Almost everywhere convergence of weighted ergodic averages

Christopher Michael Wedrychowicz
University at Albany, State University of New York, cw4347@albany.edu

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Physical Sciences and Mathematics Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive. Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
Almost Everywhere Convergence
Of Weighted Ergodic Averages

by

Christopher M. Wedrychowicz

A Dissertation
Submitted to the University at Albany, State University of New York
in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy

College of Arts and Sciences
Department of Mathematics & Statistics

2009
ALMOST EVERYWHERE CONVERGENCE OF WEIGHTED ERGODIC AVERAGES

CHRISTOPHER M. WEDRYCHOWICZ

ABSTRACT. Let $(X, B, \lambda, T)$ be a dynamical system and $\text{Log}_{(n)} x$ be the $n$–times iterated logarithm. In the first half of this thesis we will prove that given $p > 0$, and $n \in \mathbb{N}$ there exists an increasing sequence of non negative integers $n_k$ and a function $f \in L\text{Log}_{(n)} L(X)$ such that $A_N f(x) = \frac{1}{N} \sum_{k=0}^{N-1} f(T^{n_k} x)$ fails to converge a.e. but $A_N g(x)$ converges a.e. for all $g \in L\text{Log}_{(n)}^q L(X)$ with $q > p$. In the second half of this thesis we extend a theorem of Bellow and Calderón, which states that the sequence of convolution powers $\mu_n f(x) = \sum_{k \in \mathbb{Z}} \mu^n(k) f(T^k x)$ converges a.e, when $\mu$ is a strictly aperiodic probability measure on $\mathbb{Z}$ such that the expectation $E(\mu) = 0$ and the second moment of $\mu$, $m_2(\mu) < \infty$. 
I have been a student at Albany since I was 19 years old, first as an Undergraduate then as a Graduate. Before I came to Albany I was unsure as to what I would study. It was in the introductory Math courses where I decided what I would like to do. While an Undergraduate I developed the basic skills I would need to pursue my Ph.D. The courses I took as a graduate exposed me to a large variety of Mathematics. There were many interesting possibilities and it was difficult to choose a topic for a dissertation. Therefore I would like to acknowledge all of the faculty members who have taught and exposed me to many interesting areas of Mathematics. I would like to single out Dr. Hildebrand. The one year probability sequence that he taught was very useful and interesting. Additionally he and Dr. Reinhold organized the PERTH seminar in which over a number of weeks Anna and I presented parts of our theses in complete detail. I am grateful to him for serving on both my oral, and dissertation committees, he has supplied valuable suggestions throughout the process.

I would like to thank the remaining members of my defense committee, Professors Stessin and O’Neil who provided useful suggestions during the review process. I also thank Professors Wilken and Thomas for sitting on my Oral committee.

About three years ago I began working with my adviser Dr. Reinhold. I knew very little of Ergodic Theory when we began. From the very beginning she kept things focused and made sure that Anna and I learned the most important topics connected to the work we would be doing. Most importantly, however, are the questions that she gave me for my dissertation. I was lucky to have been given problems that forced me to do my best work in order to obtain a solution. I would like to thank her for all of the time spent with me helping me learn the background material, and the useful suggestions regarding my thesis. I wish her and her daughter Fiona all the best.
I would like to thank staff members JoAnna Aveyard, Rose Bellanger, Stacy Newman, Joan Mainwaring, and Ellen Fisher for all of the help they have given me over the years. Of course I must give a special mention to my office mate of four years Tim Clark who graduated last year. Outside of my wife there is no one with whom I have spent so much time. We had many conversations covering all topics of life, mainly about sports and current events. I also learned a lot about the process of graduate studies from Tim especially with the dissertation.

I thank my parents Mary and Michael Wedrychowicz for being great parents and making sure that I received a good education. They supported me throughout my Undergraduate years and were encouraging during my graduate studies. There are so many things throughout the years that they have given me, certainly I would not have made it to this point without them. My in-laws Konstantinos and Herikleia Savvopoulos have been very supportive. I have spent every summer with them since Anna and I have been together.

The last word is for my wife Anna Savvopoulou. We started here together in 2003. We have been together almost the entire time. Now we graduate together, a married couple, and we start a new phase of our lives. Throughout this process she has been there for me in every way imaginable. Through all times, good and bad. No man could have a better wife. I love you so much Anna. παντα θα ειμαι μαζι σου για να σου δωσω αγαπη στα δυσκολα.
## CONTENTS

<table>
<thead>
<tr>
<th>Sections</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>1. Convergence and Divergence of averages along subsequences in certain Orlicz spaces</td>
<td>1</td>
</tr>
<tr>
<td>1.1. Preliminaries</td>
<td>1</td>
</tr>
<tr>
<td>1.2. Sawyer’s Theorem in Orlicz Spaces and Conze’s Principle</td>
<td>3</td>
</tr>
<tr>
<td>1.3. Perturbations of Block Sequences</td>
<td>10</td>
</tr>
<tr>
<td>1.4. Monotone functions and rotations of the circle</td>
<td>17</td>
</tr>
<tr>
<td>1.5. Examples of separating sequences</td>
<td>23</td>
</tr>
<tr>
<td>2. Almost everywhere convergence of convolution powers</td>
<td>28</td>
</tr>
<tr>
<td>2.1. Preliminaries</td>
<td>28</td>
</tr>
<tr>
<td>2.2. Main Results</td>
<td>33</td>
</tr>
<tr>
<td>References</td>
<td>43</td>
</tr>
</tbody>
</table>
1. Convergence and Divergence of Averages along Subsequences in Certain Orlicz Spaces

1.1. Preliminaries.

Definition 1.1.1. Let \((X, \mathcal{B}, \mu)\) be a measure space. Let \(T : X \to X\) be a one-to-one, onto map such that \(\mu(T^{-1}A) = \mu(A) \forall A \in \mathcal{B}\). Then \(T\) is called a measure preserving transformation and \((X, \mathcal{B}, \mu, T)\) is called a dynamical system.

Example 1.1. An example of central importance to this work is when \(X = [0, 1)\), \(\mu\) is Lebesgue measure, \(\mathcal{B}\) is the \(\sigma\)-algebra of Borel sets and \(T\) is defined by \(T(x) = x + \alpha \mod(1)\) where \(\alpha \in [0, 1)\). It is equivalent to realizing \([0, 1)\) as the unit circle and \(T\) as a rotation by \(2\pi\alpha\).

Definition 1.1.2. If the averages

\[
\frac{1}{N} \sum_{k=0}^{N-1} \mu(T^{-k}A \cap B) \xrightarrow{N \to \infty} \mu(A)\mu(B) \quad \forall A, B \in \mathcal{B},
\]

Then \(T\) is called ergodic.

Example 1.1.3. If \(\alpha\) is irrational then \(T\) is ergodic as defined in the previous example; if \(\alpha\) is rational then \(T\) is not ergodic, ([15]).

A theorem of fundamental importance in ergodic theory is Birkhoff’s Theorem, which is stated as follows,

Theorem 1.1.4 (Birkhoff). Let \((X, \mathcal{B}, \mu, T)\) be a dynamical system and \((X, \mathcal{B}, \mu)\) be a \(\sigma\)-finite measure space then,

\[
\frac{1}{N} \sum_{k=0}^{N-1} f(T^k x) \xrightarrow{N \to \infty} E(f|J)(x) \text{ a.e.}
\]

where \(J\) is the \(\sigma\)-algebra of invariant sets.

There have been many attempts to generalize Birkhoff’s Theorem. One in particular is connected to the topic of this thesis. Let \(\{n_k\}\) be an increasing sequence of
positive integers. One may ask the following question:

Do the averages,

$$\frac{1}{N} \sum_{k=0}^{N-1} f(T^{n_k}x)$$

converge a.e $\forall f$ in some subspace of $L_1$?

Much work has been done in this area. For example when $n_k = k^2$ Bourgain [6] has shown that the averages converge a.e. $\forall f \in L_p$ where $p > 1$. The problem of the case $p = 1$ remained open for some time. Recently it was shown that for every dynamical system there exists a function $f \in L_1$ such that the averages do not converge a.e.

In the first part of this thesis we will deal with convergence a.e of averages along subsequences in Orlicz spaces and in the second part we will look at convolution powers.

First a few definitions,

**Definition 1.1.5.** Let $S$ be a subspace of $L^1$. An increasing sequence of integers $(n_k)$ is called $S$ universally good, if for any dynamical system $(X, \mathcal{B}, \mu, T)$ the averages

$$A_N f(x) = \frac{1}{N} \sum_{k=0}^{N-1} f(T^{n_k}x)$$

converge a.e. for $x \in X$, $\forall f \in S$.

A sequence is called $S$ universally bad if for every dynamical system there exists a function $f \in S$ such that the averages $A_N f$ fail to converge a.e.

When $S = L_p$ we may ask the following questions.

**Question 1.1.** Does there exist an increasing sequence of integers $(n_k)$ that is $L_p$ universally good while $L_q$ universally bad for all $q < p$?

**Question 1.2.** Does there exist a sequence that is $L_p$ universally bad but $L_q$ universally good for all $q > p$?
The first question was answered affirmatively by Reinhold [14] while the second was answered affirmatively by Bellow [1].

**Definition 1.1.6.** The space of functions $L^s \text{Log}_p^n L$ for $p > 0$ is defined as

$$L^s \text{Log}_p^n L = \{ f \in L_1 : \int |f(x)|^s \text{Log}_p^n(|f(x)| + M)dx < \infty \}$$

where $\text{Log}_p^n(x) = \log \log \cdots \log(x)$ is $n$ times iterated logarithm, and $M$ is a value such that $\text{Log}_p^n(M) > 0$.

The notions of universally good and bad extend to the above spaces in an obvious way.

Our main result will be the following.

**Theorem 1.1.7.** Given $p$ and a dynamical system $(X, \mathcal{B}, \mu, T)$, there exists an increasing sequence of integers $(n_k)$ such that the averages $A_N f(x)$ converge a.e. for all $f \in L^s \text{Log}_p^n L$ with $q > p$ while there exists a function $f \in L^s \text{Log}_p^n L$ such that the averages $A_N f(x)$ fail to converge a.e.

1.2. Sawyer’s Theorem in Orlicz Spaces and Conze’s Principle.

**Definition 1.2.1.** Given a sequence of operators $T_n : S \to \mathcal{M}(X)$, where $S \subseteq L_1(X)$ is a Banach space and $\mathcal{M}(X)$ denotes the set of measurable functions, the maximal operator $T^*$ of $\{T_n\}$ is defined by $T^* f(x) = \sup_n |T_n f(x)|$.

**Theorem 1.2.2 (Banach’s principle).** If $T^* f(x) < \infty$ a.e. for all $f \in S$ where $S$ is a Banach space of functions contained in $L_1$ then there is a positive, decreasing function $C(\lambda)$ defined for $\lambda > 0$ that goes to zero as $\lambda \to \infty$ such that for all $f \in S$ we have

$$\mu \{ x : T^* f(x) > \lambda \| f \|_S \} \leq C(\lambda).$$
**Theorem 1.2.3.** Let \((X, \mathcal{B}, \mu)\) be a probability space and \(S \subseteq L_1\) be a Banach space. If \(\{T_n\}\) is a sequence of bounded operators such that

\[
T^* f(x) = \sup_n T_n f(x) < \infty \quad \text{a.e.}
\]

for every \(f \in S\) then the set of functions in \(S\) such that \(T_n f(x)\) converges a.e. is closed.

When \(S = L^p\) one often establishes a weak maximal inequality for the sublinear operator \(T^*\) of the form

\[
\mu(\{x : T^*(x) \geq \lambda\}) \leq C \frac{\|f\|^p}{\lambda^p}
\]

where \(C\) is a constant independent of \(f\) and \(\lambda\).

**Definition 1.2.4.** Let \(\Phi(x)\) be a function such that

1. \(\Phi\) is continuous, convex, nondecreasing, nonzero and \(\Phi(0) = 0\).
2. \(\Phi \in \Delta_2\), i.e \(\Phi(2x) \leq K\Phi(x)\) for some \(K > 0\).

Let \(L_\Phi = \left\{ f \in L_1 : \int \Phi(|f(x)|)dx < \infty \right\}\) then \(L_\Phi\) is a Banach space under the following norm ([13])

\[
\|f\|_\Phi = \inf \left\{ k : \int \Phi \left( \frac{|f|}{k} \right) dx < 1 \right\}.
\]

The space \(L_\Phi\) is an example of an Orlicz space. If \(\Phi(x) = \Phi(-x)\) then \(\Phi\) is called a Young function. If \(\Phi(x)\) is a Young function and \(\frac{\Phi(x)}{x} \to 0\) as \(x \to 0\) and \(\frac{\Phi(x)}{x} \to \infty\) as \(x \to \infty\) then \(\Phi\) is called an \(N\)-function.

**Theorem 1.2.5.** Let \(\Phi\) be a Young function such that \(\frac{\Phi(x)}{x} \to \infty\), then the following are equivalent

1. \(\Phi \in \Delta_2\)
2. There exists \(M \geq 0, p \geq 1, c > 0\) such that for \(y \geq M\) and \(x \geq 0\)

\[
\Phi(xy) \leq c\Phi(x) (\Phi(y))^p.
\]
Proof. Suppose $\Phi \in \Delta_2$. Then $\Phi(2x) \leq K\Phi(x)$ for some $K = 2^p$ with $p \geq 2$. Choose $k$ so that $2^k \leq y < 2^{k+1}$. Therefore

$$\Phi(xy) \leq \Phi(2^{k+1}x) \leq (2^p)^{k+1}\Phi(x)$$

$$= 2^p(2^k)^p\Phi(x) \leq 2^{p+1}y^p\Phi(x)$$

$$\leq 2^{p+1}(\Phi(y))^p\Phi(x)$$

Assume $y$ is chosen large enough so that $\Phi(y) \geq y$.

Now suppose that (2) is true. Choose $k \geq 2$ so that $2^k \geq M$. Then by the convexity of $\Phi$ and (2),

$$\Phi(2x) = \Phi\left(\frac{2^k x}{2^{k-1}}\right) \leq c\Phi\left(\frac{x}{2^{k-1}}\right)(\Phi(2^k))^p$$

$$\leq \frac{c}{2^{k-1}}(\Phi(2^k))^p\Phi(x) = K\Phi(x),$$

where $K = \frac{c}{2^{k-1}}(\Phi(2^k))^p$. \qed

**Proposition 1.2.6.** Suppose that $\frac{1}{N} \sum_{k=0}^{N-1} f(T^{nk}x)$ diverges for $x$ in a set of positive measure and some $f \in L_\Phi(X)$. Since $L^1(X) = \bigcup_{\Phi \text{N-function}} L_\Phi(X)$ ([13]) and

$$\int \Phi(f) \, dx < \infty$$

there exists $\Psi$ such that $\Phi(f) \in L_\Psi$ and hence $\int \Psi(\Phi(f)) < \infty$ and $f \in L_{\Psi\circ\Phi}$. Therefore the sequence is bad in a space which is smaller than $L_\Phi$ in the sense that $\frac{(\Psi \circ \Phi)(x)}{\Phi(x)} \to \infty$ as $x \to \infty$.

The following is a modification of the well-known theorem of Sawyer. To the knowledge of the author no version of the Theorem exists for Orlicz spaces. The proof is a modification of an argument found in [9]. We first give the following definition.

**Definition 1.2.7.** A family of transformations $(S_\alpha)_{\alpha \in I}$ is a mixing family of mappings, if the following holds. If $A, B \in \mathcal{B}$ and $\rho > 1$, then there exists $S_\alpha$ such that

$$\lambda(A \cup S_\alpha^{-1}(B)) \leq \rho \lambda(A)\lambda(B).$$
Theorem 1.2.8 (Sawyer’s Theorem). Let \((X, \mathcal{B}, \mu)\) be a probability measure space. Let \(\{T_k\}\) be a sequence of positive linear operators from \(L_\Phi\) to the set of measurable functions on \(X\). Assume that the \(\{T_k\}\)’s commute with a family \(\{S_\alpha\}\) of measure preserving maps from \(X\) to \(X\) that mix the measurable sets of \(X\). Assume that \(\Phi \in \Delta_2\) and \(M, p, C\) are the constants from Theorem 1.2.5 such that for \(y \geq M\), \(\Phi(xy) \leq C\Phi(x)(\Phi(y))^p\).

Then the following are equivalent,

1. \(T^*\) satisfies an inequality of the form
   \[
   \mu\{x : T^*f(x) \geq \lambda\} \leq C \int \Phi\left(\frac{f}{\lambda}\right) dx.
   \]

2. For each \(f \in L_\Phi\), \(T^*f(x) < \infty\).

Proof. The following lemma is of central importance to the proof of the theorem.

Lemma 1.2.9. Let \((X, B, \mu)\) be a probability measure space. Let \(S_\alpha : X \to X\) be a collection of measure preserving maps that mix the measurable sets of \(X\). Then if \(\{A_k\}\) is a sequence of measurable sets of \(X\) such that \(\sum \mu(A_k) = \infty\), there exists a sequence \(\{S_k\} \subseteq (S_\alpha)\) such that almost every \(x \in X\) is in infinitely many of the sets \(S_k^{-1}(A_k)\).

Assume that \(T^*\) does not satisfy an inequality of the form

\[
\mu(\{x : T^*f(x) \geq \lambda\}) \leq C \int \Phi\left(\frac{f}{\lambda}\right) dx.
\]

Then fix a sequence \(c_k\) increasing to infinity, \(c_k > 0\). Then there exists a sequence \(\{f_k\} \subseteq L_\Phi\), \(\lambda_k > 0\) such that,

\[
\mu\{T^*f_k(x) \geq \lambda_k\} > c_k \int \Phi\left(\frac{f_k}{\lambda_k}\right) dx.
\]

Call \(g_k = \frac{f_k}{\lambda_k}\), \(A_k = \{T^*g_k \geq 1\}\). Then,
\[
1 \geq \mu(A_k) \geq c_k \int \Phi(g_k) \, dx.
\]

Let \( h_k \) be natural numbers such that \( 1 \leq h_k \mu(A_k) \leq 2 \) and take \( h_k \) copies of \( A_k \) denoted by \( A_k^1, \ldots, A_k^{h_k} \). Thus \( \sum_{k=1}^\infty \sum_{j=1}^{h_k} \mu(A_k^j) = \infty \), and by the previous lemma there are \( S_k^j \in (S_\alpha) \) such that almost every \( x \in X \) is in infinitely many of the sets \( (S_k^j)^{-1}(A_k^j) \).

Define a function

\[
F(x) = \sup_{1 \leq j \leq h_k} \alpha_k S_k^j g_k^j(x)
\]

where \( g_k^j = g_k \) and the constants \( \alpha_k \), to be determined later, and will be greater than or equal to \( M \).

Then

\[
\Phi(F(x)) \leq \sum_{k=0}^{h_k} \Phi(\alpha_k S_k^j g_k^j(x))
\]

\[
\leq \sum_{k=0}^{h_k} C[\Phi(\alpha_k)]^p \Phi(S_k^j g_k^j(x))
\]

\[
\leq C \sum_{k=1}^\infty [\Phi(\alpha_k)]^p \sum_{j=1}^{h_k} \Phi(S_k^j g_k^j(x)).
\]

And so

\[
\int \Phi(F(x)) \, dx \leq C \sum_{k=1}^\infty [\Phi(\alpha_k)]^p h_k \int \Phi(g_k) \, dx
\]

\[
\leq C \sum_{k=1}^\infty [\Phi(\alpha_k)]^p \frac{\mu(A_k)}{c_k} h_k
\]

\[
\leq C \sum_{k=1}^\infty \frac{[\Phi(\alpha_k)]^p}{c_k},
\]
by (1). Given that the sequence \( \left\{ \frac{1}{c_k} \right\} \) sums, the \( \{\alpha_k\} \) may be chosen so that the above sum is finite and the \( \alpha_k \) increase to infinity. The remainder of the argument is the same as in [9].

Let \( M = e^{e^{\cdot^e}} \) so that \( \log(n) M = 1 \), \( \Phi(x) = x \log^n(x + M) \). We will need the following,

**Lemma 1.2.10.** For \( x, y > M \), \( x < y \) we have
\[
\frac{\log(n) x}{\log(n) y} > \frac{\log(n-1) x}{\log(n-1) y} > \cdots > \frac{\log x}{\log y} > \frac{x}{y}
\]

**Proof.** Let \( f(x) = \frac{\log x}{x} \). Then \( f'(x) = \frac{1 - \log x}{x^2} < 0 \) when \( x > e \). Therefore \( \frac{\log x}{x} \geq \frac{\log y}{y} \) and \( \frac{\log x}{\log y} \geq \frac{x}{y} \). The lemma follows by iteration. \( \square \)

**Lemma 1.2.11.** Suppose \( M \) is such that \( \log(n) M = 1 \) then \( \log(k) M > 2^{n-k} \) for \( 1 \leq k < n \).

**Proof.** For \( k = n - 1 \) we have, \( \log(n-1) M = e^{\log(n) M} = e^1 = e > 2 \). Suppose the result holds for \( k + 1 \) then we have \( \log(k) M = e^{\log(k+1) M} > e^{2^{n-(k+1)}} > 2 \cdot 2^{n-(k+1)} = 2^{n-k} \). \( \square \)

**Lemma 1.2.12.** Let
\[
\Psi(x) = \frac{1}{x + M} \left( 1 + \frac{1}{\log(x + M)} + \frac{1}{\prod_{i=1}^{2} \log(i)(x + M)} + \cdots + \frac{1}{\prod_{i=1}^{n-1} \log(i)(x + M)} \right)
\]
then \( \lim_{x \to \infty} x \Psi(x) = 1 \) and \( x \Psi(x) < 2 - \frac{1}{2^{n-1}} \forall x \).

**Proof.** \( \lim_{x \to \infty} x \Psi(x) = 1 \) is immediate. Now \( \frac{x}{x + M} < 1 \) and by Lemma 1.2.11
\[
\Psi(x) \leq \frac{1}{x + M} \left( 1 + \frac{1}{2^{n-1}} + \frac{1}{2^{n-2}} + \cdots + \frac{1}{2} \right).
\]
\( \square \)
Lemma 1.2.13. Let \( f(x) = \log(n)(x + M) \). Then \( |xf'(x)| < \frac{1}{2^{n-1}} \).

Proof. \( f'(x) = \frac{1}{(x + M)\log(x + M) \cdots \log(n-1)(x + M)} \). So
\[
|xf'(x)| \leq \frac{x}{x + M} \cdot \frac{1}{\log(x + M)} < \frac{1}{2^{n-1}} \text{ by Lemma 1.2.11.}
\]

\[\Box\]

Theorem 1.2.14. \( \Phi(x) = x(\log(n)(x + M))^p \) is convex \( \forall p > 0 \) and \( x \geq 0 \).

Proof.
\[
\Phi'(x) = (\log(n)(x + M))^p + px(\log(n)(x + M))^{p-1}(\log(n)(x + M))' .
\]
\[
\Phi''(x) = 2p(\log(n)(x + M))^{p-1}(\log(n)(x + M))' + p(p-1)x(\log(n)(x + M))^{p-2} ((\log(n)(x + M))')^2 + px(\log(n)(x + M))^{p-1}(\log(n)(x + M))'' .
\]
Noting that \( (\log(n)(x + M))'' = -(\log(n)(x + M))'\Psi(x) \), where \( \Psi(x) \) is defined as in Lemma 1.2.12 we have
\[
\Psi''(x) \geq 0 \Leftrightarrow 2 \log(n)(x + M) + (p - 1)x(\log(n)(x + M))' - x\log(n)(x + M)\Psi(x) \geq 0
\]
\[
\Leftrightarrow \log(n)(x + M)(2 - x\Psi(x)) + (p - 1)x(\log(n)(x + M))' \geq 0 .
\]
If \( p \geq 1 \) this follows and if \( p < 1 \) we have by Lemmas 1.2.12 and 1.2.13 that the left hand side is greater than
\[
\frac{\log(n)(x + M)}{2^{n-1}} - \frac{(1-p)}{2^{n-1}} \geq \frac{p}{2^{n-1}} > 0 , \forall x \geq 0 .
\]
\[\Box\]

Corollary 1.2.15. \( \Phi(x) = |x|(\log(n)(|x| + M))^p \) is a Young function \( \forall p > 0 \).

Corollary 1.2.16. \( \Phi(x) = |x|(\log(n)(|x| + M))^p - |x| \) is an \( \mathcal{N} \)–function \( \forall p > 0 \).

Theorem 1.2.17. \( \Phi(x) = |x|(\log(n)(|x| + M))^p \in \Delta_2 \forall p \geq 0 \).
Proof. By Lemma 1.2.10 there exists $K > 0$ such that

$$\frac{2}{K} \leq \left( \frac{|x| + M}{2|x| + M} \right)^p \leq \left( \frac{\log(n)(|x| + M)}{\log(n)(2|x| + M)} \right)^p.$$ 

Therefore $2 \left( \log(n)(2|x| + M) \right)^p \leq K \left( \log(n)(|x| + M) \right)^p$. The theorem follows after multiplying both sides by $|x|$.

If we fix a sequence $(n_k)$, to each dynamical system, we may associate a constant $C(n_k)$ such that

$$\mu\{x : T^*f > \lambda\} \leq C(n_k) \int \Phi \left( \frac{f}{\lambda} \right) \, dx.$$ 

We may then consider the minimal constant so that the inequality holds in all dynamical systems. The so-called Conze’s Principle [7] asserts a condition under which we may conclude that such a minimal constant exists and is finite.

**Theorem 1.2.18 (Conze’s Principle).** For a given sequence $(n_k)$ to have its associated minimal constant finite, it is enough that there exists a single ergodic dynamical system $(X, \mathcal{B}, \mu, T)$ such that the averages

$$\frac{1}{N} \sum_{k=0}^{N-1} f(T^{nk}x)$$

converge a.e. $\forall f \in L_\Phi$.

The main application of this theorem will be the following. Suppose that $T$ is a rotation of the circle and there exists a function $L_\Phi$ and a sequence of integers $(n_k)$ such that $(T_N f)(x) = \frac{1}{N} \sum_{k=0}^{N-1} f(T^{nk}x)$ is unbounded for all $x$ in a set of positive measure. By Sawyer’s Theorem there is no maximal inequality. Conze’s Principle then implies that $(T_N f)(x)$ can not converge a.e. $\forall f \in L_\Phi$ in all dynamical systems $(X, \mathcal{B}, \mu, T)$.

1.3. Perturbations of Block Sequences. The desired sequence will be a perturbation of a block sequence. A block of integers is a set of the form $B = [n, n + 1, \cdots, n + k - 1]$ of consecutive integers. We will let $|B| = k$ denote the number of integers in $B$ and will refer to it as the length of $B$. A block sequence is a sequence
\{n_k\} that can be arranged into blocks \(B_1, B_2, \ldots\). As a set \(\{n_k\} = \bigcup_{k=1}^{\infty} B_k\). Let \(D_k\) be an arbitrary collection of integers between \(B_k\) and \(B_{k+1}\). The collection \(\bigcup_{k=1}^{\infty} D_k\) will be referred to as a perturbation of the block sequence \(\bigcup_{k=1}^{\infty} B_k\) and the sequence whose elements are \(\bigcup_{k=1}^{\infty} B_k \cup D_k\) will be referred to as a perturbed block sequence. Theorem 1.3.2 is a generalization of a theorem of Bellow [1]. It essentially states that if we begin with a block sequence, which is universally good in a certain subspace of \(L_1\) there is a certain degree to which we may perturb it so that the resulting sequence is also universally good in that subspace. We will need a theorem from [14]. Notationwise, \(l_k = |B_k|\) and \(d_k = |D_k|\).

**Theorem 1.3.1** (Reinhold). Let \(B_k\) and \(D_k\) be a block sequence and a perturbation of that block sequence. If the sequence \(\bigcup_{k=1}^{\infty} B_k\) is universally good for \(L_\infty\) and

\[
\sup_{k} \frac{d_1 + \cdots + d_k}{l_1 + \cdots + l_k} < \infty
\]

then the sequence \(\bigcup_{k=1}^{\infty} B_k \cup D_k\) is also universally good for \(L_\infty\).

**Theorem 1.3.2.** Let \(\bigcup_{k=1}^{\infty} B_k\) be a block sequence that is universally good in the Orlicz space \(L_\Phi\), and let \(\bigcup_{k=1}^{\infty} D_k\) be as above. If

\[
\sum_{k=1}^{\infty} \frac{1}{\Phi \left( \frac{l_1 + \cdots + l_k}{d_1 + \cdots + d_k} \right)} < \infty
\]

then the sequence \(\bigcup_{k=1}^{\infty} B_k \cup D_k\) is also universally good in \(L_\Phi\).
Proof. We proceed as in [14].

Let

\[ C = \bigcup B_k \cup D_k, \]

\[ b_n = \left| \left( \bigcup_{k=1}^{\infty} B_k \right) \cap [0, n] \right| \text{ and } \]

\[ c_n = \left| \left( \bigcup_{k=1}^{\infty} D_k \right) \cap [0, n] \right|. \]

The averages \( A_n f(x) = \frac{1}{|C \cap [0, n]|} \sum_{u \in C \cap [0, n]} f(T^u x) \) can be written as the convex combination

\[ A_n f(x) = \frac{b_n}{b_n + c_n} \left( \frac{1}{b_n} \sum_{u \in \bigcup_{k=1}^{\infty} B_k \cap [0, n]} f(T^u x) \right) + \frac{c_n}{b_n + c_n} \left( \frac{1}{c_n} \sum_{u \in \bigcup_{k=1}^{\infty} D_k \cap [0, n]} f(T^u x) \right) \]

\[ = \frac{b_n}{b_n + c_n} A_B^f(x) + \frac{c_n}{b_n + c_n} A_D^f(x). \]

To establish a.e. convergence it is enough to do so on each piece separately.

First we observe that since

\[ \Phi \left( \frac{l_1 + \cdots + l_k}{d_1 + \cdots + d_k} \right) \to 0, \]

we have

\[ \Phi \left( \frac{l_1 + \cdots + l_k}{d_1 + \cdots + d_k} \right) \to \infty \]

so

\[ \frac{l_1 + \cdots + l_k}{d_1 + \cdots + d_k} \to \infty \]

and hence its reciprocal goes to 0.

This implies by the previously stated theorem that the averages of functions in \( L_\infty \) converge a.e. and thus there is convergence on a dense set.
We have

\[
\begin{array}{l}
\frac{c_n}{b_n} = \\
\quad \begin{cases} \\
\quad \frac{d_1 + \cdots + d_{k-1}}{l_1 + \cdots + l_{k-1} + s_k} & \text{if } k \text{ is the smallest integer such that } B_k \text{ is not contained in } [0, n] \text{ and } B_k \cap [0, n] \neq \emptyset, \\
\quad \frac{d_1 + \cdots + d_{k-2} + r_{k-1}}{l_1 + \cdots + l_{k-1}} & \text{if } k \text{ is the smallest integer such that } B_k \text{ is not contained in } [0, n], B_k \cap [0, n] = \emptyset \text{ and } B_{k-1} \subset [0, n] \\
\end{cases}
\end{array}
\]

where \(0 \leq r_{k-1} \leq d_{k-1}\) and \(0 \leq s_k \leq l_k\).

In either case

\[
\frac{c_n}{b_n} \leq \frac{d_1 + \cdots + d_{k-1}}{l_1 + \cdots + l_{k-1}} \to 0,
\]

and so

\[
\frac{b_n}{b_n + c_n} \to 1.
\]

Therefore

\[
\frac{b_n}{b_n + c_n}A^B_nf(x)
\]

converges a.e. since \(\bigcup_{k=1}^\infty B_k\) is universally good in \(L_\Phi\). We note this implies that

\[
\frac{c_n}{b_n + c_n}A^D_nf(x)\]

converges a.e \(\forall f \in L^\infty\).

Consider the following operator:

\[
\sup_n \frac{c_n}{b_n + c_n}A^D_nf(x) = D^*f(x)
\]
Let
\[ A^\lambda = \{ x : D^* f(x) \geq \lambda N \} \]

where \( N \) will in this instance denote the Orlicz norm of \( f \). See 1.2.4

\[ \frac{c_n}{b_n + c_n} |A_n f(x)| \leq \frac{1}{b_n + c_n} \sum_{u \in \bigcup_{i=1}^{k-1} D_i \cap [0,n]} |f(T^u x)| \]
\[ \leq \frac{1}{b_n + c_n} \sum_{u \in \bigcup_{i=1}^{k-1} D_i} |f(T^u x)| \]
\[ \leq \sum_{u \in \bigcup_{i=1}^{k-1} D_i} \frac{1}{l_1 + \cdots + l_{k-1}} |f(T^u x)| \]
\[ = R_{k-1} f(x) . \]

Let \( T^* f(x) = \sup_k R_k f(x) \) and \( A_k^\lambda = \{ x : R_k f(x) \geq \lambda \} \), therefore

\[ \mu(A^\lambda) \leq \sum_{k=1}^{\infty} \mu(A_k^\lambda) \]

Now if,

\[ I = \int \{ R_k f(x) \geq \lambda N \} = A_k^\lambda \] \[ \phi \left( \frac{1}{N_\phi(d_1 + \cdots + d_k)} \sum_{u \in \bigcup_{i=1}^{k} D_i} |f(T^u x)| \right) dx \]
\[ = \int \{ \sum_{u \in \bigcup_{i=1}^{k} D_i} T^u f(x) \geq \lambda N_\phi(l_1 + \cdots + l_k) \} \]
\[ \phi \left( \frac{1}{N_\phi(d_1 + \cdots + d_k)} \sum_{u \in \bigcup_{i=1}^{k} D_i} |f(T^u x)| \right) dx \]
\[ \geq \mu(A_k^\lambda) \phi \left( \frac{\lambda (l_1 + \cdots + l_k)}{(d_1 + \cdots + d_k)} \right) . \]
Then,

\[
\mu(A_k^\lambda) \Phi\left( \frac{\lambda(l_1 + \cdots + l_k)}{d_1 + \cdots + d_k} \right) \leq I \leq 1
\]

since

\[
\left\| \sum_{u \in \bigcup_{i=1}^k D_i} T^u f(x) \right\|_\Phi \leq N_\Phi(d_1 + \cdots + d_k).
\]

Therefore we have the inequality

\[
\mu(A_k^\lambda) \leq \frac{1}{\Phi\left( \frac{\lambda(l_1 + \cdots + l_k)}{d_1 + \cdots + d_k} \right)}.
\]

For large enough \( \lambda \) we have,

\[
\frac{1}{\Phi\left( \frac{\lambda(l_1 + \cdots + l_k)}{d_1 + \cdots + d_k} \right)} \leq \frac{1}{\Phi\left(\frac{l_1 + \cdots + l_k}{d_1 + \cdots + d_k}\right)}.
\]

Also,

\[
\frac{1}{\Phi\left( \frac{\lambda(l_1 + \cdots + l_k)}{d_1 + \cdots + d_k} \right)} \rightarrow 0
\]

monotonically as \( \lambda \rightarrow \infty \) for every \( k \).

Therefore if we have

\[
\mu(A^\lambda) \leq \sum_{k=1}^\infty \mu(A_k^\lambda) = F(\lambda).
\]

Therefore by the Lebesgue dominated convergence theorem \( F(\lambda) \) is an eventually monotone decreasing function that goes to 0 as \( \lambda \rightarrow \infty \). Since the maximal operator
satisfies a weak-maximal inequality, and \( A_n^D f(x) \) converges a.e. \( \forall f \in L^\infty, A_n^D f(x) \) converges a.e \( \forall f \in L_\Phi. \)

\( \square \)

**Proposition 1.3.3.** Let \( B_k \) and \( D_k \) be as above. Let \( l_k = |B_k| \) and \( d_k = |D_k| \).

Suppose that \( \forall k \)

\[
  l_1 + \cdots + l_k \leq C l_{k+1}
\]

\[
d_k = c_k l_k
\]

are such that \( \sum_{k=1}^{\infty} \frac{1}{\Phi \left( \frac{l_{k+1}}{l_k} \right)} < \infty \) and \( \sum_{k=1}^{\infty} \frac{1}{\Phi \left( \frac{1}{c_k} \right)} < \infty \).

Then if \( \bigcup_{k=1}^{\infty} B_k \) is universally good in \( L_\Phi \) then \( \bigcup_{k=1}^{\infty} (B_k \cup D_k) \) is universally good in \( L_\Phi \).

**Proof.** Choose \( k_0 \) so that \( c_k \leq 1 \) for all \( k \geq k_0 \). Then

\[
  \frac{d_1 + \cdots + d_k}{l_1 + \cdots + l_k} \leq \frac{d_1 + \cdots + d_{k_0-1}}{l_1 + \cdots + l_k} + \frac{d_{k_0} + \cdots + d_{k-2}}{l_1 + \cdots + l_k} + \frac{d_{k-1} + d_{k}}{l_1 + \cdots + l_k}
\]

\[
  \leq \frac{C_0}{l_k} + \frac{C l_{k-1}}{l_k} + c_{k-1} + c_k
\]

Therefore,

\[
  \frac{d_1 + \cdots + d_k}{l_1 + \cdots + l_k} \leq \frac{C_0}{l_k} + \frac{C l_{k-1}}{l_k} + c_{k-1} + c_k, \quad \text{or}
\]

\[
  \frac{l_1 + \cdots + l_k}{d_1 + \cdots + d_k} \geq \frac{1}{\frac{C_0}{l_k} + \frac{C l_{k-1}}{l_k} + c_{k-1} + c_k}
\]

\[
  \geq \frac{1}{4 \max \left( \frac{C_0}{l_k}, \frac{C l_{k-1}}{l_k}, c_{k-1}, c_k \right)} = \frac{1}{4} \min \left( \frac{1}{\frac{C_0}{l_k}}, \frac{1}{\frac{C l_{k-1}}{l_k}}, \frac{1}{c_{k-1}}, \frac{1}{c_k} \right)
\]

Therefore,

\[
  \Phi \left( \frac{l_1 + \cdots + l_k}{d_1 + \cdots + d_k} \right) \geq \Phi \left( \frac{1}{4} \min(A_k, B_k, Ck, D_k) \right)
\]

or,
\[
\frac{1}{\Phi\left(\frac{l_1 + \cdots + l_k}{d_1 + \cdots + d_k}\right)} \leq \frac{1}{\Phi\left(\frac{1}{4}\min(A_k, B_k, C_k, D_k)\right)}
\]
\[
= \max\left(\frac{1}{\Phi\left(\frac{1}{4}A_k\right)}, \frac{1}{\Phi\left(\frac{1}{4}B_k\right)}, \frac{1}{\Phi\left(\frac{1}{4}C_k\right)}, \frac{1}{\Phi\left(\frac{1}{4}D_k\right)}\right)
\]

and
\[
\sum_{k=1}^{\infty} \frac{1}{\Phi\left(\frac{l_1 + \cdots + l_k}{d_1 + \cdots + d_k}\right)} < \sum_{k=1}^{\infty} \max\left(\frac{1}{\Phi\left(\frac{1}{4}A_k\right)}, \frac{1}{\Phi\left(\frac{1}{4}B_k\right)}, \frac{1}{\Phi\left(\frac{1}{4}C_k\right)}, \frac{1}{\Phi\left(\frac{1}{4}D_k\right)}\right) < \infty
\]

\[\Box\]

### 1.4. Monotone functions and rotations of the circle.

Suppose \( f \geq 0 \) is a monotone decreasing function on \((0, 1)\), \( \{n_k\} \) is an increasing sequence of integers and \( T(x) = x + \alpha \mod(1) \) for \( \alpha \in [0, 1) - \mathbb{Q} \). Let

\[ B_\lambda = \{ x : A_N = \frac{1}{N} \sum_{k=0}^{N-1} f(T^{n_k}x) \geq \lambda \}. \]

For each \( k \) where \( 0 \leq k \leq N \), let

\[ a_k = \sup\{ x : y = T^{-n_k}(x) \in B_\lambda \text{ and } T^{n_k}(x) \leq T^{n_p}(x) \quad \forall \quad 0 \leq p \leq N - 1 \}. \]

Intuitively this is the supremum of the \( x \) values such that there is a \( y \in B_\lambda \) with \( T^{n_k}(y) = x \) and \( x \) is the smallest distance of the partial orbit \( \{T^{n_k}(y)\}_{k=1}^{N} \) to the origin. Note that distance here is the Euclidean distance on \([0, 1)\). Let \( S_k = T^{-n_k}([0, a_k]) \).

**Theorem 1.4.1.** \( \bigcup_{k=0}^{N-1} A_k = B_\lambda \) up to a set of measure zero, the union being disjoint.

**Proof.** At this point it may be convenient to view modulo 1 arithmetic as a counterclockwise rotation of the circle. We first observe that if, \( A_p = [e_p, y_p] \) where \( T^{n_p}(y_p) = a_p \), then moving counterclockwise through \( A_p \) one will begin at \( e_p \) and end at \( y_p \); this follows from the definition of \( a_p \), that \( T \) is an orientation preserving isometry, and that \( T^{n_p}(A_p) = [0, a_p] \). Now suppose \( x \in A_k \) for some \( 0 \leq k \leq N - 1 \). Then \( T^{n_k}(x) \leq a_k \) and \( 0 \leq T^{n_k}(x) \leq T^{n_p}(x) \) \( \forall \) \( 0 \leq p \leq N - 1 \) otherwise \( T^{n_p}(y_k) \leq T^{n_k}(y_k) = a_k \)
a contradiction. Assume $T^{m_k}(x) < a_k$, and let $y = T^{-n_k}(a_k)$. We must have $0 \leq T^{m_p}(x) \leq T^{m_p}(y) \forall 0 \leq k \leq N - 1$. This implies $\sum_{p=0}^{N-1} f(T^{m_p}x) \geq \sum_{p=0}^{N-1} f(T^{m_p}y)$ and the assertion that $\frac{1}{N} \sum_{p=0}^{N-1} f(T^{m_p}x) \geq \lambda$ follows. To prove disjointness suppose $A_p \cap A_q \neq \emptyset$ and $a_p \geq a_q$. The irrationality of $\alpha$ ensures that $A_p \neq A_q$, $e_p \neq e_q$ and $y_p \neq y_q$. Figure 1 exhibits the case where $A_q \subset A_p$. A contradiction of the definition of $a_p$ results when $y_p$ is brought closer to 0, measures counter-clockwise, by $T^q$. Figure 2 represents the case where $e_p$ is met first when moving from 0 counter-clockwise. Again by the definition of $a_p$. Supposing $e_q$ is met first when moving counter-clockwise from 0 will have $T^{m_p}(y)$ closer to 0 then $T^{m_q}(y_q)$ a contradiction.

\[\square\]
Let $L$ denote the measure of $B_{\lambda}$. We have that $B_{\lambda} = \bigcup_{k=0}^{N-1} I_k$, where $I_k$ is an interval, possibly empty, and if $[c_k, d_k]$ denotes such an interval then $T^{n_k}(c_k) = 0$ and $T^{n_k}(d_k) = a_k$ where $a_k$ is as above.

We now create an interval of length $L$ which consists of intervals $\{J_k\}_{k=1}^{N}$ linked at their endpoints, with $|J_k| = |I_k|$, and such that the orientation of the $\{J_k\}$ is the same as that of the $I_k$. See Figure 3. Let us call this new space $X$. Map $B_{\lambda}$ to $X$ as follows. Let $\Phi : B_{\lambda} \to X$ where $\Phi(I_k) = J_k$, where $\Phi$ is defined in the obvious way as an orientation preserving isometry when so restricted. One may imagine this as pushing the intervals together eliminating the empty space between them. We now define a sequence of measure preserving transformations $\{\Psi_k\}_{k=1}^{N}$ on the probability space $(X, \mathcal{B}, \mu_L)$, where $\mu$ is the Lebesgue measure of the unit interval.

If $J_k = [r_k, s_k]$ we let $\Psi_k(x) = x + (L - r_k) \text{mod} L$, so that $\Psi(r_k) = 0$ and $\Psi(s_k) = a_k$. See Figure 4. Now let $F_N(x) = \frac{1}{N} \sum_{k=0}^{N-1} f(\Psi_k(x))$ for $x$ in $[0, L]$.

**Proposition 1.4.2.** Let $C = \{x \in [0, L] : F_N(x) \geq \lambda\}$. Then $|C| = L$.

**Proof.** Let $x \in J_k$. Since $x \in J_k$ we have that $y = \Phi^{-1}(x) \in I_k$.

For all $0 \leq p \leq N - 1$, we have $T^{n_p}(y) \geq \Psi^p(x)$, and therefore by the monotonicity of $f$, $f(T^{n_k})(y) \leq f(\Psi^k(x))$ and hence also $F_N(x) \geq A_N(y) \geq \lambda$. The first assertion of the last line follows from the fact that the transformations $T^{n_p}$ and $\Psi_k$ map $I_k$ and $J_k$ to the interval $[0, a_p]$ respectively and therefore there will be the same number of the intervals from the collections between $[0, a_p]$ and the images of $I_k$ and $J_k$ under $T^{n_p}$ respectively, however in $X$ we have eliminated the space between the intervals and thus the distance from each point to the origin has been decreased.
Theorem 1.4.3. Let \( M_\lambda = \sup \{a : \frac{1}{a} \int_0^a f(x) \, dx > \lambda \} \). Then \( |B_\lambda| \leq M_\lambda \).

Proof. Since \( |X| = L \), as above, and \( \Psi_k \) is an m.p.t. of the space \( X \), we have

\[
\lambda \leq \frac{1}{L} \int_0^L F_N(x) \, dx = \frac{1}{L} \int_0^L \sum_{k=1}^{N} f(\Psi_k(x)) \, dx = \frac{1}{L} \int_0^L f(x) \, dx.
\]

Therefore recalling that \( L = |B_\lambda| \) and that \( f \) is monotone decreasing, \( M_\lambda \geq L = |B_\lambda| \). \( \square \)

Theorem 1.4.4. Let \( P = \sup_{n_1 < \cdots < n_k} \left| B_{\frac{A}{4}} \right| \). Then \( P = M_\lambda \). Where the supremum is understood to be taken over all finite sequences.

Proof. Let \( \eta, \delta > 0 \). Since \( f \geq 0 \) is a monotone decreasing function \( f_{\lambda|\epsilon,M_\lambda} \) is a bounded function and therefore Riemann Integrable we can choose an \( \epsilon > 0 \) such that \( \int_0^\epsilon f(x) \, dx < \eta \) and \( \epsilon < \min\{\eta, \frac{M_\lambda}{2}\} \). Therefore there exists a number \( r_k \) such that if \( [\epsilon, M_\lambda] \) is partitioned into \( r_k \) intervals of equal length \( \{I_j\}_{j=1}^{r_k} \) and \( x_i \in I_i \) we
have,
\[
\left| \sum_{i=1}^{r_k} f(x_i) |I_i| - \int_{\epsilon}^{M_\lambda} f(x) \, dx \right| < \eta.
\]

Now let \( I_j = [a_j, b_j] \), where \( I_1 \) has right endpoint \( M_\lambda \) and \( I_{r_k} \) has left endpoint \( \epsilon \).

Choose \( n_j \) so that \( T^{n_j} (M_\lambda) \in I_j \) and it’s distance from \( b_j \) is less than some number \( \beta \), with \( 0 < \beta < \frac{\delta}{4r_k} \), for all \( 1 \leq j \leq r_k \). Now choose \( n_j \), \( r_k < j \leq 2r_k \) so that \( T^{n_j}(\epsilon) \in I_j \) and is within some \( \beta \) of \( a_j \). The transformation \( T^{n_i} 1 \leq i \leq r_k \) has the effect of moving most points in the intervals \( I_j, 1 \leq j \leq r_k - i \), \((i-1)\) intervals to the left while for \( r_k < i \leq 2r_k \) most points in \( I_j, r_k - i \leq j \leq r_k \), \((i-1)\) intervals to the right. Precisely, for all \( x \) except for those which are contained in a set \( S \) whose measure is determined by \( \beta \) we have for \( x \in I_j, T^{n_i}(x) \in I_{j+i} \) for \( 1 \leq i \leq r_k - j \). Also \( T^{n_i}(x) \in I_{j-i} \) for \( r_k \leq i \leq r_k + j \).

\[
\text{Figure 5. Theorem 1.4.4}
\]

Therefore with \( x_i \in I_i \) and \( x \in [\epsilon, M_\lambda]/S \),
\[
\frac{1}{2r_k} \sum_{i=1}^{2r_k} f(T^{n_i}(x)) \geq \frac{1}{2r_k} \sum_{i=r_k+1}^{r_k+j} f(T^{n_i}(x)) + \frac{1}{2r_k} \sum_{i=1}^{r_k-j} f(T^{n_i}(x))
\]
\[
= \frac{1}{2r_k} \sum_{i=1}^{r_k} f(x_i)
\]
\[
= \frac{1}{2} \frac{1}{M_\lambda - \epsilon} \sum_{i=1}^{r_k} f(x_i)
\]
\[
= \frac{1}{2} \frac{1}{M_\lambda - \epsilon} \sum_{i=1}^{r_k} f(x_i) |I_i| = F(x, \epsilon).
\]
Now,

\[
\left| 2F(x, \epsilon) - \frac{1}{M} \int_0^{M} f(x)\,dx \right| \leq \left| 2F(x, \epsilon) - \frac{1}{M} \int_0^{M} f(x)\,dx \right| + \frac{1}{M} \int_0^{\epsilon} f(x)\,dx
\]

\[
< \left| 2F(x, \epsilon) - \frac{1}{M} \int_0^{M} f(x)\,dx \right| + \frac{\eta}{M}.
\]

Therefore,

\[
2F(x, \epsilon) > \frac{1}{M} \int_0^{M} f(x)\,dx - \frac{3\eta}{M} - \frac{2\eta}{M^2} \| f \|_1.
\]

Choosing \( \eta \) small enough gives:

\[
4 \left( \frac{1}{2r_k} \sum_{i=1}^{2r_k} f(T^{n_i}x) \right) \geq 4F(x, \epsilon) > \frac{1}{M} \int_0^{M} f(x)\,dx > \lambda.
\]

This implies that for the finite subsequence \( n_1 < \cdots < n_{2r_k} \) we have that

\[
\left\{ x : \frac{1}{2r_k} \sum_{i=1}^{2r_k} f(T^{n_i}x) > \frac{\lambda}{4} \right\} \geq M - \delta.
\]

Letting \( \delta \to 0 \) yields the result. \( \square \)

**Theorem 1.4.5.** Given any interval \( I \) of length \( M \) and any \( \delta > 0 \) there exists a finite subsequence of integers \( n_0 < n_1 < \cdots < n_{k-1} \) and a subinterval \( I_\delta \subseteq I, |I_\delta| > M - \delta \) such that
∀x ∈ I_δ
\[ \frac{1}{k} \sum_{j=0}^{k-1} f(T^n_j(x)) \geq \frac{\lambda}{4}. \]

Furthermore the sequence can be made arbitrarily long. Also the choice of \( n_0 \) can be taken arbitrarily large.

Proof. Let \( I = [a, a + M_\lambda] \). In the previous proof replace \( \epsilon \) by \( a + \epsilon \) and \( M_\lambda \) by \( a + M_\lambda \) partition it and \( [\epsilon, M_\lambda] \) into \( r_k \) intervals of equal length. Let \( I_j \) be as in the previous theorem and label the intervals in \( [a + \epsilon, a + M_\lambda] \) as \( I_j' \), where \( I_1' \) has \( a + M_\lambda \) as its right endpoint and \( I_{r_k}' \) has \( a + \epsilon \) as its left endpoint. Now choose \( n_j \) with \( 1 \leq j \leq r_k \) so that \( |T^{n_j}(a + M_\lambda) - b_j| < \beta \) where \( b_j \) and \( \beta \) are as in the previous theorem. When \( r_k + 1 \leq j \leq 2r_k \) choose \( n_j \) so that \( |T^{n_k}(a + \epsilon) - a_{(2r_k - j)}| < \beta \). The remainder of the argument will be as in the previous theorem. The ergodicity of the transformation (irrationality of \( \alpha \)) ensures the claim regarding \( n_0 \). By refining the partition one creates more intervals and the sequence can be made longer, while the error between the Riemann sum and the integral will remain small. \( \square \)

1.5. Examples of separating sequences. The following can be found in [2],

Lemma 1.5.1. Let \( B_k = [n_k, n_k + 1, \ldots, n_k + l_k - 1] \) be blocks of consecutive integers such that there exists \( r \) with \( l_k \geq n_k - r \) for all big enough values of \( k \). Then the sequence \( \bigcup_{k=1}^{\infty} B_k \) is universally good in \( L^1 \).

Theorem 1.5.2. Suppose that \( f \in L^1 \) is a monotone decreasing function on \((0,1)\) and there exists a sequence \( s_k \to \infty \) such that for

\[ a_k = \sup \{ t : \frac{1}{t} \int_0^t f(x) dx > s_k \} = M_{s_k} \]

we have \( \sum_{i=1}^{\infty} a_k = \infty \). Then, if \( c_k \) is a sequence such that \( s_k c_k \to \infty \), there exists a block sequence \( \bigcup B_k \) that is universally good in \( L^1 \) and a perturbation of this sequence
\[ \bigcup (B_k \cup D_k) \text{ where } |D_k| = c_k |B_k| \] such that the ergodic averages of \( f \) along this subsequence fail to converge a.e.

**Proof.** Let \( p_1 = 0 \) and \( p_k = \sum_{j=1}^{k-1} a_j \mod(1) \) for \( k \geq 2 \), and \( J_K = [p_k, p_{k+1}] \). Since \( \sum a_j \) diverges, each point of \([0, 1)\) is in infinitely many of the \( J_k \). We will write \( B_k = [n_k, n_{k+1}, \ldots, n_k + l_{k-1}] \) for a block of integers \( B_k \) of length \( |B_k| = l_k \). We construct the sequence inductively as follows:

1. \( l_k > n_{k-1} \)
2. \( l_k \geq k l_{k-1} \geq l_1 + \cdots + l_{k-1} \).

Thus Lemma 1.5.1 implies that \( \bigcup B_k \) is universally \( L^1 \) good. Given \( \delta_k \) by 1.4.5 choose an integer \( d_k \) large enough so that there exists a subsequence of length \( d_k \) where

\[
\frac{1}{d_k} \sum_{j=0}^{d_k-1} f(T^{n_j}x) \geq \frac{s_k}{4} \quad \forall x \in (J_k)_{\delta_k}.
\]

Now \( d_k \) and \( l_k \) may be chosen so that \( d_k = c_k l_k \) and the above conditions are satisfied. Note the fact that we may arbitrarily lengthen a subsequence is key to finding the integer \( d_k \). Let \( B_k \) consist of a block of integers starting to the right of \( D_{k-1} \), and let \( D_k \) be \( d_k \) integers to the right of \( B_k \) that yield the above inequality. Therefore,

\[
\forall x \in (J_k)_{\delta_k}
\]

\[\begin{align*}
\frac{1}{l_1 + \cdots + l_k + d_1 + \cdots + d_k} \sum_{u \in \bigcup (B_j \cup D_j)} f(T^u x) &\geq C \frac{1}{l_k} \sum_{u \in D_k} f(T^u x) \\
&\geq C \frac{d_k}{l_k} s_k \\
&= C s_k c_k \to \infty.
\end{align*}\]
If the $\delta_k$'s are chosen small enough, there will exist a set of positive measure $J$ so that each $x \in J$ is in infinitely many of the $(J_k)_{\delta_k}$. Clearly such a point will have a subsequence of averages which diverge to infinity. □

Proof. (Proof of Theorem 1.1.7) By induction, we have

\[ (\log(n)(2/x))' = -\frac{1}{x \log(2/x) \cdots \log(n-1)(2/x)}. \]

Let $t > 0$ and let

\[
\begin{align*}
v_k &= \frac{1}{k \log(k) \cdots \log(n-1)(k)}, \\
s_k &= c k \log(k) \cdots \log(n-1)(k) \frac{\log^{1-t}(k)}{\log^{2}(n-1)(k)}, \\
c_k &= \frac{\log(n+1)(k)}{s_k} ; \\
g(x) &= \log(n)(2/x) \log^r(n+1)(2/x) \chi(0, \epsilon_t)(t) 
\end{align*}
\]

with $r = 2/t$, where $\epsilon_t$ small enough so that $g(x)$ is defined and $g(x)$ and $f(x)$, defined below, are monotone decreasing. Note that $\sum v_k = \infty$ and that $s_k c_k \to \infty$,

\[
g'(x) = (\log(n)(2/x))' \log^r(n+1)(2/x) + r \log^r(n+1)(2/x) \log(n)(2/x)(\log(n+1)(2/x))'
\]

\[
= -\frac{1}{x \log(2/x) \cdots \log(n-1)(2/x)} \left( \log^r(n+1)(2/x) + r \log^{r-1}(n+1)(2/x) \right)
\]

\[
= -\log^r(n+1)(2/x) B(x) \frac{x \log(2/x) \cdots \log(n-1)(2/x)}{x \log(2/x) \cdots \log(n-1)(2/x)} ; \quad B(x) \text{ a bounded function}.
\]

Now let $f(x) = \frac{g'(x)}{(g(x))^{t+1}}$ then

\[
f(x) \log^t(n) f(x) = \frac{-c(x) \log^r(n+1)(2/x) \log^t(n)(2/x)}{x \log(2/x) \cdots \log(n-1)(2/x) \log^{t+1}(2/x) \log^{r(t+1)}(n+1)(2/x)}
\]

\[
= \frac{-c(x)}{x \log(2/x) \cdots \log(n)(2/x) \log^{r+t+r-t}(n+1)(2/x)} ; \quad \text{for } c(x) \text{ bounded}.
\]
Since \( rt > 1 \) as \( r > 1/t \), we have \( f(x) \in L(\log_{(n)} L)^t \). We will show that for \( f(t) \) there exists \( c > 0 \) such that

\[
a_k = M_{cs_k} = \left\{ \lambda : \frac{1}{\lambda} \int_0^\lambda f(t) \, dt \geq cs_k \right\} \geq v_k.
\]

Therefore, since \( \sum a_k \geq \sum v_k = \infty \) and \( cs_k c_k \to \infty \) the previous theorem implies that there exists \( \{n_k\} \) such that \( \frac{1}{N} \sum_{k=0}^{N-1} f(T^{nk} x) \) diverges a.e. Now

\[
\int_0^\lambda f(x) \, dx = \int_0^\lambda \frac{g'(x)}{(g(x))^{t+1}} \, dx = \int_\infty^{g(\lambda)} \frac{1}{u^{t+1}} \, du = \frac{1}{(g(\lambda))^t},
\]

Therefore,

\[
A_\lambda = \left| \frac{1}{\lambda} \int_0^\lambda f(x) \, dx \right| = \frac{1}{\lambda(g(\lambda))^t}
\]

and, for \( \lambda < v_k = \frac{1}{k \log(k) \cdots \log_{(n)}(k)} \), we have

\[
A_\lambda \geq \frac{ck \log(k) \cdots \log_{(n)}(k)}{\log_{(n)}(k) \log_{(n+1)}(k) \log_{(n+1)}(k)} = csk.
\]

with \( c \) independent of \( k \). However,

\[
c_k = \frac{\log_{(n+1)}(k)}{s_k} = \frac{\log_{(n+1)}^3(k)}{k \log(k) \cdots \log_{(n)}(k)^t},
\]

\[
\sum \frac{1}{c_k \log_{(n)}(1/c_k)} < \infty \quad \text{for } p > t,
\]

26
and by Proposition 1.3.3 \( \frac{1}{N} \sum_{k=1}^{N-1} f(T^{mk}x) \) converges a.e. \( \forall f \in \text{LLog}^p_{(n)} \text{L} \) when \( p > t \).

\( \square \)
2. Almost everywhere convergence of convolution powers

2.1. Preliminaries. Let \( \mu(k), k \in \mathbb{Z} \), be a probability measure. Its convolution \( \mu \ast \mu \) is defined by \( \mu \ast \mu(k) = \sum_{j \in \mathbb{Z}} \mu(j) \mu(k-j) \). The \( n \)-fold convolution \( \mu \ast \cdots \ast \mu(k) = \mu^n(k) \) is defined inductively by \( \mu^n(k) = \mu \ast \mu^{n-1}(k) \). The Fourier transform of \( \mu \) will be denoted by \( \theta(t) \) for \( t \in [-1/2, 1/2) \) and it is equal to

\[
\theta(t) = \sum_{k \in \mathbb{Z}} \mu(k) e^{2\pi i kt}.
\]

The weights may be recovered from the inversion formula

\[
\mu(k) = \int_{-1/2}^{1/2} \theta(t) e^{-2\pi i kt} dt.
\]

The \( p \)-th moment of \( \mu \) is the sum \( m_p(\mu) = \sum |k|^p \mu(k) \) and we say that \( \mu \) has a \( p \)-th moment if the above sum is finite. It will be necessary to consider positive non-integral moments \( n \). The expectation of \( \mu \) is given by \( E(\mu) = \sum_{k \in \mathbb{Z}} k \mu(k) \).

Definition 2.1. \( \mu \) is called strictly aperiodic if the support of \( \mu \) is not contained in a proper coset of the integers.

We have the following important theorem by Foguel([3]).

Theorem 2.1. \( \mu \) is strictly aperiodic \( \iff |\theta(t)| \neq 1 \forall t \neq 0 \).

Definition 2.2. A probability measure \( \mu \) on \( \mathbb{Z} \) has bounded angular ratio if \( \mu \) is strictly aperiodic and there exists \( \epsilon > 0, K < \infty \) such that if \( \lambda \in \mathbb{C}, |\lambda| = 1, \lambda \neq 1, |\lambda - 1| < \epsilon \), then \( |\hat{\mu}(\lambda) - 1|/(1 - |\hat{\mu}(\lambda)|) \leq K \).

The following theorem appears in [4].

Theorem 2.2. Let \( \mu \) be strictly aperiodic. Then

1. If \( E(\mu) = 0 \) and \( m_2(\mu) < \infty \) then \( \mu \) has bounded angular ratio.
2. If \( E(\mu) \neq 0 \) and \( m_1(\mu) < \infty \) then \( \mu \) does not have bounded angular ratio.
Now let $T : X \to X$ be a measure preserving transformation of a probability space $(X, \mathcal{B}, \lambda)$. For $f \in L^1(\lambda)$ one may define

$$\mu f(x) = \sum_{k \in \mathbb{Z}} \mu(k) f(T^k x).$$

Since

$$\|\mu f\|_p \leq \sum_{k \in \mathbb{Z}} \mu(k) \|f \circ T_k\|_p = \sum_{k \in \mathbb{Z}} \mu(k) \|f\|_p = \|f\|_p,$$

$\mu f(x)$ is finite a.e and $\mu : L^p \to L^p$, $\forall p \geq 1$. The question of whether the sequence $\{\mu^n f(x)\}$ converges for almost every $x \in X$ has been taken up by various authors. The first major results were obtained in [4]. The first result in [4] is the following;

**Theorem 2.3.** If $\mu$ is symmetric and decreasing then $(\mu^n) f(x)$ converges a.e. $\forall f \in L^1(X)$.

The major results of the theory attempt to supply sufficient conditions for convergence by imposing conditions on the number of moments $\mu$ must have. The previous theorem implies that the symmetric measure given by $\mu(k) = \frac{c}{|k| \log^2 |k|}$ yields a.e. convergence of the sequence $(\mu^n f) (x)$ even though $m_p(\mu) = \infty \ \forall p > 0$.

**Proposition 2.4 ( [4]).** If $\mu$ is strictly aperiodic, then $\lim_{n \to \infty} \|\mu^n \ast \delta_1 - \mu^n\|_{L^1(\mathbb{Z})} = 0$.

**Theorem 2.5 ( [4]).** If $\mu$ is strictly aperiodic, $\mu_n f(x)$ converges a.e. for all $f \in S = \{(f_1 \circ T - f_1) + f_2 : f_1 \in L_\infty(X), f_2(Tx) = f_2(x) \text{ a.e.}\}$, which is a dense set in $L^p$ for all $p \geq 1$.

**Proof.** For $f \in L^\infty$,

$$\limsup_{n \to \infty} \|\mu^n(f \circ T - f)\|_\infty \leq \limsup_{n \to \infty} \|\mu^n \ast \delta_1 - \mu^n\|_1 \|f\|_\infty.$$

Therefore $(\mu^n g)(x)$ converges a.e. when $g = f \circ T - f$ for some $f \in L^\infty$. If $h(Tx) = h(x)$ a.e. then clearly $(\mu^n h)(x)$ converges a.e. as well. \[29\]
Theorem 2.6. If $\mu$ has bounded angular ratio and $1 < p < \infty$, then there exists a constant $K_p < \infty$ such that $\forall f \in L^p(X)$, $\|\sup_n |\mu^n f|\|_p \leq K_p \|f\|_p$. Moreover, given $f \in L^p(X)$, there exists a unique $T$-invariant $f^* \in L^p(X)$ such that $\lim_{n \to \infty} \mu^n f(x) = f^*(x)$ for a.e. $x$. If $T$ is ergodic and $m(X) < \infty$ then $f^* = \int f \, dm$ for a.e. $x$.

The second claim of this Theorem follows from Theorem 1.2.3 by combining the first claim of the Theorem and the discussion following Theorem 2.3. From 2.2 and 2.6 we get,

Theorem 2.7. If $\mu$ is strictly aperiodic and has mean 0 and finite second moment then $(\mu^n f)(x)$ converges a.e $\forall f \in L^p(X)$ for $p > 1$.

The following results establish the necessity of the hypothesis that $\mu$ has mean 0, strict aperiodicity and bounded angular ratio. A definition first,

Definition 2.8. The sequence of measures $\mu_n$ is said to have the strong sweeping out property, if given $\epsilon > 0$, there is a set $B \in \mathcal{B}$ with $m(B) < \epsilon$ such that

$$\limsup_n \mu_n \chi_B(x) = 1 \text{ a.e. and } \liminf_n \mu_n \chi_B(x) = 0 \text{ a.e.}$$

Theorem 2.9 (Losert([10])). Suppose that $\mu$ is a probability measure on $\mathbb{Z}$ ($\mu \neq \delta_k$, i.e not concentrated in a single point) and $\hat{\mu}$ has unbounded angular ratio. Then $(\mu^n)$ has the strong sweeping out property.

From Theorem 2.9 and Theorem 2.2 we obtain;

Theorem 2.10. If $m_1(\mu) < \infty$ and $E(\mu) \neq 0$ then $\mu^n$ has the strong sweeping out property.

Thus the question of convergence in $L^1(X)$ for more general $\mu$ was left open. In [5] the following was obtained.

Theorem 2.11. If $\mu$ has expectation 0 and finite second moment then $(\mu^n f)(x)$ converges a.e. for all $f \in L^1$. 

In [5] it was shown that in order to establish weak maximal inequalities for operators on continuous spaces it is enough to establish them for an operator on $l^1(\mathbb{Z})$ that has been transferred from the continuous space. Such processes are known collectively as the Calderón Transfer Principle. In our case we make use of the following version.

**Theorem 2.12** (Calderón Transfer Principle). *Consider the dynamical system $(\mathbb{Z}, \mathcal{P}, \cdot, |\cdot|, T)$ where $|B|$ = # of elements in $B$, $\mathcal{P}$ = Power set of $\mathbb{Z}$ and $T(x) = x + 1$. For $\forall \phi \in l_1(\mathbb{Z})$ we have $\mu^n \phi(k) = (\mu^n * \phi)(k) = \sum_{j \in \mathbb{Z}} \mu^n(j) \phi(k - j)$ and $(M\phi)(k) = \sup_n (\mu^n \phi)(k)$. Then if $|k \in \mathbb{Z} : |(M\phi)(k)| \geq t| \leq C \frac{\|\phi\|_{l_1}}{t}$ we have

$$\lambda \{x \in X : |(Mf)(x)| \geq t\} \leq C \frac{\|f\|_1}{t}$$

for all $(X, \mathcal{B}, \lambda, T)$.

In light of the above the following general result, which we shall use, was obtained in [5].

**Theorem 2.13.** Let $(\mu_n)$ be a sequence of probabilities on $\mathbb{Z}$ and for $f : X \to \mathbb{R}$ define the maximal operator

$$(Mf)(x) = \sup_n |(\mu_n f)(x)|, \quad x \in X$$

Assume that there is $0 < \alpha \leq 1$ and $C'' > 0$ such that for each $n \geq 1$

$$(3) \quad |\mu_n(x + y) - \mu_n(x)| \leq C'' \frac{|y|^{\alpha}}{|x|^{1+\alpha}}, \text{ for } x, y \in \mathbb{Z}, \text{ and } 2|y| \leq |x|$$

Then the maximal operator $M$ is weak type $(1,1)$, i.e there is $C > 0$ such that for any $\lambda > 0$

$$m \{x \in X : (Mf)(x) > \lambda\} \leq \frac{C}{\lambda} \|f\|_1 \text{ for all } f \in L^1(X)$$
Translating the second moment condition into a statement concerning Fourier transforms we obtain

\[ \mu \text{ has finite second moment } \iff \theta \text{ is twice continuously differentiable} \]

Therefore we will seek a condition weaker than a continuous second derivative. The most obvious condition would be that \( \theta'(t) \in \text{Lip}[\frac{-1}{2}, \frac{1}{2}] \), however the following (\cite{11}) shows this extension to be vacuous,

**Theorem 2.14** (Moricz). Let \( f(x) = \sum_{k \in \mathbb{Z}} c_k e^{ikx} \). If \( \{c_k\} \subset \mathbb{C} \) is such that \( \sum_{|k| \leq n} |kc_k| = O(n^{1-\alpha}) \), \( n = 1, 2, \ldots \), for some \( 0 < \alpha \leq 1 \), then \( f \in \text{Lip}(\alpha) \). Conversely, let \( c_k \) be a sequence of real numbers such that \( kc_k \geq 0 \) for all \( k \in \mathbb{Z} \). If \( \sum_{k \in \mathbb{Z}} |c_k| \) is finite and \( f \in \text{Lip}(\alpha) \) for some \( 0 < \alpha \leq 1 \), then \( \sum_{|k| \leq n} |kc_k| = O(n^{1-\alpha}) \).

Although this extension is vacuous the condition \( \theta'(t) \in \text{Lip}(\alpha) \) for some \( 0 < \alpha \leq 1 \) will not be. Using this we will construct examples of non-symmetric measures \( \mu \) with \( m_2(\mu) = \infty \) with \( \{\mu^n f(x)\} \) converging a.e. In fact, given \( p > 1 \), we will give examples of non-symmetric \( \mu \) with \( m_p(\mu) = \infty \) and \( \{\mu^n f(x)\} \) converging a.e.

In \cite{5} the fact that measures having finite second moment and mean 0 also satisfy the conditions of Theorem 2.13 boiled down to showing that

\[ \sup_n \int |\theta_n''(t)||t| \, dt < \infty \]

where \( \theta_n(t) = \theta^n(t) \). The results that follow will depend on achieving a similar type of bound.

We will need the following from \cite{12}.

**Theorem 2.15.** Let \( \theta(t) \) be the Fourier transform of a measure \( \mu \) on \( \mathbb{Z} \). Then there exist positive constants \( \delta \) and \( \epsilon \) such that \( |\theta(t)| \leq 1 - \epsilon t^2 \) for \( |t| \leq \delta \). Therefore there exists a \( C \) such that \( |\theta(t)| \leq e^{-Ct^2} \forall t \in [-1/2, 1/2] \).

In terms of notation \( e(x) = e^{2\pi ix} \)
Lemma 2.16. There is a constant $C > 0$ such that, for any $x, y \in \mathbb{R}$, $0 < 2|y| < |x|$, and $t \in \mathbb{R}$
\[
\left| \frac{e((x+y)t) - 1}{(x+y)^2} - \frac{e(xt) - 1}{x^2} \right| \leq C|t||y|/|x|^2.
\]

2.2. Main Results. Throughout we suppose that all measures $\mu$ are strictly aperiodic and have $E(\mu) = 0$. Note that a constant $c$, independent of certain quantities, may change throughout an argument. Our main results are the following.

Theorem 2.2.1. Suppose $\mu$ is a strictly aperiodic measure on $\mathbb{Z}$ with $E(\mu) = 0$ and for some $0 < \alpha \leq 1$ \(\sum_{|k| \leq n} k^2 \mu(k) = O(n^{1-\alpha})\). Suppose $\theta''(t)$ exists in some neighborhood $0 < |t| < \delta$, the real part of $\text{Re}(\theta''(t)) = p(t) + O(1)$, $\text{Im}(\theta''(t)) \to 0$ as $t \to 0$, where $p(t)$ is non decreasing in this set. Then \(\{\mu^n f(x)\}\) converges a.e. for all $f \in L^1(X)$.

Corollary 2.2.2. Suppose $\mu$ is a strictly aperiodic, symmetric measure on $\mathbb{Z}$ with $E(\mu) = 0$ and for some $0 < \alpha \leq 1$ \(\sum_{|k| \leq n} k^2 \mu(k) = O(n^{1-\alpha})\). Suppose $\theta''(t)$ exists in some neighborhood $0 < |t| < \delta$, and the real part of $\theta''(t) = p(t) + O(1)$, where $p(t)$ is non decreasing in this set. Then \(\{\mu^n f(x)\}\) converges a.e. for all $f \in L^1(X)$.

The following gives examples of non symmetric measures $\mu$ with $m_2(\mu) = \infty$ and $\mu_n f(x)$ converging a.e. for all $f \in L^1(X)$.

Example 2.2.3. Let $\eta(k) = s/|k|^3$, $k \neq 0$ where $s = \left( \sum 1/(|k|^3) \right)^{-1}$. Then $\hat{\eta}(t) = \sum_{k \in \mathbb{Z}} s/|k|^3 \cos(2\pi kt)$ and therefore $\hat{\eta}''(t) = -s \sum_{k > 0} \cos(2\pi kt)$. We have $\sum_{|k| \leq n} 1/|k| = O(\log(n))$ and $\sum_{k} \frac{1}{k} \cos(2\pi kt) = \log \left( \frac{1}{|2\sin(x/2)|} \right)$. Hence, $f''(t) = -s \log \left( \frac{1}{|2\sin(t/2)|} \right)$ is monotone in a neighborhood of 0. If $\nu$ is a measure with $E(\nu) = 0$ and $m_2(\nu) < \infty$ then $\mu = a_1 \eta + a_2 \nu$ will have $\theta''(t) = p(t) + O(1)(t)$ and by Theorem 2.2.1 $(\mu^n f)(x)$ converges a.e. Note here that $m_p(\mu)$ is finite if and only if $p < 2$. 

33
Example 2.2.4. Let $\eta(k) = s/|k|^{2+\sigma}$ for some $0 < \sigma < 1$, $k \neq 0$ and $s = \left( \sum 1/|k|^{2+\sigma} \right)^{-1}$. Since $\sum 1/|k|^\sigma = O(n^{1-\sigma})$ and
$$\sum_{k=1}^{\infty} \cos(2\pi kt)/k^\alpha = \Gamma(1-\alpha) \sin(\frac{1}{2}\pi \alpha)t^{\alpha-1} + O(1)(t).$$
Then we can construct a measure $\mu = a_1 \eta + a_2 \nu$ where $\nu$ is as in the previous example such that $\mu_n f(x)$ converges a.e.

Note here that $m_p(\mu)$ is finite if and only if $p < 1 + \sigma$.

Proposition 2.2.5. Suppose that the measure $\mu$ satisfies the condition $\sum_{|k|\leq n} k^2 \mu(k) = O(n^{1-\delta})$, so that the Fourier transform $\theta(t)$ has $\theta'(t) \in \text{Lip}_\delta$, for some $0 < \delta \leq 1$, then
$$|\mu^n(x)| \leq c \left\{ \frac{\sqrt{n}}{|x|^{1+\delta}} + \frac{n^2}{|x|^2} \right\}.$$

Proof.

$$|\mu^n(x)| = \left| \int_{-1/2}^{1/2} \theta^n(t)e^{-2\pi ixt} \, dt \right| \quad \text{by integration by parts}$$
$$= \left| \frac{n}{x} \int_{-1/2}^{1/2} \theta^{n-1}(t)\theta'(t)e^{-2\pi ixt} \, dt \right|$$
$$\leq c \frac{n}{|x|} \left| \int_{-1/2}^{1/2} (\theta^{n-1}(t+h)\theta'(t+h) - \theta^{n-1}(t)\theta'(t))e^{-2\pi ixt} \, dt \right|, \quad \text{where } h = \frac{1}{2x}$$
$$= c \frac{n}{|x|} \int_{-1/2}^{1/2} |\theta^{n-1}(t+h)||\theta'(t+h) - \theta'(t)| + |\theta'(t)||\theta^{n-1}(t+h) - \theta^{n-1}(t)| \, dt$$
$$= c \frac{n}{|x|} I_1 + c \frac{n}{|x|} I_2.$$

Now we examine $I_1$ and $I_2$ separately. By the Lipschitz property of $\theta'(t)$

$$I_1 = \frac{1}{x^\delta} \int_{-1/2}^{1/2} |\theta^{n-1}(t+h)||x^\delta|\theta'(t+h) - \theta'(t)| \, dt$$
$$\leq c \frac{1}{x^\delta} \int_{-1/2}^{1/2} |\theta^{n-1}(t+h)| \, dt$$
$$\leq \frac{c}{x^\delta \sqrt{n}}, \quad \text{by 2.15}.$$
Therefore,
\[
\frac{c n}{|x|} I_1 \leq \frac{c \sqrt{n}}{|x|^{1+\delta}}.
\]

\[
I_2 = \frac{1}{|x|} \int_{-1/2}^{1/2} |\theta'(t)||x||\theta^{n-1}(t+h) - \theta^{n-1}(t)|\,dt
\leq c \frac{n-1}{|x|} \int_{-1/2}^{1/2} |\theta'(t)||\theta'(c(t))||\theta^{n-2}(c(t))|\,dt \leq \frac{c n}{|x|}
\]

for some \( c(t) \) between \( t \) and \( t+h \) by the mean value theorem.

Therefore
\[
\frac{c n}{|x|} I_2 \leq \frac{c n^2}{|x|^2}.
\]

Hence
\[
|\mu_n(x)| \leq \frac{c \sqrt{n}}{|x|^{1+\delta}} + \frac{c n^2}{|x|^2}.
\]

\[\square\]

**Corollary 2.2.6.** Let \( \sigma = \min \left\{ \frac{15 \delta}{16}, \frac{3}{4} \right\} \) then \( \sigma > 0 \) and if \( n \leq |x|^\delta \),
\[
|\mu_n(x)| \leq \frac{C}{|x|^{1+\sigma}} \leq \frac{|y|^\sigma}{|x|^{1+\sigma}},
\]
\( \forall y \in \mathbb{Z}, y \neq 0 \).

**Proposition 2.2.7.** Suppose there is a neighborhood \((-\delta, \delta)\) such that \( \theta(t) \) is twice differentiable at all points except 0, and assume that if \( \theta(t) = f(t) + ig(t) \), we have

1. \( f''(t) \to -\infty \) as \( t \to 0 \) and \( g''(t) \to 0 \), \( g''(0) = 0 \)
2. \( \left| \frac{f''(t/2)}{f''(t)} \right| \geq c \).

Then, for \( t \) in \((-\delta, \delta)\),
\[
|\theta(t)| \leq 1 - c \phi(t) t^2
\]
where \( \phi(t) = \left| \frac{f'(t)}{t} \right| \).
Proof. First note that
\[ \lim_{t \to 0} \frac{f'(t/2)}{f'(t)} = \lim_{t \to 0} \frac{\frac{1}{2}f''(t/2)}{f''(t)}. \]
Thus \( \left| \frac{f''(t/2)}{f''(t)} \right| \geq c \) implies that \( \left| \frac{f'(t/2)}{f'(t)} \right| \geq c \), and also that \( f''(t) \to -\infty \) implies \( \left| \frac{f'(t)}{t} \right| \to \infty \) as \( t \to 0 \). Let \( t_0 = t/2 \) then by Taylor's Theorem for some \( c(t) \) between \( t/2 \) and \( t \),

\[
\begin{align*}
f(t) &= f(t_0) + f'(t_0)(t - t_0) + f''(c(t))\frac{(t-t_0)^2}{2} \\
&= f(t_0) + \frac{f'(t_0)\frac{t^2}{4}}{t_0} + f''(c(t))\frac{t^2}{8} \\
&= f(t_0) + \frac{t^2}{4} \left[ \frac{f'(t_0)}{t_0} + \frac{f''(c(t))}{2} \right] \\
&= f(t_0) - \frac{t^2}{4} \left[ - \left\{ \frac{f'(t_0)}{t_0} + \frac{f''(c(t))}{2} \right\} \right] \\
&= f(t_0) - \frac{t^2}{4} H(t),
\end{align*}
\]

where
\[
H(t) = - \left\{ \frac{f'(t/2)}{t/2} + \frac{f''(c(t))}{2} \right\}. \] Since \( f'' < 0 \), \( f''(c(t)) \) has the same sign as \( \frac{f'(t/2)}{t/2} \)

\[
H(t) \geq \left| \frac{f'(t/2)}{t/2} \right| \\
= 2 \left| \frac{f'(t)}{t} \right| \left| \frac{f'(t/2)}{f'(t)} \right| \\
\geq c \left| \frac{f'(t)}{t} \right|.
\]

Therefore \( |f(t)| \leq 1 - ct^2|\frac{f'(t)}{|t|}|. \)

Now, since \( \theta(t) = 1, g(0) = 0 \) and therefore

\[
\begin{align*}
g(t) &= g(0) + g'(0)t + g''(c(t))\frac{t^2}{2} \\
&= g''(c(t))\frac{t^2}{2},
\end{align*}
\]
so that \(|g(t)| \leq \frac{c t^2}{2}\). Therefore

\[
\theta(t) = f(t) + ig(t)
\]

\[
|\theta(t)| \leq |f(t)| + |g(t)|
\]

\[
\leq 1 - ct^2 \left| \frac{f'(t)}{t} \right| + \frac{c t^2}{2}
\]

\[
= 1 - ct^2 \left| \frac{f'(t)}{t} \right| \text{ since } \left| \frac{f'(t)}{t} \right| \to \infty.
\]

\[\square\]

**Lemma 2.2.8.** Suppose that \(\text{Re}(\theta''(t)) = f''(t) = p(t) + O(1)\) where \(p(t)\) is nonincreasing as \(|t| \to 0\). Then the function \(\phi(t) = \left| \frac{f'(t)}{t} \right|\) satisfies the following properties for all \(t\) in some neighborhood \([-\delta, \delta]\),

1. \(\phi(t) = \phi(-t)\)
2. \(c\phi(t) \geq |\theta''(t)|\) for \(1 < c < 2\); \(c\phi(t) \geq \left| \frac{\theta'(t)}{t} \right|\) for \(c > 0\)
3. \(|t\phi(t)| \to 0\) as \(t \to 0\) and \(|t\phi'(t)| \leq \phi(t)\)

**Proof.** \(f(t) = \sum_{k \geq 0} c_k \cos(2\pi kt)\) for \(c_k \geq 0\) and therefore the first assertion is trivial.

Since \(f'(0) = 0\), \(|t\phi(t)| \to 0\). Suppose the second moment exists. Then \(\lim_{t \to 0} \frac{f'(t)}{t} = \lim_{t \to 0} f''(t) \neq 0\) and \(\lim_{t \to 0} \frac{f'(t)}{t f''(t)} = 1\). So \(|f''(t)| \leq \left| \frac{f'(t)}{t} \right|\) for some \(c\) as close to \(1\) as we like. Note that this may require taking a smaller \(\delta\). Now suppose that \(f''(t) = p(t) + O(1)(t)\) where \(p(t) \to -\infty\) monotonically as \(|t| \to 0\). We have as \(f'(0) = 0\) for \(c(t)\) between 0 and \(t\).

\[
\phi(t) = \left| \frac{f'(t)}{t} \right| = |f''(c(t))| = |p(c(t)) + O(1)(c(t))|
\]

\[
\geq |p(t) + O(1)(c(t))|
\]
Hence for any \( c > 1 \) since \( \phi(t) \to \infty \) as \( t \to 0 \),

\[
c\phi(t) \geq |p(t)| + (c - 1)\phi(t) - |O(1)(c(t))|
\]

\[
\geq |p(t)| + |O(1)(t)| \geq |p(t) + O(1)(t) |
\]

\[
= |f''(t)|.
\]

Since \( g''(0) = 0 \), in either case we have \( |g''(t)| \leq c\phi(t) \) for a small value of \( c' \). So \( |\theta''(t)| \leq |g''(t)| + |f''(t)| \leq c\phi(t) \) for some \( 1 < c < 2 \). Observe that,

\[
\left| \frac{\theta'(t)}{t} \right| \leq \left| \frac{g'(t)}{t} \right| + \left| \frac{f'(t)}{t} \right| = g''(c(t)) + \phi(t) \leq c\phi(t).
\]

Therefore the second assertion has been established. Now \( \phi(t) = -\frac{f''(t)}{t} \) so \( \phi'(t) = \frac{-f''(t)t + f'(t)}{t^2} \) and \( t\phi'(t) = -f''(t) + \frac{f'(t)}{t} \) which implies

\[
|t\phi'(t)| = \left| \frac{f'(t)}{t} - f''(t) \right| \\
= \left| f''(t) \right| - \left| \frac{f'(t)}{t} \right| \\
\leq |2\phi(t) - \phi(t)| = \phi(t).
\]

Thus the third assertion follows.

\[ \square \]

**Lemma 2.2.9.** For a function \( \phi(t) \) satisfying the properties of Lemma 2.2.8, the following hold.

\[
(1) \quad n \int_{-\delta}^{\delta} (1 - kt^2\phi(t))^{n-1} |t| \phi(t) < C
\]

38
\[ n^2 \int_{-\delta}^{\delta} (1 - kt^2 \phi(t))^{n-2} |t|^3 \phi^2(t) < C \]

where \( C \) is independent of \( n \).

Proof. \( 1 \)

\[
\begin{align*}
 n \int_{-\delta}^{\delta} (1 - kt^2 \phi(t))^{n-1} |t| \phi(t) \, dt &= 2n \int_{0}^{\delta} (1 - kt^2 \phi(t))^{n-1} t \phi(t) \, dt \\
&= \frac{2n}{-2k} \int_{0}^{\delta} (1 - kt^2 \phi(t))^{n-1} (-2kt\phi(t)) \, dt \\
&= \frac{n}{-k} \int_{0}^{\delta} (1 - kt^2 \phi(t))^{n-1} (-2kt\phi(t) - kt^2 \phi'(t)) \, dt \\
&+ \left( \frac{n}{k} \right) \int_{0}^{\delta} (1 - kt^2 \phi(t))^{n-1} (-kt^2 \phi'(t)) \, dt.
\end{align*}
\]

Therefore

\[
2n \int_{0}^{\delta} (1 - kt^2 \phi(t))^{n-1} t \phi(t) \, dt - \frac{n}{k} \int_{0}^{\delta} (1 - kt^2 \phi(t))^{n-1} (-kt^2 \phi'(t)) \, dt = \\
= - \frac{n}{k} \int_{0}^{\delta} (1 - kt^2 \phi(t))^{n-1} (-2kt\phi(t) - kt^2 \phi'(t)) \, dt \\
= \frac{n}{k} \int_{1-k\delta^2 \phi(\delta)}^{1} u^{n-1} \, du \leq C.
\]

However, since \( |t\phi'(t)| \leq \phi(t) \) we have

\[
2n \int_{0}^{\delta} (1 - kt^2 \phi(t))^{n-1} t \phi(t) \, dt + n \int_{0}^{\delta} (1 - kt^2 \phi(t))^{n-1} t^2 \phi'(t) \, dt \\
\geq n \int_{0}^{\delta} (1 - kt^2 \phi(t))^{n-1} t \phi(t) \, dt \\
= \frac{1}{2} n \int_{-\delta}^{\delta} (1 - kt^2 \phi(t))^{n-1} |t| \phi(t) \, dt.
\]

The claim follows.
\[ n^2 \int_{-\delta}^{\delta} (1 - kt^2\phi(t))^{n-2}t^3\phi^2(t) \, dt = \]
\[ = 2n^2 \int_{0}^{\delta} (1 - kt^2\phi(t))^{n-2}t^3\phi^2(t) \, dt \]
\[ = 2n^2 \int_{0}^{\delta} (1 - kt^2\phi(t))^{n-2}t\phi(t)t^2\phi(t) \, dt \]
\[ = \frac{n^2}{-k^2} \int_{0}^{\delta} (1 - kt^2\phi(t))^{n-2}(-2kt\phi(t) - kt^2\phi'(t))t^2\phi(t) \, dt \]
\[ + \left( \frac{n^2}{k} \right) \int_{0}^{\delta} (1 - kt^2\phi(t))^{n-2}(-kt^2\phi'(t))t^2\phi(t) \, dt \]

Therefore by similar arguments as in the first part of this lemma, we have,

\[
\left| \frac{1}{2}n^2 \int_{-\delta}^{\delta} (1 - kt^2\phi(t))^{n-2}t^3\phi^2(t) \, dt \right| \leq \]
\[
\leq -\frac{n^2}{k} \int_{0}^{\delta} (1 - kt^2\phi(t))^{n-2}(-2kt\phi(t) - kt^2\phi'(t))t^2\phi(t) \, dt \]
\[ = \frac{n^2}{k} \int_{1-k\delta^2\phi(\delta)}^{1} u^{n-2}(1 - u) \, du \]
\[ = \frac{n^2}{k^2} \int_{1-k\delta^2\phi(\delta)}^{1} u^{n-2} - u^{n-1} \, du \]
\[ = \frac{n^2}{k^2} \left( \frac{u^{n-1}}{n-1} - \frac{u^n}{n} \right|_{1-k\delta^2\phi(\delta)}^{1} \right) \]
\[ = \frac{n^2}{k^2} \left( \frac{1}{n-1} - \frac{1}{n} \right) + \frac{(1-k\delta^2\phi(\delta))^n}{n} - \frac{(1-k\delta^2\phi(\delta))^{n-1}}{n-1} \]
\[ \leq \frac{n^2}{k^2} \left( \frac{1}{(n-1)n} + O \left( \frac{(1-k\delta^2\phi(\delta))^n}{n} \right) \right) \]
\[ \leq C \frac{n^2}{k^2} \left( \frac{1}{n^2} + \frac{1}{n} O((1-k\delta^2)^n) \right) \]
\[ \leq C \frac{n^2}{k^2} \left( \frac{1}{n^2} + \frac{1}{n} O(e^{-nk\delta^2}) \right) \]
\[ \leq C \frac{n^2}{k^2} \left( \frac{1}{n^2} + O \left( \frac{1}{n^3} \right) \right) \leq C \]
Theorem 2.2.10. Suppose that $\theta(t)$ is twice differentiable in a neighborhood $[-\delta, \delta]$ of 0, except perhaps at 0, and there exists a function $\phi(t)$ satisfying the properties of Lemma 2.2.8 such that

(1) $|\theta(t)| \leq 1 - kt^2 \phi(t)$
(2) $\left| \frac{\theta'(t)}{t} \right| \leq c \phi(t)$
(3) $|\theta''(t)| \leq c \phi(t)$

on $[-\delta, \delta]$. Then

$$|\mu_n(x + y) - \mu_n(x)| \leq c \frac{|y|}{|x|^2} \text{ for } n \geq |x|^{\delta/8}$$

Proof. By the inversion formula

$$\mu_n(x) = \int_{-1/2}^{1/2} \theta^n(t) e^{-2\pi i xt} \, dt$$

$$= \int_{|t| \leq \delta} \theta^n(t) e^{-2\pi i xt} \, dt + \int_{|t| \geq \delta} \theta^n(t) e^{-2\pi i xt} \, dt$$

$$= I_{1,x} + I_{2,x}$$

Now, if $\theta^n(t) = \theta_n(t)$ we note that since $\theta'(t) = \int_{\epsilon}^{t} \theta''(t) - \theta''(\epsilon)$, letting $\epsilon \to 0$ we have $\theta'(t) = \int_{0}^{t} \theta''(t)$ and $\theta'(t)$ is absolutely continuous on $[-\delta, \delta]$ hence

$$I_{1,x} = \frac{1}{2\pi i} \left( - \frac{\theta_n(t) e^{-2\pi i xt}}{x} \right)_{-\delta}^{\delta} + \frac{1}{x} \int_{-\delta}^{\delta} \theta_n'(t) e^{-2\pi i xt} \, dt$$

$$= \frac{1}{2\pi i} \left( - \frac{\theta_n(t) e^{-2\pi i xt}}{x} \right)_{-\delta}^{\delta} + \frac{1}{4\pi^2} \left( \theta_n'(t) e^{-2\pi i xt} \right)_{-\delta}^{\delta}$$

$$- \frac{1}{4\pi^2} \left( \frac{1}{x^2} \int_{-\delta}^{\delta} \theta_n''(t) e^{-2\pi i xt} \, dt \right)$$

$$= Q_x - \frac{1}{4\pi^2 x^2} \int_{-\delta}^{\delta} \theta_n''(t) e^{-2\pi i xt} \, dt$$
Let for some $C_n\varepsilon$ 
\[
\theta(t) = f(t) + ig(t)\text{ is the Fourier transform of }\mu
\]
and $\theta''(t)$ exists at all points except possibly 0 in a neighborhood $[-\delta, \delta]$. Then if
\[ f''(t) = p(t) + O(1) \] where \( p(t) \) is monotone, there exists \( 0 < \alpha \leq 1 \) such that

\[
|\mu^n(x + y) - \mu^n(x)| \leq C \frac{|y|^\alpha}{|x|^{1+\alpha}}, \forall |y| \leq \frac{|x|}{2}.
\]

Proof. We need only consider the case where \( m_2(\mu) = \infty \). In this case \( \rho(t) \to -\infty \) as \( |t| \to 0 \). We first show that \( \left| \frac{f''(t/2)}{f''(t)} \right| \geq c > 0 \).

We have

\[
\left| \frac{f''(t/2)}{f''(t)} \right| = \frac{|p(t/2) + O(1)(t/2)|}{|p(t) + O(1)(t)|} \\
\geq \frac{|p(t/2)| - |O(1)(t/2)|}{|p(t)| + |O(1)(t)|} \\
\geq \frac{|p(t/2)|}{|p(t)| + |p(t)|} - O(|1/p(t)|) \geq 1/4
\]

Therefore Proposition 2.2.7 gives \( |\theta(t)| \leq 1 - c\phi(t)t^2 \) where \( \phi(t) \) satisfies the properties of Lemma 2.2.8. Therefore Theorem 2.2.10 gives

\[
|\mu_n(x + y) - \mu_n(x)| \leq c \frac{|y|}{|x|^2}, |y| \leq |x|/2, \text{ for } n \geq |x|^\delta/8.
\]

The condition \( \sum_{|k| \leq n} |k|^2 \mu(k) = O(n^{1-\delta}) \) gives, by Corollary 2.2.6, the inequality when \( n \leq |x|^\delta/8 \). \( \square \)

References


Department of Mathematics, University at Albany, SUNY, Albany, NY 12222

E-mail address: cw4347@albany.edu