trim

TEXTS AND READINGS IN MATHEMATICS 7

Harmonic Analysis

Second Edition

Henry Helson

HINDUSTAN BOOK AGENCY

TEXTS AND READINGS IN MATHEMATICS **7**

Harmonic Analysis

Second Edition

Harmonic Analysis
Second Edition

Henry Helson

HINDUSTAN
BOOK AGENCY

Published by
Hindustan Book Agency (India)
P 19 Green Park Extension
New Delhi 110 016
India

email: info@hindbook.com
http://www.hindbook.com

ISBN 978-93-80250-05-2 ISBN 978-93-86279-47-7 (eBook)
DOI 10.1007/978-93-86279-47-7

CONTENTS

PREFACE TO THE SECOND EDITION

Harmonic Analysis used to go by the more prosaic name Fourier Series. Its elevation in status may be due to recognition of its crucial place in the intersection of function theory, functional analysis, and real variable theory; or perhaps merely to the greater weightiness of our times. The 1950's were a decade of progress, in which the author was fortunate to be a participant. Some of the results from that time are included here.

This book begins at the beginning, and is intended as an introduction for students who have some knowledge of complex variables, measure theory, and linear spaces. Classically the subject is related to complex function theory. We follow that tradition rather than the modern direction, which prefers real methods in order to generalize some of the results to higher dimensional spaces. In this edition there is a full presentation of Bochner's theorem, and a new chapter treats the duality theory for compact and discrete abelian groups. Then the author indulges his own experience and tastes, presenting some of his own theorems, a proof by C. L. Siegel of Minkowski's theorem, applications to probability, and in the last chapter two different methods of proving the theorem of Weyl on equidistribution modulo 1 of $(P(k))$, where P is a real polynomial with at least one irrational coefficient.

This is not a treatise. If what follows is interesting and useful, no apology is offered for what is not here. The notes at the end are intended to orient the reader who wishes to explore further.

I express my warm thanks to Robert Burckel, whose expert criticism and suggestions have been most valuable. I am also indebted to Alex Gottlieb for detailed reading of the text. This edition is appearing simultaneously in India. I am grateful to Professor R. Bhatia, and to the Hindustan Book Agency for the opportunity to present it in this cooperative way.

HH

Chapter 1
Fourier Series and Integrals

1. Definitions and easy results

The unit circle $\mathbf{T}$ consists of all complex numbers of modulus 1. It is a compact abelian group under multiplication. If f is a function on $\mathbf{T}$, we can define a periodic function F on the real line $\mathbf{R}$ by setting $F(x) = f(e^{ix})$. It does not matter whether we study functions on $\mathbf{T}$ or periodic functions on $\mathbf{R}$; generally we shall write functions on $\mathbf{T}$. Everyone knows that in this subject the factor 2π appears constantly. Most of these factors can be avoided if we replace Lebesgue measure dx on the interval $(0, 2\pi)$ by $d\sigma(x) = dx/2\pi$. We shall generally omit the limits of integration when the measure is σ; they are always 0 and 2π, or another interval of the same length.

One more definition will simplify formulas: χ is the function on $\mathbf{T}$ with values $\chi(e^{ix}) = e^{ix}$. Thus χ^n represents the exponential e^{nix} for each integer n.

We construct Lebesgue spaces $\mathbf{L}^p(\mathbf{T})$ with respect to σ, $1 \le p \le \infty$. The spaces $\mathbf{L}^1(\mathbf{T})$ of summable functions and $\mathbf{L}^2(\mathbf{T})$ of square-summable functions are of most interest. Since the measure is finite these spaces are nested: $\mathbf{L}^p(\mathbf{T}) \supset \mathbf{L}^r(\mathbf{T})$ if $p < r$ (Problem 2 below). Thus $\mathbf{L}^1(\mathbf{T})$ contains all the others. For summable functions f we define *Fourier coefficients*

$$a_n(f) = \int f\chi^{-n}\, d\sigma \quad (n = 0, \pm 1, \pm 2, \ldots), \tag{1.1}$$

and then the *Fourier series* of f is

$$f(e^{ix}) \sim \sum a_n(f)\, e^{nix}. \tag{1.2}$$

We do not write equality in (1.2) unless the series converges to f.

There is a class of functions for which this is obviously the case. A *trigonometric polynomial* is a finite sum

$$P(e^{ix}) = \sum a_n\, e^{nix}. \tag{1.3}$$

Then

$$a_n(P) = \sum_k a_k \int e^{(k-n)ix}\, d\sigma(x) = a_n, \tag{1.4}$$

if we define $a_n = 0$ for values of n not occurring in the sum (1.3). Thus (1.3), which defines P, is also the Fourier series of P.

This reasoning can be carried further. Suppose that f is a function defined as the sum of the series in (1.3), now allowed to have infinitely many terms but assumed to converge *uniformly* on $\mathbf{T}$ for some ordering of the series. Then the same calculation is valid and we find that $a_n(f) = a_n$ for each n. That is, the trigonometric series converging to f is also the Fourier series of f.

From (1.1) we obviously have $|a_n(f)| \le \|f\|_1$ for all n. A more precise result can be proved in $\mathbf{L}^2(\mathbf{T})$.

Bessel's Inequality. *If f is in $\mathbf{L}^2(\mathbf{T})$, then*

$$\sum |a_n(f)|^2 \le \|f\|_2^2. \tag{1.5}$$

Part of the assertion is that the series on the left converges. For each positive integer k set

$$f_k = \sum_{-k}^{k} a_n(f)\chi^n. \tag{1.6}$$

The norm of any function is non-negative, and thus

$$0 \leq \|f - f_k\|_2^2 = \|f\|_2^2 + \|f_k\|_2^2 - 2\Re \int f\bar{f}_k \, d\sigma. \tag{1.7}$$

Since the exponentials form an orthonormal system, the second term on the right equals

$$\sum_{-k}^{k} |a_n|^2. \tag{1.8}$$

The last term is

$$-2\Re \sum_{-k}^{k} \bar{a}_n \int f\chi^{-n} \, d\sigma = -2 \sum_{-k}^{k} |a_n|^2. \tag{1.9}$$

This term combines with (1.8), and (1.7) becomes

$$\sum_{-k}^{k} |a_n|^2 \leq \|f\|^2. \tag{1.10}$$

Since k is arbitrary, (1.5) is proved.

A kind of converse to Bessel's inequality is the

Riesz-Fischer theorem. *If (a_n) is any square-summable sequence, there is a function f in $\mathbf{L}^2(\mathbf{T})$ such that $a_n(f) = a_n$ for all n, and*

$$\sum |a_n|^2 = \|f\|_2^2. \tag{1.11}$$

Define

$$f_k = \sum_{-k}^{k} a_n \chi^n \tag{1.12}$$

for each positive integer N. Then

$$\|f_{k+r} - f_k\|_2^2 = \sum_{k<|n|\leq k+r} |a_n|^2 \tag{1.13}$$

for all positive integers r. Thus (f_k) is a Cauchy sequence in $\mathbf{L}^2(\mathbf{T})$. Let f be its limit. The sequence $(a_n(f))$ converges to $(a_n(f))$ in the space $\mathbf{l}^2$, by Bessel's inequality. Therefore (1.11), which holds for each k, is valid in the limit.

Now Bessel's inequality is actually *equality* for every f in $\mathbf{L}^2(\mathbf{T})$, and this equality is called the *Parseval relation.* Computations already performed show that equality holds for all trigonometric polynomials. The Fourier transform, thought of as a mapping from trigonometric polynomials in the norm of $\mathbf{L}^2(\mathbf{T})$ into $\mathbf{l}^2$, is an isometry whose range consists of all sequences (a_n) such that $a_n = 0$ for $|n|$ sufficiently large. The range is dense in $\mathbf{l}^2$. If we knew that the family of trigonometric polynomials is dense in $\mathbf{L}^2(\mathbf{T})$, then the isometry has a unique continuous extension to a linear isometry of all of $\mathbf{L}^2(\mathbf{T})$ onto $\mathbf{l}^2$. It is obvious that this extension is the Fourier transform. In Section 4 it will be shown that trigonometric polynomials are indeed dense in $\mathbf{L}^2(\mathbf{T})$, and this will prove the Parseval relation.

When the Parseval relation is known, the Riesz-Fischer theorem can immediately be strengthened to say that *the function whose coefficients are the given sequence (a_n) is unique.* For if f and g have the same Fourier coefficients, then $f-g$ has all its coefficients 0; that is, $f-g$ is orthogonal in the Hilbert space to all trigonometric polynomials, and must be null.

Mercer's theorem. *For all f in $L^1(T)$, $a_n(f)$ tends to 0 as n tends to $\pm\infty$.*

Bessel's inequality shows that this is true if f is in $L^2(T)$. Now $L^2(T)$ is dense in $L^1(T)$. (A proof of this fact depends on the particular way in which measure theory was developed and the Lebesgue spaces defined.) Choose a sequence (f_k) of elements of $L^2(T)$ converging to f in the norm of $L^1(T)$. Then $(a_n(f_k))$ converges to $(a_n(f))$ as k tends to ∞, *uniformly in n*. It follows that the limit sequence vanishes at $\pm\infty$, as was to be proved.

Mercer's theorem is the source of theorems asserting the convergence of Fourier series. Here is the most important such result, with a simple proof suggested by Paul Chernoff.

Theorem 1. *Suppose that f is in $L^1(T)$ and that $f(e^{ix})/x$ is summable on $(-\pi, \pi)$.* Then

$$\sum_{-M}^{N} a_n(f) \to 0 \tag{1.14}$$

as M, N tend independently to ∞.

Form the function $g(e^{ix}) = f(e^{2ix})/\sin x$. The hypothesis implies that g is in $L^1(T)$. Let (a_n) be the Fourier coefficients of f, and (b_n) those of g. We have

$$f(e^{2ix}) = \frac{1}{2i}(e^{ix} - e^{-ix})\, g(e^{ix}); \tag{1.15}$$

calculating the coefficients with even indices of the functions on both sides gives

$$2ia_n = b_{2n-1} - b_{2n+1} \tag{1.16}$$

for each integer n. Hence

$$2i\sum_{-M}^{N} a_n = b_{-2M-1} - b_{2N+1}. \tag{1.17}$$

By Mercer's theorem, this quantity tends to 0 as M, N tend to ∞.

Corollary. *If f is in $\mathbf{L}^1(\mathrm{T})$ and satisfies a Lipschitz condition at a point e^{it}, then the Fourier series of f converges to $f(e^{it})$ at that point.*

It is easy to check that addition of a constant to a summable function, and translation, have the formally obvious effect on the Fourier series of the function (Problem 1 below). Therefore without loss of generality we may assume that $t = 0$ and $f(1) = 0$. Now the hypothesis means that

$$|f(e^{ix}) - f(e^{it})| \leq k|x-t|^{\alpha} \tag{1.18}$$

for some constant k, and some α satisfying $0 < \alpha \leq 1$, for all x. With f and t as just assumed, the hypothesis of the theorem is satisfied. The conclusion is that f is the sum of its Fourier series at the point 1, and the general result follows.

The results of this section are striking but elementary. In order to get further we shall have to introduce new techniques.

Problems

1. Show that if f is in $\mathbf{L}^1(\mathrm{T})$, and g is defined by

$$g(e^{ix}) = c + f(e^{i(x+s)})$$

where c is complex and s real, then $a_n(g) = a_n(f)\, e^{nis}$ for all $n \neq 0$, and $a_0(g) = a_0(f) + c$.

2. Suppose that $1 \leq p < r \leq \infty$. Use Hölder's inequality to show that (a) $\mathbf{L}^p(\mathbf{T}) \supset \mathbf{L}^r(\mathbf{T})$ and (b) $||f||_p \leq ||f||_r$ for f in $\mathbf{L}^r(\mathbf{T})$.

3. Show that if f is real, then its coefficients satisfy $a_{-n} = \bar{a}_n$ for all n. (In particular, a_0 is real.)

4. Calculate the Fourier series of these periodic functions.

(a) $f(e^{ix}) = -1$ on $(-\pi, 0)$, $= 1$ on $(0, \pi)$

(b) $g(e^{ix}) = x + \pi$ on $(-\pi, 0)$, $= x - \pi$ on $(0, \pi)$

(c) $h(e^{ix}) = (1 - re^{ix})^{-1}$, where $0 < r < 1$

(These series will be needed later; keep a record of them.)

5. Prove the *principle of localization: if f and g are in* $\mathbf{L}^1(\mathbf{T})$ *and are equal on some interval, then at each interior point of the interval their Fourier series both converge and to the same value, or else both diverge.*

6. Suppose that $f(e^{ix})$ and $(f(e^{ix}) + f(e^{-ix}))/x$ are summable on $(-\pi, \pi)$. Show that

$$\sum_{-N}^{N} a_n(f) \to 0$$

as $N \to \infty$. Is the conclusion of Theorem 1 necessarily true?

2. The Fourier transform

On the line $\mathbf{R}$ construct the Lebesgue spaces $\mathbf{L}^p(\mathbf{R})$ with respect to ordinary Lebesgue measure. The Fourier transform of a summable function f is

$$\hat{f}(y) = \int_{-\infty}^{\infty} f(x)\, e^{-ixy}\, dx. \tag{2.1}$$

Obviously $|\hat{f}(y)| \leq ||f||_1$ for all y. The function $\hat{f}$ is continuous. To see this, write

$$\hat{f}(y') - \hat{f}(y) = \int_{-\infty}^{\infty} f(x)\, (e^{-ixy'} - e^{-ixy})\, dx. \tag{2.2}$$

As y' tends to y the integrand tends to 0 for each x. Also the modulus of the integrand does not exceed $2|f(x)|$, a summable function. By Lebesgue's dominated convergence theorem, the integral tends to 0 as claimed.

The analogue of Mercer's theorem is called the *Riemann-Lebesgue lemma: $\hat{f}(y)$ tends to 0 as y tends to $\pm\infty$.* If a sequence of functions f_n converges to f in $\mathbf{L}^1(\mathbf{R})$, then the transforms $\hat{f}_n$ tend to $\hat{f}$ uniformly. Therefore, as in the proof of Mercer's theorem, it will suffice to show that the assertion is true for all functions in a dense subset of $\mathbf{L}^1(\mathbf{R})$. Let f be the characteristic function of an interval $[a, b]$. Its transform is

$$\int_a^b e^{-ixy}\,dx = \frac{e^{-iay} - e^{-iby}}{iy} \qquad (y \neq 0), \tag{2.3}$$

which tends to 0. Therefore a linear combination of characteristic functions of intervals, that is a step function, has the property, and such functions are dense. This completes the proof.

There is a version of Theorem 1 for the line.

Theorem 1′. *If f and $f(x)/x$ are summable, then*

$$\lim_{A,B\to\infty} \int_{-A}^{B} \hat{f}(y)\,dy = 0. \tag{2.4}$$

The quantity that should tend to 0 as A, B tend to ∞ is

$$\int_{-A}^{B} \int_{-\infty}^{\infty} f(x)\,e^{-ixy}\,dx\,dy. \tag{2.5}$$

The integrand is summable over the product space, so it is legitimate to interchange the order of integration. After integrating with respect to y we find

$$\int_{-\infty}^{\infty} f(x)\,(-ix)^{-1}(e^{-iBx} - e^{iAx})\,dx. \tag{2.6}$$

This tends to 0 by the Riemann-Lebesgue lemma.

There is an inversion theorem like the Corollary to Theorem 1, but a difficulty has to be met that did not arise on the circle.

Corollary. *If f is summable on the line and satisfies a Lipschitz condition at t, then*

$$f(t) = \lim_{A,B\to\infty} \frac{1}{2\pi} \int_{-A}^{B} \hat{f}(y)\, e^{ity}\, dy. \tag{2.7}$$

As on the circle, we may assume that $t = 0$. If $f(0) = 0$ there is nothing more to prove. However, we cannot now reduce the general result to this one merely by subtracting $f(0)$ from f, because nonzero constants are not summable functions. Therefore we must subtract from f a *function* g that is summable, smooth near 0, with $g(0) = f(0)$, such that we can actually calculate $\hat{g}$ and the inverse Fourier transform of $\hat{g}$. The details are in Problem 5 below.

These theorems are parallel to the ones proved on the circle group in the last section. Analysts have known for a hundred years that this similarity goes very far, although proofs on the line are often more complicated. However our axiomatic age has produced a unified theory of harmonic analysis on locally compact abelian groups, among which are the circle, the line and the integer group, but also other groups of interest in analysis and number theory. The discovery of Banach algebras by A. Beurling and I. M. Gelfand was closely associated with this generalization of classical harmonic analysis. In beginning the study of Fourier series it is well to realize that its ideas are more general than appears in the classical context. Furthermore, this commutative theory has inspired much of the theory of group representations, which is the generalization of harmonic analysis to locally compact, non-abelian groups.

There is one more classical Fourier transform. The transform of a function on **T** is a sequence, that is, a function on **Z**. A function on **R** has transform that is another function on **R**. Now let (a_n) be a summable function on **Z**. Its transform is the function on **T** defined by

$$f(e^{ix}) = \sum_{-\infty}^{\infty} a_n e^{-nix}. \tag{2.8}$$

Generally, the transform of a function on a locally compact abelian group is a continuous function on the dual of that group. The dual of **T** is **Z**, the dual of **Z** is **T**, and **R** is dual to itself.

Problems

1. Calculate the Fourier transforms of (a) the characteristic function of the interval $[-A, A]$; (b) the triangular function vanishing outside $(-A, A)$, equal to 1 at the origin, and linear on the intervals $(-A, 0)$ and $(0, A)$.

2. Express the Fourier transform of $f(x+t)$ (a function of x) in terms of $\hat{f}$.

3. Suppose that $(1+|x|)f(x)$ is summable. Show that $(\hat{f})'$ is the Fourier transform of $-ixf(x)$, as suggested by differentiating (2.1).

4. Find the Fourier transform of f' if f is continuously differentiable with compact support. (Such functions are dense in

$\mathbf{L}^1(\mathbf{R})$; this gives a new proof of the Riemann-Lebesgue lemma.)

5. Calculate the Fourier transform of $g(x) = e^{-|x|}$. Verify that

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{g}(y)\, dy = 1.$$

Use this information to complete the proof of the corollary.

6. (a) Use the calculus of residues to find

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{2}{1+y^2} e^{ixy}\, dy.$$

(b) Obtain the same result by applying the inversion theorem to the function of Problem 5.

7. Show that the Fourier transform of $\exp(-x^2)$ is $\sqrt{\pi}\exp(-y^2/4)$. [Find the real integral

$$\int_{-\infty}^{\infty} \exp(-x^2 + 2xu)\, dx.$$

Show that this is an entire function of u, and complexify u.]

3. Convolution; approximate identities; Fejér's theorem

The *convolution* of functions in $\mathbf{L}^1(\mathbf{T})$ is defined by

$$f*g(e^{ix}) = \int f(e^{it})\, g(e^{i(x-t)})\, d\sigma(t). \tag{3.1}$$

If f and g are square-summable the integral exists, and is bounded by $\|f\|_2\|g\|_2$ by the Schwarz inequality. More generally, if f is in $\mathbf{L}^p(\mathbf{T})$ and g is in $\mathbf{L}^q(\mathbf{T})$, where p and q are conjugate exponents, then the integral exists, and $|f*g(e^{ix})| \le \|f\|_p\|g\|_q$ for all x, by the Hölder inequality. Furthermore the convolution is a continuous function (Problem 4 below).

If f and g are merely assumed to be summable, their product may not be summable, and the integral may not exist. It is surprising and important that nevertheless the product under the integral sign in (3.1) is summable at *almost every* point. To prove this form the double integral

$$\int\int |f(e^{it})\, g(e^{i(x-t)})|\, d\sigma(t)\, d\sigma(x). \tag{3.2}$$

The integrand is non-negative and measurable on the product space, so the integral exists, finite or infinite. By Fubini's theorem, it equals either of the two iterated integrals. Integrating first with respect to x, and using the fact that σ is invariant under translation, we find $||f||_1||g||_1$ as the value of (3.2). Therefore the integral with respect to t

$$h(e^{ix}) = \int |f(e^{it})\, g(e^{i(x-t)})|\, d\sigma(t) \tag{3.3}$$

must be finite a.e. Moreover h is summable, and $||h||_1 \leq ||f||_1||g||_1$. It follows that the integrand in (3.1) is summable for a.e. x, the convolution (defined almost everywhere) belongs to $\mathbf{L}^1(\mathbf{T})$, and

$$||f*g||_1 \leq ||f||_1||g||_1. \tag{3.4}$$

Convolution is associative and commutative, and distributes over addition. (The proofs are calculations involving elementary changes of variable, and are asked for in Problem 1 below.) Thus $\mathbf{L}^1(\mathbf{T})$ is a *commutative Banach algebra*, and indeed this is the algebra that led Gelfand to the concept. This algebra has no identity.

(An *algebra* is a ring admitting multiplication by scalars from a specified field. Most rings in analysis are algebras over the real or complex numbers. An *ideal* in an algebra is an ideal (in the sense of rings) that is invariant under multiplication by scalars.)

Convolution is defined analogously on **R** and on **Z**. On **R**

$$f*g(x) = \int_{-\infty}^{\infty} f(t)\, g(x-t)\, dt. \tag{3.5}$$

Once more the integral is absolutely convergent if f and g are in $\mathbf{L}^2(\mathbf{R})$, or if f and g belong to complementary Lebesgue spaces, and the convolution is continuous. If the functions are merely summable, the Fubini theorem shows, exactly as on the circle, that the integrand is summable for almost every x, $f*g$ is summable, and (3.4) holds.

If f and g are in $\mathbf{l}^1$, the definition is

$$f*g(n) = \sum_{m=-\infty}^{\infty} f(m)\, g(n-m). \tag{3.6}$$

This series converges absolutely for all n, and (3.4) holds once more.

An *approximate identity* on **T** is a sequence of functions (e_n) with these properties: each e_n is non-negative, has integral 1 with respect to σ, and for every positive number $\epsilon < \pi$

$$\lim_{n\to\infty} \int_{-\epsilon}^{\epsilon} e_n\, d\sigma = 1. \tag{3.7}$$

This means that nearly all the area under the graph of e_n is close to the origin if n is large. Although $\mathbf{L}^1(\mathbf{T})$ has no identity under convolution, it has approximate identities, and these almost serve the same purpose, by this fundamental result:

Fejér's Theorem. *Let f belong to $\mathbf{L}^p(\mathbf{T})$ with $1 \le p < \infty$. For any approximate identity (e_n), $e_n * f$ converges to f in the norm of $\mathbf{L}^p(\mathbf{T})$. If f is in $\mathbf{C}(\mathbf{T})$, the space of continuous functions on* $\mathbf{T}$, *then $e_n * f$ converges to f uniformly.*

We prove the second assertion first. Let f be continuous on $\mathbf{T}$. Then

$$(3.8) \qquad f(e^{ix}) - e_n * f(e^{ix}) = \int [f(e^{ix}) - f(e^{i(x-t)})]\, e_n(e^{it})\, d\sigma(t),$$

because e_n has integral 1. Since f is continuous and $\mathbf{T}$ compact, f is uniformly continuous: given any positive number ϵ, there is a positive δ such that $|f(e^{ix}) - f(e^{i(x-t)})| \le \epsilon$ for all x, provided that $|t| \le \delta$. Denote by I the part of the integral over $(-\delta, \delta)$, and by J the integral over the complementary subset of $\mathbf{T}$. For all n

$$(3.9) \qquad |I| \le \epsilon \int e_n \, d\sigma = \epsilon.$$

If M is an upper bound for $|f|$, we have

$$(3.10) \qquad |J| \le 2M \int_{\delta}^{2\pi - \delta} e_n \, d\sigma,$$

and this quantity is as small as we please provided that n is large enough. These estimates do not depend on x. Therefore the difference (3.8) is uniformly as small as we please if n is large enough, which is what was to be proved.

Now let p satisfy $1 \le p < \infty$. For continuous f, the uniform convergence of $e_n * f$ to f implies convergence in $\mathbf{L}^p(\mathbf{T})$. In order to extend this result from continuous functions to all functions in

$\mathbf{L}^p(\mathbf{T})$, we use the following general principle, which has many applications besides this one.

Principle. *Let* X *and* Y *be normed vector spaces, and* Y *a Banach space. Let* (T_n) be *a sequence of linear operators from* X *to* Y *whose bounds* $||T_n||$ *are all less than a number* K. *Suppose that* $T_n x$ *converges to a limit we call* Tx, *for each* x *in a dense subset of* X. *Then* $T_n x$ *converges for all* x *in* X, *and the limit* Tx *defines a linear operator* T *with bound at most* K.

The easy proof is omitted.

Lemma. *For* f *in* $L^1(\mathbf{T})$ *and* g *in* $\mathbf{L}^p(\mathbf{T})$, $1 \le p$, $f*g$ *is in* $\mathbf{L}^p(\mathbf{T})$ *and* $||f*g||_p \le ||f||_1 ||g||_p$.

The lemma is trivial if $p = \infty$, so we assume p finite. Let q be the conjugate exponent, and h any function in $\mathbf{L}^q(\mathbf{T})$ with norm 1. As a linear functional on $f*g$, h has the value

$$(3.11) \qquad h*(f*g)(1) = \int\int h(e^{-ix}) f(e^{it})\, g(e^{i(x-t)})\, d\sigma(t)\, d\sigma(x).$$

The double integral exists absolutely by Hölder's inequality, as we see by integrating first with respect to x, and in modulus does not exceed $||f||_1 ||g||_p$ (because h has norm 1 in $\mathbf{L}^q(\mathbf{T})$). This proves that $f*g$ belongs to $\mathbf{L}^p(\mathbf{T})$ with norm at most equal to this quantity, and the lemma is proved.

Now we can finish the proof of Fejér's theorem. Convolution with e_n is a linear operation in $\mathbf{L}^p(\mathbf{T})$ with bound at most 1, by the lemma. This sequence of operators converges to the identity operator on each element of $\mathbf{C}(\mathbf{T})$, a dense subset of $\mathbf{L}^p(\mathbf{T})$. (This result comes from integration theory; the point is discussed further below.) The Principle asserts that the sequence of operators converges everywhere, and the limit is a bounded operator.