POINTWISE ERGODIC THEOREMS FOR ARHIMETHIC SETS

by

J. BOURGAIN[1]

WITH AN APPENDIX ON RETURN-TIME SEQUENCES

jointly with H. Furstenberg[2], Y. Katznelson[3] and D. Ornstein[3]

Institut des Hautes Etudes Scientifiques
35, route de Chartres
91440 - BURES sur YVETTE (France)

Décembre 1988

IHES/M/88/66

---

(1) IHES, Bures sur Yvette, France
(2) Hebrew University of Jerusalem, Israel
(3) Stanford University, USA

POINTWISE ERGODIC THEOREMS FOR ARITHMETIC SETS

J. Bourgain[1]

WITH AN APPENDIX ON RETURN-TIME SEQUENCES

jointly with H. Furstenberg[2], Y. Katznelson[3] and D. Ornstein[3]

1. Introduction

This paper is a development of the earlier work [B_1], [B_2], [B_3] of the author on extending Birkhoff's ergodic theorem to certain subsets of the integers. It was proved in [B_1] that given a dynamical system, (DS, for short), $(\Omega, \mathcal{B}, \mu, T)$ and a polynomial $p(x)$ with integer coefficients, then the ergodic means

$$A_N f = \frac{1}{N} \sum_{1 \leq n \leq N} T^{p(n)} f \tag{1.1}$$

converge almost surely for $N \to \infty$, assuming f a function of class $L^2(\Omega, \mu)$. Here and in the sequel, one assumes μ a probability measure and T a measure preserving automorphism. The natural problem of developing the L^p-theory for $p < 2$ was studied in [B_2] and a partial result was obtained. We continue here this line of investigation.

The approach used in [B_1],[B_2] relies on a method which may be summarized as follows:

(a) Reduction of the general problem to statements about the shift S on $\mathbb{Z}$, which are of a "finite" and "quantitative" nature (in the sense of inequalities involving finitely many iterates of the transformation).

(b) Proof of certain maximal function inequalities, relative to the shift, by Fourier Analysis methods.

(c) Use of the "major arc" description of the relevant exponential sums, similar to that in the Hardy-Littlewood circle method.

---

(1) IHES, Bures/Yvette, France
(2) Hebrew University of Jerusalem, Israel
(3) Stanford University, USA

As I observed in [B_1], this approach should be considered more general than the solution to some isolated questions.

The purpose of this paper is two-fold. First, as far as the L^2-theory is concerned, we will develop appropriate harmonic analysis methods (maximal function estimates for certain sequences of multipliers), which will make the argument less dependent on special properties of the exponential sums (which were essentially exploited in [B_1],[B_2]). Using this additional ingredient, further examples will be obtained, for instance sets of the form

$$\Lambda = \{[p(n)]\ ;\ n = 1, 2, \dots\}$$

where $p(x)$ is any polynomial with real coefficients and $[x]$ stands for the integer part. Secondly, a method will be described to cover the full L^p-range, $p > 2$. In particular, it is shown that the averages $A_N f$ given by (1.1) converge almost surely for f a function of class $L^p(\Omega, \mu)$, $p > 1$. The problem for L^1-functions remains open at the present time. The shift reduction mentioned above permits giving a new and simple proof of Birkhoff's ergodic theorem (cf. [B_3]). Our proof of the pointwise and maximal ergodic theorem is related to [K-W], although different and providing more quantitative information. In particular, in order to illustrate ideas, it will be shown how to avoid the invariance of the limit. When dealing with subsets of $\mathbb{Z}$, this invariance is indeed not available in general and the pointwise ergodic theorem is not a formal consequence of the maximal ergodic theorem (except if the linear span of the eigenfunctions of T is dense). The shift reduction applies equally well for positive isometries. Already considering the sequence of squares $\Lambda = \{n^2\}$, the L^p-result for all $p > 1$ is new, and in particular the following corollary (for $p = 2$, see [B_1]).

Let f be an L^p-function on the circle $\pi = \mathbb{R}/\mathbb{Z}$ and $\alpha \in \mathbb{R}\backslash Q$ an irrational number. Then the averages

$$\frac{1}{N}\sum_{n=1}^{N} f(x + n^2\alpha) \tag{1.2}$$

converge to the mean $\int_0^1 f(x)dx$, for almost all x.

It is tempting, especially for $p = 2$, to approach such a problem by straight Fourier Analysis, considering the Fourier expansion of the function f (cf. [S]). However, to make this method succeed, stronger information on the Fourier-coefficients of f seems needed than just their square summability. The proof of the previous statement uses indeed harmonic analysis methods, but only after reduction to a dynamical system problem. Observe that in this case only the maximal inequality needs to be proven ($p > 1$)

$$\int_0^1 \left(\sup_N \left[\frac{1}{N} \sum_{n \leq N} f(x + n^2\alpha) \right]^p \right) dx \leq c \int_0^1 f(x)^p dx \tag{1.3}$$

for $f \geq 0$.

Next, we describe the organization of the paper and state the main results.

In the next section, an approach to Birkhoff's theorem is presented along the lines explained above and some less known features of this result are pointed out.

In section 3, we consider the variation spaces v_p, where $\|x\|_{v_p}$ is defined as

$$\sup_{s; j_1 < \cdots < j_s} \left(\sum |x_{j_{s-1}} - x_{j_s}|^p \right)^{1/p} \qquad x = (x_j)_{j=1,2,\ldots} \tag{1.4}$$

These spaces are well-adapted for a quantitative formulation of convergence properties. In this context, we recall a result due to Lépingle on bounded martingales, which is of importance later on in the paper.

Section 4 is devoted to the proof of a maximal inequality for certain sequences of Fourier multipliers. These Fourier multipliers appear naturally in the "major arc" description of exponential sums. The results of section 4 are purely L^2.

In section 5, we recall some basic and well-known facts on the behaviour of exponential sums of the form

$$\varphi_N(\overline{\alpha}) = \sum_{n=0}^{N} e^{2\pi i p(n,\overline{\alpha})} \tag{1.5}$$

where

$$p(x, \overline{\alpha}) = \alpha_d x^d + \alpha_{d-1} x^{d-1} + \cdots + \alpha_1 x \quad \text{and} \quad \overline{\alpha} = (\alpha_1, \ldots, \alpha_d) \in [0,1]^d \,. \tag{1.6}$$

The information on these sums needed for our purpose is essentially the same as for solving the Waring problem by the Hardy-Littlewood circle method.

Section 6 is a new presentation of the L^2-result on polynomial ergodic averages obtained in [B_1], based on the new ingredient obtained in section 4. In this proof, we no longer need the prior estimate of A. Weil for exponential sums with prime modulus.

Section 7 of this paper contains the corresponding (new) L^r-result for all $r > 1$. Thus the following theorem is proved.

Theorem 1. *Let $(\Omega, \mathcal{B}, \mu, T)$ by a dynamical system and $p(x)$ a polynomial with integer coefficients. Then there is the maximal inequality*

$$\big\| \sup_N |A_N f| \big\|_r \leq C \|f\|_r \tag{1.7}$$

where $A_N f$ is given by (1.1), i.e.,

$$A_N f = \frac{1}{N} \sum_{1 \leq n \leq N} T^{p(n)} f$$

and $f \in L^r(\Omega, \mu)$, $r > 1$.
The constant C in (1.7) depends only on $r > 1$ and on the polynomial $p(x)$. Moreover, the averages $A_N f$ converge almost surely for $N \to \infty$. If T is weakly mixing, the limit is given by $\int f \, d\mu$.

The previous result remains valid for positive isometries on $L^r(\Omega, \mu)$. Let us point out that the proof of Theorem 1, in the case of a general polynomial $p(x)$ with integer coefficients, is essentially identical to the special case $p(x) = x^2$. Essential use is made of duality and interpolation methods.

In section 8, the results of section 4 and section 5 are used to prove the following

Theorem 2. *Let $(\Omega, \mathcal{B}, \mu, T)$ be a dynamical system and $p(x)$ an arbitrary polynomial. Then the averages*

$$A_N f = \frac{1}{N} \sum_{1 \leq n \leq N} T^{[p(n)]} f \tag{1.8}$$

for f any bounded measurable function on Ω, converge almost surely. Here $[x]$ stands for the integer part of $x \in \mathbb{R}$.

It is possible to obtain L^r-results, $r > 1$, relative to the averages (1.8), at the price of additional technicalities, based on the method of proof for Theorem 1. This further development is not worked out in the paper and is left to the reader.

Section 9 contains various comments and remarks on almost sure convergence in general, related to [B5].

Summary

The paper has an Appendix on return time sequences, in joint work with H. Furstenberg, Y. Katznelson and D. Ornstein, simplifying an earlier exposition [B$_4$] (cf. also [B$_6$]).

2. Birkhoff's Theorem Revisited

Let $(\Omega, \mathcal{B}, \mu, T)$ be a dynamical system. In this section, we consider the usual ergodic averages $A_N f = \frac{1}{N} \sum_{1 \leq n \leq N} T^n f$, appearing in Birkhoff's ergodic theorem. We discuss their convergence properties, partly keeping in mind possible extensions to certain subsets of $\mathbb{Z}$.

A. Mean Convergence.

The sequence of complex polynomials $p_N(z) = \frac{1}{N} \sum_{n=1}^{N} z^n$ pointwise converges on the unit circle (to 0 except for $z = 1$). Consequently, by general spectral theory of unitary operators, $A_N f$ converges in $L^2(\mu)$ whenever $f \in L^2(\mu)$. The main point here is the existence of a spectral measure. The Herglotz-Bochner theorem assures indeed the existence of a positive Radon measure ν on the circle T, such that

$$\langle T^n f, f \rangle = \hat{\nu}(n) \equiv \int_0^1 e^{-2\pi i n \theta} \nu(d\theta) \tag{2.1}$$

implying that the map $L^2(\Pi, \nu) \longrightarrow L^2(\Omega, \mu)$ mapping the nth character $e^{2\pi i n \Theta}$ on $T^n f$ is an isometry. Thus the convergence of $A_N f$ in $L^2(\Omega, \mu)$ is equivalent to the convergence of $p_N(z)$ in $L^2(\Pi, \nu)$.

This is clearly an L^2-theory. In general, given a subset Λ of the positive integers, the pointwise convergence on the unit circle of the sequence of polynomials

$$p_N(z) = \frac{1}{|\Lambda \cap [1,N]|} \sum_{\substack{1 \le n \le N \\ n \in \Lambda}} z^n \tag{2.2}$$

is equivalent with a mean ergodic theorem for the set Λ. In the case of "arithmetic sets", this test is particularly useful since the convergence of $p_N(z)$ given by (4) is closely related to phenomena of uniform distribution. For instance, if Λ is the set of squares $\{n^2 \mid n = 1, 2, \dots\}$, we have

$$p_N(e^{2\pi i\alpha}) \longrightarrow 0 \quad \text{if } \alpha \text{ is irrational}$$

and

$$p_N(e^{2\pi i\alpha}) \longrightarrow S(q,a) \equiv \frac{1}{q} \sum_{r=0}^{q-1} e^{2\pi i \frac{a}{q} r^2} \quad \text{for} \quad \alpha = \frac{a}{q}$$

(the Gauss-sums).

It is not surprising that the (stronger) almost-sure convergence properties result from a finer analysis of these exponential sums and the class of L^2-functions appears as the natural function space in these problems. A sequence $\Lambda \subset \mathbb{Z}_+$ is "ergodic" provided $p_N(z) \longrightarrow 0$ for $z \in \mathbb{T} - \{1\}$. The property implies mean convergence of $A_N f$ to $\int_\Omega f \, d\mu$, assuming T ergodic (this is the case for $\Lambda = \mathbb{Z}_+$ but not if $\Lambda = \{n^2 \mid n = 1, 2, \dots\}$ for instance).

B. Weiss [W] observed that sequence Λ obtained by taking suitable unions of disjoint intervals are ergodic but may fail to satisfy the pointwise ergodic theorem, even with respect to bounded measurable functions.

B. Maximal Ergodic Theorems.

Let again

$$A_n f = \frac{1}{N} \sum_{n=1}^{N} T^n f$$

and define the "maximal function"

$$f^* = \sup_{N=1,2,\dots} |A_N f| \, .$$

There are the L^p-inequalities $(1 < p \le \infty)$

$$\|f^*\|_{L^p(\Omega,\mu)} \le C(p) \|f\|_{L^p(\Omega,\mu)} \tag{2.3}$$

and the weak-type inequality

$$\|f^*\|_{L^{1,\infty}(\Omega,\mu)} \leq C\|f\|_{L^1(\Omega,\mu)} \tag{2.4}$$

where $\|g\|_{L^{1,\infty}} = \sup_{\lambda>0} \lambda\mu\big[|g| > \lambda\big]$ and $C, C(p)$ are absolute constants.

Let us give a simple proof of (2.3),(2.4) by deriving them from the shift model $(\mathbb{Z}, S)$. In the case of the shift, the weak-type property (2.4) easily follows from geometric covering properties of integer-intervals, similarly as for the Hardy-Littlewood maximal function on the real line. Once (2.4) is obtained, the L^p-inequalities follow from the Marcinkiewicz interpolation theorem. Consider now the case of the general dynamical system (Ω, μ, T). Of course it suffices to prove inequalities (2.3), (2.4) (with fixed constants) for a "restricted" maximal function

$$\overline{f} = \sup_{1\leq N\leq \overline{N}} A_N f \qquad (f \geq 0) \tag{2.5}$$

where $\overline{N}$ is an arbitrarily chosen positive integer. Take an integer $J \gg \overline{N}$ and for fixed $x \in \Omega$, consider the orbit

$$x, Tx, T^2x, \ldots, T^Jx\ .$$

For the function f, define the function φ on $\mathbb{Z}$ as follows

$$\left.\begin{aligned} \varphi(j) &= f(T^jx) \quad \text{if} \quad 0 \leq j \leq J \\ &= 0 \qquad \text{otherwise}\ . \end{aligned}\right\} \tag{2.6}$$

Thus $A_N\varphi(j) = A_N f(T^jx)$ provided $0 \leq j < J - N$ and hence, with the definition (2.5)

$$\overline{\varphi}(j) = \overline{f}(T^jx) \quad \text{for} \quad 0 \leq j < J - \overline{N}\ . \tag{2.7}$$

The inequality $\|\overline{\varphi}\|_{\ell^p(\mathbb{Z})} \leq \|\varphi^*\|_{\ell^p(\mathbb{Z})} \leq C(p)\|\varphi\|_{\ell^p(\mathbb{Z})}$ then immediately implies by (2.6), (2.7)

$$\sum_{0\leq j<J-\overline{N}} \|\overline{f}(T^jx)\|^p \leq C(p)^p \sum_{0\leq j\leq J} |f(T^jx)|^p\ . \tag{2.8}$$

Integrating (2.8) in $x \in \Omega$ with respect to the measure μ yields

$$\sum_{0\leq j<J-\overline{N}} \|T^j\overline{f}\|_p^p \leq C(p)^p \sum_{0\leq j\leq J} \|T^jf\|_p^p$$

and since T is measure preserving, one gets

$$\|\overline{f}\|_p \leq C(p)\frac{J}{J-\overline{N}}\|f\|_p$$

hence

$$\|f^*\|_p \leq C(p)\|f\|_p .$$

One can deal similarly with the weak-type inequality (2.4). Assume $f \in L^1(\Omega, \mu)$, $\lambda > 0$ and let $\Omega_\lambda = [\overline{f} > \lambda]$, χ its indicator function. Given $x \in \Omega$, let φ be defined as above and let $|I|$ stand for the cardinality of a (finite) subset I of $\mathbb{Z}$. The shift inequality gives thus

$$\|\overline{\varphi}\|_{\ell^{1,\infty}(\mathbb{Z})} \leq C\|\varphi\|_{\ell^1(\mathbb{Z})}$$

and, by (2.7)

$$\lambda\big|\{0 \leq j < J - \overline{N} \mid \overline{f}(T^j x) > \lambda\}\big| \leq C \sum_{0 \leq j \leq J} f(T^j x)$$

hence

$$\sum_{0 \leq j < J - \overline{N}} \chi(T^j x) \leq \frac{C}{\lambda} \sum_{0 \leq j \leq J} f(T^j x) . \tag{2.9}$$

Again integrating

$$\lambda\mu(\Omega_\lambda) \leq C \frac{J}{J - \overline{N}} \|f\|_1$$

from where (2.4) easily follows.

Presently, the covering argument leading to weak-type inequalities seems not available when dealing with particular subsets of $\mathbb{Z}$, such as the squares or the primes. In these cases, we were unable so far to develop an L^1-theory. The L^2 and L^p-inequalities ($p > 1$) are obtained by making essential use of Fourier-transform methods. This is a similar approach as in differentiation problems in real analysis involving lower dimensional manifolds.

C. Almost-sure Convergence.

By the maximal inequality and a standard truncation argument, the almost sure convergence of $A_N f$ for f in $L^1(\Omega, \mu)$ reduces to bounded functions. Denote F the L^2-limit of $(A_N f)$ and, for given $\varepsilon > 0$, let N_ε satisfy

$$\|F - A_{N_\varepsilon} f\|_2 < \varepsilon .$$

By the invariance of the limit (since the ergodic means relate to the full set of positive integers) and the maximal inequality

$$\big\|\sup_N |F - A_N(A_{N_\varepsilon} f)\big\|_2 < C\varepsilon . \tag{2.10}$$

Since

$$|A_N(A_{N_\epsilon}f) - A_Nf| \leq 2\frac{N_\epsilon}{N}\|f\|_\infty$$

it follows from (2.10)

$$\left\|\overline{\lim}|F - A_Nf|\right\|_2 < C\epsilon \ , \text{ hence } \overline{\lim_N}|F - A_Nf| = 0 \text{ almost sure} .$$

This discussion completes the proof of Birkhoff's theorem. It is clear that the preceding argument does not apply when dealing with the more general averages

$$A_Nf = \frac{1}{|\Lambda \cap [1,N]|}\sum_{n\in\Lambda, n\leq N} T^n f \tag{2.11}$$

corresponding to a subset Λ of $\mathbb{Z}_+$.

If the eigenfunctions of T generate a dense subspace of L^2, the almost sure convergence of A_Nf for f of class L^p, $p \leq 2$, is implied by the pointwise convergence of the sequence $p_N(z)$, $|z| = 1$ given by (2.2) and the maximal inequality

$$\|f^*\|_p \leq C\|f\|_p \ ; \quad f^* = \sup|A_Nf| \ .$$

This is the case for instance for the model $(\Omega, T) = (\mathbb{T}, R_a)$, $R_a x = x + a$.

In the remainder of this section, an alternative method is explained for the purpose of proving the theorems stated in the introduction.

Take f in $L^\infty(\Omega,\mu)$, $|f| \leq 1$. For $\epsilon > 0$, consider the subset

$$Z_\epsilon = \left\{ \left[(1+\epsilon)^n\right] \mid n = 1,2,\dots \right\} \tag{2.12}$$

of $\mathbb{Z}_+$. Clearly, for each $N \in \mathbb{Z}_+$, there is $N' \in Z_\epsilon$ such that

$$|A_Nf - A_{N'}f| \leq 2\epsilon \ .$$

Thus to prove the almost sure convergence of (A_Nf), it suffices to show that there is no $\epsilon > 0$ and no sequence of positive integers N_j, $N_{j+1} > 2N_j$, such that

$$\|\mathcal{M}_jf\|_2 > \epsilon \quad \text{where} \quad \mathcal{M}_jf = \sup_{\substack{N_j \leq N \leq N_{j+1} \\ N \in Z_\epsilon}} |A_Nf - A_{N_j}f| \ . \tag{2.13}$$

In fact, a more quantitative statement is shown

$$\sum_{1\leq j\leq J} \|\mathcal{M}_j f\|_2 \leq o(J)\|f\|_2 \tag{2.14}$$

for J large (depending on ε appearing in the definition of $\mathcal{M}_j$). Since (2.14) only involves finitely many iterates of T, the general case reduces again to the shift (Z, S). For the sets $\{p(n) \mid n = 1,2,\dots\}$ (resp. $\{[p(n)] \ ; \ n = 1,2,\dots\}$) considered in Theorem 1 (resp. Theorem 2), the inequality (2.14) follows easily from the proof of the L^2-maximal inequality. In the context of Theorem 1, this argument was carried out in [B$_1$]. The method will be repeated in section 6 of this paper, for the sake of completeness.

3. Variation Spaces and Variational Inequalities

We start by recalling the definition of the variation norm v_s $(1 \leq s \leq \infty)$ for scalar sequences $\overline{x} = (x_n)_{n=1,2,\dots}$

$$\|\overline{x}\|_{v_s} = \sup\left\{\left(\sum_{j=1}^{J} |x_{n_j} - x_{n_{j+1}}|^s\right)^{1/s} \mid J = 1,2,\dots \quad \text{and} \quad n_1 < n_2 < \cdots < n_J\right\}. \tag{3.1}$$

The sequence space v_s consists then of those sequence $\overline{x}$ for which $\|\overline{x}\|_{v_s} < \infty$. We will also use the notation $\| \quad \|_{v_s}$ for continuously indexed systems $\overline{x} = (x_t)_{t>0}$, where now

$$\|\overline{x}\|_{v_s} = \sup\left\{\left(\sum_{j=1}^{J} |x_{t_j} - x_{t_{j+1}}|^s\right)^{1/s} \mid J = 1,2,\dots \quad \text{and} \quad t_1 < t_2 < \cdots < t_J\right\}. \tag{3.2}$$

These spaces v_s are frequently used in probability theory in questions about convergence. In this context, some known inequalities about martingales are needed for our purpose. More precisely, we will use the following result due to Lépingle [Lé] (cf. also [P-X]).

Lemma 3.3. *Let $\mathbb{E}_n$ $(n = 1,2,\dots)$ be the sequence of expectation operators with respect to an increasing sequence of σ-algebras on a probability space and $f_n = \mathbb{E}_n f$ an associated scalar martingale. Then, for $s > 2$, we have the inequality*

$$\|\{f_n\}\|_{L^2_{v_s}} \leq c(s-2)^{-1}\|f\|_{L^2} \tag{3.4}$$

where $L^2_{v_s}$ refers to the v_s-valued L^2-space.

This result may be seen as the quantitative form the martingale convergence theorem. Inequality (3.4) fails for $s = 2$ (this is a well-known feature of the Brownian martingale, related to the law of the iterated logarithm). In fact, the dependence in s stated in (3.4) will be of relevance later on and we include a fast proof here.

Proof of (3.4). For $\lambda > 0$, denote $N_\lambda(\omega)$ the number of λ-jumps in the sequence $\{f_n(\omega)\}$, where f_n is defined as above. One has the following inequality for $1 < r < \infty$

$$\|\lambda(N_\lambda)^{1/2}\|_r \leq c_r\|f\|_r\,, \qquad \forall\lambda > 0\,. \tag{3.5}$$

(3.5) is a form of Doob's oscillation lemma for martingales (see [Nev]) and is obtained by methods of stopping times and square functions. We use interpolation to derive (3.4) from (3.5). First we prove ($L^{p,1}$ denoting the Lorentz space)

$$\|\{f_n\}\|_{L^p_{v_s}} \leq c(s-2)^{-1/p}\|f\|_{L^{p,1}} \quad \text{for} \quad \frac{3}{2} \leq p \leq s < \frac{5}{2}\,, \qquad s > 2\,. \tag{3.6}$$

Let thus $f = \chi_A$, $A \subset \Omega$ a measurable set of measure $\mu(A) = \epsilon$, hence $\|f\|_{p,1} = \epsilon^{1/p}$. Estimate pointwise, for N_λ defined as above from the function f

$$\|\{f_n(\omega)\|_{v_s} \leq \left[\sum_{k=0}^\infty 2^{-ks}N_{2^{-k}}(\omega)\right]^{1/s}\,. \tag{3.7}$$

Hence, since $p \leq s$

$$\|\{f_n\}\|_{L^p_{v_s}} \leq 2\left[\sum_{k=0}^\infty 2^{-kp}\int_\Omega (N_{2^{-k}})^{p/s}d\omega\right]^{1/p} \leq c\left[\sum 2^{-kp\left(1-\frac{2}{s}\right)}\|f\|_r^r\right]^{1/p} \tag{3.8}$$

applying (3.5) with $r = \frac{2p}{s}$, thus $\frac{6}{5} \leq r \leq 2$ from hypothesis on p, s, and $\lambda = 2^{-k}$. Since $\|f\|_r^r = \epsilon$, (3.6) is immediate from (3.8). Writing L^2 as interpolation space between $L^{s,1}$ and $L^{3/2}$, (3.6) is easily seen to imply (3.4).

We will now derive from Lemma 3.3 a real analysis version of (3.4). For a function f on $\mathbb{R}$, denote $f_t(x) = \frac{1}{t}f\left(\frac{x}{t}\right)$. Denote also

$$\mathcal{F}f(\lambda) \equiv \hat{f}(\lambda) = \int_{-\infty}^\infty f(x)e^{-2\pi i\lambda x}dx \tag{3.9}$$

the Fourier transform of f. Thus

$$\widehat{f}_t(\lambda) = \widehat{f}(t\lambda) . \tag{3.10}$$

Lemma 3.11. *Let $\chi = \chi_{[0,1]}$ be the indicator function of the interval $[0,1]$. Then, for $f \in L^2(\mathbb{R})$ and $s > 2$*

$$\left\| \{ f * \chi_t \mid t > 0 \} \right\|_{L^2_{v_s}(\mathbb{R})} \leq c(s-2)^{-1} \|f\|_2 \tag{3.12}$$

where v_s stands here for $v_s(\mathbb{R}_+)$, with the norm given by (3.2).

As usual, $f * g$ denotes the convolution of f and g.

Denote $(P_t)_{t>0}$ the Poisson semi-group on $\mathbb{R}$. Thus if $P_t f = f * P_t$, one has $\widehat{P}_t(\lambda) = e^{-t|\lambda|}$. Considering the Brownian martingale associated to the harmonic function $u(x,t) = (f * P_t)(x)$ on the upper half-plane or, alternatively, invoking Rota's dilation theorem, inequality (3.4) relative to martingales implies

$$\left\| \{ P_t f \mid t > 0 \} \right\|_{L^2_{v_s}} \leq c(s-2)^{-1} \|f\|_2 . \tag{3.13}$$

Proof of Lemma 3.11.

By (3.13), (3.12) will be a consequence of following inequality

$$\left\| \{ f * K_t \mid t > 0 \} \right\|_{L^2_{v_2}} \leq c \|f\|_2 \tag{3.14}$$

where K stands for the function $\chi - P_1$, hence satisfying the Fourier transform estimates

$$|\lambda| \cdot \left| (\widehat{K})'(\lambda) \right| < c \quad \text{and} \quad \left| \widehat{K}(\lambda) \right| \leq c \min\left(|\lambda|, |\lambda|^{-1} \right) . \tag{3.15}$$

There is clearly the pointwise estimate

$$\left\| \{ f * K_t \mid t > 0 \} \right\|_{v_2} \leq \Big(\sum_{k \in \mathbb{Z}} |f * K_{2^k}|^2 \Big)^{1/2} + \Big(\sum_{k \in \mathbb{Z}} \left\| \{ f * K_t \mid 2^k \leq t \leq 2^{k+1} \} \right\|_{v_2}^2 \Big)^{1/2} . \tag{3.16}$$

By Parseval's identity, the L^2-norm of the first term in (3.16) is bounded by

$$\Big[\sum_{k \in \mathbb{Z}} \int_{-\infty}^{\infty} |\widehat{f}(\lambda)|^2 |\widehat{K}(2^k \lambda)|^2 d\lambda \Big]^{1/2} \leq c \cdot \Big[\int |\widehat{f}(\lambda)|^2 d\lambda \Big] = c \|f\|_2 \tag{3.17}$$

invoking (3.15).

Next estimate the contribution of the second term

$$\left\{ \sum_{k\in \mathbb{Z}} \|\{f * K_t \mid 2^k \le t \le 2^{k+1}\}\|^2_{L^2_{v_2}} \right\}^{1/2} . \tag{3.18}$$

Let $0 < \eta < 1$ be a function supported by $[\frac{1}{2}, 2] \cup [-2, -\frac{1}{2}]$, $|\eta'| < C$, such that

$$\sum_{\alpha \in \mathbb{Z}} \eta(2^\alpha \lambda) = 1 .$$

Defining K_α by $\widehat{K_\alpha}(\lambda) = \widehat{K}(\lambda)\eta(2^\alpha\lambda)$, one has that $K = \sum\limits_\alpha K_\alpha$ and (3.18) may be estimated by triangle inequality as

$$\sum_{\alpha\in \mathbb{Z}} \left\{ \sum_{k\in \mathbb{Z}} \|\{f * (K_\alpha)_t \mid 2^k \le t \le 2^{k+1}\}\|^2_{L^2_{v_2}} \right\}^{1/2} . \tag{3.19}$$

From (3.15)

$$|\lambda| \big|(\widehat{K}_\alpha)'(\lambda)\big| < c \quad \text{and} \quad \big|\widehat{K}_\alpha(\lambda)\big| < c2^{-|\alpha|} . \tag{3.20}$$

Fix $\alpha \in \mathbb{Z}$. For $k \in \mathbb{Z}$, consider a net $2^k = u_1 < u_2 < \cdots < u_N = 2^{k+1}$ of $N = N_\alpha$ equidistributed points. The number N_α will be specified later. Estimate

$$\big\|\{f * (K_\alpha)_t \mid 2^k \le t \le 2^{k+1}\}\big\|_{v_2} \le \left[\sum_{\ell=1}^N \left|f * (K_\alpha)_{u_\ell}\right|^2\right]^{1/2} + \tag{3.21}$$

$$+ \left\{ \sum_{\ell=1}^N \left[\int_{u_\ell}^{u_{\ell+1}} |\partial_t[f * (K_\alpha)_t]| dt \right]^2 \right\}^{1/2} \tag{3.22}$$

majorizing $v_2([u_\ell, u_{\ell+1}])$ by $v_1([u_\ell, u_{\ell+1}])$.

Again by Parseval's identity, the L^2-norm of (3.21) is bounded by

$$\left[\sum_{\ell=1}^N \int_{-\infty}^{\infty} |\widehat{f}(\lambda)|^2 |\widehat{K}_\alpha(u_\ell\lambda)|^2 d\lambda \right]^{1/2} \le C N_\alpha^{1/2} 2^{-|\alpha|} \left[\int_{|\lambda| \sim 2^{-\alpha-k}} |\widehat{f}(\lambda)|^2 d\lambda \right]^{1/2} \tag{3.23}$$

by the definition of K_α and (3.20). Here $|\lambda| \sim \rho$ stands for $\frac{1}{4}\rho < |\lambda| < 4\rho$. Similarly, the L^2-norm of (3.22) is bounded by

$$\left[\sum_{\ell=1}^N (u_{\ell+1} - u_\ell) \int_{u_\ell}^{u_{\ell+1}} \left[\int_{-\infty}^{\infty} |\widehat{f}(\lambda)|^2 |\lambda|^2 |(\widehat{K_\alpha})'(t\lambda)|^2 d\lambda \right] dt \right]^{1/2} \le$$

$$C\left[\sum_{\ell=1}^{N}\left(\frac{2^k}{N}\right)^2 4^{-k}\Big(\int_{|\lambda|\sim 2^{-k-\alpha}}|\hat f(\lambda)|^2 d\lambda\Big)\right]^{1/2} = CN_\alpha^{-1/2}\Big[\int_{|\lambda|\sim 2^{-k-\alpha}}|\hat f(\lambda)|^2 d\lambda\Big]^{1/2}. \quad (3.24)$$

Substitution of estimates (3.23), (3.24) in (3.19) give finally the bound

$$\left[\sum_{\alpha,k\in\mathbb{Z}}(N_\alpha 4^{-|\alpha|}+N_\alpha^{-1})\Big(\int_{|\lambda|\sim 2^{-k-\alpha}}|\hat f(\lambda)|^2 d\lambda\Big)\right]^{1/2} \leq C\|\hat f\|_2 = C\|f\|_2$$

letting $N_\alpha = 2^{|\alpha|}$.

Summation of (3.17),(3.18) yields (3.14), which proves Lemma 3.11.

Let us point out one corollary of Lemma 3.11 to the convergence of the averages

$$A_N f = \frac{1}{N}\sum_{1\leq n\leq N} T^n f$$

in Birkhoff's theorem.

Corollary 3.25. *Let $(\Omega,\mathcal{B},\mu,T)$ be a DS and $f\in L^2(\mu)$. Then for $s>2$*

$$\left\|\left\{\frac{1}{N}\sum_{n\leq N}T^n f \mid N=1,2,\dots\right\}\right\|_{L^2_{v_s}} \leq c(s)\|f\|_2 . \quad (3.26)$$

The last result does not seem to appear in the literature. It refines the results discussed in the previous section (related to almost sure convergence). The proof of (3.26) reduces to the particular case of the shift model $(\mathbb{Z}, S)$, following the procedure described in section 2 of this paper. In the context of the shift, (3.26) is just a discrete version of (3.12).

Writing

$$\varphi = -\int_0^\infty \chi_t\cdot\varphi'(t)t\,dt \qquad \chi_t = \frac{1}{t}\chi_{[0,t]} \quad (3.27)$$

for a smooth function φ on $[0,\infty[$, vanishing at ∞, it follows from (3.12) and convexity

Lemma 3.28. *Let φ be a differentiable function on $\mathbb{R}$, vanishing at ∞. Then, for $s > 2$*

$$\left\|\{f * \varphi_t \mid t > 0\}\right\|_{L^2_{v_s}} \leq c \cdot (s-2)^{-1}\left(\int_{-\infty}^{\infty} |\varphi'(x)||x|dx\right)\|f\|_2 . \tag{3.29}$$

We conclude this section with a corollary of (3.28) which will be of importance in the proof of certain Fourier-multiplier maximal inequalities considered in the next section.

Let H be a Hilbert space. If A is a subset of H, denote $M_\lambda(A)$ the λ-entropy number of A, $\lambda > 0$. By λ-entropy number, we mean the minimal number ($\leq \infty$) balls (with respect to the H-norm) of radius λ, needed to cover A. We let $M_\lambda = 0$ if diam $A < \lambda$. The following result relates to H-valued functions on $\mathbb{R}$.

Lemma 3.30. *Let φ be as in (3.28), $s > 2$ and H a Hilbert space. Then for $f \in L^2_H(\mathbb{R})$*

$$\|\sup_{\lambda>0}(\lambda M_\lambda^{1/s})\|_2 \leq c_\varphi(s-2)^{-1}\|f\|_2 \tag{3.31}$$

where one defines pointwise $M_\lambda(x) = M_\lambda\big(\{(f * \varphi_t(x) \mid t > 0\}\big)$ and $C_\varphi = \int |\varphi'(x)||x|dx$.

Proof: Observe first the pointwise inequality

$$\lambda M_\lambda(x)^{1/s} \leq \left\{\sum_j \|(f*\varphi_{t_j})(x) - (f*\varphi_{t_{j-1}})(x)\|_H^s\right\}^{1/s} \leq \left\|\{(f*\varphi_t)(x)\}\right\|_{v_s} \tag{3.32}$$

where $\bar{t} = (t_j)$ is defined by putting

$$t_j = \min\left\{t > t_{j-1} \mid \|(f*\varphi_t)(x) - (f*\varphi_{t_{j-1}})(x)\|_H > \lambda\right\} .$$

(Since we are concerned with a priori inequalities, we may take the sequence $\bar{t} = (t_j)$ of bounded length.)

Writing $f = \sum f_\alpha e_\alpha$, $f_\alpha = \langle f, e_\alpha\rangle$, where $\{e_\alpha\}$ is an orthonormal basis for H, it follows from (3.32), (3.29) and convexity ($s > 2$)

$$\|\sup_{\lambda>0}(\lambda M_\lambda^{1/s})\|_2 \leq \left[\sum_\alpha \|\{f_\alpha * \varphi_t\}\|^2_{L^2_{v_s}}\right]^{1/2} \leq c_\varphi(s-2)^{-1}\Big(\sum_\alpha \|f_\alpha\|_2^2\Big)^{1/2} .$$

This proves (3.31).

There is the following corollary.

Lemma 3.33. *Let φ be as in (3.28) and H a Hilbert space. Then, with the notations of (3.30), for $f \in L^2_H(\mathbb{R})$ and $K > 0$*

$$\left\| \int_0^\infty \min\left(K, M_\lambda(x)\right)^{1/2} d\lambda \right\|_2 \leq c'_\varphi (\log K)^2 \|f\|_2 \,. \tag{3.34}$$

Proof: Let, with the notations of the proof of Lemma 3.30,

$$f^*_\alpha = \sup_{t>0} |f_\alpha * \varphi_t| \quad ; \quad f_\alpha = \langle f, e_\alpha \rangle$$

for which

$$\|f^*_\alpha\|_2 \leq c_\varphi \|f_\alpha\|_2 \tag{3.35}$$

by the Hardy-Littlewood maximal inequality. Define

$$F = \left[\sum (f^*_\alpha)^2\right]^{1/2} \tag{3.36}$$

and write, letting $s > 2$

$$\int_0^\infty \min\left(K, M_\lambda(x)\right)^{1/2} d\lambda \leq F(x) + \int_{K^{-1/2}F(x)}^{F(x)} K^{\frac{1}{2}-\frac{1}{s}} M_\lambda(x)^{1/s} d\lambda \leq$$

$$\leq F(x) + K^{\frac{1}{2}-\frac{1}{s}} (\log K) \sup_{\lambda>0} \lambda \cdot M_\lambda(x)^{1/s} \,.$$

(3.24) then follows from (3.35), (3.36) and (3.31), letting $\frac{1}{2} - \frac{1}{s} = (\log K)^{-1}$.

4. Maximal Inequalities for Certain Sequences of Fourier Multipliers

Proving the L^2-maximal inequality in Theorems 1 and 2 in the context of the shift $(\mathbb{Z}, S)$ by harmonic analysis methods leads to Fourier multipliers given by exponential sums (which properties will be recalled in the next section). In this section a rather general estimate is obtained, specially motivated by the major arc description of these exponential sums.

The dual group of $\mathbb{Z}$ is the circle group $\Pi = \mathbb{R}/\mathbb{Z}$, which will be identified with $[0,1]$ (identifying 0 and 1).

The main result of this section is contained in

Lemma 4.1. *Assume $\lambda_1 < \cdots < \lambda_K \in \Pi$ and define for $j \in \mathbb{Z}_+$ the neighborhoods*

$$R_j = \left\{ \lambda \in \Pi \mid \min_{1 \leq k \leq K} |\lambda - \lambda_k| \leq 2^{-j} \right\}. \tag{4.2}$$

Then

$$\left\| \sup_j \left| \int_{R_j} \widehat{f}(\lambda) e^{2\pi i \lambda x} d\lambda \right| \right\|_{\ell^2(\mathbb{Z})} \leq C(\log K)^2 \|f\|_{\ell^2(\mathbb{Z})} \tag{4.3}$$

for functions f on $\mathbb{Z}$.

Remark. It is an interesting question whether there needs to be a dependence on the number K of base points in (4.3). The logarithmic dependence will suffice largely for our purpose.

In order to simplify notations, we denote by $\mathcal{F}$ (resp. $\mathcal{F}^{-1}$) the Fourier-transform (resp. inverse Fourier-transform) for functions on either $\mathbb{R}$ or $\mathbb{Z}$.

For the sake of completeness , we include the following known argument to derive the corresponding inequality for $\mathbb{Z}$ from the $\mathbb{R}$-case. It is indeed often more appealing to prove the result on $\mathbb{R}$, because of the presence of the dilation structure.

Lemma 4.4. *Let Φ be a set of multipliers on $[0,1]$ satisfying*

$$\| \sup_{\varphi \in \Phi} |\mathcal{F}^{-1}[\varphi \mathcal{F} f]| \|_{L^2(\mathbb{R})} \leq B \|f\|_{L^2(\mathbb{R})} . \tag{4.5}$$

Then

$$\| \sup_{\varphi \in \Phi} |\mathcal{F}^{-1}[\varphi \mathcal{F} f]| \|_{\ell^2(\mathbb{Z})} \leq CB \|f\|_{\ell^2(\mathbb{Z})} \tag{4.6}$$

where C is an absolute constant.

Proof: Denote B_1 the best constant satisfying (4.6). Writing for $x \in \mathbb{Z}$, $u \in [0, \rho]$ ($\rho < 1$ to be specified later)

$$\mathcal{F}^{-1}[\varphi \mathcal{F} f](x) = \mathcal{F}^{-1}[\varphi \mathcal{F} f](x+u) + \mathcal{F}^{-1}\left[(1 - e^{2\pi i \lambda u}) \varphi \mathcal{F} f\right](x)$$

and averaging in u gives

$$\| \sup_{\varphi} |\mathcal{F}^{-1}[\varphi \mathcal{F} f]| \|_{\ell^2(\mathbb{Z})} \leq \rho^{-1/2} \| \sup_{\varphi} |\mathcal{F}^{-1}[\varphi \mathcal{F} f]| \|_{L^2(\mathbb{R})} \tag{4.7}$$

$$+ \sup_{0<u<\rho} \| \sup_{\varphi} |\mathcal{F}^{-1}[(1 - e^{2\pi i \lambda u}) \varphi \mathcal{F} f] | \|_{\ell^2(\mathbb{Z})} . \tag{4.8}$$

By (4.5), (4.7) is clearly bounded by

$$\rho^{-1/2} B \|\mathcal{F} f\|_{L^2[0,1]} = \rho^{-1/2} B \|f\|_{\ell^2(\mathbb{Z})} \, . \tag{4.9}$$

By definition of B_1, (4.8) is bounded by

$$\begin{aligned} B_1 \left\| f * \mathcal{F}^{-1}[1 - e^{2\pi i \lambda u}] \right\|_{\ell^2(\mathbb{Z})} &= B_1 \left\| \mathcal{F} f \cdot [1 - e^{2\pi i \lambda u}] \right\|_{L^2[0,1]} \\ &\leq C\rho B_1 \|\mathcal{F} f\|_{L^2[0,1]} \\ &= C\rho B_1 \|f\|_{\ell^2(\mathbb{Z})} \, . \end{aligned} \tag{4.10}$$

Hence, from (4.9),(4.10), $B_1 \leq \rho^{-1/2} B + C\rho B_1$, thus $B_1 \leq C' B$ by choosing ρ small enough.

By Lemma 4.4, Lemma 4.1 may be restated as

Lemma 4.11. *Let $\lambda_1, \ldots, \lambda_K \in \mathbb{R}$ and R_j stand for the 2^{-j}-neighborhood of the set $\Lambda = \{\lambda_1, \ldots, \lambda_K\}$, for $j \in \mathbb{Z}$. Then*

$$\left\| \sup_j |\mathcal{F}^{-1}[\chi_{R_j} \mathcal{F} f]| \right\|_2 \leq C (\log K)^2 \|f\|_2 \, . \tag{4.12}$$

The proof is mainly based on Lemma 3.33 of the previous section and will be presented in several steps.

Lemma 4.13. *Let $\lambda_1, \ldots, \lambda_K \in \mathbb{R}$ satisfy $|\lambda_k - \lambda_{k'}| > \tau > 0$ for $k \neq k'$. Let $0 \leq \varphi \leq 1$ be a smooth function such that $\operatorname{supp} \widehat{\varphi} \subset [-1, 1]$. Then*

$$\left\| \sup_{t > \tau^{-1}} \left| \sum_{k=1}^{K} e^{2\pi i \lambda_k x} (f_k * \varphi_t) \right| \right\|_2 \leq C (\log K)^2 \left(\sum_{k=1}^{K} \|f_k\|_2^2 \right)^{1/2} . \tag{4.14}$$

Proof: Observe first that

$$\left\| \sum_{k \leq K} a_k e^{2\pi i \lambda_k u} \right\|_{L^2[0, \tau^{-1}]} \leq C \tau^{-1/2} \left(\sum_{k=1}^{k} |a_k|^2 \right)^{1/2} \tag{4.15}$$

for all scalar sequences $\overline{a} = (a_k)_{1 \leq k \leq K}$. This is an easy consequence of the separation hypothesis of the λ_k's and we leave the verification to the reader.

Since $\operatorname{supp} \widehat{\varphi_t} \subset [-\tau, \tau]$ for $t > \tau^{-1}$, there is no restriction in assuming

$$\operatorname{supp} \widehat{f_k} \subset [-\tau, \tau] \quad \text{for} \quad 1 \leq k \leq K \, . \tag{4.16}$$

For $u \in \mathbb{R}$, denote σ_u the translation operator, thus $\sigma_u f(x) = f(x+u)$. It follows them from (4.16) and Parseval's identity that

$$\|f_k - \sigma_u f_k\|_2 < \frac{1}{2}\|f_k\|_2 \quad \text{for} \quad |u| < \frac{1}{100}\tau^{-1} \, . \tag{4.17}$$

Denoting B the best constant fulfilling (4.14) ($CK^{1/2}$ will certainly do) one gets from (4.17) for $0 \leq u \leq \frac{1}{100}\tau^{-1}$

$$\begin{aligned} &\left\| \sup_{t > \tau^{-1}} \left| \sum_{k=1}^{K} e^{2\pi i \lambda_k x} (f_k * \varphi_t) \right| \right\|_2 \leq \\ &\left\| \sup_{t > \tau^{-1}} \left| \sum_{k=1}^{K} e^{2\pi i \lambda_k x} \sigma_u (f_k * \varphi_t) \right| \right\|_2 \\ &\qquad + \tfrac{1}{2} B \Big(\sum \|f_k\|_2^2 \Big)^{1/2} \, . \end{aligned} \tag{4.18}$$

Integrating (4.18) in u on $\left[0, \frac{\tau^{-1}}{100}\right]$ permits replacing (4.18) by

$$C \left\| \tau^{1/2} \left\| \sup_{t>0} \left| \sum_{k=1}^{K} e^{-2\pi i \lambda_k u} e^{2\pi i \lambda_k x} (f_k * \varphi_t)(x) \right| \right\|_{L^2([0,\tau^{-1}], du)} \right\|_{L^2(dx)} \, . \tag{4.19}$$

By the preceding, it will suffice to bound (4.19) by $C(\log K)^2 \cdot \left(\sum \|f_k\|_2^2\right)^{1/2}$ in order to prove Lemma 4.13.

Fixing $x \in \mathbb{R}$, consider the set

$$A = A_x = \left\{ \left((f_1 * \varphi_t)(x), \ldots, (f_K * \varphi_t)(x) \right) \mid t > 0 \right\} \tag{4.20}$$

as subset of the K-dimensional Hilbert space ℓ_K^2. For $\lambda > 0$, denote again by $M_\lambda = M_\lambda(x)$ the entropy numbers of A. There exists a sequence B_s ($s \in \mathbb{Z}$) of finite subsets of the difference set $A - A$ such that

$$|\overline{b}| \leq 2.2^s \quad \text{for} \quad \overline{b} \in B_s \tag{4.21}$$

$$\# B_s \leq M_{(2^s)} \tag{4.22}$$

and each element $\overline{a} \in A$ has a representation

$$\overline{a} = \sum_{s \in \mathbb{Z}} \overline{b}_s \quad \text{with} \quad \overline{b}_s \in B_s \tag{4.23}$$

(# stands for "cardinality" and $|\bar{a}|$ refers to $\left(\sum_{k=1}^{K}|a_k|^2\right)^{1/2}$). In writing (4.23), we make the implicit assumption that $A = A_x$ is bounded, which is clearly no restriction.

Estimate

$$\sup_{t>0}\left|\sum_{k=1}^{K} e^{-2\pi i\lambda_k u} e^{2\pi i\lambda_k x}(f_k * \varphi_t)(x)\right| \le \sum_{s\in\mathbb{Z}} \max_{\bar{b}\in B_s}\left|\sum_{k=1}^{K} e^{-2\pi i\lambda_k u} e^{2\pi i\lambda_k x} b_k\right|$$

and the $L^2([0,\tau^{-1}], du)$-norm by

$$\sum_{s\in\mathbb{Z}}\left\|\max_{\bar{b}\in B_s}\left|\sum_{k=1}^{K} e^{-2\pi i\lambda_k u} e^{2\pi i\lambda_k x} b_k\right|\right\|_{L^2([0,\tau^{-1}])} . \tag{4.24}$$

For given s, consider the following bounds

$$\max_{\bar{b}\in B_s}|\cdots| \le \min\left\{2^{s+1}K^{1/2}\ ,\ \left[\sum_{\bar{b}\in B_s}\left|\sum_{k=1}^{K} e^{-2\pi i\lambda_k u} e^{2\pi i\lambda_k x} b_k\right|^2\right]^{1/2}\right\} .$$

Hence, invoking (4.15), (4.24) is bounded by

$$\sum_{s\in\mathbb{Z}} \min\left\{\tau^{-1/2}2^{s+1}K^{1/2}\ ,\ C\tau^{-1/2}2^{s+1}(\#B_s)^{1/2}\right\} \sim C\tau^{-1/2}\int_0^\infty \min\left(K, M_\lambda(x)\right)^{1/2} d\lambda \tag{4.25}$$

using (4.21), (4.22).

Taking the $L^2(dx)$-norm of (4.25), the required bound on (4.19) is obtained from (3.34). This proves (4.14).

Lemma 4.26. *Assume $\lambda_1,\ldots,\lambda_K \in \mathbb{R}$ satisfying $|\lambda_k - \lambda_{k'}| > 2^{-s}$ for $k \ne k'$. Then with previous notations*

$$\left\|\sup_{j\ge s}\left|\mathcal{F}^{-1}[\chi_{R_j}\mathcal{F}f]\right|\right\|_2 \le C(\log K)^2\|f\|_2 . \tag{4.27}$$

Proof: (4.27) is derived from (4.14) by a standard square function argument. Take φ as in Lemma 4.13 satisfying $\widehat{\varphi} = 1$ on $\left[-\frac{1}{2}, \frac{1}{2}\right]$. Estimate

$$\sup_{j\ge s}\left|\mathcal{F}^{-1}[\chi_{R_j}\mathcal{F}f]\right| \le \sup_{j\ge s}\left|\sum_{k=1}^{K} e^{2\pi i\lambda_k x}\left[(f\cdot e^{-2\pi i\lambda_k x}) * \varphi_{2^j}\right]\right| \tag{4.28}$$

$$+ \quad \left\{\sum_{j\ge s}\left|\mathcal{F}^{-1}\left[\left(\chi_{R_j} - \sum_{k=1}^{K}\widehat{\varphi_{2^j}}(\lambda-\lambda_k)\right)\mathcal{F}f\right]\right|^2\right\}^{1/2} . \tag{4.29}$$

By hypothesis on φ, $g * \varphi_{2^j} = (g * \varphi_{2^{s-1}}) * \varphi_{2^j}$ for $j \geq s$. Hence, applying (4.13) with $f_k = (f \cdot e^{-2\pi i \lambda_k x}) * \varphi_{2^{s-1}}$, (4.14) gives following bound on (4.28)

$$C(\log K)^2 \left\{ \int \left[\sum_{k=1}^{K} |\hat{f}(\lambda + \lambda_k)|^2 |\hat{\varphi}(2^{s-1}\lambda)|^2 \right] d\lambda \right\}^{1/2} \leq C(\log K)^2 \|f\|_2 \tag{4.30}$$

invoking the separation hypothesis of the λ_k's and the fact that $\operatorname{supp} \hat{\varphi} \subset [-1,1]$.

By Parseval's identity, (4.29) is bounded by

$$\left\{ \sum_{j \geq s} \int |\hat{f}(\lambda)|^2 \left[\chi_{R_j}(\lambda) - \sum_{k=1}^{K} \hat{\varphi}(2^j(\lambda - \lambda_k))\right]^2 d\lambda \right\}^{1/2} \leq$$

$$\sup_{\lambda \in \mathbb{R}} \left[\sum_{j \geq s} \left| \chi_{R_j}(\lambda) - \sum_{k=1}^{K} \hat{\varphi}(2^j(\lambda - \lambda_k)) \right| \right] \cdot \|f\|_2 \, . \tag{4.31}$$

Since $\chi_{R_j}(\lambda) - \sum\limits_{k=1}^{K} \hat{\varphi}(2^j(\lambda - \lambda_k))$ is bounded and 0 if either $\operatorname{dist}(\lambda, \Lambda) < 2^{-j-1}$ or $\operatorname{dist}(\lambda, \Lambda) > 2^{-j}$, the first factor in (4.31) is clearly bounded.

(4.27) is then implied by (4.30), (4.31).

Remark. In a later application, $\Lambda = \{\lambda_1, \ldots, \lambda_K\}$ will be typically a set of rationals $\frac{a}{q}$, $(a, q) = 1$, with $q \leq Q$ and the neighborhoods (major-arcs) considered $\ll Q^{-2}$. Thus the more restrictive Lemma 4.26 actually already suffices for our purpose. The statement of Lemma 4.11 is simpler, however, and the result may be of independent interest.

In the remainder of this section, we complete the proof of (4.12).

Lemma 4.32. *Let again*

$$R_j = \{\lambda \in \mathbb{R} \mid \min_{1 \leq k \leq K} |\lambda - \lambda_k| \leq 2^{-j}\} \quad \text{for} \quad j \in \mathbb{Z} \, .$$

Then

$$\left\| \sup_{j \in S} |\mathcal{F}^{-1}[\chi_{R_j} \mathcal{F} f]| \right\|_2 \leq (\log |S|) \|f\|_2 \tag{4.33}$$

for S a finite subset of $\mathbb{Z}$.

Proof: The argument is inspired from the Burkholder-Davis-Gundi-Stein (cf. [Ga]) dual version of Doob's maximal inequality. The only difference here is that the operators are not positive. We only use the fact that the R_j's are decreasing. Assume thus, redefining R_j,

$$R_{j+1} \subset R_j \, , \quad 1 \leq j \leq 2^s \quad \text{where} \quad s \sim \log |S| \, .$$

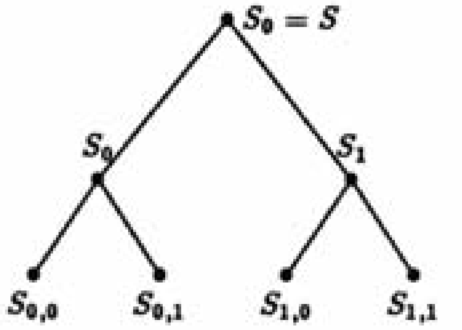
Denote B the best constant satisfying the inequality

$$\Big\| \sup_{1\leq j\leq 2^s} |\mathcal{F}^{-1}[\chi_{R_j}\mathcal{F}f]|\Big\|_2 \leq B\|f\|_2$$

or equivalently (by dualization)

$$\Big\| \sum_{j\leq 2^s} \mathcal{F}^{-1}[\chi_{R_j}\mathcal{F}g_j]\Big\|_2 \leq B\Big\| \sum_{j\leq 2^s} |g_j|\Big\|_2 . \tag{4.34}$$

Identify S and $\{1,2,\ldots,2^s\}$ and let $(S_c)_{|c|\leq s}$ be a diadic partitioning of S



Denoting $\widetilde{g}_j = \mathcal{F}^{-1}[\chi_{R_j}\mathcal{F}g_j]$, clearly

$$\langle \widetilde{g}_j, \widetilde{g}_k\rangle = \langle g_j, \widetilde{g}_k\rangle \quad \text{for} \quad j\leq k . \tag{4.35}$$

Using this fact and Hölder's inequality, one gets from definition of B

$$\begin{aligned}
\Big\|\sum_{j\in S}\widetilde{g}_j\Big\|_2^2 &= \sum \|\widetilde{g}_j\|_2^2 + 2\sum_{j<k}\langle\widetilde{g}_j,\widetilde{g}_k\rangle \\
&\leq \sum \|g_j\|_2^2 + 2\sum_{|c|<s}\Big|\Big\langle \sum_{j\in S_{c,0}} g_j \,,\, \sum_{k\in S_{c,1}}\widetilde{g}_k\Big\rangle\Big| \\
&\leq \sum \|g_j\|_2^2 + 2B\sum_{|c|<s}\Big\|\sum_{j\in S_{c,0}}|g_j|\Big\|_2\Big\|\sum_{J\in S_{c,1}}|g_j|\Big\|_2 \\
&\leq (1+2Bs)\Big\|\sum |g_j|\Big\|_2^2 .
\end{aligned}$$

Consequently, $B^2 \leq 1 + 2Bs \Longrightarrow B \leq cs$, proving (4.33).

Proof of (4.12). Define

$$S = \{j \in \mathbb{Z} \mid K^{-1}2^{-j} < |\lambda_k - \lambda_{k'}| < K2^{-j} \quad \text{for some} \quad 1 \le k \ne k' \le K\} .$$

Thus obviously

$$|S| \le K^3 . \tag{4.36}$$

Define further

$$Z_r = \{j \in \mathbb{Z} \backslash S | R_j \quad \text{has} \quad r \quad \text{components}\}$$

for $1 \le r \le K$. Hence

$$Z_1 \prec Z_2 \prec \cdots \prec Z_K \tag{4.37}$$

where Z_1, Z_K are half-lines and Z_r a finite segment, for $1 < r < K$. For $r > 1$, let $j_r = \min Z_r$. By construction, there is a set $\Lambda_r \subset \{\lambda_k\}$ satisfying

$$|\lambda - \lambda'| > 2^{-j_r} \quad \text{for} \quad \lambda \ne \lambda' \quad \text{in} \quad \Lambda_r \tag{4.38}$$

$$\bigcup_{\lambda \in \Lambda_r} [\lambda - 2^{-j}, \lambda + 2^{-j}] \subset R_j \subset \bigcup_{\lambda \in \Lambda_r} [\lambda - 2^{-j+1}, \lambda + 2^{-j+1}] \quad \text{for} \quad j \in Z_r . \tag{4.39}$$

To prove (4.12) we proceed again by duality and estimate the best B fulfilling

$$\Big\| \sum_j \tilde{g}_j \Big\|_2 \le B \Big\| \sum |g_j| \Big\|_2 \qquad \tilde{g}_j = \mathcal{F}^{-1}[\chi_{R_j} \mathcal{F} g_j] .$$

Estimate using (4.32) and (4.36)

$$\Big\| \sum_j \tilde{g}_j \Big\|_2 \le \Big\| \sum_{j \in S} \tilde{g}_j \Big\|_2 + \Big\| \sum_r \big(\sum_{j \in Z_r} \tilde{g}_j \big) \Big\|_2 \le (\log K) \Big\| \sum |g_j| \Big\|_2 + \Big\| \sum_r \tilde{G}_r \Big\|_2$$

denoting

$$G_r = \sum_{j \in Z_r} g_j \quad \text{and} \quad \tilde{G}_r = \sum_{j \in Z_r} \tilde{g}_j .$$

Because $Z_r \prec Z_{r'}$ for $r < r'$, we have for $j \in Z_r$, $j' \in Z_{r'}$

$$\langle \tilde{g}_j, \tilde{g}_{j'} \rangle = \langle g_j, \tilde{g}_{j'} \rangle .$$

Hence

$$\left\langle \tilde{G}_r, \tilde{G}_{r'} \right\rangle = \left\langle G_r, \tilde{G}_{r'} \right\rangle$$

and

$$\left\| \sum \widetilde{G}_r \right\|_2^2 = \sum \|\widetilde{G}_r\|_2^2 + 2 \sum_{r<r'} \left\langle G_r, \widetilde{G}_{r'} \right\rangle .$$

The same argument as in (4.32) then shows that

$$B^2 \leq (\log K)^2 + B_1^2 + B(\log K) \tag{4.40}$$

where B_1 has to satisfy

$$\left\| \sup_{j \in Z_r} |\mathcal{F}^{-1}[\chi_{R_j} \mathcal{F} f]| \right\|_2 \leq B_1 \|f\|_2 . \tag{4.41}$$

In order to estimate B_1, apply (4.26) with $\Lambda = \Lambda_r$, $s = j_r$, taking in account (4.38) and $j \geq j_r$ for $j \in Z_r$. Invoking then (4.39) and a square function argument such as in (4.31), it follows that $B_1 < C(\log K)^2$. Substitution in (4.40) yields that $B < C(\log K)^2$. This proves (4.12), hence Lemma's 4.11 and 4.1.

5. Behaviour of Exponential Sums

In analyzing the Fourier multipliers appearing in proving Theorems 1 and 2, information is needed on exponential sums (1.5), i.e.,

$$\varphi_N(\overline{\alpha}) = \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i p(n,\overline{\alpha})} \tag{5.1}$$

where

$$p(x, \overline{\alpha}) = \alpha_1 x + \cdots + \alpha_d x^d \quad \text{and} \quad \overline{\alpha} = (\alpha_1, \ldots, \alpha_d) \in [0,1]^d . \tag{5.2}$$

In this section, some well-known results and procedures are summarized. The estimates required are mainly provided by H. Weyl's basic lemma

Lemma 5.3. *Let $f(x) = \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_d x^d$ and $\left|\alpha_d - \frac{a}{q}\right| < \frac{1}{q^2}$ where $(a,q) = 1$. Then* $(\forall \varepsilon > 0)$

$$\left| \sum_{m=1}^{n} e^{2\pi i f(m)} \right| \leq C_\varepsilon n^{1+\varepsilon} [q^{-1} + n^{-1} + q n^{-d}]^\rho \quad \text{where} \quad \rho = \frac{1}{2^{d-1}} \tag{5.4}$$

(cf. [Vaug] or [Vin] for a proof).

Denote Q the set of rationals. For $\delta = \delta(d) > 0$ and $\theta_1, \ldots, \theta_d \in [0,1] \cap Q$ with common denominator $q < N^\delta$, define the "major box" in the d-dimensional torus

$$\mathcal{M}(\theta_1, \ldots, \theta_d) = \left\{ \overline{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \Pi^d \mid |\alpha_j - \theta_j| < N^{-j+\delta} \quad (1 \leq j \leq d) \right\} . \tag{5.5}$$

The following fact may be found in [Vin] (ch. IV, Th. 3) and can be proved by iterated applications of Lemma 5.3 combined with Dirichlet's principle.

Lemma 5.6. *If $\overline{\alpha}$ does not belong to some major box as defined above, then*

$$\left|\varphi_N(\overline{\alpha})\right| < CN^{-\delta'} . \tag{5.7}$$

Here $\varphi_N(\overline{\alpha})$ is defined by (5.1) and $C, \delta' > 0$ depend on d.

One may describe the shape of $\varphi_N(\overline{\alpha})$ on $\mathcal{M}(\theta_1, \ldots, \theta_d)$. Let $\theta_j = \frac{a_j}{q}$, $\alpha_j = \theta_j + \beta_j$ and $|\beta_j| < N^{-j+\delta}$. Writing $n = qs + r$ where $0 \le s < \frac{N}{q}$ and $r = 0, 1, \ldots, q-1$, one has for $j = 1, \ldots, d$

$$\alpha_j n^j = (\theta_j + \beta_j)(qs + r)^j \in \mathbb{Z} + \theta_j r^j + \beta_j q^j s^j + o(N^{-1+2\delta}) \tag{5.8}$$

since $q < N^\delta$. Hence, clearly

$$\varphi_N(\overline{\alpha}) = \left\{\frac{1}{q}\sum_{r=0}^{q-1} e^{2\pi i(r\theta_1 + \cdots + r^d\theta_d)}\right\}\left\{\frac{q}{N}\sum_{s=0}^{N/q} e^{2\pi i(\beta_1 qs + \cdots + \beta_d q^d s^d)}\right\} + o(N^{-1/2}) . \tag{5.9}$$

For $(a_1, \ldots, a_d, q) = 1$ and $\theta_j = \frac{a_j}{q}$, define

$$S(q, a_1, \ldots, a_d) = \frac{1}{q}\sum_{r=0}^{q-1} e^{2\pi i(r\theta_1 + \cdots + r^d\theta_d)} . \tag{5.10}$$

Letting

$$V_N(\overline{\beta}) = \frac{1}{N}\int_0^N e^{2\pi i(\beta_1 y + \beta_2 y^2 + \cdots + \beta_d y^d)} dy . \tag{5.11}$$

(5.9) and the estimates $|\beta_j| < N^{-j+\delta}$ easily yield, replacing the second factor in (5.9) by its continuous substitute

Lemma 5.12. *For $\overline{\alpha} \in \mathcal{M}(\overline{\theta})$, $\overline{\alpha} = \overline{\theta} + \overline{\beta}$*

$$\varphi_N(\overline{\alpha}) = S(q, a_1, \ldots, a_d)V_N(\overline{\beta}) + O(N^{-1/2}) \tag{5.13}$$

where $\theta_j = \frac{a_j}{q}$.

Recall also

Lemma 5.14. *If* $(q, a_1, \ldots, a_d) = 1$, *then*

$$\left|S(q, a_1, \ldots, a_d)\right| \leq cq^{-\delta'} \tag{5.15}$$

where $\delta' = \delta(d) > 0$,
which is clearly a consequence of (5.3).

In this work, we will not need finer information on the $S(q, a_1, \ldots, a_d)$, such as multiplicativity properties and A. Weil's estimate for q a prime number.

Finally, some estimates on the function

$$V_N(\overline{\beta}) = \int_0^1 e^{2\pi i(\beta_1 N y + \beta_2 N^2 y^2 + \cdots + \beta_d N^d y^d)} dy\ . \tag{5.16}$$

Lemma 5.17.

$$\left|1 - V_N(\overline{\beta})\right| < C \sum_{j=1}^{d} |\beta_j| N^j \tag{5.18}$$

$$\left|V_N(\overline{\beta})\right| < C \left[1 + \sum_{j=1}^{d} |\beta_j| N^j\right]^{-1/d} \tag{5.19}$$

where $C = c(d)$.

The first estimate (5.18) is obvious and the second (5.19) results from van der Corput's estimate on oscillatory integrals.

6. Ergodic Theorems in L^2

In this section, we prove Theorem 1 for functions of class L^2. This result appears in [B$_1$]. The argument presented here uses less structure. According to the discussion in section 1, the maximal inequality and convergence problem for the averages

$$A_N f = \frac{1}{N} \sum_{n=1}^{N} T^{p(n)} f \tag{6.1}$$

$$p(x) = b_1 x + b_2 x^2 + \cdots + b_d x^d \qquad b_j \in \mathbb{Z} \quad \text{and} \quad b_d > 0 \tag{6.2}$$

where reduced to proving certain inequalities for the shift model $(\mathbb{Z}, S)$. In the case of the shift

$$A_N\mathcal{F} = f * K_N \quad \text{where} \quad K_N = \frac{1}{N}\sum_{n=1}^{N} \delta_{\{p(n)\}} \tag{6.3}$$

and δ_x stands for the Dirac measure at $x \in \mathbb{Z}$. Hence, introducing Fourier transform

$$A_N f = \mathcal{F}^{-1}\left[\mathcal{F}[K_N]\cdot\mathcal{F}[f]\right] \tag{6.4}$$

where for $\alpha \in \Pi \simeq [0,1]$

$$\mathcal{F}[K_N](\alpha) = \frac{1}{N}\sum_{n=1}^{N} e^{-2\pi i p(n)\cdot\alpha} = \varphi_N(-b_1\alpha, \ldots, -b_d\alpha) \equiv \varphi_N(-\alpha\cdot\bar{b}) \ . \tag{6.5}$$

For $s \geq 0$, define an exhaustion of the rationals in $\mathbb{T}$

$$\mathcal{R}_s = \left\{\theta \in Q \cap [0,1] \mid \theta = \frac{a}{q} \ , \ (a,q) = 1 \quad \text{and} \quad 2^s \leq q < 2^{s+1}\right\} \tag{6.6}$$

which is considered as subset of Π. Thus $\mathcal{R}_0 = \{0 \equiv 1\}$.

Denote ζ a smooth function on $\mathbb{R}$ with $\zeta = 1$ on $\left[-\frac{1}{10}, \frac{1}{10}\right]$ and $\zeta = 0$ outside $\left[-\frac{1}{5}, \frac{1}{5}\right]$. (The smoothness of ζ will be irrelevant for the L^2-theory but has importance when considering L^r-estimates for $r < 2$ in the next section).

Define

$$\psi_{s,N}(\alpha) = \sum_{\theta\in\mathcal{R}_s} S(\theta) w_N(\alpha - \theta)\zeta\big(10^s(\alpha-\theta)\big) \tag{6.7}$$

where, with the notations (5.10), (5.11) of section 5

$$S(\theta) = S(q', a_1', \ldots, a_d') \quad \text{where} \quad -\theta\cdot b_j \equiv \frac{a_j'}{q'} (\operatorname{mod} 1) \quad \text{and} \quad (a_1', \ldots, a_d', q') = 1 \tag{6.8}$$

$$w_N(\beta) = V_N(-\beta b_1, \ldots, -\beta b_d) \ . \tag{6.9}$$

Thus it follows from Lemma 5.12 that if $\theta = \frac{a}{q}$, $q < N^\delta$

$$\mathcal{F}[K_N](\alpha) = S(\theta) w_N(\alpha-\theta) + O(N^{-1/2}) \quad \text{if} \quad |\alpha - \theta| < N^{-d+\delta} \ . \tag{6.10}$$

Also, since $q' > \frac{q}{b_d}$, if $\theta = \frac{a}{q}$, $(a,q) = 1$, one has by (5.15) with notation (6.8)

$$|S(\theta)| < C2^{-s\delta'} \qquad \text{for} \qquad \theta \in \mathcal{R}_s \ . \tag{6.11}$$

From (6.9) and (5.17)

$$\left|1 - w_N(\beta)\right| < C|\beta|\cdot N^d \tag{6.12}$$

$$\left|w_N(\beta)\right| < C\left[1 + |\beta| N^d\right]^{-1/d} \ . \tag{6.13}$$

Observe also that the summands in (6.7) are disjointly supported, by definition of $\mathcal{R}_s$ and ζ.

Lemma 6.14. *There exists $\delta_1 > 0$ such that the uniform estimate*

$$\left|\mathcal{F}[K_N](\alpha) - \sum_{s\geq 0} \psi_{s,N}(\alpha)\right| < CN^{-\delta_1} \tag{6.15}$$

holds.

This lemma permits the replacement of $\mathcal{F}[K_N]$ in (6.4) by more explicit multipliers which will be taken care of by Lemma 4.1.

Proof of (6.14). Redefine major arcs in Π by letting

$$\mathcal{M}(\theta) = \{\alpha \in \Pi \mid |\alpha - \theta| < N^{-d+\delta}\} \tag{6.16}$$

for θ a rational $\frac{a}{q}$, $1 \leq a \leq q$, $(a,q) = 1$ with $q < N^\delta$.

Case 1. α belongs to an arc $\mathcal{M}(\theta_0)$.

Assume $\theta_0 \in \mathcal{R}_{s_0}$, thus $2^{s_0} < N^\delta$. Let s_1 be a positive integer (depending on N), to be specified later. Estimate using (6.10), (6.11)

$$\left|\mathcal{F}[K_N](\alpha) - \sum \psi_{s,N}(\alpha)\right| \leq |1 - \zeta(10^{s_0}(\alpha - \theta_0))| + \sum_{s \leq s_1} \sup_{\substack{\theta \in \mathcal{R}_s \\ \theta \neq \theta_0}} |w_N(\alpha - \theta)| + C2^{-s_1\delta'} + CN^{-1/2} . \tag{6.17}$$

Since $10^{s_0} < N^{4\delta}$ and $|\alpha - \theta_0| < N^{-d+\delta} < N^{-1}$, the first term in (6.17) vanishes. Letting $2^{s_1} \sim N^\delta$ and writing $|\alpha - \theta| \geq |\theta - \theta_0| - |\alpha - \theta_0|$, $|\theta - \theta_0| > \frac{1}{2}q^{-1}2^{-s_1} \geq \frac{1}{4}N^{-2\delta}$ for $\theta \in \mathcal{R}_s$, $s \leq s_1$, $\theta \neq \theta_0$ and $|\alpha - \theta_0| < N^{-1}$, it follows that $|\alpha - \theta| > \frac{1}{2}|\theta - \theta_0|$. Thus the second term of (6.17) is bounded by $(\log N) \cdot N^{-1+\frac{2\delta}{d}}$, invoking (6.13). Hence (6.15) holds.

Case 2. α does not belong to a major arc.

Clearly, from definition (5.5) and (5.7), we have $|\mathcal{F}[K_N](\alpha)| < CN^{-\delta'}$, by (6.5). Write for $2^{s_1} < \frac{1}{2}N^\delta$, $2^{s_1} \sim N^\delta$

$$\left|\sum \psi_{s,N}(\alpha)\right| \leq \sum_{s \leq s_1} \sup_{\theta \in \mathcal{R}_s} |w_N(\alpha - \theta)| + C2^{-s_1\delta'} . \tag{6.18}$$

By definition of $\mathcal{M}(\theta)$, it follows from the hypothesis on α that $|\alpha - \theta| > N^{-d+\delta}$ whenever $\theta \in \mathcal{R}_s$, $s \leq s_1$. Hence $|w_N(\alpha - \theta)| < CN^{-\delta/d}$ by (6.13) and (6.18) implies again (6.15).

This proves Lemma (6.14).

It is clear that when proving the maximal inequality

$$\left\| \sup_N |f * K_N| \right\|_{\ell^2(\mathbb{Z})} \le C \|f\|_{\ell^2(\mathbb{Z})} \tag{6.19}$$

the function f may be taken positive and hence the supremum taken over the set $Z_1 = \{2^k \mid k = 1, 2, \ldots\}$. Denoting

$$\psi_N = \sum_S \psi_{S,N} \tag{6.20}$$

estimate by (6.15) and Parseval

$$\begin{aligned} \left\| \sup_{N \in Z_1} |f * K_N| \right\|_2 &\le \left\| \sup_{n \in Z_1} |\mathcal{F}^{-1}[\psi_N \mathcal{F} f]| \right\|_2 + \Big(\sum_{n \in Z_1} \|\mathcal{F}[K_N] - \psi_N\|_\infty^2 \Big)^{1/2} \|f\|_2 \\ &\le \sum_{s=0}^{\infty} \left\| \sup_{Z_1} |\mathcal{F}^{-1}[\psi_{s,N} \mathcal{F} f]| \right\|_2 + C\|f\|_2 \, . \end{aligned} \tag{6.21}$$

To estimate the contribution of the first terms, define

$$\tilde{\psi}_{s,N}(\alpha) = \sum_{\theta \in \mathcal{R}_s} S(\theta) \chi\big(N^d(\alpha - \theta)\big) \zeta\big(10^s(\alpha - \theta)\big) \tag{6.22}$$

where $\chi = \chi_{[-1,1]}$ considered as function on $\mathbb{R}$. It easily follows from (6.11),(6.12),(6.13) that there is a uniform estimate

$$\sum_{N \in Z_1} |\psi_{s,N} - \tilde{\psi}_{s,N}| \le C 2^{-s\delta'} \, . \tag{6.23}$$

Therefore, again by a square function argument

$$\left\| \sup_{Z_1} |\mathcal{F}^{-1}[\psi_{s,N} \mathcal{F} f]| \right\|_2 \le \left\| \sup_{Z_1} |\mathcal{F}^{-1}[\tilde{\psi}_{s,N} \mathcal{F} f]| \right\|_2 + C 2^{-s\delta'} \|f\|_2 \, . \tag{6.24}$$

For $N \in Z_1$, write $N^d = 2^j$ and R_j the 2^{-j}-neighborhood of $R_s \subset \Pi$. Thus, letting

$$\mathcal{F}[g_s] = \mathcal{F}[f] \sum_{\theta \in \mathcal{R}_s} S(\theta) \zeta\big(10^s(\alpha - \theta)\big) \tag{6.25}$$

$$\tilde{\psi}_{s,N} \mathcal{F} f = \mathcal{F}[g_s] \cdot \chi_{R_j} \tag{6.26}$$

it follows from inequality (4.3) in Lemma 4.1 that

$$\begin{aligned} \left\| \sup_{N \in Z_1} |\mathcal{F}^{-1}[\tilde{\psi}_{s N} \cdot \mathcal{F} f]| \right\|_2 &\le \left\| \sup_{j \in \mathbb{Z}_+} |\mathcal{F}^{-1}[\mathcal{F}[g_s] \chi_{R_j}]| \right\|_2 \\ &\le C (\log |\mathcal{R}_S|)^2 \|g_s\|_2 \end{aligned} \tag{6.27}$$

By its definition, $|\mathcal{R}_s| < 4^s$ and it follows from (6.11) and Parseval that $\|g_S\|_2 \le C2^{-s\delta'}\|f\|_2$. Substitution in (6.24) yields the bound

$$\left\| \sup_{Z_1} |\mathcal{F}^{-1}[\psi_{sN}\mathcal{F}f]| \right\|_2 \le C \cdot s^2 2^{-s\delta'} \|f\|_2 \tag{6.28}$$

and hence, substituting in (6.21)

$$\left\| \sup_{N \in Z_1} |f * K_N| \right\|_2 \le C\left(\sum_{s=0}^{\infty} s^2 2^{-s\delta'} \right) \|f\|_2 \le C\|f\|_2 \tag{6.29}$$

proving the maximal inequality

$$\left\| \sup_N |A_N f| \right\|_2 \le C\|f\|_2 \, . \tag{6.30}$$

Next, we verify the almost sure convergence using the method described in section 2 of this paper. Thus we prove an inequality (2.14) in $(\mathbb{Z}, S)$

$$\sum_{j=1}^{J} \|\mathcal{M}_j f\|_2 \le o(J)\|f\|_2 \tag{6.31}$$

denoting

$$\mathcal{M}_j f = \sup_{\substack{N_j < N < N_{j+1} \\ N \in Z_\varepsilon}} |f * (K_N - K_{N_j})| \tag{6.32}$$

where $Z_\varepsilon = \{[(1+\varepsilon)^n] \; ; \; n = 1, 2, \dots\}$, $\varepsilon > 0$ fixed, and N_j is any rapidly increasing sequence ($N_{j+1} > 2N_j$).

We again apply the Fourier transform method. With previous definitions, it again follows from (6.15) that $f * (K_N - K_{N_j})$ may be replaced by $\mathcal{F}^{-1}[(\psi_N - \psi_{N_j})\mathcal{F}f]$ when defining $\mathcal{M}_j f$. Fixing s_0, it follows from the previous inequality (6.28) that then

$$\|\mathcal{M}_j f\|_2 \le \sum_{s \le s_0} \left\| \sup_{\substack{N_j < N < N_{j+1} \\ N \in Z_\varepsilon}} |\mathcal{F}^{-1}[(\psi_{s,N} - \psi_{s,N_j})\mathcal{F}f]| \right\|_2 + C\varepsilon^{-1} 2^{-\delta'' s_0} \|f\|_2 \tag{6.33}$$

where the second term in (6.33) will be $o(\|f\|_2)$ for appropriate s_0. Thus it suffices to verify (6.31), defining now

$$\mathcal{M}_j f = \sup_{\substack{N_j < N < N_{j+1} \\ N \in Z_\varepsilon}} |\mathcal{F}^{-1}[(w_N - w_{N_j})\mathcal{F}f]| \tag{6.34}$$

where w_N is given by (6.9). The reader will indeed verify that summing the first terms of (6.33) over $j = 1, \dots, J$ will only introduce an additional factor (depending on s_0).

Let $\chi = \chi_{[0,1]}$ and

$$\widetilde{\mathcal{M}}_j f = \sup_{N_j < N < N_{j+1}} \left| f * (\chi_{(N^d)} - \chi_{(N_j^d)}) \right| \tag{6.35}$$

where $\chi_t = \frac{1}{t}\chi_{[0,t]}$. Since pointwise clearly (with v_S as in section 3)

$$\left\{ \sum_{j=1}^{J} (\widetilde{\mathcal{M}}_j f)^2 \right\}^{1/2} \le J^{1/4} \left\| \{ f * \chi_N \mid N = 1, 2, \dots \} \right\|_{v_4} \tag{6.36}$$

it follows from (6.36) and (3.26) that

$$\sum_{j=1}^{J} \|\widetilde{\mathcal{M}}_j f\|_2^2 \le C J^{1/2} \|f\|_2^2$$

hence

$$\sum_{j=1}^{J} \|\mathcal{M}_j f\|_2^2 \le C \sum_{N \in Z_s} \|\mathcal{F}^{-1}[(w_N - \mathcal{F}(\chi_{N^d}))\mathcal{F} f\|_2^2 + C J^{1/2} \|f\|_2^2 \,. \tag{6.37}$$

This first term in (6.37) is bounded by

$$\sup_\alpha \left[\sum_{N \in Z_s} \left| w_N(\alpha) - \widehat{\chi}(N^d \alpha) \right|^2 \right] \|f\|_2^2 < C_\epsilon \|f\|_2^2 \tag{6.38}$$

using the fact that, by (6.12), (6.13),

$$\left| w_N(\alpha) - \widehat{\chi}(N^d \alpha) \right| \le C \min\left(|\alpha| N^d, (|\alpha| N^d)^{-1/d} \right) . \tag{6.39}$$

Hence, for $\mathcal{M}_j f$ defined by (6.34), by (6.37),(6.38)

$$\sum \|\mathcal{M}_j f\|_2 \le C_\epsilon J^{3/4} \|f\|_2 \tag{6.40}$$

independently of the choice of the sequence $N_1 \ll N_2 \ll \dots \ll N_J$. The proof of (6.31) is now completed, and so is the proof of Theorem 1 for L^2-functions.

Observe finally that if T is weakly mixing, then $A_N f \longrightarrow \int f \, d\mu$ in L^2 (hence a.s.). Indeed T has no point spectrum as unitary operator and $\frac{1}{N} \sum_{n \le N} z^{p(n)} \xrightarrow{N \to \infty} 0$ for $z \in C_1 = \{ z \in C_1 \mid |z| = 1 \}$, except on a countable set.

7. Ergodic Theorems in L^p, $p > 1$

The purpose of this section is to extend the L^2-theory to L^p, $p > 1$. Of course, only the maximal inequality

$$\Big\| \sup_N |A_N f| \Big\|_p \le c \|f\|_p \tag{7.1}$$

needs to be shown. Once (7.1) is obtained, the a.s. convergence for functions f of class $L^p(\mu)$ reduces to bounded functions and hence is taken care of by the L^2 result, obtained in the previous section.

The partial result was obtained in [B$_2$] $\left(p > \frac{1+\sqrt{5}}{2}\right)$.

Considering again the shift model $(\mathbb{Z}, S)$, (7.1) becomes

$$\| \sup |f * K_N| \|_p \le C \|f\|_p \, ; \qquad K_N = \frac{1}{N} \sum_{n=1}^{N} \delta_{\{p(n)\}} \, . \tag{7.2}$$

The proof of (7.2) by Fourier Analysis methods is more delicate than in the L^2-case because the Fourier multipliers involved in the argument need to have good bounds on L^p.

We use the notations of the previous section. Thus in particular

$$S(\theta) = \frac{1}{q} \sum_{r=0}^{q-1} e^{-2\pi i p(r)\theta} \quad \text{for} \quad \theta = \frac{a}{q} \quad \text{and} \quad w_N(\beta) = \int_0^1 e^{-2\pi i p(Ny)\beta} dy \, . \tag{7.3}$$

Denote again ζ a smooth function on $\mathbb{R}$, $0 \le \zeta \le 1$, $\operatorname{supp} \zeta \subset \left[-\frac{1}{2}, \frac{1}{2}\right]$ and $\zeta = 1$ on $\left[-\frac{1}{4}, \frac{1}{4}\right]$.

The following lemma will be useful when comparing $L^p(\mathbb{R})$ and $\ell^p(\mathbb{Z})$-norms.

Lemma 7.4. *For $1 < q < \varepsilon D$, $\varepsilon = o(1)$, one has*

$$\Big\| \int F(\beta) e^{2\pi i \beta q y} \zeta(D\beta) d\beta \Big\|_{L^p(\mathbb{R})} \sim \Big\| \int F(\beta) e^{2\pi i \beta q y} \zeta(D\beta) d\beta \Big\|_{\ell^p(\mathbb{Z})} . \tag{7.5}$$

Proof: Observe first that by Bernstein's inequality and the hypothesis for $0 \le u \le 1$

$$\Big\| \int F(\beta) [e^{2\pi i q \beta u} - 1] e^{2\pi i q \beta y} \zeta(D\beta) d\beta \Big\|_{L^p} \le C\varepsilon \Big\| \int F(\beta) e^{2\pi i q \beta y} \zeta(D\beta) d\beta \Big\|_{L^p} . \tag{7.6}$$

We first prove the inequality $\| \quad \|_{\ell^p(\mathbb{Z})} \le \rho \| \quad \|_{L^p(\mathbb{R})}$ in (7.5), for some bounded ρ. Let $0 \le u < 1$ and write

$$\begin{aligned} \Big\| \int F(\beta) e^{2\pi i \beta q y} \zeta(D\beta) d\beta) \Big\|_{\ell^p} \le & \Big\| \int F(\beta) e^{2\pi i \beta q (y+u)} \zeta(D\beta) d\beta \Big\|_{\ell^p} + \\ & + \Big\| \int F(\beta) [1 - e^{2\pi i \beta q u}] e^{2\pi i \beta q y} \zeta(D\beta) d\beta \Big\|_{\ell^p} . \end{aligned} \tag{7.7}$$

Integrating the pth power of the first term in (7.7) in u, the $L^p(\mathbb{R})$-norm is obtained. Let ρ be an a priori constant satisfying the above inequality, the second term in (7.7) may be estimated for fixed u

$$\rho\left\| \int F(\beta)[1-e^{2\pi i\beta q u}]e^{2\pi i\beta q y}\zeta(D\beta)d\beta\right\|_{L^p} \leq C\epsilon\rho\left\|\int F(\beta)e^{2\pi i q\beta y}\zeta(D\beta)d\beta\right\|_{L^p}$$

invoking (7.6). Thus it follows that $\rho \leq 1 + C\epsilon\rho$, from where the boundedness of ρ.

To prove the converse inequality in (7.5), write

$$\|\quad\|_{L^p(\mathbb{R})} \leq \|\quad\|_{\ell^p(\mathbb{Z})} + \left\{\int_0^1 \left\|\int F(\beta)e^{2\pi i\beta q y}[1-e^{2\pi i\beta q u}]\zeta(D\beta)d\beta\right\|^p_{\ell^p(dy)} du\right\}^{1/p} \tag{7.8}$$

and apply the $\|\quad\|_{\ell^p} \leq \rho\|\quad\|_{L^p}$ inequality and (7.6) to estimate

$$\left\|\int F(\beta)e^{2\pi i\beta q y}[1-e^{2\pi i\beta q u}]\zeta(D\beta)d\beta\right\|_{\ell^p(dy)} \leq C\epsilon\rho\left\|F(\beta)e^{2\pi i q\beta y}\zeta(D\beta)d\beta\right\|_{L^p} \tag{7.9}$$

for $0 \leq u \leq 1$. Since $C\epsilon\rho < \frac{1}{2}$ for ϵ taken small enough, substitution of (7.9) in (7.8) yields the converse inequality, proving (7.5).

Lemma 7.10. *For $S(\theta)$ defined by (7.3), the $\ell^1(\mathbb{Z})$-norm of the Fourier transform of the function on Π*

$$\sum_{0\leq a\leq q} S\left(\frac{a}{q}\right) F\left(\alpha - \frac{a}{q}\right) \tag{7.11}$$

is bounded by

$$q\sum_{j\in\mathbb{Z}} \sup_{0\leq x<q} \left|\mathcal{F}F(jq+x)\right| . \tag{7.12}$$

Proof: By definition of $S(\theta)$, the Fourier transform of (7.11) at the point $x \in \mathbb{Z}$ equals

$$\sum_{0\leq a<q} S\left(\frac{a}{q}\right) e^{2\pi i\frac{a}{q}x}\mathcal{F}F(x) = (\#\{0\leq r<q \mid x-p(r)\in q\mathbb{Z}\})\mathcal{F}F(x) .$$

Thus the $\ell'(\mathbb{Z})$-norm is bounded by $\sum_{r=0}^{q-1}\sum_{j\in\mathbb{Z}}\left|\mathcal{F}F\big(jq+p(r)\big)\right|$, hence by (7.12)

Lemma 7.13. *Let $1<q<D$. Then with the notations (7.3)*

$$\left\|\sum_{0\leq a<q} S\left(\frac{a}{q}\right)\int w_N\left(\alpha-\frac{a}{q}\right)\zeta\left(D\left(\alpha-\frac{a}{q}\right)\right)e^{2\pi i\alpha x}d\alpha\right\|_{\ell^1(\mathbb{Z})} < C . \tag{7.14}$$

Proof: Apply (7.10) with $F(\beta) = w_N(\beta)\zeta(D\beta)$. It follows from (7.3) that $w_N(\beta)$ is the Fourier transform of the image measure ν_N under the mapping $p(Ny) : [0,1] \longrightarrow \mathbb{R}$. Hence $\mathcal{F}F = \nu_N * \left(\mathcal{F}^{-1}[\zeta]\right)_D$ and (7.12) is bounded by

$$\frac{q}{D} \sum_{j \in \mathbb{Z}} \sup_{0 \le x < q} \int_0^1 \left| \mathcal{F}^{-1}[\zeta] \left(\frac{x + jq - p(Ny)}{D} \right) \right| dy \, . \tag{7.15}$$

Since $\left|\mathcal{F}^{-1}[\zeta](t)\right| < C(1+t^2)^{-1}$, $\sum_{j \in \mathbb{Z}} \sup_{0 \le x < q} \left|\mathcal{F}^{-1}[\zeta]\left(\frac{x+jq}{D}\right)\right| < C\left(\frac{D}{q}+1\right)$. Substitution in (7.15),(7.14) follows.

There is the following real Analysis maximal inequality

Lemma 7.16. *For $p > 1$ and $f \in L^p(\mathbb{R})$*

$$\left\| \sup_N \left| \mathcal{F}^{-1}[w_N \mathcal{F} f] \right| \right\|_{L^p(\mathbb{R})} \le C \|f\|_{L^p(\mathbb{R})} \, . \tag{7.17}$$

Proof: As observed eariler w_N is the Fourier transform of the image measure ν_N under the mapping $p(y) : [0,N], \frac{dy}{N} \longrightarrow \mathbb{R}$. Thus we have to estimate $\left\| \sup_N |f * \nu_N| \right\|_p$. For t sufficiently large, one has that $\left.\frac{d\nu_N}{ds}\right|_{s=t} = \frac{1}{Np'\left(p^{-1}(t)\right)}$ which is of the order $\frac{1}{N} t^{-1+\frac{1}{d}}$ in size. Thus the problem reduces to show that

$$\left\| \sup_N \left| f * \left[\frac{1}{N} t^{-1+\frac{1}{d}} \chi_{[0,N^d]}(t) \right] \right| \right\|_p \le C \|f\|_p \, . \tag{7.18}$$

Defining $k(t) = t^{-1+\frac{1}{d}} \chi_{[0,1]}$, $\frac{1}{N} t^{-1+\frac{1}{d}} \chi_{[0,N^d]}(t) = k_{(N^d)}$ where $k_s(t) = \frac{1}{s} k\left(\frac{t}{s}\right)$, $s > 0$. The fact that

$$\left\| \sup_{s > 0} |f * k_s| \right\|_p \le C \|f\|_p \tag{7.19}$$

follows from the Hardy-Littlewood maximal function boundedness on $\mathbb{R}$. This proves the lemma.

Next, we prove a discrete maximal inequality

Lemma 7.20. *Let $1 < q < \epsilon D$, $\epsilon = o(1)$. Then for $p > 1$*

$$\left\| \sup_N \left| \sum_{0 \le a < q} \int w_N(\beta) \mathcal{F} f \left(\frac{a}{q} + \beta \right) \zeta(D\beta) e^{2\pi i x\left(\frac{a}{q}+\beta\right)} d\beta \right| \right\|_{L^p(\mathbb{Z})} \le C_p \|f\|_{\ell^p(\mathbb{Z})} \, . \tag{7.21}$$

Proof: The main ingredient will be (7.16) and the problem is to pass from $\mathbb{R}$ to $\mathbb{Z}$. Writing $x \in \mathbb{Z}$ as $x = yq + z$, $z = 0, 1, \ldots, q-1$, the left member of (7.21) equals

$$\left\{ \sum_{0 \le z < q} \left\| \sup_N \left| \int w_N(\beta) F_z(\beta) \zeta(D\beta) e^{2\pi i \beta q y} d\beta \right| \right\|_{\ell^p(dy)}^p \right\}^{1/p} \tag{7.22}$$

denoting

$$F_z(\beta) = \sum_{0 \le a < q} \mathcal{F} f\left(\frac{a}{q} + \beta\right) e^{2\pi i z\left(\frac{a}{q}+\beta\right)} . \tag{7.23}$$

As in the proof of (7.4), denote ρ the apriori best constant in the inequality

$$\left\| \sup_N \left| \int w_N(\beta) F(\beta) \zeta(D\beta) e^{2\pi i \beta q y} d\beta \right| \right\|_{\ell^p} \le \rho \left\| \int F(\beta) \zeta(D\beta) e^{2\pi i \beta q y} d\beta \right\|_{\ell^p} . \tag{7.24}$$

For $0 \le u < 1$, write

$$\sup_N \left| \int w_N(\beta) F(\beta) \zeta(D\beta) e^{2\pi i \beta q y} d\beta \right| \le$$

$$\sup_N \left| \int w_N(\beta) F(\beta) \zeta(D\beta) e^{2\pi i \beta q (y+u)} d\beta \right| + \sup_N \left| \int w_N(\beta) F(\beta) \zeta(D\beta) [e^{2\pi i \beta q u} - 1] e^{2\pi i \beta q y} d\beta \right| . \tag{7.25}$$

Integrating the pth power of the first term of (7.25) in $u \in [0,1]$ gives by (7.16),(7.4)

$$q^{-1/p} \left\| \sup_N |\mathcal{F}^{-1}[w_N F \zeta(D \cdot)]| \right\|_{L^p} \le \tag{7.26}$$

$$C q^{-1/p} \| \mathcal{F}^{-1}[F\zeta(D\cdot)] \|_{L^p} =$$

$$C \left\| \int F(\beta) \zeta(D\beta) e^{2\pi i \beta q y} d\beta \right\|_{L^p(dy)} \sim \left\| \int F(\beta) \zeta(D\beta) e^{2\pi i q y} d\beta \right\|_{\ell^p(dy)} .$$

By definition of ρ, the ℓ^p-norm of the second term in (7.25) is for fixed $u \in [0,1]$ bounded by

$$\rho \left\| \int F(\beta) [e^{2\pi i \beta q u} - 1] \zeta(D\beta) e^{2\pi i \beta q y} d\beta \right\|_{\ell^p} . \tag{7.27}$$

Apply consecutively (7.5),(7.6),(7.5) to estimate (7.27) by

$$C \varepsilon \rho \left\| \int F(\beta) \zeta(D\beta) e^{2\pi i \beta q y} d\beta \right\|_{\ell^p} . \tag{7.28}$$

From (7.26),(7.27),(7.28), it follows that $\rho \le C + C\varepsilon\rho \Longrightarrow \rho < C$ assuming ε small enough. This yields (7.24).

Applying (7.24) with $F = F_z$ and substitution of (7.23) yields

$$\left\|\sup_N\left|\int w_N(\beta)F_z(\beta)\zeta(D\beta)e^{2\pi i\beta qy}d\beta\right|\right\|_{\ell^p(dy)} \leq$$

$$C\left\|\sum_{0\leq a<q} e^{2\pi i z\frac{a}{q}}\left[\int \mathcal{F}f\left(\frac{a}{q}+\beta\right)\zeta(D\beta)e^{2\pi i\beta(qy+z)}d\beta\right]\right\|_{\ell^p(dy)} =$$

$$C\left\|\sum_{0\leq a<q}\mathcal{F}^{-1}\left[\mathcal{F}f\cdot\zeta\left(D\left(\cdot-\frac{a}{q}\right)\right)\right](qy+z)\right\|_{\ell^p(dy)}$$

and summation over $z = 0,\ldots,q-1$ gives the following estimate on (7.22)

$$\left\|f*\mathcal{F}^{-1}\left[\sum_{0\leq a<q}\zeta\left(D\left(\cdot-\frac{a}{q}\right)\right)\right]\right\|_{\ell^p} < C\|f\|_{\ell^p}\ .$$

This completes the proof of Lemma 7.20.

Lemma 7.29. *Under the hypothesis of Lemma (7.20), for $p > 1$*

$$\left\|\sup_N\left|\sum_{0\leq a<q} S\left(\frac{a}{q}\right)\int w_N(\beta)\mathcal{F}f\left(\frac{a}{q}+\beta\right)\zeta(D\beta)e^{2\pi i x\left(\frac{a}{q}+\beta\right)}d\beta\right|\right\|_{\ell^p} \leq C_p\|f\|_{\ell^p}\ . \tag{7.30}$$

Proof: Apply (7.21) to the function g given by

$$\mathcal{F}g(\alpha) = \left[\sum_{0\leq a<q} S\left(\frac{a}{q}\right)\zeta\left(\frac{D}{4}\left(\alpha-\frac{a}{q}\right)\right)\right]\mathcal{F}f(\alpha)\ . \tag{7.31}$$

Observe that $\zeta(\frac{D}{4}\beta)\zeta(D\beta) = \zeta(D\beta)$ and the first factor in (7.31) is the Fourier-transform of an $\ell^1(\mathbb{Z})$-function, by taking $w_N = 1$ in (7.14). Inequality (7.30) now follows.

The following lemma in an important new ingredient in proving (7.1).

Lemma 7.32. *There is the following restricted maximal inequality*

$$\left\|\sup_{N_0<N<N_0^2}|f*K_N|\right\|_p \leq C_p(\log\log N_0)\|f\|_p\ , \quad \text{for } p > 1\ . \tag{7.33}$$

This is a problem about positive functions and hence N may be taken of the form $N = 2^k$, $k_0 \leq k \leq 2k_0$. Instead of considering the $\ell^p(\mathbb{Z})$-inequality, we will rather deal with functions f taken on a finite cyclic group $G \equiv \mathbb{Z}_J = \mathbb{Z}/J\mathbb{Z}$, where J is taken large enough (depending on N_0). The measure on G is the normalized counting measure and $f * K_N$ is the convolution on G of f and $\frac{1}{N}\sum\limits_{1\le n\le N}\delta_{\{p(n)\}}$. (7.33) is equivalent to

$$\Big\| \sup_{k_0\le k\le 2k_0} |f * K_{2^k}| \Big\|_{L^p(G)} \le C_p(\log k_0)\|f\|_{L^p(G)} \,. \tag{7.34}$$

The reason for this set-up is to invoke Stein's extrapolation theorem [St] according to which the inequalities (7.34) for $p > 1$ follow from the weaker inequalities

$$\Big\| \sup_{k_0\le k\le 2k_0} |f * K_{2^k}| \Big\|_{L^1(G)} \le C_p(\log k_0)\|f\|_{L^p(G)} \,. \tag{7.35}$$

Since (7.35) weakens for increasing p, one may assume $q = p' = \frac{p}{p-1}$ an integer. We replace (7.35) by its dual version

$$\Big\| \sum_{k=k_0}^{2k_0} (g_k * K_{2^k}) \Big\|_q < C_q(\log k_0) \tag{7.36}$$

whenever

$$g_k \ge 0 \quad , \quad \sum g_k \le 1 \,. \tag{7.37}$$

Let M (to be specified later) satisfy

$$M \sim \log k_0 \tag{7.38}$$

and put $L_k = K_{2^{Mk}}$ for simplicity. By splitting in sub-sums, (7.36) will clearly follow from

$$\Big\| \sum_{k_0<k<2k_0} (g_k * L_k) \Big\|_q < C_q \tag{7.39}$$

whenever $\{g_k\}$ fulfils (7.37). Denote ρ the smallest constant C_q satisfying (7.39).

In the sequel, let C stand for a constant depending on q.

Expanding the qth power of a sum and integrating

$$\Big\| \sum_{k_0<k<2k_0} (g_k * L_k) \Big\|_q^q \le$$

$$C \sum_{k_* < k_1 < \cdots < k_q < 2k_0} \int_G (g_{k_1} * L_{k_1}) \cdots (g_{k_q} * L_{k_q}) \tag{7.40}$$

$$+C\int_G \Big[\sum_{k_0<k<2k_0} (g_k * L_k) \Big]^{q-1} \tag{7.41}$$

where (7.41) is bounded by ρ^{k-1}.

Choosing M appropriately, we will achieve the estimate

$$\left\| [(g_{k_2} * L_{k_2}) \cdots (g_{k_q} * L_{k_q})] * (L_{k_1} - L_{k_0}) \right\|_{L^2(G)} \leq k_0^{-q} \tag{7.42}$$

whenever $k_0 < k_1 < k_2 < \cdots < k_q < 2k_0$.

Once (7.42) obtained, write

$$\left| \int_G (g_{k_1} * L_{k_1})(g_{k_2} * L_{k_2}) \cdots (g_{k_q} * L_{k_q}) - \int_G (g_{k_1} * L_{k_0})(g_{k_2} * L_{k_2}) \cdots (g_{k_q} * L_{k_q}) \right| < k_0^{-q}$$

and estimate (7.40) by

$$C + \sum_{k_0 < k_2 < \cdots < k_q < 2k_0} \int_G \left[\left(\sum_{k_0 < k < 2k_0} g_k \right) * L_{k_0} \right] (g_{k_2} * L_{k_2}) \cdots (g_{k_q} * L_{k_q}) \,. \tag{7.43}$$

Since the first factor in the integrand is 1-bounded, by (7.37), (7.43) turns out to be bounded by (7.41), thus by $C\rho^{k-1}$. Consequently, one gets $\rho^k < C + C\rho^{k-1}$, hence $\rho < C$, proving (7.39), thus (7.36) and (7.33). It remains to obtain (7.42).

Proof of (7.42). This is an L^2-problem and we use the Fourier transform method. Denote g_{k_r} by g_r and let $N_r = 2^{Mk_r}$.

We keep the notation $\mathcal{F}$ for the Fourier transform on $\mathbb{Z}$ and identify G with the integer interval $[0, J]$ endowed with normalized counting measure.

For each r, let s_r be increasing integers to be specified later. Define with the notations of the previous section and D to be specified

$$\Omega_r = \sum_{s \leq s_r} \sum_{\theta \in \mathcal{R}_{s_r}} S(\theta) w_{N_r}(\alpha - \theta) \zeta\big(10^s(\alpha - \theta)\big) \zeta\big(D^{-1} N_r^d (\alpha - \theta)\big) \,. \tag{7.44}$$

It follows from (6.15),(6.11),(6.13) that then for some $\delta > 0$

$$\left| \mathcal{F}[L_{k_r}](\alpha) - \Omega_r(\alpha) \right| < C(N_r^{-\delta} + 2^{-s_r \delta} + D^{-1/d}) \,. \tag{7.45}$$

Since from the definition (6.6) of $\mathcal{R}_s$, clearly

$$\left\| \mathcal{F}^{-1}(\Omega_r) \right\|_{L^1(\mathbb{Z})} \leq C 4^{s_r} \tag{7.46}$$

there is a uniform estimate

$$\left| \mathcal{F}^{-1} [\Omega_r \cdot \mathcal{F}(g_r)] \right| \leq C 4^{s_r} \,. \tag{7.47}$$

It also follows from (7.45) that

$$\|(g_r * L_{k_r}) - \mathcal{F}^{-1}[\Omega_r \mathcal{F}(g_r)]\|_{L^2(G)} \leq C(N_r^{-\delta} + 2^{-s_r\delta} + D^{-1/d}) . \tag{7.48}$$

Observe that the Fourier transform of the function $\mathcal{F}^{-1}[\Omega_2\mathcal{F}(g_2)] \cdots \mathcal{F}^{-1}[\Omega_q\mathcal{F}(g_q)]$ vanishes outside a DN_2^{-d} neighborhood Γ of $\{\frac{a}{b} \in \Pi \cap Q \mid b \leq 2^{qs_q}\}$.

Estimate the left member of (7.42) as

$$\|(g_2 * L_{k_2}) - \mathcal{F}^{-1}[\mathcal{F}(g_2)\Omega_2]\|_{L^2(G)} +$$

$$\|\mathcal{F}^{-1}[\mathcal{F}(g_2)\Omega_2]\|_\infty \|(g_3 * L_{k_3}) - \mathcal{F}^{-1}[\mathcal{F}(g_3)\Omega_3]\|_{L^2(G)} + \cdots +$$

$$\|\mathcal{F}^{-1}[\mathcal{F}(g_2)\Omega_2]\|_\infty \cdots \|\mathcal{F}^{-1}[\mathcal{F}(g_{q-1})\Omega_{q-1}]\|_\infty \|(g_q * L_{k_q}) - \mathcal{F}^{-1}[\mathcal{F}(g_q\Omega_q]\|_{L^2(G)} \tag{7.49}$$

$$+ \|\{\mathcal{F}^{-1}[\mathcal{F}(g_2)\Omega_2] \ldots . \mathcal{F}^{-1}[\mathcal{F}(g_q)\Omega_q]\} * (L_{k_1} - L_{k_0})\|_{L^2(G)} . \tag{7.50}$$

By (7.47),(7.48),(7.49) is bounded by

$$C\sum_{r=2}^{q} 4^{s_2+\cdots+s_{r-1}}(N_r^{-\delta} + 2^{-s_r\delta} + D^{-1/d}) . \tag{7.51}$$

Making the appropriate choice of the numbers $s_r \sim \log k_0$ permits then the estimation

$$(7.49) \leq \frac{1}{10} k_0^{-q} + k_0^C \cdot (N_2^{-\delta} + D^{-1/d}) . \tag{7.52}$$

By the remark on the support of the Fourier transform made above, (7.50) is clearly bounded by

$$\|\mathcal{F}^{-1}[\mathcal{F}(g_2)\Omega_2] \ldots . \mathcal{F}^{-1}[\mathcal{F}(g_q)\Omega_q]\|_{L^2(G)} \cdot \sup_{\alpha\in\Gamma} |\mathcal{F}(L_{k_1} - L_{k_0})(\alpha)| . \tag{7.53}$$

Again from (7.49), the first factor in (7.53) is bounded by $C4^{s_2+\cdots+s_q} < k_0^C$. By definition of Γ, (6.10),(6.12), one easily verifies that the second factor in (7.53) is at most

$$CD\left(\frac{N_1}{N_2}\right)^d + CN_0^{-1/2} \tag{7.54}$$

provided

$$D2^{qs_q} < N_0^{\delta'} \tag{7.55}$$

which is obviously satisfied for $D < 2^{\delta k_0}$.

By definitions of N_r and since $k_1 < k_2$,

$$(7.53) \leq k_0^C [D 2^{-Md} + 2^{-\frac{1}{2}Mk_0}] \,. \tag{7.56}$$

Collecting estimates (7.52), (7.56), the left member of (7.42) is bounded by

$$\frac{1}{10} k_0^{-q} + k_0^C [D^{-1/d} + D 2^{-Md} + 2^{-\delta k_0}] < k_0^{-q}$$

for a suitable choice of D, $\log D \sim \log k_0$ and $M \sim \log k_0$, (7.38). This completes the proof of (7.42) and hence Lemma 7.32).

The proof of (7.2) is mainly based on L^2-estimates, (7.29),(7,32) and interpolation.

Proof of (7.2). Denote $\| \quad \|_r$ the $\ell^r(\mathbb{Z})$-norm in what follows. Define for $s = 1, 2, \ldots$

$$Q_s = 2^s! \tag{7.57}$$

$$\mathcal{K}_s = \{k \in \mathbb{Z} \mid 4^s \leq k < 4^{s+1}\} \tag{7.58}$$

and with previous notations, let

$$\Omega_{k,s'} = \sum_{0 \leq a < Q_{s'}} S\left(\frac{a}{Q_{s'}}\right) w_{2^k}\left(\alpha - \frac{a}{Q_{s'}}\right) \zeta\left(Q_{s'}^2 \left(\alpha - \frac{a}{Q_{s'}}\right)\right) \tag{7.59}$$

for $s' \leq s$, $k \in \mathcal{K}_s$.

It follows from (6.15),(6.11),(6.13) that for $s' \leq s$, $k \in \mathcal{K}_s$

$$\left|\mathcal{F}[K_{2^k}](\alpha) - \Omega_{k,s'}(\alpha)\right| < 2^{-\delta' s'} \tag{7.60}$$

and by (7.13)

$$\left\|\mathcal{F}^{-1}[\Omega_{k,s'}]\right\|_1 < C \,. \tag{7.61}$$

Fix $1 < p_0 < p < 2$. It follows from (7.30) that

$$\left\| \sup_k \left|\mathcal{F}^{-1}[\Omega_{k,s'} \mathcal{F} f]\right| \right\|_{p_0} \leq C \|f\|_{p_0} \,. \tag{7.62}$$

For $k \in \mathcal{K}_s$, write

$$f * K_{2^k} = \mathcal{F}^{-1}[\Omega_{k,1} \mathcal{F} f] + \mathcal{F}^{-1}\big[(\Omega_{k,2} - \Omega_{k,1}) \mathcal{F} f\big] + \cdots + \mathcal{F}^{-1}\big[(\Omega_{k,s} - \Omega_{k,s-1}) \mathcal{F} f\big] +$$

$$\{(f * K_{2^k}) - \mathcal{F}^{-1}[\Omega_{k,s}\mathcal{F}f]\}$$

hence

$$\sup_k |f * K_{2^k}| \le \sum_{s'} \sup_{k\ge 4^{s'}} \left|\mathcal{F}^{-1}[(\Omega_{k,s'} - \Omega_{k,s'-1})\mathcal{F}f]\right| + \sum_s \sup_{k\in\mathcal{K}_s} \left|(f * K_{2^k}) - \mathcal{F}^{-1}[\Omega_{k,s}\mathcal{F}f]\right| . \tag{7.63}$$

By (7.62)

$$\left\| \sup_{k\ge 4^{s'}} \left|\mathcal{F}^{-1}[(\Omega_{k,s'} - \Omega_{k,s'-1})\mathcal{F}f]\right| \right\|_{p_0} < C\|f\|_{p_0} \tag{7.64}$$

while by (7.33) and (7.62) also

$$\left\| \sup_{k\in\mathcal{K}_s} \left|(f * K_{2^k}) - \mathcal{F}^{-1}[\Omega_{k,s}\mathcal{F}f]\right| \right\|_{p_0} \le Cs\|f\|_{p_0} . \tag{7.65}$$

Our purpose is to interpolate (7.64),(7.65) with better ℓ^2-estimates. Estimate using (6.15)

$$\left\| \sup_{k\ge 4^{s'}} \left|(f * K_{2^k}) - \mathcal{F}^{-1}[\Omega_{k,s'}\mathcal{F}f]\right| \right\|_2 \le$$

$$C\sum_{k\ge 4^{s'}} 2^{-k\delta_1} + \left\| \sup_{k\ge 4^{s'}} \left| \sum_{0\le r\le s'} \mathcal{F}^{-1}[\psi_{r,2^k}\cdot\mathcal{F}f] - \mathcal{F}^{-1}[\Omega_{k,s'}\cdot\mathcal{F}f] \right| \right\|_2 +$$

$$\sum_{r>s'} \left\| \sup_k \left|\mathcal{F}^{-1}[\psi_{r,2^k}\cdot\mathcal{F}f]\right| \right\|_2 \tag{7.66}$$

where $\psi_{r,N}$ is given by (6.7).

By (6.28), the last term of (7.66) is bounded by $C\cdot 2^{-s'\delta'}\|f\|_2$.

Write

$$\Omega_{k,s'} - \sum_{r\le s'} \psi_{r,2^k} = \sum_{r\le s'} \sum_{\theta\in\mathcal{R}_r} S(\theta) w_{2^k}(\alpha-\theta)\left[\zeta\left(Q_{s'}^2(\alpha-\theta)\right) - \zeta\left(10^r(\alpha-\theta)\right)\right] \tag{7.67}$$

$$+ \sum_{\substack{q|Q_{s'} \\ q\ge 2^{s'}+1}} \sum_{\substack{1\le a\le q \\ (a,q)=1}} S\left(\frac{a}{q}\right) w_{2^k}\left(\alpha-\frac{a}{q}\right)\zeta\left(Q_{s'}^2\left(\alpha-\frac{a}{q}\right)\right) . \tag{7.68}$$

There is a uniform estimate on (7.67) for $k\ge 4^{s'}$ by

$$C\cdot 4^{s'}\cdot \sup_{|\beta|>Q_{s'}^{-2}} \left|w_{2^k}(\beta)\right| < C2^{-k/2} \tag{7.69}$$

by (6.13) and (7.57). Thus (7.67) contributes to the maximal function for at most $C2^{-s'}$.

As done in section 6 to prove (6.28),(6.29), one estimates the maximal function contribution of (7.68) (in ℓ^2) by $C \cdot 2^{-\delta' s'}$.

Collecting estimates yields the bound $C \cdot 2^{-\delta' s'}$ on (7.66). Hence also, by subtraction

$$\left\| \sup_{k \geq 4^{s'}} \left| \mathcal{F}^{-1}\left[(\Omega_{k,s'} - \Omega_{k,s'-1})\mathcal{F}f\right]\right| \right\|_2 \leq C \cdot 2^{-\delta' s'} \|f\|_2 \tag{7.70}$$

while

$$\left\| \sup_{k \in \mathcal{K}_s} \left|(f * K_{2^k}) - \mathcal{F}^{-1}[\Omega_{k,s} \cdot \mathcal{F}f]\right| \right\|_2 \leq C \cdot 2^{-\delta' s} \|f\|_2 \,. \tag{7.71}$$

Interpolating (7.64),(7.70) at $p_0 < p < 2$ yields the corresponding ℓ^p-inequality with constant $C \cdot 2^{-\delta_p s'}$. Similarly when interpolating (7.65),(7.71), an ℓ^p-estimate $C \cdot 2^{-\delta_p s}$ is found. Here $\delta_p > 0$ depends on $p > 1$. Substitution of these bounds in (7.63) yields

$$\left\| \sup_k |f * K_{2^k}| \right\|_p \leq C \sum_{s'} 2^{-\delta_p s'} + C \sum_s 2^{-\delta_p s} < C$$

completing the proof of (7.2).

8. Integer Parts of Polynomial Sequences

Consider a polynomial with real coefficients ($d \geq 1$)

$$p(x) = b_0 + b_1 x + \cdots + b_d x^d\,, \qquad b_d > 0 \tag{8.1}$$

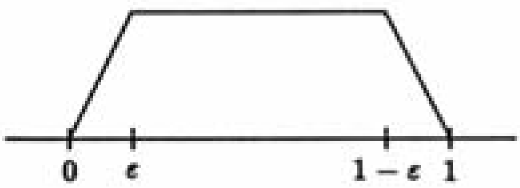
and for a given DS $(\Omega, \mathcal{B}, \mu, T)$ the averages

$$A_N f = \frac{1}{N} \sum_{n=1}^{N} T^{[p(n)]} f \tag{8.2}$$

where $[y]$ stands for the integer part of y. Here we let f be of class $L^\infty(\Omega, \mu)$. In proving the a.s. convergence of (8.2), one may assume at least one of the coefficients $b_1, \dots, b_d$ irrational. Otherwise, if $b_j = \frac{a_j}{q}$ $(1 \leq j \leq q)$, write $n = mq + r$ $(0 \leq r < q)$ and $A_N f = \frac{1}{q} \sum_{0 \leq r < q} \frac{1}{Nq^{-1}} \sum_{m \leq Nq^{-1}} T^{p_1(m)}(T^{[p(r)]} f)$ where $p_1(m) = p(mq + r) - p(r)$ has integer coefficients. The a.s. convergence of the $A_N f$ is thus implied by Theorem 1 of this paper.

Assuming thus $b_1, \ldots, b_d$ not all rational, the sequence $\{p(n) - [p(n)] \mid n = 1, 2, \ldots\}$ is uniformly distributed in $[0,1]$.

Fix $\varepsilon > 0$ and consider the function $\tau = \tau_\varepsilon$ on $\mathbb{R}$



Denote

$$\widetilde{A}_N f = \frac{1}{N} \sum_{n=1}^{N} \sum_{m \in \mathbb{Z}} \tau\big(p(n) - m\big) T^m f \ . \tag{8.3}$$

Clearly, invoking the uniform distribution property, there is the pointwise inequality

$$|A_N f - \widetilde{A}_N f| \le \frac{\|f\|_\infty}{N} \#\{1 \le n \le N \mid \operatorname{dist}\big(p(n), \mathbb{Z}\big) < \varepsilon\} \le 3\varepsilon \|f\|_\infty \tag{8.4}$$

for N large enough.

Thus, it suffices to show the a.s. convergence of (8.3) (for a fixed $\varepsilon > 0$, assuming $f \in L^2(\mu)$ (the hypothesis $f \in L^\infty$ is only of relevance when replacing $A_N f$ by $\widetilde{A}_N f$).

The proof of this follows the same method as used in section 6.

The relevant exponential sums are given by

$$\mathcal{F}[K_N](\alpha) = \frac{1}{N} \sum_{n \le N} \sum_{m \in \mathbb{Z}} \tau\big(p(n) - m\big) e^{-2\pi i m \alpha} = \sum_{k \in \mathbb{Z}} \widehat{\tau}(k - \alpha) \left\{ \frac{1}{N} \sum_{n \le N} e^{2\pi i (k-\alpha) p(n)} \right\} \tag{8.5}$$

where

$$K_N = \frac{1}{N} \sum_{n \le N} \sum_{m \in \mathbb{Z}} \tau\big(p(n) - m\big) \delta_{\{m\}} \ . \tag{8.6}$$

Let $\varphi_N(\overline{\alpha})$ be given by (5.1)

$$\varphi_N(\overline{\alpha}) = \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i (\alpha_1 n + \alpha_2 n^2 + \cdots + \alpha_d n^d)} \ .$$

Then

$$\mathcal{F}[K_N](\alpha) = \sum_{k \in \mathbb{Z}} \widehat{\tau}(k - \alpha) e^{2\pi i (k-\alpha) b_0} \varphi_N\big(b_1(k - \alpha), \ldots, b_d(k - \alpha)\big) \ . \tag{8.7}$$

Observe also the decay property

$$\left|\widehat{\tau}(\lambda)\right| < \frac{C}{1+\varepsilon^2\lambda^2} . \tag{8.8}$$

Define for $k \in \mathbb{Z}$

$$\mathcal{R}_{s,k} = \Big\{\theta \in [0,1] \mid b_j(k-\theta) \equiv \frac{a_j}{q} (\mathrm{mod}\ 1) \quad \text{where} \quad (q, a_1, \ldots, a_d) = 1 \quad \text{and} \quad 2^s \leq q \leq 2^{s+1}\Big\} \tag{8.9}$$

and for $\theta \in \mathcal{R}_{s,k}$, let with notation (5.10)

$$S(\theta) = S(q, a_1, \ldots, a_d) . \tag{8.10}$$

Define also, with notation (5.11)

$$w_N(\beta) = v_N(-b_1\beta, \ldots, -b_d\beta) . \tag{8.11}$$

Denote further

$$\psi_{s,k,N}(\alpha) = \sum_{\theta \in \mathcal{R}_{s,k}} S(\theta) w_N(\alpha-\theta)\zeta\left(10^s b_d(\alpha-\theta)\right) \tag{8.12}$$

$$\psi_{s,N}(\alpha) = \sum_{k\in\mathbb{Z}} \widehat{\tau}(k-\alpha)e^{2\pi i(k-\alpha)b_0}\psi_{s,k,N}(\alpha) . \tag{8.13}$$

Notice that different elements of $\mathcal{R}_{s,k}$ are at least $4^{-s-1}b_d^{-1}$-separated and the summands in (8.12) are thus supported by disjoint arcs in Π (not necessarily centered around rational points).

Using (8.8), one then has the analogue of (6.14)

$$\left|\mathcal{F}[K_N] - \sum_{s\geq 0}\psi_{s,N}\right| < C \cdot N^{-\delta_1} . \tag{8.14}$$

We leave the reader to make the verification.

Proceeding similarly to in the proof of (6.28) in section 6, one gets

$$\Big\| \sup_{N\in Z_1} |\mathcal{F}^{-1}[\psi_{s,k,N}\mathcal{F}f]| \Big\|_2 \leq C(\log|\mathcal{R}_{s,k}|)^2 2^{-\delta' s}\|f\|_2 \sim Cs^2 2^{-\delta' s}\|f\|_2 . \tag{8.15}$$

Hence, by (8.8)

$$\Big\| \sup_{N\in Z_1} |\mathcal{F}^{-1}[(\sum_{s\geq 0}\psi_{s,N})\mathcal{F}f]| \Big\|_2 \leq C\sum_{s\geq 0}\sum_{k\in\mathbb{Z}} \frac{1}{1+\varepsilon^2k^2} s^2 2^{-\delta' s}\|f\|_2 \leq C\|f\|_2$$

yielding the maximal inequality

$$\left\| \sup |\widetilde{A}_N f| \right\|_2 \leq C_\varepsilon \|f\|_2 . \tag{8.16}$$

With this information, the proof of a maximal variational inequality (6.31) is essentially identical to the argument given in section 6 and will therefore not be elaborated here.

This completes the proof of Theorem 2 (for L^∞-functions).

Remark. The main additional item in proving the L^r-verion , $r > 1$, of the previous result for sets $\Lambda = \{[p(n)]\}$ is a more detailed analysis of the approximation of $A_N f$ by $\widetilde{A}_N f$, based on rational approximation of the coefficients $b_1, \dots, b_d$ of $p(x)$.

9. Further Comments and Remarks on Almost Sure Convergence

(1) In [B₃], the author considered the sequence Λ of prime numbers and proved that the averages

$$A_N f = \frac{1}{|\Lambda_N|} \sum_{n \in \Lambda_N} T^n f \qquad \Lambda_N = \{\text{primes} \leq N\} \tag{9.1}$$

converge a.s. for f a function of class L^2. Letting

$$K_N = \frac{1}{N} \sum_{\substack{p \leq N \\ p \text{ prime}}} (\log p) \delta_{\{p\}} , \tag{9.2}$$

it is well known that

$$\mathcal{F}[K_N](\alpha) = \frac{\mu(q)}{\phi(q)} \frac{1}{N} \left(\sum_{k=1}^{N} e^{2\pi i k (\alpha - \frac{a}{q})} \right) + O(e^{-c\sqrt{\log N}}) \tag{9.3}$$

for $\left|\alpha - \frac{a}{q}\right| < (\log N)^c N^{-1}$, $1 \leq a \leq q$, $(a, q) = 1$ and $q < (\log N)^c$. Here μ denotes the Moebius function and $\phi(q)$ the number of Dirichlet characters to the modulus q.

Prior to this work, it has been shown in [ ] that $A_N f$ given by (9.1) converge a.s. for a function of class L^r, whenever $r > 1$. The argument is based on special properties of the expression in the right member of (9.3) and does not seem adaptable for the set of the squares for instance. The reader will easily check that the method described in section 7 of this paper applies equally well to the primes.

(2) Both Theorem's 1 and 2 of this paper generalize to positive (not necessarily invertible) isometries on L^r, $r > 1$. It was indeed pointed out in section 2 that this situation reduces also to the shift. Thus in particular, there is the following generalization of the Riesz-Raikov result (cf. [Ra], [Ri]), where $p(x) = x$:

Let $p(n)$ be a polynomial mapping positive integers to positive integers and f a function on the circle Π of class L^r, $r > 1$. Then $\frac{1}{N} \sum_{1 \leq n \leq N} f(2^{p(n)} x)$ converge a.s. to $\int_\Pi f$.

Recall in this context Marstrand's counterexample to the Khinchine conjecture [Ma], according to which there are bounded measurable functions f on Π such that $\frac{1}{N} \sum_{1 \leq n \leq N} f(nx)$ does not converge a.s.

(3) Let T_n $(n = 1, 2, \ldots)$ be a sequence of *commuting* positive isometries on $L^2(\mu)$ and define $A_N f = \frac{1}{N} \sum_{n \leq N} T_n f$. It follows from the results of [B_5] that the following property is a necessary condition for a.s. convergence of $A_N f$, $N \to \infty$, even restricting to functions $f \in L^\infty(\mu)$:

For each $\delta > 0$, there is a bound $C(\delta) < \infty$ on the δ-metrical entropy number (in the sense of section 3)

$$M_\delta\big(\{A_N f \mid N = 1, 2, \ldots\}\big) < C(\delta) \tag{9.4}$$

of the subset $\{A_N f\}$ of $L^2(\mu)$. This bound (9.4) has to be uniform when f ranges in the unit ball of $L^2(\mu)$.

As pointed out through several applications (including Marstrand's example mentioned above) in [B_5], the previous criterion is often effective in disproving the a.s. convergence of such averages.

APPENDIX

Return Times of Dynamical Systems

Let $(X, \mathcal{B}, \mu, T)$ be an ergodic system and $A \in \mathcal{B}$ of positive measure $\mu(A) > 0$. For $x \in X$, consider the return time sequence $\Lambda_x = \{n \in \mathbb{Z}_+ \mid T^n x \in A\}$. By Birkhoff's pointwise ergodic theorem, the sequence Λ_x has positive density, for μ-almost all $x \in X$. This fact refines the classical Poincaré recurrence principle (cf. [Fu]). An even stronger statement is given by the Wiener-Wintner theorem: there is a set X' of X of full measure such that the sums

$$\frac{1}{N} \sum_{1 \leq n \leq N} \chi_A(T^n x) z^n$$

converge, for all z in the unit circle $\mathbb{G} = \{z \in \mathbb{C} \mid |z| = 1\}$ and $x \in X'$. Thus from general theory of unitary operators, this fact may be reinterpreted by saying that almost all sequences Λ_x satisfy the L^2, hence the mean ergodic theorem. Our purpose here is to prove the following fact, answering a question open for some time.

Theorem. *With the notations above, Λ_x satisfies almost surely the pointwise ergodic theorem, i.e., the averages*

$$\frac{1}{N} \sum_{\substack{1 \leq n \leq N \\ n \in \Lambda_x}} S^n g$$

converge almost surely, for any measure preserving system $(Y, \mathcal{D}, \nu, S)$ and $g \in L^1(Y)$.

The argument given next actually yields a more precise condition on the point x.

Let $f \in L^\infty(X)$ be obtained by projecting χ_A on the orthogonal complement of the eigenfunctions of T. It clearly suffices to prove that for almost all $x \in X$, $\{f(T^n x)\}$ is a "summing sequence". i.e.,

$$\frac{1}{N} \sum_{1 \leq n \leq N} f(T^n x) g(S^n y) \to 0 \quad \text{a.e.} \quad y \in Y \tag{$*$}$$

for any measure preserving system $(Y, \mathcal{D}, \nu, S)$ and $g \in L^\infty(Y)$. (The contribution of eigenfunctions is taken care of by Birkhoff's theorem.)

Observe the equivalence of the following statements:

i) f has continuous spectral measure

ii) $\langle T^n f, f\rangle = \widehat{\sigma}_f(n)$, σ a continuous measure

iii) $\frac{1}{N}\sum_1^N f(T^n x) f(T^n \xi) \to 0$ a.e. (x,ξ) as $N \to \infty$.

Proof of ii) $\Rightarrow$ iii). Write $F = \lim \frac{1}{N}\sum_1^N f(T^n x) f(T^n \xi)$, exists by the ergodic theorem, $\|F\|^2 = \lim \frac{1}{N^2} \sum_{m,m=1}^{N} \left(\widehat{\sigma}_f(n-m)\right)^2 = 0$.

Proposition. *Assume x generic for f and $\frac{1}{N}\sum f(T^n x) f(T^n \xi) \to 0$, ξ a.e. (!). Then $\{f(T^n x)\}$ is a summing sequence.*

Proof: I) Assume that for some $(Y, \mathcal{D}, \nu, S)$ and $g \in L^\infty(Y)$ there is a set B^* of positive measure for which the limsup of $(*)$ is positive. Then there exists $a > 0$, $B \subset B^*$, $\nu(B) > 0$ and a sequence of intervals $R_j = (L_j, M_j)$ (called "ranges") such that for every $y \in B$ and every j there exists $n_j \in \mathbf{R}_j$ $(n_j = n_j(y))$ such that

$$\sum_{n=1}^{n_j} f(T^n x) g(S^n y) > a n_j \,. \tag{$**$}$$

II) Given $\delta > 0$, there exists $K = K(N, \delta)$ such that

$$\nu\Big(\bigcup_1^K S^j B\Big) > 1 - \delta \,.$$

III) Write $\varphi = 1_{\bigcup_1^K S^j B}$. If M_0 is large enough, and we denote by G the set $G = \{y : |\frac{1}{n}\sum_1^n \varphi(S^j y) - 1| < 2\delta \text{ for all } n > M_0\}$, then $\nu(G) > 1 - \delta$.

IV) For notational convenience we assume that f has a finite range, and denote by B_n the set of all n-blocks for f, i.e., the (set of) words $w_k^{(n)} = (f(T^{k+1}x), \ldots, f(T^{k+n}x))$; $w_k^{(n)}$ appears with density $p(w_k^{(n)})$.

Given $\delta > 0$ (δ can be chosen once and for all as a function of a and $\nu(B)$ in I) let N_δ be such that for each set $A_\delta \subset X$, $\mu(A_\delta) > 1 - \delta$, $\left|\frac{1}{N}\sum f(T^n x) f(T^n \xi)\right| < \delta$ for all $\xi \in A_\delta$ and $N > N_\delta$ (cf. assumption (!)).

Given a range (L, M) with $L > N_\delta$, set $N = N(M)$ so that in any interval on the integers of length $\geq N$ the statistics of the n-blocks (for f) with $n \leq M$ is correct. Denote by B_n^* the n-blocks that have the form $\left(f(T\xi), \ldots, f(T^n \xi)\right)$ with $\xi \in A_\delta$ (we are interested in

$n \in (L, M)$). For $L < n < M$ the total probability (=density) of the blocks in B_n^* exceeds $1-\delta$ (in any interval of length $\geq N(M)$). Notice also that heads of M-blocks which are in B_M^* are in the appropriate B_n^*.

V) A sequence of ranges $\{(L_j, M_j)\}$ is *properly spaced* if $L_{j+1} > N(M_j)$. (We also assume $L_1 > N_\delta$. Another assumption on L_1 is that it is $> M_0$ (recall the definition of G in III) and assume that K (II)) is $\ll L_1$.) *Going back to* I), we select a properly spaced sequence of ranges $\{(L_j, M_j)\}_{j=1}^{J}$ (J depending on a) and N large enough so that $N \gg N(M_J)$.

Recall B from I) and G from III).

For any $y \in B \cap G$ we define a sequence $\{c_n(y)\}_{n=1}^{N}$ which is a sum of J sequences (layers) $\{c_n^j(y)\}$ having the following properties:

(α) For all j, n, y, $c_n^j(y)$ is in the range of f (in particular unif. bounded)

(β) For $j_1 \neq j_2$, $\left|\frac{1}{N}\sum\limits_{n=1}^{N} c_n^{j_1}(y)c_n^{j_2}(y)\right| < \delta$

(γ) $\frac{1}{N}\sum\limits_{n=1}^{N} c_n^j(y)g(S^n y) > a - \delta$, $j = 1, \ldots, J$

(α) and (β) together imply $\left[\frac{1}{N}\sum\left(c_n(y)\right)^2\right]^{1/2} = O(\sqrt{J} + \delta J)$. ($\gamma$) implies $\frac{1}{N}\sum\limits_{1}^{N} c_n(y)g(S^n y) > J(a-\delta)$. Contradiction.

We construct $\{c_n^j\}$ in reverse order on j. $c_n^J(y)$ is defined as follows: $l_1(y)$ is the first index $k > 0$ such that $S^k y \in B$; on the interval $\left(l_1(y), l_1(y) + n_J(S^{l_1(y)}y)\right)$ we set

$$c_n^J(y) = f(T^{n-l_1(y)}x) \; .$$

$l_2(y)$ is the index of the first point in the S-orbit of y, after $l_1(y) + n_J(S^{l_1(y)}y)$, which is in B, and on the interval $l_2(y), l_2(y) + n_J(S^{l_2(y)}y)$ we copy again $\{f(T^k x)\}_{k=1}^{n_J(S^{l_2(y)}y)}$ etc. The intervals on which we copy those starting n_J blocks fill most of $[1, N]$. We refer to these on the basic intervals of the J-layer. Outside of these, set $c_n^J(y)$ arbitrarily.

We now define $c_n^{J-1}(y)$ in a similar manner within every basic interval of the J-layer, with the additional restriction on the starting place of the new basic blocks that (in addition to the fact that the corresponding point in the orbit of y is in B) the matching piece of the basic J-layer block is in B^*, i.e., more or less orthogonal to the "new" basic block; see IV). Since the "orthogonal" blocks have density $> 1 - \delta$, the new basic blocks cover more than $1 - 3\delta$ of $[1, N]$. We continue with $c_n^{J-2}(y), \ldots, c_n^1(y)$, working each time within the basic blocks of the previous level and introducing blocks which are "orthogonal" to all previous levels.

Remarks.

(i) The condition that $\frac{1}{N}\sum_{1}^{N} f(T^n x) f(T^n \xi) \to 0$ a.e. in ξ(!) is a special case of $(*)$ and hence necessary. One can construct examples showing that it is not a consequence of the genericity of x.

(ii) One may construct a sequence $\Lambda = \{k_n\}$, $k_n = o(n)$, and a weakly mixing system (Y, S) such that $\frac{1}{N}\sum_{1}^{N} g(S^{k_n} y)$ does not average a.e., for some $g \in L^\infty(Y)$. (This question was considered in [Fu], p. 96.)

References

[B_1] J. BOURGAIN: On the maximal ergodic theorem for certain subsets of the integers, Israel J. Math. Vol 61, N1, 1988, 39-72.

[B_2] J. BOURGAIN: On the pointwise ergodic theorem on L^p for arithmetic sets, idem, 73-84.

[B_3] J. BOURGAIN: An approach to pointwise ergodic theorems, Springer LNM, Vol 1317, 204-223.

[B_4] J. BOURGAIN: Return time sequences of dynamical systems, preprint IHES 3/1988.

[B_5] J. BOURGAIN: Almost sure convergence and bounded entropy, Israel J. Math. Vol 63, N1, 1988, 79-97.

[B_6] J. BOURGAIN: Temps de retour pour les systèmes dynamiques, CR A Sc Paris, t. 306, Ser I, 1988, 483-485.

[Fu] H. FURSTENBERG: Recurrence in Ergodic theory and Combinatorial Number theory, Princeton UP, 1987.

[Ga] A. GARSIA: Martingale Inequalities, Benjamin 1970.

[K-W] Y. KATZNELSON, B. WEISS: A simple proof of some ergodic theorems, Israel J. Math., Vol 42, N4, 1982.

[Lé] D. LEPINGLE: La variation d'ordre p des semi-martingales, Z. Wahrscheinlichkeitstheorie verw. Gebiete 36, 1976, 295-316.

[Ma] J.M. MARSTRAND: On Khinchine's conjecture about strong uniform distribution, Proc. London Math. Soc. 21, 1970, 540-556.

[Nev] J. NEVEU: Martingales à temps discret, Masson 1972.

[P-X] G. PISIER, Q. XU: The strong p-th variation of martingales and orthogonal series, preprint Paris VI.

[Ra] D.A. RAIKOV: On some arithmetical properties of summable functions, Mat. Sb. 1(43), 1936, 377-384.

[Ri] F. RIESZ: Sur la Thèorie ergodique, Comment. Math. Helv. 17, 1945, 221-239.

[S] R. SALEM: Collected works.

[St] E. STEIN: On limits of sequences of operators, Ann. Math. 74, 1961, 140-170.

[Vaug] R.C. VAUGHAN: The Hardy-Littlewood method, Cambridge Tracts, 70 (1981).

[Vin] VINOGRADOV: The method of trigonometrical sums in the theory of numbers, Interscience N-Y, 1954.

[W] B. WEISS: private communications.