Jean Bourgain

# Pointwise ergodic theorems for arithmetic sets

*Publications mathématiques de l'I.H.É.S.*, tome 69 (1989), p. 5-41
<http://www.numdam.org/item?id=PMIHES_1989__69__5_0>

# POINTWISE ERGODIC THEOREMS FOR ARITHMETIC SETS

*by* Jean BOURGAIN

## With an appendix on return-time sequences

*jointly with* Harry FURSTENBERG, Yitzhak KATZNELSON
*and* Donald S. ORNSTEIN

### 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) $(\Omega, \mathscr{B}, \mu, T)$ and a polynominal $p(x)$ with integer coefficients, then the ergodic means

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

converge almost surely for $N \to \infty$, assuming $f$ a function of class $L^2(\Omega, \mu)$. Here and in the sequel, one denotes by $\mu$ a probability measure and by $T$ 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 this line of investigation here.

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 **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.

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 (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, $p > 1$. 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 allows one to give 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], but it is different and provides 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 **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 for 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 and $L^p$-function on the circle $\pi = \mathbf{R}/\mathbf{Z}$ and $\alpha \in \mathbf{R}\backslash\mathbf{Q}$ an irrational number. Then the averages

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

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 forward 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 ($p > 1$)

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

for $f \geqslant 0$.

Next, we describe the organisation 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 considerer the variation spaces $v_p$, where $\| x \|_{v_p}$ is defined as

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

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

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

where

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

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 [B₁], based on the new ingredient obtained in section 4. In this proof, we no longer need the *a priori* 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* $(\Omega, \mathscr{B}, \mu, T)$ *by a dynamical system and* $p(x)$ *a polynomial with integer coefficients. Then there is the maximal inequality*

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

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

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

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

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

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

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.

Section 9 contains various comments and remarks on almost sure convergence in general, related to $[B_5]$.

CONTENTS



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 $(\Omega, \mathscr{B}, \mu, T)$ 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 $\mathbf{Z}$.

### 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 indeed ensures the existence of a positive Radon measure $\nu$ on the circle T, such that

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

implying that the map $L^2(\Pi, \nu) \to L^2(\Omega, \mu)$ mapping the $n$th 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 $\Lambda$ of the positive integers, the pointwise convergence on the unit circle of the sequence of polynomials

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

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

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

and

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

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 $\Lambda \subset \mathbf{Z}_+$ is " ergodic " provided $p_N(z) \to 0$ for $z \in \mathbf{T} - \{ 1 \}$. The property implies mean convergence of $A_N f$ to $\int_\Omega f\, d\mu$, assuming T ergodic (this is the case for $\Lambda = \mathbf{Z}_+$ but not if $\Lambda = \{ n^2 \mid n = 1, 2, \dots \}$ for instance).

B. Weiss [W] observed that sequences $\Lambda$ 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 $(1 < p \leq \infty)$

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

and the weak-type inequality

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

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

Let us give a simple proof (2.3), (2.4) by deriving them from the shift model (**Z**, S). In the case of the shift, the weak-type property (2.4) easily follows from geometric covering properties of integer-intervals, in the some way 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 $(\Omega, \mu, T)$. Of course it suffices to prove inequalities (2.3), (2.4) (with fixed constants) for a " restricted " maximal function

$$\bar{f} = \sup_{1 \leqslant N \leqslant \bar{N}} A_N f \quad (f \leqslant 0) \tag{2.5}$$

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

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

For the function $f$, define the function $\varphi$ on **Z** as follows

$$\begin{cases} \varphi(j) = f(T^j x) & \text{if } 0 \leqslant j \leqslant J \\ \phantom{\varphi(j)} = 0 & \text{otherwise.} \end{cases} \tag{2.6}$$

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

$$\bar{\varphi}(j) = \bar{f}(T^j x) \quad \text{for } 0 \leqslant j < J - \bar{N}. \tag{2.7}$$

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

$$\sum_{0 \leqslant j < J - \bar{N}} | \bar{f}(T^j x) |^p \leqslant C(p)^p \sum_{0 \leqslant j \leqslant J} | f(T^j x) |^p. \tag{2.8}$$

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

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

and since T is measure-preserving, one gets

$$\| \bar{f} \|_p \leqslant C(p) \frac{J}{J - \bar{N}} \| f \|_p,$$

hence

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

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

$$\| \bar{\varphi} \|_{\ell^{1,\infty}(\mathbf{Z})} \leqslant C \, \| \varphi \|_{\ell^1(\mathbf{Z})}$$

and, by (2.7),

$$\lambda \, | \{ 0 \leqslant j < J - \bar{N} \mid \bar{f}(T^j x) > \lambda \} | \leqslant C \sum_{0 \leqslant j \leqslant J} f(T^j x),$$

hence

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

Integrating again, we have

$$\lambda \mu(\Omega_\lambda) \leqslant C \frac{J}{J - \bar{N}} \| f \|_1,$$

from which (2.4) easily follows.

At present, the covering argument leading to weak-type inequalities does not seem to be available when dealing with particular subsets of **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 an approach similar to that 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 by F the $L^2$-limit of $(A_N f)$ and, for given $\varepsilon > 0$, let $N_\varepsilon$ satisfy

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

By the invariance of the limit (since the ergodic means relates to the full set of positive integers) and the maximal inequality, we have

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

Since

$$| A_N(A_{N_\varepsilon} f) - A_N f | \leqslant 2 \frac{N_\varepsilon}{N} \| f \|_\infty,$$

it follows from (2.10) that

$$\| \overline{\lim} | F - A_N f | \|_2 < C\varepsilon, \quad \text{hence } \overline{\lim_N} | F - A_N f | = 0 \text{ almost surely.}$$

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

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

corresponding to a subset $\Lambda$ of $\mathbf{Z}_+$.

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

$$\|f^*\|_p \leqslant C \|f\|_p; \qquad f^* = \sup |A_N f|.$$

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

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 $L^\infty(\Omega, \mu)$, $|f| \leqslant 1$. For $\varepsilon > 0$, consider the subset

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

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

$$|A_N f - A_{N'} f| \leqslant 2\varepsilon.$$

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

$$\|\mathcal{M}_j f\|_2 > \varepsilon \quad \text{where } \mathcal{M}_j f = \sup_{\substack{N_j \leqslant N \leqslant N_{j+1} \\ N \in Z_\varepsilon}} |A_N f - A_{N_j} f|. \tag{2.13}$$

In fact, a more quantitative statement is shown, namely

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

for J large (depending on $\varepsilon$ 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 shifs (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 $v_s$ $(1 \leqslant s \leqslant \infty)$ for scalar sequences $\bar{x} = (x_n)_{n=1,2,\dots}$

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

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

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

These spaces $v_s$ are frequently used in probability theory when studying 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* $\mathbf{E}_n$ $(n = 1, 2, \ldots)$ *be the sequence of expectation operators with respect to an increasing sequence of* $\sigma$*-algebras on a probability space and* $f_n = \mathbf{E}_n f$ *an associated scalar martingale. Then, for* $s > 2$, *we have the inequality*

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

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

This result may be seen as the quantitative form of the martingale convergence theorem. The 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 by $N_\lambda(\omega)$ the number of $\lambda$-jumps in the sequence $\{f_n(\omega)\}$, where $f_n$ is defined as above. One has the following inequality for $1 < r < \infty$:

$$\| \lambda (N_\lambda)^{1/2} \|_r \leqslant c_r \|f\|_r \quad \text{for all } \lambda > 0. \tag{3.5}$$

This 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):

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

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

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

Hence, since $p \leqslant s$,

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

applying (3.5) with $r = 2p/s$ (which implies $6/5 \leqslant r < 2$ in view of the hypotheses made on $p, s$) and $\lambda = 2^{-k}$. Since $\|f\|_r^r = \varepsilon$, (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 a real analysis version of (3.4) from Lemma 3.3. For a function $f$ on $\mathbf{R}$, set $f_t(x) = \frac{1}{t} f\left(\frac{x}{t}\right)$. Denote also by

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

the Fourier transform of $f$. Thus

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

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

$$\|\{f * \chi_t \mid t > 0\}\|_{L^2_{v_s}(\mathbf{R})} \leqslant c(s-2)^{-1} \|f\|_2, \tag{3.12}$$

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

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

Denote by $(P_t)_{t>0}$ the Poisson semi-group on $\mathbf{R}$. Thus if $P_t f = f * P_t$, one has $\hat{P}_t(\lambda) = e^{-t|\lambda|}$. 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

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

*Proof of Lemma* **3.11.** — By (3.13), (3.12) will be a consequence of the following inequality

$$\|\{f * K_t \mid t > 0\}\|_{L^2_{v_s}} \leqslant c \|f\|_2, \tag{3.14}$$

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

$$|\lambda| . |(\hat{K})'(\lambda)| < c \quad \text{and} \quad |\hat{K}(\lambda)| \leqslant c \min(|\lambda|, |\lambda|^{-1}). \tag{3.15}$$

We clearly have the pointwise estimate

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

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

$$\left[\sum_{k \in \mathbf{Z}} \int_{-\infty}^{\infty} |\hat{f}(\lambda)|^2 \, |\hat{K}(2^k \lambda)|^2 \, d\lambda\right]^{1/2} \leqslant c . \left[\int |\hat{f}(\lambda)|^2 \, d\lambda\right] = c \|f\|_2, \tag{3.17}$$

invoking also (3.5).

Next, we estimate the contribution of the second term

$$\Big\{\sum_{k \in \mathbf{Z}} \|\{f * K_t \mid 2^k \leqslant t \leqslant 2^{k+1}\}\|_{L^2_{v_s}}^2\Big\}^{1/2}. \tag{3.18}$$

Let $0 < \eta < 1$ be a function supported by $\left[\frac{1}{2}, 2\right] \cup \left[-2, -\frac{1}{2}\right]$, $|\eta'| < C$, such that $\sum_{\alpha \in \mathbf{Z}} \eta(2^\alpha \lambda) = 1$.

Defining $K_\alpha$ by $\hat{K}_\alpha(\lambda) = \hat{K}(\lambda) \, \eta(2^\alpha \lambda)$, one has that $K = \sum_\alpha K_\alpha$, and (3.18) may be estimated by the triangle inequality as

$$\sum_{\alpha \in \mathbf{Z}} \Big\{\sum_{k \in \mathbf{Z}} \|\{f * (K_\alpha)_t \mid 2^k \leqslant t \leqslant 2^{k+1}\}\|_{L^2_{v_s}}^2\Big\}^{1/2}. \tag{3.19}$$

From (3.15),

$$|\lambda| \, |(\hat{K}_\alpha)'(\lambda)| < c \quad \text{and} \quad |\hat{K}_\alpha(\lambda)| < c 2^{-|\alpha|}. \tag{3.20}$$

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

$$\| \{ f * (K_\alpha)_t \mid 2^k \leqslant t \leqslant 2^{k+1} \} \|_{v_2} \leqslant \tag{3.21}$$

$$\left[ \sum_{\ell=1}^{N} | f * (K_\alpha)_{u_\ell} |^2 \right]^{1/2} +$$

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

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

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

by the definition of $K_\alpha$ 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

$$\left[ \sum_{\ell=1}^{N} (u_{\ell+1} - u_\ell) \int_{u_\ell}^{u_{\ell+1}} \left[ \int_{-\infty}^{\infty} |\hat{f}(\lambda)|^2 \, |\lambda|^2 \, |(\hat{K}_\alpha)'(t\lambda)|^2 \, d\lambda \right] dt \right]^{1/2} \leqslant \tag{3.24}$$

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

$$= C N_\alpha^{-1/2} \left[ \int_{|\lambda| \sim 2^{-k-\alpha}} |\hat{f}(\lambda)|^2 \, d\lambda \right]^{1/2}.$$

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

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

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

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

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

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

in Birkhoff's theorem.

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

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

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 $(\mathbf{Z}, S)$, 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

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

for a smooth function $\varphi$ on $[0, \infty]$, vanishing at $\infty$, the following lemma is a consequence of (3.12) and the convexity.

*Lemma* **3.28.** — *Let $\varphi$ be a differentiable function on* **R**, *vanishing at* $\infty$. *Then, for* $s > 2$,

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

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 by $M_\lambda(A)$ the $\lambda$-entropy number of A, $\lambda > 0$. By $\lambda$-entropy number, we mean the minimal number ($\leqslant \infty$) of balls (with respect to the H-norm) of radius $\lambda$, needed to cover A. We set $M_\lambda = 0$ if diam $A < \lambda$. The following result relates to H-valued functions on **R**.

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

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

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

*Proof.* — Observe first the pointwise inequality

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

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

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

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

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

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

This proves (3.31).

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

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

*Proof.* — With the notation of the proof of Lemma 3.30, set

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

so that

(3.35) $$\| f_\alpha^* \|_2 \leqslant c_\varphi \| f_\alpha \|_2$$

by the Hardy-Littlewood maximal inequality. Define

(3.36) $$\mathrm{F} = [\Sigma (f_\alpha^*)^2]^{1/2}$$

and, for $s > 2$, write

$$\int_0^\infty \min(\mathrm{K}, \mathrm{M}_\lambda(x))^{1/2}\, d\lambda \leqslant \mathrm{F}(x) + \int_{\mathrm{K}^{-1/s}\mathrm{F}(x)}^{\mathrm{F}(x)} \mathrm{K}^{\frac{1}{2}-\frac{1}{s}}\, \mathrm{M}_\lambda(x)^{1/s}\, d\lambda$$
$$\leqslant \mathrm{F}(x) + \mathrm{K}^{\frac{1}{2}-\frac{1}{s}} (\log \mathrm{K}) \sup_{\lambda > 0} \lambda . \mathrm{M}_\lambda(x)^{1/s}.$$

Now, (3.34) follows from (3.35), (3.36) and (3.31), letting $\frac{1}{2} - \frac{1}{s} = (\log \mathrm{K})^{-1}$.

## 4. Maximal Inequalities for Certain Sequences of Fourier Multipliers

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

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

The main result of this section is contained in

*Lemma* **4.1**. — *Assume* $\lambda_1 < \ldots < \lambda_\mathrm{K} \in \Pi$ *and, for* $j \in \mathbf{Z}_+$, *define the neighborhoods*

(4.2) $$\mathrm{R}_j = \{ \lambda \in \Pi \mid \min_{1 \leqslant k \leqslant \mathrm{K}} | \lambda - \lambda_k | \leqslant 2^{-j} \}.$$

*Then*

(4.3) $$\left\| \sup_j \left| \int_{\mathrm{R}_j} \hat{f}(\lambda)\, e^{2\pi i \lambda x}\, d\lambda \right| \right\|_{\ell^2(\mathbf{Z})} \leqslant \mathrm{C} (\log \mathrm{K})^2 \| f \|_{\ell^2(\mathbf{Z})}$$

*for functions* $f$ *on* $\mathbf{Z}$.

*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 for our purpose.

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

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

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

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

*Then*

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

*where* C *is an absolute constant.*

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

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

and averaging in $u$ gives

$$\| \sup_{\varphi} | \mathscr{F}^{-1}[\varphi \mathscr{F} f] | \|_{\ell^2(\mathbf{Z})} \leqslant$$

$$\rho^{-1/2} \| \sup_{\varphi} | \mathscr{F}^{-1}[\varphi \mathscr{F} f] | \|_{L^2(\mathbf{R})} + \tag{4.7}$$

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

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

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

By definition of $B_1$, (4.8) is bounded by

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

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

By Lemma 4.4, Lemma 4.1 may be restated as

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

$$\| \sup_j | \mathscr{F}^{-1}[\chi_{R_j} \mathscr{F} f] | \|_2 \leqslant C (\log K)^2 \| f \|_2. \tag{4.12}$$

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

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

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

*Proof.* — Observe first that

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

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

Since supp $\hat{\varphi}_t \subset [-\tau, \tau]$ for $t > \tau^{-1}$, there is no restriction in assuming that

$$(4.16)\qquad \operatorname{supp} \hat{f}_k \subset [-\tau, \tau] \quad \text{for } 1 \leqslant k \leqslant K.$$

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

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

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

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

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

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

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

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

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

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

$$(4.21)\qquad |\bar{b}| \leqslant 2.2^s \quad \text{for } \bar{b} \in B_s$$

$$(4.22)\qquad \# B_s \leqslant M_{(2^s)}$$

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

$$(4.23)\qquad \bar{a} = \sum_{s\in \mathbf{Z}} \bar{b}_s \quad \text{with } \bar{b}_s \in B_s$$

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

Estimate

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

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

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

For given $s$, consider the following bounds

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

They imply, invoking (4.15), that (4.24) is bounded by

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

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 that* $\lambda_1, \dots, \lambda_K \in \mathbf{R}$ *satisfy* $|\lambda_k - \lambda_{k'}| > 2^{-s}$ *for* $k \neq k'$. *Then, with previous notation,*

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

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

$$\sup_{j\geqslant s} | \mathcal{F}^{-1}[\chi_{R_j} \mathcal{F} f] | \leqslant$$

$$\sup_{j\geqslant s} \Big| \sum_{k=1}^{K} e^{2\pi i\lambda_k x}[(f.e^{-2\pi i\lambda_k x}) * \varphi_{2^j}] \Big| \tag{4.28}$$

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

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

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

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

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

$$\Big\{ \sum_{j\geqslant s} \int |\hat f(\lambda)|^2 \Big[\chi_{R_j}(\lambda) - \sum_{k=1}^{K} \hat\varphi(2^j(\lambda - \lambda_k))\Big]^2 d\lambda \Big\}^{1/2} \leqslant \sup_{\lambda\in\mathbf{R}} \Big[ \sum_{j\geqslant s} \Big| \chi_{R_j}(\lambda) - \sum_{k=1}^{K} \hat\varphi(2^j(\lambda - \lambda_k)) \Big| \Big] . \|f\|_2. \tag{4.31}$$

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

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

*Remark.* — In a later application, $\Lambda = \{\lambda_1, \ldots, \lambda_K\}$ will be typically a set of rational numbers $a/q$, $(a, q) = 1$, with $q \leqslant Q$ 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*

$$R_j = \{\lambda \in \mathbf{R} \mid \min_{1 \leqslant k \leqslant K} |\lambda - \lambda_k| \leqslant 2^{-j}\} \quad \textit{for } j \in \mathbf{Z}.$$

*Then*

$$\textbf{(4.33)} \qquad \| \sup_{j \in S} | \mathscr{F}^{-1}[\chi_{R_j} \mathscr{F} f] | \|_2 \leqslant (\log |S|) \, \|f\|_2$$

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

*Proof.* — The argument is inspired by the Burkholder-Davis-Gundy-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 $R_j$, that

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

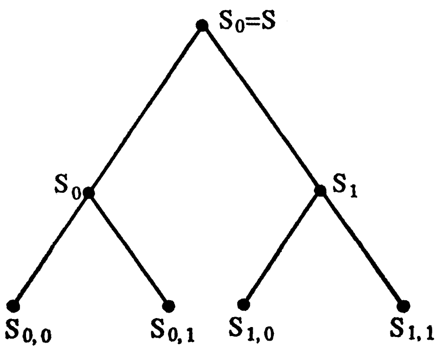
Denote by B the best constant satisfying the inequality

$$\| \sup_{1 \leqslant j \leqslant 2^s} | \mathscr{F}^{-1}[\chi_{R_j} \mathscr{F} f] | \|_2 \leqslant B \, \|f\|_2$$

or equivalently (by dualization)

$$\textbf{(4.34)} \qquad \| \sum_{j \leqslant 2^s} \mathscr{F}^{-1}[\chi_{R_j} \mathscr{F} g_j] \|_2 \leqslant B \, \| \sum_{j \leqslant 2^s} |g_j| \, \|_2.$$

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



Set $\tilde{g}_j = \mathscr{F}^{-1}[\chi_{R_j}\mathscr{F}g_j]$; clearly

(4.35) $$\langle \tilde{g}_j, \tilde{g}_k \rangle = \langle g_j, \tilde{g}_k \rangle \quad \text{for } j \leqslant k.$$

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

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

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

*Proof of* **(4.12)**. — Define

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

Obviously

(4.36) $$|S| \leqslant K^3.$$

Define further

$$Z_r = \{j \in \mathbf{Z}\backslash S \mid R_j \text{ has } r \text{ components}\}$$

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

(4.37) $$Z_1 \prec Z_2 \prec \ldots \prec Z_K$$

where $Z_1$, $Z_K$ are half-lines and $Z_r$ is 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

(4.38) $$|\lambda - \lambda'| > 2^{-j_r} \quad \text{for } \lambda \neq \lambda' \text{ in } \Lambda_r$$

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

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

$$\|\sum_j \tilde{g}_j\|_2 \leqslant B\|\Sigma |g_j|\|_2 \quad \text{for } \tilde{g}_j = \mathscr{F}^{-1}[\chi_{R_j}\mathscr{F}g_j].$$

Using (4.32) and (4.36) and setting $G_r = \sum_{j\in Z_r} g_j$ and $\tilde{G}_r = \sum_{j\in Z_r} \tilde{g}_j$ we have

$$\|\sum_j \tilde{g}_j\|_2 \leqslant \|\sum_{j\in S} \tilde{g}_j\|_2 + \|\sum_r (\sum_{j\in Z_r} \tilde{g}_j)\|_2 \leqslant (\log K)\|\Sigma |g_j|\|_2 + \|\sum_r \tilde{G}_r\|_2.$$

Since $Z_r \prec Z_{r'}$ 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

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

and $$\|\Sigma \tilde{G}_r\|_2^2 = \Sigma \|\tilde{G}_r\|_2^2 + 2\sum_{r<r'} \langle G_r, \tilde{G}_{r'} \rangle.$$

The same argument as in (4.32) than shows that

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

where $B_1$ has to satisfy

$$\| \sup_{j \in Z_r} | \mathscr{F}^{-1}[\chi_{B_j} \mathscr{F} f] | \|_2 \leqslant B_1 \|f\|_2. \tag{4.41}$$

In order to estimate $B_1$, apply (4.26) with $\Lambda = \Lambda_r$, $s = j_r$, taking into account (4.38) and $j \geqslant j_r$ 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 the exponential sums (1.5), i.e.,

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

where

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

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* $f(x) = \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_d x^d$ *and* $|\alpha_d - (a/q)| < 1/q^2$, *where* $(a, q) = 1$. *Then for all* $\varepsilon > 0$,

$$\Big| \sum_{m=1}^{n} e^{2\pi i f(m)} \Big| \leqslant C_\varepsilon n^{1+\varepsilon}[q^{-1} + n^{-1} + qn^{-d}]^\rho, \quad \textit{where } \rho = \frac{1}{2^{d-1}} \tag{5.4}$$

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

Denote by $\mathbf{Q}$ the set of rational numbers. For

$$\delta = \delta(d) > 0 \quad \text{and} \quad \theta_1, \ldots, \theta_d \in [0, 1] \cap \mathbf{Q}$$

with common denominator $q < N^\delta$, define the " major box " in the $d$-dimensional torus as

$$\mathscr{M}(\theta_1, \ldots, \theta_d) = \{\bar{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \Pi^d \mid |\alpha_j - \theta_j| < N^{-j+\delta} \ (1 \leqslant j \leqslant d)\}. \tag{5.5}$$

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* $\bar{\alpha}$ *does not belong to some major box as defined above, then*

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

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

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

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

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

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

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

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

Set

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

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

*Lemma* **5.12**. — *For* $\bar{\alpha} \in \mathscr{M}(\bar{\theta})$, $\bar{\alpha} = \bar{\theta} + \bar{\beta}$, *one has*

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

*where* $\theta_j = a_j/q$.

Recall also

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

$$|S(q, a_1, \ldots, a_d)| \leqslant c q^{-\delta'} \tag{5.15}$$

*where* $\delta' = \delta(d) > 0$.

This 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 the multiplicativity properties and A. Weil's estimate for $q$ a prime number.

Finally, we give some estimates on the function

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

*Lemma* **5.17**.

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

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

*where* $C = c(d)$.

The first estimate (5.18) is obvious and the second (5.19) follows 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

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

(6.2) $$p(x) = b_1 x + b_2 x^2 + \ldots + b_d x^d, \quad b_j \in \mathbf{Z} \text{ and } b_d > 0,$$

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

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

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

(6.4) $$A_N f = \mathscr{F}^{-1}[\mathscr{F}[K_N].\mathscr{F}[f]],$$

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

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

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

(6.6) $$\mathscr{R}_s = \{ \theta \in \mathbf{Q} \cap [0, 1] \mid \theta = a/q, \quad (a, q) = 1 \text{ and } 2^s \leqslant q < 2^{s+1} \}$$

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

Denote by $\zeta$ a smooth function on $\mathbf{R}$ 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 $\zeta$ will be irrelevant for the $L^2$-theory but has importance when considering $L^r$-estimates for $r < 2$ in the next section.)

Define

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

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

(6.8) $$S(\theta) = S(q', a'_1, \ldots, a'_d)$$
$$\text{where } -\theta.b_j \equiv a'_j/q' (\text{mod } 1) \text{ and } (a'_1, \ldots, a'_d, q') = 1,$$

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

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

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

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

(6.11) $$|S(\theta)| < C 2^{-s\delta'} \quad \text{for } \theta \in \mathscr{R}_s.$$

From (6.9) and (5.17)

**(6.12)** $$| 1 - w_N(\beta) | < C | \beta | . N^d,$$

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

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

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

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

*holds.*

This lemma allows the replacement of $\mathscr{F}[K_N]$ 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 $\Pi$ by letting

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

for $\theta$ a rational $a/q$, $1 \leq a \leq q$, $(a, q) = 1$ with $q < N^\delta$.

**Case 1.** — $\alpha$ belongs to an arc $\mathscr{M}(\theta_0)$.

Assume $\theta_0 \in \mathscr{R}_{s_0}$, 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),

**(6.17)** $$| \mathscr{F}[K_N] (\alpha) - \sum \psi_{s,N}(\alpha) | \leq | 1 - \zeta(10^{s_0}(\alpha - \theta_0)) | + \sum_{s \leq s_1} \sup | w_N(\alpha - \theta) | + C2^{-s_1\delta'} + CN^{-1/2},$$

where the sup is extended over all $\theta \in \mathscr{R}_s$ different from $\theta_0$.

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

$$| \alpha - \theta | \geq | \theta - \theta_0 | - | \alpha - \theta_0 |, \qquad | \theta - \theta_0 | > \frac{1}{2} q^{-1} 2^{-s_1} \geq \frac{1}{4} N^{-2\delta} \quad \text{for } \theta \in \mathscr{R}_s,$$

$s \leq s_1$, $\theta \neq \theta_0$ 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 $(\log N) . N^{-1+(2\delta/d)}$, invoking (6.13). Hence (6.15) holds.

**Case 2.** — $\alpha$ does not belong to a major arc.

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

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

By definition of $\mathcal{M}(\theta)$, it follows from the hypothesis on $\alpha$ that $|\alpha - \theta| > N^{-d+\delta}$ whenever $\theta \in \mathcal{R}_s$, $s \leqslant s_1$. 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

**(6.19)** $$\| \sup_N |f * K_N| \|_{\ell^2(Z)} \leqslant C \|f\|_{\ell^2(Z)},$$

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

**(6.20)** $$\psi_N = \sum_S \psi_{S,N},$$

estimate by (6.15) and Parseval

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

To estimate the contribution of the first terms, define

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

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

**(6.23)** $$\sum_{N \in Z_1} | \psi_{s,N} - \tilde{\psi}_{s,N} | \leqslant C 2^{-s\delta'}.$$

Therefore, again by a square function argument

**(6.24)** $$\| \sup_{Z_1} | \mathcal{F}^{-1}[\psi_{s,N} \mathcal{F} f] | \|_2 \leqslant \| \sup_{Z_1} | \mathcal{F}^{-1}[\tilde{\psi}_{s,N} \mathcal{F} f] | \|_2 + C 2^{-s\delta'} \|f\|_2.$$

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

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

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

it follows from inequality (4.3) in Lemma 4.1 that

**(6.27)** $$\begin{aligned} \| \sup_{N \in Z_1} | \mathcal{F}^{-1}[\tilde{\psi}_{s,N} \cdot \mathcal{F} f] | \|_2 &\leqslant \| \sup_{j \in Z_+} | \mathcal{F}^{-1}[\mathcal{F}[g_s]\, \chi_{R_j}] | \|_2 \\ &\leqslant C (\log |\mathcal{R}_s|)^2 \| g_s \|_2. \end{aligned}$$

By definition, $|\mathcal{R}_s| < 4^s$, and it follows from (6.11) and Parseval that

$$\| g_s \|_2 \leqslant C 2^{-s\delta'} \|f\|_2.$$

Substitution in (6.24) yields the bound

**(6.28)** $$\| \sup_{Z_1} | \mathcal{F}^{-1}[\psi_{s,N} \mathcal{F} f] | \|_2 \leqslant C . s^2\, 2^{-s\delta'} \|f\|_2$$

and hence, substituting in (6.21),

$$\text{(6.29)} \qquad \| \sup_{N \in Z_1} |f * K_N| \|_2 \leqslant C\Big(\sum_{s=0}^{\infty} s^2 2^{-s\delta'}\Big) \|f\|_2 \leqslant C \|f\|_2$$

which proves the maximal inequality

$$\text{(6.30)} \qquad \| \sup_N | A_N f | \|_2 \leqslant C \|f\|_2.$$

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 $(\mathbf{Z}, S)$

$$\text{(6.31)} \qquad \sum_{j=1}^{J} \| \mathscr{M}_j f \|_2 \leqslant o(J) \|f\|_2$$

setting

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

where $Z_\varepsilon = \{[(1+\varepsilon)^n]; n = 1, 2, \dots\}$ for $\varepsilon > 0$ 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 $\mathscr{F}^{-1}[(\psi_N - \psi_{N_j}) \mathscr{F} f]$ when defining $\mathscr{M}_j f$. Fixing $s_0$, it follows from the previous inequality (6.28) that then

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

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

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

where $w_N$ is given by (6.9). The reader will indeed verify that summing up 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

$$\text{(6.35)} \qquad \widetilde{\mathscr{M}}_j f = \sup_{N_j < N < N_{j+1}} |f * (\chi_{(N^d)} - \chi_{(N_j^d)})|,$$

where $\chi_t = \dfrac{1}{t} \chi_{[0,t]}$. Since the following inequality clearly holds pointwise (with $v_s$ as in section 3),

$$\text{(6.36)} \qquad \Big\{ \sum_{j=1}^{J} (\widetilde{\mathscr{M}}_j f)^2 \Big\}^{1/2} \leqslant J^{1/4} \| \{ f * \chi_N \mid N = 1, 2, \dots \} \|_{v_4},$$

is follows from (6.36) and (3.26) that

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

hence

$$(6.37) \qquad \sum_{j=1}^{J} ||\mathcal{M}_j f||_2^2 \leqslant C \sum_{N \in Z_\varepsilon} || \mathcal{F}^{-1}[(w_N - \mathcal{F}(\chi_{N^d})) \mathcal{F} f ||_2^2 + CJ^{1/2} ||f||_2^2.$$

This first term in (6.37) is bounded by

$$(6.38) \qquad \sup_\alpha [\sum_{N \in Z_\varepsilon} | w_N(\alpha) - \hat{\chi}(N^d \alpha) |^2] \, ||f||_2^2 < C_\varepsilon \, ||f||_2^2$$

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

$$(6.39) \qquad | w_N(\alpha) - \hat{\chi}(N^d \alpha)| \leqslant C \min(|\alpha| N^d, (|\alpha| N^d)^{-1/d}).$$

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

$$(6.40) \qquad \Sigma \, ||\mathcal{M}_j f||_2 \leqslant C_\varepsilon J^{3/4} \, ||f||_2$$

independently of the choice of the sequence $N_1 \leqslant N_2 \leqslant \ldots \leqslant N_J$. 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 $A_N f \to \int f \, d\mu$ in $L^2$ (hence a.s.). Indeed T has no point spectrum as unitary operator and $\frac{1}{N} \sum_{n \leqslant N} z^{p(n)} \xrightarrow{N \to \infty} 0$ for

$$z \in \mathbf{C}_1 = \{ z \in \mathbf{C} \mid |z| = 1 \},$$

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

$$(7.1) \qquad || \sup_N | A_N f | \, ||_p \leqslant c \, ||f||_p$$

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 ($\mathbf{Z}$, S), (7.1) becomes

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

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 goods bounds on $L^p$.

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

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

Denote again by $\zeta$ a smooth function on $\mathbf{R}$, $0 \leqslant \zeta \leqslant 1$, $\operatorname{supp} \zeta \subset [-1/2, 1/2]$ and $\zeta = 1$ on $[-1/4, 1/4]$.

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

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

$$\left\| \int F(\beta)\, e^{2\pi i \beta q v}\, \zeta(D\beta)\, d\beta \right\|_{L^p(\mathbf{R})} \sim \left\| \int F(\beta)\, e^{2\pi i \beta q v}\, \zeta(D\beta)\, d\beta \right\|_{\ell^p(\mathbf{Z})}. \tag{7.5}$$

*Proof.* — Observe that, by Bernstein's inequality and the hypothesis,

$$\left\| \int F(\beta)\, [e^{2\pi i q \beta u} - 1]\, e^{2\pi i q \beta v}\, \zeta(D\beta)\, d\beta \right\|_{L^p} \leqslant C\varepsilon \left\| \int F(\beta)\, e^{2\pi i q \beta v}\, \zeta(D\beta)\, d\beta \right\|_{L^p} \tag{7.6}$$

for $0 \leqslant u \leqslant 1$.

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

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

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

$$\begin{aligned} \rho \left\| \int F(\beta)\, [1 - e^{2\pi i \beta q u}]\, e^{2\pi i \beta q v}\, \zeta(D\beta)\, d\beta \right\|_{L^p} & \\ \leqslant C\varepsilon\rho \left\| \int F(\beta)\, e^{2\pi i q \beta v}\, \zeta(D\beta)\, d\beta \right\|_{L^p}, & \end{aligned}$$

invoking (7.6). Thus it follows that $\rho \leqslant 1 + C\varepsilon\rho$, hence the boundedness of $\rho$.

To prove the converse inequality in (7.5), write

$$\| \quad \|_{L^p(\mathbf{R})} \leqslant \| \quad \|_{\ell^p(\mathbf{Z})} + \left\{ \int_0^1 \left\| \int F(\beta)\, e^{2\pi i \beta q v} [1 - e^{2\pi i \beta q u}]\, \zeta(D\beta)\, d\beta \right\|_{\ell^p(dv)}^p du \right\}^{1/p} \tag{7.8}$$

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

$$\left\| \int F(\beta)\, e^{2\pi i \beta q v} [1 - e^{2\pi i \beta q u}]\, \zeta(D\beta)\, d\beta \right\|_{\ell^p(dv)} \leqslant C\varepsilon\rho \, \| F(\beta)\, e^{2\pi i q \beta v}\, \zeta(D\beta)\, d\beta \|_{L^p} \tag{7.9}$$

for $0 \leqslant u \leqslant 1$. Since $C\varepsilon\rho < 1/2$ for $\varepsilon$ 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* $\ell^1(\mathbf{Z})$*-norm of the Fourier transform of the function on* $\Pi$

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

*is bounded by*

$$q \sum_{j \in \mathbf{Z}} \sup_{0 \leqslant x < q} | \mathscr{F} F(jq + x) |. \tag{7.12}$$

*Proof.* — By definition of $S(\theta)$, the Fourier transform of (7.11) at the point $x \in \mathbf{Z}$ equals

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

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

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

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

*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 $\nu_N$ under the mapping $p(Ny) : [0, 1] \to \mathbf{R}$. Hence $\mathscr{F}F = \nu_N * (\mathscr{F}^{-1}[\zeta])_D$ and (7.12) is bounded by

$$\textbf{(7.15)} \qquad \frac{q}{D} \sum_{j \in \mathbf{Z}} \sup_{0 \leq x < q} \int_0^1 \left| \mathscr{F}^{-1}[\zeta]\left(\frac{x + jq - p(Ny)}{D}\right) \right| dy.$$

Since $| \mathscr{F}^{-1}[\zeta](t) | < C(1 + t^2)^{-1}$, one has

$$\sum_{j \in \mathbf{Z}} \sup_{0 \leq x < q} \left| \mathscr{F}^{-1}[\zeta]\left(\frac{x + jq}{D}\right) \right| < C\left(\frac{D}{q} + 1\right).$$

Substituting in (7.15), (7.14) follows.

One has the following real Analysis maximal inequality:

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

$$\textbf{(7.17)} \qquad \| \sup_N | \mathscr{F}^{-1}[w_N \mathscr{F}f] | \|_{L^p(\mathbf{R})} \leq C \| f \|_{L^p(\mathbf{R})}.$$

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

$$\textbf{(7.18)} \qquad \| \sup_N | f * [(1/N)\, t^{-1+(1/d)} \chi_{[0, N^d]}(t)] | \|_p \leq C \| f \|_p.$$

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

$$\textbf{(7.19)} \qquad \| \sup_{s > 0} | f * k_s | \|_p \leq C \| f \|_p$$

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

Next, we prove a discrete maximal inequality:

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

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

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

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

with

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

As in the proof of (7.4), denote by $\rho$ the *a priori* best constant in the inequality

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

For $0 \leqslant u < 1$, write

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

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

$$\begin{aligned} q^{-1/p} \|\sup_{N} | \mathscr{F}^{-1}[w_N F\zeta(D.)]| \|_{L^p} \leqslant \\ Cq^{-1/p} \| \mathscr{F}^{-1}[F\zeta(D.)]\|_{L^p} = \\ C \left\| \int F(\beta)\, \zeta(D\beta)\, e^{2\pi i \beta q y}\, d\beta \right\|_{L^p(dy)} \sim \left\| \int F(\beta)\, \zeta(D\beta)\, e^{2\pi i q y}\, d\beta \right\|_{\ell^p(dy)}. \end{aligned} \tag{7.26}$$

By definition of $\rho$, the $\ell^p$-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 $\rho \leqslant C + C\varepsilon\rho$ implies $\rho < C$, assuming $\varepsilon$ small enough. This yields (7.24).

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

$$\Big\| \sup_N \Big| \int w_N(\beta)\, F_z(\beta)\, \zeta(D\beta)\, e^{2\pi i\beta q y}\, d\beta \Big| \Big\|_{\ell^p(dy)} \leqslant$$

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

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

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

$$\Big\| f * \mathscr{F}^{-1} \Big[ \sum_{0 \leqslant a < q} \zeta\Big(D\Big(\cdot - \frac{a}{q}\Big)\Big)\Big] \Big\|_{\ell^p} < C\, \|f\|_{\ell^p}.$$

This completes the proof of Lemma 7.20.

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

**(7.30)** $$\Big\| \sup_N \Big| \sum_{0 \leqslant a < q} S\Big(\frac{a}{q}\Big) \int w_N(\beta)\, \mathscr{F}f\Big(\frac{a}{q} + \beta\Big)\, \zeta(D\beta)\, e^{2\pi i x(\frac{a}{q}+\beta)}\, d\beta \Big| \Big\|_{\ell^p} \leqslant C_p\, \|f\|_{\ell^p}.$$

*Proof.* — Apply (7.21) to the function $g$ given by

**(7.31)** $$\mathscr{F}g(\alpha) = \Big[ \sum_{0 \leqslant a < q} S\Big(\frac{a}{q}\Big)\, \zeta\Big(\frac{D}{4}\Big(\alpha - \frac{a}{q}\Big)\Big)\Big]\, \mathscr{F}f(\alpha).$$

Observe that $\zeta\Big(\frac{D}{4}\beta\Big)\zeta(D\beta) = \zeta(D\beta)$ and that the first factor in (7.31) is the Fourier-transform of an $\ell^1(\mathbf{Z})$-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.** — *One has the following restricted maximal inequality*:

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

This is a problem about positive functions and hence N may be taken of the form $N = 2^k$, $k_0 \leqslant k \leqslant 2k_0$. Instead of considering the $\ell^p(\mathbf{Z})$-inequality, we will rather deal with functions $f$ taken on a finite cyclic group $G \equiv \mathbf{Z}_J = \mathbf{Z}/J\mathbf{Z}$, 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 $\frac{1}{N} \sum_{1 \leqslant n \leqslant N} \delta_{\{p(n)\}}$. The inequality (7.33) is equivalent to

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

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

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

Since (7.35) weakens for increasing $p$, one may assume that $q = p' = p/(p-1)$ is an integer. We replace (7.35) by its dual version

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

whenever

$$\text{(7.37)} \qquad g_k \geqslant 0, \qquad \Sigma g_k \leqslant 1.$$

Let M (to be specified later) satisfy

$$\text{(7.38)} \qquad M \sim \log k_0$$

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

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

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

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

Expanding the $q$th power of a sum and integrating, we have

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

$$\text{(7.40)} \qquad C \sum_{k_0 < k_1 < \dots < k_q < 2k_0} \int_G (g_{k_1} * L_{k_1}) \dots (g_{k_q} * L_{k_q})$$

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

where (7.41) is bounded by $\rho^{k-1}$.

Choosing M appropriately, we will achieve the estimate

$$\text{(7.42)} \qquad \| [(g_{k_2} * L_{k_2}) \dots (g_{k_q} * L_{k_q})] * (L_{k_1} - L_{k_0}) \|_{L^2(G)} \leqslant k_0^{-q}$$

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

Once (7.42) is obtained, write

$$\Big| \int_G (g_{k_1} * L_{k_1}) (g_{k_2} * L_{k_2}) \dots (g_{k_q} * L_{k_q}) - \int_G (g_{k_1} * L_{k_0}) (g_{k_2} * L_{k_2}) \dots (g_{k_q} * L_{k_q}) \Big| < k_0^{-q}$$

and estimate (7.40) by

$$\text{(7.43)} \qquad C + \sum_{k_0 < k_2 < \dots < k_q < 2k_0} \int_G \Big[ \Big( \sum_{k_0 < k < 2k_0} g_k \Big) * L_{k_0} \Big] (g_{k_2} * L_{k_2}) \dots (g_{k_q} * L_{k_q}).$$

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 $\mathscr{F}$ for the Fourier transform on **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. With the notation of the previous section and D to be specified, define

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

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

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

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

$$|| \mathscr{F}^{-1}(\Omega_r) ||_{\ell^1(\mathbf{Z})} \leqslant C4^{s_r}, \tag{7.46}$$

there is a uniform estimate

$$| \mathscr{F}^{-1}[\Omega_r . \mathscr{F}(g_r)] | \leqslant C4^{s_r}. \tag{7.47}$$

It also follows from (7.45) that

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

Observe that the Fourier transform of the function $\mathscr{F}^{-1}[\Omega_2\, \mathscr{F}(g_2)] \ldots \mathscr{F}^{-1}[\Omega_q\, \mathscr{F}(g_q)]$ vanishes outside a $DN_2^{-d}$ neighborhood $\Gamma$ of $\{(a/b) \in \Pi \cap \mathbf{Q} \mid b \leqslant 2^{qs_q}\}$.

Estimate the left member of (7.42) as

$$\begin{aligned} & || (g_2 * L_{k_2}) - \mathscr{F}^{-1}[\mathscr{F}(g_2)\, \Omega_2] ||_{L^2(G)} + \\ & \quad || \mathscr{F}^{-1}[\mathscr{F}(g_2)\, \Omega_2] ||_\infty\, || (g_3 * L_{k_3}) - \mathscr{F}^{-1}[\mathscr{F}(g_3)\, \Omega_3] ||_{L^2(G)} + \ldots + \\ & \quad || \mathscr{F}^{-1}[\mathscr{F}(g_2)\, \Omega_2 ||_\infty \ldots || \mathscr{F}^{-1}[\mathscr{F}(g_{q-1})\, \Omega_{q-1}] ||_\infty\, || (g_q * L_{k_q}) \\ & \qquad - \mathscr{F}^{-1}[\mathscr{F}(g_q)\, \Omega_q] ||_{L^2(G)} \end{aligned} \tag{7.49}$$

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

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

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

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

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

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

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

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 $\Gamma$, (6.10) and (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 that

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

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

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

$$(7.53) \leqslant k_0^C[D2^{-Md} + 2^{-\frac{1}{2}Mk_0}]. \tag{7.56}$$

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

$$\frac{1}{10}k_0^{-q} + k_0^C[D^{-1/d} + D2^{-Md} + 2^{-\delta k_0}] < k_0^{-q}$$

for a suitable choice of D, $\log D \sim \log k_0$ and $M \sim \log k_0$ (cf. (7.38)). This completes the proof of (7.42) and hence of 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 by $\| \ \|_r$ the $\ell^r(\mathbf{Z})$-norm in what follows. For $s = 1, 2, \ldots$, define

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

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

and with previous notation, let

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

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

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

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

and by (7.13)

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

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

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

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

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

so that

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

By (7.62)

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

while by (7.33) and (7.62) also

$$\| \sup_{k \in \mathscr{K}_s} |(f * K_{2^k}) - \mathscr{F}^{-1}[\Omega_{k,s}\,\mathscr{F}f]| \|_{p_0} \leqslant C.s \|f\|_{p_0}. \tag{7.65}$$

Our purpose is to interpolate (7.64), (7.65) with better $\ell^2$-estimates. Using (6.15), estimate

$$\| \sup_{k \geqslant 4^{s'}} |(f * K_{2^k}) - \mathscr{F}^{-1}[\Omega_{k,s'}\,\mathscr{F}f]| \|_2 \leqslant$$

$$C \sum_{k \geqslant 4^{s'}} 2^{-k\delta_1} + \| \sup_{k \geqslant 4^{s'}} | \sum_{0 \leqslant r \leqslant s'} \mathscr{F}^{-1}[\psi_{r,2^k}.\mathscr{F}f] - \mathscr{F}^{-1}[\Omega_{k,s'}.\mathscr{F}f]| \|_2 + \sum_{r > s'} \| \sup_k |\mathscr{F}^{-1}[\psi_{r,2^k}.\mathscr{F}f]| \|_2, \tag{7.66}$$

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

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

Write

$$\Omega_{k,s'} - \sum_{r \leqslant s'} \psi_{r,2^k} =$$

$$\sum_{r \leqslant s'} \sum_{\theta \in \mathscr{R}_r} S(\theta)\, w_{2^k}(\alpha - \theta)\, [\zeta(Q_{s'}^2(\alpha - \theta)) - \zeta(10^r(\alpha - \theta))] \tag{7.67}$$

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

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

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

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

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

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

$$\| \sup_{k \geqslant 4^{s'}} |\mathscr{F}^{-1}[(\Omega_{k,s'} - \Omega_{k,s'-1})\,\mathscr{F}f]| \|_2 \leqslant C.2^{-\delta' s'} \|f\|_2 \tag{7.70}$$

while

$$\| \sup_{k \in \mathscr{K}_s} |(f * K_{2^k}) - \mathscr{F}^{-1}[\Omega_{ks}.\mathscr{F}f]| \|_2 \leqslant C.2^{-\delta' s} \|f\|_2. \tag{7.71}$$

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

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

completing the proof of (7.2).

## 8. Integer Parts of Polynomial Sequences

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

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

and for a given $DS(\Omega, \mathscr{B}, \mu, T)$ 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 $b_1, \ldots, b_d$ irrational. Otherwise, if $b_j = (a_j/q)$ $(1 \leqslant j \leqslant q)$, write $n = mq + r$ $(0 \leqslant r < q)$ and

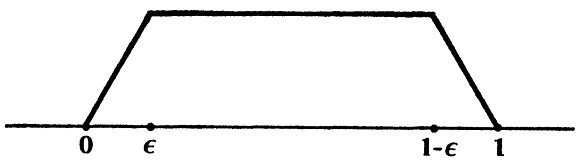
$$A_N f = \frac{1}{q} \sum_{0 \leqslant r < q} \frac{1}{Nq^{-1}} \sum_{m \leqslant Nq^{-1}} T^{p_1(m)}(T^{[p(r)]} f)$$

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 that $b_1, \ldots, b_d$ are 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 $\mathbf{R}$




Set

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

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

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

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 uses the same method as in section 6. The relevant exponential sums are given by

$$(8.5)\qquad \mathscr{F}[K_N](\alpha) = \frac{1}{N} \sum_{n \leqslant N} \sum_{m \in \mathbf{Z}} \tau(p(n) - m)\, e^{-2\pi i m\alpha} = \sum_{k \in \mathbf{Z}} \hat{\tau}(k - \alpha) \left\{ \frac{1}{N} \sum_{n \leqslant N} e^{2\pi i (k-\alpha) p(n)} \right\},$$

where

$$(8.6)\qquad K_N = \frac{1}{N} \sum_{n \leqslant N} \sum_{m \in \mathbf{Z}} \tau(p(n) - m)\, \delta_{\{m\}}.$$

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

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

Then

$$(8.7)\qquad \mathscr{F}[K_N](\alpha) = \sum_{k \in \mathbf{Z}} \hat{\tau}(k - \alpha)\, e^{2\pi i (k-\alpha) b_0}\, \varphi_N(b_1(k - \alpha), \ldots, b_d(k - \alpha)).$$

Observe also the decay property

$$(8.8)\qquad |\hat{\tau}(\lambda)| < \frac{C}{1 + \varepsilon^2 \lambda^2}.$$

For $k \in \mathbf{Z}$, define

$$(8.9)\qquad \mathscr{R}_{s,k} = \left\{ \theta \in [0, 1] \mid b_j(k - \theta) \equiv \frac{a_j}{q} \pmod 1 \text{ where } (q, a_1, \ldots, a_d) = 1 \text{ and } 2^s \leqslant q \leqslant 2^{s+1} \right\}$$

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

$$(8.10)\qquad S(\theta) = S(q, a_1, \ldots, a_d).$$

Define also, with the notation (5.11),

$$(8.11)\qquad w_N(\beta) = v_N(-b_1 \beta, \ldots, -b_d \beta).$$

Set further

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

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

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

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

$$(8.14)\qquad |\,\mathscr{F}[\mathrm{K}_\mathrm{N}] - \sum_{s\geqslant 0} \psi_{s,\mathrm{N}}\,| < \mathrm{C}.\mathrm{N}^{-\delta_1}.$$

We leave the verification to the reader.

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

$$(8.15)\qquad \|\sup_{\mathrm{N}\in \mathrm{Z}_1} |\,\mathscr{F}^{-1}[\psi_{s,k,\mathrm{N}}\,\mathscr{F}f]\,|\,\|_2 \leqslant \mathrm{C}(\log|\,\mathscr{R}_{s,k}\,|)^2\, 2^{-\delta' s}\,\|f\|_2 \sim \mathrm{C}s^2\, 2^{-\delta' s}\,\|f\|_2.$$

Hence, by (8.8)

$$\|\sup_{\mathrm{N}\in \mathrm{Z}_1} |\,\mathscr{F}^{-1}[(\sum_{s\geqslant 0}\psi_{s,\mathrm{N}})\,\mathscr{F}f]\,|\,\|_2 \leqslant \mathrm{C}\sum_{s\geqslant 0}\sum_{k\in\mathbf{Z}} \frac{1}{1+\varepsilon^2 k^2}\, s^2\, 2^{-\delta' s}\,\|f\|_2 \leqslant \mathrm{C}\,\|f\|_2$$

yielding the maximal inequality

$$(8.16)\qquad \|\sup|\,\widetilde{\mathrm{A}}_\mathrm{N} f\,|\,\|_2 \leqslant \mathrm{C}_\varepsilon\,\|f\|_2.$$

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 $\mathrm{L}^\infty$-functions).

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

## 9. Further Comments and Remarks on Almost Sure Convergence

(1) In [B$_3$], the author considered the sequence $\Lambda$ of prime numbers and proved that the averages

$$(9.1)\qquad \mathrm{A}_\mathrm{N} f = \frac{1}{|\,\Lambda_\mathrm{N}\,|}\sum_{n\in\Lambda_\mathrm{N}} \mathrm{T}^n f;\quad \Lambda_\mathrm{N} = \{\text{primes} \leqslant \mathrm{N}\}$$

converge a.s. for $f$ a function of class $\mathrm{L}^2$. Setting

$$(9.2)\qquad \mathrm{K}_\mathrm{N} = \frac{1}{\mathrm{N}}\sum_{\substack{p\leqslant \mathrm{N}\\ p\text{ prime}}} (\log p)\, \delta_{\{p\}},$$

it is well known that

$$(9.3)\qquad \mathscr{F}[\mathrm{K}_\mathrm{N}]\,(\alpha) = \frac{\mu(q)}{\varphi(q)}\,\frac{1}{\mathrm{N}}\,(\sum_{k=1}^{\mathrm{N}} e^{2\pi i k(\alpha - \frac{a}{q})}) + \mathrm{O}(e^{-c\sqrt{\log \mathrm{N}}})$$

for $|\,\alpha - (a/q)\,| < (\log \mathrm{N})^c\,\mathrm{N}^{-1}$, $1 \leqslant a \leqslant q$, $(a, q) = 1$ and $q < (\log \mathrm{N})^c$. Here $\mu$ denotes the Moebius function and $\varphi(q)$ the number of Dirichlet characters to the modulus $q$.

Prior to this work, it has been shown in [W1] that $A_N f$ given by (9.1) converges 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 to 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 Theorems 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, one has 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 $\Pi$ of class $L^r$, $r > 1$. Then $(1/N) \sum_{1 \leqslant n \leqslant N} f(2^{p(n)} x)$ converges 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 $\Pi$ such that $(1/N) \sum_{1 \leqslant n \leqslant N} f(nx)$ 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 $A_N f = (1/N) \sum_{n \leqslant N} T_n f$. 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 $\delta$-metrical entropy number (in the sense of section 3)

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

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 if often effective in disproving the a.s. convergence of such averages.