asymptotically convex* if there is a sequence $\{K_k\}$ of compact, convex sets such that $\lim_{k\to\infty} \lambda(E_k \triangle K_k)/\lambda(E_k) = 0$.

*Regularity.* We define here a number of regularity conditions on $\{\tau_k\}$ which we will be considering. They will be referred to by number throughout the paper.

($R_1$) $\tau_k$ are measurable, non-singular mappings with measurable, non-singular inverses.

($R_2$) $\tau_k$ and $\tau_k^{-1}$ take bounded sets to bounded sets.

https://doi.org/10.4153/CJM-1995-044-0 Published online by Cambridge University Press

(R$_3$) *($L_\infty$ boundedness)* The Radon-Nikodym derivatives $d\lambda \circ \tau_k / d\lambda$ and $d\lambda \circ \tau_k^{-1} / d\lambda$ are $L_\infty(\mathbb{R}^n)$. Henceforth, we will denote these two functions by $\delta_k$ and $\delta_k^{\sim}$ respectively. We require that

$$C_\delta \stackrel{\text{def}}{=} \sup_k (\|\delta_k\|_\infty \cdot \|\delta_k^{\sim}\|_\infty) < \infty.$$

(R$_4$) *(Super-regularity)* For any ball $B$, $\{\tau_k^{-1}B\}$ is super-regular.

(R$_5$) *($\beta$-regularity)* For some ball $B$, $\{\tau_k^{-1}B\}$ is a $\beta$-sequence.

(R$_5$)$^c$ There is a sequence of balls $\{B_i\}_{i\in\mathbb{N}}$ such that $B_i \subset B_{i+1}, \bigcup_{i=1}^\infty B_i = \mathbb{R}^n$, and such that, for all $i$, $\{\tau_k^{-1}B_i\}_{k\in\mathbb{N}}$ is not a $\beta$-sequence.

(R$_6$) *(Localizing regularity)* $\|\delta_k^{\sim}\|_\infty$ and $|\tau_k^{-1}(0)|$ both converge to zero.

(R$_7$) *(Globalizing regularity)* $\|\delta_k^{\sim}\|_\infty$ converges to infinity. For any ball $B$, $\{\tau_k^{-1}B\}$ is asymptotically convex, and

$$\lim_{k\to\infty} \int_{\tau_k^{-1}B} \left| \delta_k(v) - \frac{\lambda(B)}{\lambda(\tau_k^{-1}B)} \right| \lambda(dv) = 0.$$

The last property will be called *asymptotic flatness* of $\{\delta_k\}$.

These conditions on $\{\tau_k\}$ will not all be imposed simultaneously—in fact the last two are mutually exclusive. The first two, however, will always be assumed. We prove a Hardy-Littlewood type maximal theorem using regularity conditions (R$_1$)–(R$_5$). Then, to obtain pointwise convergence of the averages $\mathcal{A}^{\varphi_k} f(x)$, we impose the additional requirement that the sequence $\{\tau_k\}$ satisfy either (R$_6$) or (R$_7$). Conversely, we obtain divergence when the $\{\tau_k\}$ fail to satisfy a particular one of these regularity conditions, the $\beta$-regularity property (R$_5$).

*Principal results.*

THEOREM 1.1. *Assume $\{T_v\}$ is an aperiodic, ergodic $\mathbb{R}^n$ flow. Suppose $\{\tau_k\}$ satisfy (R$_1$), (R$_2$), and (R$_5$)$^c$. Then, for every weight function $\varphi$ on $\mathbb{R}^n$, there is a bounded function $f$ on $X$ such that $\mathcal{A}^{\varphi_k} f(x)$ diverges on a set of positive measure.*

THEOREM 1.2. *Assume that $\{\tau_k\}$ satisfy (R$_1$)–(R$_5$) and either (R$_6$) or (R$_7$). Then $\mathcal{A}^{\varphi_k} f(x)$ converges a.e. for any bounded function $f$ and any weight function $\varphi$.*

Before presenting the next two theorems, we need one more definition. Let $(Y, \mathcal{E}, \nu)$ be a measure space, not necessarily of finite measure, and let $f\colon Y \longrightarrow \mathbb{R}^+$. We define the *distribution* of $f$ to be the measure $D_f$ on the Borel sets of $\mathbb{R}$ given by $D_f(E) = \nu \circ f^{-1}\big(E \cap (0,\infty)\big)$. If this measure is finite for all sets of the form $E = (a, \infty)$, where $a > 0$, then there is an a.e.-unique function $\xi\colon \mathbb{R} \longrightarrow \mathbb{R}^+$ which satisfies

i) $\xi(t) = 0$ for all $t \leq 0$

ii) $\xi$ is non-increasing on $(0,\infty)$, and $\lim_{t\to\infty} \xi(t) = 0$

iii) $D_f = D_\xi$.

Any function $\xi$ satisfying (i) and (ii) will be called a *re-arrangement*, and if $\xi$ also satisfies (iii) for some $f$, then we call $\xi$ the re-arrangement of $f$ and denote it by $f^* = \xi$.

THEOREM 1.3. *Assume that $\{\tau_k\}$ satisfy ($R_1$)–($R_5$) and either ($R_6$) or ($R_7$). Let $\varphi$ be a compact weight function. Then $\mathcal{A}^{\varphi_k}f(x)$ converges a.e. for all functions $f$ on $X$ such that*

$$\int_{\mathbb{R}} |f|^* \varphi^* \, d\ell < \infty.$$

We also examine convergence for a slightly more general sequence of weight functions. Let $\{\theta_k\}$ be a sequence of weight functions which are all bounded by a fixed function $\Theta \in L_1(\mathbb{R}^n)$ of bounded support, and which converge a.e. to a function $\varphi$. $\varphi$ is thus necessarily a compact weight function. Define

$$\psi_k(v) = \delta_k(v) \cdot \theta_k \circ \tau_k(v).$$

THEOREM 1.4. *Assume that $\{\tau_k\}$ satisfy ($R_1$)–($R_5$) and either ($R_6$) or ($R_7$). Let $\varphi$, $\{\theta_k\}$, and $\{\psi_k\}$ be as above. Then $\mathcal{A}^{\psi_k}f(x)$ converges a.e. for all functions $f$ on $X$ such that*

$$\int_{\mathbb{R}} |f|^* \Theta^* \, d\ell < \infty.$$

*Applications and examples.* We may obtain many of the results found in [3], [7], [12], and [13] as a subcase of those found here. In [7], the authors deal with the class of linear transformations $\tau\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ satisfying $\|\tau\| \cdot \|\tau^{-1}\| \leq K$, and examine the weight functions $|\det \tau|^{-1} \varphi \circ \tau^{-1}$. For affine transformations, note that

$$|\det \tau| = \left\| \frac{d\lambda \circ \tau}{d\lambda} \right\|_{\infty} = \left\| \frac{d\lambda \circ \tau^{-1}}{d\lambda} \right\|_{\infty}^{-1},$$

so $C_\delta = 1$. Thus ($R_1$)–($R_3$) are satisfied trivially. If $I$ is the unit ball centred at 0, then for any sequence of linear transformations, $\{\tau_k^{-1} I\}$ is necessarily a $\beta$-sequence with constant $\beta < 1$, because all members of the sequence contain the common point 0. Thus ($R_5$) is easily satisfied. The condition $\|\tau\| \cdot \|\tau^{-1}\| \leq K$ guarantees that

$$\frac{\mathcal{R}^n(\tau B)}{\lambda(\tau B)} \leq \frac{K^n}{\lambda(I)},$$

and so ($R_4$) is satisfied as well. Conditions ($R_6$) and ($R_7$) are also trivial in this setting.

In [3], the conditions ($R_1$)–($R_4$) are again trivial. (($R_4$) holds because there we have only intervals in $\mathbb{R}^1$). ($R_5$) is a direct generalization of the $\beta$-sequence condition defined by Akçoğlu-Déniel. The same techniques used there to apply their results to the discrete flow $\{T^k\}$ may also be used in our setting. Here however, the non-linearity of $\tau_k$ allows us to obtain stronger results. To begin with, we restrict our attention to the case $n = 1$, that is, we consider only $\tau_k\colon \mathbb{R} \longrightarrow \mathbb{R}$. We will use this to investigate the averages

$$\mathcal{A}_k f(x) = \frac{1}{\# E_k} \sum_{i \in E_k} f(T^i x)$$

for quite general sequences of sets of integers $\{E_k\}$. Given an ergodic, invertible measure preserving transformation $T: X \longrightarrow X$, we consider the auxiliary system $\tilde{X} = X \times [0,1)$ (with the product measure $\mu \times \lambda$), and $\tilde{T}_t(x,s) = (T^{\lfloor t+s \rfloor}x, [[t+s]])$. This is the standard flow under a ceiling function of unit height. Let $\{E_k\}$ be a sequence of sets of integers and set $l_k = \#E_k$. To be specific suppose

$$E_k = \{j_{k1} < j_{k2} < j_{k3} < \cdots < j_{kl_k}\}$$

and define $\tilde{E}_k = \bigcup_{i=1}^{l_k} [j_{ki}, j_{ki} + 1)$ and $F_k = (-\infty, j_{k1}) \bigcup \tilde{E}_k$. Define $\tau_k: \mathbb{R} \longrightarrow \mathbb{R}$ as

$$\tau_k(t) = \mathbf{1}_{F_k^c}(t) + \frac{1}{l_k} \int_{j_{k1}}^{t} \mathbf{1}_{F_k}(t)\mathbf{1}_{F_k}(s) + \mathbf{1}_{F_k^c}(t)\mathbf{1}_{F_k^c}(s)\, ds$$

and finally let $\varphi = \mathbf{1}_{[0,1)}$. With these definitions, we see that

$$\mathcal{A}^{\varphi_k} f(x) = \mathcal{A}_k f(x)$$

where the definitions for $\mathcal{A}^{\varphi_k} f$ are implicitly assumed to be made with the elements of our auxiliary system. Thus we have reduced the question of convergence of $\mathcal{A}_k f(x)$ to examining the transformations $\{\tau_k\}$ and the question of whether or not $(R_1)$–$(R_5)$ and $(R_7)$ are satisfied. As usual, $(R_1)$ and $(R_2)$ are trivial, and we assume $l_k \longrightarrow \infty$. $(R_3)$ is also trivial, because $\|\delta_k\|_\infty = \|\delta_k^{\sim}\|_\infty^{-1} = l_k^{-1}$. By examining $\{\tau_k^{-1} I\} = \{\tilde{E}_k\}$, we see that $(R_5)$ holds if and only if the sequence $\{[j_{k1}, j_{kl_k}+1)\} = \{[j_{k1}, j_{k1}+\Omega(\tilde{E}_k))\}$ is a $\beta$-sequence, and so Theorem 1.1 implies that we get divergence whenever this fails. Conversely, to obtain convergence, both $(R_4)$ and the asymptotic convexity of $(R_7)$ are guaranteed if we impose the condition $\lim_{k\to\infty} \Omega(\tilde{E}_k)/\#E_k = 1$, although we can weaken this somewhat. Indeed, in this context $(R_4)$ and the asymptotic convexity in $(R_7)$ are very similar to Tempel'man's [22] regularity conditions.

## 2. **Regularity.**

This section examines more closely the definitions and regularity conditions introduced in Section 1. The restrictions $(R_1)$ and $(R_2)$ need no comment.

Intuitively, the first statement of $(R_3)$ says that, for fixed $k$, $\tau_k$ cannot stretch or shrink sets by arbitrarily large factors. The second statement says that the variation between the maximum amount $\tau_k$ stretches sets to the maximum amount it shrinks (or "doesn't stretch") sets must be bounded uniformly for all $k$. We will denote this bound by $C_\delta = \sup_k(\|\delta_k\|_\infty \cdot \|\delta_k^{\sim}\|_\infty)$.

*Super-Regularity.* Intuitively, a sequence $\{E_k\}$ is super-regular if none of the sets in the sequence are too dispersed. More precisely, a sequence is super-regular if each member $E_k$ of the sequence occupies at least a certain fixed proportion of the measure of some ball $B_k$ which contains it. As applied to the sequence $\{\tau_k^{-1}B\}$, where $B$ is a fixed ball, this implies that we can choose a sequence of open balls $\{B_k\}$ and a $\Gamma > 0$ such that $\tau_k^{-1}B \subset B_k$ and $\lambda(B_k) < \Gamma\lambda(\tau_k^{-1}B)$. Thus $(R_4)$ is essentially a strengthening the requirement that $\tau_k^{-1}$ take bounded sets to bounded sets.

Similar types of "regularity" conditions appear throughout the literature. With notation as above, Krengel [14, p. 209] defines a sequence of convex sets $\{E_k\}$ to be regular if there is a *nested* sequence of intervals $B_k \subset \mathbb{R}^n$ such that $E_k \subset B_k$ and such that $\lambda(B_k) \leq \Gamma\lambda(E_k)$. He also defines the notion of restricted convergence along a sequence of intervals $[0, v_k]$ where the corners $v_k$ *remain in a sector*. This notion is directly related to super-regularity. Tempel'man [22] defines regularity similarly to Krengel, but the sequence $\{B_k\}$ is any nested sequence satisfying a number of other conditions. Broise, Déniel and Derriennic's [7] condition $\|\tau\| \cdot \|\tau^{-1}\| < K$ for all $\tau$ corresponds precisely to our condition of super-regularity on the sequence $\{\tau_k^{-1}B\}$ when applied to linear transformations.

*$\beta$-Sequences.* Let $\mathfrak{B}$ denote the collection of all open balls in our chosen norm on $\mathbb{R}^n$, and let $\{E_k\}_{k\in\mathbb{N}}$ be a sequence of bounded, measurable subsets of $\mathbb{R}^n$. Recall that this sequence is called a *$\beta$-sequence* if there exists a constant $\beta \in \mathbb{R}^+$ such that, for any ball $B \in \mathfrak{B}$,

$$\lambda(\{v \in \mathbb{R}^n \mid \exists k, v + E_k \subset B\}) \leq \beta\lambda(B). \tag{1}$$

We enquire as to how this definition depends on the collection $\mathfrak{B}$ of measurable sets from which we choose $B$. Suppose $\mathfrak{E}$ and $\mathfrak{F}$ are two collections of measurable sets, and suppose there exists a constant $C$ such that, for any set $E \in \mathfrak{E}$, there is an $F \in \mathfrak{F}$ with $E \subset F$ and $\lambda(F) < C\lambda(E)$. It is clear that if (1) holds for any $B \in \mathfrak{F}$ (with a particular constant $\beta$), then it also holds for any $B \in \mathfrak{E}$ (with the constant $\beta' = C\beta$—see Lemma 2.2 below). Thus our definition of $\beta$-sequence is independent of the particular norm $|\cdot|$ which we use. For most sequences $\{E_k\}$, the determination of whether or not the sequence is a $\beta$-sequence would change significantly if we widened our collection of test sets $\mathfrak{B}$ from just the open balls to the whole Borel $\sigma$-algebra. However, when $\{E_k\}$ is a sequence of balls, it would not change. Indeed, we have the following lemma.

LEMMA 2.1. *For any sequence of balls $\{B_k\}_{k\in\mathbb{N}}$, the following are equivalent:*

*i) $\{B_k\}_{k\in\mathbb{N}}$ is a $\beta$-sequence (with constant $\beta$).*

*ii) There is a constant $\beta'$, such that for any measurable set $E \subset \mathbb{R}^n$,*

$$\lambda(\{v \in \mathbb{R}^n \mid \exists k, v + B_k \subset E\}) \leq \beta'\lambda(E). \tag{2}$$

*iii) If $B_k = B_{u_k}(r_k)$, then there is a constant $C$ such that, for all $t > 0$,*

$$\lambda\big(\{v \mid \exists k, |v - u_k| \leq (t - r_k)\}\big) \leq Ct^n.$$

REMARK. In the proof of this lemma, we use the following result. It is known that any norm on $\mathbb{R}^n$ satisfies the following decomposition property. There exist constants $L \in \mathbb{N}$ and $l > 0$ such that any non-trivial open set $O \subset \mathbb{R}^n$ can be written as a union of open balls $O = \bigcup_{k=1}^{\infty} B_k$ satisfying

a) $lB_k \cap O^c \neq \emptyset$;

b) for all $v \in O$, $\#\{k \mid v \in B_k\} \leq L$, *i.e.* the collection $\{B_k\}$ is "$L$-overlapping".

This form of decomposition, called a *Whitney decomposition*, is taken from [10, pp. 66–71]. The pair of constants $l$, $L$ will be called *Whitney constants*. With this definition, we may add the following statement to the lemma:

*The minimum possible constants for (i) and (ii) are related by*

$$\beta \leq \beta' \leq L(2+l)^n \beta$$

*where $l$ and $L$ are Whitney constants for the norm $|\cdot|$.*

PROOF. ii) $\Rightarrow$ i) trivially with $\beta = \beta'$.

i) $\Rightarrow$ ii). Given $\varepsilon > 0$, find an open set $O$ such that $E \subset O$ and $\lambda(O \setminus E) < \varepsilon$. Then $\{v \mid \exists k, v + B_k \subset E\} \subset \{v \mid \exists k, v + B_k \subset O\}$, so we need only show (ii) for $O$, since then

$$\begin{aligned}\lambda(\{v \mid \exists k, v + B_k \subset E\}) &\leq \lambda(\{v \mid \exists k, v + B_k \subset O\}) \\ &\leq \beta' \lambda(O) \\ &\leq \beta' \lambda(E) + \varepsilon \beta'.\end{aligned}$$

Write $O$ as a Whitney decomposition $O = \bigcup_{i=1}^{\infty} O_i$. Suppose $v + B_k \subset \bigcup_{i=1}^{\infty} O_i$. If $B_k = B_{u_k}(r_k)$, then $v + u_k \in O_i$ for some $i$. Now, $lO_i \cap O^c \neq \emptyset$, so also $lO_i \cap (v + B_k)^c \neq \emptyset$, and hence $v + B_k \subset (l+2)O_i$. Thus $\{v \mid \exists k, v + B_k \subset O\} \subset \bigcup_{i=1}^{\infty} \{v \mid \exists k, v + B_k \subset (2+l)O_i\}$, so

$$\begin{aligned}\lambda(\{v \mid \exists k, v + B_k \subset O\}) &\leq \sum_{i=1}^{\infty} \lambda\big(\{v \mid \exists k, v + B_k \subset (2+l)O_i\}\big) \\ &\leq \sum_{i=1}^{\infty} \beta \lambda\big((2+l)O_i\big) \\ &\leq \beta(2+l)^n L \lambda(O).\end{aligned}$$

Setting $\beta' = \beta(2+l)^n L$ yields the result.

For the equivalence of iii), see, *e.g.*, [3]. Condition iii) is the "cone" condition, used by Bellow, Jones, Rosenblatt, and Nagel, Stein, and Sueiro in [4, 17, 21]. ∎

We state the following easy facts without proof. In fact we have used this result in the discussion above.

LEMMA 2.2. *i) If $\{E_k\}$ is a $\beta$-sequence and $v \in \mathbb{R}^n$, then $\{v + E_k\}$ is a $\beta$-sequence.*

*ii) If $\{E_k\}$ is a $\beta$-sequence and $E_k \subset F_k$ for all $k$, then $\{F_k\}$ is a $\beta$-sequence. (We assume all $F_k$ are measurable and bounded.)*

LEMMA 2.3. *$(R_5)^c$ holds if and only if $(R_5)$ is false.*

*Localizing regularity.* This property ensures that the averages $\mathcal{A}^{\varphi_k} f(x)$ are "local averages" about the point $x$. The following lemma gives some equivalent formulations.

LEMMA 2.4. *Assume $\tau_k$ satisfy conditions $(R_1)$–$(R_4)$. Then the following are equivalent.*

*i)* $\lim_{k\to\infty} \tau_k^{-1}(0) = 0$, $\lim_{k\to\infty} \|\delta_k^{\sim}\|_\infty = 0$.

*ii) Let $B$ be any ball. Then, for any $\varepsilon > 0$, there is a $K$ such that for all $k \geq K$, $\tau_k^{-1} B \subset B_0(\varepsilon)$.*

*iii)* *For $\lambda$-a.e. $v \in \mathbb{R}^n$, $\tau_k^{-1} v$ converges to zero.*

*iv)* *There is a set $E$ of non-zero measure such that, for any $v \in E$, $\tau_k^{-1} v$ converges to zero.*

It may seem odd that ($R_6$) should depend on the values of the $\tau_k$ at the single point $v = 0$. This is merely a convenience which avoids many unnecessary "a.e." arguments. Furthermore, it emphasizes the parallels between this condition and similar conditions in the literature. Clearly we could replace ($R_6$) by (ii) above. Then in the definition of $\mathcal{R}(E)$, we would replace the containment $E \subset B_v(r)$ by containment a.e. The arguments would merely become more cumbersome. Instead it is simpler to work with the definitions made, and then note that all the convergence and maximal results clearly hold for any sequence $\{\tau_k'\}$ whose members are equal a.e. to a sequence satisfying the given regularity conditions.

*Globalizing regularity.* The discussion of the consequences of ($R_7$) is postponed until Section 6 where it is needed.

3. **Divergence of averages.** In this section, we prove Theorem 1.1. We need some preliminary lemmas.

The next result requires some notation. Let $Q \subset \mathbb{R}^n$. We say $Q$ *tiles* $\mathbb{R}^n$ iff $Q$ is measurable with compact closure, and if there exists a closed additive subgroup $H \subset \mathbb{R}^n$ such that $\mathbb{R}^n/H$ is compact and such that the projection mapping $\pi: \mathbb{R}^n \longrightarrow \mathbb{R}^n/H$ is bijective on $Q$. If $F \subset X$, let $T_Q F \stackrel{\text{def}}{=} \bigcup_{v \in Q} T_v F$. $T_Q F$ is called *disjoint* iff $\{T_v F\}_{v \in Q}$ is disjoint. The following lemma is known (see [15]).

LEMMA 3.1. *Let $\{T_v\}_{v \in \mathbb{R}^n}$ be an aperiodic $\mathbb{R}^n$-flow on $(X, \mathcal{F}, \mu)$. Then for any $Q$ which tiles $\mathbb{R}^n$, and any $\varepsilon > 0$, there is a measurable set $E \subset X$ such that $T_Q E$ is disjoint and measurable, and such that $\mu(T_Q E) > (1-\varepsilon)\mu(X)$. Furthermore, on $T_Q E$ the measure $\mu$ is the completed product of a measure $\mu_E$ on $E$ with $\lambda$ on $Q$.*

REMARK. The last statement in this lemma implies the following. Let $Q$ and $E$ be as in the lemma. If $P \subset Q$, $F \subset E$ are measurable, then

$$\mu(T_P F) = \frac{\lambda(P)}{\lambda(Q)} \mu(T_Q F). \tag{1}$$

We will not need the condition $\mu(T_Q E) > (1-\varepsilon)\mu(X)$, but we require that any bounded set be contained in a set $Q$ which tiles $\mathbb{R}^n$. This is trivial, since every cube $Q = [-a, a)^n$ tiles $\mathbb{R}^n$.

LEMMA 3.2. *Let $\{T_v\}_{v \in \mathbb{R}^n}$ be an aperiodic $\mathbb{R}^n$-flow on $(X, \mathcal{F}, \mu)$. Suppose that ($R_1$), ($R_2$), and $(R_5)^c$ hold. Let $\varphi$ be any weight function. Then for each $M > 0$, $K \in \mathbb{N}$, and $\varepsilon > 0$, we can find two sets $C$ and $D$ in $X$ such that*

$$0 < M\mu(C) < \mu(D)$$

*and such that*

$$\sup_{k\geq K} \mathcal{A}^{\varphi_k}\mathbf{1}_C(x) \geq 1-\varepsilon, \quad \forall x \in D.$$

PROOF. Find a ball $B$ such that the integral of $\varphi$ over $B$ is greater than $1-\varepsilon$. Then for any $k$,

$$\int_{\tau_k^{-1}B} \varphi_k\, d\lambda = \int_{\tau_k^{-1}B} \varphi\circ\tau_k(v)\delta_k(v)\lambda(dv) = \int_B \varphi(v)\lambda(dv) > 1-\varepsilon.$$

By Lemma 2.3, the hypotheses imply that $\{\tau_k^{-1}B\}$ is not a $\beta$-sequence for any ball $B$, so nor is its tail $\{\tau_k^{-1}B\}_{k\geq K}$ for any $K$. Thus we can find a ball $A$ such that

$$\lambda(\{v\in\mathbb{R}^n \mid \exists k\geq K, v+\tau_k^{-1}B\subset A\}) > M\lambda(A).$$

Then there is an integer $J$ such that

$$\lambda(\{v\in\mathbb{R}^n \mid \exists k, K\leq k\leq J, v+\tau_k^{-1}B\subset A\}) > M\lambda(A).$$

Let $P=\{v\in\mathbb{R}^n \mid \exists k, K\leq k\leq J, v+\tau_k^{-1}B\subset A\}$, so $\lambda(P) > M\lambda(A)$. Since $A$ and $P$ are both bounded, we can find $Q$ such that $A\cup P\subset Q$ and such that $Q$ tiles $\mathbb{R}^n$. Finally, let $E$ be chosen as in Lemma 3.1, and let $C=T_A(E)$ and $D=T_P(E)$. Then we see that $C$ and $D$ satisfy the requirements of the lemma as follows.

We have

$$0 < M\frac{\lambda(A)}{\lambda(Q)}\mu(T_QE) < \frac{\lambda(P)}{\lambda(Q)}\mu(T_QE)$$

and so by (1), $0 < M\mu(C) < \mu(D)$, satisfying the first requirement.

Next, let $x\in D$, so $x=T_v(y)$ for some choice of $v\in P$, $y\in E$. Then there is a $k$, $K\leq k\leq J$, such that $v+\tau_k^{-1}B\subset A$. Thus for all $u\in\tau_k^{-1}B$, we have $v+u\in A$, and so $T_u(x)=T_{v+u}(y)\in T_A(E)=C$. Then for this $x$, with this choice of $k$,

$$\mathcal{A}^{\varphi_k}\mathbf{1}_C(x) \geq \int_{\tau_k^{-1}B} \mathbf{1}_C(T_ux)\varphi_k(u)\lambda(du) = \int_{\tau_k^{-1}B}\varphi_k(u)\lambda(du) > 1-\varepsilon,$$

and so for all $x\in D$, $\sup_{k\geq K}\mathcal{A}^{\varphi_k}\mathbf{1}_C(x)\geq 1-\varepsilon$. ∎

The following lemma follows easily from the ergodicity of the flow and the mean ergodic theorem.

LEMMA 3.3. *Let $\{D_k\}$ be a sequence of measurable subsets of $X$ which satisfy*

$$\sum_{k=1}^{\infty}\mu(D_k)=\infty.$$

*If $\{T_v\}_{v\in\mathbb{R}^n}$ is an ergodic $\mathbb{R}^n$-flow on $X$, then there is a sequence of points $\{v_k \mid v_k\in\mathbb{R}^n\}$ such that*

$$\mu\Big(\bigcup_{k=K}^{\infty} T_{v_k}D_k\Big) = \mu(X)$$

*for any* $K \in \mathbb{N}$.

PROOF OF THEOREM 1.1. Given $\varepsilon, \eta > 0$, we may choose a sequence, $\{(C_k, D_k)\}$, of pairs of non-null subsets of $X$, and a sequence of integers $K_k$ converging to infinity, such that the following conditions are satisfied:

i) $\sum_{k=1}^{\infty} \mu(C_k) < \eta\mu(X)$.

ii) $\sum_{k=1}^{\infty} \mu(D_k) = \infty$.

iii) $\sup_{i \geq K_k} \mathcal{A}^{\varphi_i} \mathbf{1}_{C_k}(x) \geq 1 - \varepsilon, \forall x \in D_k$.

Briefly, we see this as follows. Apply Lemma 3.2 repeatedly to generate the sequence $\{(C_k, D_k)\}$, choosing successively larger ratios $M$ in such a way as to guarantee (i). To satisfy (ii), we can repeat pairs in our sequence whenever $\mu(D) < \mu(X)/2$, and this can be done without violating (i). The sequence thus chosen will also satisfy (iii).

Conditions (i) through (iii) are unaffected if we replace $\{C_k\}$ and $\{D_k\}$ by $\{C'_k = T_{v_k} C_k\}$ and $\{D'_k = T_{v_k} D_k\}$, so by the previous lemma, we may also assume $\mu(\bigcup_{k=K}^{\infty} D'_k) = \mu(X)$ for all $K \in \mathbb{N}$. Let $E = \bigcup_{k=1}^{\infty} C'_k$, and define $f = \mathbf{1}_E$. Then

$$\int_X f \, d\mu < \eta\mu(X),$$

yet

$$\sup_{k \geq K} \mathcal{A}^{\varphi_k} f(x) \geq 1 - \varepsilon$$

for $\mu$-a.e. $x$. If we choose $\eta + \varepsilon < 1$, then $\mathcal{A}^{\varphi_k} f(x)$ must diverge on a set of positive measure. ∎

4. **Re-arrangements.** This section contains a number of elementary, but very technical results. These results are not readily available in the literature in the form that we require, and hence we present them here, primarily for the reader's convenience. Their proofs can be found in [16].

Let $f$ be a non-negative, measurable function on some non-atomic measure space $(Y, \mathcal{E}, \nu)$. Define the distribution of $f$ to be the measure $D_f$ on the Borel sets of $\mathbb{R}$ given by

$$D_f(E) = \nu\Big(f^{-1}\big(E \cap (0, \infty)\big)\Big).$$

Define $f'(t) = D_f(t, \infty)$ for $t > 0$, and zero otherwise. Then $f'$ is non-increasing on $(0, \infty)$, but may be infinite at all points $t$ in some interval $(0, a)$. Whenever this is not the case, *i.e.* whenever $f' < \infty$, define the re-arrangement of $f$ to be $f^* = f''$. We call any non-negative measurable function $\xi$ on $\mathbb{R}$ a *re-arrangement* if it is zero on $(-\infty, 0]$, non-increasing on $(0, \infty)$, and $\lim_{t \to \infty} \xi(t) = 0$. Let $L_*(Y)$ denote the set of all functions $f$ for which $|f|^*$ is defined. As usual $L_*^+(Y)$ denotes all non-negative $f \in L_*(Y)$.

We note the following facts.

LEMMA 4.1. *Let* $f \colon Y \longrightarrow \mathbb{R}^+$ *be* $\nu$*-measurable.*

*i) If* $f' < \infty$, *then* $f^* < \infty$, $\lim_{t \to \infty} f^*(t) = 0$ *(so we are justified in calling* $f^*$ *a re-arrangement) and*

$$D_f = D_{f^*}.$$

https://doi.org/10.4153/CJM-1995-044-0 Published online by Cambridge University Press

*ii) Any two re-arrangements with the same distribution are equal except possibly on a set of $\ell$-measure zero.*

*iii) $f^*$ is defined (and finite) for any non-negative $f$ with support of finite measure. In this case the support of $f^*$ is a finite interval of length $\nu(S_f)$, beginning at* 0.

*iv) $f^*$ is defined (and finite) for any non-negative $f \in L_p(Y)$, $1 \leq p < \infty$. Furthermore $(f^p)^* = (f^*)^p$ a.e., $f^* \in L_p(\mathbb{R})$ and in fact $\|f\|_p = \|f^*\|_p$.*

*v) If $f \in L_\infty$ is such that $f^*$ is defined (and finite), then $\|f\|_\infty = \|f^*\|_\infty$.*

We remark that (iv) implies, for $f \in L_p(X)$, $1 \leq p < \infty$,

$$f^*(t) \leq \frac{\|f\|_p}{t^{1/p}} \quad \forall t > 0,$$

and thus if $f_a \longrightarrow f$ in $L_p(X)$, $1 \leq p \leq \infty$, then $|f - f_a|^* \longrightarrow 0$ pointwise.

LEMMA 4.2. *Let $f \in L_*^+(\mathbb{R}^n)$. Let $\tau: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a measurable, non-singular mapping, and let $\delta^\sim = d\lambda \circ \tau^{-1} / d\lambda$ be the Radon-Nikodym derivative of the measure $\lambda \circ \tau^{-1}$. If $\delta^\sim \in L_\infty(\mathbb{R}^n)$, then $(f \circ \tau)^*(t) \leq f^*(t / \|\delta^\sim\|_\infty)$.*

LEMMA 4.3. *For any $f \in L_*^+(\mathbb{R}^n)$, there is a measure preserving map $\tau: S_f \longrightarrow S_{f^*}$ such that $f(v) = f^*(\tau v)$ for $\lambda$-a.e. $v$. $\tau$ is surjective but not necessarily injective.*

LEMMA 4.4. *Let $f \in L_*^+(\mathbb{R}^n)$. Given a re-arrangement function $\xi$ on $\mathbb{R}$, there is a function $g$ on $\mathbb{R}^n$ such that $g^* = \xi$ and such that*

$$\int_{\mathbb{R}^n} fg\, d\lambda = \int_{\mathbb{R}} f^* g^*\, d\ell = \int_{\mathbb{R}} f^* \xi\, d\ell.$$

*Furthermore, if $E \subset \mathbb{R}^n$ is such that $S_f \subset E$ and $S_\xi \subset \big(0, \lambda(E)\big)$, then the support of $g$ can be taken within $E$.*

LEMMA 4.5. *Let $f, g \in L_*^+(Y)$, and let $A \subset Y$ be measurable. Then*

$$\begin{aligned} \int_A fg\, d\nu &\leq \int_{\mathbb{R}} (f\mathbf{1}_A)^* (g\mathbf{1}_A)^*\, d\ell \\ &\leq \int_{\mathbb{R}} (f\mathbf{1}_A)^* g^*\, d\ell \\ &\leq \int_0^{\nu(A)} f^* g^*\, d\ell. \end{aligned}$$

Let $(X, \mathcal{F}, \mu)$ be another non-atomic measure space, and let $F$ be a function from $X \times Y$ to $\mathbb{R}^+$ which is measurable with respect to $\mu \times \nu$. Define $F_y(x) = F(x,y)$, and suppose that, for $\nu$-a.e. $y \in Y$, $F_y \in L_*^+(X)$. Define $\tilde{F}: \mathbb{R} \times Y \longrightarrow \mathbb{R}^+$ by

$$\tilde{F}(s,y) = F_y^*(s).$$

LEMMA 4.6. *If $F \in L_*^+(X \times Y)$ or $\tilde{F} \in L_*^+(\mathbb{R} \times Y)$, then $F^* = \tilde{F}^*$.*

Suppose $F$ has the form $F(x,y) = f(x,y)g(y)$ where $f: X \times Y \longrightarrow \mathbb{R}^+$ and $g: Y \longrightarrow \mathbb{R}^+$ satisfy: $f_y^*(s) = \xi\big(s/c(y)\big)$ for $c(y) > 0$ such that $cg \in L_1(Y)$, and $\xi$ such that $\xi = \xi^*$.

LEMMA 4.7. *$F^*(s) = \xi(s/C)$, where $C = \int_Y c(y)g(y)\,dy$.*

We will use the following applications of Lemma 4.7.

COROLLARY 4.8. *i) Let $\{f_i\}_{i=1}^m$ be a collection of functions in $L_*^+(X)$ with disjoint supports $\{A_i\}_{i=1}^m$, satisfying $f_i^*(t) = \xi(t/c_i)$ for some fixed $\xi = \xi^*$ and $c_i > 0$. If $f = \sum_{i=1}^m f_i$, then $f^* = \xi(t/C)$ where $C = \sum c_i$.*

*ii) Let $Y = \mathbb{R}^n$, and let $E \subset \mathbb{R}^n$ be of non-zero, finite measure. Let $X$ be a finite measure space, let $\{T_v\}$ be an $\mathbb{R}^n$-flow on $X$, and let $f \in L_*^+(X)$. If $F(x,v) = f(T_v x)\mathbf{1}_E(v)$, then $F^*(t) = f^*\big(t/\lambda(E)\big)$. In alternate notation, if $\psi_x(v) = f(T_v x)\mathbf{1}_E(v)$, then $\psi^*(t) = f^*\big(t/\lambda(E)\big)$.*

*iii) Let $\varphi^* \in L_*^+(\mathbb{R})$, let $Y$ be a finite measure space, and let $F(t,y) = \varphi^*(t)$. Then $F^*(t) = \varphi^*\big(t/\nu(Y)\big)$.*

LEMMA 4.9. *Let $f, g \in L_*^+(X \times Y)$. Then $f_y, g_y \in L_*^+(X)$ $\nu$-a.e. $y$ and*

$$\begin{aligned}\int_{X\times Y} fg &\le \int_Y \int_{\mathbb{R}} f_y^*(s) g_y^*(s)\,ds\,dy\\ &\le \int_{\mathbb{R}} f^* g^*\,d\ell.\end{aligned}$$

The following easy lemma will be used occasionally.

LEMMA 4.10. *If $\xi = \xi^*$, $\eta = \eta^*$, and $a, b > 0$, then*

$$\int_{\mathbb{R}} \xi(t/a)\eta(t/b)\,dt \le \max(a,b) \int_{\mathbb{R}} \xi(t)\eta(t)\,dt.$$

## 5. **Maximal estimates.**

In this section we prove a Hardy-Littlewood type maximal theorem:

THEOREM 5.1. *Suppose $\{\tau_k\}$ satisfy ($R_1$) through ($R_5$). Given any ball $B$, there exists a constant $C_B$ such that if $\varphi$ is any weight function with $S_\varphi \subset B$, then*

$$\mu\big(\{x \mid \exists k, \mathcal{A}^{\varphi_k} f(x) > \alpha\}\big) \le \frac{C_B}{\alpha} \int_{\mathbb{R}} f^*(t)\varphi^*(t)\,d\ell$$

*for all $\alpha > 0$ and all $f: X \longrightarrow \mathbb{R}^+$.*

Once we have proved some preliminary lemmas, we prove a maximal-type theorem for the measure space $\mathbb{R}^n$ under the "flow" $T_v(u) = v + u$. Then we use Calderón's transference principle to get a maximal estimate for the general case $(X, \mathcal{F}, \mu, T_v)$.

For the remainder of this section, we will assume that $\{\tau_k\}$ satisfies regularity conditions ($R_1$) through ($R_5$). Recall that if $B' \subset B$ and $\{\tau_k^{-1}B'\}$ is a $\beta$-sequence, then so is $\{\tau_k^{-1}B\}$. Thus our regularity conditions guarantee that given any $B'' \supset S_\varphi$ we can choose a ball $B$, a constant $\Gamma > 1$, and a sequence of balls $\{B_k\}$ such that

i) $S_\varphi \subset B'' \subset B$;

ii) $\tau_k^{-1}B$ is a $\beta$-sequence;
iii) $\tau_k^{-1}B \subset B_k$, so $\{B_k\}$ is necessarily also a $\beta$-sequence;
iv) $\lambda(B_k) < \Gamma\lambda(\tau_k^{-1}B)$.

Fix $B$, $\Gamma$ and $\{B_k\}$ in this manner, and fix any $\psi \in L_*^+(\mathbb{R}^n)$. Then for any Borel set $E \subset \mathbb{R}^n$ of finite, non-zero measure, define

$$\Xi(E) = \frac{\lambda(B)}{\lambda(E)} \int_{\mathbb{R}} (\psi \mathbf{1}_E)^*(t)\varphi^*\Big(t\frac{\lambda(B)}{\lambda(E)}\Big)\,dt.$$

By Lemma 4.4, there is a $\theta_E \in L_*(\mathbb{R}^n)$ with support in $E$, such that $\theta_E^*(t) = \varphi^*\big(t\lambda(B)/\lambda(E)\big)$, and such that

$$\Xi(E) = \frac{\lambda(B)}{\lambda(E)} \int_{\mathbb{R}^n} \psi\theta_E\,d\lambda.$$

For $\psi \in L_*(\mathbb{R}^n)$ we abuse our notation slightly and define

$$\mathcal{A}^{\varphi_k}\psi(u) = \int_{\mathbb{R}^n} \psi(u+v)\varphi_k(v)\lambda(dv).$$

This is just a special case of our previous definition of $\mathcal{A}^{\varphi_k}$, in which $X = \mathbb{R}^n$ and $T_u(v) = u + v$.

LEMMA 5.2. *If $\mathcal{A}^{\varphi_k}\psi(u) > \alpha$, then $\Xi(u + \tau_k^{-1}B) > \alpha/C_\delta$, where $C_\delta = \sup_k \|\delta_k\|_\infty \|\delta_k^\sim\|_\infty$.*

PROOF. We begin by remarking that $\|\delta_k^\sim\|_\infty \lambda(B)/\lambda(u + \tau_k^{-1}B) \geq 1$ and that Lemma 4.2 implies $\varphi_k^*(t) \leq \|\delta_k\|_\infty \varphi^*(t/\|\delta_k^\sim\|_\infty)$. Since $S_{\varphi_k} \subset \tau_k^{-1}B$, we may calculate as follows.

$$\begin{aligned}
\alpha &< \int_{\mathbb{R}^n} \psi(u+v)\varphi_k(v)\lambda(dv) \\
&\leq \int_{\mathbb{R}} (\psi\mathbf{1}_{u+\tau_k^{-1}B})^*(t)\varphi_k^*(t)\ell(dt) \\
&\leq \|\delta_k\|_\infty \int_{\mathbb{R}} (\psi\mathbf{1}_{u+\tau_k^{-1}B})^*(t)\varphi^*(t/\|\delta_k^\sim\|_\infty)\ell(dt) \\
&\leq \|\delta_k\|_\infty \max\Big(1, \|\delta_k^\sim\|_\infty \frac{\lambda(B)}{\lambda(u+\tau_k^{-1}B)}\Big) \times \\
&\qquad \int_{\mathbb{R}} (\psi\mathbf{1}_{u+\tau_k^{-1}B})^*(t)\varphi^*\Big(t\frac{\lambda(B)}{\lambda(u+\tau_k^{-1}B)}\Big)\ell(dt) \\
&\leq C_\delta \Xi(u + \tau_k^{-1}B).
\end{aligned}$$

■

COROLLARY 5.3. *Suppose $\tau_k^{-1}B \subset B_k$ and $\lambda(B_k) \leq \Gamma\lambda(\tau_k^{-1}B)$. If $\mathcal{A}^{\varphi_k}\psi(u) > \alpha$, then*

$$\Xi(u + B_k) > \frac{\alpha}{C_\delta \Gamma}.$$

PROOF.

$$\begin{aligned}
\Xi(u + B_k) &\geq \frac{\lambda(B)}{\Gamma\lambda(u+\tau_k^{-1}B)} \int_{\mathbb{R}} (\psi\mathbf{1}_{u+\tau_k^{-1}B})^*(t)\varphi^*\Big(t\frac{\lambda(B)}{\lambda(u+\tau_k^{-1}B)}\Big)\,dt \\
&= \frac{\Xi(u+\tau_k^{-1}B)}{\Gamma} > \frac{\alpha}{C_\delta\Gamma}.
\end{aligned}$$

■

LEMMA 5.4. *Let $\alpha > 0$, $R > 0$ be fixed. Let $H$ be the union of all open balls $A \subset B_0(R)$ for which $\Xi(A) > \alpha$. Then*

$$\lambda(H) < \frac{3^n\lambda(B)}{\alpha}\int_{\mathbb{R}}\psi^*(t)\varphi^*\Big(t\frac{\lambda(B)}{\lambda\big(B_0(R)\big)}\Big)\,dt. \tag{2}$$

PROOF. From the collection of balls $\{A_i\}$ whose union gives $H$ we can find a finite, disjoint collection $\{A_i\}_{i=1}^m$ such that $\lambda(H) \leq 3^n\sum_{i=1}^m \lambda(A_i)$ (see [20], p. 164). Now, $\alpha < \Xi(A_i)$ implies $\lambda(A_i) < \lambda(B)/\alpha\int_{\mathbb{R}^n}\psi(v)\theta_{A_i}(v)\,dv$, and thus

$$\begin{aligned}
\lambda(H) &< \frac{3^n\lambda(B)}{\alpha}\int_{\mathbb{R}^n}\psi(v)\Big(\sum_{i=1}^m\theta_{A_i}\Big)(v)\,dv\\
&\leq \frac{3^n\lambda(B)}{\alpha}\int_{\mathbb{R}}\psi^*(t)\Big(\sum_{i=1}^m\theta_{A_i}\Big)^*(t)\,dt\\
&= \frac{3^n\lambda(B)}{\alpha}\int_{\mathbb{R}}\psi^*(t)\varphi^*\Big(t\frac{\lambda(B)}{\sum_{i=1}^m\lambda(A_i)}\Big)\,dt\\
&\leq \frac{3^n\lambda(B)}{\alpha}\max\Big(1,\frac{\sum_{i=1}^m\lambda(A_i)}{\lambda\big(B_0(R)\big)}\Big)\int_{\mathbb{R}}\psi^*(t)\varphi^*\Big(t\frac{\lambda(B)}{\lambda\big(B_0(R)\big)}\Big)\,dt\\
&= \frac{3^n\lambda(B)}{\alpha}\int_{\mathbb{R}}\psi^*(t)\varphi^*\Big(t\frac{\lambda(B)}{\lambda\big(B_0(R)\big)}\Big)\,dt.
\end{aligned}$$

The two equalities hold by virtue of the disjointedness of $\{A_i\}_{i=1}^m$, Corollary 4.8(i) and of course $A_i \subset B_0(R)$. ∎

For the proof of the main theorem, we expand the notation slightly. Instead of having a fixed $\psi \in L^+_*(\mathbb{R}^n)$, we will now deal with a $\psi_x \in L^+_*(\mathbb{R}^n)$ which depends on $x \in X$. Thus our function $\Xi$ depends on $x$, and we denote this by $\Xi(E,x)$. The set $H$ also depends on $x$ and on the value of $\alpha$, so we denote this by $H(\alpha,x)$.

PROOF OF THEOREM 5.1. Recall that for $\varphi$ with compact support, we have chosen balls $B$, $\{B_k\}_{k\in\mathbb{N}}$ and constant $\Gamma > 1$ to satisfy

1) $S_\varphi \subset B$;
2) $\{\tau_k^{-1}B\}$ is a $\beta$-sequence (with constant $\beta$);
3) $\tau_k^{-1}B \subset B_k$;
4) $\lambda(B_k) < \Gamma\lambda(\tau_k^{-1}B)$ for all $k \in \mathbb{N}$.

Let $f \in L^+_*(X)$. Fix $K \in \mathbb{N}$ and define

$$E_K = \{x \mid \exists k \leq K, \mathcal{A}^{\varphi_k}f(x) > \alpha\}$$

Choose $m > 0$ such that $B \subset B_0(m)$ and $B_k \subset B_0(m)$ for all $k \leq K$. Fix arbitrary $M > 0$, and let $R = M + m$, so

$$\tau_k^{-1}B + B_0(M) \subset B_k + B_0(M) \subset B_0(m) + B_0(M) \subset B_0(R).$$

For all $x \in X$, define $\psi_x \in L_*^+(\mathbb{R}^n)$ by

$$\psi_x(u) = f(T_u x)\mathbf{1}_{B_0(R)}(u).$$

Let $W = \{(x,u) \in X \times \mathbb{R}^n \mid u \in B_0(M), \exists k \leq K \ni \mathcal{A}^{\varphi_k}\psi_x(u) > \alpha\}$. For $k \leq K$, $S_{\varphi_k} \subset \tau_k^{-1}B \subset B_0(m)$, and so if $u \in B_0(M)$

$$\begin{aligned}
\mathcal{A}^{\varphi_k}\psi_x(u) &= \int_{\mathbb{R}^n} f(T_{u+v}x)\mathbf{1}_{B_0(R)}(u+v)\varphi_k(v)\,dv \\
&= \int_{B_0(m)} f(T_{u+v}x)\mathbf{1}_{B_0(M+m)}(u+v)\varphi_k(v)\,dv \\
&= \int_{\mathbb{R}^n} f(T_{u+v}x)\varphi_k(v)\,dv \\
&= \mathcal{A}^{\varphi_k}f(T_u x).
\end{aligned}$$

Thus

$$W = \{(x,u) \mid u \in B_0(M), T_u x \in E_K\}, \tag{3}$$

which is clearly measurable since $E_K$ is measurable. Let $W_x = \{u \mid (x,u) \in W\}$. If $u \in W_x$, then Corollary 5.3 implies that for some $k \leq K$, $\Xi(u + B_k, x) > \alpha/(C_\delta\Gamma)$. Now $u + B_k \subset B_0(R)$, so $u + B_k \subset H\big(\alpha/(C_\delta\Gamma), x\big)$, and thus $W_x \subset \{u \mid \exists k, u + B_k \subset H(\alpha/(C_\delta\Gamma), x)\}$. Since $\{B_k\}$ is a $\beta$-sequence of balls, Lemma 2.1 implies that

$$\lambda(W_x) \leq \beta'\lambda\Big(H\Big(\frac{\alpha}{C_\delta\Gamma}, x\Big)\Big) \leq \beta'\frac{3^n\lambda(B)}{\alpha/(C_\delta\Gamma)}\int_{\mathbb{R}}\psi_x^*(t)\varphi^*\Big(t\frac{\lambda(B)}{\lambda\big(B_0(R)\big)}\Big)\,dt$$

where $\beta' \leq L(2+l)^n\beta$. Now we calculate $\mu \times \lambda(W)$ and use Fubini's Theorem to get an upper bound. On the one hand

$$\begin{aligned}
\mu \times \lambda(W) &= \int_{\mathbb{R}^n}\int_X \mathbf{1}_W(x,v)\mu(dx)\lambda(dv) \\
&= \int_{B_0(M)}\int_X \mathbf{1}_{E_K}(T_v x)\mu(dx)\lambda(dv) \\
&= \mu(E_K)\lambda\big(B_0(M)\big)
\end{aligned}$$

by (3) and the fact that $T_v$ is measure preserving. On the other hand

$$\begin{aligned}
\mu \times \lambda(W) &= \int_X \lambda(W_x)\mu(dx) \\
&\leq \frac{\beta' 3^n C_\delta\Gamma\lambda(B)}{\alpha}\int_X\int_{\mathbb{R}}\psi_x^*(t)\varphi^*\Big(t\frac{\lambda(B)}{\lambda\big(B_0(R)\big)}\Big)\ell(dt)\mu(dx) \\
&\leq \frac{\beta' 3^n C_\delta\Gamma\lambda(B)}{\alpha}\int_{\mathbb{R}} f^*\Big(\frac{t}{\lambda\big(B_0(R)\big)}\Big)\varphi^*\Big(t\frac{\lambda(B)}{\mu(X)\lambda\big(B_0(R)\big)}\Big)\ell(dt) \\
&\leq \frac{\beta' 3^n C_\delta\Gamma\lambda(B)}{\alpha}\max\Big(\lambda\big(B_0(R)\big), \frac{\mu(X)\lambda\big(B_0(R)\big)}{\lambda(B)}\Big)\int_{\mathbb{R}} f^*\varphi^*\,d\ell \\
&= \lambda\big(B_0(R)\big)\frac{\beta' 3^n C_\delta\Gamma\max\big(\lambda(B),\mu(X)\big)}{\alpha}\int_{\mathbb{R}} f^*\varphi^*\,d\ell.
\end{aligned}$$

Dividing the resulting inequality by $\lambda\big(B_0(M)\big)$ gives a bound on $\mu(E_K)$. Letting $M \longrightarrow \infty$ and then $K \longrightarrow \infty$, we get

$$\mu\big(\{x \mid \exists k, \mathcal{A}^{\varphi_k} f(x) > \alpha\}\big) \leq \frac{3^n(2+l)^n L\beta C_\delta \Gamma \max\big(\lambda(B), \mu(X)\big)}{\alpha} \int_{\mathbb{R}} f^* \varphi^* \, d\ell. \qquad \blacksquare$$

6. **Convergence.** In this section, we consider the point-wise convergence of the averages $\mathcal{A}^{\varphi_k} f(x)$ under the assumption of regularity conditions ($R_1$)–($R_5$) and either localizing regularity ($R_6$) or globalizing regularity ($R_7$). We begin with two lemmas which allow us to obtain the a.e. convergence of $\mathcal{A}^{\varphi_k} f$ from the a.e. convergence of various sorts of approximations.

For bounded functions $f$, we will use the following simple lemma.

LEMMA 6.1. *If $\varphi$ and $\varphi'$ are two weight functions, and $f \in L_\infty(X)$, then*

$$|\mathcal{A}^{\varphi_k} f(x) - \mathcal{A}^{\varphi_k} f'(x)| \leq \|\varphi - \varphi'\|_1 \|f\|_\infty.$$

*Thus, given a bounded function $f$, if $\mathcal{A}^{\varphi_k} f$ converges a.e. for a class of weight functions that approximate any weight function in the $L_1(\mathbb{R}^n)$ norm, then $\mathcal{A}^{\varphi_k} f$ converges a.e. for all weight functions.*

The next lemma is where the maximal estimate of Section 5 is applied. Let $\varphi$ be a compact weight function. Define a maximal function $\mathcal{M}^\varphi$ by

$$\mathcal{M}^\varphi f(x) = \sup_k |\mathcal{A}^{\varphi_k} f(x)|.$$

If $\varphi$ has bounded support, then the maximal inequality in Theorem 5.1 implies that there is a constant $C_\varphi$ such that

$$\mu(\{x \mid \mathcal{M}^\varphi f > \alpha\}) \leq \frac{C_\varphi}{\alpha} \int_{\mathbb{R}} |f|^* \varphi^* \, d\lambda.$$

LEMMA 6.2. *Suppose ($R_1$)–($R_5$) hold. Let $\varphi$ be any compact weight function and let $f \in L_*(X)$. Suppose $\{f_a : X \longrightarrow \mathbb{R}\}$ is a family of functions satisfying*

*i) $\{f_a\}$ approximates $f$ in $L_p(X)$ for some $p \in [1, \infty]$, i.e. $\exists p \in [1, \infty]$, $\forall \varepsilon > 0, \exists a$ such that $\|f - f_a\|_p < \varepsilon$.*

*ii) There is a re-arrangement $\xi$ such that $\int_{\mathbb{R}} \xi \varphi^* \, d\ell < \infty$ and such that for any $a$, $|f - f_a|^* < \xi$.*

*iii) For any $a$, $\mathcal{A}^{\varphi_k} f_a(x)$ converges for $\mu$-a.e. $x$.*

*Then $\mathcal{A}^{\varphi_k} f$ converges $\mu$-a.e.*

PROOF. If $\mathcal{A}^{\varphi_k} f$ does not converge a.e. on $X$, then there is an $\alpha > 0$ and a set $E \subset X$ of non-zero measure such that $\limsup_k \mathcal{A}^{\varphi_k} f(x) - \liminf_k \mathcal{A}^{\varphi_k} f(x) > \alpha$ on $E$. Since $\mathcal{A}^{\varphi_k} f_a$ converges a.e. for any $a$, we have

$$\limsup_k \mathcal{A}^{\varphi_k}(f - f_a)(x) - \liminf_k \mathcal{A}^{\varphi_k}(f - f_a)(x) > \alpha$$

for all $x \in E$. Thus $E \subset \{x \mid \mathcal{M}^{\varphi}|f - f_a|(x) > \alpha/2\}$, for any $a$. On the other hand, the maximal inequality in Theorem 5.1 says that the measure of the last set is less than

$$\frac{2C_{\varphi}}{\alpha}\int_{\mathbb{R}} |f - f_a|^* \varphi^*\, d\ell.$$

Since $\lim_a f_a = f$ in the $L_p$ norm, then by the remarks following Lemma 4.1, $\lim_a |f - f_a|^*(t) = 0$ for all $t$. By hypothesis $|f - f_a|^* \varphi^* \leq \xi\varphi^*$ which is integrable on $\mathbb{R}$. Thus by Lebesgue's Dominated Convergence Theorem, this last integral can be made arbitrarily small by suitable choice of $a$. This contradicts $\mu(E) > 0$, so $\mathcal{A}^{\varphi_k} f$ converges $\mu$-a.e. on $X$. ∎

*Local convergence.*

LEMMA 6.3. *Suppose $(R_1)$–$(R_6)$ hold, and let $\varphi$ be any weight function. Then $\mathcal{A}^{\varphi_k} f$ converges $\mu$-a.e. for any function $f$ of the form $f(x) = \int_{B_0(a)} g(T_v x)\lambda(dv)$, where $g \in L_\infty(X)$ and $a > 0$.*

PROOF. The function $F(u) = f(T_u x) = \int_{B_u(a)} g(T_v x)\lambda(dv)$ is a bounded and continuous function of $u$ for all $x \in X$. Thus it is enough to show that if $F\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ is a bounded and continuous function then

$$\int_{\mathbb{R}^n} F(v_0 + v)\varphi_k(v)\lambda(dv)$$

converges to $F(v_0)$ for each $v_0 \in \mathbb{R}^n$. We may assume $F(v_0) = 0$, without loss of generality. Thus given $\varepsilon > 0$, choose $\eta > 0$ such that $|F(u)| < \varepsilon$ whenever $|u - v_0| < \eta$. By Lemma 6.1, we may assume $\varphi$ has compact support, so choose $B$ such that $S_\varphi \subset B$. Then by Lemma 2.4, we may choose a $K$ such that for all $k \geq K$, $\tau_k^{-1}B \subset B_0(\eta)$, so that $S_{\varphi_k} \subset B_0(\eta)$. Then, for $k \geq K$,

$$\begin{aligned}\left|\int_{\mathbb{R}^n} F(v_0 + v)\varphi_k(v)\lambda(dv)\right| &= \left|\int_{B_0(\eta)} F(v_0 + v)\varphi_k(v)\lambda(dv)\right| \\ &\leq \int_{B_0(\eta)} |F(v_0 + v)|\varphi_k(v)\lambda(dv) \\ &< \varepsilon.\end{aligned}$$

∎

THEOREM 6.4. *Suppose $\{\tau_k\}$ satisfies regularity conditions $(R_1)$–$(R_5)$ and $(R_6)$. Let $\varphi$ be any weight function. Then $\mathcal{A}^{\varphi_k} f$ converges $\mu$-a.e. for every $f \in L_\infty(X)$.*

PROOF. Again by Lemma 6.1, we may assume without loss of generality that $\varphi$ has compact support.

Let $f$ be a bounded function with $M = \|f\|_\infty$. Then the family of functions $f_\varepsilon(x) = 1/\lambda\big(B_0(\varepsilon)\big)\int_{B_0(\varepsilon)} f(T_v x)\lambda(dv)$ converge to $f$ in $L_1(X)$ as $\varepsilon$ approaches $0^+$. Also, this family is bounded by the same bound $M$, and by the previous lemma $\mathcal{A}^{\varphi_k} f_\varepsilon(x)$ converges for a.e. $x$ and any $\varepsilon > 0$. Thus by applying Lemma 6.2 with $\xi = 2M\mathbf{1}_{(0,\mu(X)]}$, $\mathcal{A}^{\varphi_k} f(x)$ must converge for a.e. $x$. ∎

*Global convergence.* In this part, we will be concerned with the convergence of the averages $\mathcal{A}^{\varphi_k} f$ when the sets $\tau_k^{-1}B$ become very large as $k \longrightarrow \infty$. To ensure convergence, we will impose the global regularity condition $(R_7)$. Specifically, we will prove the following.

THEOREM 6.5. *Suppose $\{\tau_k\}$ satisfies regularity conditions $(R_1)$–$(R_5)$ and $(R_7)$. Let $\varphi$ be any weight function. Then $\mathcal{A}^{\varphi_k}f(x)$ converges $\mu$-a.e. for every $f \in L_\infty(X)$.*

We will begin by proving the assertion for weight functions $\varphi$ of the form $\varphi(v) = 1/\lambda(B)\mathbf{1}_B(v)$, where $B$ is a ball in $\mathbb{R}^n$. For this type of weight function, we see

$$\mathcal{A}^{\varphi_k}f(x) = \frac{1}{\lambda(B)}\int_{\tau_k^{-1}B} f(T_v x)\delta_k(v)\lambda(dv)$$

Asymptotic convexity in $(R_7)$ says that we can approximate $\tau_k^{-1}B$ by compact, convex sets $K_k$. Combined with the asymptotic flatness in $(R_7)$ of $\delta_k$, we get the following approximation for $\mathcal{A}^{\varphi_k}f(x)$.

LEMMA 6.6. *Suppose $\tau_k$ satisfy regularity conditions $(R_1)$–$(R_2)$ and $(R_7)$. Let $\varphi$ be as above, let $f \in L_\infty^+(X)$, and define $\mathcal{A}^{K_k}f(x) = 1/\lambda(K_k)\int_{K_k} f(T_v x)\lambda(dv)$ where $K_k$ are the compact, convex sets from the asymptotic convexity in $(R_7)$. Then*

$$\lim_{k\to\infty}\left|\mathcal{A}^{\varphi_k}f(x) - \mathcal{A}^{K_k}f(x)\right| = 0 \quad \mu\text{-a.e. } x.$$

PROOF.

$$\begin{aligned}|\mathcal{A}^{\varphi_k}f(x) - \mathcal{A}^{K_k}f(x)| \le{}& \frac{\|f\|_\infty}{\lambda(B)}\int_{\tau_k^{-1}B}\left|\delta_k(v) - \frac{\lambda(B)}{\lambda(\tau_k^{-1}B)}\right|\lambda(dv).\\ &+ \|f\|_\infty\left|1 - \frac{\lambda(\tau_k^{-1}B)}{\lambda(K_k)}\right|\\ &+ \|f\|_\infty\frac{\lambda(K_k \triangle \tau_k^{-1}B)}{\lambda(K_k)}\end{aligned}$$

By the asymptotic flatness in $(R_7)$ of $\{\delta_k\}$, the first term on the right goes to zero as $k \to \infty$, and by asymptotic convexity, each of the last two terms go to zero as $k \to \infty$. ∎

Thus we need only consider the averages $\mathcal{A}^{K_k}f(x)$. We need some lemmas concerning compact, convex sets. Let $\mathcal{K}$ be the set of all compact, convex $K \subset \mathbb{R}^n$ with non-empty interior. For $K \in \mathcal{K}$, recall

$$\mathcal{R}(K) = \inf\{r \mid \exists v, K \subset B_v(r)\}.$$

We can make a similar definition regarding the radii of balls contained within $K$, namely

$$\varrho(K) = \sup\{r \mid \exists v, B_v(r) \subset K\}.$$

Corresponding to the notion of super-regularity, we have the notion of sub-regularity. We will say a sequence of sets $\{K_k\} \subset \mathcal{K}$ is *sub-regular* if there is a $\gamma > 1$ such that $\lambda(K_k) < \gamma\big(\varrho(K_k)\big)^n$ for all $k \in \mathbb{N}$. To establish a connection between super-regularity and sub-regularity, we need the following lemma.

LEMMA 6.7. *There is a constant $C$ such that, for all $K \in \mathcal{K}$,*

$$\lambda(K) \leq C\big(\mathcal{R}(K)\big)^{n-1}\varrho(K).$$

PROOF. We need only show that this holds for the $\ell_2$-norm, as the general case follows easily from this. Let $P_1$ and $P_2$ be two parallel linear hyperplanes of dimension $n-1$ which "sandwich" $K$, in the sense that $K$ lies between $P_1$ and $P_2$. If $d(P_1, P_2)$ is the distance between any two such planes, we define the width of $K$, $\omega(K)$, to be

$$\omega(K) = \inf\{d(P_1, P_2) \mid P_1 \text{ and } P_2 \text{ sandwich } K, \text{ and } P_1 \| P_2\}$$

If $K$ is contained within a ball $B$ of radius $R$, then it is also contained within the intersection of $B$ and the space between any two planes which sandwich $K$. Thus if $C_1$ is the Lebesgue measure of a solid unit sphere in $\mathbb{R}^{n-1}$, then we have $\lambda(K) \leq C_1 R^{n-1}\omega(K)$, and so $\lambda(K) \leq C_1\big(\mathcal{R}(K)\big)^{n-1}\omega(K)$. It is known that $\omega(K) \leq 2\sqrt{n+1}\varrho(K)$ (see, *e.g.*, [11], p. 112, Theorem 50), and thus $\lambda(K) \leq 2C_1\sqrt{n+1}\big(\mathcal{R}(K)\big)^{n-1}\varrho(K)$. ■

The following complementary result is of interest, although we do not need it.

LEMMA 6.8. *There is a constant $c$ such that, for all $K \in \mathcal{K}$,*

$$c\big(\varrho(K)\big)^{n-1}\mathcal{R}(K) \leq \lambda(K).$$

COROLLARY 6.9. *Given any $\Gamma > 1$, there exists $\gamma > 1$ such that the following is true. For any $K \in \mathcal{K}$, if $\mathcal{R}^n(K) \leq \Gamma\lambda(K)$, then $\lambda(K) \leq \gamma\varrho^n(K)$. We may take $\gamma = (C\Gamma)^n\lambda(I)$, where $I$ is the closed unit ball and $C$ is the constant of Lemma 6.7. Similarly, given any $\gamma' > 1$, there exists a $\Gamma' > 1$ such that the following is true. For any $K \in \mathcal{K}$, if $\lambda(K) \leq \gamma'\varrho^n(K)$, then $\mathcal{R}^n(K) \leq \Gamma'\lambda(K)$. We may take $\Gamma' = (\gamma'/c)^n/\lambda(I)$, where $I$ is the closed unit ball and $c$ is the constant of Lemma 6.8. Thus a sequence of convex sets is super-regular if and only if it is sub-regular.*

REMARK. In the literature, the condition $\varrho(A_k) \longrightarrow \infty$ is often imposed on a sequence of compact, convex sets, in combination with some type of regularity condition. It seems more natural to require $\lambda(A_k) \longrightarrow \infty$, which the previous corollary allows us to do in certain cases, as for example in Theorem 6.10 below.

For any $K \in \mathcal{K}$ and $h > 0$, define the "$h$-frame of $K$" to be

$$\partial_h K = \{v \in \mathbb{R}^n \mid d(v, \partial K) \leq h\}$$

where $\partial K$ is the boundary of $K$. Let $F = [0, 1]^n$, a closed unit cube, and define

$$\begin{gathered} {}_{\mathbb{Z}}K = \{u \mid u \in \mathbb{Z}^n, F + u \subset K\} \\ \underset{\sim}{K} = \bigcup\{F + u \mid u \in {}_{\mathbb{Z}}K\} \\ \tilde{K} = \bigcup\{F + u \mid u \in \mathbb{Z}, (F + u) \cap K \neq \emptyset\} \end{gathered}$$

The following facts are known (see, *e.g.*, [18]). As $\varrho(K) \longrightarrow \infty$, we obtain

1) $\lambda(\underset{\sim}{K})/\lambda(K) \longrightarrow 1$;
2) $\lambda(\tilde{K} \setminus \underset{\sim}{K})/\lambda(K) \longrightarrow 0$;
3) $\lambda(\partial_h K)/\lambda(K) \longrightarrow 0$;
4) $\#(K \cap \mathbb{Z}^n)/\lambda(K) \longrightarrow 1$.

We combine these facts with the previous corollary to get the following.

THEOREM 6.10. *Let $\{K_k\}$ be any super-regular sequence of sets in $\mathcal{K}$, and let $h > 0$. If $\lim_{k\to\infty} \lambda(K_k) = \infty$, then*

*i)* $\lim_{k\to\infty} \lambda(\tilde{K}_k)/\lambda(K_k) = 1$;
*ii)* $\lim_{k\to\infty} \lambda(\tilde{K}_k \setminus \underset{\sim}{K}_k)/\lambda(K_k) = 0$;
*iii)* $\lim_{k\to\infty} \lambda(\partial_h K_k)/\lambda(K_k) = 0$;
*iv)* $\lim_{k\to\infty} \#({}_{\mathbb{Z}}K_k)/\lambda(K_k) = 1$.

LEMMA 6.11. *Suppose $\{\tau_k\}$ satisfies regularity conditions ($R_1$)–($R_5$) and ($R_7$). Let $B$ be any ball, and let $\varphi = 1/\lambda(B)\mathbf{1}_B$. Then*

*i) If $f \in L_\infty(X)$ satisfies $f(T_u x) = f(x)$ for all $u \in \mathbb{Z}^n$, then $\mathcal{A}^{\varphi_k} f(x)$ converges a.e. to*

$$\int_{[0,1]^n} f(T_v x)\lambda(dv).$$

*ii) If $f$ is of the form $f(x) = g(x) - g(T_{e_i}x)$, where $g \in L_\infty(X)$ and $e_i$ is the unit basis vector in the i-th co-ordinate direction, then $\mathcal{A}^{\varphi_k} f(x)$ converges a.e. to 0.*

PROOF. $\{\tau_k^{-1}B\}$ is super-regular, so choose $\{B_k\}$ and $\Gamma$ such that $\tau_k^{-1}B \subset B_k$ and $\lambda(B_k) < \Gamma\lambda(\tau_k^{-1}B)$. $\{\tau_k^{-1}B\}$ is asymptotically convex, so choose $\{K_k\} \subset \mathcal{K}$ such that $\lambda(K_k \bigtriangleup \tau_k^{-1}B)/\lambda(\tau_k^{-1}B) \longrightarrow 0$. We may assume that $K_k \subset \overline{B_k}$ because each member of $\{K_k \cap \overline{B_k}\}$ is in $\mathcal{K}$ and this sequence still satisfies all the conditions of asymptotic convexity. With this assumption, $\{K_k\}$ must also be super-regular. Furthermore, ($R_7$) implies that $\lim_{k\to\infty} \lambda(\tau_k^{-1}B) = \infty$, and so the same holds for $\{K_k\}$. By Lemma 6.6, we need only show that $\mathcal{A}^{K_k} f(x)$ converges to the given values in (i) and (ii). All integrals are with respect to $\lambda(dv)$.

(i)

$$\begin{aligned}
\left|\int_F f(T_v x) - \mathcal{A}^{K_k} f(x)\right| &\le \left|\int_F f(T_v x) - \frac{1}{\lambda(K_k)}\int_{\underset{\sim}{K}_k} f(T_v x)\right| + \frac{1}{\lambda(K_k)}\int_{K_k\setminus \underset{\sim}{K}_k} |f(T_v x)| \\
&\le \left|\int_F f(T_v x) - \frac{1}{\lambda(K_k)}\sum_{u\in {}_{\mathbb{Z}}K_k}\int_{F+u} f(T_v x)\right| + \frac{\lambda(K_k \setminus \underset{\sim}{K}_k)}{\lambda(K_k)}\|f\|_\infty \\
&\le \left|1 - \frac{\#({}_{\mathbb{Z}}K_k)}{\lambda(K_k)}\right| \cdot \left|\int_F f(T_v x)\right| + \frac{\lambda(\tilde{K}_k \setminus \underset{\sim}{K}_k)}{\lambda(K_k)}\|f\|_\infty.
\end{aligned}$$

By the previous theorem, both of the last terms go to zero as $k \longrightarrow \infty$.

(ii)

$$\begin{aligned}
|\mathcal{A}^{K_k}f(x)| &= \left|\frac{1}{\lambda(K_k)}\int_{K_k} g(T_v x) - g(T_{v+e_i})\right| \\
&= \frac{1}{\lambda(K_k)}\left|\int_{K_k\setminus K_k+e_i} g(T_v x) - \int_{K_k+e_i\setminus K_k} g(T_v x)\right| \\
&\leq \frac{1}{\lambda(K_k)}\int_{K_k\triangle(K_k+e_i)} |g(T_v x)| \\
&\leq \|g\|_\infty \frac{\lambda\big(K_k \triangle (K_k+e_i)\big)}{\lambda(K_k)} \\
&\leq \|g\|_\infty \frac{\lambda(\partial_1 K_k)}{\lambda(K_k)}.
\end{aligned}$$

By the previous theorem, this goes to zero as $k \longrightarrow \infty$. ∎

PROOF OF THEOREM 6.5. To begin with, let $B$ be any ball, and let $\varphi$ be of the particular form $1/\lambda(B)\mathbf{1}_B$. If $f \in L_\infty(X)$, then we can approximate $f$ in the $L_1(X)$ norm by functions of the form

$$f(x) \approx h(x) + \sum_{i=1}^{n} g(x) - g(T_{e_i}x)$$

where $h$ and $g$ are bounded in $L_\infty(X)$ by $\|f\|_\infty$, and $h$ satisfies $h(x) = h(T_u x)$ for all $u \in \mathbb{Z}^n$ (see, *e.g.*, [14], p. 205). Then, by Lemma 6.11 and Lemma 6.2, $\mathcal{A}^{\varphi_k}f(x)$ must converge for $\mu$-a.e. $x$. If $\varphi$ and $\varphi'$ are any two weight functions and $a, a' > 0$, then $\mathcal{A}^{(a\varphi+a'\varphi')_k} = a\mathcal{A}^{\varphi_k} + a'\mathcal{A}^{\varphi'_k}$. Thus $\mathcal{A}^{\varphi_k}f(x)$ converges a.e. for any positive linear combination of weight functions of the particular form above. Since any weight function can be approximated in the $L_1$ norm by step functions of this type, applying Lemma 6.1 yields the required result. ∎

*Extensions.* We begin by noting that Theorem 1.2 is merely the combination of Theorems 6.4 and 6.5.

PROOF OF THEOREM 1.3. We may assume $f > 0$, and the hypothesis implies that $f$ is integrable. Define $f_m(x) = \min\big(f(x), m\big)$, so $f_m$ approaches $f$ in the $L_1(X)$ norm. By Theorem 1.2 we know that $\mathcal{A}^{\varphi_k}f_m$ converges a.e., and $(f - f_m)^* < f^*$. Then Lemma 6.2 implies that $\mathcal{A}^{\varphi_k}f(x)$ converges a.e. on $X$. ∎

THEOREM 6.12. *Suppose $\{\tau_k\}$ satisfy ($R_1$)–($R_5$), and either ($R_6$) or ($R_7$). Let $\{\theta_k\}$ be a sequence of compact weight function which are dominated by an $L_1(\mathbb{R}^n)$ function $\Theta$ with bounded support. Suppose this sequence converges $\lambda$-a.e. to a function $\varphi$, which is thus necessarily a compact weight function. Let*

$$\psi_k = \delta_k \cdot \theta_k \circ \tau_k$$

*and as before*

$$\varphi_k = \delta_k \cdot \varphi \circ \tau_k.$$

https://doi.org/10.4153/CJM-1995-044-0 Published online by Cambridge University Press

*If $f: X \longrightarrow \mathbb{R}$ satisfies*

$$\int_{\mathbb{R}} |f|^* \Theta^* \, d\ell < \infty$$

*then*

$$F_k(x) = |\mathcal{A}^{\varphi_k} f(x) - \mathcal{A}^{\psi_k} f(x)|$$

*converges to zero $\mu$-a.e. on $X$.*

PROOF. The result will be obtained by applying our maximal theorem to the $L_1(\mathbb{R}^n)$ function

$$\Phi_m(v) = \sup_{i \geq m} |\varphi(v) - \theta_i(v)|.$$

To apply Theorem 5.1 with $\Phi_m$ taking the role of $\varphi$, we define

$$\Phi_{m,k}(v) = \delta_k(v) \cdot \Phi_m(\tau_k v)$$

in analogy with the definition of $\varphi_k$. Now, for $k \geq m$,

$$\begin{aligned} F_k(x) &= |\mathcal{A}^{\varphi_k} f(x) - \mathcal{A}^{\psi_k} f(x)| \\ &\leq \int_{\mathbb{R}^n} |f|(T_v x) |\varphi_k - \psi_k|(v) \lambda(dv) \\ &\leq \int_{\mathbb{R}^n} |f|(T_v x) \Phi_{m,k}(v) \lambda(dv) \\ &= \mathcal{A}^{\Phi_{m,k}} |f|(x) \end{aligned}$$

and thus the maximal inequality in Theorem 5.1 says that, for any $\alpha > 0$, we have

$$(1) \qquad \mu\Big(\{x \mid \sup_{k \geq m} F_k(x) > \alpha\}\Big) \leq \mu\Big(\{x \mid \sup_{k \geq m} \mathcal{A}^{\Phi_{m,k}} |f|(x) > \alpha\}\Big) \leq \frac{C_{\Phi_m}}{\alpha} \int_{\mathbb{R}} |f|^* \Phi_m^* \, d\ell.$$

Since the supports of $\Phi_m$ are all contained within the support of $\Theta$, the constants $C_{\Phi_m}$ can be taken independent of $m$. $\Phi_m$ is dominated by $\Theta$ and converges to zero pointwise, so it also converges to zero in $L_1(\mathbb{R}^n)$. Thus $\Phi_m^*$ is dominated by $\Theta^*$, so by hypothesis this last integral is dominated by

$$\int_{\mathbb{R}} |f|^* \Theta^* \, d\ell < \infty,$$

and $\Phi_m^*$ converges to zero pointwise. Thus, by Lebesgue's Dominated Convergence Theorem, (1) converges to zero as $m$ goes to infinity, which completes our proof. ∎

Theorem 1.4 follows immediately from this.

ACKNOWLEDGMENTS. The problem of examining moving weighted averages in $\mathbb{R}^n$ was suggested to me by Mustafa Akçoğlu. I would like to express my thanks for the many useful and enjoyable discussions I had with him. Thanks also to Y. Déniel for further discussions and useful advice.

## REFERENCES


**1.** M. A. Akçoğlu, and A. Del Junco, *Convergence of Averages of Point Transformations*, Proc. Amer. Math. Soc. **49**(1975), 265–266.

**2.** M. A. Akçoğlu, A. Del Junco, and W. M. F. Lee, *A Solution to a Problem of A. Bellow*, Collection: Almost everywhere convergence, II, Evanston, Illinois, 1989, 1–7, Academic Press, Boston, Massachusetts, 1991.

**3.** M. A. Akçoğlu, and Y. Déniel, *Moving Weighted Averages*, Canad. J. Math. **45**(1993), 449–469.

**4.** A. Bellow, R. Jones, and J. Rosenblatt, *Convergence for Moving Averages*, Ergodic Theory Dynamical Systems **10**(1990), 43–62.

**5.** ———, *Harmonic Analysis and Ergodic Theory*, Collection: Almost everywhere convergence, Columbus, Ohio, 1988, 73–98, Academic Press, Boston, Massachusetts, 1989.

**6.** ———, *Almost Everywhere Convergence of Powers*, Collection: Almost everywhere convergence, Columbus, Ohio, 1988, 99–120, Academic Press, Boston, Massachusetts, 1989.

**7.** M. Broise, Y. Déniel and Y. Derriennic, *Réarrangement, Inégalités Maximales et Théorèmes Ergodiques Fractionnaires*, Ann. Inst. Fourier (Grenoble) **39**(1989), 689–714.

**8.** ———, *Maximal Inequalities and Ergodic Theorems for Cesàro-$\alpha$ or Weighted Averages*, Collection: Almost everywhere convergence, II, Evanston, Illinois, 1989, 93–107, Academic Press, Boston, Massachusetts, 1991.

**9.** A. P. Calderón, *Ergodic theory and translation-invariant operators*, Proc. Nat. Acad. Sci. **59**(1968), 349–353.

**10.** R. Coifman, and G. Weiss, *Analyse Harmonique Non-commutative sur Certains Espaces Homogènes*, Lecture Notes in Math. **242**, Springer-Verlag, 1971.

**11.** H. G. Eggleston, *Convexity*, Cambridge Tracts in Math. and Mathematical Physics, **47**, Cambridge Univ. Press, Cambridge, 1958.

**12.** R. L. Jones and J. Olsen, *Multi-Parameter Moving Averages*, Collection: Almost everywhere convergence, II, Evanston, Illinois, 1989, 127–149, Academic Press, Boston, Massachusetts, 1991.

**13.** ———, *Multiparameter Weighted Ergodic Theorems*, preprint.

**14.** U. Krengel, *Ergodic Theorems*, de Gruyter Studies in Math. **6**, W. de Gruyter, 1985.

**15.** D. A. Lind, *Locally Compact Measure Preserving Flows*, Adv. in Math. **15**(1975), 175–193.

**16.** D. I. McIntosh, *Ergodic Averages For Weight Functions Moved by Non-Linear Transformations on $\mathbb{R}^n$*, Ph.D. Thesis, University of Toronto, 1993.

**17.** A. Nagel, and E. M. Stein, *On Certain Maximal Functions and Approach Regions*, Adv. in Math. **54**(1984), 83–106.

**18.** X. X. Nguyen and H. Zessin, *Ergodic Theorems for Spatial Processes*, Z. Wahrscheinlichkeitstheorie verw. Geb. **48**(1979), 133–158.

**19.** J. Rosenblatt, and M. Wierdl, *A New Maximal Inequality and its Applications*, Ergodic Theory Dynamical Systems **12**(1992), 509–558.

**20.** W. Rudin, *Real and Complex Analysis*, second edition, Tata McGraw-Hill Publishing, New Delhi, India, 1974.

**21.** J. Sueiro, *A note on maximal operators of Hardy-Littlewood type*, Math. Proc. Cambridge Philos. Soc. **102**(1987), 131–134.

**22.** A. A. Tempel'man, *Ergodic Theorems for General Dynamical Systems*, Trudy Moskov. Mat. Obshch. **26**(1972), Trans. Moscow Math. Soc. **26**(1972), 94–132.



*Department of Mathematics and Statistics*
*McGill University*
*Montreal, Quebec*
*H3A 2K6*