**Real Analysis Exchange**
Vol. 30(2), 2004/2005, pp. 755–766

Sergey Kostyukovsky and Alexander Olevskii, School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel.
email: `sergey@post.tau.ac.il` and `olevskii@post.tau.ac.il`

# COMPACTNESS OF FAMILIES OF CONVOLUTION OPERATORS WITH RESPECT TO CONVERGENCE ALMOST EVERYWHERE

**Abstract**

For a given sequence of measures $\mu_n$ on the circle $\mathbb{T}$ weakly convergent to the Dirac measure, we ask, is it possible to extract a subsequence $n(j)$ such that for any $f$ in the space $L^1(L^2, L^\infty)$ the convolutions $f * \mu_{n(j)}$ converge to $f$ almost everywhere. We show that it is crucial whether the measures are absolutely continuous, discrete or singular (non-atomic).

## 1 Introduction.

Let $\mu_n$ $(n = 1, 2, \dots)$ be a sequence of probability measures on the circle $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ or in $\mathbb{R}^d$. Consider the convolution operators

$$f \to (f * \mu_n)(x) := \int f(x-t)\, d\mu_n(t).$$

One may ask whether they converge in a functional space $F \subset L^1$, meaning that there is a measure $\mu$, such that for any $f \in F$,

$$(f * \mu_n)(x) \to (f * \mu)(x) \ (n \to \infty)$$

almost everywhere (a.e.) with respect to Lebesgue measure.

The same question can be asked for a family $\mu_t$, depending on a continuous parameter $t$ . A number of well-known results can be viewed in this framework.

Key Words: almost everywhere convergence, convolutions
Mathematical Reviews subject classification: 42A05, 42A45
Received by the editors January 13, 2005
Communicated by: Clifford E. Weil

In particular, if $\mu_t$ is probability measure uniformly distributed on the interval $(-t,t)$ (or in the ball in $\mathbb{R}^d$ of radius $t$ centered at zero), then for $f \in L^1$,

$$(f * \mu_t)(x) \to f(x) \text{ } a.e. \text{ } (t \to 0). \tag{1}$$

This is the classical Lebesgue differentiation theorem.

A more general situation appears when each $\mu_t$ is an absolutely continuous measure, with density $K_t$ obtained from a single function $K$ by a "contraction". Let $K$ be supported on $[-1,1]$, non-negative and $\int K = 1$. Let

$$K_t(x) = \frac{1}{t} K(\frac{x}{t}), \ d\mu_t = K_t(x)\, dx. \tag{2}$$

If one requires convergence of convolutions (1) at every Lebesgue point $x$, then a necessary and sufficient condition follows from the Fadeev-Romanovskii theorem. $K$ must have a "humpbacked" majorant (that is, increasing for $x < 0$ and decreasing for $x > 0$) belonging to $L^1$ (see [N]).

But in general, to recognize whether a given family $\{\mu_t\}$ satisfies the convergence property (1) is a difficult problem, even for families (2). A natural restriction is that the measures $\mu_t$ are concentrated near the origin (as in (2)); that is, for any ball $B$ centered at zero,

$$\mu_t(B) \to 1 \text{ as } t \to 0. \tag{3}$$

In this case one says that $\mu_t$ is *an approximate identity* ($\mu_t \in$(AI)).

Clearly this implies (1) (uniformly) for any continuous function $f$. But convergence for $f \in L^1(\mathbb{T})$ and even for $f \in L^\infty$ does not follow.

There are many generalizations and versions of the Lebesgue theorem, see [Br] for a comprehensive survey. Notice especially an interesting result of Nagel and Stein ([NS], see also [S] ch.2). If the measures $\mu_n$ are uniformly distributed on intervals $I_n = [a(n), b(n)]$, $0 < a(n) < b(n), a(n+1) = o(|I_n|)$, then $f * \mu_n \to f$ a.e. for $f \in L^1$.

A surprising phenomenon was discovered for singular measures in $\mathbb{R}^d$, $d > 1$ ( see [S] ch. 11). If instead of averaging over balls one averages over spheres, then the corresponding analog of the Lebesgue theorem holds for $f \in L^p$, $p > d/(d-1)$.

The case of discrete measures is also of great interest. The Birkhoff ergodic theorem gives a classic example. On the other hand, Bourgain proved in [B] that, given a sequence of numbers $a(n) = o(1)$, $a(n) \neq 0$, one can construct a function $f \in L^\infty$ such that the averages

$$\frac{1}{N} \sum_{n=1}^{N} f(t + a(n))$$

do not converge a.e.

Having a sequence of probability measures $\mu_n$ satisfying the (AI) condition (3), we ask whether it is possible to select an increasing sequence of integers $\{n(j)\}$ such that, for any $f$

$$f * \mu_{n(j)} \to f \text{ } a.e. \tag{4}$$

(with respect to the Lebesgue measure).

First we consider absolutely continuous measures. We prove that, in this case, the answer is positive (Theorem 1). However the sequence in general must be sparse (Theorem 2).

For discrete measures the answer is negative (Theorem 3). In this case convergence (4) never happens, unless the measures are not sitting at the origin.

For families of non-atomic singular measures both situations are possible; they may behave like absolutely continuous and like discrete ones. This depends on the arithmetical nature of supports rather than on the metrical sizes. One can see this from Theorems 4 and 5, where some Fourier Analysis conditions are imposed.

Below the sign $\widehat{\ }$ stands for the Fourier transform of functions and measures.

## 2 Absolutely Continuous Measures.

We will prove the following compactness theorem:

**Theorem 1.** *Consider an (AI) kernel $\{K_n\}$:*

$$\begin{gathered} K_n(t) \geq 0, \int K_n \, dt = 1, \\ \int_{|t|>d} K_n \to 0 \text{ for every } d > 0 \; (n \to \infty). \end{gathered} \tag{5}$$

*Then there exists an increasing sequence of integers $\{n(j)\}$ such that for any $f \in L^1$*

$$(f * K_{n(j)})(x) \to f(x) \text{ } a.e. \; (j \to \infty). \tag{6}$$

This result is inspired by the Nagel-Stein theorem above, which covers the case when the $K_n$ are indicators of segments (normalized in $L^1$).

PROOF. Denote by $\omega(f, \delta)$ the integral modulus of continuity.

$$\omega(f, \delta) := \sup_{|t| \leq \delta} \int |f(x-t) - f(x)| \, dx.$$

Proceed by a simple induction. Suppose integers $0 < n(1) < \ldots n(s-1)$ and numbers $1 = d(0) > d(1) \cdots > d(s-1)$ have already been defined with conditions (for $k < s$):

$$\begin{gathered} \omega(K_{n(k)}, d(k)) < 2^{-k}, \\ \int_{|t|>d(k-1)} K_{n(k)} < 2^{-(k+1)}. \end{gathered} \tag{7}$$

Using (5) choose $n(s) > n(s-1)$ so that

$$\int_{|t|>d(s-1)} K_{n(s)} < 2^{-(s+1)}$$

and then a number $d(s)$ so that $\omega(K_{n(s)}, d(s)) < 2^{-s}$. So we get monotone sequences $\{n(k)\}$, $\{d(k)\}$ satisfying the conditions (7) for each $k$. Now for any $h$, $|h| < 1$, we have the estimate:

$$\sum \int_{|t|>2|h|} |K_{n(k)}(t-h) - K_{n(k)}(t)| < 3. \tag{8}$$

Indeed, fix $h \neq 0$ and find $s$ such that $d(s+1) < |h| \leq d(s)$. Then, for $k < s+1$,

$$\begin{aligned} \int_{|t|>2|h|} |K_{n(k)}(t-h) - K_{n(k)}(t)| &\leq \omega(K_{n(k)}; 2|h|) \leq \omega(K_{n(k)}; d(s)) \\ &\leq \omega(K_{n(k)}; d(k)) < 2^{-k}. \end{aligned}$$

For $k > s+1$,

$$\begin{aligned} \int_{|t|>2|h|} |K_{n(k)}(t-h) - K_{n(k)}(t)| &\leq 2 \int_{|t|>|h|} K_{n(k)}(t) \\ &\leq 2 \int_{|t|>d(s+1)} K_{n(k)}(t) < 2^{-k}. \end{aligned}$$

Summing up over $k$ and using (5), we get (8), which implies

$$\int_{|t|>2|h|} \sup_k |K_{n(k)}(t-h) - K_{n(k)}(t)| < 3.$$

Now, as in [S] p. 74, we can apply the Banach version of Calderon-Zygmund theorem. The integral operator

$$f \to \{f * K_{n(k)}\}\ k = 1, 2, \ldots,$$

as an operator from $L^1$ into $L^1(l_\infty)$, satisfies the condition of this theorem and, therefore, it is of weak type $(1,1)$. This means that the maximal operator

$$Mf(x) = \sup_k |f * K_{n(k)}(x)|$$

satisfies the inequality

$$\text{mes}\{Mf(x) > \lambda\} \leq \frac{C}{\lambda}\|f\|_{L^1}. \tag{9}$$

In addition (6) holds for a dense set (of continuous functions $f$), and it is well known that these two properties imply (6) for any $f \in L^1$. □

In general, the subsequence in Theorem 1 must be sparse. Let $K$ be a non-negative function on $\mathbb{R}$, $\int K = 1$, supported on the segment $[-1,1]$. Consider "contracted" functions on $\mathbb{T}$ defined as

$$K_n(x) = nK(nx) \text{ for } |x| \leq \pi.$$

Clearly this is an (AI) kernel. If $K$ decreases on $[0,1]$ and increases on $[-1,0]$, then $f * K_n \to f$ a.e. (actually at all Lebesgue points) according to Fadeev-Romanovskii theorem. On the other hand, the following result is true.

**Theorem 2.** *If $K$ is essentially unbounded near some point $d \neq 0$, then for any subsequence $\{n(k)\}$ satisfying*

$$n(k+1)/n(k) \to 1 \ (k \to \infty), \tag{10}$$

*there is a function $f \in L^1(\mathbb{T})$ such that convolutions $f * K_{n(k)}$ diverge almost everywhere.*

Proof. It is well known (see [S], p. 441) that it is enough to disprove the weak type inequality (9). Suppose $d > 0$. Fix a large number $C$. Find a segment $J = [a,b]$, $b > a > d/2$, such that

$$\text{mes}\{x \in J : K(x) > C\} > |J|/2.$$

Due to (10), for sufficiently large $N$, one can select from the sequence $\{n(k)\}$ numbers

$$l(0) \leq 2^N < l(1) < l(2) < \dots l(m) < 2^{(N+1)} \leq l(m+1),$$

so that for every $j$, $0 \leq j \leq m$,

$$\frac{b}{a} < \frac{l(j+1)}{l(j)} < 1 + \frac{2|J|}{a}.$$

The left inequality means that segments $J/l(j)$ are pairwise disjoint. The right one implies $(1+2|J|/a)^m > 2$, so $m > 2ca/|J| > cd/|J|$ ($c > 0$ is an absolute constant).

Now, for each $j = 1, 2, \dots, m$,

$$\operatorname{mes}\{x \in J/l(j) : K_{l(j)}(x) > Cl(j)\} > \frac{1}{2}|J|/l(j);$$

so

$$\operatorname{mes}\{x : \max_j K_{l(j)}(x) > C2^N\} > m|J|2^{-N-2} > cd/2^{N+2}.$$

Thus for $\lambda = C2^N$, we have

$$\operatorname{mes}\{x : \max_j K_{l(j)}(x) > \lambda\} > C\frac{cd}{4\lambda}.$$

Taking the approximation of the $\delta$-function $f = \frac{1}{h}1_{[0,h]}$ with sufficiently small $h$, and remembering that all $l(j)$ belong to the sequence $\{n(k)\}$, we obtain

$$\operatorname{mes}\{Mf(x) > \lambda\} > \frac{Ccd}{\lambda}\|f\|.$$

Since $C$ is arbitrarily large, the maximal operator does not satisfy the weak type inequality. $\square$

**Remark 1.** The condition (10) in the theorem is essentially sharp. Indeed, take $K$ of sufficiently slow growth at some point $d \neq 0$, such that $\omega(K;\delta) = O(\delta^\alpha)$ (for some $\alpha > 0$). Then, due to [S], p. 75, $f * K_{2^n}$ converges to $f$ a.e. for any $f \in L^1$.

## 3 Discrete Measures.

In this section we prove the following assertion

**Theorem 3.** *Let $\{\mu_n\}$ be an (AI) sequence of discrete probability measures on the circle $\mathbb{T}$, $\mu_n(0) = 0$. Then there is a function $f \in L^\infty(\mathbb{T})$ such that convolutions $f * \mu_n$ do not converge a.e.*

This result generalizes Bourgain's theorem stated in the introduction. Actually, our observation is that his approach still works if Cesaro averages of translates are replaced by convolutions with arbitrary discrete measures.

The following proposition is a consequence of Bourgain's "bounded entropy principle.

**Proposition 1.** *If for any $r$ one can choose numbers $n(1) < n(2) < \cdots < n(r)$ such that, for any vector $a = \{a(1), a(2), \dots, a(r)\}$, $a(j) = 0$ or 1, there is $m = m(a)$ satisfying*

$$\begin{aligned} &|\widehat{\mu_{n(j)}}(m) - 1| < c/10 \text{ if } a(j) = 0 \\ &|\widehat{\mu_{n(j)}}(m) - 1| > c \text{ if } a(j) = 1, \end{aligned} \tag{11}$$

*($c > 0$ is a constant), then the conclusion of Theorem 3 holds.*

For the proof one needs basically to repeat the argument on p. 95 of [B].

PROOF OF THEOREM 3. Denote by $A_n$ the support of $\mu_n$. One may suppose that it is a finite set for every $n$ (otherwise, replace $\mu_n$ by a sequence of probability measures $\{\mu'_n\}$ with finite supports such that $Var(\mu_n - \mu'_n) < 1/2^n$, which implies equiconvergence of the convolutions almost everywhere).

Notice that, in order to satisfy the condition (11), it is enough to have two increasing sequences of integers $\{n(j)\}$ and $\{m(j)\}$ such that

$$|e^{im(s)x} - 1| < \frac{1}{100} 2^{-|j-s|} \; x \in A_{n(j)}, \; j \neq s, \tag{12}$$

and

$$|\widehat{\mu_{n(j)}}(m(j)) - 1| > 1/3. \tag{13}$$

To check this implication fix $r$ and a vector $a$. Let $m = \sum_{s=1}^{r} a(s)m(s)$. We will use the following elementary inequality for complex numbers $\{z(s)\}$ belonging to the unit disc.

$$|z(1) \dots z(r) - 1| \leq \sum_{s=1}^{r} |z(s) - 1|, \tag{14}$$

which comes by induction.

$$|z(1) \dots z(r) - 1| = |z(1) \dots z(r) - z(1) + z(1) - 1| \leq |z(2) \dots z(r) - 1| + |z(1) - 1|.$$

Now, take $j$ such that $a(j) = 0$. (12) and (14) imply, for $x \in A_{n(j)}$, that $|e^{imx} - 1| < 1/10$ which gives the first condition (11) with $c = 1/10$.

In the same way if $a(j) = 1$ we get $|e^{imx} - e^{im(j)x}| < 1/10$ on $A_{n(j)}$, and (13) implies the second condition (11).

Now, to finish the proof of Theorem 3, we will select inductively subsequences $\{n(j)\}$ and $\{m(j)\}$ satisfying (12) and (13). Suppose that the required numbers are already defined for $j < r$. We have to choose $n(r)$ and $m(r)$ so that the following estimates hold:

(i) $|e^{im(s)x} - 1| < \frac{1}{100} 2^{-|r-s|}, \ x \in A_{n(r)}, \ s < r;$

(ii) $|e^{im(r)x} - 1| < \frac{1}{100} 2^{-|r-j|}, \ x \in A_{n(j)}, \ j < r;$

(iii) $|\widehat{\mu}_{n(r)}(m(r)) - 1| > \frac{1}{3}.$

For $(i)$ it is enough to choose $n(r)$ sufficiently large.

Now, the set $E := \cup_{j<r} A_{n(j)}$ is finite, so for any positive $\epsilon$, one can find $M$ such that any segment of length $M$ contains an integer frequency $m$ satisfying the inequality

$$|e^{imx} - 1| < \epsilon \text{ on } E. \tag{15}$$

Set $\epsilon = 2^{-r}/100$ and fix the corresponding $M$.

Now one can find $d(M) > 0$, such that for any probability measure $\mu$ supported on $(-d, d)$, the oscillation of the Fourier transform on any segment of length $M$ is smaller than $1/10$. Fix $n(r)$ so large that the last property holds for measure $\mu_{n(r)}$. The support of this measure does not contain zero; so averages of its Fourier coefficients $1/N \sum_{0<m\leq N} \widehat{\mu_{n(r)}}(m)$ tend to zero. Therefore, one can find $m'$ for which $|\widehat{\mu_{n(r)}}(m') - 1| > 1/2$. Due to the property above we have the same inequality (with $1/2$ replaced by $1/3$) for any $m \in [m', m'+M]$. Taking in this segment $m$ satisfying (15), we put $m(r) := m$. This choice settles (ii), (iii), and the theorem is proved. □

## 4 Singular Non-Atomic Measures.

First we note that there is a sequence of continuous measures $\{\mu_n\} \in$(AI) with no compactness property. We state this result in a more general form, using the concept of Kroneker set .

A compact $E$ is called a *Kroneker set* if any $f \in C(\mathbb{T})$, $|f(t)| = 1$, admits uniform approximation on $E$ by characters $\{e^{imt}\}$.

**Theorem 4.** *Let $E \subset \mathbb{T}$ be a Kroneker set, and $\{\mu_n\}$ be a sequence of probability measures supported on disjoint portions of $E$. Then there is an $f \in L^\infty$ such that convolutions $f * \mu_n$ do not converge a.e.*

This is a direct consequence of Proposition 1. Indeed, let $\mu_n$ be supported on portions $E_n$ which are pairwise disjoint. Fix $N$ and a vector $a$. Define an unimodular function $f \in C(E)$ , so that $f|_{E_j} = (-1)^{a(j)}$ for all $j \leq N$. Clearly $\int_{\mathbb{T}} f \, d\mu_j = (-1)^{a(j)}$. Now, approximate $f$ by an exponential $e^{imt}$ with a small error in $C(E)$ and we get the result.

**Remark 2.** This shows that an (AI) sequence of "thick" continuous measures may not satisfy the compactness property. Indeed, using the well-known construction of Kroneker sets (see [K], ch. 7) one can produce E such that:

(i) $E$ is a perfect set containing zero;

(ii) By removing from $E$ an arbitrary small neighborhood of zero one gets a Kroneker set;

(iii) Each non-empty portion of $E$ has Hausdorff dimension 1.

Now, it is enough to take a sequence of disjoint portions of $E$ tending to zero and to distribute on each one a probability measure.

Finally we prove a positive result in which the compactness property for one-dimensional singular measures does hold. This result involves the important concept of Fourier Analysis which goes back to D. E. Menshov who first constructed a singular probability measure $\mu$ on the circle satisfying the property

$$\widehat{\mu}(n) = o(1) \ (|n| \to \infty) \tag{16}$$

Such measures are called *Rajchman measures* , see [KL].

**Remark 3.** Having a Rajchman measure $\mu$ it is easy to construct an (AI) family of such measures. Indeed, suppose that the support of $\mu$ contains zero (otherwise we translate). If a function $g$ on $\mathbb{T}$ is smooth, then the measure $\nu$ defined by equality $d\nu = g\,d\mu$ also satisfies the condition (16), due to elementary estimate of the corresponding convolution on $\mathbb{Z}$ . So it is enough to multiply $\mu$ by smooth $g_n$ supported on $(-1/n, 1/n)$, and then to normalize the obtained measures.

**Theorem 5.** *Let $\{\mu_n\}$ be an AI sequence of Rajchman probability measures on $\mathbb{T}$ . Then one can choose $n(1) < n(2) < \dots$ so that, for every $f \in L^2(\mathbb{T})$,*

$$f * \mu_{n(j)} \to f \ a.e. \ (j \to \infty).$$

*Stronger, the maximal operator*

$$(Mf)(x) := \sup_j |(f * \mu_{n(j)})(x)|$$

*is bounded in $L^2$ : $\|Mf\| \leq C\|f\|$ .*

Proof. Consider $\{\widehat{\mu}_(s)\}$, the Fourier coefficients of $\mu_n$.

The (AI) condition implies that, for $|s| < s(n)$, the coefficients are close to 1, and $s(n)$ tends to $\infty$ together with $n$ . On the other hand, due to (16), the coefficients are small for $|s| > s'(n)$. So, by a simple induction, we can define increasing sequences of integers $n(j)$ and $s(j)$ so that

$$
\begin{aligned}
|1 - \widehat{\mu_{n(j)}}(s)| &< 2^{-j} \text{ for } |s| \leq s(j) \\
|\widehat{\mu_{n(j)}}(s)| &< 2^{-j} \text{ for } |s| > s(j+1) \qquad (17) \\
s(j+1)/s(j) &> 2. \qquad (18)
\end{aligned}
$$

Now take $f \in L^2, \|f\| < 1$, and denote by $S_k$ its partial Fourier sums. Consider the convolution $f * \mu_{n(j)}$. We have

$$(f * \mu_{n(j)})(x) = \sum \widehat{f}(s)\widehat{\mu_{n(j)}}(s)e^{isx} = \sum_{|s| \leq s(j)} + \sum_{s(j) < |s| \leq s(j+1)} + \sum_{|s| > s(j+1)} .$$

Then (17) implies that the first sum equals to $S_{s(j)}$ up to an error with the norm $< 2^{-j}$, and the norm of the last one $< 2^{-j}$. So,

$$
\begin{aligned}
f * \mu_{n(j)} =& S_{s(j)} + T_j + R_j, \\
(T_j)(x) :=& \sum_{s(j) < |s| \leq s(j+1)} \widehat{\mu_{n(j)}}(s)\widehat{f}(s)e^{isx}, \ \|R_j\| < 2^{-j}.
\end{aligned} \qquad (19)
$$

Since

$$\|T_j\|^2 = \sum_{s(j) < |s| \leq s(j+1)} |\widehat{\mu_{n(j)}}(s)|^2 |\widehat{f}(s)|^2 \leq \sum_{s(j) < |s| \leq s(j+1)} |\widehat{f}(s)|^2,$$

the series $\sum T_j$ converges in $L^2$. Denote its sum by $F$ . Clearly $\|F\| \leq 1$.

Condition (18) allows one to apply the Kolmogorov theorem on lacunary subsequences (see [Z], ch.13 ), which gives

$$\| \sup_j |S_{s(j)}(x)\| \leq C\|f\| \text{ and } \| \sup_j |T_j|\| \leq C\|F\|.$$

Finally, the last inequality in(19) gives

$$\| \sup_j |R_j|\| \leq \| \sum |R_j|\| \leq 2.$$

So $\|M(f)\| \leq C$, which ends the proof. □

Note that measures $\{\mu_j\}$ in Theorem 5 could be very "thin", since it is known from [I] that for any Hausdorff scaling function $h$, there is a Rajchman measure supported by a compact of $h$-measure zero.

**Remark 4.** The following questions seem to be open:

(i) (D. Preiss; private communication, 1993): to find (if possible) a singular measure $\mu$ on $\mathbb{R}$ (different from the Dirac measure) such that contracted measures $\mu_t$ defined by the equality $\mu_t(A) = \mu(A/t)$, satisfy (1) for $f \in L^2$ .

(ii) To find (if possible) a nontrivial sequence of singular measures $\mu_n \in$(AI) such that $f * \mu_n \to f$ a.e. for any $f \in L^1$.

Note that some further results on the subject above are obtained in [Ko].

## References

[B] J. Bourgain, *Almost sure convergence and bounded entropy*, Israel Journal of Math., **63, No.1** (1988), 79–97.

[Br] A. M. Bruckner, *Differentiation of integrals*, AMS Math. Monthly, **78, No. 9** (1971), Part II, ii+51.

[I] O. S. Ivashev-Musatov, *M-sets and Hausdorff measure* (in Russian), DAN USSR, **142, No. 5** (1962), 1001–1004.

[K] J-P. Kahane, *Series de Fourier Absolument Convergentes*, Springer-Verlag, 1970.

[KL] A. Kechris, A. Louveau, *Descriptive set theory and the structure of sets of uniqueness*, Cambridge University Press, 1987.

[Ko] S. Kostyukovsky, *Convergence almost everywhere of convolution operators*, (in preparation).

[N-S] A. Nagel, E. M. Stein, *On certain maximal functions and approach regions*, Advances in Math., **54** (1984), 83–106.

[N] I. P. Natanson, *Theory of functions of a real variable*, Rev. Ed. N-Y, Ungar, 1961.

[S] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton, N.J., Princeton Univ. Press, 1993.

[Z] A. Zygmund, *Trigonometric series*, Cambridge Univ. Press, 1959.