London Mathematical Society Lecture Note Series. 114

Lectures on Bochner-Riesz Means

Katherine Michelle Davis
The University of Texas at Austin

Yang-Chun Chang
The University of Texas at Austin and
Beijing Normal University

The right of the University of Cambridge to print and sell all manner of books was granted by Henry VIII in 1534. The University has printed and published continuously since 1584.

CAMBRIDGE UNIVERSITY PRESS

Cambridge

New York New Rochelle Melbourne Sydney

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
32 East 57th Street, New York, NY 10022, USA
10, Stamford Road, Oakleigh, Melbourne 3166, Australia

First published 1987

Library of Congress cataloging in publication data available

British Library cataloguing in publication data

Davis, Katherine Michelle
Lectures on Bochner-Riesz means. – (London Mathematical Society lecture note series, ISSN 0076-0552; 114)
1. Fourier series 2. Convergence
3. Bochner-Riesz means
I. Title II. Chang, Yang-Chun III. Series
515'.2433 QA404

ISBN 0 521 31277 9

CONTENTS

CHAPTER 0

0.1 Introduction

In this section we give a quick discussion of the main topic of the book: when does a Fourier series converge? We let T^n denote the n-torus $[0,1)^n$, with elements $\theta = (\theta_1, \theta_2 \ldots, \theta_n) \in \mathrm{T}^n$, and $k = (k_1, k_2, \ldots, k_n)$ an n-tuple of integers. If f is in $L^p(\mathrm{T}^n)$, then the Fourier transform and the formal Fourier series associated to f are defined as

$$\hat{f}(k) = \int f(\theta) e^{-2\pi i k\theta} d\theta$$

$$f = \sum_{k \in \mathbb{Z}^n} \hat{f}(k) e^{2\pi i k\theta}.$$

The central topic of this book is an analysis of the sense in which the formal Fourier series of f actually converges to f. Of course we have to take finite sums, and then a limit. Since the sum is over all n-tuples of integers, that is, over all $k \in \mathbb{Z}^n$, we need some method to order the lattice points of $\mathbb{Z}^n$. The simplest is to include them all in ever-expanding spheres; define $|k|^2 = k_1^2 + k_2^2 \ldots + k_n^2$, and then define the R^{th} spherical partial sum of the Fourier series of f as

$$S_R f(\theta) = \sum_{|k| < R} \hat{f}(k) e^{2\pi i k\theta}.$$

With this definition, the basic question we want to study is: when does $S_R f$ converge to f in L^p?

In one dimension, the answer is classical, and was found by M. Riesz in 1910. (See Zygmund [**57**], Chapter 7, 2.4). Convergence is valid for all f in L^p if and only if $1 < p < \infty$.

In higher dimensions, the question was open until very recently. Carl Herz showed in 1954 that a necessary condition for convergence is that $\frac{2n}{n+1} < p < \frac{2n}{n-1}$. Charles Fefferman showed in 1972 that convergence in L^p holds if and only if $p = 2$. Fefferman's proof and its consequences are the main subject of this book.

Fefferman's result shows that this is not the right way to sum Fourier series. What other ways are there? Well, the points of $\mathbb{Z}^n$ could be grouped in some order other than taking whatever is inside a sphere. Concentric polygons are an obvious thing to try, but this turns out to be no more interesting than repeating several one-dimensional results. It doesn't give any new mathematics, and it avoids having to think deeply about Fefferman's result. To avoid thinking about a subject is almost always a mistake; at best you are in for some big suprises later on.

A classical analyst would quickly tell you the alternative to using spherical partial sums: use a summability method. We shall analyse a method introduced by S. Bochner, which itself was a variation of a summability method of Riesz: Set $(\xi)_+ = \xi$ if $\xi > 0$; let it be 0 if $\xi < 0$. Then we define Bochner-Riesz means of order α by

$$S_R^\alpha f(\theta) = \sum_{|k|<R} \left[\left(1 - \frac{|k|^2}{R^2}\right)_+ \right]^\alpha \hat{f}(k) e^{2\pi i k\theta}.$$

The point here is that if $\alpha = 0$, $S_R^\alpha f = S_R f$, so that by studying the limiting behaviour as α tends to zero, we can hope to understand what's wrong with S_R.

Very well; what is known about S_R^α for α near zero? The behaviour turns out to be very complicated in high dimensions, and in fact the complete answer is known only in two dimensions. Instead of trying to write down a formula for the answer, we'll draw a picture. Figures 1 and 2 below show the L^p boundedness of S_R^α for $1 \leq p \leq \infty$. The vertical axis is indexed by α; the horizontal by $\frac{1}{p}$. Dotted lines and open circles represent points of known unboundedness; shaded regions known boundedness.

The purpose of this book is to give detailed proofs of the results in the pictures.

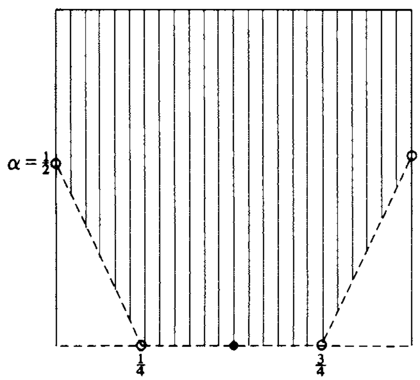

Figure 1

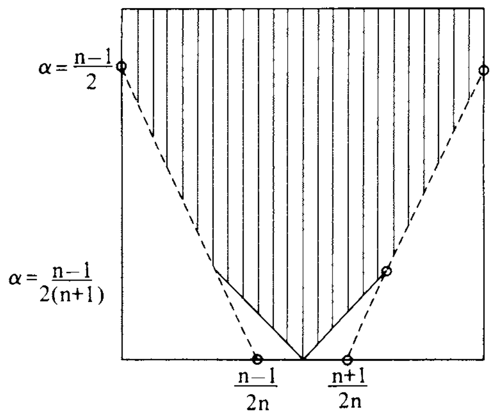

Figure 2

0.2 The Role of Functional Analysis

In this section we give an idea how convergence questions can be attacked. The best way to do this is through a fast review of a standard result: there is a continuous function in T whose Fourier series diverges at $\theta = 0$. Of course this function can be explcitly constructed, but in general dimension explicit computations are unmanageable. We turn for relief to the methods of functional analysis.

The abstract problem is this. If I take a finite Fourier series, the partial sums of its Fourier series simply stop after a while, and I get the original function back. So, partial sums of Fourier series converge on the so-called trigonometric polynomials. The trig polynomials are dense in the continuous functions (Stone-Weierstrass Theorem) and the continuous functions are dense in L^p. I need to pass from information about convergence of partial sums on a dense subset to convergence on a whole space of functions: I need to interchange limits. Standard real analysis tells me that some sort of uniform convergence allows interchange of limits; functional analysis tells me that uniformity is a necessary condition. This is the point of the uniform boundedness theorem.

0.1 Theorem. *There is an* $f \in C(\mathbf{T})$ *such that* $\sup_N |S_N f(0)| = \infty$.

Proof: We begin by changing this into a problem about operators on function spaces. Define the operators $S_N : C(\mathbf{T}) \to \mathbb{C}$ by $S_N f = \sum_{|k| \leq N} \hat{f}(k)$. Since the sum is finite, each S_N is a countinous linear functional on $C(\mathbf{T})$. Every such linear functional is given by integration against a finite Borel measure, in this case,

$$\sum_{|k|\leq N} \hat{f}(k) = \sum \int e^{-2\pi i k\theta} f(\theta) d\theta = \int K_N(\theta) f(\theta) d\theta;$$

here

$$K_N(\theta) = \sum_{|k|\leq N} e^{-2\pi i k\theta} = \frac{\sin 2\pi(N+\frac{1}{2})\theta}{\sin(\pi\theta)}.$$

So in this case, the measure giving S_N is absolutely continuous, and its total variation norm is just $\|K_N\|_1$. This is not even very difficult to compute; in Zygmund [**57**] the precise computation is given as

$$\|K_N\|_1 = \frac{2}{\pi}\log N + O(1).$$

The real problem is to relate how S_N acts on individual functions with how it behaves as an operator on the space $C(\mathbf{T})$. This is what the uniform boundedness theorem tells us. Either: i) The S_N are uniformly bounded in operator norm; or, ii) $\sup_N |S_N f| = \infty$ for a dense set of f. Since the first conclusion does not hold, the second does.

This result is absolutely typical of what we will do in the rest of the book. There will be some functional analysis trickery that reduces convergence questions to questions on the boundedness of operators on function spaces. The functional analysis is followed by a computation of specific operators; and the analysis concludes with a detailed computation.

0.3 Background

In the next chapter, we will present the functional analysis needed to analyze convergence of Fourier series. The best trick will be to transfer problems from Fourier series to Fourier integrals. This is good because it is easier to compute an integral explicitly than to sum a series in a closed form. On the other hand, this is bad because the integrals defining Fourier transforms do not converge absolutely. The modern solution to this difficulty is to use a dense subset on which everything does converge, and then pass to a limit. This is the point of the theory of rapidly decreasing functions, which we summarize in this section. A more detailed treatment is given in Stein and Weiss [**50**].

For $f \in L^1(\mathbb{R}^n)$, the Fourier transform $\hat{f}$ is defined as

$$\hat{f}(\xi) = \int f(x) e^{-2\pi i x\xi} dx.$$

The integral converges absolutely, and $\|\hat{f}\|_\infty \leq \|f\|_1$. To define $\hat{f}$ for $f \in L^2$, we need to use trickery and deceit.

The space of rapidly decreasing or Schwartz functions is denoted by $\mathcal{S}$ and is defined as the class of all smooth functions f on $\mathbb{R}^n$ for which the seminorms $\sup_x |x^\alpha D^\beta \phi(x)|$ are finite. We've used multiindex notation here; if we let α denote the multiindex $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$; then $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. As these seminorms vary with α, they give rise to a topology on $\mathcal{S}$. The dual of $\mathcal{S}$ is then called the space of tempered distributions. Then it is not hard to prove:

0.2 THEOREM.

a) $\mathcal{S}$ is dense in L^p if $1 \leq p < \infty$.

b) The Fourier transform is a continuous, one to one map of $\mathcal{S}$ onto $\mathcal{S}$.

c) If the convolution of two functions f and g in $\mathcal{S}$ is defined as

$$f \star g = \int f(x-y)g(y)dy,$$

then this function is again in $\mathcal{S}$ and $\widehat{f \star g} = \hat{f}\hat{g}$.

d) If f and g are in $\mathcal{S}$, then $\int f\bar{g} = \int \hat{f}\bar{\hat{g}}$

e)

$$\|f\|_2 = \|\hat{f}\|_2$$

f) The Fourier transform of a distribution u is defined by $\hat{u}(f) = u(\hat{f})$. Then $\hat{u}$ is again a distribution.

g) The translation operator τ_y is defined by $(\tau_y f)(x) = f(x-y)$. If convolution of a distribution u and a function in $\mathcal{S}$ is defined as $(u \star f)(x) = u(\tau_x \tilde{f})$, where $\tilde{f}(y) = f(-y)$, then $u \star f$ is a smooth function.

The space $\mathcal{S}$ is a technical tool which makes it easy to do computations with integrals that might otherwise be infinite. We will also need some notation for the rest of the book. If T is a bounded linear operator from L^p to L^p, we denote its operator norm by ${}_p\|T\|$. Recall then that

$${}_p\|T\| = \sup\{\|Tf\|_p \mid \|f\|_p = 1\}$$

and that

$${}_p\|T\| = \sup\{\int Tfg \mid \|f\|_p = \|g\|_{p'} = 1\},$$

where $\frac{1}{p} + \frac{1}{p'} = 1$.

0.4 NOTES FOR CHAPTER 0

0.1): Bochner-Riesz means are a variant of Riesz means, $(1-|n|^2)_+^\alpha$; we shall see that the two summability methods give the same results in L^p, but that Bochner-Riesz means are much easier to compute with. The original interest of the S_R^α was in conection with pointwise summability and the localization of Fourier series (a topic now in disrepute due to its difficulty). Bochner's paper [**1**] showed that n-dimensional problems are deeper than one-dimensional; localization did not hold for all α, in contrast to one-dimensional results. The essential ideas are also treated in Stein and Weiss [**50**] Chapter 7 Theorem 4.2.

There are good reasons for grouping Fourier coefficients together in concentric spheres rather than in concentric polygons. In Bochner's words([**1**] pps. 179-80):

"The elementary exponentials $u(x) = e^{i(n_1x_1+\dots+n_kx_k)}$(all $n_1,\dots,n_k$ integers) are a complete set of regular solutions of the characteristic value problem

$$\Delta u(x) = -\lambda u(x),$$

if this system is being considered on the (closed) torus

$$0 \leq x_1 \leq 2\pi,\dots,0 \leq x_k \leq 2\pi,$$

and Δ is the Laplace operator with respect to the Euclidean metric on the torus, namely

$$\frac{\partial^2 u}{\partial x_1^2} + \dots + \frac{\partial^2 u}{\partial x_k^2}.$$

Since $\lambda = n_1^2 + \dots + n_k^2$ our way of writing series . . . . satisfies the very natural principal of ordering the terms in a series according to the magnitude of the characteristic eigenvalues λ."

To paraphrase Bochner, in looking at spherical summability methods, we are really looking at the question of the convergence of the eigenfunction expansion of a differential operator on a compact Riemannian manifold. C. Fefferman's result has been used to show that if $n > 1$, such expansions never converge. This is our main point: the divergence of spherical means are not an abberation caused by a crazy choice of summability methods, but is a fundamental fact of high dimension geometry.

The results summarized in Figures 1 and 2 are due to many individuals. See Herz [**27**], C. Fefferman [**23**], Carleson and Sjolin [**5**], and Tomas [**51**] for a summary.

0.2): The computation of the $\|K_R\|_1$, which are called the Lebesgue constants, is due to Fejer (1910). These sort of techniques fail miserably in higher dimensions, because the series for $K_R(\theta)$ cannot be summed. It is possible to show that $\|K_R\|_1 = 0(R^{\frac{n-1}{2}})$ if $n \geq 2$, which is the best possible estimate; see Shapiro [**44**] for the proof and references to earlier results. In trying to get an explicit formula for the partial sums K_R, you quickly realize that a really good formula would allow you to compute exactly the number of integer points inside a ball of radius R in $\mathbb{R}^n$. Unfortunately, getting good estimates on this is one of the hardest problems in number theory. It is a serious problem because for some R, there are no pairs of integers (m, n) with $m^2 + n^2 = R^2$. But for a slightly different R, there will be lots of pairs. You can't expect any really regular expression that would follow from a good formula for K_R.

The classical, detailed proofs of the divergence of Fourier series at a point are best found in Zygmund [**57**], Chapter 7 Volume 1.

CHAPTER 1 MULTIPLIER THEORY

SECTION 1.1 MULTIPLIERS ON L^p

There are problems "summing" Fourier integrals; the purpose of this section is to see what the problems are and how they can be overcome.

If $f \in L^2$, the expected inverse Fourier transform

$$\check{f} = \int f(\xi) e^{2\pi i x \xi} d\xi$$

simply does not converge. The traditional remedy is to look instead at "partial sums" : integrals over bounded regions; this forces the integral to converge. One then hopes to take a limit and have convergence of the limit. We define the partial sums operators S_R by

$$S_R f(x) = \int_{|\xi| \leq R} \hat{f}(\xi) e^{2\pi i x \xi} d\xi,$$

which does converge absolutely. We hope that for all $f \in L^2$, $lim_{R \to \infty} S_R f$ exists in L^2, and that the limit is f. The first problem is that we don't even know that $S_R f$ is in L^2, let alone convergent in L^2. So we begin with the functional analysis of the operators S_R: do they take L^2 functions into L^2 functions? We let B denote the unit ball in $\mathbb{R}^n$, and let $\chi_B\left(\frac{\xi}{R}\right)$ be the characteristic function of the ball of radius R. It follows that

$$S_R f(x) = \int \chi_B \left(\frac{\xi}{R}\right) \hat{f}(\xi) e^{2\pi i x \xi} d\xi.$$

These are the operators we intend to study.

1.1 DEFINITION: A Fourier multiplier operator on $L^p(\mathbb{R}^n)$ is a linear operator T bounded on L^p for which there exists a $\mu \in L^\infty(\mathbb{R}^n)$ satisfying

$$Tf(x) = \int \mu(\xi) \hat{f}(\xi) e^{2\pi i x \xi} d\xi$$

for all Schwartz functions f. In this case, μ is called the Fourier multiplier associated to the operator T.

REMARK: The operator is really defined by a functional equation: $\widehat{Tf} = \mu \hat{f}$, which is very indirect. To actually work with multiplier operators, we need characterizations which stay on one side: either the function side or the Fourier transform side.

1.2 PROPOSITION. *Assume T is a Fourier multiplier of L^p. Then there is a distribution K for which*

$$Tf(x) = K \star f(x),$$

for all $f \in \mathcal{S}$. K is called the convolution kernel of T.

PROOF: To see what is going on, take $\mu \in \mathcal{S}$. Then

$$Tf(x) = \int \mu(\xi) e^{2\pi i x\xi} \hat{f}(\xi) d\xi = \int \mu(\xi) e^{2\pi i x\xi} \int e^{-2\pi i \xi y} f(y) dy$$

$$= \int \int \mu(\xi) e^{2\pi i \xi(x-y)} d\xi f(y) dy = \int \check{\mu}(x-y) f(y)\, dy = \check{\mu} \star f(x).$$

Because everything in sight was a Schwartz function, we could rearrange the integrals, and we took the distribution K to be the function $\check{\mu}$. To give a general proof, we need a lot of distribution facts. First of all, if $f \in \mathcal{S}$, $\hat{f} \in \mathcal{S}$ and, since $\mu \in L^\infty$, $\mu \hat{f} \in L^1$, so that

$$Tf(x) = \int e^{2\pi i x\xi} \mu(\xi) \hat{f}(\xi)\, d\xi$$

exists for all x.

The distribution K is defined by $K(f) = (T\tilde{f})(0)$. The integral representation for f shows that this is a distribution: $f \to \hat{f}$ is a continuous map of $\mathcal{S}$ into $\mathcal{S}$, and integration against an L^∞ function is well known to be a distribution. To finish, notice that

$$(K \star f)(x) \equiv K(\tau_x \tilde{f}) \equiv T(\widetilde{\tau_x \tilde{f}})(0)$$

$$= T(\tau_{-x} f)(0) = \int \mu(\xi) \widehat{\tau_{-x} f}(\xi) d\xi = \int \mu(\xi) e^{2\pi i x\xi} \hat{f}(\xi) d\xi \equiv Tf(x);$$

the next to the last equality holds because

$$\widehat{\tau_x f}(\xi) = \int e^{-2\pi i y\xi} (\tau_x f)(y) dy = \int e^{2\pi i x\xi} \hat{f}(y).$$

1.3 REMARKS:

a) As distributions, $\hat{K}$ is μ and $\check{\mu}$ is K.

b) The distribution K could be very bad; look at the operator $Tf = f$. Here the Fourier inversion formula tells me that $\mu \equiv 1$. What distribution gives this? The distribution corresponding to T is the Dirac delta measure, $\delta(f) = f(0)$. Working out the formalities,

$$(\delta \star f)(x) \equiv \delta(\tau_x \tilde{f}) = (\tau_x \tilde{f})(0) = \tilde{f}(-x) = f(x).$$

The moral is that even trivial operators can generate distributions that are not even given by integration against functions. To make things worse, we will now sketch an example where K is not even given by integration against a finite measure. Let μ be the L^∞ function $isign(\xi)$, then we will compute in 2.13 that $\check{\mu}(x) = \frac{1}{\pi x}$. But let's face it, $\check{\mu}(f) = \frac{1}{\pi} \int \frac{1}{x} f(x) dx$ normally does not converge. Properly speaking, we have to take a principal value integral integral,

$$\lim_{\epsilon \to 0} \frac{1}{\pi} \int_{|x| \geq \epsilon} \frac{1}{x} f(x) dx.$$

This sort of distribution should be expected when the multiplier is not an integrable function.

1.4 PROPOSITION. *If T is a multiplier of L^p and $f \in \mathcal{S}$, then*

$$\widehat{(Tf)}(\xi) = \mu(\xi)\hat{f}(\xi)$$

PROOF: To see this, we have to see that as distributions, $\widehat{Tf}$ and $\mu(\xi)\hat{f}(\xi)$ act the same. Pointwise equality follows immediately, since both distributions are given by integration against functions; that was the whole point of taking $f \in \mathcal{S}$; then $\hat{f} \in \mathcal{S}$ and $\mu\hat{f} \in L^1$. It follows that Tf is even continuous. Ok, checking equality by integrating against a function $g \in \mathcal{S}$,

$$\int \widehat{Tf}g = \int Tf\hat{g},$$

from 0.2 extended to all of L^2. But

$$\int \mu\hat{f}g = \int\int \mu\hat{f}(g) = \int(\mu\hat{f})\check{g} = \int Tf\hat{g}.$$

The triple (T, K, μ), is called a multiplier triple. T is the multiplier operator, K the convolution kernel, and μ the multiplier. In the rest of the book we will often use phrases like " Let T be an operator with multiplier μ", or, " If μ is a multiplier with kernel K".

1.5 PROPOSITION. *Let T be a multiplier operator on L^2, with multiplier μ. Then*

$${}_2\|T\| = \|\mu\|_\infty.$$

PROOF: From the Plancherel theorem,

$$\|Tf\|_2 = \|\widehat{Tf}\|_2 = \|\mu\hat{f}\|_2$$

$$\leq \|\mu\|_\infty\|\hat{f}\|_2 = \|\mu\|_\infty\|f\|_2 \ ,$$

so that ${}_2\|T\| \leq \|\mu\|_\infty$. Now we'll show that if $\epsilon > 0$, then ${}_2\|T\| \geq \|\mu\|_\infty - \epsilon$.

From the definition of L^∞, the set $E = \{\xi \mid |\mu(\xi)| \geq \|\mu\|_\infty - \epsilon\}$ has positive measure; we can choose a subset S of E with positive but finite measure. Then we can compute the norm of this function: $\chi_S \in L^2$, and $\|\chi_S\|_2 = \|\hat{\chi}_S\|_2 = |S|^{\frac{1}{2}}$. We also can compute the norm of T applied to this function:

$${}_2\|T\|\ \|\check{\chi}_S\|_2 = {}_2\|T\|\ |S|^{\frac{1}{2}} \geq \|T(\check{\chi}_S)\|_2 = \|\mu\chi_S\|_2$$

$$\geq \inf|\mu(\xi)\chi_S(\xi)|\|\chi_S\|_2 \geq (\|\mu\|_\infty - \epsilon)|S|^{\frac{1}{2}} = (\|\mu\|_\infty - \epsilon)\ \|\hat{\chi}_S\|_2.$$

1.6 TECHNICAL REMARK: We used the result that $\widehat{T(\check{\chi}_S)} = \mu\chi_S$. We really only proved this if $f \in \mathcal{S}$, but the characteristic function of a set S is certainly not in $\mathcal{S}$. This is typical of the petty technical problems that plague this subject, all due to the fact that our multipliers are only defined on a dense set of L^p, and extended to the remainder by continuity. In this remark we want to show how to treat a typical petty problem.

For $f \in L^2$, choose $f_j \in \mathcal{S}$ such that f_j converges to f in L^2. Since the Fourier transform is continuous on L^2, $\hat{f}_j$ converges to $\hat{f}$ in L^2. Since T is a bounded(that is, continuous) operator on L^2, Tf_j converges in L^2 to Tf, and $\widehat{Tf_j}$ converges to $\widehat{Tf}$ in L^2. We may take a subsequence j_k such that $\widehat{Tf_{j_k}}$ converges and $\hat{f}_{j_k}$ converges pointwise ae to $\widehat{Tf}$ and to $\hat{f}$. Then, for almost all ξ,

$$\widehat{Tf}(\xi) = \lim_{k\to\infty} \widehat{Tf}_{j_k}(\xi) = \lim_{k\to\infty} \mu(\xi)\hat{f}_{j_k}(\xi) = \mu(\xi)\hat{f}(\xi).$$

1.7 PROPOSITION. *T is a multiplier of L^1 if and only if the convolution kernel K is a finite measure, dm, and in that case*

$${}_1\|T\| = \|dm\|.$$

PROOF: This result has a simple intuition. The Dirac delta function δ is almost an L^1 function, and it also acts like the identity operator. Then ${}_1\|T\|$ should be like $\|K \star \delta\|_1 = \|K\|_1$ This proof shows how to handle the "minor" technical annoyance that δ is not a function in L^1. If we assume that K acts as a finite measure, then

$$\|K \star f\|_1 = \|K(\tau_x \tilde{f})\|_1 = \| \int (\tau_x \tilde{f})(y) dm(y)\|_1$$

$$\leq \int \|\tau_x \tilde{f}\|_{1,y} d|m|(y) = \int \|f(y-x)\|_1 d|m|(y) = \|f\|_1 \|dm\|.$$

It follows that ${}_1\|T\| \leq \|dm\|$.

For the other direction, I need to get at the δ function even though it is not an L^1 function. We'll use an absolutely standard technique: instead of using δ, we approximate it by functions which are in L^1. We choose Gaussian kernels:

$$\omega_t(x) = (2\pi t)^{-n} \exp\left[-(|x|/2t)^2\right];$$

we rigged it so that $\|\omega_t\|_1 = 1$, $\omega_t \in \mathcal{S}$, and so that for every $f \in \mathcal{S}, \omega_t \star f$ converges in $\mathcal{S}$ to f as t tends to zero. Of course we get

$$\|T(\omega_t)\|_1 \leq {}_1\|T\| \|\omega_t\|_1 = {}_1\|T\|;$$

the question is whether we get $T(\omega_t) = K \star \omega_t$ converging to dm. If K were a function in $\mathcal{S}$, this would be no problem; here all we know is that the $T(\omega_t)$ are uniformly bounded in L^1. To get from this sort of information(boundedness) to convergence information is clearly some sort of compactness condition. Unfortunately, the unit ball in L^1 is not sequentially compact, so we resort to trickery and deceit.

The $T(\omega_t)$ are in L^1, and they can be viewed as finite measures with total variation norm bounded by 1. Amazingly enough there is a topology on the space of finite measures, $\mathcal{M}$ under which closed, bounded sets are sequentially compact. We view $\mathcal{M}$ as the dual of $\mathcal{C}_0$, the space of countinuous functions vanishing at infinity. The topology we put on $\mathcal{M}$ is called the weak-$\star$ topology: μ_j converges to μ means that $\mu_j(f)$ converges to $\mu(f)$ for all $f \in \mathcal{C}_0$. Then Alogolu's theorem (Rudin [**43**]) asserts the needed compactness: there is a subsequence $T(\omega_{t_j})$ which converges weak- $\star$ to some finite measure dm in the ball of radius ${}_1\|T\|$. We've gotten as far as knowing what dm is, and we even know that the total variation norm of dm is bounded by ${}_1\|T\|$. The only little detail lacking in this sweet picture is the statement that the distribution K is given by integration against dm. This is the role of the weak-$\star$ convergence. Since $\mathcal{S} \subset \mathcal{C}_0$, for every $f \in \mathcal{S}$,

$$dm(f) \equiv \int f dm = \lim_{j\to\infty} \int f T(\omega_{t_j}) = \lim_{j\to\infty} \int f K \star \omega_{t_j}$$

$$= \lim_{j\to\infty} \int f(x) K \left(\tau_x \tilde{\omega}_{t_j}(y)\right) dx = \lim_{j\to\infty} K \left(\int f \omega_{t_j}(y-x) dx \right).$$

The last equality is true because the integral can be approximated by a Riemann sum. The approximation will converge in $\mathcal{S}$ and so we interchanged the approximation and the distribution. Finally,

$$\lim_{j\to\infty} K\left(\omega_{t_j}\right) \star f = K \left(\lim_{j\to\infty} \omega_{t_j} \star f \right) = K(f).$$

This is called a "regularity" result; at first we only knew that K was some potentialy awful distribution. In fact, it had to be a much nicer object, a finite measure. How did a distribution suddenly get forced to be a measure? The boundedness of an L^1 norm forced a sequence of smooth approximations to converge. So this is our first example of how L^p boundedness of operators forces regularity; once again note the role of functional analysis.

1.8 LEMMA. *If T is a multiplier of L^p, then it is a multiplier of $L^{p'}$, where $\frac{1}{p}+\frac{1}{p'}=1$.*

PROOF: To see the idea here, remember that $L^{p'}$ is the dual space of L^p, so that the boundednes of the operator T on L^p implies boundedness of the adjoint operator $T^\star$ on $L^{p'}$. Our goal is to relate T and $T^\star$. If T were given by convolution with some sort of reasonable function, say a nice Schwartz function K, then we could compute:

$$\int Tf\bar{g} = \int K\star f\bar{g} = \int\int K(x-y)f(y)\bar{g}(x)dydx$$

$$= \int\int K(x-y)\bar{g}(x)f(y)dydx = \int\left[K\star(\tilde{\bar{g}})\right]f(y)dy$$

$$= \int f(y)\widetilde{K\star(\tilde{\bar{g}})}(y)dy.$$

This computation means that if we define the isometry J of L^p by $Jf=\bar{\tilde{f}}$, then $T^\star = JTJ$. Since J is an isometry, we get

$${}_p\|T\| = {}_{p'}\|T^\star\| = {}_{p'}\|T\|.$$

The real problem will be in changing the order of integration when K is not a nice Schwartz function. We will do some unpleasant computations with distributions, following the intuition we just gave.

$$\int Tf\bar{g} = \int K\star f(x)\bar{g}(x)dx = \int K\left(\tau_x(\tilde{f})(y)\right)\bar{g}(x)dx$$

$$= K\left(\int(\tau_x\tilde{f})(y)\bar{g}(x)dx\right) = K\left(\int f(x-y)\bar{g}(x)dx\right)$$

$$= K\left(\int f(x)\bar{g}(x+y)dx\right) = K\left(\int f(x)(\tau_{-x}\bar{g})(y)dx\right)$$

$$= \int f(x)\left[K\star\tilde{\bar{g}}(x)\right]dx.$$

Etc, etc.

1.10 LEMMA. *If $K\in L^1$, then $\|K\star f\|_p \le \|K\|_1\|f\|_p$.*

PROOF: As in the proof of 1.7,

$$\|K\star f\|_p = \|\int K(y)f(x-y)dy\|_p$$

$$\le \int|K(y)|\ \|f(x-y)\|_{p,x}dy = \|K\|_1\|f\|_p.$$

1.11 REMARK: We now have a pretty complete understanding of multipliers on L^1 and L^2; we would like to get a better understanding of L^p for $1<p<2$. By duality, that is, by 1.8, we would get a grasp of multipliers of L^p for $2<p<\infty$.

Think about the identity operator, bounded on all L^p. If we understand a function in L^1 and in L^2, why should that tell us anything about the function in L^p for $1 < p < 2$? What is needed here is some sort of convexity: if you think of things graphically, the operator norm of T on L^p should lie below the norm for T on the line between L^1 and L^2. This happens not to be true. Look at Holder's inequality:

$$\int |f|^{tp_0+(1-t)p_1} \leq \left(\int |f|^{p_0}\right)^t \left(\int |f|^{p_1}\right)^{1-t} ;$$

this tells me that $log\left(\int |f|^p\right)$ is a convex function of p.

1.12 THEOREM. *Let $S = \{z \in \mathbb{C} \mid 0 \leq Re(z) \leq 1\}$. Let F be bounded and continuous on S, and analytic in the interior of S. Let $k_x = \sup_y |F(x+iy)|$. Then* $\log k_x$ *is a convex function of* x.

IDEA OF PROOF: This is a standard result from complex variable theory; the first thing to do is to rescale F to be bounded by 1 on $0+iy$ and on $1+iy$; the rescaled function is $G(z) = F(z)k_0^{-z}k_1^{1-z}$. Then we try to prove $|G| \leq 1$ everywhere. If G vanishes as $|y| \to \infty$, we rig up a rectangle on whose boundary $|G| \leq 1$, outside of which $|G| \leq 1$. Inside, we apply the maximum principle. Now if G does not vanish at infinity, we use the boundedness of G to multiply by an exponential tending to zero. As we let the exponential go to infinity more slowly, we recover G. It follows that $G(z)$ is everywhere bounded by 1, and therefore that $|F(z)| \leq k_0^z k_1^{z-1}$. Taking the supremum and then the log gives the right result.

1.13 THEOREM. *Let T be a linear operator bounded on L^{p_i}, $i = 0, 1$. Then* $\log\ {}_p\|T\|$ *is a convex function on* $\left[\frac{1}{p_0}, \frac{1}{p_1}\right]$. *That is, if $\frac{1}{p} = \frac{t}{p_0} + \frac{1-t}{p_1}$ for some t with $0 \leq t \leq 1$, then*

$${}_p\|T\| \leq {}_{p_0}\|T\|^t\ {}_{p_1}\|T\|^{1-t} .$$

PROOF: Clearly we want to use 1.12, and take $F(z) = \int |Tf|^z$. This runs into technical problems at any x for which $Tf(x) = 0$, and I can't raise $T(f)$ above 0 by adding on ϵ without destroying finiteness of the integrals. Then too, this doesn't get me a look at the ${}_p\|T\|$. I fix all these problems in one step: I'll test the size of $T(f)$ by integrating it against an $L^{p'}$ function, and I'll avoid the zeroes problem by evaluating T on simple functions. Thus, let f, g be simple functions with

$$\|f\|_p = \|g\|_{p'} = 1,$$

$$f(x) = \sum |a_j| e^{i\theta_j} \chi_{E_j}; \quad g(x) = \sum |b_k| e^{i\phi_k} \chi_{F_k}.$$

Here E_j, F_k are sets of finite measure; we will show that

$$\int Tfg \leq {}_{p_0}\|T\|^t\ {}_{p_1}\|T\|^{1-t}.$$

First let $\alpha(z) = \frac{z}{p_0} + \frac{1-z}{p_1}$. This maps the $\frac{1}{p}$ interval into $[0, 1]$. Then let

$$F(z) = \int T\left(\sum |a_j|^{p\alpha} e^{i\theta_j} \chi_{E_j}\right) \left(\sum |b_k|^{p'(1-\alpha)} e^{i\phi_k} \chi_{F_k}\right).$$

This is not the same as raising Tf to the z power, but it is an analytic function of z in the strip S, and bounded and continuous. Note that

$$|F(x+iy)| \leq {}_{p_0}\|T\| \; \|\sum |a_j|^{p\alpha} e^{i\theta_j} \chi_{E_j} \|_{p_0} \; \|\sum |b_k|^{p'(1-\alpha)} e^{i\phi_k} \chi_{F_k} \|_{p'}$$

But

$$\|\sum |a_j|^{p\alpha} e^{i\theta_j} \chi_{E_j} \|_{p_0} = \|\sum |a_j|^{Re(p\alpha)} \chi_{E_j} \|_{p_0}$$

because of the disjointness of the sets E_j. Now

$$Re(\alpha p) = pRe\left[\frac{1-iy}{p_0} + \frac{iy}{p_0}\right];$$

it follows that

$$\|\sum |a_j|^{Re(p\alpha)} \chi_{E_j} \|_{p_0} = \|\sum |a_j|^{\frac{p}{p_0}} \chi_{E_j} \|_{p_0}$$

$$= \|\sum |a_j| \chi_{E_j} \|_p^{\frac{p}{p_0}} = \|f\|_p^{\frac{p}{p_0}} = 1.$$

Similar estimates apply to the second term, and it follows that

$$|F(0+iy)| \leq {}_{p_0}\|T\|; \; |F(1+iy)| \leq {}_{p_1}\|T\|.$$

The Theorem now follows from 1.12.

REMARK: This result is called the Riesz-Thorin interpolation theorem.

1.14 THEOREM. *If* $1 \leq p \leq 2$, $\|\hat{f}\|_{p'} \leq \|f\|_p$.

PROOF:

$$\|Tf\|_\infty = \|\hat{f}\|_\infty \leq \|f\|_1$$

$$\|Tf\|_2 = \|\hat{f}\|_2 = \|f\|_2$$

These two cases show that T maps L^p into $L^{p'}$ for the special cases of $p = 1$ and $p = 2$. The p, p' relationship is linear; interpolation preserves it. The map has norm bounded by 1 on each of the endpoint spaces, and interpolation preserves this norm as well.

1.15 REMARKS:

a) Initially, the Fourier transform was defined only for L^1 functions, because only there did the integrals converge absolutely. The Plancherel theorem allowed us to extend the operator to L^2, but if $f \in L^2$, $\hat{f}$ is only defined almost everywhere. The result above is called the Hausdorff-Young theorem, and it allows us to extend the definition to L^p for $1 \leq p \leq 2$. Thus, the Fourier transform of a function in L^p is the limit in $L^{p'}$ of the Fourier transform of Schwartz functions.

b) The Hausdorff-Young inequality is pretty much the best we can do, at least if size is measured by L^p norms. The following trick demonstrates it: let $f_\delta(x) = f(\delta x)$. Then for a fixed f,

$$\|f_\delta\|_p = \|f\|_p \delta^{-\frac{n}{p}} = c\delta^{-\frac{n}{p}}$$

$$\hat{f}_\delta(\xi) = \delta^{-n} \hat{f}(\delta^{-1}\xi)$$

$$\|\hat{f}_\delta\|_q = c'\delta^{-\frac{n}{q'}}.$$

Therefore, $\|\hat{f}\|_q \leq C\|f\|_p$ can only happen for some f if

$$c'\delta^{-\frac{n}{q'}} \leq c\delta^{-\frac{n}{p}}$$

for all $\delta > 0$. Thus $q' = p$.

1.16 LEMMA. *Let (T, K, μ) be a multiplier of L^p. Then*

$$\|\mu\|_\infty \leq {}_p\|T\|.$$

PROOF: Note that ${}_p\|T\| = {}_{p'}\|T\|$, so that we can interpolate between L^p and $L^{p'}$. L^2 always lies in between, so the operator is always bounded on L^2. Moreover, $\frac{1}{2} = \frac{1}{2}\left(\frac{1}{p} + \frac{1}{p'}\right)$, whence

$$\|\mu\|_\infty = {}_2\|T\| \leq {}_p\|T\|^{\frac{1}{2}}\ {}_{p'}\|T\|^{\frac{1}{2}} = {}_p\|T\|.$$

1.17 REMARK: Notice that ${}_p\|T\| \leq \|K\|_1$, as remarked in 1.10, although for a multiplier of L^p the L^1 norm of the convolution kernel is probably not finite. But we now do have a relationship amongst all three components of a multiplier triple, (T, K, μ):

$$\|\mu\|_\infty \leq {}_p\|T\| \leq \|K\|_1.$$

Of course, this relation is worthless for most purposes: we can expect the estimate to be good only when the extreme sides are equal: $\|\mu\|_\infty = \|K\|_1$. But this means $\|\hat{K}\|_\infty = \|K\|_1$. But $\|\hat{K}\|_\infty \leq \|K\|_1$ in general. On the other hand, if K is a positive function,

$$\|K\|_1 = \int K = \int K(x)e^{2\pi i 0 x}dx = \hat{K}(0) \leq \|\hat{K}\|_\infty\ ;$$

so that for positive K,

$$\|\hat{K}\|_\infty = \|K\|_1 = {}_p\|T\|.$$

For positive convolution kernels, then, the whole story of multipliers is clear: they yield multipliers of an L^p if and only if they yield multipliers of all L^p, and in this case their L^1 norm is exactly their L^p operator norm. On the other hand, we saw in the Introduction that really the most interesting multiplier operators have terribly unpositive convolution kernels. For such multipliers, the gap between $\|K\|_1$ and $\|\hat{K}\|_\infty$ is large, but we will need to control both of these to control the operator norm of T.

This is one of the central problems in Fourier analysis: the need for simultaneous control of K and $\hat{K}$.

1.18 REMARK: There are other means of controlling functions in L^p for $1 < p < 2$. The simplest is to chop an L^p function into pieces, each of which is in some other Lebesgue class. The basic example of this is setting

$$G = \{x|\ |f(x)| \leq 1\};\ B = \{x|\ |f(x)| > 1\}.$$

Then $f = f\chi_G + f\chi_B = g + b$. This is a decomposition of f into a "good" function g and a "bad" function b; the good part is that

$$\|g\|_2^2 = \int_G |f|^2 = \int_G |f|^{2-p}|f|^p \leq \int_G |f|^p \leq \|f\|_p^p;$$

$$\|b\|_1 = \int |b| = \int_B |f| \leq |B|^{\frac{1}{p}}\|f\|_p.$$

So we see that an arbitrary function in L^p for p between 1 and 2 can be written as a sum of functions, one in L^1 and the other in L^2. If T is a linear operator, we have some hope of following the L^p boundedness of T simply knowing the L^1 and L^2 behaviour. Now we'll do all this more carefully.

1.19 DEFINITION: A measurable function is in weak L^p if

$$|\{x|\ |f(x)| > \lambda\}| \leq C\lambda^{-p}$$

where C is independent of λ, $0 < \lambda < \infty$.

An operator T is said to be weak (p, p) if

$$|\{x|\ |Tf(x)| > \lambda\}| \leq C\frac{\|f\|_p^p}{\lambda^p}$$

where C is independent of λ and f.

1.20 REMARKS:

a) Usually we write $|\{f > \lambda\}|$ as a shorthand for $|\{x|\ |f(x)| > \lambda\}|$; since we may often take positive functions, it turns out this never causes notational problems. This is called the distribution function of f, and often written as $\lambda(f)$.

b) A typical function which is in weak L^1 but not in L^1 is $f(x) = \frac{1}{x}$. A standard computation shows L^p functions are always weak L^p:

$$\lambda^p|\{x|\ |f(x)| > \lambda\}| = \int_{\{f>\lambda\}} \lambda^p dx$$

$$\leq \int_{\{f>\lambda\}} |f(x)|^p dx \leq \int |f|^p = \|f\|_p^p.$$

c) If weak L^p is going to be any good in analyzing serious L^p functions, distribution functions need to be connected with L^p norms. The relation is:

$$\|f\|_p^p = p\int_0^\infty |\{x|\ |f(x)| > \lambda\}|\ \lambda^{p-1}d\lambda.$$

The simplest proof uses simple functions, and this reduces the whole thing to what happens for $f = \chi_I$, where I is an interval. Then $\|f\|_p^p = |I|$, and $|\{x|\ |f(x)| > \lambda\}| = 0$ if $\lambda > 1$, and equals $|I|$ if $0 < \lambda \leq 1$. Then the integral is

$$p|I|\int_0^1 \lambda^{p-1}d\lambda = |I|.$$

1.21 THEOREM. *Assume T is a linear operator which is weak (p_0, p_0) and is weak (p_1, p_1). Then T is a bounded operator on L^p for p between p_0 and p_1.*

PROOF: To make it easy on the authors, we do only the case $1 \leq p_0 < p < p_1 < \infty$. As in Remark 1.18, we let:

$$G = \{|f| \leq \frac{\lambda}{2}\},\ B = \{|f| > \frac{\lambda}{2}\}.$$

$$g = f\chi_G;\ b = f\chi_B$$

and we note that $|Tf| = |Tg + Tb| \leq |Tg| + |Tb|$. Unfortunately, distribution functions are not linear, so this decomposition of Tf is not immediately useful. But there is a substitute for linearity:

$$|\{|Tf| > \lambda\}| \leq |\{|Tg| + |Tg| > \lambda\}|$$

$$\leq |\{|Tg| > \frac{\lambda}{2}\}| + |\{|Tb| > \frac{\lambda}{2}\}|$$

$$\leq C_0\lambda^{-p_0}\int |b|^{p_0} + C_1\lambda^{-p_1}\int |g|^{p_1}.$$

Now we compute:

$$\|Tf\|_p^p = p\int |\{|Tf| > \lambda\}|\lambda^{p-1}d\lambda$$

and find that it is bounded by the sum of two terms; the first term is:

$$C_0 p\int_0^\infty \lambda^{p-p_0-1}\int_{\{|f|\geq\frac{\lambda}{2}\}} |f|^{p_0}dx d\lambda$$

$$= C_0 p\int\int_0^{2|f|} \lambda^{p-p_0-1}d\lambda|f(x)|^{p_0}dx = \frac{2^{p-1}pC_0}{p-p_0}\|f\|_p^p.$$

Similarly, the second term is bounded by:

$$\frac{2^{p-1}pC_1}{p_1-p}\|f\|_p^p.$$

REMARK: This result is due to Marcinkiewicz; notice that as p tends to p_0 or to p_1, the bounds on the multiplier get very bad. The rate at which they get bad provides extra information at the endpoints. The basic example to keep in mind is $\frac{1}{x}$; here the L^1 norm is not finite, but the divergence is measured by $\log x$. A similar result holds in the Marcinkiewicz interpolation theorem; see the notes at the end of the Chapter.

1.22 LEMMA. *If f, g are in $\mathcal{S}$, then*

a) $\widehat{\bar{f}}(\xi) = \overline{\hat{f}(-\xi)}$.
b) If $f_\delta(x) = f(\delta x)$, *then* $\hat{f}_\delta(\xi) = \delta^{-n}\hat{f}(\delta^{-1}\xi)$.
c) If R denotes a rotation of $\mathbb{R}^n$, *and* $f_R(x) = f(Rx)$, *then* $\hat{f}_R(\xi) = \hat{f}(R^{-1}\xi)$.
d) $\widehat{(\tau_y f)}(\xi) = e^{2\pi i y\xi}\hat{f}(\xi)$.
e) $\widehat{(fg)}(\xi) = \hat{f}\star\hat{g}(\xi)$.
f) $\widehat{(f\star g)}(\xi) = \hat{f}(\xi)\hat{g}(\xi)$.
g) $D_{\xi_j}\hat{f}(\xi) = -2\pi i\widehat{(x_j f)}(\xi)$.

PROOF: All of these follow from changes of variables and the definitions. A reader who has not seen these results before is urged to work out the details. This is making friends with the Fourier transform.

1.23 THEOREM. *Assume (T,μ,K) is a Fourier multiplier triple for $L^p(\mathbb{R}^n)$. Then:*

a) $Re(\mu)$ and $Im(\mu)$ are L^p multipliers.

b) Let T_δ be the operator with multiplier $\mu(\delta\xi)$. Then ${}_p\|T_\delta\| = {}_p\|T\|$.

c) Let T_η be the operator with multiplier $\mu(\xi-\eta)$. Then ${}_p\|T_\eta\| = {}_p\|T\|$.

d) If $\check{\psi} \in L^1$, let T_ψ be the operator with multiplier $\psi\mu$. Then ${}_p\|T_\psi\| \le \|\psi\|_1\, {}_p\|T\|$.

e) If $\psi \in L^1$, let T_ψ be the operator with multiplier $\psi \star \mu$. Then ${}_p\|T_\psi\| \le \|\psi\|_1\, {}_p\|T\|$.

f) Let $\xi_1, \xi_2, \ldots, \xi_{n-k}$ be fixed, and define

$$\mathbb{R}^{k_0} = \{\xi \in \mathbb{R}^n | \xi_1 \cdot \xi + \ldots + \xi_{n-k} \cdot \xi = \lambda\}.$$

Let $\mu_0 = \mu|_{\mathbb{R}^{k_0}}$ and let T_0 be the operator on $L^p(\mathbb{R}^{k_0})$ with multiplier μ_0. Then for almost all λ, ${}_p\|T_0\| \le {}_p\|T\|$.

PROOF: The proofs of a) - d) are all easy, and all make use of the fact that the Fourier transform behaves predictably under conjugation, dilation, rotation, translation, or what-have-you. We do a typical proof to suggest that the reader have the same experiences:

d)

$$T_\eta f(x) = \int \mu(\xi-\eta) e^{2\pi i x\xi} \hat{f}(\xi) d\xi$$

$$= \int \mu(\xi) e^{2\pi i x(\xi+\eta)} \hat{f}(\xi+\eta) d\xi$$

$$= e^{2\pi i x\xi\eta} \int \mu(\xi) e^{2\pi i x\eta} \left(e^{-2\pi i x\eta} f(x)\right)(\xi) d\xi = JTJ^{-1}(f)$$

where J is the isometry of multiplication by $e^{2\pi i x\eta}$.

e) We will use duality, a very standard trick in the Fourier multiplier game. Choose f, g, in $\mathcal{S}$ with $\|f\|_p = \|g\|_{p'} = 1$, and compute:

$$\int T_\psi f \bar{g} = \int \widehat{T_\psi f}\ \widehat{\bar{g}} = \int \psi \star \mu \hat{f}\ \widehat{\bar{g}}$$

$$= \int\int \psi(\eta)\mu(\xi-\eta)\hat{f}(\xi)\ \widehat{\bar{g}}(\xi) d\eta d\xi = \int \psi(\eta) \int \mu(\xi-\eta)\hat{f}(\xi)\ \widehat{\bar{g}}(\xi) d\xi d\eta$$

$$\le \|\psi\|_1\ \|\int \mu(\xi-\eta)\hat{f}(\xi)\ \widehat{\bar{g}}(\xi) d\xi\|_{\infty,\eta} \|g\|_{p'} \le \|\psi\|_1\ {}_p\|T\|,$$

by part d) above.

f) We will give the proof only for the case $n=2, k=1$, and even restrict attention to $\xi_1 = (0,1)$. Higher dimensional cases are very similar, except they involve a lot more notation, while more general cases in $\mathbb{R}^2$ can be obtained by rotations and translations of this one.

In our special case, $\mathbb{R}^{1_0} = \{(-,\lambda)\}$, and we must show that for almost all λ, $\mu(\xi_1,\lambda)$ is a multiplier of $L^p(\mathbb{R}^1)$. Since this λ business could cause real friction with previous notation, we'll prove that $\mu(\xi_1,\xi_2)$ is a multiplier of $L^p(\mathbb{R}^1)$ for almost all ξ_2.

The first pain is that μ is only measurable, and the restriction of μ to one-dimensional hyperplanes is not very well defined. We will use another standard trick in multiplier theory. Assume first that μ is continuous. Let $f_i, g_i,\ i = 1, 2$ be chosen with $\|f_i\|_p = \|g_i\|_{p'} = 1$; let $f(x_1, x_2) = f_1(x_1)f(x_2)$, and similarly for g. Then:

$$\int \left[\int \mu(\xi_1, \xi_2)\hat{f}_1(\xi_1)\hat{g}(\xi_1)d\xi_1 \right] \hat{f}_2(\xi_2)\hat{g}(\xi_2)d\xi_2 = \int \mu(\xi)\hat{f}(\xi)\hat{g}(\xi)d\xi$$

$$= \int \widehat{Tf}\hat{g} = \int Tfg \leq \|Tf\|_p \|g\|_{p'} \leq \ _p\|T\|.$$

Define

$$\sigma(\xi_2) = \int \mu(\xi_1, \xi_2)\hat{f}_1(\xi_1)\hat{g}_1(\xi_1)d\xi_1.$$

Rewriting the above computation gives us:

$$\int \sigma(\xi_2)\hat{f}(\xi_2)\hat{g}_2(\xi_2)d\xi_2 \leq \ _p\|T\|.$$

Therefore, σ is a multiplier of $L^p(\mathbb{R}^1)$. But Lemma 1.16 told us that $\|\sigma(\xi_2)\|_\infty \leq \ _p\|T\|$. Since μ is continuous, σ is continuous, and $\sigma(\xi_2) \leq \ _p\|T\|$, for all ξ_2. Rewriting,

$$|\int \mu(\xi_1, \xi_2)\hat{f}_1(\xi_1)\hat{g}(\xi_1)d\xi_1| \leq \ _p\|T\|$$

$$|\int \mu_0(\xi_1)\hat{f}_1(\xi_1)\hat{g}(\xi_1)d\xi_1| \leq \ _p\|T\|$$

$$|\int T_0 f_1 g_1| \leq \ _p\|T\|$$

$$_p\|T_0\| \leq \ _p\|T\|.$$

We assumed μ was continuous; we'll use a standard trick to replace continuity assumptions by almost everywhere conditions. Let S denote the unit square, and let $\psi_\delta(\xi) = \delta^{-2}\chi_S(\delta^{-1}\xi)$. Then $\psi_\delta \in L^1$ and $\mu \in L^\infty$, so that $\psi_\delta \star \mu$ is continuous, whence, for every ξ_2,

$$\int \psi_\delta \star \mu(\xi_1, \xi_2)\hat{f}_1(\xi_1)\hat{g}_1(\xi_1)d\xi_1 \leq \ _p\|T\psi_\delta\|$$

$$\leq \ _p\|T\|\|\psi_\delta\|_1 = \ _p\|T\|.$$

Now, for almost every ξ_2, it happens that almost every ξ_1 is a Lebesgue point of μ_0. These ξ_2 are the ones referred to in the theorem; fix one. The definition of Lebesgue point ξ_1 is that

$$\lim_{\delta \to 0} \psi_\delta \star \mu(\xi_1, \xi_2) = \lim_{\delta \to 0} \delta^{-2} \int_{-\delta}^{\delta} \int_{-\delta}^{\delta} \mu(\eta_1 - \xi_1, \eta_2 - \xi_2)d\eta_1 d\eta_2$$

$$= \mu(\xi_1, \xi_2).$$

For this choice of ξ_2, $\psi_\delta \star \mu(\xi_1, \xi_2)$ converges almost everywhere in ξ_1 to $\mu(\xi_1, \xi_2) = \mu_0(\xi_1)$. Of course $(\psi_\delta \star \mu)\hat{f}_1 \hat{g}_1$ also converges; to get the integrals to converge, I need dominated convergence. But

$$|(\psi_\delta \star \mu)\hat{f}_1 \hat{g}_1| \leq \|\psi_\delta \star \mu\|_\infty |\hat{f}_1||\hat{g}_1|$$

$$\leq \|\psi_\delta\|_1 \|\mu\|_\infty |\hat{f}_1 \hat{g}_1|.$$

Therefore,

$$\int \mu_0(\xi_1)\hat{f}_1(\xi_1)\hat{g}(\xi_1)d\xi_1 = \lim_{\delta \to 0} \int \psi_\delta \star \mu \hat{f}_1 \hat{g}_1 \leq {}_p\|T\|.$$

and ${}_p\|T_0\| \leq {}_p\|T\|$ for almost all ξ_2.

1.24 REMARKS: These results are important because they let us deform multipliers into things we can handle. For example, if ψ is in $\mathcal{S}$, and μ is a multiplier, then $\psi\mu$ and $\psi \star \mu$ are multipliers. We already used this in the proof of the above theorem: we changed an arbitrary multiplier into a continuous function, with no side-effects. In general, μ is just in L^∞ , and K is some distribution, but if we alter by a compactly supported function ψ, then K_ψ, the inverse Fourier transform of $\psi\mu$, is a continuous function. In the following, changing an awful distribution into a nice regular function is going to be our favorite trick.

The real importance is philosophical. The L^p boundedness of (T, K, μ) is influenced by the size of K, which reflects the smoothness of $\check{\mu} = K$. But the computation of μ is a global average of all of μ against some exponentials, so it is difficult to understand how a local property of μ can affect the Fourier transform K. This is the point of having around $\psi\mu$ or $\psi \star \mu$. The first changes μ into a compactly supported function; the second changes μ into a smoother function.