INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Almost everywhere convergence of weighted averages

Reinhold-Larsson, Karin Beatriz, Ph.D.

The Ohio State University, 1991
Almost Everywhere Convergence of Weighted Averages

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Karin B. Reinhold-Larsson,

*** *** *

The Ohio State University

1991

Dissertation Committee:
Joseph Rosenblatt
Boris Pittel
Gerald Edgar

Approved by

Joseph Rosenblatt
Adviser
Department of Mathematics
To my parents

and

to the memory of Paul Larsson.
I would like to express my gratitude to my advisor, Professor Joseph Rosenblatt. Throughout these years, he has set an example of excellence as a mathematician. This work is the result of our many invaluable discussions and I am indebted to him for his constant support and encouragement during the many stages of its preparation. Furthermore, I am appreciative of his perspective on life and mathematics.

Other mathematicians have contributed to my education. I had many enthusiastic discussions with Dr. Mâte Wierdl while he was at OSU. He is always willing to listen and discuss any problem down to the very technical details. Professor Roger Jones read preliminary versions of this work and encouraged its development. Professor Vitaly Bergelson introduced me to ergodic theory and dynamical systems and Professors Boris Pittel and Ranco Bojanic provided me with helpful discussions. Professor Peter March and I had many interesting discussions. His is a friendly and inviting way to look at problems. I admire Professor Boris Mityagin’s unique way of discussing mathematics and for making the Department of Mathematics a warm place to be. I want to express my gratitude to Professor Cora Sadosky who encouraged me to pursue a PhD in the USA.

Finally, I thank my parents who inculcated in me the love for knowledge and, despite their personal loss, encouraged me to pursue my career far away from them.
Also, I give thanks to my brother, Carlos, and my friend, Nora Figueroa, for their words of support.
**VITA**

<table>
<thead>
<tr>
<th>Year(s)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>May 30, 1963</td>
<td>Born in Mar del Plata, Buenos Aires, Argentina</td>
</tr>
<tr>
<td>1984–1986</td>
<td>Teaching Associate, Department of Mathematics, University of Buenos Aires,</td>
</tr>
<tr>
<td></td>
<td>Buenos Aires, Argentina</td>
</tr>
<tr>
<td>1986–1987</td>
<td>Graduate Teaching Associate, Department of Mathematics, The Ohio State</td>
</tr>
<tr>
<td></td>
<td>University</td>
</tr>
<tr>
<td>1987–1988</td>
<td>Graduate Research Associate, Mathematical Clinic (Institute for Mathematics</td>
</tr>
<tr>
<td></td>
<td>and its Applications), Department of Mathematics, The Ohio State University</td>
</tr>
<tr>
<td>1988–1989</td>
<td>Graduate Research Associate, Non Destructive Evaluation Laboratory,</td>
</tr>
<tr>
<td></td>
<td>Department of Welding Engineering, The Ohio State University</td>
</tr>
<tr>
<td>Fall 1989</td>
<td>Graduate Teaching Assistant, Department of Mathematics, The Ohio State</td>
</tr>
<tr>
<td></td>
<td>University</td>
</tr>
</tbody>
</table>

**FIELDS OF STUDY**

Major Field: Mathematics
Studies in Ergodic Theory with Professor Joseph Rosenblatt
# Table of Contents

ACKNOWLEDGMENTS ......................................................... iii

VITA ......................................................................................... v

CHAPTER

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1.1</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>5</td>
</tr>
<tr>
<td>II</td>
<td>12</td>
</tr>
<tr>
<td>2.1</td>
<td>13</td>
</tr>
<tr>
<td>2.2</td>
<td>34</td>
</tr>
<tr>
<td>2.3</td>
<td>40</td>
</tr>
<tr>
<td>III</td>
<td>45</td>
</tr>
<tr>
<td>3.1</td>
<td>47</td>
</tr>
<tr>
<td>IV</td>
<td>63</td>
</tr>
<tr>
<td>4.1</td>
<td>64</td>
</tr>
<tr>
<td>4.2</td>
<td>86</td>
</tr>
<tr>
<td>V</td>
<td>91</td>
</tr>
<tr>
<td>5.1</td>
<td>92</td>
</tr>
</tbody>
</table>

APPENDICES

vi
A Appendix to Chapter 1 ......................................................... 106
B Appendix to Chapter 2 ......................................................... 108
C Appendix to Chapter 3 ......................................................... 114
D Appendix to Chapter 4 ......................................................... 117

BIBLIOGRAPHY ............................................................................ 122
CHAPTER I

Introduction to Ergodic Theory

This chapter is a brief introduction to ergodic theory and to the theory of weighted averages.

Before starting I would like to mention the etymology of the word “ergodic”. “Ergo” comes from the Greek “ergon” which means work. An erg is the centimeter-gram-second unit of work or energy and is equal to the work done by a force of one dyne when its point of application moves through a distance of one centimeter in the direction of the force (10^{-7} Joules) (Webster’s Dictionary). The term “ergodic” was originally coined to describe particle systems in which orbits penetrate all corners of the space. In mathematics, ergodic theory is the study of the long-term average behavior of systems.

1.1 Definitions and Main Results

We begin with the classical ergodic theorems for measure preserving transformations. Let $(X, \beta, m)$ be a probability space and $\tau : X \rightarrow X$ a measure preserving transformation ($m(\tau^{-1}A) = m(A)$ for all $A \in \beta$). We call $\tau$ ergodic (with respect to $m$) if for any set $A \in \beta$ with positive measure such that $m(A \Delta \tau^{-1}A) = 0$ then $A = X$ or $A = \emptyset$ up to sets of measure 0. In other words, $\tau$ is ergodic if it does not have any
non-trivial invariant subsets.

The following theorems are the corner stones of ergodic theory.

Lemma 1.1.1 Let $U$ be an operator on $L^2(X)$ with $\|U\| \leq 1$. Then

$$L^2(X) = \{f \in L^2(X) : Uf = f\} \oplus \{f - Uf : f \in L^2(X)\}.$$ 

Let $f$ be a measurable function and denote by $A_nf$ the time averages of $f$,

$$A_nf(x) = \frac{1}{n} \sum_{k=0}^{n-1} f(\tau^k x).$$

Theorem 1.1.2 Mean Ergodic Theorem (von Neumann, 1931)

Let $(X, \beta, m)$ be a probability space and $\tau$ a measure preserving transformation on it. Then, for all $p \geq 1$ and any $f \in L^p(X)$, there exists $f^* \in L^p(X)$ with $f^* \circ \tau = f^*$ such that $\lim_{n \to \infty} \|A_nf - f^*\|_p = 0$.

This theorem was first proved for ergodic transformations for which the set of $\tau$-invariant functions consists precisely of the constant functions. Therefore, in the ergodic case, $f^* = \int f dm$, which means that the time average of a function equals its spatial average.

Proof. Since $L^2 \cap L^p$ is dense in $L^p$ for all $p \geq 1$; then, by the above lemma, it suffices to prove the theorem for functions of the form $f = g - g \circ \tau$ for some $g \in L^p(X)$.

$$\|A_nf\|_p = \frac{1}{n} \|\sum_{k=0}^{n-1} (g - g \circ \tau) \circ \tau^k\|_p$$

$$= \frac{1}{n} \|g - g \circ \tau^n\|_p$$

$$\leq \frac{2}{n} \|g\|_p \to 0 \quad \text{as} \quad n \to \infty.$$
The Ergodic Theorem of von Neumann was strengthened by Birkhoff who proved a pointwise version of the theorem. Notice that if in the above proof we take \( g \in L^\infty(X) \) instead of \( L^p(X) \),

\[
|A_n(g - g \circ \tau)(x)| = \frac{1}{n} \left| \sum_{k=0}^{n-1} g(\tau^k x) - g(\tau^{k+1} x) \right|
\]

\[
= \frac{1}{n} |g(x) - g(\tau^n x)|
\]

\[
\leq \frac{2}{n} \|g\|_\infty \to 0
\]

as \( n \to \infty \). Therefore, \( \lim_{n \to \infty} A_n f(x) \) exists a.e. for all

\[
f \in \{ f \in L^p(X) : f \circ \tau = f \} \oplus \text{cl}-\{ f - f \circ \tau : f \in L^\infty(X) \},
\]

which is a dense subspace of \( L^p(X) \).

**Theorem 1.1.3 Maximal Ergodic Theorem (Yosida–Kakutani, 1939)**

Let \((X, \beta, m)\) be a probability space and \( \tau \) a measure preserving transformation on it. Then for any \( f \in L^1(X) \),

\[
\int_{\{\sup_{n \in \mathbb{N}} A_n f > 0\}} f(x) dm(x) > 0.
\]

From this theorem it follows that the maximal function \( \sup_{n \in \mathbb{N}} |A_n f| \) is a weak \((1,1)\) operator; that is,

\[
m\{ x : \sup_{n \in \mathbb{N}} |A_n f(x)| > \lambda \} \leq C \frac{\|f\|_1}{\lambda}.
\]

This maximal estimate gives the theorem on almost everywhere (a.e.) convergence.
Theorem 1.1.4 Pointwise Ergodic Theorem (Birkhoff, 1931)
Let $(X, \beta, m)$ be a probability space and $\tau$ a measure preserving transformation on it. Then, for all $p \geq 1$, $\lim_{n \to \infty} A_n f(x) = f^*(x)$ for all $x$ except possibly for a set of measure zero, where $f^* \circ \tau = f^*$ is a $\tau$-invariant function.

The Pointwise Ergodic Theorem can be proved by noticing that there is convergence a.e. for a dense subspace of functions and that a maximal estimate holds. Then, appealing to the Banach Principle (Garsia [22]), convergence for the whole space follows. The proof of this theorem does not need to be necessarily based on this principle, although we mention it because this is precisely the technique used to prove a.e. convergence theorems in the subsequent chapters.

Theorem 1.1.5 Banach Principle
Let $(X, \beta, m)$ be a measure space and $\{T_n\}$ a sequence of linear, continuous operators on $L^p(X) \ (1 \leq p)$. Then,

(a) If $\sup_{n \in \mathbb{N}} |T_n f(x)| < \infty$ a.e. for all $f \in L^p(X)$, there exists a positive decreasing function $C(\lambda) \to 0$ as $\lambda \to \infty$ such that for all $f \in L^p(X)$ we have

$$m\{x : \sup_{n \in \mathbb{N}} |T_n f(x)| > \lambda \|f\|_p\} \leq C(\lambda)$$

for all $\lambda > 0$.

(b) If there exists a positive decreasing function $C(\lambda)$ as above such that for all $f \in L^p(X)$ we have

$$m\{x : \sup_{n \in \mathbb{N}} |T_n f(x)| > \lambda \|f\|_p\} \leq C(\lambda)$$

for all $\lambda > 0$ then the set of functions $f \in L^p(X)$ for which $\lim_{n \to \infty} T_n f(x)$ exists a.e. is closed.
In the following chapters, we use the notation $A_n f$ to denote different types of averages. In chapter 2, we refer to $A_n f$ to denote two-sided averages rather than one-sided averages as above. That is

$$A_n f(x) = \frac{1}{2n+1} \sum_{k=-n}^{n} f(\tau^k x).$$

For these averages, the transformation $\tau$ also needs to be invertible.

$n$-dimensional averages are also considered. Let $\tau_1, \ldots, \tau_d$ be measure preserving, invertible and commuting transformations on $(X, \beta, m)$. We define the $d$-dimensional averages $A_n f$, induced by $\tau_1, \ldots, \tau_d$, by

$$A_n f(x) = \frac{1}{(2n+1)^d} \sum_{k_1 \in [-n,n]} \cdots \sum_{k_d \in [-n,n]} f(\tau_1^{k_1} \cdots \tau_d^{k_d} x).$$

In the case of continuous time, if $\tau_1, \ldots, \tau_d$ are measure preserving, invertible and commuting flows on $(X, \beta, m)$, $(\tau^{u} \tau_j = \tau^{u+v}$ and $\tau_j^0 = id$ for all $j$), then the averages they induced are defined as

$$A_n f(x) = \frac{1}{(2n)^d} \int_{[-n,n]^d} f(\tau_1^{u_1} \cdots \tau_d^{u_d} x) du.$$

The Mean and Pointwise Ergodic Theorems also hold for these averages (Krengel [28]).

1.2 Weighted Averages

Let $(X, \beta, m)$ be a probability space and $\tau$ a measure preserving transformation on it. Let $\{a_n^N\}$ be sequences of positive numbers such that $\sum_n a_n^N = 1$ for all $N$. Averages of the form

$$S_N f(x) = \sum_n a_n^N f(\tau^n x)$$


are called **weighted averages**. Two different types of weighted averages will be considered. The first ones are defined by measures in \( Z \). If \( \mu \) is a probability measure on \( Z \), it induces an average whose weights are given by the measure,

\[
\mu f(x) = \sum_{k \in \mathbb{Z}} \mu(k)f(\tau^kx).
\]

The second type are averages induced by sequences of positive numbers \( \{n_k\} \subset \mathbb{N} \),

\[
A_N f(x) = \frac{1}{N} \sum_{k=1}^{N} f(\tau^{n_k}x).
\]

We refer to the latter as averages along the subsequence \( \{n_k\} \).

The study of weighted averages arose as a generalization of the classical averages, with connections to questions in harmonic analysis. Some of the most important work in the area is that of A. Bellow, J. Bourgain, S. Connes, W. Emerson, R. Jones, U. Krengel, S. Losert, J. Rosenblatt and M. Wierdl.

This work is divided into four parts. The first two treat averages induced by probability measures on \( \mathbb{R}^d \), the third treats averages on amenable groups and the last deals with averages along subsequences. A brief introduction to these topics is presented below.

Averages induced by probability measures were studied extensively by A. Bellow, R. Jones and J. Rosenblatt. Given a sequence of probabilities \( \{\mu_n\} \) on \( Z \), they give sufficient conditions, in terms of the Fourier transform of the measures, for the convergence a.e. on all \( L^p(X) \) \( (p > 1) \) of the operators \( \mu_n f(x) \) as \( n \to \infty \) (see [4]). In case the full sequence fails to converge, they also give sufficient conditions for the convergence of subsequences. In the special case in which the sequence of measures consists of the convolution powers of a single measure, more could be said.
Theorem 1.2.1 (Bellow–Jones–Rosenblatt [5])

Let $1 < p < \infty$. If $\mu$ is a strictly aperiodic probability measure on $\mathbb{Z}$ with bounded angular ratio, that is,

$$\sup_{|\gamma| = 1, \gamma \neq 1} \frac{|\hat{\mu}(\gamma) - 1|}{1 - |\hat{\mu}(\gamma)|} < \infty,$$

then for all $f \in L^p(X)$ there exists a unique $\tau$-invariant function $f^* \in L^p(X)$ such that

$$\lim_{n \to \infty} \mu^n f(x) = f^*(x) \text{ a.e.}$$

Bellow–Jones–Rosenblatt also showed that $\mu$ has bounded angular ratio if $m_2(\mu) < \infty$, $E(\mu) = 0$ and $\mu$ is strictly aperiodic. In the same paper, they give a sufficient condition for the operators $\mu^n f$ to diverge:

$$\lim_{\gamma \to 1} \frac{|\hat{\mu}(\gamma) - 1|}{1 - |\hat{\mu}(\gamma)|} = \infty.$$

With this criteria, they constructed an example of a strictly aperiodic probability measure with only a finite first moment for which the averages diverge. This is a very important example, showing the relevance of the finiteness of some moment, even in the case where $E(\mu) = 0$.

In Chapter 2, it is shown that under similar conditions, the averages $\mu^n f(x)$ converge a.e. in $L^1(X)$. Bellow–Jones–Rosenblatt’s method is based on Fourier transform techniques which prove the result on $L^2(X)$ and admits to be extended to $L^p(X)$, $p > 1$, but not to $L^1(X)$. Instead, in this chapter a probabilistic argument is used to handle the case $p = 1$. Since the operators are defined by convolution powers of measures, the theory of limiting distributions can be applied to its treatment. We
compare the powers of $\mu$ with those of the normal distribution. This is a very simple idea, but it gives a clearer understanding of the behavior of these averages. In this sense, our method is stronger than that of Bellow–Jones–Rosenblatt, but unfortunately, we can not get as a strong a result in terms of the moment condition. We prove the following

**Theorem 1.2.2** Let $\mu$ be a strictly aperiodic probability measure on $\mathbb{Z}$ with $E(\mu) = 0$ and $m_{2+\delta}(\mu) < \infty$ for some $\delta > (\sqrt{17} - 3)/2$. Then, for any $f \in L^1(X)$, there exists a unique $\tau$-invariant function $f^* \in L^1(X)$ such that $\lim_{n\to\infty} \mu^n f(x) = f^*(x)$ a.e.

Related questions of a.e. convergence are also discussed for measures which do not satisfy the conditions of the theorem as well as properties of the maximal function. In Chapter 3, the $L^1$ convergence result is extended for measures on $\mathbb{Z}^d$ and $\mathbb{R}^d$. The dimension of the space affects the moment condition required on $\mu$.

**Theorem 1.2.3** Let $\mu$ be a strictly aperiodic, spread out probability measure on $\mathbb{Z}^d$. If $d$ and $\mu$ satisfy one of the following three conditions, then $\lim_{n\to\infty} \mu^n f$ exists for all $f \in L^1(X)$.

1. $d = 1, 2$ and $m_{2+\delta}(\mu) < \infty$, for some $[\sqrt{(2+d)^2 + 8d - (d+2)}/2 \leq \delta$,

2. $d = 3$ and $m_3(\mu) < \infty$,

3. $d > 3$, $\mu$ is symmetric and $m_{3+\delta}(\mu) < \infty$ for some $\delta$ in the range $[\sqrt{(2+d)^2 + 8d - (d+2)}/2 \leq \delta < 1$.

For measures on $\mathbb{R}^d$, stronger conditions are required in terms of the relation between $\mu$ and the Lebesgue measure.
Chapter 4 deals with averages on amenable groups. Given \( \{F_n\} \), a sequence of sets in \( G \) with positive measure and a measurable action \( T : G \times X \rightarrow X, T(g, x) = T_g(x) \), define the averages \( A_n f \) by

\[
A_n f(x) = \frac{1}{|F_n|} \int_{F_n} f(T_g x) d\lambda_G(g).
\]

We discuss averages along Fσ-ınner sequences \( \{F_n\} \). This is the corresponding generalization of the usual averages on general groups. The Mean Ergodic Theorem is immediate for averages along Fσ-ınner sequences, however, the Pointwise Ergodic Theorem along Fσ-ınner sequences requires some care. The problem of a.e. convergence was studied by Tempel’man, Emerson, Greenleaf, Ornstein and Weiss.

If the Fσ-ınner sequence is not centered \( (F_n \subset F_{n+1}) \), then the measure of the set on which the average is performed may not be the correct normalization to obtain a.e. convergence. A. Tempel’man was the first to discuss such issues. He found a normalization for which the averages do converge a.e. but he did not know whether his normalization was the right one. That is, there could be a sequence of numbers increasing at a rate slower than his normalization for which the averages also converge a.e. J. Rosenblatt and M. Wierdl in [36] proved a similar result for block sequences on \( \mathbb{Z} \) and also showed that Tempel’man’s normalization is the correct one for blocks of increasing length. In such cases, they called the normalization the “correct factors” associated to the sequence. In this chapter we prove that under certain regularity conditions on the Fσ-ınner sequence, Tempel’man’s normalization is also the correct factor.

Lastly, we treat pathological behaviors of averages along subsequences. We con-
struct sequences that are good universal in some $L^p(X)$ spaces and bad universal in others. U. Krengel [27] was the first to introduce the concept of a bad universal sequence. He showed that any sequence has a further subsequence which is bad universal in $L^\infty(X)$; that is, the convergence fails even for bounded functions. A. Bellow and V. Losert [7] later proved that any sufficiently lacunary sequence is bad universal, for example $n_k = 2^k$. J. Rosenblatt [18] developed the criteria of $\delta$-sweeping out for subsequences. A sequence $\{n_k\}$ is $\delta$-sweeping out if for all ergodic dynamical systems and all $\epsilon > 0$, there is a measurable set $E$ with $m(E) < \epsilon$ such that 

$$\limsup_{n \to \infty} A_n 1_E \geq \delta \text{ a.e.}$$

This is a weaker version of the strong-sweeping out property ($\delta = 1$) which was previously introduced by A. Bellow. Although A. Bellow and J. Rosenblatt were the first to use the strong-sweeping out property to show bad universal properties of sequences, it was A. Bellow who coined the term and used it systematically. J. Rosenblatt and del Junco [18] showed that if a sequence has the strong-sweeping out property, then for sets $E$ in a residual class, 

$$\limsup_{n \to \infty} A_n 1_E = 1$$

and 

$$\liminf_{n \to \infty} A_n 1_E = 0.$$ 

This illustrates how badly the averages along the subsequence behave. Rosenblatt [35] showed that sequences with only the $\delta$-sweeping out property are bad universal in $L^\infty$ and gave examples of sequences that are $\delta$-sweeping out but not strong-sweeping out. Sufficiently lacunary sequences were known to have the strong sweeping out property. In addition to this technique, J. Bourgain [9] developed an entropy criterion which is helpful in determining whether a sequence is bad universal in $L^\infty$. J. Rosenblatt [35] showed that the entropy of the averages is related to the sweeping out behavior of the sequence.
The above are all negative results for subsequences whose averaging behavior is so bad that they are already bad for bounded functions. With regard to positive results, it was shown by J. Bourgain [8] and M. Wierdl [40] that the sequence of primes is good universal in $L^p$ for all $p > 1$. J. Bourgain [10, 11] also proved that the sequence of squares is good universal in $L^p$, $p > 1$. Moreover, he proved that for any polynomial $p(x) \in \mathbb{Z}[X]$, the sequence $\{p(n)\}$ is good universal $L^p$, $p > 1$. M. Wierdl has other examples of sequences whose gaps increase to infinity which are good universal in $L^p$ for all $p > 1$. By the nature of their technique their examples can not be extended to $L^1$. Wierdl has conjectured, however, that any such sequence is bad universal in $L^1$.

All the examples above give positive results in $L^p$ for all $p > 1$ or negative results in $L^\infty(X)$ and hence in all $L^p$, $p \geq 1$. But results that would distinguish among $L^p$ spaces were not known. W. Emerson and A. Bellow showed that one can in fact construct sequences that, for a fixed $1 \leq p < \infty$, are good universal for all $L^q$ for all $q \geq p$ and bad universal in $L^q$, $1 \leq q < p$. But they did not give examples for the extreme cases. It was not clear whether Wierdl’s conjecture was even possible; that is, it was not known the existence of a sequence which is good universal in all $L^p$, $p > 1$, but bad universal in $L^1$. Or even more surprisingly, if there could be sequences which are only good universal in $L^\infty$ and bad universal in all other $L^p$, $p < \infty$. Chapter 5 answers both of these questions.
In this chapter we study the behavior of weighted averages induced by a probability measure on the integers. Let $(X, \beta, m)$ be a probability space and $\tau : X \rightarrow X$ an invertible measure preserving point transformation. A probability measure $\mu$ on $\mathbb{Z}$, the integers, gives rise to the weighted average

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

The powers of $\mu f$ are, defined by the convolution powers of $\mu$:

$$\mu^n f(x) = \sum_{k=-\infty}^{+\infty} \mu^n(k)f(\tau^k x)$$

where, on the right hand side, $\mu^n(k)$ denotes the $n$th convolution power of $\mu$ evaluated at $k$. Note that since

$$(\int |\mu^n f(x)|^p dm(x))^{\frac{1}{p}} \leq \sum_{k \in \mathbb{Z}} \mu^n(k)(\int |f(\tau^k x)|^p dm(x))^{\frac{1}{p}} = \|f\|_p,$$

these operators are well defined a.e. and are positive contractions in all $L^p(X)$, $1 \leq p \leq \infty$. Bellow-Jones-Rosenblatt [4], [3], [6] studied these types of averaging operators as well as more general types of weighted averages. They proved that these operators converge in norm whenever the support of $\mu$ is not contained in a coset of a proper
subgroup of \( \mathbb{Z} \). In addition, they proved ([3]) that if the measure is centered and has finite second moment, then there is convergence almost everywhere in \( L^p(X) \) for \( p > 1 \). For this they used Fourier techniques that could not be extended to \( p = 1 \).

V. I. Oseledec [32] proved convergence almost everywhere on \( L \log L \) for symmetric measures without any moment condition. His proof is based on Doob's Martingale Convergence Theorem. We show a.e. convergence on \( L^1(X) \), via estimations of the probability distribution of the convolution powers \( \mu^n(k), k \in \mathbb{Z} \).

This chapter is divided into three parts. In the first, we prove convergence a.e. of \( \mu^n f \) in \( L^1(X) \) for measures that are centered, satisfy some moment condition and so that their support has an appropriate distribution. In the second, we analyze the convergence of \( \mu^n f \) along subsequences for measures \( \mu \) which are not centered. For such measures, convergence of the whole sequence does not hold, but a moment condition is still required for a subsequence result. And in the third, we prove several other properties of the maximal function \( \sup_{n \in \mathbb{N}} \mu^n f \) for centered measures with a moment condition higher than two.

### 2.1 Convergence for centered measures

Initially, we will focus on the properties of the measure \( \mu \), beginning with some definitions and well-known facts.

**Definition 2.1.1** Let \( \mu \) be a probability measure on \( \mathbb{Z} \). We say that \( \mu \) is adapted if its support generates \( \mathbb{Z} \), and that \( \mu \) is strictly aperiodic if its support is not contained in a coset of a proper subgroup of \( \mathbb{Z} \).
For $\alpha > 0$, the $\alpha$-moment of $\mu$ is defined as $\sum_{k=-\infty}^{\infty} |k|^\alpha \mu(k)$, and is denoted by $m_\alpha(\mu)$. The expected value of $\mu$ is $\sum_{k=-\infty}^{\infty} k \mu(k)$, and is denoted by $E(\mu)$.

The measure $\mu$ is called centered if it has expected value zero.

The following are useful characterizations of strictly aperiodic probabilities.

**Proposition 2.1.2** (FOGUEL [21]). Let $\mu$ be a probability measure on $\mathbb{Z}$.

i) If $\mu$ is adapted, then $\mu$ is strictly aperiodic if and only if
\[ \lim_{n \to \infty} \| \mu^{n+1} - \mu^n \|_{M(\mathbb{Z})} = 0. \]

ii) $\mu$ is strictly aperiodic if and only if $|\hat{\mu}(\lambda)| < 1$ for all $\lambda \in C$ with $|\lambda| = 1, \lambda \neq 1$.

The strict aperiodicity of the measure will be needed to prove convergence a.e. for all functions on $L^1(X)$.

**Proposition 2.1.3** Let $1 \leq p < \infty$. Then
\[ \{ f \in L^p(X) : \mu f = f \} + \text{cl} \{ f - \mu f : f \in L^p(X) \} \]

is a dense subspace of $L^p(X)$. Also, if $\mu$ is strictly aperiodic and $\tau$ is ergodic, $\text{cl} \{ f - \mu f : f \in L^p(X) \}$ is the subspace of mean zero functions in $L^p(X)$.

**Proof.** Since $L^2(X) \cap L^p(X)$ is dense in $L^p(X)$ and Lemma 1.1.1 holds on $L^2$,
\[ \{ f \in L^p(X) : \mu f = f \} + \text{cl} \{ f - \mu f : f \in L^p(X) \} \]

is dense in $L^p(X)$.

Assume now that $\mu$ is strictly aperiodic and $\tau$ is ergodic. Then, by the Spectral Theorem, there exists a positive measure $m_f$ on the unit circle such that
\[ < \tau^n f, f > = \int_{|\gamma| = 1} \gamma^n dm_f \text{ for all } n \in \mathbb{Z}. \]
If in addition $\mu f = f$, then

$$\|f\|^2_2 = \lim_{n \to \infty} <\mu^n f, f> = \lim_{n \to \infty} \int_{|\gamma|=1} \hat{\mu}^n(\gamma) d\mu_f = \int_{|\gamma|=1} \lim_{n \to \infty} \hat{\mu}^n(\gamma) d\mu_f = m_f\{1\}.$$ 

However, $m_f\{1\} \neq 0$ if and only if $f \circ \tau = f$. So if $\mu f = f$, $f \in L^2(X)$ and $f \neq 0$ then $m_f\{1\} = \|f\|_2^2$ and this implies that $f \circ \tau = f$ and since $\tau$ is ergodic, $f = \text{constant}$. We have proved

$$\{f \in L^2(X) : \mu f = f\} = \{f \in L^2(X) : f \circ \tau = f\} = \{f \in L^2(X) : f \text{ is constant}\}$$

For $p \neq 2$, the statement follows by approximating by functions on $L^2(X)$. Indeed, if $p > 2$, $L^p \subset L^2$. Therefore, $\mu f = f$ implies that $f$ is constant. If $1 \leq p < 2$ and $\mu f = f$, we can assume, without loss of generality, that $\int f d\mu = 0$. Given $\epsilon > 0$ arbitrarily small, there exists $h \in L^2$ such that $\|f - h\|_p \leq \epsilon$. By changing $\epsilon$ for $2\epsilon$, we can assume $\int h d\mu = 0$ since $|\int h d\mu| = |\int (h - f) d\mu| \leq \|h - f\| \leq \epsilon$, so $\|f - (h - \mu h)\|_p \leq \|f - h\|_p + \|h - \mu h\|_p \leq \|f - h\|_p + \epsilon$. But $\int h d\mu = 0$ implies that $h \in \text{cl}_{\|f\|_p} \{g - \mu g : g \in L^2(X)\}$. Thus $\lim_{n \to \infty} \|\mu^n h\|_2 = 0$. But

$$\epsilon > \|f - h\|_p \geq \|\mu^n(f - h)\|_p = \|f - \mu^n h\|_p \geq \|f\|_p - \|\mu^n h\|_p \geq \|f\|_p - \|\mu^n h\|_2 \to \|f\|_p$$
as $n \to \infty$. Since $\epsilon$ is arbitrary, $f = 0$. 

\[\square\]

NOTE: This proposition is, in fact, true under weaker hypothesis. If $\mu$ only is adapted and $\tau$ is ergodic then $\text{cl} \{f - \mu f : f \in L^p(X)\}$ is the subspace of mean zero functions in $L^p(X)$, for all $1 \leq p < \infty$. (Bellow–Jones–Rosenblatt [3]).
Corollary 2.1.4 If $\mu$ is strictly aperiodic, then for every $f \in L^1(X)$ there exists a unique $\mu$-invariant function $f^* \in L^1(X)$ with $\lim_{n \to \infty} \|\mu^n f - f^*\|_1$. In particular, if $\mu$ is strictly aperiodic and $\tau$ is ergodic, then $\lim_{n \to \infty} \|\mu^n f - \int f dm\|_1 = 0$.

Proof. Since by 2.1.3 $\{f \in L^1(X) : \mu f = f\} \oplus \{f - \mu f : f \in L^1(X)\}$ is dense in $L^1(X)$, it suffices to prove the corollary on this subspace. Let $f = g + (h - \mu h)$ with $\mu g = g$ and $g, h \in L^1(X)$. Therefore,

$$\|\mu^n f - g\|_1 = \int |\mu^n f - g| dm = \int |\mu^n h - \mu^{n+1} h| dm \leq \|\mu^n - \mu^{n+1}\|_{L^1(Z)} \|h\|_1.$$ 

Thus, by Proposition 2.1.2 $\lim_{n \to \infty} \mu^n f = g$ in the $L^1$-norm.

Notice that $\int f dm = \int g dm$ because $\tau$ is measure preserving. Indeed,

$$\int \mu(h) dm = \sum_{k=-\infty}^{\infty} \mu(k) \int h \circ \tau^k dm = \sum_{k=-\infty}^{\infty} \mu(k) \int h dm = \int h dm.$$ 

But if $\tau$ is ergodic, then $g$ must be a constant by Proposition 2.1.3. Moreover $g = \int f dm$ and

$$\lim_{n \to \infty} \mu^n f = \int f dm$$ 

in $L^1$-norm for all $f \in L^1(X)$.

Finally, we obtain

Proposition 2.1.5 If $\mu$ is strictly aperiodic, then for $f$ in a dense subspace of $L^1(X)$,

$$\lim_{n \to \infty} \mu^n f(x)$$ 

exist for almost every $x$. 
PROOF. As in the previous propositions,

\[ \{f - \mu f : f \in L^{\infty}(X)\} + \{f \in L^1(X) : \mu f = f\} \]

is a dense subspace of \( L^1(X) \). Clearly, the limit exists on this subspace since

\[ \|\mu^n(f - \mu f)\|_1 \leq \|\mu^n - \mu^{n+1}\|_{L^1(Z)} \|f\|_\infty \to 0 \]

as \( n \to \infty \). (Proposition 2.1.2).

From this last proposition and the Banach Principle, it follows that in order to obtain convergence for every \( f \in L^1 \), it suffices to prove that the maximal operator

\[ \sup_{n \in \mathbb{N}} |\mu^n f(x)| \]

is of weak type \((1,1)\). That is, it suffices to show that there is a constant \( C > 0 \) so that

\[ m\{x \in X : \sup_{n \in \mathbb{N}} |\mu^n f(x)| > \lambda\} \leq C \frac{\|f\|_1}{\lambda} \]

for all \( f \in L^1(X) \). We prove that this weak \((1,1)\) maximal inequality holds by comparing the distribution of \( \mu^n \) with the \( n^{th} \)-convolution power of the Gaussian distribution.

**Theorem 2.1.6** Let \( \mu \) be a strictly aperiodic probability measure on the integers with \( m_{2+\delta}(\mu) < \infty \) for some \( 0 < \delta \leq 1 \). Then

\[ \sup_{k \in \mathbb{Z}} |\mu^n(k) - \frac{1}{\sigma \sqrt{2\pi n}} \exp\left(-\frac{(k - an)^2}{2\sigma^2 n}\right)| \leq \frac{C}{n^{(1+\delta)/2}} \]

where \( a \) denotes the expected value of \( \mu \), \( a = E(\mu) \).
If the measure has a finite moment higher than three, then it also has a finite third moment, and the theorem applies with $\delta = 1$.

This limit theorem is a classical result in the theory of infinitely divisible distributions. A complete exposition of limit theorems for probabilities in the domain of attraction of infinite divisible distributions can be found in Ibragimov–Linnik [24]. The authors provide a proof of the above theorem for any discrete valued strictly aperiodic measure with a finite third moment and mention that the above result holds for measures with a finite moment $2 + \delta$ for $0 < \delta < 1$. We present here a proof which is based on the techniques used in [24] because the estimates are essential to our purposes.

Several lemmas are required to prove this theorem.

**Lemma 2.1.7** Let $g(x) = \int_{-\infty}^{\infty} e^{ixt}d\nu(t)$ where $\nu$ is a complex-valued measure. Let $0 < \delta \leq 1$. If $\int_{-\infty}^{\infty} |t|^\delta d|\nu| < \infty$ then $g \in \text{Lip}(\delta)$. Moreover,

$$|g(x) - g(y)| \leq 3\int_{-\infty}^{\infty} |t|^\delta d|\nu|(t)|x - y|^\delta.$$  

**Proof.** Fix $N$ to be determined and let $C = \int_{-\infty}^{\infty} |t|^\delta d|\nu|(t)$.

$$|g(x) - g(y)| = \left| \int_{|t| < N} (e^{ixt} - e^{iyt})d\nu(t) \right| + \left| g(x) - \int_{|t| < N} e^{ixt}d\nu(t) \right| + \left| g(y) - \int_{|t| < N} e^{iyt}d\nu(t) \right|.$$

The last two terms can be estimated by

$$\left| g(x) - \int_{|t| < N} e^{ixt}d\nu(t) \right| = \int_{|t| \geq N} |e^{ixt}d\nu(t)| \leq \int_{|t| \geq N} |t|^\delta d|\nu|(t) \leq \frac{1}{N^\delta} \int |t|^\delta d|\nu|(t) = \frac{C}{N^\delta}.$$
But we can estimate
\[ |e^{ixt} - e^{iyt}| = |e^{it(x-y)} - 1| = | \int_0^1 i(t(x-y)e^{i(x-y)ts}ds| \leq |t||x-y|. \]

Then,
\[
\int_{|t|<N} |e^{ixt} - e^{iyt}|d|\nu|(t) \leq \int_{|t|<N} |t|^{\delta}d|\nu|(t) |x-y|
= N^{1-\delta} C|x-y|.
\]

Collecting all the terms, we conclude that
\[ |g(x) - g(y)| \leq \frac{2C}{N^{\delta}} + CN^{1-\delta}|x-y|. \]

By choosing \( N = \lfloor 1/|x-y| \rfloor \) we finally prove that
\[ |g(x) - g(y)| \leq 3C|x-y|^\delta. \]

\[ \blacksquare \]

To prove Theorem 2.1.6 we need to consider \( \mu^n \) as the probability distribution of \( n \) independent identically distributed random variables and use the existing theory of infinitely divisible distributions.

We use the same notation as in Ibragimov-Linnik [24]. Let \( \{X_i\}_{i=1}^\infty \) be a sequence of i.i.d. random variables with distribution \( \mu \); i.e. \( p(X_i = k) = \mu(k) \). Let \( \sigma = \sqrt{\text{var}(X_1)} \) and \( a = E(X_1) \) \( (\sigma^2 = E(X_1^2) - a^2) \). If we let \( S_n = X_1 + \cdots + X_n \), then \( p(S_n = k) = \mu^n(k) \). Let \( F_n \) be the distribution function of \( S_n/\sigma \sqrt{n} \),
\[ F_n(x) = p(S_n \leq x\sigma \sqrt{n}) = \sum_{\{k \in \mathbb{Z} : k \leq x\sigma \sqrt{n}\}} \mu^n(k). \]
Let \( f_n \) be the characteristic function of \( F_n \),
\[
f_n(x) = \int e^{ixt} dF_n(t).
\]

\( S_n/\sigma \sqrt{n} \) is a discrete valued random variable with values in the lattice \((1/\sigma \sqrt{n})\mathbb{Z}\). So we can identify \( F_n \) with \( \sum_{k \in \mathbb{Z}} \mu^n(k) \delta_{k/\sigma \sqrt{n}} \). Thus,
\[
f_n(x) = \sum_{k \in \mathbb{Z}} \mu^n(k) \int \exp(\text{i}xt) \delta_{k/\sigma \sqrt{n}}(dt) = \sum_{k \in \mathbb{Z}} \mu^n(k) \exp(\frac{ixk}{\sigma \sqrt{n}}).
\]

Let \( F \) be the distribution of \( X_1 \) and \( f \) its characteristic function. Then
\[
F(x) = p(X_1 \leq x) = \sum_{\{k \in \mathbb{Z} : k \leq x\}} \mu(k).
\]

So \( F \) is identified with \( \sum_{k \in \mathbb{Z}} \mu(k) \delta_k \). Hence
\[
f(x) = \sum_{k \in \mathbb{Z}} \mu(k) \exp(\text{i}xk).
\]

Since \( \mu^n \) is the \( n^{\text{th}} \) convolution power of \( \mu \), the characteristic function of \( \mu^n \) is \( f^n \). So,
\[
f^n(x) = \sum_{k \in \mathbb{Z}} \mu^n(k) \exp(\text{i}xk).
\]

It follows that \( f_n(x) = f^n(x/\sigma \sqrt{n}) \).

Now, by the inversion formula for Fourier transforms (Rudin [37]), or by a straightforward computation based on the last formula for \( f^n(x) \), we get
\[
\mu^n(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f^n(t) \exp(-\text{i}tk) dt
\]
\[
= \frac{1}{2\sigma \sqrt{n} \pi} \int_{-\sigma \sqrt{n} \pi}^{\sigma \sqrt{n} \pi} \frac{t}{\sigma \sqrt{n}} f^n(\frac{t}{\sigma \sqrt{n}}) \exp(-\frac{ikt}{\sigma \sqrt{n}}) dt.
\]

(iv)
Let $N(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, the density function for the normal distribution with mean zero and variance 1. Then $\phi(x) = e^{-x^2/2}$ is its characteristic function and

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(t)e^{-ixt}dt.$$  \hfill (v)

**Lemma 2.1.8** If $m_{2+\delta}(\mu) = E(|X_1|^{2+\delta}) < \infty$ for some $0 < \delta \leq 1$ and $a = E(X_1)$, there exists $T$ depending only on $m_{2+\delta}(\mu)$ and $\delta$ such that, for all $|t| \leq Tn = T\sigma \sqrt{n}$,

$$|f_n(t) - \phi(t)\exp\left(\frac{ia\sqrt{n}t}{\sigma}\right)| \leq K \frac{|t|^{2+\delta}}{\sigma^{2+\delta}n^{\delta/2}} \exp\left(-t^2/4\right)$$

where $K$ is a universal constant.

Note that if $\mu$ has a higher moment than 3, then it also has a finite third moment and this lemma holds with $\delta = 1$.

**Proof.** Since $m_{2+\delta}(\mu) < \infty$, the characteristic function $f$ of $\mu$ has a continuous first and second derivatives (Chow–Teicher [15], page 272). Since for any $g \in C^1(\mathbb{R})$,

$$g(x) - g(0) = \left(\int_0^1 g'(sx)ds\right)x = \left(\int_0^1 \text{Re}(g')(sx)ds + i \int_0^1 \text{Im}(g')(sx)ds\right)x,$$

we can write

$$f(t) = 1 + f'(0)t + \left(\int_0^1 (1-s)f''(st)ds\right)t^2.$$

Since $m_{2+\delta}(\mu) < \infty$, Lemma 2.1.7 gives that

$$f''(u) = \int_{-\infty}^{\infty} (ix)^2 \exp(-iux)dF(x) \text{ is Lip}(\delta).$$

Moreover, Lemma 2.1.7 also gives the following bound,

$$|f''(u) - f''(0)| \leq 3m_{2+\delta}(\mu)|u|^{\delta} \leq 3m_{2+\delta}(\mu)|u|^{\delta}.$$

Therefore,

$$\left|\int_0^1 (1-s)(f''(st) - f''(0))ds\right| \leq 3m_{2+\delta}(\mu) \int_0^1 (1-s)|s|^{\delta}ds \leq 3m_{2+\delta}(\mu)|t|^{\delta}.$$
Notice that \( f'(0) = i \sum_{k \in \mathbb{Z}} k \mu(k) = a \) and \( f''(0) = -\sum_{k \in \mathbb{Z}} k^2 \mu(k) \). But

\[
\sum_{k \in \mathbb{Z}} k^2 \mu(k) = E(X_1^2) = E((X_1 - a)^2) + a^2 = \sigma^2 + a^2.
\]

Then, \( f \) becomes

\[
f(t) = 1 + f'(0)t + \frac{f''(0)}{2} t^2 + R(t)|t|^{2+\delta} = 1 + iat - \frac{\sigma^2 + a^2}{2} t^2 + R(t)|t|^{2+\delta}
\]

where

\[
R(t) = \begin{cases} 
\frac{1}{|t|^\delta} \int_0^1 (1 - s)[f''(st) - f''(0)]ds & \text{if } t \neq 0 \\
0 & \text{if } t = 0
\end{cases}
\]

By the above considerations, the remainder term \( R(t) \) is bounded by \( |R(t)| \leq Cm_{2+\delta}(\mu) \).

Since \( f \) is continuous and \( f(0) = 1 \), for any \( \varepsilon > 0 \) there exists a small \( T \) such that if \( |t| < T \) then \( |f(t) - 1| < \varepsilon \). Notice that at this step, the choice of \( T \) depends on the derivatives of \( f \) and the bound on \( R \); that is, it depends on \( a \), \( \sigma \) and \( m_{2+\delta}(\mu) \). On that small neighborhood, \( |t| < T \), we can consider \( \log f(t) \). From the above expression,

\[
\log f(t) = iat - \frac{(\sigma^2 + a^2)t^2}{2} + R(t)|t|^{2+\delta} - \left( \frac{iat - \frac{(\sigma^2 + a^2)t^2}{2} + R(t)|t|^{2+\delta}}{2} \right)^2
\]

\[
+ \sum_{k=3}^{\infty} \frac{(-1)^{k+1}}{k} \left( iat - \frac{(\sigma^2 + a^2)t^2}{2} + R(t)|t|^{2+\delta} \right)^k
\]

\[
= iat - \frac{\sigma^2}{2} t^2 + r(t)|t|^{2+\delta},
\]

where

\[
r(t) = R(t) - \frac{|t|^{2-\delta}}{2} \left( \frac{(- (\sigma^2 + a^2) + R(t)|t|^{\delta})^2}{2} \right)
\]

\[
+ i a \frac{t^3}{2} \frac{|t|^{2+\delta}}{2} \left( R(t)|t|^{\delta} \right)
\]

\[
+ \sum_{k=3}^{\infty} \frac{(-1)^{k+1}}{k} \frac{t^k}{|t|^{2+\delta}} \left[ i a - \frac{(\sigma^2 + a^2)t}{2} + R(t)|t|^{1+\delta} \right]^k.
\]
That is, \( r(t) = O(|R(t)|) + O(|t|^{1-\delta}). \) (If \( \delta = 1 \), we understand the last term as \( O(1) \).)

Therefore, since we can take \( T \) as small as we want, there exist constants \( K_1 \) and \( K_2 \) such that if \( T < K_1 \) then \( |r(t)| < K_2 \). Then,

\[
\log f_n(t) = n \log f \left( \frac{t}{\sigma \sqrt{n}} \right) = \frac{ia \sqrt{nt}}{\sigma} - \frac{1}{2} t^2 + r \left( \frac{t}{\sigma \sqrt{n}} \right) \frac{t^{2+\delta}}{n^{\delta/2} \sigma^{2+\delta}}
\]

and

\[
\left| f_n(t) - \exp \left( -\frac{t^2}{2} \right) \exp \left( \frac{ia \sqrt{nt}}{\sigma} \right) \right| = \left| \exp(\log f_n(t)) - \exp \left( -\frac{t^2}{2} \right) \exp \left( \frac{ia \sqrt{nt}}{\sigma} \right) \right|
\]

\[
= \exp \left( -\frac{t^2}{2} \right) \left| \exp \left( \frac{r \left( \frac{t}{\sigma \sqrt{n}} \right) t^{2+\delta}}{n^{\delta/2} \sigma^{2+\delta}} \right) - 1 \right|.
\]

Since for all \( x \) and \( \alpha \in C \), \( |e^{\alpha x} - 1| \leq |\alpha x||e^{\alpha x} - 1| \), we can estimate the above by

\[
\leq \exp \left( -\frac{t^2}{2} \right) \frac{|r \left( \frac{t}{\sigma \sqrt{n}} \right)| |t|^{2+\delta}}{\sigma^{2+\delta} n^{\delta/2}} \exp \left( \frac{|r \left( \frac{t}{\sigma \sqrt{n}} \right)| |t|^{2+\delta}}{\sigma^{2+\delta} n^{\delta/2}} \right)
\]

\[
= \frac{|r \left( \frac{t}{\sigma \sqrt{n}} \right)|}{\sigma^{2+\delta} n^{\delta/2}} |t|^{2+\delta} \exp \left( -\frac{t^2}{2} + \frac{|r \left( \frac{t}{\sigma \sqrt{n}} \right)| |t|^{2+\delta}}{\sigma^{2+\delta} n^{\delta/2}} \right).
\]

Now, if \( |\frac{|t|}{(\sigma \sqrt{n})} \leq T \),

\[
\frac{t^2}{2} - \frac{|r \left( \frac{t}{\sigma \sqrt{n}} \right)| |t|^{2+\delta}}{\sigma^{2+\delta} n^{\delta/2}} > t^2 \left( \frac{1}{2} - \frac{K_1}{\sigma^2} \left( \frac{|t|}{\sigma \sqrt{n}} \right)^\delta \right) > t^2 \left( \frac{1}{2} - \frac{K_1^T}{\sigma^2} \right) > \frac{t^2}{4} \quad \text{(vi)}
\]

if \( T < (\sigma^2/4K_2)^{1/\delta} \). Thus, for such \( T \) and \( |\frac{|t|}{(\sigma \sqrt{n})} \leq T \),

\[
|f_n(t) - \exp \left( -\frac{t^2}{2} \right) \exp \left( \frac{ia \sqrt{nt}}{\sigma} \right)| \leq \frac{K_2 |t|^{2+\delta}}{\sigma^{2+\delta} n^{\delta/2}} \exp \left( -\frac{t^2}{4} \right).
\]

From the bound for \( r(t) \), the above condition for \( T \) and the first choice of \( T \), it follows that \( T \) only depends on \( a, \sigma^2 \) and \( m_{2+\delta}(\mu) \).
Note: The fact that \( \delta > 0 \) is crucial to carry out the estimate (vi) and bound \( |f_n(t) - \exp(-\frac{t^2}{2}) \exp\left(\frac{ia\sqrt{n}t}{\sigma}\right)| \) by a function that is integrable on \( t \). This is an important issue in the next proof.

Proof of Theorem 2.1.6:

From the previous discussion, we can express \( \mu \) and the normal distribution in terms of their Fourier transforms (iv, v) (inversion formula),

\[
\mu^n(k) = \int_{-\pi}^{\pi} f^n(t) e^{-ikt} dt = \frac{1}{\sqrt{n}} \int_{-\alpha}^{\alpha} f^n\left(\frac{t}{\sqrt{n}}\right) \exp\left(-\frac{ikt}{\sigma}\right) dt
\]

and

\[
\frac{1}{\sqrt{2\pi\sigma\sqrt{n}}} \exp\left(-\frac{(k - an)^2}{2\sigma^2n}\right) = \frac{1}{\sigma\sqrt{n}} N\left(\frac{k - an}{\sigma\sqrt{n}}\right)
\]

\[= \frac{1}{\sigma\sqrt{n}} \int_{-\infty}^{\infty} \phi(t) \exp\left(\frac{ia\sqrt{n}t}{\sigma}\right) \exp\left(-\frac{ikt}{\sqrt{n}\sigma}\right) dt.
\]

Using these inversion formulas, we can estimate the difference between \( \mu^n \) and the normal distribution essentially by comparing their characteristic functions.

\[
\mu^n(k) - \frac{1}{\sqrt{2\pi\sigma\sqrt{n}}} \exp\left(-\frac{(k - an)^2}{2\sigma^2n}\right) = \frac{1}{\sigma\sqrt{n}} \int_{-T_n}^{T_n} \{f^n\left(\frac{t}{\sqrt{n}}\right) - \phi(t) \exp\left(\frac{ia\sqrt{n}t}{\sigma}\right)\} \exp\left(-\frac{ikt}{\sigma\sqrt{n}}\right) dt
\]

\[+ \frac{1}{\sigma\sqrt{n}} \int_{T_n < |t| < \sigma\sqrt{n}} \left(f^n\left(\frac{t}{\sqrt{n}}\right) \exp\left(-\frac{ikt}{\sigma\sqrt{n}}\right) dt
\]

\[+ \frac{1}{\sigma\sqrt{n}} \int_{|t| > T_n} \phi(t) \exp\left(-\frac{i(k - an)t}{\sigma\sqrt{n}}\right) dt
\]

\[= \frac{1}{\sigma\sqrt{n}} \int_{-T_n}^{T_n} \{f_n(t) - \phi(t) \exp\left(\frac{ia\sqrt{n}t}{\sigma}\right)\} \exp\left(-\frac{ikt}{\sigma\sqrt{n}}\right) dt
\]

\[+ \int_{T_n < |t| < \sigma\sqrt{n}} f^n(t) e^{-ikt} dt + \frac{1}{\sigma\sqrt{n}} \int_{|t| > T_n} \phi(t) \exp\left(-\frac{i(k - an)t}{\sigma\sqrt{n}}\right) dt
\]

\[= I_1 + I_2 + I_3.
\]
where $I_1$, $I_2$ and $I_3$ are the integrals in the above line in the same order.

I. Estimation of $I_1$:

By Lemma 2.1.8, if $|t| \leq T_n$,

$$|f_n(t) - \phi(t)\exp(i\sqrt{n}t)\exp(-\frac{ikt}{\sigma})| \leq K\frac{|t|^{2+\delta}}{\sigma^{2+\epsilon}n^{\delta/2}}\exp(-\frac{t^2}{4}).$$

So

$$\left|\int_{-T_n}^{T_n} (f_n(t) - \phi(t)\exp(i\sqrt{n}t)\exp(-\frac{ikt}{\sigma}))\exp(-\frac{ikt}{\sigma\sqrt{n}})dt\right| \leq \frac{K}{\sigma^{2+\delta}n^{\delta/2}}\int_{-T_n}^{T_n} |t|^{2+\delta}\exp(-\frac{t^2}{4}) < \frac{K'}{n^{\delta/2}}$$

because the integrand is an integrable function on $\mathbb{R}$. We can then estimate $I_1$ by

$$|I_1| \leq \frac{K''}{n^{(1+\delta)/2}}.$$

II. Estimation of $I_2$:

Since $\mu$ is strictly aperiodic, $|f(t)| < 1$ for all $t \neq 0$, $t \in [-\pi, \pi]$. By compactness of $T \leq |t| \leq \pi$, there exists $0 < a < 1$ such that $|f(t)| \leq a$ for all $t \in [-\pi, -T] \cup [T, \pi]$. Therefore,

$$|I_2| \leq 2(\pi - t)a^n \to 0$$

as $n \to \infty$ faster than any power of $n$.

III. Estimation of $I_3$:

$$|I_3| \leq \frac{1}{\sigma\sqrt{n}} \int_{|t| > T\sqrt{n}} e^{-t^2/2}dt = \int_{|t| > T} e^{-(na^2t^2/4)}e^{-(na^2t^2/4)}dt$$

$$\leq e^{-T^2\sigma^2n/4} \int_{|t| > T} e^{-t^2na^2/4}dt < Ce^{-T^2\sigma^2n/4} \to 0$$
as $n \to \infty$ faster than any power of $n$.

From I, II and III, we obtain

$$\sup_{k \in \mathbb{Z}} |\mu^n(k) - \frac{1}{\sqrt{2\pi}\sqrt{n}} \exp\left(-\frac{(k - an)^2}{2\sigma^2n}\right)| \leq \frac{c}{n^{(1+\delta)/2}}$$

for some constant $c$.

\[\blacksquare\]

**Note:** (1) From the proof of the above Theorem 2.1.6, we see that the behavior of the characteristic functions near 0 determines the rate of decay of the difference between the convolution powers of $\mu$ and the convolution powers of the normal distribution. Formally, one might think that if $\mu$ had a moment higher than 3 this decay would be faster. But this is not the case, since in Lemma 2.1.8 the technique is to use Taylor's formula

$$f(t) = f(0) + f'(0)t + \frac{f''(0)}{2}t^2 + R(t)t^{3+\delta}$$

where $0 < \delta \leq 1$ and $m_{2+\delta}(m) < \infty$. If we had a higher finite moment of $\mu (3 + \delta)$, $f$ would have a third derivative and we could write

$$f(t) = f(0) + f'(0)t + \frac{f''(0)}{2}t^2 + R(t)t^{3+\delta}. $$

In order to capitalize on this higher finite moment of $\mu$ we would then need $f'''(0) = 0$. Otherwise we would not be able to compare the powers of $\mu$ with those of the normal distribution. Symmetric measures with a finite third moment have a vanishing third derivative but we want to avoid, if possible, imposing such a restrictive condition.

(2) Another issue is to see what the behavior of the differences is when $\delta = 0$. From
the above proof it is not clear whether there is an estimate of order \( n^{-1/2} \) for measures with only a finite second moment. According to the note following Lemma 2.1.8, this technique can not give such an estimate. Instead, Chung and Erdős [16] have the following result.

**Lemma 2.1.9** Let \( \mu \) be a strictly aperiodic probability measure on \( \mathbb{Z} \) such that \( \{ k : \mu(k) > 0 \} \) do not have all the same sign. If \( m_1(\mu) < \infty \), \( \sup_{k \in \mathbb{Z}} \mu^n(k) \leq C/\sqrt{n} \) where \( C \) does not depend on \( n \).

Now we are ready to prove that the maximal function (i) is a weak \((1,1)\) operator.

**Theorem 2.1.10** If \( \mu \) has finite support and \( E(\mu) = 0 \), then \( \sup_{n \in \mathbb{N}} |\mu^n f(x)| \) is a weak \((1,1)\) operator.

**Proof.** Suppose \( \text{supp}(\mu) \subseteq [-N, N] \), for some positive integer \( N \). Let \( \phi_n \) be a discrete version of the Gaussian distribution, i.e.

\[
\phi_n(k) = \begin{cases} 
\frac{1}{\sigma \sqrt{2\pi n}} \exp\left(-\frac{k^2}{2\sigma^2 n}\right) & \text{for } k \in [-nN, nN] \cap \mathbb{Z} \\
0 & \text{otherwise.}
\end{cases}
\]

Then from Theorem 2.1.6 we get

\[
|\mu^n(k) - \phi_n(k)| \leq \begin{cases} 
\frac{C}{n} & \text{for } k \in [-nN, nN] \\
0 & \text{otherwise.}
\end{cases}
\]

(Theorem 2.1.6 was used with \( \delta = 1 \), since we are assuming that \( \mu \) has finite support and certainly finite third moment). In other words,

\[
|\mu^n(k) - \phi_n(k)| \leq \frac{C}{n} \sum_{j=-nN}^{nN} \delta_{(j)}(k) = C\frac{2nN + 1}{n} A_{nN}(k)
\]
where

\[ A_n = \frac{1}{2n+1} \sum_{j=-n}^{n} \delta(j) \]

is the measure corresponding to the usual averages. Now, since \( A_n \) satisfies a maximal inequality, so does \(|\mu^n - \phi_n|\). Indeed,

\[
m \{ x \in X / \sup_{n \in \mathbb{N}} |(\mu^n - \phi_n)f(x)| > \lambda \} \\
\leq m \{ x \in X / \sup_{n \in \mathbb{N}} |\mu^n - \phi_n||f||x| > \lambda \} \\
\leq m \{ x \in X / \sup_{n \in \mathbb{N}} A_n|f|(x) > \frac{\lambda}{2C(2N + 1)} \} \\
\leq m \{ x \in X / \sup_{n \in \mathbb{N}} A_n|f|(x) > \frac{\lambda}{2C(2N + 1)} \} \\
\leq C'2C(2N + 1) \frac{\|f\|_1}{\lambda}.
\]

Writing \( \mu^n = (\mu^n - \phi_n) + \phi_n \), \( \mu^n \) satisfies a weak (1,1) maximal inequality if \( \phi_n \) does.

Now, \( \phi_n \) can be rewritten as a combination of usual averages:

\[
\phi_n = \sum_{k=0}^{nN} a^n_k A_k \quad \text{where} \quad \sum_{k=0}^{nN} a^n_k \leq 2 \quad \text{and} \quad a^n_k > 0 \quad \text{for} \quad 0 \leq k \leq nN.
\]

This is possible since the \( \phi_n \)'s are symmetric and decreasing in \([0, \infty] \).

\((a^n_k = [\phi_n(k) - \phi_n(k + 1)](2k + 1) \) and \( \phi_n(0) + 2 \sum_{k=1}^{nN} \phi_n(k) < 2 \).

Then, we can estimate

\[
\sup_{n \in \mathbb{N}} |\phi_n f(x)| \leq \sup_{n \in \mathbb{N}} \sum_{k=0}^{nN} a^n_k |A_k f(x)| \leq \sup_{n \in \mathbb{N}} \sum_{k=0}^{nN} a^n_k \sup_{j \in \mathbb{N}} |A_j f(x)| \leq 2 \sup_{n \in \mathbb{N}} |A_n f(x)|.
\]

And consequently,

\[
m \{ x \in X / \sup_{n \in \mathbb{N}} |\phi_n f(x)| > \lambda \} \leq m \{ x \in X / \sup_{n \in \mathbb{N}} |A_k f(x)| > \lambda \} \leq C' \frac{\|f\|_1}{\lambda}.
\]
NOTE: The condition $E(\mu) = 0$ was used to make the $\phi_n$'s and the $A_n$'s centered at the origin. The maximal inequality for either $\phi_n$ or $A_n$ may not hold otherwise.

Theorem 2.1.10 can be extended to measures which do not have finite support. However, in such cases, a moment condition is required. This condition arises from the estimation of large deviation probabilities. There has been a great deal of study of such problems. For our purpose, we will use the following theorem of Baum and Katz [1] which gives the lowest moment condition.

**Theorem 2.1.11 (Baum–Katz [1]).** Let $\{X_i\}$ be a family of i.i.d. random variables with $E(X_1) = 0$. If $t > 1$, $r > 1$ and $1/2 < r/t \leq 1$, the following are equivalent:

(i) $E|X_1|^t < \infty$

(ii) $\sum_{n=1}^{\infty} n^{r-2} P(|S_n| > n^{r/2} \epsilon) < \infty$ for all $\epsilon > 0$.

(See the appendix for a proof of (i) $\Rightarrow$ (ii)).

**Proposition 2.1.12** If $E(\mu) = 0$ and $m_{2+\delta}(\mu) < \infty$ for some $\delta > \sqrt{\frac{\sqrt{2}}{2}}$; then $\sup_{n \in \mathbb{N}} |\mu^n f|$ is a weak $(1,1)$ operator.

**Proof.** Let $\alpha = \frac{1+\delta}{2}$. Split $\mu^n$ into two pieces, $\mu^n = \nu_n + \omega_n$ where $\nu_n = \mu^n|_{[-n^{\alpha},n^{\alpha}]}$ and $\omega_n = \mu^n - \nu_n$. The center piece, $\nu_n$, is handled in the same way as $\mu^n$ was in the previous theorem. Let $\phi_n$ denote the discrete version of the Gaussian distribution,
now with support in $[-n^\alpha, n^\alpha]$. By Theorem 2.1.6
\[
|\nu_n(k) - \phi_n(k)| \leq \frac{1}{n^\alpha}
\]
for all $|k| \leq n^{\alpha}$. Therefore,
\[
|\nu_n f(x)| \leq |\phi_n f(x)| + |\nu_n f(x) - \phi_n f(x)| \leq |\phi_n f(x)| + C A_{n^\alpha} f(x).
\]
This shows that $\sup_{n \in \mathbb{N}} |\nu_n f|$ satisfies the weak $(1,1)$ inequality because $\sup_{n \in \mathbb{N}} |\phi_n f|$ does and
\[
\sup_{n \in \mathbb{N}} |A_{n^\alpha} f(x)| \leq \sup_{n \in \mathbb{N}} |A_n f(x)|.
\]
The tails, $\omega_n$, are controlled by estimating
\[
\sup_{n \in \mathbb{N}} |\omega_n f(x)| \leq \sum_{n=0}^{\infty} |\omega_n f(x)|
\]
and certainly,
\[
m\{x \in X/ \sup_{n \in \mathbb{N}} |\omega_n f(x)| > \lambda\} \leq \frac{1}{\lambda} \|f\|_1 \sum_{n=0}^{\infty} \|\omega_n\|_1.
\]
This is the point where we need the moment condition for $\mu$. Thinking of $\mu^n$ as the distribution of the sum of $n$ i.i.d. random variables, say $\{X_i\}$, each with distribution $\mu$, the $\ell_1$ norms of the tails are
\[
\|\omega_n\|_1 = \sum_{|k| > [n^\alpha]} \mu^n(k) = P(|X_1 + X_2 + \ldots + X_n| > [n^\alpha]) = P(|S_n| > [n^\alpha]).
\]
Therefore, the problem becomes how to estimate the large deviation probabilities. In view of Baum and Katz's theorem, we need to take $r = 2$ and $r/t = \alpha = (1 + \delta)/2$. Therefore, $t = 4/(1+\delta)$. Since we have only $m_{2+\delta}(\mu) < \infty$ we must restrict $t$ such that
t \leq 2 + \delta. That is, \delta must satisfy \(4/(1 + \delta) \leq 2 + \delta\) or, equivalently, \(\delta^2 + 3\delta - 2 \geq 0\).

From this quadratic equation we finally obtain

\[
\delta \geq (\sqrt{17} - 3)/2 \Rightarrow \sum_{n=0}^{\infty} \|\omega_n\|_1 < \infty.
\]

\[\blacksquare\]

REMARK. By Lemma 2.1.9, if the measure has only a finite second moment, we could still carry out the first part of the proof without having to compare the center pieces with the normal distribution. However, the handling of the tail pieces needs a sufficiently high moment condition.

This last proposition is the last piece we need to conclude what we announced in the introductory chapter. Proposition 2.1.5, together with the Banach Principle and Proposition 2.1.12, proves convergence a.e. in \(L^1\).

**Theorem 2.1.13** Let \(\mu\) be a strictly aperiodic probability measure on \(\mathbb{Z}\) with \(E(\mu) = 0\) and \(m_{2+\delta}(\mu) < \infty\) for some \(\delta > (\sqrt{17} - 3)/2\). Then, for any \(f \in L^1(\mathbb{X})\), there exists a unique \(\tau\)-invariant function \(f^* \in L^1(\mathbb{X})\) such that \(\lim_{n \to \infty} \mu^n f(x) = f^*(x)\) a.e.

For measures with a finite moment bigger than two, there is convergence along subsequence.

**Proposition 2.1.14** Let \(\mu\) be a strictly aperiodic probability measure on \(\mathbb{Z}\) with \(m_{2+\delta}(\mu) < \infty\) for some \(\delta > 0\) and \(E(\mu) = 0\). Then, there exists \(k, N \in \mathbb{N}\), \(k = k(\delta)\) and \(N = N(k)\), such that \(\lim_{n \to \infty} \mu^n f(x)\) exists a.e. for all \(f \in L^1(\mathbb{X})\). Moreover, \(\lim_{n \to \infty} \mu^n f(x)\) exists a.e. for all \(f \in L \log^{N-1} L\).
PROOF. Let $\omega_n$ denote the tail of $\mu^n$ as in the previous proposition. We need only to prove the existence of an integer $k = k(\delta)$ such that

$$\sum_{n=1}^{\infty} \|\omega_n^k\|_1 < \infty.$$ 

Since $E(|S_n|^2) \leq nE(|X_1|^2)$, then

$$\|\omega_n^k\|_1 = P(|S_n^k| > (n^k)^{(1+\epsilon)/2})$$

$$\leq \frac{1}{n^kn^{(1+\epsilon)}} E(|S_n^k|^2)$$

$$\leq K \frac{n^k}{n^{k \epsilon}}$$

Therefore, if $k > 1/\delta$ then $\sum_{n=1}^{\infty} \|\omega_n^k\|_1 < \infty$.

For the second statement, observe that for a fixed $k$, there exists $N = N(k)$, such that any positive integer can be written as the sum of at most $N$ powers of $k$, i.e. $n = n_1^k + n_2^k + \cdots + n_r^k$ with $r \leq N$, $n, n_1, \cdots, n_r \in \mathbb{N}$. This is the Waring-Hilbert Theorem (see Hua [23], chapters 18-19). By replacing $f$ with $|f|$, we can assume $f \geq 0$. Then,

$$\sup_{n \in \mathbb{N}} \mu^n f(x) \leq \sup_{n_1 \geq 0} \sup_{n_2 \geq 0} \cdots \sup_{n_N \geq 0} \mu^{n_1} \mu^{n_2} \cdots \mu^{n_N} f(x).$$

But for any $M$, if $f \in L \log^M L$ then $\sup_{n \in \mathbb{N}} A_n f \in L \log^{M-1} L$ where $A_n f$ denotes the usual average as in Theorem 2.1.10. (See Krengel [28], page 54). Let $\nu_n$ be the central part of $\mu^n$ as in the previous proposition. Then $\sup_{n \in \mathbb{N}} \nu_n f$ is dominated by $\sup_{n \in \mathbb{N}} A_n f$ and $\sup_{n \in \mathbb{N}} \omega_n f$ is integrable. Therefore, if $f \in L \log^{N-1} L$, then

$$\sup_{n_2 \geq 0} \cdots \sup_{n_N \geq 0} \mu^{n_2} \cdots \mu^{n_N} f \in L^1.$$
and

\[ \int \sup_{n_2 \geq 0} \cdots \sup_{n_N \geq 0} \mu^{n_2} \cdots \mu^{n_N} f(x) \, dx \leq C \int f(x) \log^{N-1}(f(x)) \, dx. \]

Consequently,

\[
m\{ x : \sup_{n_1 \geq 0} \sup_{n_2 \geq 0} \cdots \sup_{n_N \geq 0} \mu^{n_1} \mu^{n_2} \cdots \mu^{n_N} f(x) > \lambda \} \\
\quad \leq C \int f(x) \log^{N-1}(f(x)) \, dx.
\]

REMARK. With the same techniques one can prove that if \( \mu \) is a strictly aperiodic probability measure on \( \mathbb{Z} \) with \( m_{2+\delta}(\mu) < \infty \) for some \( \delta > 0 \) and \( E(\mu) = 0 \), then \( \lim_{n \to \infty} \mu^{2n} f(x) \) exists a.e. for all \( f \in L^1(X) \).
2.2 Convergence for non-centered measures

There are examples that show that when the measure $\mu$ does not have expected value zero, the whole sequence $\mu^n f(x)$ fails to converge even though the support of the measure generates $Z$. The simplest example is $\mu = \frac{1}{2}(\delta_0 + \delta_1)$ (see [6]). However, even in such cases, convergence results may be obtained by restricting oneself to a subsequence $\mu^{n_k} f(x)$. Results of this type were studied in [4] and [6]. In [4] the authors show that if $\mu$ has finite support and is strictly aperiodic then $\lim_{n \to \infty} \mu^{2^n} f(x)$ exists almost everywhere for any $f \in L^p(X), 1 < p \leq \infty$. We prove the same result holds for $p=1$ and characterize the sequences $\{n_k\}$ for which $\lim_{K \to \infty} \mu^{n_k} f(x)$ exists almost everywhere for any $f \in L^1(X)$.

Let $\mu$ be a probability measure on the integers. Denote by $\phi_n$ a discrete version of the Gaussian distribution centered at zero; that is,

$$\phi_n(k) = \begin{cases} \frac{1}{\sigma \sqrt{2\pi n}} \exp\left(-\frac{k^2}{2\sigma^2 n}\right) & \text{for } k \in [-n,n] \cap \mathbb{Z} \\ 0 & \text{otherwise.} \end{cases}$$

Throughout this section, $[x]$ denotes the integer part of $x$. Let $a = E(\mu)$ and $\sigma = \sqrt{m_2(\mu)}$.

**Proposition 2.2.1** If $\mu$ is a strictly aperiodic probability measure on $\mathbb{Z}$ with $m_{2+\delta}(\mu) < \infty$ for some $\delta > 0$ and $\{n_k\}_{k=0}^\infty$ is a sequence in $\mathbb{Z}^+$ satisfying $n_{k+1} > \gamma n_k$ for some $\gamma > 1$, then for any $f \in L^1(X)$, $\mu^{n_k} f(x)$ converges a. e. if and only if $\phi_{n_k} f([\lceil a n_k \rceil])$ does.

**Proof.** By calculations using Theorem 2.1.6 and $m_{2+\delta}(\mu) < \infty$ one can prove

$$\|\mu^n - \phi_n * \delta_{[an]}\|_1 \leq \frac{C}{n^\delta}$$
for some $\beta > 0$ (Ibragimov–Linnik [24]). Therefore, for any lacunary sequence,

$$\sum_{k\geq 1} \| \mu^{n_k} - \phi_{n_k} \ast \delta_{[a_n]} \|_1 < \infty$$

and, consequently,

$$m\{ x \in X / \sup_{k \in \mathbb{N}} | \mu^{n_k} f(x) - \phi_{n_k} f(\tau^{[a_n]} x) | > \lambda \} \leq C \frac{\|f\|_1}{\lambda}.$$  (vii)

Bellow-Jones-Rosenblatt [6] have already studied the behavior of $\phi_n f(x)$ along subsequences. Its convergence is related to the behavior of a maximal operator of block averages. We will use here the same notation. Let $\Omega$ be any subset of $\mathbb{Z} \times \mathbb{Z}^+$ and let

$$\Omega_\alpha = \{(z, w) \in \mathbb{Z} \times \mathbb{Z}^+ / \text{there exists } (s, n) \in \Omega \text{ such that } |z - s| \leq \alpha(w - n)\}.$$  (viii)

the cone of aperture $\alpha$ with vertices in $\Omega$. Denote the cross section of $\Omega_\alpha$ by $\Omega_\alpha(n) = \{ z \in \mathbb{Z} / (z, n) \in \Omega_\alpha \}$. Consider the maximal operator

$$M_n f(x) = \sup_{(k, n) \in \Omega} \frac{1}{2n+1} \sum_{s=-n}^{n} f(\tau^{k+s}(s)) = \sup_{(k, n) \in \Omega} A_n f(\tau^k x)$$

There is a close relation between the weak type of the operator $M_n f(x)$ and the growth of the cross sections $\Omega_\alpha(n)$.

**Theorem 2.2.2 (Bellow-Jones-Rosenblatt [6]).**

The maximal operator $M_n f$ is of weak type $(1,1)$ if and only if there exists an $\alpha > 0$ and a positive constant $C < \infty$ such that $|\Omega_\alpha(n)| \leq Cn$ for all $n \geq 1$; in other words, if and only if the cross sections of $\Omega_\alpha$ grow at most linearly.
Then, for the maximal operator with respect to the Gaussian distributions, we have the following relation with the maximal operator on block averages.

**Theorem 2.2.3 (Bellow-Jones-Rosenblatt [6]).**

The operator

\[ N_nf(x) = \sup_{(k,n) \in \Omega} \phi_n f(\tau^k x) \]

is of weak type \((1,1)\) if and only if the operator \(Mnf\) is of weak type \((1,1)\), where

\[ \Omega = \{(k, [\sqrt{n}]): (k, n) \in \Omega\}. \]

From Proposition (2.2.1) and the above two theorems we obtain

**Theorem 2.2.4** Let \(\mu\) be a strictly aperiodic probability measure on \(\mathbb{Z}\) with \(m_{2+\delta}(\mu) < \infty\) for some \(\delta > 0\). Let \(\tau\) be an ergodic measure preserving transformation and \(\Omega = \{(an_k, \sqrt{n_k})\}\), where \(\{n_k\}\) is an increasing sequence with \(n_k \geq \gamma n_{k-1}\) for some \(\gamma > 1\). If \(\Omega_\alpha(n)\) grows linearly for all \(\alpha\), then \(\lim_{k \to \infty} \mu^{\alpha} f(x) = \int f dm\) a.e. for all \(f \in L^1(\mathbb{X})\).

**Proof.** If the cross sections of \(\Omega\) grow linearly, then by Theorem 2.2.2, \(Mnf\) is of weak type \((1,1)\). This implies, by Theorem 2.2.3 that \(\sup_{(k,n) \in \Omega} \phi_n f(\tau^k x)\) is of weak type \((1,1)\). And finally, by proposition 2.2.1, we obtain \(\sup_{k \in \mathbb{N}} \mu^{\alpha} f\) is of weak type \((1,1)\).

The next lemma characterizes the sequences with 2.2.1 for which the cross sections of \(\Omega_\alpha\) grow linearly.
Lemma 2.2.5 If \( \{n_k\}_{k \geq 1} \) is a lacunary sequence, then the cross sections of \( \Omega = \{(an_k, \sqrt{n_k})/k \geq 1\} \) grow linearly if and only if the function \( \Psi(\lambda) = \# \{n_k : \lambda < n_k \leq \lambda^2\} \) is bounded.

Proof. Let \( (n_k)_{k \geq 1} \) be a sequence such that \( n_{k+1} \geq \gamma n_k \) for all \( k \). Without loss of generality we can assume \( a = 1 \).

\( \Rightarrow \) It suffices to consider cones with aperture \( \alpha \leq \gamma - 1 \). Then for any \( \lambda \in \mathbb{N} \), the cones with aperture \( \alpha \) and vertices in \( \Omega = \{(n_k, \sqrt{n_k})/k \geq 1\} \) at points with \( n_k \geq \lambda \) have disjoint cross sections at level \( \lambda \). Indeed, consider \( \lambda < n_{k-1} < n_k \leq \lambda^2 \). By condition (2.2.1), \( n_k - n_{k-1} \geq (\gamma - 1)n_{k-1} \). Two consecutive cones are disjoint at level \( \lambda \) if

\[
\alpha \lambda - \alpha \left(\sqrt{n_k} + \sqrt{n_{k-1}}\right) < \alpha \lambda - \alpha \left(\sqrt{n_k - n_{k-1}}\right)
\]

But \( n_{k-1} > \lambda \) and \( \alpha \leq \gamma - 1 \). Thus the cross sections in consideration are disjoint.

Now we can estimate the size of the whole cross section of \( \Omega_\alpha(\lambda) \). The contribution of the cones with vertices corresponding to \( n_k \)'s, \( n_{k-1} \leq \lambda \), does not exceed \( 3\lambda \). And, by the above discussion, the remaining part is

\[
2 \sum_{\lambda < n_k \leq \lambda^2} (\lambda - \sqrt{n_k}).
\]

We have that \( |\Omega_\alpha(\lambda)| \leq C\lambda \) for some constant \( C \), therefore

\[
\sum_{\lambda < n_k \leq \lambda^2} (\lambda - \sqrt{n_k}) \leq C\lambda
\]

or equivalently,

\[
\lambda \Psi(\lambda) - \sum_{\lambda < n_k \leq \lambda^2} \sqrt{n_k} \leq C\lambda
\]
where $\Psi(\lambda) = \#\{k : \lambda < n_k \leq \lambda^2\}$. So,

$$
\Psi(\lambda) \leq C + \frac{1}{\lambda} \sum_{\lambda < n_k \leq \lambda^2} \sqrt{n_k}.
$$

(ix)

Let $n_{j_0}$ be the first element of the sequence $\{n_k\}$ with $n_k > \lambda$. Then

$$
\{n_k : \lambda < n_k \leq \lambda^2\} = \{n_{j_0}, n_{j_0+1}, \cdots, n_{j_0+r}\}
$$

where $r = \Psi(\lambda) - 1$. By the hypothesis on the sequence $\{n_k\}$,

$$
n_{j_0} < (1/\gamma)^r n_{j_0+r}, n_{j_0+1} < (1/\gamma)^r n_{j_0+r}, \cdots, n_{j_0+r-1} < (1/\gamma)n_{j_0+r}.
$$

So the left hand side of (ix) is smaller than

$$
C + \frac{1}{\lambda} \sum_{k=0}^{r} \left( \frac{1}{\sqrt{\gamma}} \right)^k \sqrt{n_{j_0+r}} \leq C + \sum_{k=0}^{\infty} \left( \frac{1}{\sqrt{\gamma}} \right)^k = C' < \infty
$$

because $\gamma > 1$ and $n_{j_0+r} < \lambda^2$.

$\Leftarrow$) If $\Psi(\lambda) \leq M$ for all $\lambda$, then, for any $\alpha > 0$ and $\lambda > 0$,

$$
|\Omega_\alpha(\lambda)| \leq (\text{contribution of cones with vertices } \leq \lambda) + 2M\lambda
$$

$$
\leq (2 + 2M)\lambda.
$$

$\blacksquare$

**Corollary 2.2.6** If $\mu$ is a strictly aperiodic probability measure on $\mathbb{Z}$, then

$\lim_{n \rightarrow \infty} \mu^{2^n} f(x)$ and $\lim_{n \rightarrow \infty} \mu^{[2^n/n^p]} f(x)$ exist a.e. for all $f \in L^1(X)$ and all $p \geq 0$; but $\lim_{n \rightarrow \infty} \mu^{2^n} f(x)$ and $\lim_{n \rightarrow \infty} \mu^{[2^n]} f(x)$ fail to exist on a set of positive measure, for some $f \in L^1(X)$. 
PROOF. By the previous lemma, it suffices to look at the function \( \Psi \) defined above. Let \( n_k = 2^{2^k} \). If \( \lambda < n_k \leq \lambda^2 \), then \( \log_2 \log_2 \lambda < k \leq 1 + \log_2 \log_2 \lambda \) (for \( \lambda > 1 \)). Thus, for this sequence, \( \Psi(\lambda) \leq 1 \).

Let \( n_k = [2^{(2^k/p)}] \) for some \( p > 0 \). If \( \lambda < n_k \leq \lambda^2 \), then \( \log_2 \log_2 \lambda < k - p \log_2 k \leq 2 + \log_2 \log_2 \lambda \) (for \( \lambda > 1 \)). Fix \( n_k \) with such property. Then, if \( k > 3/(2^{1/p} - 1) \), \( n_{k+3} > 2 + \log_2 \log_2 \lambda \). Indeed, for such \( k \),

\[
2^{1/p} - 1 > \frac{3}{k} \\
2^{1/p} > 1 + \frac{3}{k} \\
\frac{1}{p} > \log_2(1 + \frac{3}{k}).
\]

So

\[
1 + p \log_2 k > p \log_2(k + 3) = p(\log_2 k + \log_2(1 + \frac{3}{k})) \\
k + 3 - p \log_2(k + 3) > (k - p \log_2 k) + 2 > \log_2 \log_2 \lambda + 2
\]

Therefore, there exists \( k(p) \) such that for all \( k \geq k(p) \), \( \Psi(\lambda) \leq 3 \).

Let \( n_k = 2^{2^k} \). If \( \lambda < n_k \leq \lambda^2 \), then \( \sqrt{\log_2 \lambda} < k \leq \sqrt{2 \log_2 \lambda} \) (for \( \lambda > 1 \)). Thus, for this sequence, \( \Psi(\lambda) \sim (\sqrt{2} - 1)^{\sqrt{\log_2 \lambda}} \to \infty \) as \( \lambda \to \infty \).

Let \( n_k = [2^{2^{2^k}}] \). If \( \lambda < n_k \leq \lambda^2 \), then \( \lambda < 2^{2^{2^k}} < \lambda^2 + 1 \) and \( (\log_2 \log_2 \lambda)^2 < k \leq \log_2 \log_2 (\lambda^2 + 1) = (1 + \log_2 \log_2 \lambda)^2 + \log_2 (1 + \log_2 (1 + \frac{1}{\lambda})) \) (for \( \lambda > 1 \)). Thus, for this sequence, \( \Psi(\lambda) \sim 1 + 2 \log_2 \log_2 \lambda \to \infty \) as \( \lambda \to \infty \).

\( \square \)

NOTE: In a similar way, if \( \mu \) is strictly aperiodic, one can show that \( \lim_{n \to \infty} \mu^{[a^n]} f(x) \) exists a.e. for all \( a, b > 1 \) and all integrable functions. But for all \( \alpha < 1 \), there exists
some integrable function $f$ for which $\lim_{n \to \infty} \mu^{2^na} f(x)$ fails to exists on a set of positive measure.

2.3 On the maximal function.

It is a well known fact in ergodic theory that for the usual averages $A_nf(x) = (1/(2n+1)) \sum_{k=-n}^{n} f(\tau^k x)$, the maximal operator $\sup_{n \in \mathbb{N}} |A_n f|$ is an operator of strong type $(p,p)$ for all $p > 1$, and of weak type $(1,1)$. However, the maximal function is not always integrable. It is natural to ask when is it integrable. It is a classical result that, for an ergodic $\tau$,

$$\sup_{n \in \mathbb{N}} A_n |f| \in L^1(X) \Leftrightarrow f \in L \log L.$$

For a complete exposition of this result, we refer the reader to Krengel [28]. For the operators $\mu^n f$, we can prove similar results. Bellow–Jones–Rosenblatt [3] proved that whenever $\mu$ is a strictly aperiodic centered probability measure on $\mathbb{Z}$ with $\mu_2(\mu) < \infty$, the maximal function $\sup_{n \in \mathbb{N}} |\mu^n f|$ is of strong type $(p,p)$ operator for all $p > 1$.

Under the same hypothesis of Theorem 2.1.13, we show that the maximal function is also a weak $(1,1)$ operator. As for the usual averages, we have

**Proposition 2.3.1** Let $\mu$ be a strictly aperiodic, centered probability measure on $\mathbb{Z}$, with $\mu_{2+\delta}(\mu) < \infty$ for some $\delta > 0$; and let $\tau$ be an ergodic measure preserving invertible transformation. Then:

(a) $\sup_{n \in \mathbb{N}} \mu^n |f| \in L^1(X) \Rightarrow f \in L \log L$ and

(b) $f \in L \log L$ and $\delta > \frac{\sqrt{17} - 3}{2} \Rightarrow \sup_{n \in \mathbb{N}} |\mu^n f| \in L^1(X)$. 
PROOF. We use the same notation as in Theorem 2.1.10 and Proposition 2.1.12.

(a): From Theorem 2.1.6,

\[ |\mu^n(k) - \phi_n(k)| \leq \frac{C}{n^{(1+\delta)/2}} \]

for all k. Then

\[ \mu^n(k) > \phi_n(k) - \frac{C}{n^{(1+\delta)/2}}. \]  

Let \( n_0 \) be fixed, to be determined later. Then, for \( |k| \leq \sqrt{n} \),

\[ \mu^n(k) > \phi_n(\sqrt{n}) - \frac{C}{n^{(1+\delta)/2}} \]

\[ = \frac{\exp\left(-\frac{1}{2\sigma^2}\right)}{\sqrt{2\pi \sigma}} - \frac{C}{n^{(1+\delta)/2}} > 0 \]  

if \( n \geq n_0 \) and \( n_0 \) is chosen such that

\[ \frac{\exp\left(-\frac{1}{2\sigma^2}\right)}{\sqrt{2\pi \sigma}} > \frac{C}{n_0^{\delta/2}} \]

i.e. \( n_0 > [C \exp\left(\frac{1}{2\sigma^2}\right)\sqrt{2\pi \sigma}]^{2/\delta} \). Hence, for all \( n \geq n_0 \), \([-\sqrt{n}, \sqrt{n}] \subset \text{supp}(\mu^n) \) and, for \( f \geq 0 \),

\[ \mu^n f(x) > \sum_{-\sqrt{n}}^{\sqrt{n}} \mu^n(k) f(\tau^k x) > \left( \min_{k \in [-\sqrt{n}, \sqrt{n}]} \mu^n(k) \right) \sum_{-\sqrt{n}}^{\sqrt{n}} f(\tau^k x) \geq C_n A_{[n]} f(x) \]

where

\[ C_n = \sqrt{n} \left( \min_{k \in [-\sqrt{n}, \sqrt{n}]} \mu^n(k) \right). \]

By (x) and (xi),

\[ C_n > \frac{\exp\left(-\frac{1}{2\sigma^2}\right)}{\sqrt{2\pi \sigma}} - \frac{C}{n_0^{\delta/2}} \geq \frac{\exp\left(-\frac{1}{2\sigma^2}\right)}{\sqrt{2\pi \sigma}} - \frac{C}{n_0^{\delta/2}} = K. \]
Therefore,

\[
\sup_{n \geq n_0} \mu^n f > K \sup_{n \geq n_0} A_{\sqrt{n}} f > K \sup_{n \geq \sqrt{n_0}} A_n f.
\]

Then

\[
\sup_{n \in \mathbb{N}} \mu^n f \in L^1(X) \Rightarrow \sup_{n \in \mathbb{N}} A_n f \in L^1(X)
\]

and since \( \tau \) is ergodic,

\[
\sup_{n \in \mathbb{N}} A_n f \in L^1(X) \Rightarrow f \in L \log L.
\]

(b): We can split \( \mu \) into \( \mu^n = \nu_n + \omega_n \), where \( \nu_n = \mu^n|_{[-n\alpha,n\alpha]} \) are the center pieces and \( \omega_n = \mu^n|_{[-n\alpha,n\alpha]^{c}} \) are the tails. Here \( \alpha = (1 + \delta)/2 \). As in the proof of Theorem 2.1.10 and Proposition 2.1.12, we have

\[
\sup_{n \in \mathbb{N}} |\nu_n f| \leq c \sup_{n \in \mathbb{N}} |A_n f|.
\]

Thus, if \( f \in L \log L \), then \( \sup_{n \in \mathbb{N}} |\nu_n f| \) is integrable.

The maximal function of the tails is, indeed, integrable because of the condition on \( \delta \) (Theorem 2.1.12)

\[
\|\sup_{n \in \mathbb{N}} |\omega_n f||_1 \leq \sum_{n=1}^{\infty} \|\omega_n||_1 \|f||_1 \leq C \int |f| \log^+ |f| dm.
\]

Thus, if \( f \in L \log L \), then \( \sup_{n \in \mathbb{N}} |\mu^n f| \) is integrable.

\[\blacksquare\]

**Lemma 2.3.2** Let \( \mu \) be a strictly aperiodic probability measure on \( \mathbb{Z} \) with \( E(\mu) = 0 \) and \( m_{2+\delta}(\mu) < \infty \) for some \( \delta \geq 0 \). Then, for all \( \alpha \geq \min(0, (2 - \delta^2 - 3\delta)/(6 + 2\delta)) \),

\[
\sup_{n \in \mathbb{N}} \frac{|\mu^n f|}{n^\alpha}
\]

is a weak \((1,1)\) operator.
Proof. Write \( \mu = \nu_n + \omega_n \) where
\[
\nu_n = \sum_{|k| \leq n^{\alpha+(1+\delta)/2}} \mu^n(k) \delta_k \quad \text{and} \quad \omega_n = \mu^n - \nu_n.
\]

If \( \delta > 0 \), the center pieces are handled as in Proposition 2.1.12 by comparing them to a discrete version of the convolution powers of the normal distribution. Let
\[
\phi_n(k) = \begin{cases} 
\frac{1}{\sigma\sqrt{2\pi n}} \exp\left(-\frac{k^2}{2\sigma^2 n}\right) & \text{for } k \in [-n^{\alpha+(1+\delta)/2}, n^{\alpha+(1+\delta)/2}] \cap \mathbb{Z} \\
0 & \text{otherwise.}
\end{cases}
\]

Then
\[
\left| \frac{\nu_n f - \phi_n f}{n^\alpha} \right| \leq c \frac{1}{n^{\alpha+(1+\delta)/2}} \sum_{|k| \leq n^{\alpha+(1+\delta)/2}} |f \circ \tau^k|.
\]

Thus, as before, \( \sup_{n \in \mathbb{N}} |\nu_n f - \phi_n f|/n^\alpha \) is a weak \((1,1)\) operator. Also, since the \( \phi_n/n^\alpha \) are symmetric and decreasing on \((0, \infty)\), \( \sup_{n \in \mathbb{N}} |\phi_n f|/n^\alpha \) is a weak \((1,1)\) operator too.

If \( \delta = 0 \), there is no need to compare the center pieces with the normal distributions because one can apply Lemma 2.1.9 to get the same result.

The tail pieces are also handled as in Proposition 2.1.12.
\[
\left\| \sup_{n \in \mathbb{N}} \frac{|\omega_n f|}{n^\alpha} \right\|_1 \leq \sum_{n=1}^{\infty} \frac{1}{n^\alpha} \left\| \omega_n \right\|_1 \left\| f \right\|_1
\]

and
\[
\sum_{n=1}^{\infty} \frac{1}{n^\alpha} \left\| \omega_n \right\|_1 = \sum_{n=1}^{\infty} n^{-\alpha} p(|S_n| > n^{\alpha+(1+\delta)/2}).
\]

Applying Theorem B.0.10, if we take \( r \geq 2 - \alpha \) and \( r/t = \alpha + (1+\delta)/2 \), the sum is finite. This yields the following conditions on \( \alpha \),
\[
\frac{2 - \delta^2 - 3\delta}{6 + 2\delta} \leq \alpha \leq \frac{1 - \delta}{2}
\]
and $\alpha \geq 0$. But the condition $\alpha > (1 - \delta)/2$ can be drop, since in that case

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha} \|\omega_n\|_1 \leq \sum_{n=1}^{\infty} \frac{1}{n^{(1-\delta)/2}} \|\omega_n\|_1.$$ 

It is worth to notice that, when $\delta = 0$, $\alpha$ can be made smaller that $1/2$, which is the factor that, after Theorem 2.1.6 and Lemma 2.1.9, one would have expected to be the right normalization.

**Corollary 2.3.3** Let $\mu$ be a strictly aperiodic probability measure in $\mathbb{Z}$ with $m_{2+\delta}(\mu) < \infty$ for some $\delta \geq 0$. Then $\lim_{n \to \infty} \mu^n f(x)/n^\alpha = 0$ for all $f \in L^1(X)$ and all $\alpha \geq \min(0, (2 - \delta^2 - 3\delta)/(6 + 2\delta)).$

**Proof.** The statement is true for all $f \in L^\infty(X)$ and, by Lemma 2.3.2, we have a maximal inequality. Therefore the corollary follows from the Banach Principle.

$\blacksquare$
CHAPTER III

Extensions

The methods employed in Chapter 1 can be used to extend the results to locally compact subgroups of $\mathbb{R}^d$.

Let $G$ be a locally compact abelian group (LCA group) with Haar measure $\lambda_G$. Slightly abusing the notations, we denote by $|A|$ the Haar measure of a Borel subset $A \subset G$, and we also use $dg$ instead of $d\lambda_G(g)$. Let $(X, \beta, m)$ be a probability space such that $G$ acts on $X$ by a jointly measurable transformation $G \times X \rightarrow X$, $(g, x) \mapsto T_g x$.

Each $T_g : x \mapsto gx$ $(g \in G)$ is a measure preserving invertible transformation and $T_g T_h = T_{g+h}$, $T_0 = Id$. Let $\mu$ be a regular probability measure on $G$ and define the weighted averages

$$\mu^n f(x) = \int_G f(T_g x) d\mu^n(g).$$

In case $G$ is discrete, these averages take the form

$$\mu^n f(x) = \sum_{g \in G} \mu^n(g) f(T_g x).$$

As in the previous case, these operators are positive contractions in all $L^p(X)$, $1 \leq p \leq \infty$. Indeed

$$\int |\mu^n f|^p dm(x) \leq \int_G \int |f(T_g x)|^p dm(x) d\mu^n(g) = \|f\|_p^p.$$
Definition 3.0.4 Let \( \mu \) be a probability measure on \( G \). We say that \( \mu \) is adapted if its support generates \( G \), and that \( \mu \) is strictly aperiodic if its support is not contained in a coset of a proper closed subgroup of \( G \). Also, \( \mu \) is called spread out if there exists \( n \) such that \( \mu^n \) and \( \lambda_G \) are not mutually singular.

Notice that if \( G \) is discrete, then \( \mu \) is always spread–out.

For subgroups of \( \mathbb{R}^d \), we can define the expected value and the moments of a measure. Let \( \|x\| \) denote the Euclidean norm in \( \mathbb{R}^d \) and \((x,y)\) the standard dot product in \( \mathbb{R}^d \).

Definition 3.0.5 If \( G \) is a subgroup of \( \mathbb{R}^d \) and \( \alpha > 0 \), the \( \alpha \)-moment of \( \mu \) is defined as \( \int_G \|g\|^\alpha d\mu(g) \), and is denoted by \( m_\alpha(\mu) \). If \( g_i \) is the \( i \)-th coordinate of \( g \in \mathbb{R}^d \), \( g = (g_1, \ldots, g_d) \), define \( a_i = \int_G g_i d\mu(g) \) to be the partial expected values of \( \mu \). Then define \( a = (a_1, \ldots, a_d) \) as the vector–expected value of \( \mu \). We say that \( \mu \) is centered if \( a = 0 \).

NOTATION: If \( \zeta \in \mathbb{R} \) and \( x \in \mathbb{R}^d \) then \( \zeta x \) denotes \((\zeta x_1, \ldots, \zeta x_d)\).

Proposition 3.0.6 (FOGUEL [21]). Let \( \mu \) be a probability measure on \( G \) and \( \Gamma \) the Pontryagin dual group of \( G \).

1. If \( \mu \) is strictly aperiodic and spread out then \( \lim_{n \to \infty} \|\mu^{n+1} - \mu^n\|_{L^1(G)} = 0 \).

2. \( \mu \) is strictly aperiodic if and only if \( |\hat{\mu}(\lambda)| < 1 \) for all \( \lambda \in \Gamma \) with \( \lambda \neq 1 \).
Proposition 3.0.7 If $\mu$ is a strictly aperiodic, spread out probability measure on a LCA group $G$ then, for all $1 \leq p$ and all $f \in L^p(X)$, \( \lim_{n \to \infty} \|\mu^n f - P_t f\|_p = 0 \), where $P_t$ is the projection onto the $G$-invariant functions, and \( \lim_{n \to \infty} \mu^n f = P_t f \) exists a.e. for $f$ in a dense subspace of $L^p(X)$.

**Proof.** By Lemma 1.1.1, \( \{f \in L^p(X) : \mu f = f\} + \text{cl} \{f - \mu f : f \in L^p(X)\} \) is a dense subspace of $L^p(X)$. Therefore, the first statement in Proposition 2.1.3 holds in this context. Since $\mu$ is spread out, then the first statements in Corollary 2.1.4 and 2.1.5 also hold, giving convergence in norm and in a dense subspace of $L^p(X)$.

From this proposition and the Banach Principle, it follows that to prove an a.e. convergence theorem on $L^1(X)$, we only need to prove that the maximal operator $\sup_{n \in \mathbb{N}} |\mu^n f(x)|$ is of weak type $(1,1)$.

### 3.1 Probability measures on $\mathbb{R}^d$

In this section we consider $G = \mathbb{Z}^d$ or $\mathbb{R}^d$ and we prove that under suitable conditions on the measure $\mu$, the weighted averages $\mu^n f$ converges a.e. in $L^1(X)$. The idea of proof is basically the same one we used for the group $G = \mathbb{Z}$. The dimension of the group changes the problem and forces stronger conditions on the measure.

We intend to compare the convolution powers of $\mu$ with those of an appropriate version of the normal distribution on $\mathbb{R}^d$. Repeating the same argument of the previous chapter, we prove first the appropriate limit theorem 2.1.6 for $d$-dimensional measures. We recur to some elements of probability theory.
Let \( \{X_i\}_{i=1}^{\infty} \) be a sequence of i.i.d. random variables on \( G \) with distribution \( \mu \); i.e. 
\[ p(X_i \in A) = f_A d\mu(g) . \]
If we let \( S_n = X_1 + \cdots + X_n \), then 
\[ p(S_n \in A) = f_A d\mu^n(g) . \]
Let \( F \) be the distribution of \( X_1 \) and \( f \) its characteristic function, defined by 
\[ f(x) = \int_G \exp(i(x.g)) d\mu(g) \]
for all \( x \in \mathbb{R}^d \). Define also \( f_n(x) = f^n(x/\sqrt{n}) \), the characteristic function of the distribution corresponding to the random variable \( S_n/\sqrt{n} = (X_1 + \cdots + X_n)/\sqrt{n} \).

Notice that, unlike the definition of \( S_n \) in the previous chapter, the normalizing factor is simply \( \sqrt{n} \), without any variance-type factor. The reason is that now we have a covariance matrix instead which will play its role. However, we include it in the definition of the multidimensional normal distribution.

Assume that \( m_2(\mu) < \infty \). Then \( f \) is twice differentiable (all second order partial derivatives exist and are continuous). Let \( B \) be the symmetric matrix formed by their values at \( 0 \),
\[ B_{ij} = \int_{x \in G} x_i x_j d\mu(x) = b_{i,j} ; \]
and let \( A \) be the symmetric matrix determined by the first partial derivatives evaluated at \( 0 \),
\[ A_{ij} = \left[ \int_{x \in G} x_i \mu(x) \right] \left[ \int_{x \in G} x_j d\mu(x) \right] . \]
With the definition we used in the introduction, \( A_{i,j} = a_i a_j \). To avoid new notations, we also denote by \( B \) the quadratic form determined by \( B \), and by \( \hat{B} \), the quadratic form defined by \( B-A \).

If, in addition, \( m_3(\mu) < \infty \), then let \( C \) be the cubic form defined by the third order
partial derivatives of $\mu$ evaluated at 0,

$$C(x) = \sum_{1 \leq j_1, j_2, j_3 \leq d} \left[ \int_{y \in G} y_{j_1} y_{j_2} y_{j_3} d\mu(y) \right] x_{j_1} x_{j_2} x_{j_3}.$$  

Notice that if $\mu$ is symmetric ($\mu(A) = \mu(-A)$), then $C = 0$. If we assume that $\mu$ is adapter to $G$, then the form $\hat{B}$ is positive definite (Appendix C, Lemma C.0.11), all its eigenvalues are positive. Therefore we can define

$$N(x) = \frac{1}{(2\pi)^{d/2} \eta} \int_{\mathbb{R}^d} \exp(-i(x,t)) \exp(-\frac{\hat{B}(t)}{2}) dt$$  

where $\eta$ is an appropriate constant that makes $N$ into a probability measure on $\mathbb{R}^d$. $N$ is the "Multivariate Normal Distribution" whose associated matrix is $\hat{B}$ (Karlin [26], Breiman [12]). Since $B > 0$, $N(x) = C \exp(-D(x)/2)$ for some positive definite quadratic form $D$ (Appendix C, Lemma C.0.12).

**NOTE:** Since $\hat{B}$ is positive definite it defines a norm in $\mathbb{R}^d$. Therefore, $\sqrt{\hat{B}(x)}$ is equivalent to the usual norm $\|x\|$. That is, there exist two positive constants $b_1, b_2$ such that

$$b_1 \leq \frac{\hat{B}(x)}{\|x\|^2} \leq b_2$$

for all $x \in \mathbb{R}^d$, $x \neq 0$.

We now prove the d-dimensional version of Lemma 2.1.8. For technical reasons that will appear later on, we are interested in estimates for measures with a moment condition higher than three. Notice that in the formulas below, we do not make reference to the variance of $\mu$ because it already is incorporated into the definition of $\hat{B}$. 
Lemma 3.1.1 Let \( \mu \) be a strictly aperiodic, spread out probability measure on \( G \) and assume that \( m_\delta(\mu) = \mathbb{E}(\|X_1\|^{2+\delta}) < \infty \) for some \( \delta \in (0,2] \). There exists \( T \) depending only on \( m_{2+\delta}(\mu) \) such that, for all \( \|t\| \leq T_n = T \sqrt{n} \),

\[
|f_n(t) - \exp(-\tilde{B}(t)/2)\exp(i\sqrt{n}(a.t))| \leq K \frac{\|t\|^{2+\delta}}{n^{\delta/2}} \exp(-b_1\|t\|^2/4),
\]

if \( \delta \in (0,1] \), and

\[
|f_n(t) - \exp(-B(t)/2)\exp(i\frac{\|a.t\|}{\sqrt{n}})| \leq K \frac{\|t\|^{2+\delta}}{n^{\delta/2}} \exp(-b_1\|t\|^2/4)
\]

if \( \delta \in (1,2] \). In both cases \( K \) is a universal constant.

PROOF. Since Lemma 2.1.7 still holds in this set-up, we can repeat the proof of Lemma 2.1.8. For any \( g : \mathbb{R}^d \rightarrow \mathbb{C} \) with continuous partial derivatives, one has

\[
g(x) - g(0) = \int_0^1 (\nabla g(sx).x)ds = \int_0^1 (\sum_{j=1}^d \frac{\partial g}{\partial x_j}(sx)x_j)ds.
\]

Thus, we can write

\[
f(t) = \begin{cases}
1 + i(a.t) - \frac{1}{2}B(t) + R(t)\|t\|^2 + \delta & \text{if } \delta \in (0,1] \\
1 + i(a.t) - \frac{1}{2}B(t) - i\delta C(t) + R(t)\|t\|^2 + \delta & \text{if } \delta \in (1,2]
\end{cases}
\]

In both cases, \( |R(t)| \leq Km_{2+\delta}(\mu) \), where \( K \) is a constant independent of \( n \). For \( \epsilon \) fixed, there exists \( T \) such that if \( \|t\| < T \) then \( |f(t) - 1| < \epsilon \) because \( f(0) = 1 \) and \( f \) is continuous. Then, on \( \|t\| < T \), we can consider

\[
\log f(t) = \sum_{k=1}^\infty \frac{(-1)^{k+1}}{k}(f(t) - 1)^k.
\]
From the above expression we get

\[
\log f(t) = \begin{cases} 
  i(a \cdot t) - \frac{1}{2} \tilde{B}(t) + r(t)\|t\|^{2+\delta} & \text{if } \delta \in (0, 1] \\
  i(a \cdot t) - \frac{1}{2} \tilde{B}(t) - \frac{i}{2} \left[ \frac{C(t)}{3} - (a \cdot t)B(t) \right] + r(t)\|t\|^{2+\delta} & \text{if } \delta \in (1, 2].
\end{cases}
\]

where

\[
r(t) = R(t) - \frac{1}{2} \left[ \frac{1}{2} B(t) + R(t)\|t\|^{1+\delta/2} \right]^2 - i \frac{(a \cdot t)}{\|t\|^2} \left[ \frac{B(t)}{2\|t\|^2} + R(t) \right]
\]

\[
+ \sum_{k=3}^{\infty} \frac{(-1)^{k+1}}{k} \frac{\|t\|^k}{\|t\|^{2+\delta}} \left[ i \frac{(a \cdot t)}{\|t\|^2} - \frac{B(t)}{2\|t\|^2} + R(t)\|t\|^{1+\delta} \right]^k
\]

\[
= R(t) + O(\|t\|^{1-\delta})
\]

if \( \delta \in (0, 1] \); and

\[
r(t) = R(t) - \frac{1}{2} \left[ \frac{i}{6} C(t)\|t\|^{1+\delta/2} + R(t)\|t\|^{1+\delta/2} \right]^2
\]

\[
+ i \frac{(a \cdot t)}{\|t\|^2} \left[ \frac{i}{6} C(t)\|t\|^{1+\delta} + R(t)\|t\| \right] + \frac{i}{2} B(t) \left[ \frac{i}{6} C(t)\|t\|^{2+\delta} + R(t) \right] + \frac{1}{8} \frac{B(t)^2}{\|t\|^{2+\delta}}
\]

\[
+ \sum_{k=3}^{\infty} \frac{(-1)^{k+1}}{k} \frac{\|t\|^k}{\|t\|^{2+\delta}} \left[ i \frac{(a \cdot t)}{\|t\|^2} - \frac{B(t)}{2\|t\|^2} + R(t)\|t\|^{1+\delta} \right]^k
\]

\[
= R(t) + O(\|t\|^{2-\delta})
\]

if \( \delta \in (1, 2] \) since \( |C(t)| = O(\|t\|^3) \).

Let \( C'(t) = C(t)/3 - (a \cdot t) B(t) \). There are two constants \( K_1 \) and \( K_2 \) such that if \( T \leq K_1 \) then \( |r(t)| < K_2 \) for all \( \|t\| \leq T \) where \( K_2 \) does not depend on \( T \). Now,

\[
\log f_n(t) = n \log f \left( \frac{t}{\sqrt{n}} \right)
\]

\[
= \begin{cases} 
  i\sqrt{n}(a \cdot t) - \frac{1}{2} \tilde{B}(t) + r(t/\sqrt{n})\|t\|^{2+\delta}/n^{\delta/2} & \text{if } \delta \in (0, 1] \\
  i\sqrt{n}(a \cdot t) - \frac{1}{2} \tilde{B}(t) - \frac{1}{2\sqrt{n}} C'(t) + r(t/\sqrt{n})\|t\|^{2+\delta}/n^{\delta/2} & \text{if } \delta \in (1, 2].
\end{cases}
\]

Let

\[
I = \begin{cases} 
  |f_n(t) - \exp \left( -\frac{\tilde{B}(t)}{2} \right) \exp \left( i\sqrt{n}(a \cdot t) \right) | & \text{if } \delta \in (0, 1] \\
  |f_n(t) - \exp \left( -\frac{\tilde{B}(t)}{2} \right) \exp \left( i\frac{n(a \cdot t) - C'(t)/2}{\sqrt{n}} \right) | & \text{if } \delta \in (1, 2].
\end{cases}
\]
Then

\[ I = \exp\left(-\frac{\mathcal{B}(t)}{2}\right)\exp\left(\frac{r\left(\frac{t}{\sqrt{n}}\right)\|t\|^{2+\delta}}{n^{\delta/2}}\right) - 1. \]

Since for all \(x\), \(|e^{(\alpha \cdot x)} - 1| \leq c\|\alpha\|\|x\|e^{(\alpha \cdot x)}\|\), the above can be estimated by

\[ \leq \exp\left(-\frac{\mathcal{B}(t)}{2}\right)\frac{|r\left(\frac{t}{\sqrt{n}}\right)\|t\|^{2+\delta}}{n^{\delta/2}} \exp\left(\frac{|r\left(\frac{t}{\sqrt{n}}\right)\|t\|^{2+\delta}}{n^{\delta/2}}\right) \]

\[ = \frac{|r\left(\frac{t}{\sqrt{n}}\right)|\|t\|^{2+\delta}}{n^{\delta/2}} \exp\left(-\frac{\mathcal{B}(t)}{2} + \frac{|r\left(\frac{t}{\sqrt{n}}\right)\|t\|^{2+\delta}}{n^{\delta/2}}\right). \]

Now, if \(\|t\|/\sqrt{n} \leq T < (b_1/8K_2)^{1/\delta}\),

\[ \frac{\mathcal{B}(t)}{2} - \frac{|r\left(\frac{t}{\sqrt{n}}\right)|\|t\|^{2+\delta}}{n^{\delta/2}} > \|t\|^2\left(\frac{\mathcal{B}(t)}{2\|t\|^2} - K_2\left(\frac{\|t\|}{\sqrt{n}}\right)^{\delta}\right) \]

\[ > \|t\|^2\left(\frac{b_1}{2} - K_2T^\delta\right) \]

\[ > \frac{b_1\|t\|^2}{4}. \]

Thus, for such \(T\) and \(\|t\|/\sqrt{n} \leq T\), we obtain the desired estimate. By a careful inspection of the estimates, one sees that the choice of \(T\) does not depend on \(n\).

\[ \blacksquare \]

Now we can prove the corresponding version of the limit theorem 2.1.6. Recall that the variance of the measure is incorporated in the definition of multivariate normal distribution \(N\).

Case \(G = \mathbb{Z}^d\):

**Theorem 3.1.2** Let \(\mu\) be a strictly aperiodic probability measure on \(\mathbb{Z}^d\). Assume that \(m_{2+\delta}(\mu) < \infty\) for some \(\delta \in (0, 1]\) or \(\delta \in (1, 2]\) and \(\mu\) is symmetric. Then

\[ \sup_{k \in \mathbb{Z}^d} |\mu^n(k) - \frac{1}{(2\pi n)^{d/2}} \mathcal{N}\left(\frac{k - na}{\sqrt{n}}\right)| \leq \frac{C}{n^{(d+\delta)/2}} \]

\(\text{for some constant } C.\)
PROOF. Note that, because of the conditions on \( \mu \), the term \( C(t)/3 - (a.t)B(t) \) in Lemma 3.1.1 vanishes. By the definition of \( f \), we can express the convolution powers of \( \mu \) in terms of the powers of \( f \):

\[
\mu^n(k) = \int_{[-\pi,\pi]^d} f^n(t) \exp(-it.k) dt \\
= \frac{1}{n^{d/2}} \int_{[-\sqrt{n\pi},\sqrt{n\pi}]^d} f\left(\frac{t_1}{\sqrt{n}}, \ldots, \frac{t_d}{\sqrt{n}}\right)^n \exp\left(-\frac{i(k.t)}{\sqrt{n}}\right) dt \\
= \frac{1}{n^{d/2}} \int_{[-\sqrt{n\pi},\sqrt{n\pi}]^d} f_n(t) \exp\left(-\frac{i(k.t)}{\sqrt{n}}\right) dt;
\]

and by the choice of \( N \),

\[
\frac{1}{n^{d/2}} N\left(\frac{(k - an)}{\sqrt{n}}\right) = \frac{1}{n^{d/2}} \int_{\mathbb{R}^d} \exp\left(-\frac{\hat{B}(t)}{2}\right) \exp(i\sqrt{n}(a.t)) \exp\left(-\frac{i(k.t)}{\sqrt{n}}\right) dt.
\]

Applying Lemma 3.1.1, the rest of the proof follows exactly as in the case \( G = \mathbb{Z} \).

Let \( T_n \) be as in Lemma 3.1.1, then

\[
|\mu^n(k) - \frac{1}{n^{d/2}} N\left(\frac{(k - an)}{\sqrt{n}}\right)| \\
\leq \frac{1}{n^{d/2}} \left| \int_{[-T_n,T_n]^d} f_n(t) - \exp\left(-\frac{\hat{B}(t)}{2}\right) \exp(i\sqrt{n}(a.t)) \right| \exp\left(-\frac{i(k.t)}{\sqrt{n}}\right) dt | \\
+ \frac{1}{n^{d/2}} \left| \int_{\{t < ||t||_{\infty} \leq \sqrt{n}\pi\}} f_n(t) \exp\left(-\frac{i(k.t)}{\sqrt{n}}\right) dt | \\
+ \frac{1}{n^{d/2}} \left| \int_{\{||t||_{\infty} > T_n\}} \exp\left(-\frac{\hat{B}(t)}{2}\right) \exp(i\sqrt{n}(a.t)) \exp\left(-\frac{i(k.t)}{\sqrt{n}}\right) dt | \\
\leq \frac{1}{n^{d/2}} \int_{[-T_n,T_n]^d} |f_n(t) - \exp\left(-\frac{\hat{B}(t)}{2}\right) \exp(i\sqrt{n}(a.t)) |dt \\
+ \int_{\{t < ||t||_{\infty} \leq \pi\}} |f_n(t)|d(t) + \int_{\{||t||_{\infty} > T\}} \exp\left(-\frac{n}{2}\right) dt | \\
= I_1 + I_2 + I_3
\]

where \( I_1, I_2 \) and \( I_3 \) are each of the terms in the above inequality in the given order.
To estimate $I_1$, we have that from Lemma 3.1.1, if $||t|| \leq T_n$,

$$|f_n(t) - \exp(-\tilde{B}(t)/2)\exp(i\sqrt{n}(a.t))| \leq K \frac{||t||^{d+\varepsilon}}{n^{d/2}} \exp(-b_1||t||^2/4).$$

Thus

$$I_1 \leq \frac{K}{n^{(d+\varepsilon)/2}} \int_{[-T_n, T_n]^d} ||t||^{2+\varepsilon} \exp(-b_1||t||^2/4)dt$$

$$\leq \frac{K'}{n^{(d+\varepsilon)/2}}$$

because the integrand is an integrable function on $\mathbb{R}^d$.

To estimate $I_2$, notice that since $\mu$ is strictly aperiodic, $|f(t)| < 1$ if $t \neq 0, t \in [-\pi, \pi]^d$. Then, by compactness of $\{T < ||t||_\infty \leq \pi\}$, there exists $0 < a < 1$ such that $|f(t)| < a$ on that set. Hence

$$I_2 < a^n2^d(\pi - T)^d \to 0$$

faster than any power of $n$.

Finally

$$I_3 = \int_{\{||t||_\infty > T\}} \exp(-n\frac{\tilde{B}(t)}{4})\exp(-n\frac{\tilde{B}(t)}{4})dt$$

$$\leq \exp(-n\frac{\tilde{B}(t_0)}{4}) \int_{\{||t||_\infty > T\}} \exp(-n\frac{\tilde{B}(t)}{4})dt$$

$$\leq c \exp(-n\frac{\tilde{B}(t_0)}{4}) \to 0$$

faster than any power of $n$, and where $t_0$ is a point in the set $\{||t||_\infty = T\}$ at which $\tilde{B}$ attains its minimum.

Therefore,

$$|\mu^n(k) - \frac{1}{n^{d/2}}N\left(\frac{k - an}{\sqrt{n}}\right)| \leq C \frac{1}{n^{(d+\varepsilon)/2}}$$

for some constant $C$ and for all $k$. 

Case $G = \mathbb{R}^d$:

**Theorem 3.1.3** Let $\mu$ be a strictly aperiodic, spread out probability measure on $\mathbb{R}^d$ such that there exists $n_0 \in \mathbb{N}$ for which $d\mu^{n_0}x = \ell(x)dx$ and $\ell \in L^p(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. In case $m_{2+\delta}(\mu) < \infty$ for some $0 < \delta \leq 1$, let $\alpha = (d + \delta)/2$, and in case $m_{3+\delta}(\mu) < \infty$ for some $0 < \delta \leq 1$ and $\mu$ is symmetric, let $\alpha = (d + 1 + \delta)/2$. Then, for any $h \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$,

$$
|\int_{\mathbb{R}^d} h(x)d\mu^n(x) - \frac{1}{(2\pi n)^{d/2}} \int_{\mathbb{R}^d} h(x)N\left(\frac{x-na}{\sqrt{n}}\right)dx| \leq \frac{C}{n^\alpha} \|h\|_1
$$

for some constant $C$.

It is unfortunate that such strong conditions must be imposed on the measure $\mu$ but they are necessary because of technical reasons.

**Proof.** Without loss of generality, we can assume that $\mu$ is absolutely continuous with respect to the Lebesgue measure and that if $\ell$ is the Radon–Nikodym derivative of $\mu$, then $\ell \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. Then, for any $h \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, we can apply the Plancharel's formula (Rudin [37]),

$$
\int_{\mathbb{R}^d} h(x)d\mu^n(x) = \int_{\mathbb{R}^d} \hat{h}(-t)\hat{f}^n(t)dt
$$

(iii)

$$
= \frac{1}{n^{d/2}} \int_{\mathbb{R}^d} \hat{h}\left(-\frac{t}{\sqrt{n}}\right)f\left(\frac{t}{\sqrt{n}}\right)^n dt.
$$

(iv)

The multivariate normal distribution $N$ was defined in (i) in terms of its Fourier transform. Recall that

$$
N(x) = \frac{1}{(2\pi)^{d/2} \eta} \int_{\mathbb{R}^d} \exp(-i(x.t))\exp\left(-\frac{\bar{B}(t)}{2}\right)dt.
$$
Then (see Appendix C), its Fourier transform is

\[ \int_{\mathbb{R}^d} N(x) \exp(i x \cdot t) dx = \exp\left(-\frac{\hat{B}(t)}{2}\right). \]

By applying Plancharel's formula,

\[ \frac{1}{(2\pi n)^{d/2}} \int_{\mathbb{R}^d} h(x) N\left(\frac{x-an}{\sqrt{n}}\right) dx = \int_{\mathbb{R}^d} \hat{h}(t) \exp\left(-\frac{\hat{B}(t)_n}{2}\right) \exp(in(a \cdot t)) dt \]

\[ = \frac{1}{n^{d/2}} \int_{\mathbb{R}^d} \hat{h}\left(\frac{t}{\sqrt{n}}\right) \exp(-\frac{\hat{B}(t)}{2}) \exp(i \sqrt{n}(a \cdot t)) dt. \]

Then

\[ |\int_{\mathbb{R}^d} h(x) \mu^n(x) - \frac{1}{n^{d/2}} \int_{\mathbb{R}^d} h(x) N\left(\frac{x-an}{\sqrt{n}}\right) dx| = \]

\[ |\frac{1}{n^{d/2}} \int_{\mathbb{R}^d} \hat{h}\left(\frac{t}{\sqrt{n}}\right) f^n\left(\frac{t}{\sqrt{n}}\right) - \exp\left(-\frac{\hat{B}(t)_n}{2}\right) \exp(i \sqrt{n}(a \cdot t)) | dt| \]

\[ \leq \frac{1}{n^{d/2}} \int_{\||t|| \leq T \sqrt{n}} |\hat{h}\left(\frac{t}{\sqrt{n}}\right) f^n\left(\frac{t}{\sqrt{n}}\right) - \exp\left(-\frac{\hat{B}(t)_n}{2}\right) \exp(i \sqrt{n}(a \cdot t)) dt| \]

\[ + |\int_{\||t|| > T \sqrt{n}} \hat{h}(t) f^n(t) dt| + |\int_{\||t|| > T \sqrt{n}} \hat{h}(t) \exp\left(-\frac{\hat{B}(t)_n}{2}\right) \exp(in(a \cdot t)) dt| \]

\[ = I_1 + I_2 + I_3. \]

In the first case, put \( r = \delta \) and in the second, \( r = 1 + \delta \). Then, By Lemma 3.1.1,

\[ I_1 \leq \frac{1}{n^{d/2}} \int_{\||t|| \leq T \sqrt{n}} \|\hat{h}\|_{L^\infty} K \|t\|^{2+r} \exp\left(-\frac{b_1 \|t\|^2}{4}\right) \exp(i \sqrt{n}(a \cdot t)) | dt \]

\[ \leq \frac{c}{n^{(d+r)/2}} \|h\|_1. \]

To estimate \( I_2 \), notice that, by the Riemann–Lebesgue Theorem, \( \lim_{|t| \to \infty} \hat{\ell}(t) = 0 \).

Since \( \mu \) is strictly aperiodic, Proposition 2.1.2 also gives \( |f(t)| < 1 \) for all \( t \neq 1 \).

Therefore, there exists \( \epsilon \in (0,1) \) such that \( |f^n(t)| < \epsilon \) for all \( |t| > T \). Then

\[ I_2 \leq \epsilon^{n-2} \int_{\||t|| > T \sqrt{n}} |\hat{h}(t)||f(t)|^2 dt \]

\[ \leq \epsilon^{n-2} \|h\|_1 \int_{\||t|| > T \sqrt{n}} |f(t)|^2 dt \]
But $e^{-2} \to 0$ faster than any power of $n$, in particular, $n^{(d+r)/2}$.

Finally,

\[ I_3 \leq \int_{(\|t\| > T)} |\hat{h}(t)| \exp\left(-\frac{\hat{B}(t)n}{2}\right)dt \leq \|h\|_1 \exp\left(-\frac{n\hat{B}(t_0)}{4}\right) \int_{(\|t\| > T)} \exp\left(-\frac{B(t)n}{4}\right)dt \]

\[ = \|h\|_1 \exp\left(-\frac{n\hat{B}(t_0)}{4}\right)c. \]

where $t_0$ is a point in the set $\{\|t\| > T\}$ where $\hat{B}$ attains its minimum. By combining the estimates for $I_1$, $I_2$ and $I_3$, the theorem is proved.

\[ \blacksquare \]

**Proposition 3.1.4** Let $\mu$ be a strictly aperiodic, spread out probability measure on $\mathbb{Z}^d$ or $\mathbb{R}^d$ with $a = 0$. In the second case, $\mu$ must also satisfy that there exists $n_0 \in \mathbb{N}$ for which $d\mu^{n_0}x = \ell(x)dx$ and $\ell \in L^p(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. If $d$ and $\mu$ satisfy one of the following two cases then $\sup_{n \in \mathbb{N}} |\mu^n f|$ is a weak $(1,1)$ operator.

1. $d = 1, 2$ and $m_2(\mu) < \infty$, for some $\sqrt{(2 + d)^2 - 8d - (d + 2)}/2 \leq \delta$,
2. $d = 3$ and $m_3(\mu) < \infty$,
3. $d > 3$, $\mu$ is symmetric and $m_{3+\delta}(\mu) < \infty$ for some $\delta$ in the range $\left[\sqrt{(2 + d)^2 - 8d - (d + 2)}/2 \leq \delta < 1. \right.$

If the measure has a moment condition higher than the one specified above for each case, then the theorem holds but the technicalities are carried out with a moment condition in the above range.

**Proof.** By the different nature of the groups $\mathbb{Z}^d$ and $\mathbb{R}^d$, we have to treat each separately, although the idea of the proof is the same.
Let
\[ \epsilon = \begin{cases} 
\delta & \text{in case (1),} \\
1 & \text{in case (2),} \\
1 + \delta & \text{in case (3)} 
\end{cases} \]
and \( \alpha = (d + \epsilon)/2d \).

\( G = \mathbb{Z}^d \):

Split \( \mu^n \) into two pieces, \( \mu^n = \nu_n + \omega_n \) where
\[ \nu_n(k) = \begin{cases} 
\mu^n(k) & \text{if } \|k\|_\infty \leq n^\alpha \\
0 & \text{otherwise,} 
\end{cases} \]
\[ \omega_n = \mu^n - \nu_n. \]

Define
\[ N_n(x) = \begin{cases} 
\frac{1}{(2\pi n)^{d/2}} N(\frac{x}{\sqrt{n}}) & \text{if } \|x\|_\infty \leq n^\alpha, x \in \mathbb{Z}^d \\
0 & \text{otherwise.} 
\end{cases} \]

Then, \( N_n \) are discrete versions of \( N \).

By Theorem 3.1.2,
\[ |\nu_n f(x) - N_n f(x)| \leq \frac{C}{n^{(d+\epsilon)/2}} \sum_{\|k\|_\infty \leq n^\alpha} |f(T_k x)| = C A_{[n^\alpha]}|f|(x). \]

\( A_n \) denotes the usual averages on \( \mathbb{Z}^d \),
\[ A_n f(x) = \frac{1}{(2n + 1)^d} \sum_{\|k\|_\infty \leq n} f(T_k x). \]

Therefore, \( \sup_{n \in \mathbb{N}} |(\nu_n - N_n) f| \) is a weak (1,1) operator since
\[ \sup_{n \in \mathbb{N}} |A_{[n^\alpha]} f(x)| \leq \sup_{n \in \mathbb{N}} |A_n f(x)|. \]

To treat \( \sup_{n \in \mathbb{N}} |N_n f| \), recall that \( N \) is of the form \( c \exp(-D(x)/2) \) where \( D \) is a positive definite quadratic form in \( \mathbb{R}^d \). Then \( E = \{ x \in \mathbb{R}^d : D(x) \leq 1 \} \) is an ellipsoid in \( \mathbb{R}^d \). There exists an increasing sequence of positive integers \( \{r_n\} \) such that, if
\[ E_n = r_n E \cap \mathbb{Z}^d = \{r_n x : x \in E\} \cap \mathbb{Z}^d, \]
then $N_n$ can be decomposed as

$$N_n(x) = \sum_{j=1}^{\infty} a_n^j \frac{1}{|E_n|} \chi_{E_n}(x)$$

where $a_n^j = 0$ for all but finitely many $j$'s, $a_n^j \geq 0$ and $\sum_{j=1}^{\infty} a_n^j \leq R$ for some constant $R$ independent of $n$. Thus,

$$\sup_{n \in \mathbb{N}} |N_n f| \leq \sup_{n \in \mathbb{N}} \sum_{j=1}^{\infty} a_n^j \sup_{m \in \mathbb{N}} \frac{1}{|E_m|} \sum_{k \in E_m} |f(T_k x)|$$

$$\leq R \sup_{n \in \mathbb{N}} \frac{1}{|E_n|} \sum_{k \in E_n} |f(T_k x)|.$$

And $\sup_{n \in \mathbb{N}} (1/|E_n|) \sum_{k \in E_n} |f(T_k x)|$ is a weak $(1,1)$ operator. (See Chapter 4, Theorem 4.1.5).

Finally $\sup_{n \in \mathbb{N}} \nu_n f$ is a weak $(1,1)$ operator. It remains to prove the same for the supremum over the tails $\omega_n$. Generally

$$\int \sup_{n \in \mathbb{N}} |\omega_n f(x)| dx \leq \sum_{n=0}^{\infty} \|\omega_n\|_1 |f|_1.$$

Thinking of $\mu^n$ as the distribution of the sum of $n$ i.i.d. $d$-dimensional random variables, say $\{X_i\}$ each with distribution $\mu$, $X_i = (X_{i1}, \ldots, X_{id})$, the 1-norms of the tails are

$$\|\omega_n\|_1 = \sum_{\|k\|_\infty > n^\alpha} \mu^n(k) = P(\|X_1 + X_2 + \ldots + X_n\|_\infty > n^\alpha)$$

$$= P(\|S_n\|_\infty > n^\alpha) = P(\{|S_n^1| > n^\alpha\} \cup \ldots \cup \{|S_n^d| > n^\alpha\})$$

$$\leq \sum_{j=1}^{d} P(|S_n^j| > n^\alpha)$$

where the $S_n^j$ are the coordinates of $S_n$. Therefore, it is enough to estimate large deviation probabilities of each coordinate sum $S_n^j$. In view of Baum and Katz's
theorem, we need to take \( r = 2 \) and \( r/t = \alpha = (d + \epsilon)/2d \). Therefore, take \( t = 4d/(d + \epsilon) \). Since we have only \( m_{2+\epsilon}(\mu) < \infty \), we must assume \( t \leq 2 + \epsilon \). This gives the following condition on \( d \) and \( \epsilon \):

\[
0 \leq \epsilon^2 + (d + 2)\epsilon - 2d.
\]

Let

\[
\Delta = \frac{(2 + d)^2 + 8d - (d + 2)}{2}.
\]

We need \( \epsilon \geq \Delta \). But \( \Delta \) satisfies the following

\[
0 < \Delta \leq 1 \quad \text{if } d = 1, 2;
\]

\[
\Delta = 1 \quad \text{if } d = 3;
\]

\[
1 < \Delta < 2 \quad \text{if } d > 3.
\]

Therefore, \( d \) and \( \delta \) must satisfy

\[
\begin{align*}
\Delta < \delta < 1 & \quad \text{if } d = 1, 2; \\
\delta = 1 & \quad \text{if } d = 3; \\
0 < \Delta - 1 < \delta < 1 & \quad \text{if } d > 3;
\end{align*}
\]

which gives the conditions of the theorem.

Case \( G = \mathbb{R}^d \):

Split \( \mu^n \) into two pieces, \( \mu^n = \nu_n + \omega_n \), where

\[
\nu_n(A) = \mu^n(A \cap \{x : \|x\|_{\infty} \leq n^\alpha\})
\]

\[
\omega_n(A) = \mu^n(A \cap \{x : \|x\|_{\infty} > n^\alpha\})
\]

Define

\[
N_n(x) = \begin{cases} 
(1/n^{d/2})N(x/\sqrt{n}) & \text{if } \|x\|_{\infty} \leq n^\alpha \\
0 & \text{otherwise}.
\end{cases}
\]

For each fixed \( x \in X \), consider the measurable function \( F_x(t) = f(T_t x) \). Then,

\[
F_x x_{\{\|x\|_{\infty} \leq n^\alpha\}} \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)
\]
and
\[ \|F_x \chi_{\{\|t\| \leq n^a\}}\|_1 = \int_{\{\|t\| \leq n^a\}} |F_x(t)| \, dt. \]

By Theorem 3.1.3,
\[ |\int f(T_t x) d\nu_n(t) - \int_{\{\|t\| \leq n^a\}} f(T_t x) N_n f(t) \, dt| \leq \frac{C}{n^{(d+1+\delta)/2}} \int_{\{\|t\| \leq n^a\}} |f(T_t x)| \, dt \]
\[ = CA_{[n^a]} |f(x)|. \]

\( A_n \) denotes the usual averages on \( \mathbb{R}^d \),
\[ A_n f(x) = \frac{1}{m} \int_{\{\|t\| \leq m\}} f(T_t x) \, dt. \]

Therefore, \( \sup_{n \in \mathbb{N}} |(\nu_n - N_n)f| \) is a weak (1,1) operator since
\[ \sup_{n \in \mathbb{N}} |A_{[n^a]} f(x)| \leq \sup_{n \in \mathbb{N}} |A_n f(x)|. \]

To treat \( \sup_{n \in \mathbb{N}} |N_n f| \), decompose \( N_n \) in a similar way as in the previous case. Recall that \( N(x) = e^{\exp(-D(x)/2)} \) for some positive definite quadratic form \( D \). Then, the \( N_n \)'s can be seen as ellipsoid-decreasing functions (as opposed to radially decreasing). As before, \( E = \{ x \in \mathbb{R}^d : D(x) \leq 1 \} \) is an ellipsoid in \( \mathbb{R}^d \). Let \( E_n = n E = \{ x \in \mathbb{R}^d : D(x) \leq n \} \). Then
\[ N_n(x) = N_n(x) \chi_{E_1}(x) + \sum_{k=2}^{\infty} N_n(x) \chi_{E_k \setminus E_{k-1}}(x) \]
\[ \leq \frac{1}{(2\pi n)^{d/2} \eta} [\chi_{E_1}(x) + \sum_{k=2}^{\infty} \exp(-\frac{k-1}{2n}) \chi_{E_k \setminus E_{k-1}}(x)] \]
\[ = \frac{1}{(2\pi n)^{d/2} \eta} (e^{1/2n} - 1) [\sum_{k=1}^{\infty} \exp(-\frac{k}{2n}) \chi_{E_k}(x)] \]
\[ = \sum_{k=1}^{\infty} a_n \frac{1}{|E_k|} \chi_{E_k}(x). \]
Here $a_k^n \geq 0$ for all $k$ and $n$; but only finitely many of them are not 0. Also $\sum_{k=1}^{\infty} a_k^n \leq C$, $C$ does not depend on $n$. Thus,

$$\sup_{n \in \mathbb{N}} |N_n f(x)| \leq \sup_{n \in \mathbb{N}} \int |f(T_k x) N_n(t)| dt \leq C \sup_{n \in \mathbb{N}} \frac{1}{|E_k|} \int_{E_k} |f(T_k x)| dt.$$ 

Since the maximal function on the right hand side is a weak (1,1) operator (See Chapter 4, Theorem 4.1.5), then $\sup_{n \in \mathbb{N}} |N_n f(x)|$ is also weak (1,1).

Finally,

$$\| \sup_{n \in \mathbb{N}} |\omega_n f| \| \leq \| f \|_1 \sum_{n=1}^{\infty} \| \omega_n \|_1.$$ 

And

$$\| \omega_1 \| = \int_{\| \cdot \|_\infty > n^\alpha} d\mu^n(t) = P(\| S_n \|_\infty > n^\alpha).$$

Therefore, as in the case $G = \mathbb{Z}^d$, if $d$ and $\delta$ satisfy one of the hypothesis of the theorem ((1), (2) or (3)), then $\sum_{n=1}^{\infty} \| \omega_n \|_1 < \infty.$
CHAPTER IV

Correct Factor and Amenable Groups.

Let \((X, \beta, m, T)\) be a measure preserving dynamical system. Let \(B_k = [n_{k,1}, n_{k,2}, \cdots, n_{k,k}]\) be blocks of positive integers and consider the averages

\[
A_n f(x) = \frac{1}{r_n} \sum_{u \in B_n} T^u f(x).
\]

It is known that in this generality, the averages \(A_n f\) fail to converge a.e. in \(L^1(X)\). However, Rosenblatt and Wierdl \[36\] proved that if \(Q_n = \bigcup_{k=1}^n (B_n - B_k)\), then the averages

\[
A_{n,|Q_n|} f(x) = \frac{1}{|Q_n|} \sum_{u \in B_n} T^u f(x)
\]

do converge a.e. in \(L^1(X)\). Indeed, they proved that if the blocks \(B_k\) consist of consecutive numbers and their length is non-decreasing, then the \(|Q_n|\)'s are the "correct factors" for the a.e. convergence in the sense that if \(r_n\) are a set of weights with \(\lim_{n \to \infty} r_n / |Q_n| = 0\), then the averages

\[
A_{n,r_n} f(x) = \frac{1}{r_n} \sum_{u \in B_n} T^u f(x)
\]

fail to converge a.e. in \(L^1(X)\).
In this chapter we prove that the notion of correct factor can be extended to actions of amenable groups and discuss some application of the correct factor to obtain a result for sequences on $\mathbb{Z}$.

### 4.1 Amenable Groups.

Let $G$ be an amenable group with left invariant Haar measure $\lambda_G$, acting measurably on a probability space $(X, \beta, m)$ by left translations. For any $\lambda_G$-measurable $F \subset G$, denote its Haar measure by $|F| (= \lambda_G(F))$. A measurable function $f$ on $(X, \beta, m)$ is called $G$-invariant if $f(T_g x) = f(x)$ a.e. for all $g \in G$.

**Definition 4.1.1** Let $\{F_n\}$ be a sequence of subsets of $G$ such that $0 < |F_n| < \infty$ and

$$\lim_{n \to \infty} \frac{|gF_n \Delta F_n|}{|F_n|} = 0.$$ 

Such a sequence is called a (left) $F_0$-Følner sequence.

Throughout this chapter, we will refer to $F_0$-Følner sequences without mentioning each time that they are left $F_0$-Følner sequences.

The existence of $F_0$-Følner sequences is equivalent to the amenability of the group.

**Lemma 4.1.2** Let $\{F_n\}$ be a $F_0$-Følner sequence of subsets of $G$ such that $0 < |F_n| < \infty$. Then, for all $f \in L^p(X)$, $1 \leq p \leq \infty$, there exists a $G$-invariant function $f \in L^p(X)$ such that

$$\lim_{n \to \infty} \| \frac{1}{|F_n|} \int_{F_n} f(T_g x) d\lambda(g) - f^*(x) \|_p = 0.$$

**Proof.** Denote $L_g f(x) = f(gx)$, the left translation of $f$ by $g \in G$. From Lemma 1.1.1, it follows that the subspace

$$\{f \in L^p(X) : f \circ T_g = f \quad \text{a.e. for all} \quad g \in G\} + \text{cl-span}\{f - L_g f : f \in L^\infty(X), g \in G\}$$
is dense in $L^p(X)$. Therefore, it suffices to prove the lemma for functions in the second term. Let $f \in L^\infty(X)$. By the left invariance of $\lambda_G$,

$$\frac{1}{|F_n|} \left| \int_{F_n} [f(T_g x) - L_{g_0} f(T_{g_0} x)] d\lambda(g) \right| = \frac{1}{|F_n|} \left| \int_{F_n} f(T_g x) d\lambda(g) - \int_{F_n} f(T_{g_0} x) d\lambda(g) \right|$$

$$= \frac{1}{|F_n|} \left| \int_{F_n} f(T_g x) d\lambda(g) - \int_{g_0 F_n} f(T_g x) d\lambda(g) \right|$$

$$\leq \frac{1}{|F_n|} \int_{F_n \Delta g_0 F_n} f(T_g x) d\lambda(g)$$

$$\leq \frac{|F_n \Delta g_0 F_n|}{|F_n|} \|f\|_\infty \to \infty$$

as $n \to \infty$.

Therefore, if $G$ is a $\sigma$-compact, locally compact amenable group, there always exists a sequence which is good for mean convergence. However, the question of whether the averages $(1/|F_n|) \int_{F_n} f(T_g x) d\lambda(g)$ converge for almost every $x \in X$ is a more complicated one.

**Definition 4.1.3** Let $G$ be an amenable group with left invariant Haar measure. Let $\{F_n\}$ be a left Følner sequence consisting of compact subsets of $G$ such that $F_n \subset F_{n+1}$ for all $n$ and $\bigcup_n F_n = G$. Then $\{F_n\}$ is called a summing sequence for $G$.

**Theorem 4.1.4** (Emerson [19])

Let $G$ be a $\sigma$-compact, locally compact group. Then $G$ is amenable if and only if there exists a summing sequence for $G$.

**NOTE:** If $G$ is only a locally compact group, then $G$ is amenable if and only if there exists a summing net for $G$. 
The following theorem is due to Tempel'man [39] and Emerson [20].

**Theorem 4.1.5** Let $G$ be a $\sigma$-compact, locally compact amenable group with left invariant Haar measure and $\{F_n\}$, a summing sequence in $G$ such that

$$\sup_{n \in \mathbb{N}} |F_n F_n^{-1}|/|F_n| = C < \infty.$$ 

Then

$$\lim_{n \to \infty} \frac{1}{|F_n|} \int_{F_n} f(T_g x) d\lambda(g)$$

exists a.e. for all $f \in L^1(X)$.

This is the analog of the classical Birkhoff Ergodic Theorem for groups which admit such special summing sequence. See Jenkins [25] for a discussion of examples of such groups.

**Note:** Emerson [20] showed that the nesting condition $(F_n \subset F_{n+1})$ is not necessary. Indeed, he proved that one can slightly perturb a sequence by adding some extra points to each term $F_n$ so that it will loose the nesting property but it will still be good for pointwise averaging. (See Theorem 5.1.3 in the next chapter.)

A more general statement of pointwise convergence holds.

**Theorem 4.1.6 (Tempel'man)**

Let $G$ be a $\sigma$-compact, locally compact amenable group with left invariant Haar measure and $\{F_n\}$ a left Følner sequence. Let $Q_n = \bigcup_{k=1}^n F_k^{-1}F_n$. Then

$$\lim_{n \to \infty} \frac{1}{|Q_n|} \int_{F_n} f(T_g x) d\lambda(g)$$

exists a.e. for all $f \in L^1(X)$. 
To prove Theorem 4.1.6 we need the following Transfer Principle. (See Appendix D for a proof of it).

**Theorem 4.1.7 (Emerson [20])**

Let $G$ be a $\sigma$-compact, locally compact group. Let $\{F_n\}$ be a sequence of compact subsets of $G$ and $\{a_n\}$ a sequence of positive numbers. If

$$M\phi(g) = \sup_{n \in \mathbb{N}} \frac{1}{a_n} \int_{F_n} \phi(hg)dh$$

is a weak $(1,1)$ operator on $L^1(G)$, then

$$Mf(x) = \sup_{n \in \mathbb{N}} \frac{1}{a_n} \int_{F_n} f(T_hx)dh$$

is also a weak $(1,1)$ operator on $L^1(X)$.

Here is a simple proof of Theorem 4.1.6 for $\sigma$-compact groups. The idea of this proof is contained in Rosenblatt and Wierdl [36], where this theorem is proved with $F_n$ being blocks of consecutive numbers in $\mathbb{Z}$ with increasing length. Ornstein-Weiss [30] used a similar technique to prove Theorem 4.1.5 for discrete groups by means of covering lemma (cf. Emerson), but by avoiding the Transfer Principle. We, on the other hand, do use the transfer principle since it translates the problem to the more natural situation where the group $G$ acts on $L^1(G)$ by left translations.

**Proof of Theorem 4.1.6.**

Arguing as in Lemma 4.1.2, one can prove that the averages converge on a dense subspace of $L^1(X)$. Thus, we only need to prove a maximal estimate. Using the Transfer Principle, it suffices to show that the corresponding maximal inequality on
$L^1(G)$ holds:
\[
\left| \{ g \in G : \sup_{n \in \mathbb{N}} \frac{1}{|Q_n|} \int_{F_n} |\phi(hg)|dh > \lambda \} \right| \leq C \frac{\|\phi\|_1}{\lambda}
\]
for all $\phi \in L^1(G)$ with compact support. Let $\phi \in L^1(G)$ positive with compact support and, for fixed $\lambda > 0$, consider the set $A = \{ g \in G : \max_{1 \leq n \leq N} \frac{1}{|Q_n|} \int_{F_n} |\phi(hg)|dh > \lambda \}$. If we prove the above inequality for this set with a constant $C$ independent of $N$, we would prove the theorem.

Note that since support of $\phi$ and all the $F_n$'s are compact subsets of $G$, the mapping $g \rightarrow (1/|Q_n|) \int_{F_n} \phi(hg)dh$ is a continuous function on $G$ whose support, $F_n^{-1}\text{supp} (\phi)$, is also compact. Thus, since we are taking supremum over a finite number of $n$'s, $A$ is an open set in $G$ with compact closure. Also, since $|F_n| > 0$ for all $n$, each of the symmetric sets $F_n^{-1}F_n$ contains a neighborhood of the identity $e$. This is due to the fact that the map $g \rightarrow |Fg \cap F|$ is continuous. By the definition of the sets $Q_n$, we have $Q_n \supset F_n^{-1}F_n$. Therefore, each $Q_n$ also contains a neighborhood of $e$.

Let $n_1$ be the largest integer $\leq N$ such that there exists an element $g_1$ in $A$ with
\[
\frac{1}{|Q_{n_1}|} \int_{F_{n_1}} |\phi(hg_1)|dh > \lambda
\]
and $|A \cap Q_{n_1}g_1| > 0$. The later property is possible because $A$ is open and $Q_{n_1}$ contains a neighborhood of $e$. Then, by properties of the modular function,
\[
|Q_{n_1}g_1| < \frac{1}{\lambda} \int_{F_{n_1}g_1} |\phi(h)|dh.
\]
Let $A_1 = A \setminus Q_{n_1}g_1$. Since $Q_{n_1}g_1$ is compact, $A_1$ is an open set. If $A_1 \neq \emptyset$, let $n_2$ be the largest integer $\leq N$ such that there exists an element $g_2$ in $A_1$ with

$$\frac{1}{|Q_{n_2}|} \int_{F_{n_2}} |\phi(h^g_2)|dh > \lambda$$

such that $|A_1 \cap Q_{n_2}| > 0$. This is possible since $A_1$ is open and $Q_{n_2}$ contains a neighborhood of $e$. Again

$$|Q_{n_2}g_2| < \frac{1}{\lambda} \int_{F_{n_2}g_2} |\phi(h)|dh.$$

Observe that $n_2 \leq n_1$, so by the construction of $Q_1$, $F_{n_2}g_2 \cap F_{n_1}g_1 = \emptyset$ because, otherwise, $g_2 \in F_{n_2}^{-1}F_{n_1}g_1 \subset Q_{n_1}g_1$. Let $A_2 = A_1 \setminus Q_{n_2}g_2$ and continue the process.

By induction, we obtain a non-increasing sequence $\{n_k\}$ of integers and a sequence $\{g_k\}$ of elements of $A$ such that

$$g_{k+1} \in A \setminus (Q_{n_1}g_1 \cup \cdots \cup Q_{n_k}g_k),$$

$$|A \setminus (Q_{n_1}g_1 \cup \cdots \cup Q_{n_k}g_k) \cap Q_{n_{k+1}}g_{k+1}| > 0,$$

$\{F_{n_k}g_k\}$ is a sequence of pairwise disjoint sets and

$$|Q_{n_k}g_k| < \frac{1}{\lambda} \int_{F_{n_k}g_k} |\phi(h)|dh.$$

By the properties of the sets $F_{n}^{-1}F_n$, the sequence $\{g_k\}$ does not have a cluster point.

To see this, let $V$ be a neighborhood of $e$ such that $V \subset F_{n}^{-1}F_n$ for all $n \leq N$. If $U$ is another neighborhood of $e$ which is symmetric and $U^2 \subset V$, then the sets $\{Ug_k\}$ are pairwise disjoint. Otherwise, if $u_1g_k = u_2g_j$ with $u_i \in U$, $i = 1, 2$ and $k < j$; then $g_j = u_2^{-1}u_1g_k \in U^2g_k \in Q_{n_k}g_k$, which is a contradiction. By compactness of $A$, the
sequence \( \{g_k\} \) must be finite. It follows that \(|A \setminus \cup_k Q_{n_k}g_k| = 0\) because otherwise we could continue the construction.

From the above properties we finally obtain,

\[
|A| \leq \sum_k |Q_{n_k}g_k| \leq \frac{1}{\lambda} \sum_k \int_{F_{n_k}g_k} |\phi(h)|dh \leq \frac{1}{\lambda} \|\phi\|_1.
\]

Since the right hand side is independent of \(N\), by taking limit as \(N\) tends to infinity we obtain the desired maximal inequality;

\[
\left| \{ g \in G : \sup_{n \in \mathbb{N}} \frac{1}{|Q_n|} \int_{F_n} |\phi(hg)|dh > \lambda \} \right| \leq \frac{1}{\lambda} \|\phi\|_1.
\]

Theorem 4.1.5 now follows as a corollary of the above. Since \(F_n \subset F_{n+1}\) for all \(n\), then \(Q_n = F_n^{-1} F_n\). And since, by hypothesis, \(|F_n^{-1} F_n| \leq C |F_n|\) for all \(n\), the theorem follows. In case that the action of \(G\) on \(X\) is ergodic (if \(B \in \beta\) and \(m(B \Delta T_g B) = 0\) for all \(g \in G\) then \(B = X, \emptyset \mod 0\)), we can identify the limit. As in the classical Ergodic Theorem,

\[
\lim_{n \to \infty} \frac{1}{|F_n|} \int_{F_n} f(T_gx)dg = \int_X f dm
\]
a.e. for all \(f \in L^1(X)\).

**Corollary 4.1.8** Let \(G\), \(\{F_n\}\) and \(\{Q_n\}\) be as in the above theorem. If \(\lim_{n \to \infty} |F_n|/|Q_n| = 0\), then, for all \(f \in L^1(X)\), \(\lim_{n \to \infty} \frac{1}{|Q_n|} \int_{F_n} f(T_gx)d\lambda(g) = 0\)

**Definition 4.1.9** Given a sequence \(\{F_n\}\) of subsets of \(G\), the positive numbers \(\{q_n\}\) are called the correct factors (associated to the \(F_n's\)) if

\[
\lim_{n \to \infty} \frac{1}{q_n} \int_{F_n} f(T_gx)d\lambda(g)
\]
exists a.e. for all $f \in L^1(X)$ and if $\{c_n\}$ is any sequence such that $\lim_{n \to \infty} c_n/q_n = 0$ then $\lim_{n \to \infty} (c_n/q_n) \int_{F_n} f(T_g x) d\lambda(g)$ fails to exist for some function $f \in L^1(X)$.

Our aim is to prove that if $\{F_n\}$ satisfies the hypothesis of the Theorem 4.1.6, and $Q_n = \bigcup_{k=1}^n F_n^{-1} F_n$, then $|Q_n|$ are the correct factors associated to the sequence $\{F_n\}$.

REMARK. The sequence $\{|Q_n|\}$ is not always the correct factor.

Example 1:
Here is an example of a sequence that violates the regularity condition $|F^{-1}_n F_n| \leq C|F_N|$ for which the $|Q_n|$'s are not the correct factor.

Consider $G = \mathbb{Z}$ and $F_n = \{1, 4, 9, \ldots, n^2\}$. Then $|Q_n|$ is $|F^{-1}_N F_N| \sim n^2$, but

$$\lim_{N \to \infty} \frac{1}{N^{1+\epsilon}} \sum_{n=1}^N T^{n^2} f(x) = 0$$

because this limit exists and equals 0 for all bounded functions and

$$\sup_{n \in \mathbb{N}} \frac{1}{N^{1+\epsilon}} \sum_{n=1}^N T^{n^2} f$$

is integrable. Indeed,

$$\| \sup_{n \in \mathbb{N}} \frac{1}{N^{1+\epsilon}} \sum_{n=1}^N T^{n^2} f \|_1 \leq \sum_{k=0}^\infty \| \sup_{2^k \leq N < 2^{k+1}} \frac{1}{N^{1+\epsilon}} \sum_{n=1}^N T^{n^2} f \|_1$$

$$\leq \sum_{k=0}^\infty \frac{1}{2^{k+\epsilon}} \| \sum_{n=1}^{2^{k+1}} |T^{n^2} f| \|_1$$

$$\leq 2 \sum_{k=0}^\infty \frac{1}{(2^k)^\epsilon} \|f\|_1 = c\|f\|_1.$$

Example 2:
Here is an example in which there is no sequence $\{x_n\}$ such that $F_n x_n \subset F_{n+1} x_{n+1}$
and such that the $|Q_n|$'s are not the correct factors.

Let $B_N = \{(k,j) \in \mathbb{Z}^2 : |k| \leq N, |j| \leq \sqrt{N}\}$ and $T$ and $S$ two commuting measure-preserving invertible transformations on a probability space $(X, \beta, m)$. Then, by Theorem 4.1.5,

$$
\lim_{N \to \infty} M_N f(x) = \frac{1}{|B_N|} \sum_{(k,j) \in B_N} T^k S^j f(x)
$$

exists a.e. for all $f \in L^1(X)$. Now define $F_{2N} = B_N$ and $F_{2N+1} = \{(k,j) \in \mathbb{Z}^2 : |k| \leq \sqrt{N}/2, |j| \leq 2N\}$. Then $\{F_j\}$ is a $F_\sigma$-linear sequence that satisfies $|F_j| \not\to \infty$ and $|F_j - F_{j+1}| \leq C|F_j|$ for all $j$. But there is no sequence $\{x_j\}$ of points in $\mathbb{Z}^2$ such that $F_j + x_j \subset F_{j+1} + x_{j+1}$. For this sequence,

$$
Q_{2N} = (F_{2N} - F_{2N}) \cup (F_{2N} - F_{2N+1})
$$

$$
= \{(k,j) : |k| \leq 2N, |j| \leq 2\sqrt{N}\} \cup \{(k,j) : |k| \leq 2N + \frac{\sqrt{N}}{2}, |j| \leq 2N + \sqrt{N}\}.
$$

So, $|Q_{2N}| \geq (2N + \sqrt{N}/2)(2N + \sqrt{N}) \geq N^2$. Similarly, $|Q_{2N+1}| \geq N^2$. But then

$$
\lim_{N \to \infty} |Q_j|/|F_j| = \infty.
$$

However, the $|Q_j|$'s are not the correct factor because

$$
\lim_{j \to \infty} (1/|B_j|) \sum_{(k_1,k_2) \in B_j} T^{k_1} S^{k_2} f
$$

exists a.e. and equals $P_T P_S f$, the projections onto the subspace of invariant functions under $T$ and $S$ respectively.

\[\blacksquare\]

**Definition 4.1.10** Let $\{F_n\}$ be a $F_\sigma$-linear sequence of compact subsets of a group $G$ such that there exists $\{x_n\}$ with $F_n x_n \subset F_{n+1} x_{n+1}$ for all $n$ (nesting condition),

$$
\bigcup_{n+1}^\infty F_n x_n = G \text{ and } \sup_{n \in \mathbb{N}} |F_{n-1} F_n|/|F_n| = C < \infty \text{ (left regularity)}. Then $\{F_n\}$ is called a regular summing sequence.
A regular summing sequence is a generalization of a summing sequence since \( \{F_n x_n\} \) is a summing sequence for which Theorem 4.1.5 holds.

**Proposition 4.1.11** Let \( G \) be a \( \sigma \)-compact, locally compact amenable group, \( \{F_n\} \) a regular summing sequence and \( Q_n = \bigcup_{k=1}^n F_n^{-1} F_n \). Assume also that

(A) \( \sup_{n \in \mathbb{N}} |F_n F_n^{-1} F_n^*| = C < \infty \) (right regularity) and

(B) \( \sup_{n \in \mathbb{N}} \left| \bigcup_{i=k}^n x_i x_n^{-1} F_n^{-1} F_n \right| \left| \bigcup_{i=k}^n x_i x_n^{-1} F_n^{-1} \right| = C < \infty \),

where \( \{x_i\} \) is a sequence with respect to which \( \{F_n\} \) is a regular summing sequence.

If \( \{c_n\} \subset \mathbb{N} \) is a sequence with \( \lim_{n \to \infty} c_n = \infty \) then

\[
\sup_{n \in \mathbb{N}} \frac{c_n}{|Q_n|} \int_{F_n} f(gh) d\lambda_G(g)
\]

is not a weak \((1,1)\) operator on \( L^1(G) \).

The conditions imposed on the Følner sequence \( \{F_n\} \) force it to behave in a fashion similar to sequences of blocks of consecutive numbers in \( \mathbb{Z} \). Indeed, a sequence \( \{F_n\} \) satisfying all the hypothesis of the theorem is formed by shifting a summing sequence for which Theorem 4.1.5 holds, and controlling the shifts so that we still obtain convergence a.e. with appropriate weights.

The nesting condition imposed on the sequence \( \{F_n\} \) plays the role of the condition Rosenblatt and Wierdl had for block sequences on \( \mathbb{Z} \) in in [36]; that is, that the length of the blocks must be increasing.

**Proof.** If a group has a summing sequence with right regularity, then it is a unimodular group, Chatard [14]. In our case, \( \{F_n x_n\} \) is such a sequence. Therefore, \( G \) is unimodular.
Case sup
N∈N |Qn|/|Fn| = ∞.

Take N big enough such that

\[ |Q_n|/|F_n| \geq |Q_n|/|F_n| \text{ for all } n \leq N \text{ and } |Q_n|/|F_n| > \alpha M, \quad (i) \]

where \( \alpha \) is to be determined and M is such that \( c_n > c_j \) for all \( n > M \) and \( j \) fixed.

Let

\[ \phi(g) = \begin{cases} \frac{|Q_n|}{c_j |F_n|} & \text{if } g \in F_N F_N^{-1} \\ 0 & \text{otherwise.} \end{cases} \]

Then \( \|\phi\|_1 = |Q_n|/|F_N F_N^{-1}|/c_j |F_N| \). Since \( \{x_n\} \) is a sequence with a nesting property, i.e. \( F_n x_n \subset F_{n+1} x_{n+1} \) for all \( n \), then if \( g \in x_i x_N^{-1} F_N^{-1} \), \( g = x_i x_N^{-1} g' \) for some \( g' \in F_N^{-1} \) and then

\[ F_i g = F_i x_i x_N^{-1} g' \subset F_N x_N x_N^{-1} g' \subset F_N F_N^{-1} \]

if \( i \leq N \). Thus,

\[
M \phi(g) = \sup_{n \in \mathbb{N}} c_n \int_{F_n} \phi(hg)dh \geq \frac{c_i}{|Q_i|} \int_{F_i} \phi(hg)dh \\
= \frac{c_i}{|Q_i g|} \int_{F_i} \phi(h)dh = \frac{c_i}{|Q_i g|} \frac{|Q_N|}{c_j |F_N| |F_i|} |F_i g| \\
= \frac{c_i}{c_j |F_N| |Q_i|} \geq \frac{c_i}{c_j} > 1
\]

if \( i \geq M \) by (i). Therefore

\[ M \phi > 1 \text{ on } \bigcup_{i=M}^{N} x_i x_N^{-1} F_N^{-1}. \]

Let us estimate the size of this set. Let

\[
K_1 = \sup_{N \in \mathbb{N}} \frac{|F_N^{-1} F_N|}{|F_N|}, \quad K_2 = \sup_{N \in \mathbb{N}} \frac{|F_N F_N^{-1}|}{|F_N|} \quad \text{and} \quad K_3 = \sup_{N \in \mathbb{N}} \frac{\bigcup_{i=k}^{N} x_i x_N^{-1} F_N^{-1} F_N}{|\bigcup_{i=k}^{N} x_i x_N^{-1} F_N^{-1}|}.\]
Take $K = \max(K_1, K_2, K_3)$. Let $R = \bigcup_{i=1}^{N} x_i x_{-1}^{i} F_{N}^{-1}$. We want to see that the Haar measure of $R$ is a fraction of the Haar measure of $Q_{N}$.

$$Q_{N} = \bigcup_{i=1}^{N} F_{i}^{-1} F_{N} \subset \bigcup_{i=1}^{M-1} F_{i}^{-1} F_{N} \cup RF_{N}.$$  

Then, by the left invariance of the Haar measure,

$$| \bigcup_{i=1}^{M-1} F_{i}^{-1} F_{N} | \leq \sum_{i=1}^{M-1} | F_{i}^{-1} F_{N} | \leq \sum_{i=1}^{M-1} | x_{i} x_{-1}^{i} F_{N}^{-1} F_{N} | < M | F_{N}^{-1} F_{N} | .$$

By (i) we have

$$| \bigcup_{i=1}^{M-1} F_{i}^{-1} F_{N} | \leq \frac{|Q_{N}|}{\alpha |F_{N}|} | F_{N}^{-1} F_{N} | \leq \frac{K}{\alpha} |Q_{N}| .$$

(Here, the regularity condition was employed). Choosing $\alpha = 4K$,

$$|RF_{N}| \geq |Q_{N}| - \bigcup_{i=1}^{M-1} F_{i}^{-1} F_{N} \geq \frac{3}{4} |Q_{N}| .$$

But the hypothesis (B) means that $|RF_{N}| \leq K |R|$. Therefore, $|R| > 3|Q_{N}|/4K$.

Finally,

$$|\{g \in G : M \phi(g) > 1 \}| \geq |R| > \frac{3}{4K} |Q_{N}|$$

$$= \frac{3}{4K} \frac{c_{j} |F_{N}|}{|F_{N} F_{N}^{-1}|} \|\phi\|_{1} > \frac{3c_{j}}{4K^2} \|\phi\|_{1}$$

(here we use hypothesis (A)). Since $j$ is arbitrary, the right hand side tends to infinity as $j \to \infty$. Therefore, for any constant $C > 0$, there exists $\phi \in L^{1}(G)$ such that

$$|\{g \in G : M \phi(g) > 1 \}| > C \|\phi\|_{1} .$$

Case $\sup_{n \in \mathbb{N}} |Q_{n}| / |F_{n}| = C < \infty$.

The argument is similar to the previous case. Let $c_{j}$, $M$ and $K$ be chosen as above.
and let
\[ \phi(g) = \begin{cases} 
(1/c_j) & \text{if } g \in F_NF_N^{-1} \\
0 & \text{otherwise.} 
\end{cases} \]

Then \( \|\phi\|_1 = |F_NF_N^{-1}|/c_j \leq |Q_N|/c_j \). If \( g \in x_i x_N^{-1} F_N^{-1}, i \leq N \), then
\[ M\phi(g) = \sup_{n \in \mathbb{N}} \frac{c_n}{|Q_n|} \int_{F_n} \phi(hg)dh \geq \frac{c_i}{C|F_i|} \int_{F_i} \phi(hg)dh \]
\[ = \frac{c_i}{Cc_j} \geq \frac{1}{C}. \]

Therefore
\[ M\phi > 1/C \quad \text{on } \bigcup_{i=M}^{N} x_i x_N^{-1} F_N^{-1}. \]

But the same argument gives,
\[ |\bigcup_{i=M}^{N} x_i x_N^{-1} F_N^{-1}| \geq \frac{3}{4K} |Q_N| \geq \frac{3}{4K} c_j \|\phi\|_1. \]

Since \( j \) is arbitrary, this finishes the proof.

REMARKS:

(i) If \( G = \mathbb{R}^d \) or \( \mathbb{Z}^d \), there are sequences satisfying the hypothesis of the theorem. Indeed, \( F_n = [0, n]^d \) will do.

(ii) If \( G \) is the Heisenberg group, there are also sequences satisfying the hypothesis of the theorem. Identify \( G \) with \( \mathbb{R}^3 \) together with the operation \((x, y, z)(x', y', z') = (x + x', y + y', z + z' + xy')\). Take \( F_n = \{(x, y, z) : x \in [0, n], y \in [0, n], z \in [0, n^2]\} \). This sequence is a right invariant Følner sequence, \( |F_n^{-1}F_n| \leq 16|F_n| \) and \( |F_nF_n^{-1}| \leq 16|F_n| \).

To prove condition (B) requires more work. But the idea is essentially the same as in the previous case.
Claim: $F_n^{-1}F_n \subset \cup_{j=1}^{r} A_j$ where $r$ is independent of $n$, and $A_j \subset (a, b, c)F_n$ for appropriate $(a, b, c)$.

If this were true, then for any sequence of points $\{\omega_i\}$ in $G$, $|\cup_{i=1}^{n} \omega_i F_n^{-1}F_n| \leq C |\cup_{i=1}^{n} \omega_i F_n^{-1}|$. Indeed,

$$|\cup_{i=1}^{n} \omega_i F_n^{-1}F_n| = |\cup_{i=1}^{n} F_n^{-1}F_n \omega_i^{-1}| \leq |\cup_{j=1}^{r} \cup_{i=1}^{n} A_j \omega_i^{-1}| \leq |\cup_{j=1}^{r} \cup_{i=1}^{n} (a_j, b_j, c_j) F_n \omega_i^{-1}| \leq \sum_{j=1}^{r} |(a_j, b_j, c_j) \cup_{i=1}^{n} F_n \omega_i^{-1}| \leq r |\cup_{i=1}^{n} F_n \omega_i^{-1}| \leq r |\cup_{i=1}^{n} \omega_i F_n^{-1}|.$$

It is not difficult to see that

$$F_n^{-1}F_n \subset E_n = \{(x, y, z): |x| \leq n, |y| \leq n, |z| \leq 2n^2\}$$

and $E_n = \cup_{j=1}^{16} B_j$ where $B_j = \{(x + a, y + b, z + c): (x, y, z) \in F_n\}$ and $(a, b, c)$ is an appropriate element of the group. In fact $a = n$ or $0$, $b = n$ or $0$ and $c = n, 2n, -n$ or $0$. In case $a = 0$, let $A_j = B_j = (0, b, c)F_n$. If $a \neq 0$, then

$$(x + a, y + b, z + c) = (x + a, y + b, z' + c + ax) = (a, b, c)(x, y, z')$$

where $z' = z - ay$. If $(x, y, z) \in F_n$, then $-an \leq z' \leq n^2$. But $a = n$. Thus $|z'| \leq n^2$.

Then either $z' \in F_n$ or $z' = -n^2 z''$ for some $z'' \in F_n$. Then

$$(a + x, b + y, c + z) = \begin{cases} (a, b, c)(x, y, z') & \text{for some } (x, y, z') \in F_n \\ (a, b, c - n)(x, y, z'') & \text{for some } (x, y, z'') \in F_n. \end{cases}$$

Therefore, $B_j \subset A_j \cup A'_j$ where $A_j = (a, b, c)F_n$ and $A'_j = (a, b, c - n)F_n$. 

\[\blacksquare\]
The question of which locally compact groups admit Fσ lner sequences satisfying the entire hypothesis of Proposition 4.1.11 remains open. It is clear, from the above example, that if a group \( G \) admits a nested left Fσ lner sequence \( \{ F_n \} \) (\( F_n \subset F_{n+1} \)) with right regularity condition \( \sup_{n \in \mathbb{N}} |F_n F_n^{-1}|/|F_n| < \infty \) such that \( F_n^{-1} F_n \subset \bigcup_{i=1}^r F_n y_i \) for any \( y_i \)'s \( \in G \) and where \( r \) is independent of \( n \), this sequence would satisfy the hypothesis of Proposition 4.1.11. The above example suggests that this should be possible for nilpotent groups. We conjecture that it should be true for all groups of polynomial growth because all these groups have Fσ lner sequences which tile the group.

The construction in Proposition 4.1.11 can be carried out on the space \( (X, \beta, m) \) if the action is ergodic. To prove such a statement we need to employ a Rohlin–Kakutani tower construction and to reproduce the argument using the levels of the tower. Although for general groups the Rohlin–Kakutani Lemma [34] may not hold, the whole strength of that lemma is not needed. Instead we have a less ambitious lemma. (See Ornstein–Weiss [31]). For this lemma we need the space \( (X, \beta, m) \) to be a Lebesgue space. Such a space can be thought of as being the unit interval with the Lebesgue measure.

**Lemma 4.1.12** Let \( G \) be a locally compact group, \( (X, \beta, m) \), a non-atomic Lebesgue space and \( T : G \times X \to X \) an ergodic measure–preserving action of \( G \) on \( X \). For any compact neighborhood \( K \) of \( e \) in \( G \), there is a set \( V \in \beta \) such that \( T_K V = \bigcup_{g \in K} T_g V \) is also measurable, both are of positive measure and the sets \( \{ T_g V \}_{g \in K} \) are pairwise
disjoint. Moreover, the σ-algebra β restricted to $T_K V$ is the product σ-algebra of the completed Borel sets on $K$ and the σ-algebra of subsets $V' \subset V$ with the property that $T_K V' \in \beta$. And, on $T_K V$, the restriction of the measure $m$ coincides with the product measure of the Haar measure on $K$ and some measure on $V$.

A consequence of Lemma 4.1.12 is that on the space $(X, \beta, m)$, one can find arbitrarily small almost invariant sets. (See Rosenblatt-del Junco [18]).

**Lemma 4.1.13** Under the hypothesis of the above lemma, given $\epsilon > 0$ and a compact set $K$ in $G$, there exists $A \in \beta$ such that $m(A) < \epsilon$ and $m(T_g A \Delta A)/m(A) < \epsilon$ for all $g \in K$.

**PROOF.** Let $F$ be a compact neighborhood of $e$ almost $K$ invariant, i.e. $|gF \Delta F|/|F| < \epsilon$ for all $g \in K$. Consider the compact set $KF$ and perform Lemma 4.1.12 with it. There exists a measurable set $V \in X$ with positive measure such that $\{T_h V\}_{h \in KF}$ are disjoint and the restriction of $\beta$ to $K_{KF} V$ coincides with the product σ-algebra of the Borel subsets of $KF$ times the σ-algebra of subsets $V'$ of $V$ such that $T_{KF} V' \in \beta$. Since $X$ has no atoms, it is possible to find $V_0 \subset V$ such that $T_h V_0 \in \beta$ for all Borel subsets $H \in KF$, such that $m(T_F V_0) < \epsilon$. If we let $A = T_F V_0$, then it satisfies all the desired properties because of the choice of $V_0$ and the invariance of $F$.

Lemma 4.1.12 allows us to transfer the behavior of the averages for the regular action of the group on $L^1(G)$ back to the averages of an ergodic action of $G$ on a probability space $(X, \beta, m)$. 
Corollary 4.1.14 Let $G$ be a locally compact, $\sigma$–compact amenable group and $\{F_n\}$ any sequence satisfying the hypothesis of Proposition 4.1.11. Then, for any space $(X, \beta, m)$ on which $G$ acts ergodically, the $|Q_n| = \bigcup_{k=1}^{n} F_k^{-1} F_n$ are the correct factors associated to the $F_n$’s.

PROOF. We sketch the proof for the case $\sup_{n\in\mathbb{N}} |Q_n|/|F_n| = \infty$. The other case follows in a similar way. Following the notation of Proposition 4.1.11, let $\{c_n\}$ be a sequence of positive number so that $\lim_{n\to\infty} c_n = \infty$ and $\lim_{n\to\infty} c_n/|Q_n| = 0$. Also, let $N$, $j$ and $M$ be as in the proof of the proposition and take the compact set $K$ large enough so that it contains

$$F_N F_N^{-1} \cup (\bigcup_{i=1}^{N} x_i x_N^{-1} F_N^{-1}).$$

Let $V$ be such that $T_K V$ has the properties of Lemma 4.1.12. Define

$$f(x) = \begin{cases} \frac{|Q_N|}{c_j |F_N|} & \text{if } x \in T(F_N F_N^{-1})V \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$M f(x) = \sup_{n\in\mathbb{N}} \frac{c_n}{|Q_n|} \int_{F_n} f(T_g x) dg > 1$$

on $R = T(\bigcup_{i=M}^{N} x_i x_N^{-1} F_N^{-1})V$. By the above, there exists a measure $\mu$ on $V$ such that, for any measurable subset $K' \subset K$, $m(T_{K'} V) = |K'| \mu(V)$. Following the rest of the argument in Proposition 4.1.11, we obtain

$$m(\{x \in X : M f(x) > 1\}) \geq m(R) = |\bigcup_{i=M}^{N} x_i x_N^{-1} F_N^{-1}| \mu(V) \geq C |Q_n| \mu(V) = C \frac{c_j |F_N| \mu(V)}{m(T_{FNF_N^{-1}} V) \|f\|_1}$$
Since $j$ is arbitrary and $\lim_{j \to \infty} c_j = \infty$, we conclude that for any constant $C > 0$ we can construct a function such that

$$m\{x : \sup_{n \in \mathbb{N}} \frac{c_n}{|Q_n|} \int F_n f(T_n x) d\lambda(g) > 1\} > C\|f\|_1.$$ 

Therefore, $M f(x) = \sup_{n \in \mathbb{N}} [c_n/|Q_n|] \int F_n f(T_n x) d\lambda(g)$ is not a weak $(1,1)$ operator. From the Banach Principle, there exists $f \in L^1(X)$ such that $M f(x) = \infty$ on a set of positive measure. Since this holds for any sequence $\{c_n\}$ with the above properties, it follows, by the definition of correct factor, that the sequence $\{|Q_n|\}$ is the correct factor corresponding to the $F_n$'s.

The following proposition is a relative of Theorem 4.1.6 concerning tail distributions. For a complete exposition on large deviation problems for sequences on $\mathbb{Z}$, we refer the reader to Rosenblatt–Wierdl [36].

**Proposition 4.1.15** Let $G$ be a $\sigma$-compact, locally compact unimodular group and $\{F_n\}$ be a regular summing sequence. Then

$$\sum_{n=1}^{\infty} |\{h : \frac{1}{|F_n|} \int_{F_n} \phi(gh)dg > n\}| < C\|\phi\|_1$$

for all $\phi \in L^1(G)$ with compact support.
PROOF. Let \( \Omega_n = \{ h : \frac{1}{|F_n|} \int_{F_n} \phi(g h)dg > n \} \). One can find inductively a sequence \( \{y_n\} \) of elements of \( G \) so that \( \{y_n \Omega_n\} \) are pairwise disjoint. By the left invariance of the Haar measure, \( |y_n \Omega_n| = |\Omega_n| \). Thus,

\[
\sum_{n=1}^{\infty} |\Omega_n| = \sum_{n=1}^{\infty} |y_n \Omega_n| = |\{\sup_{n \in \mathbb{N}} \frac{1}{|F_n|} \int_{F_n} \phi(gy_n^{-1}h)dg > 1\}| = |\{\sup_{n \in \mathbb{N}} \frac{1}{|F_n y_n^{-1}|} |F_n y_n^{-1}| \int_{F_n y_n^{-1}} \phi(g h)dg > 1\}|.
\]

Let \( E_n = F_n y_n^{-1} \). Then, by unimodality of the Haar measure, \( \{E_n\} \) is also a regular summing sequence. Let \( Q_n = \bigcup_{j=1}^{n} E_j^{-1} E_n \). Then, by right and left invariance,

\[
|Q_n| \leq \sum_{j=1}^{n} |E_j^{-1} E_n| \leq n |F_n^{-1} F_n| \leq C n |F_n|.
\]

From the proof of Theorem 4.1.6, it follows that

\[
\sup_{n \in \mathbb{N}} \frac{1}{|F_n y_n^{-1}|} |F_n y_n^{-1}| \int_{F_n y_n^{-1}} \phi(g h)dg|
\]

is a weak \((1,1)\) operator. Therefore,

\[
\sum_{n=1}^{\infty} |\Omega_n| \leq C \|\phi\|_1
\]

for some constant \( C \) and all \( \phi \in L^1(G) \) with compact support.

\[\blacksquare\]

This result can be transferred to \( L^1(X) \), although we need a different version of the Transfer Principle.
Let $G$ be an amenable group. Let $\{F_n\}$ be a sequence of compact subsets of $G$ and \{a_n\} a sequence of positive numbers. If there exists a constant $C$ such that
\[
\sum_{n=1}^{\infty} \left| \left\{ g : \frac{1}{|F_n|} \int_{F_n} \phi(hg) dh > a_n \right\} \right| \leq C \|\phi\|_1
\]
for all $\phi \in L^1(G)$ with compact support, then
\[
\sum_{n=1}^{\infty} m(\left\{ x : \frac{1}{|F_n|} \int_{F_n} f(T_h x) dh > a_n \right\}) \leq C \|f\|_1
\]
for all $f \in L^1(X)$ with the same constant.

This transference principle follows by using the same method of proof as in Theorem 4.1.7. (See Appendix D)

**Corollary 4.1.16** Let $G$ be a $\sigma$-compact, locally compact unimodular amenable group and $\{F_n\}$ be a regular summing sequence. Then
\[
\sum_{n=1}^{\infty} m(\left\{ x : \frac{1}{|F_n|} \int_{F_n} f(T_h x) dg > n \right\}) < C \|f\|_1
\]
for all $f \in L^1(X)$.

**Note:** J. Rosenblatt pointed out (private communication) that, for block sequences on $Z$, if $\{c_n\}$ is any sequence of positive numbers with $\lim_{n \to \infty} c_n = 0$, there exists a function $f \in L^1(X)$ for which
\[
\sum_{n=1}^{\infty} m(\left\{ x : \frac{1}{|F_n|} \int_{F_n} f(T_h x) dg > nc_n \right\}) = \infty.
\]

**Proposition 4.1.17** Let $(X, \beta, m)$ and $G$ satisfy the hypothesis of Lemma 4.1.12. If $\{F_n\}$ is a regular summing sequence and $\{c_n\}$ is a sequence of positive numbers with $\lim_{n \to \infty} c_n = 0$, then there exists $f \in L^1(X)$ for which
\[
\sum_{n=1}^{\infty} m(\left\{ x : \frac{1}{|F_n|} \int_{F_n} f(T_h x) dg > nc_n \right\}) = \infty.
\]
PROOF. First we prove that, given $K$ large and $\delta > 0$ small, there exists $f \in L^1(X)$ such that
\[
\sum_{n=1}^{\infty} m(\{x : \frac{1}{|F_n|} \int_{F_n} f(T_g x) dg > nc_n\}) > K \|f\|_1
\]
and $\|f\|_1 < \delta$.

Let $M = 2(K + 1)$, and let $N_0$ be such that for all $n \geq N_0$, $c_n \leq 1/M$, and let $N = 2N_0$. Choose $\epsilon$ small enough so that $\epsilon < \delta M/N$ $2\epsilon \sum_{n=N_0}^{N} 1/(c_nn) \leq 1$. By Lemma 4.1.13, there exists a measurable set $A$ such that $m(A) < \epsilon$ and
\[
\frac{m(T_g^{-1}A \Delta A)}{m(A)} < \epsilon \text{ for all } g \in F_n, n \leq N.
\]

Let $f = (N/M)1_A$. Then, $\|f\|_1 = Nm(A)/M < \epsilon N/M < \delta$. The idea here is that since the set $A$ is almost invariant, then $A_nf(x) = (1/|F_n|) \int_{F_n} f(T_g x) dg$ is very close to $f$. Indeed,
\[
m(\{x : f(x) > c_nn\}) \leq m(\{x : A_nf(x) > \frac{c_nn}{2}\}) + m(\{x : |A_nf(x) - f(x)| > \frac{c_nn}{2}\})
\]
\[
\leq m(\{x : A_nf(x) > \frac{c_nn}{2}\}) + \frac{2}{c_nn} \|A_nf - f\|_1.
\]

But, by Fubini's theorem,
\[
\|A_nf - f\|_1 \leq \int_X \frac{N}{M|F_n|} \int_{F_n} |1_A(T_g x) - 1_A(x)||dgdx
\]
\[
\leq \frac{N}{M|F_n|} \int_{F_n} \int_X |1_A(T_g x) - 1_A(x)||dxdg
\]
\[
= \frac{N}{M|F_n|} \int_{F_n} m(T_g^{-1}A \Delta A)dg
\]
\[
\leq \frac{\epsilon N}{M} m(A) = \epsilon \|f\|_1.
\]

Therefore, from (ii) and (iii),
\[
\sum_{n=N_0}^{N} m(\{x : f(x) > c_nn\}) \leq \sum_{n=N_0}^{N} m(\{x : A_nf(x) > \frac{c_nn}{2}\})
\]
By the choice of $\epsilon$, the last term is small.

$$\sum_{n=N_0}^{N} m(\{x : f(x) > c_n n\}) \leq \sum_{n=1}^{N} m(\{x : A_n f(x) > c_n n/2\}) + \|f\|_1.$$  

To estimate the left hand side, note that by the choice of $N_0$, if $x \in A$, $f(x) = N/M \geq nc_n$ for all $N_0 \leq n \leq N$. Hence, since $N = 2N_0$,

$$\sum_{n=N_0}^{N} m(\{x : f(x) > c_n n\}) \geq \sum_{n=1}^{N} m(A) \geq \frac{N}{2} m(A) = \frac{M}{2} \|f\|_1.$$  

Finally, from (iv),

$$\sum_{n=1}^{N} m(\{x : A_n f(x) > c_n n/2\}) > \left(\frac{M}{2} - 1\right) \|f\|_1 = K \|f\|_1.$$  

From this fact, one can construct via an inductive procedure an increasing sequence $n_k < n_{k+1}$ and step functions $f_k$ such that

1. $\|f\|_1 \leq 2^{-k}$
2. 

$$\sum_{n=n_k+1}^{n_k} m(\{x : A_n f_k(x) > c_n n\}) > k2^k \|f_k\|_1$$

Let $f = \sum_{k=1}^{\infty} f_k$. Then, $\|f\|_1 \leq 1$ and

$$\sum_{n=1}^{\infty} m(\{x : A_n f(x) > c_n\}) \geq \sum_{k=1}^{\infty} \sum_{n=n_k+1}^{n_k} m(\{x : A_n f_k(x) > c_n n\}) > \sum_{k=1}^{\infty} k2^k \|f_k\|_1 = \infty$$
because, by Abel's summation method, if \( S_k = \sum_{j=1}^{k} 2^j \| f_j \|_1 \), then

\[
\sum_{k=1}^{\infty} k 2^k \| f_k \|_1 = \sum_{k=1}^{\infty} (S_k - S_{k-1})k = \sum_{k=1}^{\infty} S_k.
\]

Since \( S_k \) is increasing, the sum must be infinite.

Therefore, \( n \) is the right factor in the above corollary.

4.2 An application of the correct factor

The theorem proved in this section is somewhat disconnected with the topics treated in this chapter. However, it was included here because the proof relies on the correct factor and because it is connected with the topics discussed in the next chapter.

Let \( \{v_n\} \) and \( \{l_n\} \) be sequences in \( \mathbb{N} \) with \( v_n + l_n < v_{n+1} \), and let \( B_n \) be the block

\[
B_n = \{v_n, v_n + 1, \ldots, v_n + l_n - 1\}.
\]

Define

\[
B_n f(x) = \frac{1}{l_1 + \cdots + l_n} \sum_{j \in \cup_{i=1}^{l_n} B_i} T^j f(x) \quad \text{(vi)}
\]

\[
M_m f(x) = \frac{1}{|[1, m] \cap \cup_{i=1}^{\infty} B_i|} \sum_{j \in ([1, m] \cap \cup_{i=1}^{\infty} B_i)} T^j f(x). \quad \text{(vii)}
\]

The first type of average is an average on the whole blocks \( B_k \), whereas the second average stops inside the blocks. The literature contains work primarily studying the first type of averages. Our question is whether the two types of averages are equivalent or not.
Theorem 4.2.1 Let the blocks $B_n$ be defined as above. If either $l_1 + \cdots + l_{n-1} \leq Kl_n$ for all $n$ or $l_n \leq K(l_1 + \cdots + l_{n-1})$ for all $n$, then $\sup_{n \in \mathbb{N}} |B_n f|$ is a weak (1,1) operator if and only if $\sup_{n \in \mathbb{N}} |M_n f|$ is also.

PROOF. It is enough to prove that if $\sup_{n \in \mathbb{N}} |B_n f|$ is a weak (1,1) operator, $\sup_{n \in \mathbb{N}} |M_n f|$ is also weak (1,1) because if $m$ is the endpoint of $B_n$ then $M_m f(x) = B_n f(x)$.

Without loss of generality, assume $f \geq 0$.

Case 1: $\sup_{n \in \mathbb{N}} l_n/(l_1 + \cdots + l_{n-1}) = K < \infty$.

In this case, given $m$, choose $n$ so that $v_n < m < v_{n+1}$. Note that if $v_n + l_n < m < v_{n+1}$ then $B_n f(x) = M_m f(x)$. Then

$$M_m f(x) \leq \frac{1}{l_1 + \cdots + l_{n-1}} \sum_{j \in \mathbb{N}} T^j f(x) \leq \frac{l_1 + \cdots + l_n}{l_1 + \cdots + l_{n-1}} B_n f(x).$$

By hypothesis,

$$\frac{l_1 + \cdots + l_n}{l_1 + \cdots + l_{n-1}} = 1 + K < 1 + K.$$

Case 2: $\sup_{n \in \mathbb{N}} l_1 + \cdots + l_{n-1}/l_n = K < \infty$.

Given $m$, let $n$ be as above. Thus,

$$M_m f(x) = \frac{1}{l_1 + \cdots + l_{n-1} + (m - v_n)} \sum_{j \in \mathbb{N}} T^j f(x)$$

$$+ \frac{1}{l_1 + \cdots + l_{n-1} + (m - v_n)} \sum_{j = v_n}^m T^j f(x)$$

$$\leq B_{n-1} f(x) + \frac{1}{l_1 + \cdots + l_{n-1} + (m - v_n)} \sum_{j = v_n}^m T^j f(x).$$

Therefore, it suffices to prove that

$$\sup_{\{(n,s) : 0 < s \leq l_n\}} \frac{1}{l_1 + \cdots + l_{n-1} + s} \sum_{j = v_n}^{v_n + s - 1} T^j f(x)$$
is a weak (1,1) operator.

Notice that

\[ B_n f(x) = \frac{l_1 + \cdots + l_{n-1}}{l_1 + \cdots + l_n} B_{n-1} f(x) + \frac{1}{l_1 + \cdots + l_n} \sum_{j \in B_n} T^j f(x). \]

Hence the fact that

\[ \sup_{n \in \mathbb{N}} B_n f(x) \]

is a weak (1,1) operator implies that

\[ \sup_{n \in \mathbb{N}} \frac{1}{l_1 + \cdots + l_n} \sum_{j \in B_n} T^j f(x) \]

is one also. In particular, the correct factor corresponding to the \( B_n \)'s has to satisfy

\[ |Q_n| \leq C(l_1 + \cdots + l_n) \quad (viii) \]

because of the hypothesis.

\[ |Q_n| = \left| \bigcup_{i=1}^{n} (v_n - v_i - l_i, v_n - v_i + l_n) \right| = \left| \bigcup_{i=1}^{n} (-v_i - l_i, -v_i + l_n) \right| \\
= \left| \bigcup_{i=1}^{n} (v_i - l_n, v_i + l_i) \right| = \left| \bigcup_{i=1}^{n} (v_i, v_i + l_i + l_n) \right|. \]

By (viii),

\[ (l_1 + \cdots + l_n) \leq |Q_n| \leq C(l_1 + \cdots + l_n). \]

Write \( Q_{n-1} \) as union of disjoint intervals

\[ \bigcup_{i=1}^{n-1} (v_i, v_i + l_i + l_{n-1}) \quad (ix) \]

\[ = (v_1, v_1 + l_1 + l_{n-1}) \cup (v_2, v_2 + l_2 + l_{n-1}) \cup \cdots \cup (v_k, v_k + l_k + l_{n-1}) \quad (x) \]
for some \( v_1 < v_{i_1} < v_{i_2} < \cdots < v_{i_k} > v_1 \). Then

\[
|Q_{n-1}| = v_1 + l_1 - v_1 + v_{i_2} - v_{i_1 + 1} + l_{i_2} + \cdots + v_{i_k} - v_{i_{k-1} + 1} + l_k + kl_{n-1}
\]

\[
\leq C(l_1 + \cdots + l_{n-1}).
\]

This implies that

\[
k_{l_{n-1}} \leq |Q_n| \leq C(l_1 + \cdots + l_{n-1}).
\]

Thus, \( k \) is bounded by

\[
k \leq C \frac{(l_1 + \cdots + l_{n-1})}{l_{n-1}}.
\]

But since \( (l_1 + \cdots + l_n) \leq Kl_n \) for all \( n, k \leq CK \).

The correct factors associated to

\[
\sup_{(n,s), 0 \leq s \leq l_n} \frac{1}{l_1 + \cdots + l_{n-1} + s} \sum_{j=vn}^{v_{n+s-1}} T^j f(x)
\]

are

\[
\tilde{Q}_m = \left| \bigcup_{i=1}^{n-1} \bigcup_{j=1}^{l_i} (v_i, v_i + j + s) \cup \bigcup_{j=1}^{s-1} (v_n, v_n + j + s) \right|
\]

\[
= \left| \bigcup_{i=1}^{n-1} (v_i, v_i + l_i + s) \cup (v_n, v_n + 2s - 1) \right|
\]

where \( m = v_n + s - 1, 1 \leq s \leq l_n \). From (x), if \( s \leq l_{n-1} \), \( \tilde{Q}_m \leq |Q_{n-1}| + 2s \); otherwise,

\[
\tilde{Q}_m \leq |Q_{n-1}| + (k + 2)s.
\]

In both cases we obtain

\[
\tilde{Q}_m \leq C(l_1 + \cdots + l_{n-1}) + (k + 2)s \leq k'(l_1 + \cdots + l_{n-1} + s).
\]
Therefore, since

\[ \sup_{n \in \mathbb{N}} \left| \frac{1}{Q_m} \sum_{j \in (1, m) \cap \bigcup B_i} T^j f(x) \right| \]

is a weak (1,1) operator, then

\[ \sup_{n \in \mathbb{N}} \left| \frac{1}{(l_1 + \cdots + l_{n-1} + s)} \sum_{j \in (1, m) \cap \bigcup B_i} T^j f(x) \right| \]

is also a weak (1,1) operator.
CHAPTER V

Discrepancy of behavior in $L^p$ spaces of perturbed sequences

In this chapter, sequences of natural numbers with the property that the averages along them diverge in some $L^p$ spaces and converge in others are given. Emerson [20] developed a useful tool to produce such examples. A. Bellow used it in [2] to construct sequences $\{n_k\}_{k \in \mathbb{N}}$ that, for a fixed $p$, $1 < p < \infty$, are universally bad in $L^q$ for $1 \leq q < p$ but universally good for all $q \geq p$. The question, however, of whether it is possible to produce sequences that are bad in $L^1(X)$ but good in all $L^p(X)$ for $p > 1$ and sequences which are bad in all $L^p(X)$ for all $p < \infty$ but good in $L^\infty(X)$ was not answered. We use Emerson's tool to produce such examples. Indeed, we show three examples.

1. For any $p$, $1 \leq p < \infty$, there are sequences $\{n_k\}_{k \in \mathbb{N}}$ which are universally bad on $L^p$ but universally good in $L^q$, $q > p$.

2. There are sequences $\{n_k\}_{k \in \mathbb{N}}$ which are universally bad in $L^p$ for all $1 \leq p < \infty$ but universally good in $L^\infty$.  

91
3. For fixed $p$, $1 < p < \infty$, there are sequences $\{n_k\}_{k \in \mathbb{N}}$ which are universally bad on $L^q$, $1 \leq q \leq p$, but universally good on $L^q$, $q > p$ so that $\lim_{k \to \infty} n_{k+1} - n_k = \infty$.

From 1) we obtain sequences which are universally bad on $L^1$ but universally good in $L^q$ for all $q > 1$. Example 2) is the most curious because one can get convergence of the averages for bounded functions but not for non-bounded ones. The first two examples plus that of A. Bellow were constructed by starting with a sequence of blocks of consecutive numbers which is universally good in $L^1$ and spoiling it. Example 3) shows that, the block averages, in fact, are not essential to the construction. Moreover, the sequence can be chosen so that the gap between consecutive terms increases to $\infty$.

5.1 Examples

Definition 5.1.1 A sequence $\{n_k\}_{k \in \mathbb{N}} \in \mathbb{N}$ is called universally good in $L^p$ if for all aperiodic (ergodic) dynamical systems $(X, \beta, m, T)$, the averages

$$A_k f(x) = \frac{1}{k} \sum_{j=1}^{k} T^{n_j} f(x)$$

converges a.e. for all $f \in L^p$; and is called universally bad in $L^p$ if for all aperiodic (ergodic) dynamical systems, the averages $A_k f$ do not converge a.e. for some $f \in L^p$.

Recall that a dynamical system $(X, \beta, m, T)$ is "aperiodic" if the set of periodic points have measure zero.

The definitions of good and bad universal are related to the Conze's Principle. Fix a sequence $n = \{n_k\}_{k \in \mathbb{N}}$ and $1 \leq p < \infty$. By Sawyer's version of the Banach Principle
(Sawyer [38]), for any dynamical system \((X, \beta, m, T)\) the averages \(A_k f\) converge a.e. in \(L^p(X)\) if and only if there exists a finite constant \(C\) such that

\[
m(x \in X : \sup_{N \in \mathbb{N}} \frac{1}{N} \sum_{k=0}^{N-1} T^{nk} f(x) > \lambda) \leq C \frac{\|f\|_p}{\lambda^p}
\]

for all \(f \in L^p(X)\). Therefore, one can associate to \(n\) a minimal constant \(0 < C(n, p) \leq \infty\) such that

\[
m(x \in X : \sup_{N \in \mathbb{N}} \frac{1}{N} \sum_{k=0}^{N-1} T^{nk} f(x) > \lambda) \leq C(n, p) \frac{\|f\|_p}{\lambda^p}
\]

for all \(f \in L^p(X), \lambda > 0\) and all aperiodic (ergodic) dynamical systems \((X, \beta, m, T)\).

**Theorem 5.1.2** (Conze's Principle, [17]). For any given sequence \(n = \{n_k\}_{k \in \mathbb{N}}\), the associated minimal constant \(C(n, 1)\) is finite if and only if there exists an aperiodic (ergodic) dynamical system \((X, \beta, m, T)\) for which

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} T^{nk} f(x)
\]

exists a.e. for each \(f \in L^1(X)\).

**NOTE:** It follows from this principle that a sequence \(\{n_k\}_{k \in \mathbb{N}}\) is universally bad in \(L^p\) if and only if \(C(n, p) = \infty\). It is clear, after Sawyer's Banach Principle, that if \(n\) is universally bad in \(L^p\), then \(C(n, p) = \infty\). Vice versa, if \(C(n, p) = \infty\), the constant associated to each ergodic system must be infinite, because otherwise, if it is finite for one of them, one can transfer the \(L^p\) estimate to actions on \(L^p(\mathbb{Z})\) (Transfer Principle) and transfer it back, via a Rohlin tower construction, to any other ergodic system.
We next introduce some notation that will be used later on. Let $T_\theta:[0,1] \to [0,1]$ defined by $T_\theta(x) = x + \theta \mod 1$, where $\theta$ is an irrational number. Define

$$p_k = \sum_{j=1}^{k} \frac{1}{p} \mod 1 \quad \text{and} \quad J_k = [\tilde{p}_k, p_{k+1}] \mod 1.$$ (i)

Then, for every $x \in [0,1]$, there exist infinitely many $k$'s such that $x \in J_k$.

Given a family of pairwise disjoint blocks $B_k$, consider the elements of $\bigcup_{k=1}^{\infty} B_k$ arranged in increasing order without repetitions. This set can be thought of as a sequence and will be referred to as the "block sequence $\bigcup_{k=1}^{\infty} B_k$". Suppose $\{D_k\}$ is another sequence of pairwise disjoint blocks and disjoint from the $B_k$'s such that both sequences intertwine (...[......]...[.]...[......]...). We can then form the sequence $\bigcup_{k=1}^{\infty} (B_k \cup D_k)$ which is called a perturbation of the original.

Let $\{B_k\}$ and $\{D_k\}$ be as above and let $l_k = |B_k|$ and $d_k = |D_k|$.

**Theorem 5.1.3 (BELLOW [2].)** Let $B_k$ and $D_k$ be as above. If the sequence $\bigcup_{k=1}^{\infty} B_k$ is a good universal sequence in $L^p(X)$ and

$$\sum_{k=1}^{\infty} \left( \frac{d_1 + \ldots + d_k}{l_1 + \ldots + l_k} \right)^p < \infty,$$

then the sequence $\bigcup_{K=1}^{\infty} (B_k \cup D_k)$ is also good universal in $L^p(X)$.

This theorem illustrates how much a sequence can be perturbed such that the resulting sequence has the same behavior. Emerson [20] proved a similar theorem with the difference that Bellow's version proves convergence for the whole sequence allowing to stop inside the blocks $B_k$ and $D_k$ whereas Emerson's version only considers averages on the whole blocks $B_k$ and $D_k$, without considering whether one could stop inside
the blocks. In section 4.2, we discussed when the convergence of anyone of these two types of averages implies convergence of the other.

PROOF. (Here is a simpler proof of Bellow’s Theorem).

Let \( C = \bigcup_{k=1}^{\infty} (B_k \cup D_k) \), \( b_n = |\bigcup_{k=1}^{\infty} B_k \cap [0,n]| \) and \( c_n = |\bigcup_{k=1}^{\infty} D_k \cap [0,n]| \). The averages

\[
A_n f(x) = \frac{1}{|C \cap [0,n]|} \sum_{u \in C \cap [0,n]} T^u f(x)
\]

can be written as a convex combination of averages on the \( B_k \)’s and on the \( D_k \)’s,

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

\[
= \frac{b_n}{b_n + c_n} A_n^B f(x) + \frac{c_n}{b_n + c_n} A_n^D f(x).
\]

Observe that

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

\[
\frac{d_1 + \cdots + d_{k-2} + r_{k-1}}{l_1 + \cdots + l_{k-1}} & \text{if } k \text{ is the smallest integer such that } \\
B_k \text{ is not contained in } [0,n], \\
B_k \cap [0,n] = \emptyset \text{ and } B_{k-1} \subset [0,n]
\]

where \( 0 \leq r_{k-1} \leq d_{k-1} \) and \( 0 \leq s_k \leq l_k \). In the first situation we have the picture

\[
\begin{array}{cccccccc}
\cdots & [\cdots] & \cdots & [\cdots] & \cdots & [\cdots] & \cdots & [\cdots] \\
D_{k-2} & B_{k-1} & D_{k-1} & B_k & \cdots \cdots \cdots \cdots & B_k & D_k
\end{array}
\]

and in the second

\[
\begin{array}{cccccccc}
\cdots & [\cdots] & \cdots & [\cdots] & \cdots & [\cdots] & \cdots & [\cdots] \\
D_{k-2} & B_{k-1} & D_{k-1} & B_k & \cdots \cdots \cdots \cdots & B_k & D_k
\end{array}
\]

In both cases we have,

\[
\frac{c_n}{b_n} = \frac{d_1 + \cdots + d_{k-1}}{l_1 + \cdots + l_{k-1}} \to 0
\]
by the hypothesis of the theorem. Therefore,
\[ \lim_{N \to \infty} \frac{b_n}{b_n + c_n} = 1. \]  
(ii)

Since \( \lim_{N \to \infty} A_n^B f(x) \) exists a.e. for all \( f \in L^p(X) \), then
\[ \lim_{N \to \infty} \frac{b_n}{b_n + c_n} A_n^B f(x) \]
also exists and has the same limit a.e. Also, by (ii), \( [c_n/(b_n + c_n)] A_n^D f(x) \) converges a.e. for all bounded functions because \( \lim_{n \to \infty} c_n/(b_n + c_n) = 0 \). Therefore, to prove that it converges for all functions in \( L^p(X) \), it suffices to prove a maximal estimate (Banach’s Principle). Notice that if \( n \) is a point which is not in any \( B_k \) and \( D_k \), then
\[ \frac{c_n}{b_n + c_n} A_n^D f = \frac{c_m}{b_m + c_m} A_m^D f \]
where \( m \) is the closest point in \( \bigcup (B_k \cup D_k) \) to the left of \( n \). Therefore, it is sufficient to consider points in \( \bigcup (B_k \cup D_k) \). In case that \( n \in D_{k-1} \cup B_k \), \( c_n \) and \( b_n \) can be written as \( c_n = d_1 + \cdots + d_{k-2} + r_{k-1} \) where \( r_{k-1} = |D_{k-1} \cap [0, n]| \) and \( b_n = l_1 + \cdots + l_{k-1} + s_k \) where \( s_{k-1} = |B_k \cap [0, n]| \). Then,
\[ \frac{c_n}{b_n + c_n} |A_n^D f| \leq \frac{1}{b_n + c_n} \sum_{u \in \bigcup_{j=1}^{k-1} D_j} |T^u f| \]
\[ \leq \frac{1}{l_1 + \cdots + l_{k-1}} \sum_{u \in \bigcup_{j=1}^{k-1} D_j} |T^u f| \]
\[ \leq \left[ \frac{d_1 + \cdots + d_{k-1}}{l_1 + \cdots + l_{k-1}} \right] \frac{1}{d_1 + \cdots + d_{k-1}} \sum_{u \in \bigcup_{j=1}^{k-1} D_j} |T^u f|. \]

And
\[ \left[ \sup_{n \in \mathbb{N}} \frac{c_n}{b_n + c_n} |A_n^D f| \right]^p \leq \sup_{k \in \mathbb{N}} \left[ \sup_{n \in (D_{k-1} \cup B_k)} \frac{c_n}{b_n + c_n} |A_n^D f| \right]^p \]
\[ \leq \sup_{k \in \mathbb{N}} \left[ \left( \frac{d_1 + \cdots + d_{k-1}}{l_1 + \cdots + l_{k-1}} \right) \frac{1}{d_1 + \cdots + d_{k-1}} \sum_{u \in \bigcup_{j=1}^{k-1} D_j} |T^u f| \right]^p. \]
Thus,

$$\left\| \sup_{n \in \mathbb{N}} \frac{c_n}{b_n + c_n} |A_n^f||p \right\|_p \leq \sum_{k=1}^{\infty} \left( \frac{d_1 + \cdots + d_{k-1}}{l_1 + \cdots + l_{k-1}} \right)^p \frac{1}{(d_1 + \cdots + d_{k-1})} \sum_{u \in \bigcup_{j=1}^{k-1} D_j} |T^u f||p$$

$$\leq \sum_{k=1}^{\infty} \left( \frac{d_1 + \cdots + d_{k-1}}{l_1 + \cdots + l_{k-1}} \right)^p \|f\|_p = C\|f\|_p$$

which finishes the proof.

Proposition 5.1.4 Let $B_k$ and $D_k$ be as above. Let $l_k = |B_k|$ and $d_k = |D_k|$. Suppose $l_1 + \cdots + l_k \leq C l_{k+1}$ for all $k$, $d_k = c_k l_k$ and that both sums are finite $\sum_{k=1}^{\infty} (l_k/l_{k+1})^p < \infty$, $\sum_{k=1}^{\infty} c_k^p < \infty$. If $\bigcup_k B_k$ is a good universal sequence in $L^p$ then the perturbed sequence $\bigcup_{k=1}^{\infty} (B_k \cup D_k)$ is also good universal in $L^p$.

PROOF. This proposition is an application of the previous theorem.

Since $\sum_{k=1}^{\infty} c_k^p < \infty$, $\lim_{k \to \infty} c_k = 0$. Choose $k_0$ so that $c_k < 1$ for all $k \geq k_0$. Then

$$\frac{d_1 + \cdots + d_k}{l_1 + \cdots + l_k} \leq \frac{d_1 + \cdots + d_{k-1}}{l_1 + \cdots + l_{k-1}} + \frac{d_{k-1} + d_k}{l_1 + \cdots + l_{k-1}} \leq \frac{C_0}{l_k} + \frac{C l_{k-1}}{l_k} + \frac{d_k}{l_k} \leq \frac{C_0}{l_k} + \frac{C l_{k-1}}{l_k} + c_k$$

Since $\sum_{k=1}^{\infty} (l_k/l_{k+1})^p < \infty$ and $\sum_{k=1}^{\infty} c_k^p < \infty$, then

$$\sum_{k=1}^{\infty} \left( \frac{d_1 + \cdots + d_k}{l_1 + \cdots + l_k} \right)^p < \infty.$$

The result follows from Theorem 5.1.3.
Lemma 5.1.5 If \( B_k = [n_k, n_{k+1}, \ldots, n_{k+l_k-1}] \) and there exists \( r \) such that \( l_k \geq n_{k-r} \) for all big enough values of \( k \), then the sequence \( \bigcup_{k=1}^{\infty} B_k \) is a good universal sequence in \( L^1 \).

PROOF. By the Moving Averages Theorem 2.2.2, it suffices to prove that there exists a constant \( C \) such that

\[
| \bigcup_{k=1}^{N} [n_k - \lambda + l_k, n_k + \lambda - l_k] | \leq C\lambda
\]

where \( N \) is the largest integer with \( N \gamma \leq \lambda \). Indeed,

\[
| \bigcup_{k=1}^{N} [n_k - \lambda + l_k, n_k + \lambda - l_k] | \leq | \bigcup_{k=1}^{N-r} [n_k - \lambda + l_k, n_k + \lambda - l_k] |
\]

\[
+ \sum_{N-r+1}^{N} | [n_k - \lambda + l_k, n_k + \lambda - l_k] |
\]

\[
\leq | [-\lambda, n_{N-r} + \lambda - l_{N-r}] | + \sum_{N-r+1}^{N} 2(\lambda - l_k) + 1
\]

\[
\leq -n_{N-r} + 2\lambda + 2r + r
\]

\[
\leq l_N + 2(r+1)\lambda + r + 1 \leq (3r+4)\lambda
\]

since \( \lambda \geq 1 \) (\( \lambda \geq l_N \)). This proves (iii) with \( C = (3r+4) \).

This is the basic construction of a good universal sequence in \( L^1(X) \) and a perturbation that preserves that property. The construction of the blocks \( B_k \) is similar to an example of moving averages in B–J–R [6], Corollary 3. The same construction is in Bellow [2] and is already present in the work of Emerson [20].

Fix \( q > 1 \) and a sequence \( \{c_k\} \) with \( \sum_{k=1}^{\infty} c_k^q < \infty \). Then construct the blocks \( B_k \) and
$D_k$ as follows:

Let $n_1 = 1$, $l_1 = 1$, $u_1 = 1$, $d_1 = 1$ and proceed by induction. Suppose $n_1, \ldots, n_{k-1}$; $l_1, \ldots, l_{k-1}$; $D_1, \ldots, D_{k-1}$; $d_1, \ldots, d_{k-1}$ have been already defined. The block $B_k$ will consist of $l_k$ consecutive numbers, $B_k = [n_k, n_k + 1, \ldots, n_k + l_k - 1]$, where $n_k$ is chosen to the right of $D_{k-1}$ and $l_k$ is taken so large that it satisfies:

(a) $l_k \geq n_{k-1}$

(b) $l_k \geq k \cdot l_{k-1} \geq l_1 + \ldots + l_{k-1}$

Condition 1.a) guarantees that the sequence defined by the blocks $B_k$ is good universal in $L^1$. (This is similar to the construction in B-J-R [6].) As in the proposition, set $d_k = c_k l_k$, and let $D_k$ consist of $d_k$ numbers to the right of $n_k + l_k$. Then, the sequence $\bigcup_{k=1}^\infty (B_k \cup D_k)$ is good universal in $L^q$. (This construction is in Emerson [20] and Bellow [2].)

**Proposition 5.1.6** Let $p \geq 1$ and let $\{B_k\}$ be a family of blocks such that the sequence defined by them is good universal in $L^p(X)$. Let $l_k = |B_k|$ and assume that $l_1 + \cdots + l_{k-1} \leq C l_k$. Suppose that in between the blocks $B_k$'s one can insert blocks $D_k$'s of length $d_k$ such that $d_k = c_k l_k$ where $c_k < 1$ for all $k$ but $\lim_{k \to \infty} c_k (k/\log k^2)^{1/p} = \infty$, and such that for all $u \in D_k$, $T^p_\theta(p_k) < 1/k$ for some $\theta$ irrational. Then the perturbed sequence $\bigcup_{k=1}^\infty (B_k \cup D_k)$ is bad universal in $L^p$.

**Proof.** By Conze's principle, it is enough to consider the dynamical systems $(X, \beta, m, T)$ where $X = [0, 1)$ with the Lebesgue measure and $T(x) = x + \theta \mod 1$, $\theta$ irrational. Let $f(x) = (1/x \log^2(x/2))^{1/p} \chi_{(0,1)}(x)$. Then $f \in L^p(X)$. By construction of the $J_k$'s and $p_k$'s in i, if $x \in J_k$ and $u \in D_k$, $T^p_\theta(x) < 2/k$, so $T^p_\theta f(x) \geq (k/2 \log^2 k)^{1/p}$. 

Since $d_k < l_k$ and $l_1 + \cdots + l_{k-1} \leq Cl_k$, we have

$$l_1 + \cdots + l_k + d_1 + \cdots + d_k \leq 2(l_1 + \cdots + l_k) \leq 2(C + 1)l_k = C'l_k.$$  

Then

$$\frac{1}{(l_1 + \cdots + l_k + d_1 + \cdots + d_k)} \sum_{m \in ((B_l \cup D_l) \cup \cdots \cup (B_k \cup D_k))} T_{\delta}^m f(x)$$

$$> \frac{1}{C'I_k} \sum_{m \in D_k} T_{\delta}^m f(x)$$

$$\geq c \frac{d_k}{l_k} \left( \frac{k}{2 \log^2 k} \right)^\frac{1}{p} = cC_k \left( \frac{k}{\log^2 k} \right)^\frac{1}{p}.$$  

Since $\lim_{k \to \infty} c_k(k/\log^2 k)^{\frac{1}{p}} = \infty$ and every point in $(0,1)$ is in infinitely many $J_k$'s,

$$\sup_{k \in \mathbb{N}} \frac{1}{(l_1 + \cdots + l_k + d_1 + \cdots + d_k)} \sum_{m \in ((B_l \cup D_l) \cup \cdots \cup (B_k \cup D_k))} T_{\delta}^m f(x) = \infty$$

for every $x$ in $(0,1)$. Thus, $C(\bigcup_k (B_k \cup D_k), p) = \infty$ in Conze's Principle and the sequence $\bigcup_k (B_k \cup D_k)$ is bad universal in $L^p(X)$.

**Example 1:**

In the basic construction above, let the blocks $D_k$ consist of the first $d_k$ numbers to the left of $B_k$ such that $T_{\delta}^m(p_k) < 1/k$. For a fixed $p \geq 1$ take $c_k = (\log^3 k/k)^{\frac{1}{p}}$. Then, for all $q > p$, $\sum_{k=1}^\infty c_k^q < \infty$. Therefore, $\bigcup_{k=1}^\infty (B_k \cup D_k)$ is a good universal sequence in $L^q$. But since $\lim_{k \to \infty} c_k(k/\log^2 k)^{\frac{1}{p}} = \infty$, the new sequence is bad universal in $L^p$.

**Theorem 5.1.7** Let $\{B_k\}$ be a sequence of pairwise disjoint, consecutive blocks such that the sequence they define is good universal in $L^\infty$. Let $\{D_k\}$ be a perturbation of
the $B_k$'s and denote $l_k = |B_k|$, $d_k = |D_k|$. If
\[
\lim_{k \to \infty} \frac{d_1 + \ldots + d_k}{l_1 + \ldots + l_k} = 0
\]
then the perturbed sequence is also good universal in $L^\infty$.

**Proof.** As in the proof of Theorem 5.1.3, let $C = \bigcup_{k=1}^\infty (B_k \cup D_k)$, $b_n = |\bigcup_{k=1}^n B_k \cap [0, n]|$ and $c_n = |\bigcup_{k=1}^n D_k \cap [0, n]|$. The averages
\[
A_n f(x) = \frac{1}{|C \cap [0, n]|} \sum_{u \in C \cap [0, n]} T^u f(x)
\]
can be written as a convex combination of averages on the $B_k$'s and on the $D_k$'s,
\[
A_n f(x) = \frac{b_n}{b_n + c_n} \left[ \frac{1}{b_n} \sum_{u \in B_k \cap [0, n]} T^u f(x) \right] + \frac{c_n}{b_n + c_n} \left[ \frac{1}{c_n} \sum_{u \in D_k \cap [0, n]} T^u f(x) \right]
\]
\[
= \frac{b_n}{b_n + c_n} A^B_n f(x) + \frac{c_n}{b_n + c_n} A^D_n f(x).
\]
In the proof of Theorem 5.1.3, it was shown that
\[
\lim_{k \to \infty} \frac{c_n}{b_n} = 0
\]
which gives
\[
\lim_{k \to \infty} \frac{b_n}{b_n + c_n} = 1 \quad \text{and} \quad \lim_{k \to \infty} \frac{c_n}{c_n + b_n} = 0.
\]
Hence, $\lim_{n \to \infty} A_n f(x)$ exists a.e. for all $f \in L^\infty(X)$.

**Example 2:**

Build blocks $B_k$'s and $D_k$'s as in example 1, but now let $d_k = l_k / \log k$. 
Then, \( \lim_{k \to \infty} d_k/l_k = 0 \). So, by the above theorem, the sequence \( \bigcup_{k=1}^{\infty} (B_k \cup D_k) \) is good universal in \( L^\infty \). However, for any \( \alpha > 0 \),

\[
\lim_{k \to \infty} \frac{d_k}{(l_1 + \ldots + l_k + d_1 + \ldots + d_k)^{k^\alpha}} = \infty.
\]

This forces the sequence to be universally bad in all \( L^p \), \( p < \infty \). Indeed, let \( 0 < \alpha < 1 \) and

\[
f(x) = \begin{cases} 
\frac{1}{x^\alpha} & \text{if } x \in (0,1] \\
0 & \text{if } x = 0
\end{cases}
\]

Then \( f \in L^p(X) \) for all \( 1 \leq p < \frac{1}{\alpha} \). By the construction of the blocks \( D_k \), \( f(T^u(x)) \geq (\frac{k}{2})^\alpha \) for all \( x \in J_k, u \in D_k \). Therefore,

\[
\frac{1}{l_1 + \ldots + l_k + d_1 + \ldots + d_k} \sum_{u \in \bigcup_{j=1}^k (B_j \cup D_j)} f(T^u(x)) \geq \frac{d_k}{(l_1 + \ldots + l_k + d_1 + \ldots + d_k)} (\frac{k}{2})^\alpha.
\]

And by the above, this diverges as \( k \to \infty \). Consequently the maximal function for \( f \) diverges a.e.

\[\blacksquare\]

From Theorem 5.1.3 it is clear that the blocks \( B_k \) need not consist of consecutive numbers. However, the sequence \( \bigcup_{k \in \mathbb{N}} B_k \) should be good in some \( L^p(X) \) space. Following this idea we can construct sequences with gaps going to infinity with the same behavior as the sequences in A. Bellow's example [2].

**Proposition 5.1.8** Let \( \theta \in (0,1) \) be irrational. Then there is a constant \( C > 0 \) depending only on \( \theta \) such that the gaps of the sequence \( \{m_j\}_{j \in \mathbb{N}} = \{n \in \mathbb{N} : \{n\theta\} < \frac{1}{k}\} \) grow at most linearly in \( k \), that is, \( m_{j+1} - m_j \leq C \ k \) for all \( j \).
PROOF. Denote \{x\} the fractional part of x, i.e. \{x\} = x - \lfloor x \rfloor.

Let \( M \in \mathbb{N} \) be the first number so that \( \{M\theta\} \in (1/2k, 1/k) \) and denote \( M_0 = -\lfloor M\theta \rfloor \in \mathbb{Z} \). We will estimate \( m_{j+1} - m_j \). Fix \( j \) and let \( N = -\lfloor m_j\theta \rfloor \). Then

\[
M_0 + M\theta \in \left( \frac{1}{2k}, \frac{1}{k} \right) \quad \text{and} \quad N + m_j\theta \in (0, \frac{1}{k}).
\]

Then, for any \( R \in \mathbb{N} \),

\[
\frac{R}{2k} < N + m_j\theta + RM_0 + RM\theta < \frac{R + 1}{k}.
\]

If \( R < k - 1 \), then \( N + m_j\theta + RM_0 + RM\theta < 1 \). But if \( R > 2k \), then \( N + m_j\theta + RM_0 + RM\theta > 1 \).

Let \( R_0 \) be the first integer number in \([k-1, 2k]\) such that \( N + m_j\theta + R_0 M_0 + R_0 M\theta > 1 \). But then

\[
1 < N + m_j\theta + R_0 M_0 + R_0 M\theta
\]

\[
= N + m_j\theta + (R_0 - 1)M_0 + (R_0 - 1)M\theta + M_0 + M\theta
\]

\[
< 1 + M_0 + M\theta
\]

\[
< 1 + \frac{1}{k}.
\]

This means that \( m_j + R_0 M \) is an element of the sequence \((m_j)\). Therefore \( m_{j+1} - m_j \leq R_0 M \leq 2Mk \). Put \( C = 2M \).

Example 3:
Consider the sequence of cubes \( \{n^3\} \) which is good universal in \( L^p \) for all \( p > 1 \). (See
Divide this sequence into blocks $B_k$ of length $l_k$, $B_k = \{n_k^3, (n_k+1)^3, \ldots, (n_k+l_k-1)^3\}$, and perturb it by inserting blocks $D_k$ of "bad" points of length $d_k$ in between the $B_k$'s. Choose the lengths $l_k$'s so that

$$l_k \geq k l_{k-1}$$

and let

$$d_k = l_k \left( \frac{\log^3 k}{k} \right)^{\frac{1}{p}}.$$ 

The blocks $D_k$ consist of $d_k$ numbers $u$ in the interval $I_k = [(n_k + l_k - 1)^3, (n_k + l_k)^3]$ such that $T^{u}(p_k) \in [0, \frac{1}{k}]$, and such that $d(u, \{n^3\}) > k/4$. (It could happen that for a few $k$'s we have less than $d_k$ such numbers. In that case we take as many as we can.) This is possible to do because from Proposition 5.1.8, for fixed $\theta$, the gaps of the sequence $(m_j) = \{n \in N : \{n\} < \frac{1}{k}\}$ are smaller than $C_k$ for some constant $C$ depending on $\theta$. By (a), the $l_k$'s are growing at a rate faster than $k!$. Therefore, since the interval $I_k$ has length $\geq l_k^2$, $|I_k|/k \geq (k - 1)! \cdot l_k$. Then, for $k$ big enough, there is certainly room to choose $(\log^3 k/k)^{\frac{1}{2}} l_k$ points inside $I_k$ with the desired property. Moreover, the elements of $D_k$ can be chosen so that they are $k/4$ apart from each other and from the sequence of cubes. Choose $u_1$ to be the first point with the desired property to the right of $n_k^3$ which is at least at distance $k$ from the sequence of cubes. Suppose you have already chosen $u_1, u_2, \ldots, u_j$; take $u_{j+1}$ to the right of $u_j$, at distance $k$ from it. By Lemma 5.1.5, $u_{j+1} \in (u_j, u_j + (C + 1)k]$. And by 5.1, for $k$ large enough, it is possible to continue the process till we have chosen $d_k$ numbers with the desired property in $I_k$.

With the choices of $B_k$ and $D_k$ we have $d_k = c_k l_k$, $c_k = \left( \frac{\log^3 k}{k} \right)^{\frac{1}{p}}$ and $\sum_{n=0}^{\infty} c_k^q < \infty$ for all $q > p$. By Proposition 5.1.4, the perturbed sequence
∪_{k=1}^{∞}(B_k \cup D_k) is universally good in \( L^q \) for all \( q > p \). Also,

\[
\lim_{k \to \infty} c_k \left( \frac{k}{2 \log^2(k)} \right)^{\frac{1}{p}} = \infty.
\]

Then, by Proposition 5.1.6, the perturbed sequence \( ∪_{k=1}^{∞}(B_k \cup D_k) \) is bad universal in \( L^p \).

Note: The argument used here to construct the blocks \( D_k \) cannot be repeated for the sequence of squares.

From this example it follows that the sequence of cubes, which is good universal in \( L^p(X) \) for all \( p > 1 \), loses this property when slightly modified. The really important question is whether this sequence is bad universal in \( L^1(X) \). More generally, Wierdl has conjectured that if \( \{n_k\}_{k \in \mathbb{N}} \) is a sequence whose gaps increase to infinity, then the sequence is bad universal in \( L^1(X) \). Unfortunately, the methods used in this chapter do not answer this question. It is not possible to perturb the sequence of cubes so that the perturbed sequence has the same \( L^1(X) \) behavior as the sequence of cubes (as in Theorem 5.1.3) but for which the perturbation makes the perturbed sequence to be bad universal in \( L^1(X) \).
Appendix A

Appendix to Chapter 1

The following lemma is a useful fact about contractions in Hilbert spaces. We give a proof for operators in $L^2$, but the same proof works for operators in any Hilbert space.

**Lemma A.0.9** Let $U$ be an operator on $L^2(X)$ with $\|U\| \leq 1$. Then

$$L^2(X) = \{ f \in L^2(X) : Uf = f \} \oplus \text{cl}\{-f - Uf : f \in L^2(X)\}.$$

**Proof.** Set

$$I = \{ f \in L^2(X) : Uf = f \} \text{ and } L = \{ f - Uf : f \in L^2(X)\}.$$  

If $f \in L^\perp$, then for all $g \in L^2(X)$, $0 = < f, g - Ug > = < f - U^*f, g >$. So $f = U^*f$.

But then $f = Uf$ also because

$$\|f - Uf\|_2^2 = \|f\|_2^2 + \|Uf\|_2^2 - < f, Uf > - < Uf, f >$$

$$= \|f\|_2^2 + \|Uf\|_2^2 - < U^*f, f > - < f, U^*f >$$

$$= \|Uf\|_2^2 - \|f\|_2^2.$$  

Since $U$ is a contraction, we obtain $\|f - Uf\|_2^2 \leq 0$.

It follows that $L^\perp \subset I$ and that, $I \subset L^\perp$ because if $f \in I$, the same argument as
above shows that \( f \) is also invariant under \( U^* \), \( U^* f = f \). Thus, for all \( g \in L^2(X) \),

\[
< f, g - U g > = < f - U^* f, g > = 0,
\]

which implies that \( f \in L^1 \). This completes the proof.
Appendix B

Appendix to Chapter 2

We present the proof of the part of the Theorem by Baum and Katz on Large Deviations that we make use of in the previous chapters.

**Theorem B.0.10** (BAUM–KATZ [1]). Let \( \{X_i\} \) be a family of i.i.d. random variables with \( E(X_1) = 0 \). If \( t > 1, r > 1 \) and \( \frac{1}{2} < \frac{r}{t} \leq 1 \), the following are equivalent:

1. \( E|X_i|^t < \infty \)

2. \( \sum_{n=1}^{\infty} n^{r-2}P(|S_n| > n^{\frac{r}{t}} \varepsilon) < \infty \) for all \( \varepsilon > 0 \).

For our purposes, we only need \((1) \Rightarrow (2)\).

**Proof.** Let \( A_n = \{|S_n| > n^{\frac{r}{t}} \varepsilon\} \). Without loss of generality we can assume \( \varepsilon = 1 \).

Let \( i \) be such that \( 2^i \leq n < 2^{i+1} \). We can include \( A_n \) into three sets

\[
A_n^1 = \{|X_k| > 2^{(i-1)r/t} \text{ for some } k \leq n\}
\]

\[
A_n^2 = \{|X_{k_1}| > n^{\gamma r/t}, |X_{k_2}| > n^{\gamma r/t} \text{ for at least two } k_1, k_2 \leq n\}
\]

where \( \gamma \) is chosen so that

\[
\frac{r+1}{2r} < \gamma < 1, \quad 1 - \gamma r < 0, \quad 1 - \frac{2\gamma r}{t} < 0 \quad \text{and} \quad \frac{1}{r} \left( \frac{t-r}{t-1} \right) \leq \gamma.
\]

108
These choices are possible because of the hypothesis on $r$ and $t$.

$$A_n^3 = \{ |\sum_{k=0}^{n'} X_k | > 2^{(i-2)r/t} \}$$

where $\sum'$ stands for the sum omitting the terms $X_k$ with $|X_k| > n^{r/t}$.

Then $A_n \subset (A_n^1 \cup A_n^2 \cup A_n^3)$. Indeed, on $(A_n^1 \cup A_n^2)^c$, $|\sum_{k=0}^n X_k|$ has at most one term, say $X_{k_0}$, with $|X_{k_0}| > n^{r/t}$. So, on $(A_n^1 \cup A_n^2)^c \cap A_n$,

$$|\sum_{k=0}^{n'} X_k | = |\sum_{k=0}^n X_k - X_{k_0} | > n^{r/t} - 2^{(i-2)r/t} > 2^{ir/t} - 2^{(i-2)r/t}$$

because $2r/t > 1$ and on $A_n^1$, $|X_{k_0}| \leq 2^{(i-2)r/t}$.

The proof of this part will be complete once we show that

$$\sum_{n=1}^{\infty} n^{r-2} P(A_n^j) < \infty \text{ for } j = 1, 2, 3.$$

For the $A_n^1$'s we have

$$\sum_{n=1}^{\infty} n^{r-2} P(A_n^1) = \sum_{n=1}^{\infty} n^{r-2} P(|X_k| > 2^{(i-1)r/t} \text{ for all } k \leq n)$$

$$= \sum_{n=1}^{\infty} n^{r-1} P(|X_1| > 2^{(i-1)r/t})$$

$$\leq \sum_{n=1}^{\infty} n^{r-1} P(|X_1| > (\frac{n}{8})^{r/t})$$

$$\leq c \int_{0}^{\infty} u^{r-1} P(|X_1| > u^{r/t}) du$$

$$= c \int_{r}^{\infty} v^{t-1} P(|X_1| > v) dv$$

$$= c E(|X_1|^t) < \infty.$$

For the $A_n^2$'s we have,

$$P(A_n^2) = P(|X_{k_1} | > n^{r/t}, |X_{k_2} | > n^{r/t} \text{ for at least two } k_1, k_2 \leq n)$$
\[
\begin{align*}
&\leq \sum_{1 \leq k_1 < k_2 \leq n} P(|X_{k_1}| > n^{\gamma r/t}, |X_{k_2}| > n^{\gamma r/t}) \\
&\leq \sum_{1 \leq k_1 < k_2 \leq n} P(|X_{k_1}| > n^{\gamma r/t})^2 \leq n^2 P(|X_{k_1}| > n^{\gamma r/t})^2 \\
&\leq n^2 \left( \frac{E(|X_1|^t)}{n^{\gamma r}} \right)^2 = cn^{2-2\gamma r}.
\end{align*}
\]

Thus,
\[\sum_{n=1}^{\infty} n^{r-2} P(A_n^2) \leq \sum_{n=1}^{\infty} n^r \frac{c}{n^{2\gamma r}} < \infty\]
since \(1 < 2\gamma r - r\) because we have chosen \(\gamma\) such that \((r+1)/2r < \gamma\).

For the \(A_n^2\)'s, let \(j\) be the smallest integer \(\geq t\) and \(M\) a positive integer to be determined later. Let
\[X_{k,n} = \begin{cases} X_n & \text{if } |X_k| \leq n^{\gamma r/t} \\
0 & \text{otherwise}, \end{cases}\]

\(\alpha_n = E(X_{k,n})\) and \(Y_{k,n} = X_{k,n} - \alpha_n\). Note that since \(r/t \geq 1\), \(\lim_{n \to \infty} \alpha_n = 0\).

\[E(\sum_{k=0}^{n} Y_{k,n})^{2M_j} \leq \sum_{k=0}^{n} E(Y_{k,n}^{2M_j}) + \cdots + \sum_{1 \leq k_1 < k_2 < \cdots < k_r} E(Y_{k_1,n}^{2}) \cdots E(Y_{k_r,n}^{2}).\]

Each term of this sum is of the form
\[\sum_{1 \leq k_1 < k_2 < \cdots < k_r} E(Y_{k_1,n}^{\beta_1}) \cdots E(Y_{k_r,n}^{\beta_r})\] (i)

where
\[\sum_{s=1}^{r} \beta_s = 2M_j.\]

The terms for which there is at least one \(k_v\) with \(\alpha_{k_v} = 1\) vanish because \(E(Y_{kn}) = 0\).

Therefore, we can assume that \(\beta_s \geq 2\) for all \(s\). Let's estimate each of the terms.
\[\sum_{k=0}^{n} E(Y_{k,n}^{2M_j}) = \sum_{k=0}^{n} E(|Y_{m,n}|^{2M_j-1}|Y_{k,n}|^t)\]
Consider, now, terms as in (i) where there is at least one \( \beta > t \). Suppose \( \beta_1, \ldots, \beta_q > t \) and \( \beta_{q+1}, \ldots, \beta_l \leq t \) with \( q + l = \tau \). Repeating the same argument as above, this term is not larger than

\[
\begin{align*}
&\leq (n^\gamma/t + |\alpha_n|)^{2Mj - t} \sum_{k=0}^{n} E(|X_{kn}| + |\alpha_n|^t) \\
&\leq cn^{2Mj\gamma/t - \tau + 1} E(|X_1|^t)
&= cn^{2Mj\gamma/t - \tau + 1}.
\end{align*}
\] (ii)

By the choice of \( \gamma \), \( (1 - 2\gamma r/t) < 0 \) and \( (1 - \gamma r) < 0 \). Since we are considering terms for which \( q \geq 1 \), the above is maximized when \( q = 1 \) and \( l = 0 \). Hence the contribution of these terms is again at most

\[
\begin{align*}
&\leq cn^{2Mj\gamma r/t - \gamma r + 1}.
&\text{(iii)}
\end{align*}
\]

Lastly, the terms for which all the \( \beta_s \leq t \) are at most \( n^\gamma \leq n^Mj \). They contribute at most

\[
\begin{align*}
&\leq cn^{Mj}.
&\text{(iv)}
\end{align*}
\]
because

$$|E(Y_{k-1,n}^\beta_1) \cdots E(Y_{k,r,n}^\beta_r)| \leq E(|X_1|^r)^p$$

where $p$ depends only on $M$, $j$ and $t$. Since we have chosen $\gamma$ with $\gamma r/t > 1/2$, we can choose $M$ so big that

$$\frac{2Mj\gamma r}{t} - \gamma r + 1 > Mj.$$ 

Then, from (ii), (iii) and (iv),

$$E(|Y_{k,n}|^{2Mj}) \leq cn^{2Mj\gamma r/t - \gamma r + 1}. \tag{v}$$

Now,

$$P(A_n^3) = P(|\sum_{k=0}^n X_k| > 2^{(i-2)r/t}) = P(|\sum_{k=0}^n X_{k,n}| > 2^{(i-2)r/t})$$

$$= P(|\sum_{k=0}^n Y_{k,n} + n\alpha_n| > 2^{(i-2)r/t}) \leq P(|\sum_{k=0}^n Y_{k,n}| > 2^{(i-2)r/t} - n\alpha_n)$$

$$\leq P(|\sum_{k=0}^n Y_{k,n}| > (\frac{n}{8})^{r/t} - n\alpha_n) = P(|\sum_{k=0}^n Y_{k,n}| > n^{r/t}[\frac{1}{8}^{r/t} - n^{1-r/t}\alpha_n])$$

but

$$|n^{1-r/t}\alpha_n| = |n^{1-r/t}(\int x dF(x) - \int_{|x| > n^{\gamma r/t}} x dF(x))|$$

$$= |n^{1-r/t}\int_{|x| > n^{\gamma r/t}} x dF(x)|$$

$$\leq \int_{|x| > n^{\gamma r/t}} |x|^{1+(t-r)/\gamma r} dF(x)$$

$$\leq \int_{|x| > n^{\gamma r/t}} |x|^t dF(x) \to 0 \text{ as } n \to \infty.$$ 

Indeed, if $t > 1$,

$$1 + \frac{(t-r)}{\gamma r} \leq t \text{ because we chose } \gamma \text{ with } \frac{1}{t-1} \leq \gamma.$$
If \( t = 1 \) then \( r = 1 \) also, so \( 1 + (t - r)/\gamma r = 1 = t \). Therefore, for \( n \) big enough,

\[
P(A_n^3) \leq P\left( |\sum_{k=0}^{n} Y_{k,n}| > cn^{r/t} \right)
\]

where \( c \) is a constant < 1. Then

\[
P(A_n^3) \leq C \frac{E(Y_{k,n})^{2Mj}}{n^{2Mjr/t}}
\]

and (v) gives

\[
n^{r-2} P(A_n^3) \leq C n^{-r-2+2Mjr/t+2Mjr/\gamma r/t-\gamma r+1} = C n^{-1+1-\gamma r+2Mj[(\gamma r/t)-1]}.
\]

The exponent of \( n \) is < -1 if \( M \) is chosen so large that

\[
\frac{r(1-\gamma)}{1-\gamma r/t} < 2Mj.
\]

In that case,

\[
\sum_{n=1}^{\infty} n^{r-2} P(A_n^3) < \infty.
\]
Appendix C

Appendix to Chapter 3

This section contains elementary facts about matrices that are needed in chapter 3.

Lemma C.0.11 If μ is a measure on a subgroup of \( \mathbb{R}^d \) with \( m_2(\mu) < \infty \), then the matrix \( \tilde{B} \) is positive definite.

Proof. It is enough to show that \( (\tilde{B}y, y) \geq 0 \) for all \( y \in \mathbb{R}^d \).

\[
(\tilde{B}y, y) = \sum_{i=1}^{d}(\tilde{B}y)_i y_i \\
= \sum_{i=1}^{d} \sum_{j=0}^{d} (B_{i,j} - a_i a_j) y_i y_j \\
= \sum_{0 \leq i,j \leq d} \int_{\mathbb{R}^d} (x_i x_j - a_i a_j) d\mu(x) y_i y_j \\
= \int_{\mathbb{R}^d} \sum_{0 \leq i,j \leq d} (x_i - a_i)(x_j - a_j) y_i y_j d\mu(x) \\
= \int_{\mathbb{R}^d} [\sum_{i=0}^{d}(x_i - a_i)y_i]^2 d\mu(x)
\]

which is positive since it is a square.

\[\blacksquare\]
Lemma C.0.12 Let $B \in M_d(\mathbb{R})$ be a positive definite matrix. Then there exists $D \in M_d(\mathbb{R})$ also positive definite, such that

$$\int_{\mathbb{R}^d} \exp(-i(x.t)) \exp(-\frac{B(t)}{2}) dt = \frac{(2\pi)^{d/2}}{\sqrt{\det B}} \exp(-\frac{D(x)}{2}).$$

**Proof.** For any matrix $K \in M_d(\mathbb{R})$, let $K(t)$ denote the quadratic form $K(t) = (Kt.t)$. Since $B$ is positive definite, there is an orthogonal matrix $L \in M_d(\mathbb{R})$ such that $M = LBL^T$ is a diagonal matrix with $m_{ii} > 0$ for all $i$. So $B(t) = (Bt.t) = (L^TMLt.t) = (MLtLt) = (Mu.u)$. Then

$$\int \exp(-\frac{B(t)}{2}) \exp(-i(x.t)) dt = \int \exp(-\frac{M(u)}{2}) \exp(-i(x.L^Tu)) |L| du.$$

Since $L$ is orthogonal, $|L| = |\det(L)| = 1$.

$$\int \exp(-\frac{B(t)}{2}) \exp(-i(x.t)) dt$$

$$= \int \exp(-\frac{M(u)}{2}) \exp(-i(Lx.u)) du$$

$$= \prod_{j=1}^{d} \int \exp(-\frac{m_{j,j}u_j^2}{2}) \exp(-i(Lx)_j u_j) du_j$$

$$= \prod_{j=1}^{d} \frac{1}{\sqrt{m_{j,j}}} \sqrt{2\pi} \exp(-\frac{1}{2m_{j,j}}(Lx)_j^2)$$

$$= \prod_{j=1}^{d} \frac{1}{\sqrt{m_{j,j}}} \exp(-\frac{1}{2} \sum_{j=1}^{d} (Lx)_j^2).$$

Thus,

$$\int \exp(-\frac{B(t)}{2}) \exp(-i(x.t)) dt = (2\pi)^{d/2} \frac{1}{\sqrt{\det B}} \exp(-\frac{1}{2} \frac{Lx}{\sqrt{m}}^2)$$

where $x/\sqrt{m} = (x_1/\sqrt{m_{1,1}}, \ldots, x_d/\sqrt{m_{d,d}})$. 


Corollary C.0.13 Let $B \in M_d(\mathbb{R})$ be a positive definite matrix. Let

$$N(x) = \frac{1}{(2\pi)^{d/2} \eta} \int_{\mathbb{R}^d} \exp(-i(x.t)) \exp(-\frac{B(t)}{2}) dt$$

where $\eta$ is a constant such that $\int_{\mathbb{R}^d} N(x) dx = 1$. Then $\int_{\mathbb{R}^d} N(x) \exp(i x.t) dx = \exp(-\frac{B(t)}{2})$.

PROOF. Let $M$ and $L$ be as in the proof of the above lemma. Notice that since $L$ is an orthogonal matrix, $L$ is an isometry, i.e. $(x.y) = (Lx, Ly)$ for all $x, y \in \mathbb{R}^d$. Then

$$\int_{\mathbb{R}^d} N(x) dx = \frac{1}{\eta \sqrt{\det B}} \int_{\mathbb{R}^d} \exp(-\frac{1}{2} \frac{Lx}{\sqrt{m}}^2) dx$$

$$= \frac{1}{\eta \sqrt{\det B}} \int_{\mathbb{R}^d} \exp(-\frac{1}{2} \frac{u}{\sqrt{m}}^2) du$$

$$= \frac{1}{\eta \sqrt{\det B}} (2\pi)^{d/2} \frac{1}{\sqrt{\det B}} = \frac{(2\pi)^{d/2}}{\eta}.$$

Then, by the hypothesis on $\eta$, $\eta = (2\pi)^{d/2}$. Now,

$$\int_{\mathbb{R}^d} N(x) \exp(i x.t) dx = \frac{1}{(2\pi \det B)^{d/2}} \int_{\mathbb{R}^d} \exp(-\frac{1}{2} \sum_{j=1}^d \frac{(Lx)^2}{m_{j,j}}) \exp(i(x.t)) dx$$

$$= \frac{1}{(2\pi \det B)^{d/2}} \int_{\mathbb{R}^d} \exp(-\frac{1}{2} \sum_{j=1}^d \frac{(Lx)^2}{m_{j,j}}) \exp(i(Lx.Lt)) dx$$

$$= \frac{1}{(2\pi \det B)^{d/2}} \int_{\mathbb{R}^d} \exp(-\frac{1}{2} \sum_{j=1}^d \frac{u_j^2}{m_{j,j}}) \exp(i(Lt.Lt)) du$$

$$= \frac{1}{(2\pi \det B)^{d/2}} \prod_{j=1}^d \int_{\mathbb{R}} \exp(-\frac{u_j^2}{2m_{j,j}}) \exp(i(u_j Lt)_j) du_j$$

$$= \frac{1}{(2\pi \det B)^{d/2}} \prod_{j=1}^d \sqrt{2\pi m_{j,j}} \exp(-\frac{m_{j,j}(Lt)_j}{2})$$

$$= \exp(-\frac{\sum_{j=1}^d m_{j,j}(Lt)_j}{2}) = \exp(-\frac{(MLt, Lt)}{2})$$

$$= \exp(-\frac{(Bt, t)}{2}) = \exp(-\frac{B(t)}{2}).$$
Appendix D

Appendix to Chapter 4

This appendix is devoted to proving the Transfer principle. The version we present is a proof by Emerson [20] which is based on the Transfer Principle of Calderon [13].

Let $\mathcal{X}$ be a measure space with a jointly measurable action $G \times \mathcal{X} \to \mathcal{X}$ of a $\sigma$-compact group for which $m$ is $G$ invariant. Let $S$ be an operator defined on $L^1_{loc}(G)$ which is

1. sublinear
2. commutes with right translations $R_g$ ($g \in G$)
3. semilocal, i.e. there is a compact $K \subset G$ such that $\text{supp}(Sf) \subset K(\text{supp}(f))$.

To such an operator we can associate another operator $\hat{S}$ on $L^\infty(X) + L^1(X)$ in the following way: given $f$, define $F_x(g) = F(g, x) = f(T_g x)$. Fubini’s Theorem insures that $F_x$ is locally integrable for a.e. $x$. Therefore, $H_x(g) = H(g, x) = S(F_x(g))$ is a well defined continuous function on $G$ for a.e. $x$. Define $\hat{S}(x) = H(e, x)$.

**Theorem D.0.14** Let $\{S_n\}$ be a sequence of operators as above and suppose that

$$\left|\{g : \sup_{n \in \mathbb{N}} |S_n\phi(g)| > \lambda\}\right| \leq C \frac{\|\phi\|_1}{\lambda}$$
for all $f \in L^1(G)$, then

$$m(\{x : \sup_{n \in \mathbb{N}} |\hat{S}_n f(x)| > \lambda\}) \leq C \frac{\|f\|_1}{\lambda}$$

for all $f \in L^1(X)$.

Assume we have only a finite number of $n$'s. The result will follow by a passage to the limit. Note that for all $g \in G$,

$$\{x \in X : |f(x)| > \lambda\} = T_g \{x \in X : |f(T_g x)| > \lambda\}$$

and since $m$ is $G$ invariant, all the functions $F(g, x)$ are equimeasurable functions of $x$. Let $M \phi = \sup_{n \in \mathbb{N}} |S_n \phi|$. Then $M$ commutes with right translations $R_g$ and

$$H(g, T_h x) = M(F(g, T_h x)) = M(F(gh, x)) = M(R_h F(g, x))$$

$$= R_h M(F(g, x)) = R_h H(g, x) = H(gh, x).$$

Thus, all $H(g, x)$ are also equimeasurable functions of $x$.

For any compact neighborhood $V$ of $G$, define $F_V(g, x) = 1_V(g) F(g, x)$ and $H_V((g, x) = M(F_V(g, x))$. Since $M$ is sublinear and semilocal (since we are taking only a finite number of $n$'s), if $K$ is the compact set defining the semilocal property of $M$ then, for any other compact set $U$, we have

$$H = M(F) \leq M(F_{K^{-1}U} + (F - F_{K^{-1}U})) \leq M(F_{K^{-1}U}) + M((F - F_{K^{-1}U})).$$

But $F - F_{K^{-1}U}$ has support outside $K^{-1}U$ for every $x$. Hence, if $g \in U$ and every $x$,

$$M(F - F_{K^{-1}U})(g, x) = 0$$

because

$$g \in \text{supp}(M(F - F_{K^{-1}U})) \subset K \text{supp}(F - F_{K^{-1}U}) \subset K(K^{-1}U)^c.$$
which implies \( k^{-1}g \in (K^{-1}U)^c \) for some \( k \in K \). Therefore, \( g \) cannot be in \( U \).

Therefore,

\[
H(g, x) \leq H_{K^{-1}U}(g, x) \quad \text{for all } g \in U \text{ and any compact } U.
\]  

(i)

Assume \( M \) is of weak type \((1,1)\). For fixed \( \lambda \), define

\[
E = \{ x \in X : H(e, x) > \lambda \} \quad \text{and} \quad E' = \{(g, x) \in G \times X : H_{K^{-1}U}(g, x) > \lambda \}.
\]

Also, let \( E'_y = E' \cap (G \times \{y\}) \) and \( E'_g = E' \cap (\{g\} \times X) \). Then, by Fubini's Theorem,

\[
\int_{G \times X} 1_{E'} = \int_X |E'_x| dx = \int_G m(E'_g) dg \geq \int_U m(E'_g) dg,
\]

and by the equimeasurability of \( H(., x) \) and (i)

\[
|U|m(E) \leq \int_X |E'_x| dx.
\]

But since \( H \) is of weak type \((1,1)\),

\[
|E'_x| \leq \frac{C}{\lambda} \int_G |F_{K^{-1}U}(g, x)| dg,
\]

which together with the above and the equimeasurability of \( F(., x) \) gives

\[
|E| \leq \frac{C}{\lambda |U|} \int_X \int_{K^{-1}U} |F(g, x)| dg dx
= \frac{C}{\lambda |U|} \int_{K^{-1}U} \int_X |F(g, x)| dx dg
= \frac{C |K^{-1}U|}{\lambda |U|} \int_X |F(e, x)| dx.
\]

By letting \( U \) run through a uniform summing sequence, the proof is completed.
Corollary D.0.15 (Transfer Principle) Let \( \{F_n\} \) be a sequence of compact subsets of the amenable group \( G \). If there exists a constant \( C \) such that

\[
|\{g : \sup_{n \in \mathbb{N}} \frac{1}{|F_n|} \left| \int_{F_n} \phi(hg)dh \right| > \lambda \}| \leq C \frac{\|\phi\|_1}{\lambda}
\]

for all \( \phi \in L^1(G) \) then, with the same constant \( C \),

\[
m(\{x : \sup_{n \in \mathbb{N}} \frac{1}{|F_n|} \left| \int_{F_n} f(T_h x)dh \right| > \lambda \}) \leq C \frac{\|f\|_1}{\lambda}
\]

for all \( f \in L^1(X) \).

**Proof.** The operators \( S_n \phi(g) = |F_n|^{-1} \int_{F_n} \phi(hg)dh \) satisfy all the hypotheses of the theorem.

\[\Box\]

Corollary D.0.16 Let \( G \) be an amenable. Let \( \{F_n\} \) be a sequence of compact subsets of \( G \) and \( \{a_n\} \) a sequence of positive numbers. If there exists a constant \( C \) such that

\[
\sum_{n=1}^{\infty} |\{g : \frac{1}{|F_n|} \left| \int_{F_n} \phi(hg)dh \right| > a_n \}| \leq C \|\phi\|_1
\]

for all \( \phi \in L^1(G) \), then

\[
\sum_{n=1}^{\infty} m(\{x : \frac{1}{|F_n|} \left| \int_{F_n} f(T_h x)dh \right| > a_n \}) \leq C \|f\|_1
\]

for all \( f \in L^1(X) \) with the same constant.

**Proof.** Consider only a finite number of \( n \)'s, say \( n \leq N \). The corollary will follow by a passage to the limit.
The operators $S_n \phi(g) = |F_n|^{-1} \int_{F_n} \phi(hg)dh$ satisfy all the hypotheses of the theorem. Proceeding as in the above theorem, let $H_n = S_n(F)$ and let $K$ be a compact set that realizes the semilocality for all $S_n, n \leq N$. Using the same notation as before, define, for any arbitrary compact set $U$,

$$E_n = \{x \in X : H(e,x) > a_n\} \text{ and } E'_n = \{(g,x) \in G \times X : H_{n,K^{-1}U}(g,x) > a_n\}.$$ 

Also, let $E'_{n,y} = E'_n \cap (G \times \{y\})$ and $E'_{n,g} = E'_n \cap (\{g\} \times X)$. Following the proof of the theorem,

$$|U|m(E_n) \leq \int_X |E'_{n,x}|dx$$

for all $n$. By the hypothesis of the corollary, we can estimate

$$\sum_{n=1}^{\infty} |E'_{n,x}| \leq C \int_G |F_{K^{-1}U}(g,x)|dg.$$ 

Again, the equimeasurability of $F(.,x)$ gives,

$$\sum_{n=1}^{N} m(E_n) \leq \frac{C}{|U|} \int_X \int_G |F_{K^{-1}U}(g,x)|dg$$

$$= C\frac{|K^{-1}U|}{|U|} \|f\|_1.$$ 

By letting $U$ run over a uniform summing sequence, this finishes the proof because the right hand side does not depend on $N$.  

\[\blacksquare\]
BIBLIOGRAPHY


