Lectures on Harmonic Analysis

Hörmander, Lars

1995

Document Version:
Other version

Link to publication

Citation for published version (APA):
Hörmander, L. (1995). *Lectures on Harmonic Analysis*.

Total number of authors:
1

Lars Hörmander

Lectures on Harmonic Analysis

Department of Mathematics, University of Lund
Box 118, S-221 00 Lund, Sweden

Lars Hörmander
Department of Mathematics
University of Lund
Box 118, S-221 00 Lund
Sweden

This is the second printing.

PREFACE

These lectures, given during the academic year 1994–1995, are intended as an introductory course in harmonic analysis for graduate students. The prerequisites assumed are some familiarity with distribution theory, Lebesgue integration and functional analysis. It should be possible to read most of the notes without knowing distribution theory by concentrating on the study of smooth functions and the L^2 theory of the Fourier transformation. However, I have chosen to build on distribution theory since it makes many arguments simpler and more transparent.

The very short Chapter I is intended to present the algebraic contents of commutative harmonic analysis in a context which is almost free of analysis. It contains in particular a discussion of the fast Fourier transform. Chapter II develops the basic facts on Fourier analysis in $\mathbf{R}^n$ starting from approximation of $\mathbf{R}^n/\mathbf{Z}^n$ by finite groups and of $\mathbf{R}^n$ by such torus groups. There is a substantial overlap with my book [1], where many of the topics are dealt with in greater depth.

In Chapter III the basic principles of Fourier analysis are illustrated by a study of wavelets, with an emphasis on wavelets of compact support. For applications and additional results the reader should turn to Daubechies [1] and Meyer [1], [2]. Chapter IV then returns to more traditional harmonic analysis. It is centered on L^p estimates for singular integral operators and the related study of the Hardy space $\mathscr{H}^1$ and the space BMO of functions of bounded mean oscillation. The methods developed are also applied to prove that wavelets with compact support give bases in L^p spaces and the Hardy space. A much more extensive discussion of these matters can be found in E. M. Stein [1].

The final Chapter V is devoted to the study of multipliers on the Fourier transform of L^p, in particular convergence and summability of the Fourier expansion of functions in L^p. In spite of much progress in the last few decades this is an area where many problems remain open.

The choice of topics for a course such as this is of course in no way uniquely determined. A guiding principle here has been to cover results and methods which are essential in the study of linear and non-linear differential equations. So far the study of wavelets may not qualify in that respect, but it gives excellent illustrations of the tools of Fourier analysis and is important in signal theory.

Lund in February 1995

Lars Hörmander

CONTENTS

CHAPTER I

FOURIER ANALYSIS ON FINITE ABELIAN GROUPS

1.1. The structure of finite abelian groups. Let G be a finite abelian group, with $|G|$ elements and with group operation denoted by $+$. Let us recall some elementary facts:

a) For every $a \in G \setminus \{0\}$ there exists some integer $n \neq 0$ with $na = 0$, for otherwise all elements $na \in G$, $n \in \mathbf{Z}$, would be different. If n is minimal then a generates a *cyclic* subgroup

$$G_a = \{\nu a; 0 \le \nu < n\} \cong \mathbf{Z}/n\mathbf{Z} = \mathbf{Z}_n$$

of G, and n is called the period of a. Since $|G| = |G_a||G/G_a| = n|G/G_a|$ it follows that n divides $|G|$ so $|G|a = 0$ for every $a \in G$.

b) If p is a prime then $G(p) = \{a \in G; p^\nu a = 0 \text{ for some } \nu\}$ is a subgroup of G. It is trivial unless p divides $|G|$. If $|G| = \prod_1^\nu p_j^{m_j}$ with different primes p_j and $m_j \ge 1$, then $G(p_j) = \{a \in G; p_j^{m_j} a = 0\}$. We have

$$G \cong G(p_1) \times \cdots \times G(p_\nu). \tag{1.1.1}$$

In fact, we can find integers $\gamma_1, \dots, \gamma_\nu$ such that

$$\sum_{\mu=1}^{\nu} \gamma_\mu \prod_{j\neq\mu} p_j^{m_j} = 1,$$

since the products have no common factor. If $a \in G$ it follows that

$$a = \sum_1^\nu a_\mu, \quad \text{where } a_\mu = \gamma_\mu \prod_{j\neq\mu} p_j^{m_j} a, \quad \text{hence } p_\mu^{m_\mu} a_\mu = 0.$$

Such a decomposition is unique, for assume that

$$\sum_1^\nu a_\mu = 0, \quad p_\mu^{m_\mu} a_\mu = 0,\ \mu = 1, \dots, \nu.$$

Then it follows for $1 \le \varrho \le \nu$ that

$$a_\varrho = \sum_1^\nu \gamma_\mu \prod_{j\neq\mu} p_j^{m_j} a_\varrho = \gamma_\varrho \prod_{j\neq\varrho} p_j^{m_j} a_\varrho = \gamma_\varrho \prod_{j\neq\varrho} p_j^{m_j} \sum_1^\nu a_\mu = 0.$$

Since (1.1.1) implies that $|G| = |G(p_1)| \dots |G(p_\nu)|$ and we shall see in a moment that $|G(p)|$ is a power of p, it follows that $|G(p_j)| = p_j^{m_j}$.

The subgroups $G(p)$ can in general be decomposed further. Since the decomposition is not unique the proof is somewhat harder than the proof of (1.1.1).

THEOREM 1.1.1. *Let G be a finite p group where p is a prime, that is, assume that $p^m G = \{0\}$ for some m. Then one can find integers $r_1 \ge r_2 \ge \cdots \ge r_\sigma \ge 1$ such that*

$$G \cong \mathbf{Z}_{p^{r_1}} \times \cdots \times \mathbf{Z}_{p^{r_\sigma}}. \tag{1.1.2}$$

The sequence $r_1, \dots, r_\sigma$ is uniquely determined although the decomposition is not.

PROOF. Let r_1 be the smallest positive integer m such that $p^m G = 0$, and choose $a \in G$ with $p^{r_1-1}a \ne 0$. Recall that the corresponding cyclic group G_a is then isomorphic to $\mathbf{Z}_{p^{r_1}}$. If $0 \ne \bar{b} \in G/G_a$ then the period p^r of $\bar{b}$ is $\le p^{r_1}$ since every element b in the residue class $\bar{b}$ has period $\le p^{r_1}$. We claim that b can be chosen so that b also has period p^r. In fact, if $p^r b = na$ it follows that $0 = p^{r_1} b = p^{r_1 - r} na$, so n must be divisible by p^r. Then $b' = b - (n/p^r)a$ is in the residue class $\bar{b}$ and $p^r b' = 0$.

By induction with respect to $|G|$ we may assume that the quotient G/G_a is the product of cyclic groups of order $p^{r_2} \ge p^{r_3} \ge \cdots \ge p^{r_\sigma}$ generated by $\bar{b}_2, \dots, \bar{b}_\sigma$. For each of these generators we choose an element $b_j \in G$ with period p^{r_j} in the residue class $\bar{b}_j$. Then

$$G \cong G_a \times G_{b_2} \times \cdots \times G_{b_\sigma}. \tag{1.1.2$'$}$$

In fact, if $g \in G$ there are integers $\gamma_2, \dots, \gamma_\sigma$ uniquely determined modulo $p^{r_2}, \dots, p^{r_\sigma}$ such that $g - \sum_2^\sigma \gamma_j b_j \in G_a$, which proves (1.1.2)$'$, hence (1.1.2).

We can also prove the uniqueness of $r_1, \dots, r_\sigma$ by induction. In fact, since

$$pG \cong \mathbf{Z}_{p^{r_1 - 1}} \times \cdots \times \mathbf{Z}_{p^{r_\sigma - 1}}$$

the numbers r_j which are > 1 are determined, and since $|G| = p^{r_1 + \cdots + r_\sigma}$ the number of exponents equal to 1 can then be calculated.

Note that the only subgroups of $\mathbf{Z}_{p^r}$ are $p^j \mathbf{Z}_{p^r} \cong \mathbf{Z}_{p^{r-j}}$, where $0 \le j \le r$. By the uniqueness in Theorem 1.1.1 it is not possible to decompose $\mathbf{Z}_{p^r}$ into the product of two groups.

EXERCISE 1.1.1. How many non-isomorphic abelian groups of order 128 are there?

In what follows we shall avoid using the structure of G provided by (1.1.1) and (1.1.2), but it is useful to keep in mind that finite abelian groups are not more general than direct products of cyclic groups, which can even be taken of prime power order.

1.2. The dual of a finite abelian group and Fourier expansion. With G still denoting a finite abelian group we shall study the group algebra $\mathbf{C}^G$ consisting of complex valued functions $f : G \to \mathbf{C}$. This is a finite dimensional complex vector space with a natural positive definite hermitian symmetric form

$$(f, g) = \sum_{x \in G} f(x)\overline{g(x)}; \quad f, g \in \mathbf{C}^G. \tag{1.2.1}$$

With τ_y denoting the translation operator

$$(\tau_y f)(x) = f(x - y); \quad x \in G,\ y \in G,$$

it is clear that (f,g) is translation invariant, that is,

$$(\tau_y f, \tau_y g) = (f,g); \quad f,g \in \mathbf{C}^G,\ y \in G.$$

Since $\tau_y\tau_z = \tau_{y+z} = \tau_z\tau_y$ the unitary operators τ_y commute. Every linear operator $T : \mathbf{C}^G \to \mathbf{C}^G$ commuting with τ_y for every $y \in G$ is a linear combination of these operators,

$$T = \sum_{y\in G} c(y)\tau_y, \quad \text{that is,} \quad (Tf)(x) = \sum_{y\in G} f(x-y)c(y).$$

In fact, the linear form $\mathbf{C}^G \ni f \mapsto (Tf)(0)$ can be written

$$(Tf)(0) = \sum_{y\in G} c(y)f(-y), \quad f \in \mathbf{C}^G,$$

which implies that

$$(Tf)(z) = (\tau_{-z}Tf)(0) = (T\tau_{-z}f)(0) = \sum_{y\in G} c(y)f(z-y) = (\sum_{y\in G} c(y)\tau_y f)(z),$$

which proves the claim.

The commuting unitary operators τ_y, $y \in G$, have a common complete orthonormal system of eigenvectors. If χ is an eigenvector for all τ_y, with eigenvalue λ_y, then

$$\tau_y\chi = \lambda_y\chi, \quad \text{that is, } \chi(x-y) = \lambda_y\chi(x);\ x,y \in G.$$

If $\chi(0) = 0$ it follows that $\chi \equiv 0$. Otherwise we can normalize χ so that $\chi(0) = 1$, which gives $\chi(-y) = \lambda_y$, so $\chi(x-y) = \chi(x)\chi(-y)$, or more symmetrically

$$\chi(x+y) = \chi(x)\chi(y); \quad x,y \in G; \quad \chi(0) = 1. \tag{1.2.2}$$

A function χ satisfying (1.2.2) is called a *group character*. Note that if $nx = 0$ then it follows that $1 = \chi(nx) = \chi(x)^n$, so $\chi(x)$ is an n^{th} root of unity, hence a $|G|^{\text{th}}$ root of unity. In particular,

$$\chi(x)^{-1} = \chi(-x) = \overline{\chi(x)}, \quad x \in G. \tag{1.2.3}$$

If χ_1 is also a group character then $\chi\chi_1$ is another and so is $1/\chi$, so the characters form an abelian group $\widehat{G}$, the *dual group of* G, with the character which is identically one as neutral element. Since

$$(\chi,\chi_1) = (\tau_y\chi, \tau_y\chi_1) = \chi(-y)\overline{\chi_1(-y)}(\chi,\chi_1), \quad y \in G,$$

we have $(\chi,\chi_1) = 0$ unless $\chi(-y)\chi_1(y) \equiv 1$, that is, $\chi = \chi_1$. Different characters are thus orthogonal, so we have proved:

THEOREM 1.2.1. *The characters on G form an abelian group $\widehat{G}$. The elements of $\widehat{G}$ divided by $\sqrt{|G|}$ are an orthonormal basis for $\mathbf{C}^G$, so $|\widehat{G}| = |G|$.*

For every $f \in \mathbf{C}^G$ we have

$$f(x) = \sum_{\chi\in\widehat{G}} \chi(x)(f,\chi)/|G| = \sum_{\chi\in\widehat{G}}\sum_{y\in G} \chi(x)f(y)\overline{\chi(y)}/|G|,$$

which means that

$$\hat{f}(\chi) = \sum_{x\in G} f(x)\overline{\chi(x)} = \sum_{x\in G} f(x)\chi(-x) \quad \text{implies}$$
$$f(x) = \frac{1}{|G|}\sum_{\chi\in\widehat{G}} \hat{f}(\chi)\chi(x).$$

This is *Fourier's inversion formula.* For the convolution $(f * g)(x) = \sum_{y\in G} f(x-y)g(y)$ where $f, g \in \mathbf{C}^G$ we have $\widehat{f * g}(\chi) = \hat{f}(\chi)\hat{g}(\chi)$, so the Fourier transformation diagonalizes all translation invariant linear operators in $\mathbf{C}^G$. If we multiply the inversion formula by $\overline{g(x)}$ and sum over x, we obtain *Parseval's formula*

$$(f,g) = \sum_{x\in G} f(x)\overline{g(x)} = \frac{1}{|G|}\sum_{\chi\in\widehat{G}} \hat{f}(\chi)\overline{\hat{g}(\chi)} = (\hat{f},\hat{g})/|G|, \quad f, g \in \mathbf{C}^G,$$

which means that the linear map $f \mapsto |G|^{-\frac{1}{2}}\hat{f}$ is unitary. By the definition of the group operation in $\widehat{G}$ the map $\widehat{G} \ni \chi \mapsto \chi(x)$, $x \in G$, is a character on $\widehat{G}$, so these functions form an orthonormal basis in $\mathbf{C}^{\widehat{G}}$ after division by $\sqrt{|\widehat{G}|}$. Thus we have a complete symmetry between G and $\widehat{G}$ apart from the fact that we have written the group operation in $\widehat{G}$ multiplicatively (and the usual change of sign in Fourier's inversion formula).

REMARK. If all the characters $\chi \in \widehat{G}$ are real then χ only takes the values ± 1 so $\chi^2 = 1$. By Theorem 1.1.1 it follows that $\widehat{G}$ is then isomorphic to $\mathbf{Z}_2^n$ for some positive integer n, which implies that G is also isomorphic to $\mathbf{Z}_2^n$. Representing the elements of G and $\widehat{G}$ by $x \in \mathbf{Z}^n$ and $\xi \in \mathbf{Z}^n$ and writing $\langle x, \xi\rangle = \sum_1^n x_j\xi_j$ we can write $\xi(x) = (-1)^{\langle x,\xi\rangle}$ and obtain if f is a complex valued function on $\mathbf{Z}_2^n$

$$\hat{f}(\xi) = \sum_{x\in\mathbf{Z}_2^n} (-1)^{\langle x,\xi\rangle} f(x), \quad f(x) = 2^{-n}\sum_{\xi\in\mathbf{Z}_2^n} (-1)^{\langle x,\xi\rangle}\hat{f}(\xi).$$

The calculation of $\hat{f}$ seems to require $2^n(2^n-1)$ additions and subtractions. However, if we do it one variable at a time we find that only $2^n n$ such operations are then required. In Section 1.3 we shall see that a similar improvement is possible for the group $\mathbf{Z}_{2^n}$ although it is far less obvious.

If $G = G_1 \times G_2$ is a direct product, then it is clear that $\widehat{G}$ can be identified with $\widehat{G}_1 \times \widehat{G}_2$, for if $\chi_j \in \widehat{G}_j$ then

$$G = G_1 \times G_2 \ni (x_1, x_2) \mapsto \chi_1(x_1)\chi_2(x_2)$$

is a character on G and all characters are of this form. To make the preceding discussion completely explicit it is therefore sufficient to discuss the cyclic group $G = \mathbf{Z}_n$ of order n, not necessarily a prime power. Let $\omega = e^{2\pi i/n}$. If $\xi \in \mathbf{Z}$ then

$$\mathbf{Z} \ni x \mapsto \omega^{x\xi} = e^{2\pi i x\xi/n}$$

defines a character on G. This identifies $\widehat{G}$ with $\mathbf{Z}_n$, so that for a function f on $\mathbf{Z}_n$

$$\begin{aligned} \hat{f}(\xi) &= \sum_{x=0}^{n-1} f(x)\omega^{-x\xi} = \sum_{x=0}^{n-1} f(x)e^{-2\pi i x\xi/n}, \\ f(x) &= \frac{1}{n}\sum_{\xi=0}^{n-1} \hat{f}(\xi)\omega^{x\xi} = \frac{1}{n}\sum_{\xi=0}^{n-1} \hat{f}(\xi)e^{2\pi i x\xi/n}. \end{aligned} \tag{1.2.4}$$

This inversion formula is completely elementary: it follows from the fact that

$$\sum_{\xi=0}^{n-1} e^{2\pi i z\xi/n} = n\delta_{z0}, \quad z = 0, \dots, n-1,$$

where δ_{jk} is the Kronecker delta, equal to 1 when $j = k$ and 0 otherwise. We shall see in Chapter II that it is easy to pass from (1.2.4) to the basic facts on Fourier series and Fourier transforms.

As pointed out above the dual group of $G_1 \times G_2$ is $\widehat{G}_1 \times \widehat{G}_2$, which reduces the Fourier analysis in $G = G_1 \times G_2$ to Fourier analysis in G_1 and in G_2. As a preparation for Section 1.3 we shall now study the more general case where we only have a subgroup H of the finite abelian group G. An example is $G = \mathbf{Z}_{p^k}$ and $H = p^j G$ with $0 < j < k$.

THEOREM 1.2.2. *If H is a subgroup of the finite abelian group G then the characters which are equal to* 1 *on H form a subgroup $H^\perp$ of $\widehat{G}$ which is the dual group of G/H. The dual group of H is $\widehat{G}/H^\perp$.*

PROOF. Let f be a function on G/H lifted to G, that is, let $f \in \mathbf{C}^G$ and $\tau_y f = f$ for $y \in H$. If $\chi \in \widehat{G}$ then

$$(f, \chi) = (\tau_y f, \chi) = (f, \tau_{-y}\chi) = \overline{\chi(y)}(f, \chi), \quad y \in H,$$

which proves that $(f, \chi) = 0$ unless $\chi(y) = 1$ for every $y \in H$, that is, $\chi \in H^\perp$. Hence

$$f = \frac{1}{|G|}\sum_{\chi \in H^\perp} (f, \chi)\chi, \tag{1.2.5}$$

which proves that $H^\perp$ is equal to the dual group of G/H and not only a subgroup, which is obvious. Hence

$$|\widehat{G}| = |G| = |G/H||H| = |H^\perp||H|,$$

which implies that $|H| = |\widehat{G}/H^\perp|$ and proves that $\widehat{G}/H^\perp$ is the dual group of H, for it is obviously a subgroup since a character χ on G restricts to a character on H which is not trivial unless $\chi \in H^\perp$.

We saw in (1.2.5) that if $f \in \mathbf{C}^G$ and $\tau_y f = f$ for every $y \in H$, then $\hat{f}$ vanishes in $\widehat{G} \setminus H^\perp$, and the restriction to $H^\perp$ is $|H|$ times the Fourier transform of the function induced in G/H. (Note that the scalar product $\sum_{x\in G} f(x)\overline{\chi(x)}$ is equal to $|H|$ times the sum with only one x chosen in each residue class mod H.)

Every $f \in \mathbf{C}^G$ can be uniquely written in the form

$$f = \frac{1}{|H|} \sum_{\nu \in \widehat{G}/H^\perp} f_\nu, \quad f_\nu = \frac{|H|}{|G|} \sum_{\chi\in\nu} (f,\chi)\chi. \tag{1.2.6}$$

Note that $\tau_y f_\nu = \overline{\nu(y)} f_\nu$ if $y \in H$. Here $\nu(y)$ is defined since ν is a character on H. We have $\hat{f}_\nu(\chi) = |H|\hat{f}(\chi)$ when $\chi \in \nu$ and $\hat{f}_\nu(\chi) = 0$ when $\chi \notin \nu$, and

$$f_\nu = \sum_{y\in H} \nu(y)\tau_y f, \quad \text{that is,} \quad f_\nu(x) = \sum_{y\in H} f(x+y)\overline{\nu(y)}, \; x \in G, \tag{1.2.7}$$

for

$$\sum_{y\in H} \nu(y)\tau_y f = \frac{1}{|H|} \sum_{\mu\in\widehat{G}/H^\perp} \sum_{y\in H} \nu(y)\tau_y f_\mu = \sum_{\mu\in\widehat{G}/H^\perp} \frac{1}{|H|} \sum_{y\in H} \nu(y)\overline{\mu(y)} f_\mu = f_\nu.$$

Thus $f_\nu(x)$ are for fixed x the Fourier coefficients of f in the fiber $x+H$ of G/H, identified with H by the map $H \ni h \mapsto x+h$.

If we choose $\hat{\nu} \in \widehat{G}$ in the residue class $\nu \in \widehat{G}/H^\perp$, then $\tilde{f}_\nu(x) = f_\nu(x)\hat{\nu}(-x)$ is invariant under the translations τ_y with $y \in H$, for $\hat{\tilde{f}}_\nu(\chi) = |H|\hat{f}(\chi\hat{\nu})$ if $\chi \in H^\perp$ and $\hat{\tilde{f}}_\nu(\chi) = 0$ if $\chi \in \widehat{G} \setminus H^\perp$. Thus it follows that

$$\begin{aligned} \hat{f}(\chi) &= \hat{f}_\nu(\chi)/|H| = \frac{1}{|H|} \sum_{x\in G} f_\nu(x)\overline{\chi(x)} \\ &= \hat{\tilde{f}}_\nu(\chi/\hat{\nu})/|H| = \frac{1}{|H|} \sum_{x\in G} \tilde{f}_\nu(x)\hat{\nu}(x)\overline{\chi(x)}, \quad \chi \in \nu \in \widehat{G}/H^\perp. \end{aligned} \tag{1.2.8}$$

The division by $|H|$ disappears if only one x in each residue class mod H is taken in the sums, which corresponds to taking the Fourier transform of $\tilde{f}_\nu$ as a function in G/H.

Summing up, we can calculate the Fourier transform of f by

(i) computing f_ν using (1.2.7) for one x in each coset in G/H, that is, calculating the Fourier transform of $|G/H|$ functions in the group H;

(ii) calculating the Fourier transforms of the $|H|$ functions $\tilde{f}_\nu$ in G/H, or equivalently, apply (1.2.8).

The importance of this remark is as follows. Computing $\hat{f}$ from first principles means letting a $|G| \times |G|$ matrix act on a vector with $|G|$ (complex) components, which requires $(|G|-1|)^2$ multiplications and $|G|(|G|-1)$ additions, since all elements in one row and one column of the matrix are equal to 1. If instead we divide the task into two steps as above, we need $|G/H||H|(|H|-1) = |G|(|H|-1)$ multiplications and additions in step (i). In step (ii) we need at most $|H||G/H|^2$ multiplications and additions, or altogether at most $|G|(|H|-1+|G/H|)$ operations of each kind. Here $2 \le |H| \le |G|/2$ so this is $\le |G|(1+|G|/2)$. There is a more drastic saving if we have a ladder of subgroups

$$\{0\} = H_0 \subset H_1 \subset H_2 \subset \cdots \subset H_N = G. \tag{1.2.9}$$

In the first step, with $H = H_1$ we have to make $|G|(|H_1|-1)$ operations, and are essentially left with $|H_1|$ functions in G/H_1. To compute their Fourier transforms we use the subgroup H_2/H_1 and find that step (i) requires $|H_1||G/H_1|(|H_2/H_1|-1) = |G|(|H_2|/|H_1|-1)$ operations. Continuing in this way until we reach $H_N = G$ so that no step (ii) is required, the number of operations used becomes altogether

$$|G|\sum_1^N(\gamma_j-1), \quad \gamma_j = |H_j|/|H_{j-1}|. \tag{1.2.10}$$

Here $\gamma_j \ge 2$, and we have $\gamma \le 2^{\gamma-1}$ when $\gamma \ge 2$. Since $\prod_1^N \gamma_j = |G|$ the bound in (1.2.10) is $\ge |G|\log_2|G|$, with strict inequality unless $\gamma_j = 2$ for every j, that is, G is a 2 group. (The case of the group $\mathbf{Z}_2^N$ is fairly trivial and was discussed in a remark after Theorem 1.2.1. The more interesting case of the group $\mathbf{Z}_{2^N}$ will be discussed below.) The bound $|G|\log_2|G|$ for the number of operations is of course much better than the bound $(|G|-1)^2$ if $|G|$ is large, which is the reason for the importance of the fast Fourier transform using the group $\mathbf{Z}_{2^N}$, which will be discussed more explicitly in the next section.

When the group G and the subgroup H are cyclic, the calculation of the Fourier transform of $f \in \mathbf{C}^G$ using (1.2.7) and (1.2.8) can be described explicitly as follows. Let $G = \mathbf{Z}_N$ where $N = ab$ for some integers $a, b \ge 2$, and let $H = a\mathbf{Z}_N$. We identify $\widehat{G}$ with $\mathbf{Z}_N$, defining the characters by $\mathbf{Z}_N \ni x \mapsto \exp(2\pi i x\xi/N)$ when $\xi \in \mathbf{Z}_N$. Then $H^\perp = b\mathbf{Z}_N$, we represent G and $\widehat{G}$ by integers x and ξ in $[0, N-1]$, and write

$$x = ay + z, \quad 0 \le z < a,\ 0 \le y < b; \quad \xi = b\eta + \zeta, \quad 0 \le \zeta < b,\ 0 \le \eta < a.$$

Then $\exp(2\pi i x\xi/N) = \exp(2\pi i z\zeta/N)\exp(2\pi i y\zeta/b)\exp(2\pi i z\eta/a)$, hence

$$\hat{f}(b\eta+\zeta) = \sum_{z=0}^{a-1} e^{-2\pi i z\eta/a}\Big(e^{-2\pi i z\zeta/N}\sum_{y=0}^{b-1} f(ay+z)e^{-2\pi i y\zeta/b}\Big). \tag{1.2.11}$$

With our earlier notation the inner sum in (1.2.11) is the Fourier transform in H calculating $f_\zeta(z)$ in (1.2.7). The product by $e^{-2\pi i z\zeta/N}$ is $\tilde{f}_\zeta(z)$, and the outer sum in (1.2.11) calculates its Fourier transform in G/H as in (1.2.8). Thus formula (1.2.11) describes in the cyclic case precisely the same procedure as before, but with the conceptual background removed. (I owe this observation to C. Bennewitz.)

1.3. The fast Fourier transform. In this section we shall examine in detail the Fourier transform of functions on the group $G = \mathbf{Z}_{2^N} = \mathbf{Z}/2^N\mathbf{Z}$ following the pattern outlined at the end of Section 1.2. Let H_j be the subgroups

$$H_j = 2^{N-j}G, \quad 0 \le j \le N. \tag{1.3.1}$$

We have as in (1.2.9)

$$\{0\} = H_0 \subset H_1 \subset \cdots \subset H_N = G, \quad H_j/H_{j-1} \cong \mathbf{Z}_2,\ 1 \le j \le N. \tag{1.3.2}$$

We can parametrize G by $\{0,1\}^N$ using binary digits,

$$\{0,1\}^N \ni (r_1, \dots, r_N) \mapsto \sum_1^N r_j 2^{j-1} \in \mathbf{Z} \to \mathbf{Z}_{2^N},$$

and we parametrize the dual $\widehat{G} \cong \mathbf{Z}_{2^N}$ similarly by $(\varrho_1, \dots, \varrho_N)$. Then the character $G \times \widehat{G} \ni (x, \xi) \mapsto \exp(2\pi i x\xi/2^N)$ becomes

$$\exp(2\pi i \sum_{j,k\ge 1, j+k\le N+1} r_j \varrho_k 2^{j+k-2-N}).$$

The subgroup H_j consisting of all $x \in \mathbf{Z}_{2^N}$ with $2^j x \equiv 0$ is defined by $r_1 = \cdots = r_{N-j} = 0$, so G/H_j is parametrized by $(r_1, \dots, r_{N-j})$. Since $H_j^\perp$ is defined by $\varrho_1 = \cdots = \varrho_j = 0$ the quotient $\widehat{G}/H_j^\perp$ is parametrized by $\varrho_1, \dots, \varrho_j$.

Let $f \in \mathbf{C}^G$. With $H = H_1$ the first step in the calculation of $\hat{f}$ is to decompose f by the two characters on $H_1 = \mathbf{Z}_2$,

$$f_0(x) = f(x) + f(x + 2^{N-1}), \quad f_1(x) = f(x) - f(x + 2^{N-1}).$$

Here $f_0(x + 2^{N-1}) = f_0(x)$ but $f_1(x + 2^{N-1}) = -f_1(x)$. The non-trivial character which gave rise to f_1 corresponds to the coset of $\widehat{G}/H_1^\perp$ defined by $\varrho_1 = 1$, and we choose in it the character in $\widehat{G}$ corresponding to $\varrho = (1, 0, \dots, 0)$. Then f_1 is modified to

$$\tilde{f}_1(x) = (f(x) - f(x + 2^{N-1}))e^{-2\pi i x/2^N}.$$

To simplify notation we drop the tilde and define now

$$f_{\varrho_1}(x) = (f(x) + (-1)^{\varrho_1} f(x + 2^{N-1}))e^{-2\pi i x \varrho_1/2^N}, \quad \varrho_1 = 0, 1. \tag{1.3.3}$$

These two functions are now defined in $\mathbf{Z}_{2^{N-1}}$. The Fourier transform of f_{ϱ_1} as a function in $\mathbf{Z}_{2^{N-1}}$ at $\sum_2^{N-1} \varrho_j 2^{j-2}$ will be the Fourier transform of f at $\sum_1^{N-1} \varrho_j 2^{j-1}$.

It is now clear how to continue the algorithm. When $f_{\varrho_1 \dots \varrho_k}(x)$ has been defined as a function in $\mathbf{Z}_{2^{N-k}}$ for $\varrho_1, \dots, \varrho_k = 0, 1$, we define if $k < N$, for $\varrho_{k+1} = 0, 1$,

$$f_{\varrho_1 \dots \varrho_{k+1}}(x) = (f_{\varrho_1 \dots \varrho_k}(x) + (-1)^{\varrho_{k+1}} f_{\varrho_1 \dots, \varrho_k}(x + 2^{N-k-1}))e^{-2\pi i x \varrho_{k+1}/2^{N-k}}, \tag{1.3.4}$$

which are functions in $\mathbf{Z}_{2^{N-k-1}}$. The last functions $f_{\varrho_1\dots\varrho_N}$ are defined in $\mathbf{Z}_1$, that is, constants, and

$$\hat{f}\big(\sum_1^N \varrho_j 2^{j-1}\big) = f_{\varrho_1\dots\varrho_N}. \tag{1.3.5}$$

Assuming that the exponentials $e^{-2\pi i\nu/2^N}$ and the binary representation of integers in $[0, 2^N)$ have been precomputed, we have here made 2^{N-1} multiplications and 2^N additions or subtractions in each step, for a total of $N2^{N-1}$ multiplications and $N2^N$ additions or subtractions. Parametrizing $x \in G$ by the binary digits $r_1, \dots, r_N$ we see that $f_{\varrho_1\dots\varrho_k}(x)$ is parametrized by $r_1 \dots r_{N-k}, \varrho_k \dots \varrho_1$ which shows that the function values calculated in each step can be naturally stored at the same places as those in the preceding step which makes the algorithm fast and easy to program.

The algorithm is often presented in a somewhat different way which is independent of the general scheme described in Section 1.2. Let $n = 2^N$. The task is to calculate the polynomial

$$p(z) = \sum_0^{n-1} f(j) z^j$$

at the n^{th} roots of unity, or equivalently, to calculate the residue classes modulo $z - \omega^j$ for $j = 0, \dots, n-1$ where $\omega = \exp(2\pi i/n)$. Since n is even the n^{th} roots of unity can be divided up into solutions of the equations $z^{n/2} = 1$ and $z^{n/2} = -1$. For roots of the first kind we can first reduce p modulo $z^{n/2} - 1$ and for the second type we first reduce modulo $z^{n/2} + 1$. Since n is a power of 2 the procedure can be continued. After $j < N$ steps we have 2^j polynomials q of degree $2^{N-j} - 1$ the values of which we want to calculate at the zeros of $z^{2^{N-j}} - \omega^{2^{N-j}\nu}$ where ν is an integer. Thus we have to compute the values of q at the zeros of $z^{2^{N-j-1}} \mp \omega^{2^{N-j-1}\nu}$ after reducing q modulo these polynomials. For the lower sign we replace ν by $\nu + 2^j$ to preserve the structure. After N steps we are left with 2^N polynomials of degree 0 to evaluate at ω^ν where $\nu = \sum_1^N \varrho_j 2^{j-1}$ and $\varrho_j = 0$ if the minus sign is chosen in the j^{th} step of the construction, $\varrho_j = 1$ otherwise. Now reducing a polynomial $\sum_0^{2\mu-1} a_j z^j$ modulo $z^\mu \mp c$ gives the polynomials

$$\sum_0^{\mu-1} (a_j \pm a_{j+\mu} c) z^j$$

which means that μ multiplications and 2μ additions or subtractions are required to calculate the new coefficients. The total computational effort is of course the same as in the first description of the algorithm.

CHAPTER II

BASIC FOURIER ANALYSIS OF (PERIODIC) FUNCTIONS IN $\mathbf{R}^n$

2.1. The one-dimensional case. In this section we shall study functions on $\mathbf{R}$, its closed subgroups $T\mathbf{Z}$ with $T > 0$, and the quotient groups $\mathbf{R}/T\mathbf{Z}$ (the circle group). We start with the latter case.

Thus let f be a continuous function on $\mathbf{R}$ with period T. For an arbitrary integer $\nu > 0$ we can restrict f to a function f_ν on $\mathbf{Z}_\nu = \mathbf{Z}/\nu\mathbf{Z}$ defined by

$$f_\nu(j) = f(Tj/\nu), \quad j \in \mathbf{Z}, \tag{2.1.1}$$

and apply the Fourier inversion formula (1.2.4) for $\mathbf{Z}_\nu$ to f_ν. The Fourier coefficients of f_ν are

$$c_\nu(k) = \sum_{j=0}^{\nu-1} f(Tj/\nu)e^{-2\pi ijk/\nu}, \quad k \in \mathbf{Z},$$

and the inversion formula states that

$$f(Tj/\nu) = \frac{1}{\nu} \sum_{-\nu/2\le k<\nu/2} c_\nu(k)e^{2\pi ijk/\nu}. \tag{2.1.2}$$

When $\nu \to \infty$ we have by the definition of the Riemann integral $c_\nu(k)/\nu \to c(k)$ where

$$c(k) = \int_0^1 f(Tx)e^{-2\pi ikx}\, dx = \frac{1}{T}\int_0^T f(x)e^{-2\pi ikx/T}\, dx. \tag{2.1.3}$$

From (2.1.2) we would obtain

$$f(x) = \sum_{-\infty}^{\infty} c(k)e^{2\pi ikx/T}, \tag{2.1.4}$$

if it is legitimate to pass to the limit when $\nu \to \infty$ and $Tj/\nu \to x$. We shall now prove that this is permissible if $f \in C^2$, thanks to the precaution of shifting the summation index k in (2.1.2) so that k/ν does not come close to any integer $\neq 0$. To estimate $c_\nu(k)$ we observe that

$$c_\nu(k)e^{\pm 2\pi ik/\nu} = \sum_{j=0}^{\nu-1} f(T(j\pm 1)/\nu)e^{-2\pi ijk/\nu}, \quad \text{hence}$$

$$c_\nu(k)(2 - e^{2\pi ik/\nu} - e^{-2\pi ik/\nu}) = \sum_{j=0}^{\nu-1} \big(2f(Tj/\nu) - f(T(j+1)/\nu) - f(T(j-1)/\nu)\big)e^{-2\pi ijk/\nu}.$$

The parenthesis in the left-hand side is $-(e^{\pi ik/\nu} - e^{-\pi ik/\nu})^2 = 4\sin^2(\pi k/\nu) \geq 16(k/\nu)^2$ when $|k| \leq \nu/2$, for $|\sin(\pi x)| \geq 2|x|$ when $|x| \leq 1/2$ by the concavity of the sin function in $[0, \pi/2]$. The second difference of f on the right is bounded by $(T/\nu)^2 \sup|f''|$, so we have

$$|c_\nu(k)/\nu| \leq (T^2/16k^2)\sup|f''|, \quad k \neq 0; \quad |c_\nu(0)/\nu| \leq \sup|f|.$$

Thus $|c_\nu(k)/\nu|$ is majorized by a convergent series which proves that the inversion formula (2.1.4) is valid.

PROPOSITION 2.1.1. *If $f \in C^\mu(\mathbf{R})$ where μ is an integer ≥ 0, and f is periodic with period $T > 0$ then the Fourier coefficients $c(k)$ defined for $k \in \mathbf{Z}$ by* (2.1.3) *have the bound*

$$|(2\pi k/T)^\mu c(k)| \leq \frac{1}{T}\int_0^T |f^{(\mu)}(x)|\,dx \leq \sup|f^{(\mu)}|. \tag{2.1.5}$$

If $\mu \geq 2$ then (2.1.4) *is valid with absolute and uniform convergence, and Parseval's formula*

$$\int_0^T |f(x)|^2\,dx = T\sum_{-\infty}^{\infty}|c(k)|^2 \tag{2.1.6}$$

is valid. Conversely, if $c(k)$, $k \in \mathbf{Z}$, is a given sequence $\in \mathbf{C}$ such that $k^\mu c(k)$ is bounded then (2.1.4) *defines a function $f \in C^{\mu-2}(\mathbf{R})$ with period T if $\mu \geq 2$, and* (2.1.3) *is valid.*

PROOF. We have just proved that (2.1.4) is valid if $f \in C^2$, and (2.1.6) follows if we multiply by $\overline{f(x)}$ and integrate, interchanging the integration and the summation. If $f \in C^\mu(\mathbf{R})$ then we obtain by μ partial integrations in (2.1.3)

$$(2\pi ik/T)^\mu c(k) = \frac{1}{T}\int_0^T f^{(\mu)}(x)e^{-2\pi ikx/T}\,dx,$$

which proves (2.1.5). On the other hand, if $c(k)$ is given with $|c(k)| \leq Ck^{-\mu}$, $\mu \geq 2$, then (2.1.4) converges to a function $f \in C^{\mu-2}$ with period T, for the series is uniformly convergent and remains so after at most $\mu-2$ differentiations. The Fourier coefficients of f can be calculated by termwise integration which gives (2.1.3) in view of the orthogonality relations

$$\frac{1}{T}\int_0^T e^{2\pi i(j-k)x/T}\,dx = \delta_{jk}.$$

REMARK. Note that the functions $\chi(x) = \exp(2\pi ikx/T)$ with $k \in \mathbf{Z}$ are bounded continuous characters on $\mathbf{R}/T\mathbf{Z}$, that is, $\chi(x+y) = \chi(x)\chi(y)$ for $x, y \in \mathbf{R}$. We leave as an exercise to prove that all bounded continuous characters are of this form. (Hint: Prove first by integration with respect to y that a continuous character is continuously differentiable and derive a differential equation for it to prove that it is an exponential.)

If f is a C^2 function on $\mathbf{R}$ which is not periodic but has compact support, we can define a function $f_T \in C^2$ with period T by pushforward of f to the quotient space $\mathbf{R}/T\mathbf{Z}$, that is,

$$f_T(x) = \sum_{k=-\infty}^{\infty} f(x+kT).$$

The Fourier coefficients of f_T are

$$c_T(k) = \frac{1}{T}\int_0^T f_T(x)e^{-2\pi ikx/T}\,dx = \frac{1}{T}\int_{-\infty}^{\infty} f(x)e^{-2\pi ikx/T}\,dx = \frac{1}{T}\hat f(2\pi k/T),$$

where $\hat f$ denotes the *Fourier transform*

$$\hat f(\xi) = \int_{\mathbf{R}} f(x)e^{-ix\xi}\,dx. \tag{2.1.7}$$

By (2.1.4) we have

$$f_T(x) = \sum_{-\infty}^{\infty} c_T(k)e^{2\pi ikx/T} = \frac{1}{T}\sum_{-\infty}^{\infty} \hat f(2\pi k/T)e^{2\pi ixk/T}\,dx. \tag{2.1.8}$$

When $T\to\infty$ the left-hand side converges to $f(x)$; in fact on any compact set it is equal to $f(x)$ when T is large enough. To find the limit of the sum in the right-hand side of (2.1.8) we note that partial integration gives

$$\xi^2\hat f(\xi) = -\int f''(x)e^{-ix\xi}\,dx, \quad \text{hence } (1+\xi^2)|\hat f(\xi)| \le \int (|f''(x)| + |f(x)|)\,dx.$$

We can regard the sum in (2.1.8) as the integral of the function of ξ which is equal to $\hat f(2\pi k/T)e^{2\pi ixk/T}$ when $|\xi - k/T| < 1/2T$. It is bounded by $C/(1+\xi^2)$ for every $T > 1$ so the dominated convergence theorem gives

$$f(x) = \int_{\mathbf{R}} \hat f(2\pi\xi)e^{2\pi ix\xi}\,d\xi = \frac{1}{2\pi}\int_{\mathbf{R}} \hat f(\xi)e^{ix\xi}\,d\xi, \quad x\in\mathbf{R}. \tag{2.1.9}$$

PROPOSITION 2.1.2. *If $f\in C^\mu(\mathbf{R})$ and $x^j f^{(k)}(x)$ is bounded for $0\le j\le\mu$, $0\le k\le\nu$, where $\mu\ge 2$, then the Fourier transform $\hat f$ of f defined by* (2.1.7) *is in $C^{\mu-2}(\mathbf{R})$, and*

$$|\xi^j \hat f^{(k)}(\xi)| \le \int_{\mathbf{R}} |(d/dx)^j(x^k f(x))|\,dx \le \pi^{-1}\sup(1+x^2)|(d/dx)^j(x^k f(x))|, \tag{2.1.10}$$

when $k\le\mu-2$ and $j\le\nu$. If $\nu\ge 2$ also then $\hat f$ is integrable, the Fourier inversion formula (2.1.9) *is valid, and so are Parseval's formula*

$$\int_{\mathbf{R}} |f(x)|^2\,dx = \frac{1}{2\pi}\int_{\mathbf{R}} |\hat f(\xi)|^2\,d\xi, \tag{2.1.11}$$

and Poisson's summation formula

$$\sum_{-\infty}^{\infty} f(kT) = \frac{1}{T}\sum_{-\infty}^{\infty}\hat f(2\pi k/T). \tag{2.1.12}$$

PROOF. If $k \le \mu - 2$ then we may differentiate (2.1.7) k times under the integral sign, which gives

$$\hat{f}^{(k)}(\xi) = \int_{\mathbf{R}} f(x)(-ix)^k e^{-ix\xi}\,dx,$$

for the integral remains uniformly convergent. If $j \le \nu$ we can multiply by ξ^j and integrate by parts j times to get

$$\xi^j \hat{f}^{(k)}(\xi) = \int_{\mathbf{R}} \big((-id/dx)^j (f(x)(-ix)^k)\big) e^{-ix\xi}\,dx,$$

which proves (2.1.10). In particular, if $\nu \ge 2$ and $\mu \ge 2$ the proof of the inversion formula given above for $f \in C^2$ of compact support remains valid, for $f_T(x) \to f(x)$ as $T \to \infty$ since $\mu \ge 2$, and the determination of the limit of the right-hand side of (2.1.8) used only a special case of (2.1.10). Parseval's formula (2.1.11) follows if we multiply (2.1.9) by $\overline{f(x)}$ and integrate, interchanging the orders of integration in the right-hand side. Poisson's summation formula (2.1.12) is the special case of (2.1.8) with $x = 0$.

The gap between the periodic and the non-periodic case can be bridged as follows. Let f again be as in Proposition 2.1.2. Then we have used that

$$f_T(x) = \sum_{-\infty}^{\infty} f(x + kT)$$

is periodic with period T, but f_T only preserves information on the spectrum of f at $2\pi\mathbf{Z}/T$. To avoid this loss of information one can premultiply by the character $e^{-ix\xi}$ and define

$$f_{T,\xi}(x)e^{-ix\xi} = \sum_{-\infty}^{\infty} f(x + kT)e^{-i(x+kT)\xi}, \quad \text{that is,}$$

$$f_{T,\xi}(x) = \sum_{-\infty}^{\infty} f(x + kT)e^{-ikT\xi}. \tag{2.1.13}$$

(Compare this with (1.2.6) and (1.2.7).) Then $F(x,\xi) = f_{T,\xi}(x)$ is continuous in $\mathbf{R}^2$ and

$$F(x, \xi + 2\pi/T) = F(x,\xi), \quad F(x + T, \xi) = e^{iT\xi}F(x,\xi). \tag{2.1.14}$$

The Fourier coefficients of F as a periodic function of ξ are $f(x - kT)$, so Parseval's formula for Fourier series gives

$$\int_0^{2\pi/T} |F(x,\xi)|^2\,d\xi = \frac{2\pi}{T} \sum_{-\infty}^{\infty} |f(x - kT)|^2,$$

and integration with respect to x yields

$$\int_0^T dx \int_0^{2\pi/T} |F(x,\xi)|^2\,d\xi = \frac{2\pi}{T} \int_{\mathbf{R}} |f(x)|^2\,dx. \tag{2.1.15}$$

Conversely, let $F \in C^2(\mathbf{R} \times \mathbf{R})$ satisfy the periodicity conditions (2.1.14) and set

$$f(x) = \frac{T}{2\pi} \int_0^{2\pi/T} F(x, \xi)\, d\xi. \tag{2.1.16}$$

Then $f \in C^2$, and since for $0 \neq \nu \in \mathbf{Z}$

$$f(x + \nu T) = \frac{T}{2\pi} \int_0^{2\pi/T} e^{i\nu T\xi} F(x, \xi)\, d\xi = -\frac{T}{2\pi}(\nu T)^{-2} \int_0^{2\pi/T} e^{i\nu T\xi} \partial^2 F(x, \xi)/\partial \xi^2\, d\xi,$$

it follows that $x^j f(x)$ is bounded for $j \leq 2$. Fourier series expansion of F gives with the Fourier coefficients just calculated

$$F(x, \xi) = \sum_{-\infty}^{\infty} f(x - \nu T) e^{i\nu T\xi} = f_{T,\xi}(x),$$

so (2.1.16) and (2.1.13) are inverse transformations. In particular, using (2.1.15) we conclude that the map from f to $[0, T] \times [0, 2\pi/T] \ni (x, \xi) \mapsto \sqrt{T/2\pi} f_{T,\xi}(x)$ extends to a unitary map from $L^2(\mathbf{R})$ to $L^2([0, T] \times [0, 2\pi/T])$. It is sometimes called the Zak transform (see Daubechies [1]), and sometimes called expansion in Bloch waves.

The preceding two propositions are precisely analogous to the Fourier analysis for finite abelian groups in Chapter I, with $(2\pi/T)\mathbf{Z}$ as dual group of $\mathbf{R}/T\mathbf{Z}$ and $\mathbf{R}$ as its own dual group. However, the local and global hypotheses are too strong in both of them and they will be relaxed later on. Before doing so we shall discuss an extension of Propositions 2.1.1 and 2.1.2 to distributions.

Distributions $f \in \mathscr{D}'(\mathbf{R})$ with period T can be identified with continuous linear forms on the C^∞ functions on $\mathbf{R}$ with period T. In fact, let L be such a linear form. We can define a distribution $f \in \mathscr{D}'(\mathbf{R})$ by setting

$$f(\varphi) = L(\Phi), \quad \varphi \in C_0^\infty(\mathbf{R}), \quad \text{where } \Phi(x) = \sum_{-\infty}^{\infty} \varphi(x - kT), \tag{2.1.17}$$

for $\Phi \in C^\infty(\mathbf{R})$ is periodic with period T. We leave as an exercise to verify that f is a distribution. It is periodic since Φ does not change if $\varphi(x)$ is replaced by $\varphi(x - T)$. Conversely, assume given $f \in \mathscr{D}'(\mathbf{R})$ with period T. We want to define $L(\Phi)$ for $\Phi \in C^\infty(\mathbf{R})$ of period T so that (2.1.17) is valid when $\varphi \in C_0^\infty(\mathbf{R})$. This is a unique definition, for if $\varphi \in C_0^\infty(\mathbf{R})$ and $\sum_{-\infty}^{\infty} \varphi(x - kT) \equiv 0$ then

$$\varphi_1(x) = \sum_{-\infty}^{0} \varphi(x - kT) = -\sum_{1}^{\infty} \varphi(x - kT)$$

is in $C_0^\infty(\mathbf{R})$ and $\varphi(x) = \varphi_1(x) - \varphi_1(x + T)$, so $f(\varphi) = f(\varphi_1) - f(\varphi_1(\cdot + T)) = 0$ by the periodicity of f. On the other hand, we can choose $\psi \in C_0^\infty(\mathbf{R})$ so that $\sum_{-\infty}^{\infty} \psi(x - kT) \equiv 1$, for if $\psi_1 \in C_0^\infty(\mathbf{R})$ and $\psi_1 \geq 0$ with strict inequality in $[0, T]$, then we can take

$\psi(x) = \psi_1(x)/\sum_{-\infty}^{\infty} \psi_1(x - kT)$. For any $\Phi \in C^\infty(\mathbf{R})$ it follows that we can take $\varphi = \psi\Phi$ in (2.1.17), so $L(\Phi) = f(\psi\Phi)$ is defined and is obviously a continuous linear form on periodic $\Phi \in C^\infty$. If $f \in L^1(\mathbf{R}/T\mathbf{Z})$, that is, f is a locally integrable function with period T, it is clear that

$$L(\Phi) = \int_{\mathbf{R}} f(x)\varphi(x)\,dx = \int_0^T f(x)\sum_{-\infty}^{\infty}\varphi(x - kT)\,dx = \int_0^T f(x)\Phi(x)\,dx. \tag{2.1.18}$$

Thus $L(\Phi)$ should be thought of as $f\Phi$ integrated over a period, and we shall usually write $\langle f, \Phi\rangle_{\mathbf{R}/T\mathbf{Z}}$ or even $\langle f, \Phi\rangle_{[0,T]}$ or $\int_0^T f\Phi\,dx$ instead of $L(\Phi)$.

In particular, the Fourier coefficients $c(k)$ can be defined for every $f \in \mathscr{D}'(\mathbf{R})$ with period T by

$$c(k) = \frac{1}{T}\langle f, e^{-2\pi i k\cdot/T}\rangle_{\mathbf{R}/T\mathbf{Z}}. \tag{2.1.3$'$}$$

If f is of order μ it follows that

$$|c(k)| \le C(1 + |k|)^\mu. \tag{2.1.19}$$

If $\Phi \in C^\infty(\mathbf{R}/T\mathbf{Z})$ has Fourier coefficients Φ_k, then it follows from Proposition 2.1.1 that

$$\overline{\Phi(x)} = \sum_{-\infty}^{\infty} \overline{\Phi_k} e^{-2\pi i k x/T}$$

where the Fourier series converges in C^∞. Hence we obtain the polarized version of Parseval's formula

$$\langle f, \overline{\Phi}\rangle_{\mathbf{R}/T\mathbf{Z}} = T\sum_{-\infty}^{\infty} c_k\overline{\Phi_k}.$$

It proves that $f = 0$ if all the Fourier coefficients of f vanish. Now (2.1.19) implies that the Fourier series

$$F = \sum_{-\infty}^{\infty} c(k)e^{2\pi i k x/T} \tag{2.1.20}$$

converges in $\mathscr{D}'$, for the series

$$\sum_{-\infty}^{\infty} c(k)(ki + 1)^{-N} e^{2\pi i k x/T}$$

converges absolutely and uniformly if N is an integer $> \mu+1$, and if we apply the differential operator $((T/2\pi)\partial/\partial x + 1)^N$, which is continuous in $\mathscr{D}'$, it follows that (2.1.20) is also convergent. That the Fourier coefficients of F defined by (2.1.20) are equal to $c(k)$ follows as in the proof of Proposition 2.1.1. Thus we have proved:

THEOREM 2.1.3. *If $f \in \mathscr{D}'(\mathbf{R})$ is periodic with period T, then the Fourier coefficients defined by* (2.1.3)$'$ *have the polynomial bound* (2.1.19) *and the Fourier series* (2.1.4) *converges to f in $\mathscr{D}'(\mathbf{R})$. Conversely, for every sequence $c(k)$ satisfying* (2.1.19) *the series* (2.1.20) *converges in $\mathscr{D}'(\mathbf{R})$ to a distribution F which is periodic with period T and has Fourier coefficients $c(k)$.*

Summing up, taking Fourier coefficients gives an isomorphism between periodic distributions and sequences of at most polynomial growth.

Theorem 2.1.3 was based on the duality between C^∞ periodic functions and periodic distributions on one hand, and between rapidly decreasing and polynomially bounded sequences on the other. Theorem 2.1.2 suggests introducing a class of test functions giving an analogue for distributions on $\mathbf{R}$.

DEFINITION 2.1.4. By $\mathscr{S}$ or $\mathscr{S}(\mathbf{R})$ we shall denote the space consisting of all $\varphi \in C^\infty(\mathbf{R})$ such that $x^j\varphi^{(k)}(x)$ is bounded for arbitrary j and k.

$\mathscr{S}$ is called the *Schwartz space*. The boundedness of $x^j(d/dx)^k\varphi$ for all j and k implies that any product of multiplication by x and differentiation with respect to x applied to $\varphi(x)$ gives a bounded function, for the commutation relation

$$(\frac{d}{dx}x - x\frac{d}{dx})\psi = \psi$$

allows us to change the order of the factors, introducing only new terms with fewer factors x and d/dx. The space $\mathscr{S}$ is a Fréchet space with the seminorms

$$\mathscr{S} \ni \varphi \mapsto \sup |x^j\varphi^{(k)}(x)|,$$

and $C_0^\infty(\mathbf{R})$ is a dense subspace of $\mathscr{S}$. We leave the proof as an exercise. (See Hörmander [1, Lemma 7.1.8] if necessary. Note that $\mathscr{S}(\mathbf{R})$ is dense in $L^p(\mathbf{R})$ for $1 \le p < \infty$ for even $C_0^\infty(\mathbf{R})$ is dense.) This implies that the dual space $\mathscr{S}'(\mathbf{R})$ of *temperate distributions* can be identified with a subspace of the space $\mathscr{D}'(\mathbf{R})$ of distributions on $\mathbf{R}$; the restriction of a continuous linear form L on $\mathscr{S}(\mathbf{R})$ to $C_0^\infty(\mathbf{R})$ is obviously a distribution, and if it is equal to 0 then $L = 0$ since $C_0^\infty(\mathbf{R})$ is dense in $\mathscr{S}(\mathbf{R})$.

The importance of the space $\mathscr{S}$ is due to the fact that by Proposition 2.1.2 the Fourier transformation $f \mapsto \hat{f}$ is continuous from $\mathscr{S}$ to $\mathscr{S}$. Since $\hat{\hat{f}}(x) = 2\pi f(-x)$ by the Fourier inversion formula (2.1.9), it is a surjective map with continuous inverse. If $f \in L^1(\mathbf{R})$ so that the Fourier transform $\hat{f}(\xi)$ can be defined by (2.1.7), then we obtain

$$\int \hat{f}(\xi)\varphi(\xi)\,d\xi = \int f(x)\hat{\varphi}(x)\,dx, \quad \varphi \in \mathscr{S},$$

if we multiply (2.1.7) by $\varphi(\xi)$ and integrate, interchanging the order of integration in the right-hand side. We can therefore extend the definition of the Fourier transformation to arbitrary $f \in \mathscr{S}'$ by defining

$$\langle \hat{f}, \varphi\rangle = \langle f, \hat{\varphi}\rangle, \quad \varphi \in \mathscr{S}, \tag{2.1.21}$$

for $\mathscr{S} \ni \varphi \mapsto \langle f, \hat{\varphi}\rangle$ is a continuous linear form on $\mathscr{S}$ since the maps $\mathscr{S} \ni \varphi \mapsto \hat{\varphi} \in \mathscr{S}$ and $\mathscr{S} \ni \hat{\varphi} \mapsto \langle f, \hat{\varphi}\rangle$ are continuous.

THEOREM 2.1.5. *The Fourier transformation defined by* (2.1.21) *is an isomorphism* $\mathscr{S}'(\mathbf{R}) \to \mathscr{S}'(\mathbf{R})$, *and Fourier's inversion formula is valid, that is,* $\hat{\hat{f}} = 2\pi\check{f}$ *where* $\check{f}$ *is the reflection in the origin defined by*

$$\langle \check{f}, \varphi\rangle = \langle f, \check{\varphi}\rangle, \quad \varphi \in \mathscr{S}(\mathbf{R}); \quad \check{\varphi}(x) = \varphi(-x).$$

We have $f \in L^2(\mathbf{R})$ *if and only if* $\hat{f} \in L^2(\mathbf{R})$, *and Parseval's formula* (2.1.11) *is then valid. When* $f \in \mathscr{S}'$ *the Fourier transform of* $-idf/dx$ *(resp.* xf*) is* $\xi\hat{f}$ *(resp.* $id\hat{f}/d\xi$*), where* x *(resp.* ξ*) denotes the variable where* f *(resp.* $\hat{f}$*) lives. The Fourier transform of* $f(\cdot + h)$ *is* $e^{ih\cdot}\hat{f}$, *and the Fourier transform of* $e^{ih\cdot}f$ *is* $\hat{f}(\cdot - h)$ *if* $h \in \mathbf{R}$. *When* $0 \neq a \in \mathbf{R}$ *the Fourier transform of* $f(a\cdot)$ *is* $|a|^{-1}\hat{f}(\cdot/a)$.

PROOF. If $f \in \mathscr{S}'$ and $\varphi \in \mathscr{S}$ then

$$\langle \hat{\hat{f}}, \varphi\rangle = \langle \hat{f}, \hat{\varphi}\rangle = \langle f, \hat{\hat{\varphi}}\rangle = 2\pi\langle f, \check{\varphi}\rangle = 2\pi\langle \check{f}, \varphi\rangle$$

by Fourier's inversion formula for $\mathscr{S}$, so it is inherited by $\mathscr{S}'$. The same is true for the other rules of computation; for example,

$$\langle -idf/dx, \hat{\varphi}\rangle = \langle f, id\hat{\varphi}/dx\rangle, \quad \langle xf, \hat{\varphi}\rangle = \langle f, x\hat{\varphi}\rangle,$$

and the proof of Proposition 2.1.2 contains a proof that $id\hat{\varphi}(x)/dx$ is the Fourier transform of $\xi \mapsto \xi\varphi(\xi)$ and that $x\hat{\varphi}(x)$ is the Fourier transform of $\xi \mapsto -id\varphi(\xi)/d\xi$. Hence

$$\langle -idf/dx, \hat{\varphi}\rangle = \langle \hat{f}, \xi\varphi(\xi)\rangle = \langle \xi\hat{f}, \varphi\rangle, \quad \langle xf, \hat{\varphi}\rangle = \langle \hat{f}, -i\varphi'\rangle = \langle i\hat{f}', \varphi\rangle.$$

The verification of the other rules is left as an exercise. If $f \in L^2$ then

$$|\langle \hat{f}, \varphi\rangle| = |\langle f, \hat{\varphi}\rangle| \leq \|f\|_{L^2}\|\hat{\varphi}\|_{L^2} = \|f\|_{L^2}\sqrt{2\pi}\|\varphi\|_{L^2}, \quad \varphi \in \mathscr{S}.$$

Since $\mathscr{S}$ is dense in L^2 and every continuous linear form on L^2 is the scalar product by a function in L^2 it follows that $\hat{f} \in L^2$ and that $\|\hat{f}\|_{L^2} \leq \sqrt{2\pi}\|f\|_{L^2}$. Hence

$$\|\hat{\hat{f}}\|_{L^2} \leq \sqrt{2\pi}\|\hat{f}\|_{L^2} \leq 2\pi\|f\|_{L^2},$$

and since $\hat{\hat{f}} = 2\pi\check{f}$ it follows that there is equality throughout, so $f \mapsto \hat{f}/\sqrt{2\pi}$ is a unitary operator in L^2.

We give a few important examples, leaving for the reader to fill in the details.

EXAMPLES. **1.** The Fourier transform of δ_a, $a \in \mathbf{R}$, is $\xi \mapsto e^{-ia\xi}$. Hence the Fourier transform of $\xi \mapsto e^{ia\xi}$ is $2\pi\delta_a$. In particular the Fourier transform of the function which is identically 1 is $2\pi\delta_0$.

2. The Fourier transform of the characteristic function of $\mathbf{R}_\pm$ is $\mp i(\xi \mp i0)^{-1}$, for it is the limit as $\varepsilon \to \pm 0$ of the Fourier transform of the product by $e^{-\varepsilon x}$. Hence the Fourier transform of $x \mapsto \operatorname{sgn} x$ is $-2i\,\mathrm{vp}(1/\xi)$, that is, the limit in $\mathscr{S}'$ of $\xi \mapsto -2i\xi/(\xi^2+\varepsilon^2)$ as $\varepsilon \to 0$. The Fourier transform of $\mathrm{vp}(1/x)$ is $\xi \mapsto -i\pi \operatorname{sgn}\xi$.

3. The Fourier transform of $P_a = \sum_{-\infty}^{\infty} \delta_{ka}$ where $a > 0$ is equal to $(2\pi/a)P_{2\pi/a}$, for Poisson's summation formula (2.1.12) gives

$$\langle P_T, \varphi\rangle = T^{-1}\langle P_{2\pi/T}, \hat\varphi\rangle, \quad \varphi \in \mathscr{S},$$

which means that the Fourier transform of $P_{2\pi/T}$ is equal to TP_T.

4. If $f \in \mathscr{D}'(\mathbf{R})$ is periodic with period T, then $f \in \mathscr{S}'(\mathbf{R})$, and if the Fourier coefficients are defined by $(2.1.3)'$ then

$$\hat f = 2\pi \sum_{-\infty}^{\infty} c(k)\delta_{2\pi k/T}.$$

This follows from Example 1 above since f is the limit in $\mathscr{S}'$ of the partial sums of the Fourier series. Note that Example 3 is a special case.

5. For the Gaussian $g(x) = e^{-x^2/2}$ we have by Cauchy's integral formula

$$\hat g(\xi) = \int_{\mathbf{R}} e^{-x^2/2 - ix\xi}\, d\xi = e^{-\xi^2/2}\int_{\mathbf{R}} e^{-(x-i\xi)^2/2}\, dx = cg(\xi), \quad c = \int_{\mathbf{R}} e^{-x^2/2}\, dx > 0.$$

By Fourier's inversion formula $2\pi g(-x) = \hat{\hat g}(x) = c\hat g(x) = c^2 g(x)$, so $c = \sqrt{2\pi}$. Hence

$$\int e^{-x^2/2 - ix\xi}\, dx = \sqrt{2\pi} e^{-\xi^2/2}, \quad \xi \in \mathbf{R}.$$

Replacing x by $x\sqrt{a}$ and ξ by $\xi/\sqrt{a}$, $a > 0$, we obtain

$$\int e^{-ax^2/2 - ix\xi}\, dx = \sqrt{2\pi/a}\, e^{-\xi^2/2a},$$

for $a > 0$ and $\xi \in \mathbf{R}$. The integral is well defined for arbitrary $a, \xi \in \mathbf{C}$ with $\operatorname{Re} a > 0$, so by analytic continuation the formula remains valid then, with the square root chosen in the right half plane. Hence the Fourier transform of the general Gaussian $x \mapsto \exp(-ax^2/2 + bx)$ is $\xi \mapsto \sqrt{2\pi/a}\exp(-(\xi+ib)^2/2a)$ for arbitrary $a, b \in \mathbf{C}$ with $\operatorname{Re} a > 0$.

For further important examples see Hörmander [1, Lemma 7.1.17].

Since $\mathscr{S}(\mathbf{R})$ is continuously embedded in $C^\infty(\mathbf{R})$, we have $\mathscr{E}'(\mathbf{R}) \subset \mathscr{S}'(\mathbf{R})$; in fact, the obvious map $\mathscr{E}' \to \mathscr{S}'$ is injective since the composition with the injection $\mathscr{S}' \to \mathscr{D}'$ is the usual injection of $\mathscr{E}'$ in $\mathscr{D}'$ (see Hörmander [1, Theorem 2.3.1]).

THEOREM 2.1.6. *If $f \in \mathscr{E}'(\mathbf{R})$ then the Fourier transform is the C^∞ function*

$$\hat f(\xi) = \langle f_x, e^{-ix\xi}\rangle, \quad \xi \in \mathbf{R}, \tag{2.1.22}$$

where f_x means that f acts as a distribution in the x variable. We can extend $\hat f$ to an entire analytic function in $\mathbf{C}$, the Fourier-Laplace transform, by letting $\xi \in \mathbf{C}$ in (2.1.22). If $\operatorname{supp} f \subset [a,b]$ where $a \le b$, and f is of order μ, then $F(\zeta) = \hat f(\zeta)$ has a bound

$$|F(\zeta)| \le C(1+|\zeta|)^\mu \exp h(\operatorname{Im}\zeta), \quad \zeta \in \mathbf{C}; \quad h(\eta) = \max(a\eta, b\eta). \tag{2.1.23}$$

Conversely, if F is an entire analytic function such that (2.1.23) *is valid, then $F = \hat{f}$ where $f \in \mathscr{E}'$ has support in $[a, b]$ and order $\le \max(0, \mu+2)$. If $g \in \mathscr{E}'(\mathbf{R})$ then the Fourier transform of the convolution $f * g$ is $\xi \mapsto \hat{f}(\xi)\hat{g}(\xi)$.*

PROOF. If $\varphi \in C_0^\infty(\mathbf{R})$ then

$$\hat{\varphi}(\xi) = \lim_{\varepsilon \to 0} \varepsilon \sum_{-\infty}^{\infty} \varphi(\varepsilon k) e^{-i\varepsilon k \xi}$$

where $|k| \le C/\varepsilon$ in the sum and the convergence is uniform for ξ in any compact set and remains so after any number of differentiations with respect to ξ under the summation sign. Hence

$$\langle \hat{f}, \varphi \rangle = \langle f, \hat{\varphi} \rangle = \lim_{\varepsilon \to 0} \varepsilon \sum_{-\infty}^{\infty} \varphi(\varepsilon k) F(\varepsilon k),$$

where $F(\zeta) = \langle f, e_\zeta \rangle$ and $e_\zeta(x) = e^{-ix\zeta}$, $\zeta \in \mathbf{C}$. Since the map $\mathbf{C} \ni \zeta \mapsto e_\zeta \in C^\infty(\mathbf{R})$ is continuous, it follows that F is continuous in $\mathbf{C}$, hence the Riemann sums converge and we obtain $\langle \hat{f}, \varphi \rangle = \int \varphi(\xi) F(\xi)\, d\xi$, $\varphi \in C_0^\infty(\mathbf{R})$, which means that $\hat{f} = F$ in $\mathbf{R}$. Since the power series expansion

$$e^{-ix\zeta} = \sum_0^\infty (-ix\zeta)^j / j!$$

converges in $C^\infty(\mathbf{R})$ as a function of x, uniformly for ζ in any compact subset of $\mathbf{C}$, it follows that

$$F(\zeta) = \sum_0^\infty (-i\zeta)^j \langle f, x^j \rangle / j!, \quad \zeta \in \mathbf{C},$$

which proves that F is an entire analytic function. The definition of the convolution shows that

$$f * e_\zeta(x) = f_y(e_\zeta(x-y)) = e_\zeta(x) f_y(e_\zeta(-y)) = \hat{f}(\zeta) e_\zeta(x), \quad \zeta \in \mathbf{C},\ x \in \mathbf{R}.$$

If g is another distribution in $\mathscr{E}'$ it follows that

$$(f * g) * e_\zeta = f * (g * e_\zeta) = \hat{f}(\zeta)\hat{g}(\zeta) e_\zeta, \quad \zeta \in \mathbf{C},$$

which proves that $\hat{f}\hat{g}$ is the Fourier-Laplace transform of $f * g$.

To prove that (2.1.23) is valid for $F = \hat{f}$ we note that

$$|f(\varphi)| \le C \sup_{a \le x \le b} \sum_0^\mu |\varphi^{(j)}(x)|, \quad \varphi \in C^\mu(\mathbf{R}),$$

for we can redefine φ outside $[a, b]$ by the Taylor expansions of order μ at a and b without changing $f(\varphi)$. (See Hörmander [1, Theorem 2.3.3].) Hence

$$|\hat{f}(\zeta)| \le C(1 + |\zeta|)^\mu \exp(\sup_{a \le x \le b} x \operatorname{Im} \zeta),$$

which proves (2.1.23).

Now assume given an entire analytic function F satisfying (2.1.23). At first we assume that $\mu < -1$. Then

$$\widehat{F}(x) = \int_{\mathbf{R}} F(\xi)e^{-ix\xi}\,dx, \quad x \in \mathbf{R},$$

is a bounded continuous function. By Cauchy's integral formula applied to a rectangle with corners at $\pm R$ and $\pm R + i\eta$ we find when $R \to \infty$ that for any $\eta \in \mathbf{R}$

$$\widehat{F}(x) = \int_{\mathbf{R}} F(\xi + i\eta)e^{-ix(\xi+i\eta)}\,dx.$$

Hence

$$|\widehat{F}(x)| \le C\sup(x\eta + h(\eta)), \quad \eta \in \mathbf{R}.$$

When $\eta \to +\infty$ this proves that $\widehat{F}(x) = 0$ if $x + b < 0$, and when $\eta \to -\infty$ it follows that $\widehat{F}(x) = 0$ if $x + a > 0$. Hence $\operatorname{supp}\widehat{F} \subset [-b, -a]$, so $\widehat{F}$ is continuous with compact support. By Fourier's inversion formula we conclude that $F = \hat{f}$ where $f(x) = \widehat{F}(-x)/2\pi$ is a continuous function with support in $[a, b]$.

To complete the proof we choose $\varphi \in C_0^\infty(\mathbf{R})$ with $\operatorname{supp}\varphi \subset [-1, 1]$ and $\int \varphi\,dx = 1$; then $\hat{\varphi}(0) = 1$ and

$$|\zeta^j\hat{\varphi}(\zeta)| \le e^{|\operatorname{Im}\zeta|}\int |\varphi^{(j)}(x)|\,dx.$$

If F satisfies (2.1.23) it follows that $F(\zeta)\hat{\varphi}(\varepsilon\zeta)$ satisfies (2.1.23) with a, b replaced by $a - \varepsilon, b + \varepsilon$ and μ replaced by any real number. Hence the Fourier transform has support in $[-b - \varepsilon, -a + \varepsilon]$. When $\varepsilon \to 0$ it follows that the Fourier transform of F has support in $[-b, -a]$, so $f = \check{F}/2\pi$ has support in $[a, b]$ and Fourier transform F. If N is a non-negative integer $> \mu + 1$, then $\xi \mapsto (i\xi + 1)^{-N}F(\xi)$ is integrable so the Fourier transform is a continuous function. Applying the differential operator $(-d/dx + 1)^N$ to the Fourier transform we conclude that $\hat{F}$ is of order N, which completes the proof.

The characterization of the Fourier transform of $\mathscr{E}'$ in Theorem 2.1.6 is a variant of the *Paley-Wiener theorem* due to Schwartz. The last statement in Theorem 2.1.6 has also several variants worth mentioning. A classical version is that if $f, g \in L^1(\mathbf{R})$ then $(f * g)(x) = \int f(x - y)g(y)\,dy$ is defined almost everywhere and belongs to $L^1(\mathbf{R})$; the Fourier transform is $\hat{f}\hat{g}$. The proof follows at once from the Lebesgue-Fubini theorem and is left as an exercise for the reader, but we shall prove:

THEOREM 2.1.7. *If $\varphi, \psi \in \mathscr{S}(\mathbf{R})$ then $(\varphi * \psi)(x) = \int \varphi(x - y)\psi(y)\,dy = (\psi * \varphi)(x)$ is in $\mathscr{S}(\mathbf{R})$ and the bilinear map $\mathscr{S} \times \mathscr{S} \ni (\varphi, \psi) \mapsto \varphi * \psi \in \mathscr{S}$ is continuous. If $f \in \mathscr{S}'$ and $\varphi \in \mathscr{S}$ then $f * \varphi \in \mathscr{S}'$ is defined by*

$$\langle f * \varphi, \psi\rangle = \langle f, \psi * \check{\varphi}\rangle, \quad \psi \in \mathscr{S}, \tag{2.1.24}$$

and the Fourier transform is $\hat{f}\hat{\varphi}$.

PROOF. Derivatives of $\varphi * \psi$ can be calculated by differentiating either factor, and if $k \geq 0$ then

$$(1+|x|)^k |(\varphi * \psi)(x)| \leq \sup_y |\varphi(x-y)|(1+|x-y|)^k \int (1+|y|)^k |\psi(y)|\, dy$$

since $1+|x| \leq (1+|x-y|)(1+|y|)$ by the triangle inequality. This proves that $\varphi*\psi \in \mathscr{S}$ and that $(\varphi, \psi) \mapsto \varphi * \psi$ is continuous in $\mathscr{S}$. Hence (2.1.24) defines a distribution $f * \varphi \in \mathscr{S}'$, and the map $\mathscr{S} \ni \varphi \mapsto f * \varphi \in \mathscr{S}'$ is continuous. Since the Fourier transformation is continuous in $\mathscr{S}'$ and the Fourier transform of $f * \varphi$ is $\hat{f}\hat{\varphi}$ if φ is in the dense subspace C_0^∞ of $\mathscr{S}$, it follows that this is true for every $\varphi \in \mathscr{S}$.

We shall now discuss some properties of distributions $f \in \mathscr{S}'$ with $\hat{f} \in \mathscr{E}'$; they are called *band limited* in signal theory. By Theorem 2.1.6 f is the restriction to $\mathbf{R}$ of an entire function; in particular $f \in C^\infty$. Let $\operatorname{supp} \hat{f} \subset [-\lambda, \lambda]$ and choose an even function $\varphi \in C_0^\infty(\mathbf{R})$ which is equal to 1 in $[-\lambda, \lambda]$. Then $\hat{f} = \hat{f}\varphi$, and since φ is the Fourier transform of $\hat{\varphi}/(2\pi) \in \mathscr{S}$, it follows from Theorem 2.1.7 that $f = f * \hat{\varphi}/(2\pi)$. When differentiating the right-hand side we can let all derivatives fall on $\hat{\varphi}$, so all derivatives are in $L^p \cap L^\infty \cap C^\infty$ if $f \in L^p$ for some $p \in [1, \infty]$. Hölder's inequality gives if $p < \infty$

$$2\pi|f(x)| \leq \int |f(x-y)||\hat{\varphi}(y)|\, dy \leq \Big(\int |f(x-y)|^p |\hat{\varphi}(y)|\, dy\Big)^{\frac{1}{p}} \Big(\int |\hat{\varphi}(y)|\, dy\Big)^{1-\frac{1}{p}}.$$

Hence $f(x) \to 0$ as $x \to \infty$, and

$$\sum_{-\infty}^{\infty} |2\pi f(kx)|^p \leq \Big(\int |f(y)|^p \sum_{-\infty}^{\infty} |\hat{\varphi}(y+kx)| dy\Big) \|\hat{\varphi}\|_{L^1}^{p-1},$$

which proves that

$$\Big(|x|/(1+|x|) \sum_{-\infty}^{\infty} |f(kx)|^p\Big)^{1/p} \leq C_\lambda \|f\|_{L^p}. \tag{2.1.25}$$

Thus $(f(kx))_{k\in\mathbf{Z}} \in l^p$ if $x \neq 0$.

A modification of the preceding argument yields a classical inequality due to S. Bernstein and some of its generalizations, with exact constants:

THEOREM 2.1.8. *If $f \in L^p(\mathbf{R})$ for some $p \in [1, \infty]$ and $\operatorname{supp} \hat{f} \subset [-\lambda, \lambda]$ then*

$$\|f' \sin\alpha/\lambda + f\cos\alpha\|_{L^p} \leq \|f\|_{L^p}, \quad \alpha \in \mathbf{R}. \tag{2.1.26}$$

PROOF. Changing scales we note that the support of the Fourier transform of $g(x) = f(x/\lambda)$ is contained in $[-1, 1]$ and that (2.1.26) is equivalent to

$$\|h\|_{L^p} \leq \|g\|_{L^p}, \quad \text{if } h = g' \sin\alpha + g\cos\alpha.$$

At first we assume that $\operatorname{supp}\hat{g} \subset (-1,1)$. We have $\hat{h}(\xi) = (i\xi\sin\alpha + \cos\alpha)\hat{g}(\xi)$. We would like to continue the function $[-1,1] \ni \xi \mapsto i\xi\sin\alpha + \cos\alpha$ to a function with period 2 on the real axis, but that would lead to a discontinuity at ± 2. However,

$$\Phi(\xi) = (i\xi\sin\alpha + \cos\alpha)e^{-i\alpha\xi}, \quad -1 \le \xi \le 1,$$

is a better choice since $\Phi(\pm 1) = 1$, and for the extension of Φ with period 2 we have the Fourier coefficients

$$c(k) = \frac{1}{2}\int_{-1}^{1}(i\xi\sin\alpha + \cos\alpha)e^{-i\alpha\xi - \pi i k\xi}\,d\xi = (-1)^k\frac{\sin^2\alpha}{(\alpha+\pi k)^2},$$

where $(\sin^2\alpha)/\alpha^2$ should be read as 1 when $\alpha = 0$. Thus the Fourier series of Φ is absolutely and uniformly convergent, $1 = \Phi(1) = \sum_{-\infty}^{\infty}\sin^2\alpha/(\alpha+k\pi)^2$. If $g \in L^1$ then $\hat{g}$ is continuous, and we have

$$\hat{h}(\xi) = \Phi(\xi)e^{i\alpha\xi}\hat{g}(\xi) = \sum_{-\infty}^{\infty}(-1)^k\frac{\sin^2\alpha}{(\alpha+k\pi)^2}e^{i(\alpha+\pi k)\xi}\hat{g}(\xi)$$

where the series converges uniformly, hence in $\mathscr{S}'$. This proves that

$$h(x) = \sum_{-\infty}^{\infty}(-1)^k\frac{\sin^2\alpha}{(\alpha+\pi k)^2}g(x+\alpha+\pi k). \tag{2.1.27}$$

If g is not in L^1 but just in L^∞ we can apply (2.1.27) to $g(x)\sin^2(\varepsilon x)/(\varepsilon x)^2$ when ε is so small that $\operatorname{supp}\hat{g} \subset (-1+2\varepsilon, 1-2\varepsilon)$. When $\varepsilon \to 0$ it follows by dominated convergence that (2.1.26) is valid for g. If we only assume that g is bounded and that $\operatorname{supp}\hat{g} \subset [-1,1]$ we can apply (2.1.27) to $x \mapsto g(tx)$ when $0 < t < 1$, and when $t \to 1$ we conclude that (2.1.27) is valid without restriction for such functions g. Hence Minkowski's inequality yields

$$\|h\|_{L^p} \le \|g\|_{L^p}\sum_{-\infty}^{\infty}\frac{\sin^2\alpha}{(\alpha+\pi k)^2} = \|g\|_{L^p}\Phi(1) = \|g\|_{L^p}.$$

When $p = \infty$ there is equality when $g(x) = \sin(x-\alpha)$ and $h(x) = \sin x$. For $p < \infty$ equality is never attained but the constant is best possible then too, an exercise for the reader, who might also wish to prove that if P is any polynomial with only real zeros, then (2.1.26) can be generalized to

$$\|P(d/dx)f\|_{L^p} \le |P(\lambda i)|\|f\|_{L^p}, \tag{2.1.26$'$}$$

for the same f as in Theorem 2.1.8.

A very similar argument will now be used to show how a band limited function can be recovered from an equidistant sampling of its values.

THEOREM 2.1.9. *If* $f \in L^p(\mathbf{R}) \cap C(\mathbf{R})$ *for some* $p \in [1, \infty)$ *and* $\operatorname{supp} \hat{f} \subset [-\lambda, \lambda]$, *then we have (with the convention* $(\sin 0)/0 = 1$*)*

$$f(x) = \sum_{-\infty}^{\infty} f(\pi k/\lambda) \frac{\sin(\lambda x - \pi k)}{\lambda x - \pi k}, \quad x \in \mathbf{R}. \tag{2.1.28}$$

Before the proof we observe that (2.1.28) is valid when $x = \pi k/\lambda$, $k \in \mathbf{Z}$. Also note that the function $f(x) = \sin \lambda x$ has the Fourier transform $\pi i(\delta_{-\lambda} - \delta_\lambda)$ and that $f(\pi k/\lambda) = 0$, $k \in \mathbf{Z}$, so the statement would not be true for $p = \infty$. However, when $f \in L^p$ and $p < \infty$ then the sum in (2.1.28) is absolutely convergent by (2.1.25).

PROOF OF THEOREM 2.1.9. Assume at first that $\operatorname{supp} \hat{f} \subset (-\lambda, \lambda)$. As usual we then define a distribution with period 2λ equal to $\hat{f}$ in $(-\lambda, \lambda)$ by

$$F = \sum_{-\infty}^{\infty} \hat{f}(\cdot - 2\lambda j).$$

The Fourier coefficients of F are

$$c(k) = (2\lambda)^{-1} \hat{f}(e^{-\pi i k \cdot/\lambda}) = (2\lambda)^{-1} \hat{\hat{f}}(\pi k/\lambda) = (\pi/\lambda) f(-\pi k/\lambda).$$

Hence

$$F = \frac{\pi}{\lambda} \sum_{-\infty}^{\infty} f(-\pi k/\lambda) e^{\pi i k \cdot/\lambda}.$$

If $\varphi \in C_0^\infty((-\lambda, \lambda))$ is equal to 1 in a neighborhood of $\operatorname{supp} \hat{f}$, it follows that

$$\hat{f} = \frac{\pi}{\lambda} \sum_{-\infty}^{\infty} f(\pi k/\lambda) \varphi e^{-\pi i k \cdot/\lambda}$$

with convergence in $\mathscr{E}'$. Inversion of the Fourier transformation gives

$$f(x) = \frac{1}{2\lambda} \sum_{-\infty}^{\infty} f(\pi k/\lambda) \hat{\varphi}(\pi k/\lambda - x). \tag{2.1.29}$$

We can take a sequence of functions $\varphi_j \in C_0^\infty((-\lambda, \lambda))$ converging to 1 in $(-\lambda, \lambda)$ such that

$$\int (|\varphi_j(\xi)| + \lambda |\varphi_j'(\xi)|) \, d\xi \leq 4\lambda, \quad \text{hence } |\hat{\varphi}_j(x)| \leq 4\lambda, |x \hat{\varphi}_j(x)| \leq 4.$$

Since $\sum |f(\pi k/\lambda)|(1 + |k|)^{-1} < \infty$ this gives a summable majorant for the series in (2.1.29) with φ replaced by φ_j, and since $\int_{-\lambda}^{\lambda} e^{-ix\xi} \, d\xi = 2(\sin(\lambda x))/x$ we conclude when $j \to \infty$ that (2.1.28) is valid. If we only know that $\operatorname{supp} \hat{f} \subset [-\lambda, \lambda]$ we can apply this formula with λ replaced by $\mu > \lambda$ and then let $\mu \to \lambda$; the detailed proof that this yields (2.1.28) is an exercise for the reader.

REMARK. If $f \in \mathscr{S}'$ and $\operatorname{supp} \hat{f}$ is compact, $f(\pi k/\lambda) = 0$ for $k \in \mathbf{Z}$, then $f(x)/\sin(\lambda x)$ is an entire function so it follows from the *theorem of supports* proved below that the convex hull of $\operatorname{supp} \hat{f}$ must contain an interval of length 2λ. If $\operatorname{supp} \hat{f} \subset (-\lambda, \lambda)$ it is therefore clear that f is determined by the values at $\pi\mathbf{Z}/\lambda$. An explicit form of this conclusion is given by the interpolation formula (2.1.29) with $\varphi \in C_0^\infty((-\lambda, \lambda))$ equal to 1 near $\operatorname{supp} \hat{f}$. Since $f(x) = O(|x|^N)$ for some N (by Theorem 2.1.6) and $\hat{\varphi} \in \mathscr{S}$ the series in (2.1.29) is rapidly convergent. The somewhat delicate point in Theorem 2.1.9 is that $\pm\lambda$ may be in $\operatorname{supp} \hat{f}$, and then the growth of f must be restricted.

In signal theory Theorem 2.1.9 is known as Shannon's theorem (cf. Daubechies [1, p. 18]), but its mathematical roots are much older.

To prove the theorem of supports announced above we have to combine Theorem 2.1.6 with a classical result in analytic function theory:

LEMMA 2.1.10. *If F is an entire function in $\mathbf{C}$ satisfying* (2.1.23) *then*

$$\log|F(\zeta)| = b_1 \operatorname{Im}\zeta + \frac{\operatorname{Im}\zeta}{\pi}\int_{\mathbf{R}} \frac{\log|F(t)|}{|t-\zeta|^2}\,dt + \sum \log\Big|\frac{\zeta - z_j}{\zeta - \bar{z}_j}\Big|, \quad \operatorname{Im}\zeta > 0, \tag{2.1.30}$$

where $a \le b_1 \le b$ and z_j are the zeros of F in the open upper half plane, repeated according to their multiplicities.

PROOF. First assume that F satisfies (2.1.23) with $C = 1$, $\mu = 0$ and $b = 0$. Let $\varphi \in C_0$, $\log|F(t)| \le \varphi(t) \le 0$, and let M be a finite subset of the index set M_+ for the zeros of F in the open upper half plane. Then

$$G(\zeta) = F(\zeta)\exp\Big(-\frac{1}{\pi i}\int_{\mathbf{R}}\frac{\varphi(t)}{t-\zeta}\,dt\Big)\prod_{j\in M}\frac{\zeta-\bar{z}_j}{\zeta - z_j}$$

is analytic in the upper half plane, and $|G(\zeta)| \le 1$ there. In fact, by hypothesis $|F(\zeta)| \le 1$ in the upper half plane. The absolute value of the product is equal to 1 on the real axis, and the real part of the exponent is the Poisson integral

$$-\frac{\operatorname{Im}\zeta}{\pi}\int_{\mathbf{R}}\frac{\varphi(t)}{|t-\zeta|^2}\,dt = -\frac{1}{\pi}\int_{\mathbf{R}}\frac{\varphi(\operatorname{Re}\zeta + t\operatorname{Im}\zeta)}{1+t^2}\,dt$$

which is $\le -\varphi(\operatorname{Re}\zeta) + o(1)$ as $\operatorname{Im}\zeta \to 0$. Hence it follows from the maximum principle that $|G(\zeta)| \le 1$ in the upper half plane. Taking sequences $\varphi_j \downarrow \log|F|$ and $M_j \uparrow M_+$ we conclude since $|G_j| \le 1$ for the corresponding functions G_j that

$$v(\zeta) = \log|F(\zeta)| - \frac{\operatorname{Im}\zeta}{\pi}\int\frac{\log|F(t)|}{|t-\zeta|^2}\,dt - \sum\log\Big|\frac{\zeta-z_j}{\zeta-\bar{z}_j}\Big| \le 0, \quad \operatorname{Im}\zeta > 0.$$

v is the limit of harmonic functions so v is harmonic. We can choose the sequence φ_j so that it is equal to 1 for large j on any compact subset K of $\mathbf{R}$ containing no zeros of F, and since G_j is continuous up to K with boundary values 0 it follows that v is continuous with boundary values 0 at K. This is also true at a zero $\lambda \in \mathbf{R}$ of F, for the Poisson integral of $\log|t-\lambda|$ is $\log|\zeta-\lambda|$. Thus v is harmonic and ≤ 0 in the upper half plane and continuous in the closure with boundary values 0. We shall prove in a moment that this implies $v(\zeta) = b_1 \operatorname{Im}\zeta$ for some real number b_1. We have $b_1 \le 0 = b$, and it is clear that $b_1 \ge a$ for otherwise Theorem 2.1.6 would prove that F is the Fourier-Laplace transform of a distribution with support $\{c\}$ for every c with $b_1 < c < a$ which is absurd.

By the Schwarz reflection principle we can extend v to a harmonic function in $\mathbf{C}$ by defining $v(\bar\zeta) = -v(\zeta)$. We can express v in terms of the values on a large circle by the Poisson integral

$$v(\zeta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1-|\zeta/R|^2}{|\zeta/R - e^{i\theta}|^2} v(Re^{i\theta})\,d\theta, \quad R > |\zeta|.$$

(This follows from the mean value property for $z \mapsto v(R(z-\zeta)/(R^2-\bar\zeta z))$ which is harmonic when $|z| \le R$.) In our case $v(Re^{i\theta})$ is an odd function of θ, and since

$$|\zeta/R - e^{-i\theta}|^2 - |\zeta/R - e^{i\theta}|^2 = 4\sin\theta \operatorname{Im}\zeta/R,$$

we obtain

$$v(\zeta) = \frac{2\operatorname{Im}\zeta}{\pi R}\int_0^{\pi} v(Re^{i\theta})\sin\theta\,d\theta(1+O(1/R))$$

using the positivity of the integrand. When $R \to \infty$ it follows that $v(\zeta) = b_1 \operatorname{Im}\zeta$ for some b_1.

It remains to remove the hypotheses $C = 1$, $\mu = 0$ and $b = 0$ made above. Multiplication of F by $C^{-1}e^{ib\zeta}$ removes the hypotheses $C = 1$ and $b = 0$. If N is a positive integer $> \mu$ then

$$F_\varepsilon(\zeta) = F(\zeta)(\sin(\varepsilon\zeta)/\varepsilon\zeta)^N$$

satisfies (2.1.23) with μ replaced by 0 and $h(\eta)$ replaced by $h(\eta) + N\varepsilon|\eta|$, if $\varepsilon > 0$. Subtracting the same identity with $F \equiv 1$ we obtain (2.1.30) for some $b_1 \in [a - 2N\varepsilon, b + 2N\varepsilon]$. Since b_1 is uniquely determined by (2.1.30) we have $b_1 \in [a, b]$.

There is of course a formula corresponding to (2.1.30) in the lower half plane. Note that it follows from (2.1.23) and (2.1.30) that $|F(\zeta)| \le C'|\zeta + i|^\mu e^{b_1 \operatorname{Im}\zeta}$, $\operatorname{Im}\zeta > 0$, for the Poisson integral of $t \mapsto \log(1+t^2)$ is $2\log|\zeta + i|$. If $F = \hat f$ as in Theorem 2.1.6 we conclude that $b_1 = \sup\{x; x \in \operatorname{supp} f\}$. Hence we obtain:

THEOREM 2.1.11. *If $f, g \in \mathscr{E}'(\mathbf{R})$ and $[a, b]$ (resp. $[c, d]$) are the smallest intervals containing the supports, then $[a + c, b + d]$ is the smallest interval containing the support of $f * g$.*

This is the famous *theorem of supports* due to Titchmarsh.

2.2. The higher-dimensional case. We set out in Section 2.1 to study functions in $\mathbf{R}$, its closed subgroups and the quotients by the closed subgroups. In $\mathbf{R}^n$ there are many more closed subgroups:

PROPOSITION 2.2.1. *If $G \subset \mathbf{R}^n$ is a closed subgroup then there is a linear subspace V_0 of $\mathbf{R}^n$ and elements $g_1, \dots, g_r \in G$, which are linearly independent modulo V_0, such that*

$$G = \{\sum_1^r k_j g_j + g_0; k_j \in \mathbf{Z}, g_0 \in V_0\}. \tag{2.2.1}$$

Conversely, (2.2.1) *defines a closed subgroup of G.*

PROOF. The last statement is obvious if we introduce coordinates such that $g_1, \dots, g_r$ are the first r basis vectors and $V_0 \subset \{x \in \mathbf{R}^n; x_1 = \dots = x_r = 0\}$. To prove that G must be of the form (2.2.1) we first observe that if two vector spaces are subsets of G then their sum is also a subset of G. Hence there is a vector space $V_0 \subset G$ containing all other vector spaces $\subset G$. Passing to the quotient $\mathbf{R}^n/V_0$ we may assume that G does not contain any vector space. Then G is *discrete*, that is, there is some $\varepsilon > 0$ such that $x \in G$, $|x| < \varepsilon$ implies $x = 0$. In fact, otherwise there would exist a sequence $x_j \in G$ with $0 \neq x_j \to 0$. Passing to a subsequence we may assume that $x_j/|x_j|$ converges to a limit $y \neq 0$ as $j \to \infty$. If $t \in \mathbf{R} \setminus 0$ and $[t/|x_j|]$ is the largest integer $\leq t/|x_j|$ it follows that $G \ni [t/|x_j|]x_j \to ty$ as $j \to \infty$, and since G is closed it follows that $\mathbf{R}y \subset G$, which is a contradiction. Thus G is discrete. Let g_1 be any element $\neq 0$ in G which is not a multiple of another element in G, for example an element of minimal norm. Changing the coordinates we may assume that $g_1 = (1, 0, \dots, 0)$. Let $G' = \{x' \in \mathbf{R}^{n-1}; (t, x') \in G\}$ for some $t \in \mathbf{R}$. It is obvious that G' is a subgroup of $\mathbf{R}^{n-1}$. To prove that G' is closed we consider a sequence $x'_\nu \in G$ with $x'_\nu \to x'$. By the definition of G' we can choose $t_\nu \in \mathbf{R}$ with $(t_\nu, x'_\nu) \in G$, and since $(1, 0, \dots, 0) \in G$ we can choose $t_\nu \in [-\frac{1}{2}, \frac{1}{2})$. If t is a limit point of the sequence t_ν it follows that $(t, x') \in G$, so $x' \in G'$. If $0 \neq x'_\nu$ but $x' = 0$ we obtain $t = 0$, which is a contradiction with the discreteness of G, so G' is discrete. If the proposition has already been proved for lower dimensions it follows that there are elements $g_j = (t_j, g'_j) \in G$, $j = 2, \dots, r$ such that $g'_2, \dots, g'_r$ are linearly independent and $G' = \{\sum_2^r k_j g'_j; k_j \in \mathbf{Z}\}$. Thus $G = \{\sum_1^r k_j g_j; k_j \in \mathbf{Z}\}$, which proves (2.2.1).

To avoid notational complications we shall only discuss the case where $G = \mathbf{R}^n$ (Fourier integrals) and the case where $G = T\mathbf{Z}^n, \mathbf{R}^n/G = \mathbf{R}^n/T\mathbf{Z}^n$ (Fourier series). This is no serious loss of generality, but there are some cases such as the Schrödinger equation in a crystal lattice where one has to respect a Euclidean geometry and work with a lattice in a general position.

We shall use the notation $\alpha = (\alpha_1, \dots, \alpha_n)$ for a *multi-index*, a vector in $\mathbf{Z}^n$ with non-negative coordinates, and we shall write

$$|\alpha| = \sum_1^n \alpha_j, \quad \alpha! = \prod_1^n \alpha_j!, \quad \partial^\alpha = \prod_1^n (\partial/\partial x_j)^{\alpha_j}, \quad D^\alpha = \prod_1^n (-i\partial/\partial x_j)^{\alpha_j}.$$

PROPOSITION 2.2.2. *If $f \in C^\mu(\mathbf{R}^n)$ is $T\mathbf{Z}^n$ periodic, that is, $f(x + Tk) = f(x)$ if $x \in \mathbf{R}^n$ and $k \in \mathbf{Z}^n$, then the Fourier coefficients*

$$c(k) = \frac{1}{T^n} \int_{\mathbf{R}^n/T\mathbf{Z}^n} f(x) e^{-2\pi i \langle x, k\rangle/T}\, dx, \quad k \in \mathbf{Z}^n, \tag{2.2.2}$$

have the bound

$$|(2\pi k/T)^\alpha c(k)| \le \frac{1}{T^n}\int_{\mathbf{R}^n/T\mathbf{Z}^n} |D^\alpha f(x)|\,dx, \quad |\alpha| \le \mu. \tag{2.2.3}$$

If $\mu > n$ *then*

$$f(x) = \sum_{k\in\mathbf{Z}^n} c(k) e^{2\pi i\langle x,k\rangle/T} \tag{2.2.4}$$

with absolute and uniform convergence, and Parseval's formula

$$\frac{1}{T^n}\int_{\mathbf{R}^n/T\mathbf{Z}^n} |f(x)|^2\,dx = \sum_{k\in\mathbf{Z}^n} |c(k)|^2 \tag{2.2.5}$$

is valid. Conversely, if $c(k) \in \mathbf{C}$ *is given,* $k \in \mathbf{Z}^n$*, and* $|k|^\mu c(k)$ *is bounded, then* (2.2.4) *defines a* $T\mathbf{Z}^n$ *periodic function* $f \in C^{\mu-n-1}(\mathbf{R}^n)$ *if* $\mu \ge n+1$*, and* (2.2.2) *holds.*

The notation $\int_{\mathbf{R}^n/T\mathbf{Z}^n} \psi(x)\,dx$ where ψ is a $T\mathbf{Z}^n$ periodic function stands for $\int_F \psi(x)\,dx$ where F is a measurable fundamental domain, that is, a set such that $\mathbf{R}^n$ is the *disjoint* union of the translates $F + Tk$ with $k \in \mathbf{Z}^n$, that is, $\sum_{k\in\mathbf{Z}^n} \chi(x+Tk) = 1$ (almost everywhere) if χ is the characteristic function of F. More generally, if $\psi \in L^1_{\text{loc}}(\mathbf{R}^n)$ is $T\mathbf{Z}^n$ periodic then we define

$$\int_{\mathbf{R}^n/T\mathbf{Z}^n} \psi(x)\,dx = \int_{\mathbf{R}^n} \psi(x)\chi(x)\,dx, \tag{2.2.6}$$

where χ is any bounded measurable function of compact support such that

$$\sum_{k\in\mathbf{Z}^n} \chi(x+Tk) = 1 \quad \text{almost everywhere.} \tag{2.2.7}$$

The definition (2.2.6) is then independent of the choice of χ. In fact, if χ_1 is another bounded measurable function of compact support satisfying (2.2.7) then

$$\begin{aligned}\int_{\mathbf{R}^n} \psi(x)\chi(x)\,dx &= \sum_{k\in\mathbf{Z}^n}\int_{\mathbf{R}^n} \psi(x)\chi(x)\chi_1(x+Tk)\,dx \\ &= \sum_{k\in\mathbf{Z}^n}\int_{\mathbf{R}^n} \psi(x)\chi(x-Tk)\chi_1(x)\,dx = \int_{\mathbf{R}^n} \psi(x)\chi_1(x)\,dx,\end{aligned}$$

by the periodicity of ψ.

Proof of Proposition 2.2.2. Interpreting $\int_{\mathbf{R}^n/T\mathbf{Z}^n}$ as the integral over $I = \{x \in \mathbf{R}^n; 0 \le x_j \le T, j = 1,\dots,n\}$, we obtain by partial integration

$$(2\pi k/T)^\alpha c(k) = \frac{1}{T^n}\int_I (D^\alpha f(x)) e^{-2\pi i\langle x,k\rangle/T}\,dx, \quad k \in \mathbf{Z}^n, \quad |\alpha| \le \mu,$$

for the contributions at the boundary cancel by the periodicity. This proves (2.2.3), which implies that $|c(k)||k|^\mu$ is bounded. If $\mu > n$ it follows that the Fourier series (2.2.4) converges absolutely and uniformly to a continuous function g. Let $\varphi_1, \dots, \varphi_n \in C^2(\mathbf{R})$ be periodic with period T. Then it follows from Proposition 2.1.1 that

$$\prod_1^n \varphi(x_j) = \sum_{k\in\mathbf{Z}^n} e^{2\pi i\langle x,k\rangle/T} \prod_1^n \frac{1}{T}\int_0^T \varphi(y_j)e^{-2\pi i y_j k_j/T}\,dy_j,$$

hence with the notation $\Phi(x) = \prod_1^n \varphi(x_j)$

$$\frac{1}{T^n}\int_I f(x)\overline{\Phi(x)}\,dx = \sum_{k\in\mathbf{Z}^n} c(k)\frac{1}{T^n}\int_I \overline{\Phi(y)}e^{2\pi i\langle y,k\rangle/T}\,dy = \frac{1}{T^n}\int_I g(y)\overline{\Phi(y)}\,dy.$$

Hence $\int_{\mathbf{R}^n/T\mathbf{Z}^n}(f(x) - g(x))\overline{\Phi(x)}\,dx = 0$ for every choice of $T\mathbf{Z}$ periodic functions φ_j, $j = 1, \dots, n$. Choose $\psi \in C_0^\infty(\mathbf{R})$ with $\int_{\mathbf{R}} \psi(t)\,dt = 1$ and set for $y \in \mathbf{R}^n$ and $\varepsilon > 0$

$$\varphi_j(x_j) = \frac{1}{\varepsilon}\sum_{k\in\mathbf{Z}} \psi((x_j - y_j - Tk)/\varepsilon).$$

Then $\varphi_j \in C^\infty(\mathbf{R})$ is $T\mathbf{Z}$ periodic. When $\varepsilon \to 0$ it follows that $f(y) = g(y)$, which proves (2.2.4). Parseval's formula follows if we multiply (2.2.4) by $\overline{f(x)}$ and integrate, interchanging the order of integration and summation. The proof of the converse statement is obvious and is left for the reader (who should be aware of the fact that $\sum_{0\neq k\in\mathbf{Z}^n}|k|^{-\gamma}$ converges if and only if $\gamma > n$).

Remark. Another sufficient condition for uniform and absolute convergence of the Fourier series is that $D^\alpha g$ is continuous when $\alpha_j \le 1$, for $j = 1, \dots, n$. Weaker sufficient conditions will be given later on.

The definition of $\int_{\mathbf{R}^n/T\mathbf{Z}^n}$ by (2.2.6), (2.2.7) can also be used to define $\langle f, \Phi\rangle_{\mathbf{R}^n/T\mathbf{Z}^n}$ if $f \in \mathscr{D}'(\mathbf{R}^n)$ and $\Phi \in C^\infty(\mathbf{R}^n)$ are $T\mathbf{Z}^n$ periodic. In fact, we can choose $\chi \in C_0^\infty(\mathbf{R}^n)$ so that (2.2.7) holds (see the corresponding discussion of (2.1.17)) and set

$$\langle f, \Phi\rangle_{\mathbf{R}^n/T\mathbf{Z}^n} = \langle f, \chi\Phi\rangle. \tag{2.2.6$'$}$$

If χ_1 is another function in $C_0^\infty(\mathbf{R}^n)$ satisfying (2.2.7) then

$$\langle f, \chi\Phi\rangle = \sum_{k\in\mathbf{Z}^n}\langle f, \chi_1(\cdot + kT)\chi\Phi\rangle = \sum_{k\in\mathbf{Z}^n}\langle f, \chi_1\chi(\cdot - kT)\Phi\rangle = \langle f, \chi_1\Phi\rangle$$

by the periodicity of f and Φ. This identifies $T\mathbf{Z}^n$ periodic distributions with continuous linear forms on $T\mathbf{Z}^n$ periodic C^∞ functions.

The Fourier coefficients of a $T\mathbf{Z}^n$ periodic distribution f can now be defined by

$$c(k) = T^{-n}\langle f, e^{-2\pi i\langle\cdot,k\rangle/T}\rangle_{\mathbf{R}^n/T\mathbf{Z}^n}, \quad k \in \mathbf{Z}^n, \tag{2.2.2$'$}$$

and we have

$$|c(k)| \le C(1+|k|)^{\mu}, \quad k \in \mathbf{Z}^n, \tag{2.2.3$'$}$$

if f is of order μ. As in the one-dimensional case it follows from the smooth case in Proposition 2.2.2 that

$$F = \sum_{k\in\mathbf{Z}^n} c(k)e^{2\pi i\langle\cdot,k\rangle/T}, \tag{2.2.4$'$}$$

with convergence in $\mathscr{D}'(\mathbf{R}^n)$. We leave for the reader to supply the details of the proof of this and other statements in the following theorem which is completely analogous to Theorem 2.1.3:

THEOREM 2.2.3. *If $f \in \mathscr{D}'(\mathbf{R}^n)$ is $T\mathbf{Z}^n$ periodic then the Fourier coefficients defined by* (2.2.2)$'$ *have the polynomial bound* (2.2.3)$'$*, and the Fourier series* (2.2.4)$'$ *converges to f in $\mathscr{D}'(\mathbf{R}^n)$. Conversely, if $c(k) \in \mathbf{C}$ is given, $k \in \mathbf{Z}^n$, and* (2.2.3)$'$ *is fulfilled then the series* (2.2.4)$'$ *converges in $\mathscr{D}'(\mathbf{R}^n)$ to a $T\mathbf{Z}^n$ periodic distribution with Fourier coefficients $c(k)$.*

As in the one-dimensional case we have obtained an isomorphism between $T\mathbf{Z}^n$ periodic distributions and functions of polynomial growth on $\mathbf{Z}^n$. We pass now to a discussion of the Fourier transformation in n dimensions.

DEFINITION 2.2.4. By $\mathscr{S}$ or $\mathscr{S}(\mathbf{R}^n)$ we shall denote the space consisting of all $\varphi \in C^\infty(\mathbf{R}^n)$ such that $x^\beta\partial^\alpha\varphi(x)$ is bounded for arbitrary multi-indices α and β.

As when $n = 1$ it is clear that the *Schwartz space* $\mathscr{S}$ is a Fréchet space with the seminorms

$$\mathscr{S} \ni \varphi \mapsto \sup|x^\beta\partial^\alpha\varphi(x)|.$$

Since $C_0^\infty(\mathbf{R}^n)$ is a dense subspace of $\mathscr{S}(\mathbf{R}^n)$ it follows that the dual space $\mathscr{S}'(\mathbf{R}^n)$ of *temperate distributions* can be identified with a subspace of $\mathscr{D}'(\mathbf{R}^n)$, in fact $\mathscr{S}'(\mathbf{R}^n) \subset \mathscr{D}'_F(\mathbf{R}^n)$, the space of distributions of finite order.

For $f \in \mathscr{S}(\mathbf{R}^n)$ and $1 \le k \le n$ we define the *partial Fourier transform* $\hat{f}_{(k)}$ by

$$\hat{f}_{(k)}(\xi', x'') = \int_{\mathbf{R}^k} f(x', x'')e^{-i\langle x',\xi'\rangle}\, dx', \tag{2.2.8}$$

where $x' = (x_1, \dots, x_k)$, $x'' = (x_{k+1}, \dots, x_n)$ and $\xi' = (\xi_1, \dots, \xi_k)$ are real variables. The map $f \mapsto \hat{f}_{(k)}$ is an isomorphism of $\mathscr{S}(\mathbf{R}^n)$ with inverse

$$f(x', x'') = (2\pi)^{-k}\int_{\mathbf{R}^k} \hat{f}_{(k)}(\xi', x'')e^{i\langle x',\xi'\rangle}\, d\xi', \tag{2.2.9}$$

and Parseval's formula is valid,

$$\iint_{\mathbf{R}^k\times\mathbf{R}^{n-k}} |f(x', x'')|^2\, dx'\, dx'' = (2\pi)^{-k}\int_{\mathbf{R}^k\times\mathbf{R}^{n-k}} |\hat{f}_{(k)}(\xi', x'')|^2\, d\xi'\, dx''. \tag{2.2.10}$$

It is sufficient to verify this when $k=1$ and then it is an immediate consequence of the case $k=n=1$ discussed in Section 2.1. When $k=n$ we shall use the notation $\hat f$ instead of $\hat f_{(n)}$ for the *Fourier transform* defined by

$$\hat f(\xi)=\int_{\mathbf{R}^n} f(x)e^{-i\langle x,\xi\rangle}\,dx,\quad f\in\mathscr{S}(\mathbf{R}^n), \tag{2.2.11}$$

with inverse given by Fourier's inversion formula

$$f(x)=(2\pi)^{-n}\int_{\mathbf{R}^n}\hat f(\xi)e^{i\langle x,\xi\rangle}\,d\xi, \tag{2.2.12}$$

that is, $\hat{\hat f}(x)=(2\pi)^n f(-x)$. If $f\in L^1(\mathbf{R}^n)$ the Fourier transform can still be defined by (2.2.11), and we obtain

$$\langle \hat f,\varphi\rangle=\langle f,\hat\varphi\rangle,\quad \varphi\in\mathscr{S}, \tag{2.2.13}$$

if we multiply (2.2.11) by $\varphi(\xi)$ and integrate, inverting the order of integration in the right-hand side. We can therefore again extend the definition of the Fourier transformation to arbitrary $f\in\mathscr{S}'(\mathbf{R}^n)$ by (2.2.13), for the right-hand side is a continuous linear function of $\varphi\in\mathscr{S}(\mathbf{R}^n)$.

THEOREM 2.2.5. *The Fourier transformation defined by* (2.2.13) *is an isomorphism of $\mathscr{S}'(\mathbf{R}^n)$, and Fourier's inversion formula is valid, that is, $\hat{\hat f}=(2\pi)^n\check f$ where $\check f$ is the reflection in the origin defined by*

$$\langle \check f,\varphi\rangle=\langle f,\check\varphi\rangle,\quad \varphi\in\mathscr{S}(\mathbf{R}^n),\quad \check\varphi(x)=\varphi(-x).$$

We have $f\in L^2(\mathbf{R}^n)$ if and only if $\hat f\in L^2(\mathbf{R}^n)$, and Parseval's formula

$$\int_{\mathbf{R}^n}|f(x)|^2\,dx=(2\pi)^{-n}\int_{\mathbf{R}^n}|\hat f(\xi)|^2\,d\xi, \tag{2.2.14}$$

is then valid. When $f\in\mathscr{S}'$ the Fourier transform of $x^\beta D^\alpha f$ is equal to $(-D)^\beta\xi^\alpha\hat f$ where x (resp. ξ) denotes the variable where f (resp. $\hat f$) lives. The Fourier transform of $f(\cdot+h)$ is $e^{i\langle h,\cdot\rangle}\hat f$, and if $a:\mathbf{R}^n\to\mathbf{R}^n$ is a linear bijection then the Fourier transform of $f\circ a$ is $|\det a|^{-1}\hat f\circ{}^ta^{-1}$.

The proof differs only marginally from that of Theorem 2.1.5 so it is left as an exercise. We give a few examples, leaving the details of verification as an exercise.

EXAMPLES. **1.** The Fourier transform of δ_a, $a\in\mathbf{R}^n$, is $\xi\mapsto e^{-i\langle a,\xi\rangle}$. The Fourier transform of $\xi\mapsto e^{i\langle a,\xi\rangle}$ is $(2\pi)^n\delta_a$. In particular, the Fourier transform of a constant C is $(2\pi)^nC\delta_0$.

2. The Fourier transform of $P_T=\sum_{k\in\mathbf{Z}^n}\delta_{Tk}$ where $T>0$ is equal to $(2\pi/T)^nP_{2\pi/T}$. This follows from the one dimensional case. Explicitly this means that

$$\sum_{k\in\mathbf{Z}^n}\hat\varphi(Tk)=(2\pi/T)^n\sum_{k\in\mathbf{Z}^n}\varphi(2\pi k/T),\quad \varphi\in\mathscr{S}(\mathbf{R}^n), \tag{2.2.15}$$

which is known as *Poisson's summation formula.*

3. If $f \in \mathscr{D}'(\mathbf{R}^n)$ is $T\mathbf{Z}^n$ periodic, then $f \in \mathscr{S}'(\mathbf{R}^n)$, and if the Fourier coefficients are defined by (2.2.2)$'$ then

$$\hat{f} = (2\pi)^n \sum_{k\in\mathbf{Z}^n} c(k)\delta_{2\pi k/T}.$$

This follows from Example 1 above since f is the limit in $\mathscr{S}'$ of the partial sums of the Fourier series. Note that Example 2 is a special case.

4. For the Gaussian $g(x) = \exp(-\frac{1}{2}\langle x, x\rangle)$ the Fourier transform is $\hat{g}(\xi) = (2\pi)^{\frac{1}{2}n} g(\xi)$. This follows at once from the one-dimensional case. If a is a real linear bijection in $\mathbf{R}^n$, it follows that the Fourier transform of $x \mapsto \exp(-\frac{1}{2}\langle Ax, x\rangle)$ is

$$\xi \mapsto (2\pi)^{\frac{1}{2}n} (\det A)^{-\frac{1}{2}} \exp(-\tfrac{1}{2}\langle A^{-1}x, x\rangle),$$

where $A = {}^t aa$ is an arbitrary positive definite real symmetric matrix. More generally, this is true for all A in the set H of complex symmetric $n \times n$ matrices such that $\operatorname{Re} A$ is positive definite. This follows by analytic continuation of the formula

$$\int e^{-\frac{1}{2}\langle Ax,x\rangle}\hat{\varphi}(x)\,dx = (2\pi)^{\frac{1}{2}n}(\det A)^{-\frac{1}{2}} \int e^{-\frac{1}{2}\langle A^{-1}\xi,\xi\rangle}\varphi(\xi)\,d\xi, \quad \varphi \in \mathscr{S}. \tag{2.2.16}$$

Note that if $\operatorname{Re} A$ is positive definite then $\operatorname{Re}\langle Az, \bar{z}\rangle > 0$ when $z \in \mathbf{C}^n \setminus \{0\}$, so A is injective in $\mathbf{C}^n$, hence $\det A \neq 0$, which makes the square root uniquely defined in the convex set H when it is chosen positive for $A = \mathrm{Id}$. Replacing z by $A^{-1}z$, we also see that $\operatorname{Re} A^{-1}$ is positive definite when $A \in H$, so both sides of (2.2.16) are analytic in H and the formula follows since it is valid when A is real, which suffices to calculate the derivatives at the identity for example. The limiting case where $\operatorname{Re} A$ is positive semidefinite and $\det A \neq 0$ can be handled as a limit of $A + \varepsilon\,\mathrm{Id}$ when $\varepsilon \to +0$. We refer to Hörmander [1, p. 85] for the limit of $(\det A)^{-\frac{1}{2}}$ then.

Recall that if $K \subset \mathbf{R}^n$ is a compact set then the *supporting function* H_K of K is defined by

$$H_K(\xi) = \sup_{x\in K}\langle x, \xi\rangle, \quad \xi \in \mathbf{R}^n, \tag{2.2.17}$$

and if $\operatorname{ch} K$ is the convex hull of K then

$$x \in \operatorname{ch} K \iff \langle x, \xi\rangle \le H_K(\xi), \forall \xi \in \mathbf{R}^n. \tag{2.2.18}$$

(See Hörmander [1, pp. 105–106].) The supporting function H_K is convex and positively homogeneous of degree 1; conversely every such function is the supporting function of exactly one convex compact set K.

THEOREM 2.2.6. *If $f \in \mathscr{E}'(\mathbf{R}^n)$ then the Fourier transform is the C^∞ function*

$$\hat{f}(\xi) = \langle f_x, e^{-i\langle x,\xi\rangle}\rangle, \quad \xi \in \mathbf{R}^n, \tag{2.2.19}$$

where f_x means that f acts as a distribution in the x variable. The Fourier transform $\hat f$ can be extended to an entire analytic function in $\mathbf{C}^n$, the Fourier-Laplace transform of f, by letting $\xi \in \mathbf{C}^n$ in (2.2.19)*. If* $\operatorname{supp} f \subset K$ *where $K \subset \mathbf{R}^n$ is a compact set, and f is of order μ, then $F(\xi) = \hat f(\xi)$ has a bound*

$$|F(\zeta)| \le C(1+|\zeta|)^\mu \exp H_K(\operatorname{Im}\zeta), \quad \zeta \in \mathbf{C}^n. \tag{2.2.20}$$

Conversely, if F is an entire analytic function such that (2.2.20) *is valid, then $F = \hat f$ where $f \in \mathscr{E}'$ has support in* $\operatorname{ch} K$ *and order* $\le \max(0, \mu+n+1)$*. If $g \in \mathscr{E}'(\mathbf{R}^n)$ then the Fourier transform of the convolution $f * g$ is $\hat f \hat g$.*

Proof. That the C^∞ function (2.2.19) defines $\hat f$ and is extended to an entire function by letting ξ be complex follows just as in the case $n = 1$. At the same time one finds that the Fourier-Laplace transform of $f * g$ is $\hat f \hat g$.

We can choose a cutoff function $\chi_\delta \in C_0^\infty(K_\delta)$, $K_\delta = \{x+y; x \in K, |y| \le \delta\}$, such that $\chi_\delta = 1$ in a neighborhood of K and $|\partial^\alpha \chi_\delta| \le C\delta^{-|\alpha|}$ when $|\alpha| \le \mu$. If $\varphi \in C^\infty(\mathbf{R}^n)$ and $0 < \delta < 1$ it follows that

$$|\langle f, \varphi\rangle| = |\langle f, \chi_\delta \varphi\rangle| \le C \sum_{|\alpha|\le\mu} \sup |D^\alpha \chi_\delta \varphi| \le C' \sum_{|\alpha|\le\mu} \sup_{K_\delta} |D^\alpha \varphi| \delta^{|\alpha|-\mu}.$$

Taking $\delta = 1/(1+|\zeta|)$ and $\varphi(x) = \exp(-i\langle x, \zeta\rangle)$ we conclude that $F(\zeta) = \hat f(\zeta)$ has the bound (2.2.20), for $\langle x, \operatorname{Im}\zeta\rangle \le H_K(\operatorname{Im}\zeta) + 1$ when $x \in K_\delta$.

Assume now given an entire analytic function F satisfying (2.2.20) with $\mu < -n$. Then

$$f(x,\eta) = \int_{\mathbf{R}^n} F(\xi + i\eta) e^{i\langle x, \xi+i\eta\rangle}\, d\xi, \quad x \in \mathbf{R}^n,\ \eta \in \mathbf{R}^n,$$

is a continuously differentiable function, and

$$\partial f(x,\eta)/\partial\eta_j = \int_{\mathbf{R}^n} i\partial(F(\xi+i\eta)e^{i\langle x,\xi+i\eta\rangle})/\partial\xi_j\, d\xi = 0,$$

so $f(x,\eta)$ is independent of η. When $x \notin \operatorname{ch} K$ we can choose η so that $\langle x, \eta\rangle > H_K(\eta)$ and obtain when $t > 0$

$$|f(x,0)| = |f(x,t\eta)| \le C \int_{\mathbf{R}^n} (1+|\xi|)^\mu \exp(t(H_K(\eta) - \langle x,\eta\rangle))\, d\xi.$$

When $t \to +\infty$ it follows that $f(x,0) = 0$. Thus $f = f(\cdot, 0)$ has support in K and Fourier transform F, which proves that F is the Fourier-Laplace transform of a function in $C_0(\operatorname{ch} K)$ when (2.2.20) is valid with $\mu < -n$. For F satisfying (2.2.20) with an arbitrary μ we conclude exactly as in the proof of Theorem 2.1.6 that F is the Fourier-Laplace transform of some $f \in \mathscr{E}'(K)$ of order $\max(0, n+1+\mu)$, which completes the proof.

The characterization (2.2.20) of the Fourier-Laplace transform of $\mathscr{E}'(K)$ when K is convex and compact is called the *Paley-Wiener-Schwartz* theorem. The *theorem of supports* is also easily extended from the one-dimensional to the n-dimensional case.

THEOREM 2.2.7. *If* $f, g \in \mathscr{E}'(\mathbf{R}^n)$ *then*

$$\operatorname{ch supp}(f * g) = \operatorname{ch supp} f + \operatorname{ch supp} g = \{x + y; x \in \operatorname{ch supp} f,\ y \in \operatorname{ch supp} g\}. \tag{2.2.21}$$

PROOF. Since $\operatorname{supp}(f * g) \subset \operatorname{supp} f + \operatorname{supp} g \subset \operatorname{ch supp} f + \operatorname{ch supp} g$, it is clear that the left-hand side of (2.2.21) is a subset of the right-hand side. It suffices to prove the opposite inclusion when $f, g \in C_0^\infty$. In fact, if $\chi \in C_0^\infty(\mathbf{R}^n)$ has support in the unit ball B, $\int \chi\, dx = 1$, and $\chi_\varepsilon(x) = \varepsilon^{-n}\chi(x/\varepsilon)$, then this special case gives

$$\operatorname{supp}(f * \varphi_\varepsilon) + \operatorname{supp}(g * \varphi_\varepsilon) \subset \operatorname{ch supp}(f * g * \varphi_\varepsilon * \varphi_\varepsilon) \subset \operatorname{ch supp}(f * g) + 2\varepsilon B.$$

When $\varepsilon \to 0$ it follows that $\operatorname{supp} f + \operatorname{supp} g \subset \operatorname{ch supp}(f * g)$ which implies that the right-hand side of (2.2.21) is a subset of the left-hand side.

Assume now that $f, g \in C_0^\infty(\mathbf{R}^n)$, set $h = f * g$ and introduce the partial Fourier transforms

$$F(\xi', x_n) = \int_{\mathbf{R}^{n-1}} e^{-i\langle x', \xi'\rangle} f(x', x_n)\, dx',\ G(\xi', x_n) = \int_{\mathbf{R}^{n-1}} e^{-i\langle x', \xi'\rangle} g(x', x_n)\, dx',$$
$$H(\xi', x_n) = \int_{\mathbf{R}^{n-1}} e^{-i\langle x', \xi'\rangle} h(x', x_n)\, dx'.$$

Then

$$H(\xi', x_n) = \int_{\mathbf{R}^{n-1}} e^{-i\langle x', \xi'\rangle} \int_{\mathbf{R}^n} f(x' - y', x_n - y_n) g(y', y_n)\, dy'\, dy_n$$
$$= \int_{\mathbf{R}} F(\xi', x_n - y_n) G(\xi', y_n)\, dy_n.$$

For *fixed* ξ' it follows from Theorem 2.1.11 that

$$\sup\{x_n; H(\xi', x_n) \neq 0\} = \sup\{x_n; F(\xi', x_n) \neq 0\} + \sup\{x_n; G(\xi', x_n) \neq 0\}.$$

Since F, G, H are analytic in ξ', the suprema are for almost all ξ' equal to the suprema when ξ' is also allowed to vary, for if say $F(\xi', x_n^0) \neq 0$ for one ξ' then this is true except for ξ' in a null set. Now $F(\xi', x_n) = 0$ for all $\xi' \in \mathbf{R}^{n-1}$ if and only if $f(x', x_n) = 0$ for all $x' \in \mathbf{R}^{n-1}$, and similarly for g and h. Hence

$$\sup\{x_n; (x', x_n) \in \operatorname{supp} h\} = \sup\{x_n; (x', x_n) \in \operatorname{supp} f\} + \sup\{x_n; (x', x_n) \in \operatorname{supp} g\}.$$

This means that $H_{\operatorname{supp} h}(\xi) = H_{\operatorname{supp} f}(\xi) + H_{\operatorname{supp} g}(\xi)$ if $\xi = (0, \dots, 0, 1)$. By a change of coordinates we conclude that this is true for every ξ, which means precisely that (2.2.21) holds. The proof is complete.

The last statement in Theorem 2.2.6 can be extended as follows:

THEOREM 2.2.8. *If $f \in \mathscr{S}'(\mathbf{R}^n)$ and $g \in \mathscr{E}'(\mathbf{R}^n)$ then $f * g \in \mathscr{S}'(\mathbf{R}^n)$ and the Fourier transform is equal to $\hat{f}\hat{g}$. If $\varphi \in \mathscr{S}$ then $f * \varphi \in \mathscr{S}'$ and the Fourier transform is equal to $\hat{f}\hat{\varphi}$.*

Note that the products $\hat{f}\hat{g}$ and $\hat{f}\hat{\varphi}$ are defined since $\hat{g} \in C^\infty$ and $\hat{\varphi} \in \mathscr{S}$.

PROOF. If $\varphi \in C_0^\infty(\mathbf{R}^n)$ then

$$\langle f * g, \varphi \rangle = \langle f, \check{g} * \varphi \rangle$$

by the definition of convolution. The right-hand side is a continuous linear form on $\mathscr{S}$, for if g is of order μ and $|y| < M$ when $y \in \operatorname{supp} g$ then

$$\sum_{|\alpha+\beta| \le N} \sup |x^\beta D^\alpha (\check{g} * \varphi)(x)| = \sum_{|\alpha+\beta| \le N} \sup |x^\beta g(D^\alpha \varphi(x + \cdot))|$$
$$\le C \sum_{|\alpha+\beta| \le N+\mu} \sup_x \sup_{|y|<M} |x^\beta D^\alpha \varphi(x+y)| \le C' \sum_{|\alpha+\beta| \le N+\mu} |x^\beta D^\alpha \varphi(x)|.$$

This proves that $f * g \in \mathscr{S}'$. If $f_j \in \mathscr{S}'$ and $f_j \to f$ in $\mathscr{S}'$ (with the weak topology) then $f_j * g \to f * g$ in $\mathscr{S}'$. Taking f_j with compact support we conclude using Theorem 2.2.6 that $\hat{f}_j \hat{g} = \widehat{f_j * g} \to \widehat{f * g}$, and since $\hat{f}_j \to \hat{f}$ in $\mathscr{S}'$ we obtain $\widehat{f * g} = \hat{f}\hat{g}$. The proof of the statement on $f * \varphi$ when $\varphi \in \mathscr{S}$ is just a repetition of the proof of Theorem 2.1.7 and it is left as an exercise.

The *spectrum* of a distribution $f \in \mathscr{S}'$ is by definition the support of the Fourier transform $\hat{f}$. When the spectrum is compact it follows from Theorem 2.2.6 that $f \in C^\infty$. The following inequality of Bernstein (see Theorem 2.1.8) is sometimes useful to estimate the derivatives of a distribution with compact spectrum.

THEOREM 2.2.9. *If $f \in L^p(\mathbf{R}^n)$ for some $p \in [1, \infty]$ and $\hat{f}$ has compact support, then*

$$\|P(\langle a, \partial \rangle) f\|_{L^p} \le |P(iH(a))| \|f\|_{L^p}, \quad a \in \mathbf{R}^n, \tag{2.2.22}$$

where $P(\tau)$ is any polynomial in one variable with only real zeros and

$$H(a) = \sup_{\xi \in \operatorname{supp} \hat{f}} |\langle a, \xi \rangle|. \tag{2.2.23}$$

PROOF. It suffices to prove the estimate (2.2.22) when P is linear and $a = (1, 0, \dots, 0)$. Set $\lambda = H(a)$ and $C = |P(i\lambda)|$. Then $P(i\lambda) = Ce^{i\alpha} = C(i \sin\alpha + \cos\alpha)$ for some $\alpha \in \mathbf{R}$, so $P(\tau) = C(\tau \sin\alpha/\lambda + \cos\alpha)$ and the theorem states that

$$\int_{\mathbf{R}^n} |\partial f(x)/\partial x_1 \sin\alpha/\lambda + f(x) \cos\alpha|^p \, dx \le \int_{\mathbf{R}^n} |f(x)|^p \, dx,$$

if $|\xi_1| \le \lambda$ when $\xi \in \operatorname{supp} \hat{f}$. Then the Fourier transform of $f(x_1, x_2, \dots, x_n)$ with respect to x_1 for fixed $x_2, \dots, x_n$ has support in $[-\lambda, \lambda]$, so the estimate follows at once from Theorem 2.1.8.

2.3. The Fourier transform of L^p spaces. If $f \in L^2(\mathbf{R}^n)$ then $\hat f \in L^2(\mathbf{R}^n)$ and $\|\hat f\|_2 = (2\pi)^{n/2}\|f\|_2$ by Parseval's formula. If $f \in L^1(\mathbf{R}^n)$ then $\hat f \in L^\infty(\mathbf{R}^n) \cap C(\mathbf{R}^n)$ and $\|\hat f\|_\infty \le \|f\|_1$. From these facts it follows that if $f \in L^p(\mathbf{R}^n)$ for some $p \in (1,2)$, then $\hat f \in L^2(\mathbf{R}^n) + L^\infty(\mathbf{R}^n) \subset L^1_{\text{loc}}(\mathbf{R}^n)$, for we can write $f = g + h$ with $g \in L^1(\mathbf{R}^n)$ and $h \in L^2(\mathbf{R}^n)$ for example by taking $h = f$ when $|f| < 1$ and $g = f$ when $|f| \ge 1$. We shall now prove a much more precise result:

THEOREM 2.3.1 (HAUSDORFF-YOUNG). *If $f \in L^p(\mathbf{R}^n)$ and $1 < p < 2$, then $\hat f \in L^{p'}(\mathbf{R}^n)$ and*

$$\|\hat f\|_{p'} \le (2\pi)^{n/p'}\|f\|_p, \quad \text{where } 1/p + 1/p' = 1. \tag{2.3.1}$$

The proof is an easy consequence of the *Riesz-Thorin interpolation theorem*:

THEOREM 2.3.2. *If T is a linear map from $L^{p_1} \cap L^{p_2}$ to $L^{q_1} \cap L^{q_2}$ where $p_j, q_j \in [1,\infty]$, such that*

$$\|Tf\|_{q_j} \le M_j\|f\|_{p_j}, \quad j = 1,2, \quad f \in L^{p_1} \cap L^{p_2}, \tag{2.3.2}$$

and if $1/p = t/p_1 + (1-t)/p_2$, $1/q = t/q_1 + (1-t)/q_2$ for some $t \in (0,1)$, then

$$\|Tf\|_q \le M_1^t M_2^{1-t}\|f\|_p, \quad f \in L^{p_1} \cap L^{p_2}. \tag{2.3.3}$$

For a proof we refer to Hörmander [1, Theorem 7.1.12].

PROOF OF THEOREM 2.3.1. From Theorem 2.3.2 it follows that (2.3.1) is valid when $f \in L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n)$. This is a dense subset of $L^p(\mathbf{R}^n)$ so the map $L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n) \ni f \mapsto \hat f \in L^{p'}(\mathbf{R}^n)$ extends uniquely to a linear map T from $L^p(\mathbf{R}^n)$ to $L^{p'}(\mathbf{R}^n)$ with norm $\le M_1^t M_2^{1-t}$. Since the map $L^p(\mathbf{R}^n) \ni f \mapsto \hat f \in \mathscr{S}'(\mathbf{R}^n)$ is continuous, it must be equal to T which proves the theorem.

The constant in (2.3.1) is not the best possible when $1 < p < 2$. We obtain a lower bound for the possible constants by taking $f(x) = e^{-a|x|^2/2}$ for some $a > 0$. Then $\hat f(\xi) = (2\pi/a)^{n/2}e^{-|\xi|^2/2a}$ so we have

$$\int\limits_{\mathbf{R}^n} |f(x)|^p\,dx = \int\limits_{\mathbf{R}^n} e^{-pa|x|^2/2}\,dx = (2\pi/pa)^{n/2}, \quad \int\limits_{\mathbf{R}^n} |\hat f(\xi)|^{p'}\,d\xi = (2\pi/a)^{np'/2}(2\pi a/p')^{n/2}.$$

Hence $\|\hat f\|_{p'}/\|f\|_p = C$ where

$$C = (2\pi)^{n/p'}(p^{1/p}/p'^{1/p'})^{n/2}. \tag{2.3.4}$$

(2.3.4) is in fact the best possible constant in (2.3.1) by a theorem of Beckner [1]. For a proof we refer to Lieb [1] where it is proved that for arbitrary operators with Gaussians kernels, such as the Fourier transformation, the best possible constants in L^p, L^q estimates can be found from the action on Gaussian functions. (The improvement of the constant (2.3.4) over the constant $(2\pi)^{n/p'}$ in (2.3.1) is at most $\sim 0.8635^{n/2}$ which occurs for $p \sim 1.19$. However, it is of course also interesting to know the extremals f.)

The Fourier transformation is not continuous in any L^p, L^q spaces apart from the cases given in Theorem 2.3.1:

PROPOSITION 2.3.3. *If $f \in L^p(\mathbf{R}^n)$ implies $\hat{f} \in L^q(\mathbf{R}^n)$ then $1 \le p \le 2$ and $\frac{1}{p} + \frac{1}{q} = 1$.*

PROOF. Since $L^p(\mathbf{R}^n) \ni f \mapsto \hat{f} \in \mathscr{S}'(\mathbf{R}^n)$ is continuous, the hypothesis means that this is a closed map into $L^q(\mathbf{R}^n)$. Hence it is continuous by the closed graph theorem, that is,

$$\|\hat{f}\|_q \le C\|f\|_p, \quad f \in L^p(\mathbf{R}^n). \tag{2.3.5}$$

If we replace $f(x)$ by $f(x/t)$ where $t > 0$ then $\hat{f}(\xi)$ is replaced by $t^n \hat{f}(t\xi)$, so (2.3.5) gives

$$t^{n(1-1/q)}\|\hat{f}\|_q \le Ct^{n/p}\|f\|_p, \quad f \in L^p(\mathbf{R}^n),\ t > 0,$$

which implies $1/p + 1/q = 1$.

As an example of the remarks after the proof of Theorem 2.3.1 on the importance of Gaussians we shall now prove that $p \le 2$ by checking (2.3.5) for Gaussians. Of course it does not suffice to use real Gaussians since they are all equivalent under changes of coordinates, so we take

$$f(x) = \exp(-a|x|^2/2), \quad \hat{f}(\xi) = (2\pi/a)^{n/2}\exp(-|x|^2/2a),$$

where $\operatorname{Re} a > 0$. As above we obtain

$$\|f\|_p^p = (2\pi/p\operatorname{Re} a)^{n/2}, \quad \|\hat{f}\|_q^q = (2\pi|a|^2/q\operatorname{Re} a)^{n/2}(2\pi/|a|)^{nq/2},$$

so (2.3.5) requires that

$$C \ge (2\pi)^{n(1+1/q-1/p)/2}(p^{1/p}/q^{1/q})^{n/2}(\operatorname{Re} a)^{n(1/p-1/q)/2}|a|^{n(1/q-1/2)}, \quad \operatorname{Re} a > 0.$$

For reasons of homogeneity we see again that this implies $1/p + 1/q = 1$, and when $\operatorname{Re} a \to 0$ while $|a| = 1$ we find that $q \ge p$, which proves the statement.

The Fourier transform of a function in L^p is usually not even in L^1_{loc} when $p > 2$:

THEOREM 2.3.4. *If k is an integer with $0 \le k < n/2$ then one can find f such that $f \in L^p(\mathbf{R}^n) \cap C(\mathbf{R}^n)$ for every $p \in (2, \infty]$ with $k < n(1/2 - 1/p)$ and $\hat{f}$ is not a distribution of order k in any open subset of $\mathbf{R}^n$. On the other hand, $\hat{f}$ is of order k for every $f \in L^p$ if $k > n(1/2 - 1/p)$.*

PROOF. The intersection $\mathscr{F}$ of $C(\mathbf{R}^n)$ and all $L^p(\mathbf{R}^n)$ with $k < n(1/2 - 1/p)$ is a Fréchet space with the seminorms $f \mapsto \|f\|_p$ for $p \in (2n/(n-2k), \infty]$. The first statement will follow if we prove that for every ball $\Omega \subset \mathbf{R}^n$ with rational center and rational radius the set M_Ω of all $f \in \mathscr{F}$ with $\hat{f}$ of order k in Ω is of the first category. In fact, the union of M_Ω for all such balls Ω is then of the first category, and all other $f \in \mathscr{F}$ have the required property.

If M_Ω is not of the first category it follows from Banach's theorem that the map $\mathscr{F} \ni f \mapsto \hat{f}|_\Omega$ is continuous from $\mathscr{F}$ to the Fréchet space $\mathscr{D}'^k(\Omega)$. If K is a compact subset of Ω, with interior points, it follows that

$$|\langle \hat{f}, \varphi\rangle| \le C_K N(f) \sum_{|\alpha| \le k} \sup |D^\alpha \varphi|, \quad \varphi \in C_0^\infty(K), \tag{2.3.6}$$

where $N(f)$ is a seminorm in $\mathscr{F}$, so $N(f) \le C(\|f\|_p + \|f\|_\infty)$ for some p with $k < n(1/2 - 1/p)$. Choose $g \in \mathscr{S} \setminus \{0\}$ so that $\hat{g} \in C_0^\infty(K)$, and define $f_t \in \mathscr{S}$, $\varphi_t \in C_0^\infty(K)$ for $t > 0$ by

$$\hat{f}_t(\xi) = \hat{g}(\xi) e^{it|\xi|^2/2}, \quad \varphi_t(\xi) = \overline{\hat{f}_t(\xi)}.$$

Then it follows that

$$\langle \hat{f}_t, \varphi_t \rangle = \int |\hat{g}(\xi)|^2 \, d\xi, \quad \sum_{|\alpha| \le k} \sup |D^\alpha \varphi_t| = O(t^k) \quad \text{as } t \to +\infty.$$

The inverse Fourier transform of $\xi \mapsto e^{it|\xi|^2/2}$ is $x \mapsto ct^{-n/2} e^{-i|x|^2/2t}$ with a constant $c \ne 0$ which is not essential now. Thus

$$f_t(x) = ct^{-n/2} \int e^{-i|x-y|^2/2t} g(y) \, dy$$

which implies that $|f_t(x)| \le |c| t^{-n/2} \|g\|_1$. By Parseval's formula $\|f_t\|_2 = \|g\|_2$, hence

$$\|f_t\|_p \le (\|f_t\|_2^2 \|f_t\|_\infty^{p-2})^{1/p} \le Ct^{n(1/p - 1/2)} = o(t^{-k}) \quad \text{as } t \to \infty$$

if $k < n(1/2 - 1/p)$. Thus the left-hand side of (2.3.6) with $f = f_t$ and $\varphi = \varphi_t$ is independent of t while the right-hand side $\to 0$ as $t \to \infty$. This is a contradiction which completes the proof of the first statement.

The second statement is much weaker than the Hausdorff-Young theorem unless $p > 2$, which we assume now. If $f \in L^p(\mathbf{R}^n)$ and $\varphi \in \mathscr{S}(\mathbf{R}^n)$ then

$$|\langle \hat{f}, \varphi \rangle| = |\langle f, \hat{\varphi} \rangle| \le \|f\|_p \|\hat{\varphi}\|_{p'}, \tag{2.3.7}$$

where $1/p + 1/p' = 1$. By Hölder's inequality with exponents $2/p'$ and q, $1/q + p'/2 = 1$, we have

$$\|\hat{\varphi}\|_{p'} \le C\Big(\int_{\mathbf{R}^n} (1 + |\xi|^2)^k |\hat{\varphi}(\xi)|^2 \, d\xi \Big)^{\frac{1}{2}}, \tag{2.3.8}$$

for $\int (1 + |\xi|^2)^{-kp'q/2} \, d\xi < \infty$ since $kp'/n > (1/p' - 1/2)p' = 1/q$. By Parseval's formula

$$\Big(\int_{\mathbf{R}^n} (1 + |\xi|^2)^k |\hat{\varphi}(\xi)|^2 \, d\xi \Big)^{\frac{1}{2}} \le C \sum_{|\alpha| \le k} \|D^\alpha \varphi\|_2. \tag{2.3.9}$$

The estimates (2.3.7), (2.3.8), (2.3.9) imply that $\hat{f}$ is of order $\le k$.

We note in particular that when $k = 0$ the theorem gives a function in $L^p(\mathbf{R}^n) \cap C(\mathbf{R}^n)$ for every $p > 2$ such that $\hat{f}$ is not even a measure in any open subset of $\mathbf{R}^n$. A minor modification of the proof, which we leave as an exercise, shows that we can choose f so that in addition $f \in C^\infty$ and all derivatives are bounded. Thus it is only the insufficient decrease at infinity which is responsible for the lack of regularity of the Fourier transform.

The proof of the second part of the theorem relied on the equality (2.3.9), which can be extended to an equivalence between the two sides. We have here encountered the simplest *Sobolev spaces*:

DEFINITION 2.3.5. If s is a real number then $H_{(s)}(\mathbf{R}^n)$ denotes the space of all $u \in \mathscr{S}'(\mathbf{R}^n)$ such that $\hat{u} \in L^2(\mathbf{R}^n, (1+|\xi|^2)^s d\xi/(2\pi)^n)$, with the norm

$$\|u\|_{(s)} = \Big((2\pi)^{-n} \int |\hat{u}(\xi)|^2 (1+|\xi|^2)^s \, d\xi\Big)^{\frac{1}{2}}.$$

From Parseval's formula it follows at once that $H_{(0)}(\mathbf{R}^n) = L^2(\mathbf{R}^n)$. We have $u \in H_{(s+1)}(\mathbf{R}^n)$ if and only if $u \in H_{(s)}(\mathbf{R}^n)$ and $D_j u \in H_{(s)}(\mathbf{R}^n)$ for $j = 1, \dots, n$, and then we have

$$\|u\|^2_{(s+1)} = \|u\|^2_{(s)} + \sum_1^n \|D_j u\|^2_{(s)}. \tag{2.3.10}$$

Repeated use of this observation shows that $H_{(s)}(\mathbf{R}^n)$ consists of all u with $D^\alpha u \in L^2(\mathbf{R}^n)$ when $|\alpha| \le s$, if s is a positive integer, and this is what (2.3.9) expressed in part. We can also work in the other direction: $u \in H_{(s)}(\mathbf{R}^n)$ if and only if u has a representation $u = v_0 + \sum_1^n D_j v_j$ where $v_j \in H_{(s+1)}(\mathbf{R}^n)$, $j = 0, 1, \dots, n$; it can be chosen so that

$$\|u\|^2_{(s)} = \sum_0^n \|v_j\|^2_{(s+1)}. \tag{2.3.11}$$

In fact, we can take $\hat{v}_0(\xi) = \hat{u}(\xi)/(1+|\xi|^2)$ and $\hat{v}_j(\xi) = \hat{u}(\xi)\xi_j/(1+|\xi|^2)$. Roughly speaking, $H_{(s)}(\mathbf{R}^n)$ consists for negative integer s of distributions which are sums of derivatives of order $\le -s$ of functions in L^2. If we describe the functions in $H_{(s)}(\mathbf{R}^n)$ for $0 < s < 1$ without reference to the Fourier transformation we shall obtain a similar description of all spaces $H_{(s)}(\mathbf{R}^n)$. First note that for $0 < s < 1$

$$\|u\|^2_{(s)} \le (2\pi)^{-n} \int_{\mathbf{R}^n} |\hat{u}(\xi)|^2 (1+|\xi|^{2s}) \, d\xi \le 2\|u\|^2_{(s)}. \tag{2.3.12}$$

This is equivalent to the inequalities

$$(1+|\xi|^2)^s \le 1 + |\xi|^{2s} \le 2(1+|\xi|^2)^s, \quad 0 < s < 1.$$

The second is trivial and the first follows since $1 \ge (1+|\xi|^2)^{s-1}$ and $|\xi|^{2s} \ge |\xi|^2(1+|\xi|^2)^{s-1}$. For $0 < s < 1$ there is a constant A_s such that

$$(2\pi)^{-n} \int_{\mathbf{R}^n} |\hat{u}(\xi)|^2 (1+|\xi|^{2s}) \, d\xi = \int_{\mathbf{R}^n} |u(x)|^2 \, dx + A_s \iint_{\mathbf{R}^n \times \mathbf{R}^n} |u(x+y) - u(x)|^2 |y|^{-n-2s} \, dx \, dy. \tag{2.3.13}$$

Thus $H_{(s)}(\mathbf{R}^n)$ consists when $0 < s < 1$ of all $u \in L^2(\mathbf{R}^n)$ such that the right-hand side of (2.3.13) is finite; this is a kind of Hölder condition in an L^2 sense. To prove (2.3.13) we note that since the Fourier transform of $x \mapsto u(x+y) - u(x)$ is $\xi \mapsto (e^{i\langle y,\xi\rangle} - 1)\hat{u}(\xi)$ the identity is equivalent to

$$A_s \int |e^{i\langle y,\xi\rangle} - 1|^2 |y|^{-n-2s}\, dy = |\xi|^{2s}. \tag{2.3.14}$$

The integral on the left-hand side converges at 0 since $s < 1$ and at ∞ since $s > 0$. An orthogonal transformation of y proves that it is a function of $|\xi|$ only, and replacing y by ty for some $t > 0$ proves that it is homogeneous in $|\xi|$ of degree $2s$, which proves (2.3.14). (It is not hard to calculate A_s in terms of the Γ function. Even without doing that it is easy to see that $2(1-s)/A_s$ converges to the volume of the unit ball when $s \to 1$ and that s/A_s converges to the area of the unit sphere when $s \to 0$.)

The spaces $H_{(s)}(\mathbf{R}^n)$ are $\mathscr{S}$ modules, that is, $\varphi \in \mathscr{S}$ and $f \in H_{(s)}(\mathbf{R}^n)$ implies $\varphi f \in H_{(s)}(\mathbf{R}^n)$. This is an easy consequence of the fact that the Fourier transform of φf is $(2\pi)^{-n}\hat{\varphi} * \hat{f}$. However, we shall prove a much more precise result.

LEMMA 2.3.6. *If χ is a bounded Lipschitz continuous function in $\mathbf{R}^n$ and $f \in H_{(s)}(\mathbf{R}^n)$ for some $s \in [0,1]$, then $\chi f \in H_{(s)}(\mathbf{R}^n)$ and*

$$\|\chi f\|_{(s)} \le \sqrt{2\sup(|\chi|^2 + |\chi'|^2)}\|f\|_{(s)}. \tag{2.3.15}$$

PROOF. The statement is obvious when $s = 0$, and when $s = 1$ it follows since $D_j(\chi f) = (D_j\chi)f + \chi D_j f$, hence

$$|\chi f|^2 + \sum_1^n |D_j(\chi f)|^2 \le 2(|\chi|^2 + |\chi'|^2)(|f|^2 + \sum_1^n |D_j f|^2).$$

When $0 < s < 1$ we shall use a complex interpolation argument close to the proof of the Riesz-Thorin interpolation theorem. For $s \in \mathbf{C}$ we shall denote by $(1+|D|^2)^s f$ the inverse Fourier transform of $\xi \mapsto (1+|\xi|^2)^s \hat{f}(\xi)$, which is a function in $\mathscr{S}$ if $f \in \mathscr{S}$ and is analytic as a function of s. With $f, g \in \mathscr{S}$ we form

$$\Phi(s) = \langle (1+|D|^2)^{s/2}\chi(1+|D|^2)^{-s/2} f, g\rangle.$$

This is a bounded analytic function for $0 \le \operatorname{Re} s \le 1$, and

$$|\Phi(s)| \le \begin{cases} \sup|\chi|\|f\|_2\|g\|_2, & \text{if } \operatorname{Re} s = 0, \\ \sqrt{2\sup(|\chi|^2 + |\chi'|^2)}\|f\|_2\|g\|_2, & \text{if } \operatorname{Re} s = 1. \end{cases}$$

This is obvious when $\operatorname{Re} s = 0$ for $(1+|D|^2)^{s/2}$ is then a unitary operator in L^2. Since

$$\|(1+|D|^2)^{-s/2} f\|_{(1)} = \|f\|_2, \quad \|(1+|D|^2)^{s/2} w\|_2 = \|w\|_{(1)}, \quad \text{if } \operatorname{Re} s = 1,$$

the estimate follows when $\operatorname{Re} s = 1$. By the maximum principle applied to $\Phi(s)/(1+\varepsilon s)$ for $\varepsilon > 0$ we obtain when $\varepsilon \to 0$ that

$$|\Phi(s)| \le \sqrt{2\sup(|\chi|^2 + |\chi'|^2)}\|f\|_2\|g\|_2, \quad 0 < \operatorname{Re} s < 1.$$

Taking s real we obtain

$$\|\chi(1+|D|^2)^{-s/2} f\|_{(s)} \le \sqrt{2\sup(|\chi|^2 + |\chi'|^2)}\|f\|_2$$

which proves (2.3.15) when f is replaced by $(1+|D|^2)^{s/2} f$. The lemma is proved.

PROPOSITION 2.3.7. *If $\chi \in C^k(\mathbf{R}^n)$ and the derivatives of order $\le k$ are bounded, then $\chi f \in H_{(s)}(\mathbf{R}^n)$ if $f \in H_{(s)}(\mathbf{R}^n)$ and $|s| \le k$; we have*

$$\|\chi f\|_{(s)} \le C_k \sum_{|\alpha|\le k} \sup |D^\alpha \chi| \|f\|_{(s)}. \tag{2.3.16}$$

PROOF. Lemma 2.3.6 proves the statement when $k = 1$ and $0 \le s \le 1$. Using (2.3.10) we obtain inductively that it is true for every positive integer k when $0 \le s \le k$. If $-k \le s < 0$ and $f \in \mathscr{S}$ then

$$\|\chi f\|_{(s)} = \sup_{0\ne g\in\mathscr{S}} |\langle \chi f, g\rangle|/\|g\|_{(-s)} = \sup_{0\ne g\in\mathscr{S}} |\langle f, \chi g\rangle|/\|g\|_{(-s)}$$
$$\le \sup_{g\in\mathscr{S}} \|f\|_{(s)}\|\chi g\|_{(-s)}/\|g\|_{(-s)} \le \|f\|_{(s)} C_k \sum_{|\alpha|\le k} \sup |D^\alpha \chi|,$$

which completes the proof.

The theorem just proved or already the simple special case where $\chi \in C_0^\infty(\mathbf{R}^n)$ leads us to define $H_{(s)}^{\mathrm{loc}}(\Omega)$ for any open subset Ω of $\mathbf{R}^n$ as

$$H_{(s)}^{\mathrm{loc}}(\Omega) = \{f \in \mathscr{D}'(\Omega); \chi f \in H_{(s)}(\mathbf{R}^n) \text{ if } \chi \in C_0^\infty(\Omega)\}. \tag{2.3.17}$$

Here $\chi f \in \mathscr{E}'(\Omega)$ is regarded as an element in $\mathscr{E}'(\mathbf{R}^n)$. It follows from Proposition 2.3.7 that $H_{(s)}(\mathbf{R}^n) \subset H_{(s)}^{\mathrm{loc}}(\mathbf{R}^n)$. If $\Omega_1 \subset \Omega_2$ then the restriction of $H_{(s)}^{\mathrm{loc}}(\Omega_2)$ to Ω_1 is in $H_{(s)}^{\mathrm{loc}}(\Omega_1)$. If $f \in H_{(s)}^{\mathrm{loc}}(\Omega)$ then there is for every $x \in \Omega$ some $f_x \in H_{(s)}(\mathbf{R}^n)$ with $f_x = f$ in a neighborhood of x, for we can choose $f_x = \chi f$ with $\chi \in C_0^\infty(\Omega)$ equal to 1 in a neighborhood of x. Conversely, every f with this property is in $H_{(s)}^{\mathrm{loc}}(\Omega)$. In fact, if $\chi \in C_0^\infty(\Omega)$ we can for every $x \in \operatorname{supp}\chi$ choose $f_x \in H_{(s)}(\mathbf{R}^n)$ equal to f in a neighborhood $O_x \subset \Omega$ of x. By the Borel-Lebesgue lemma $\operatorname{supp}\chi$ is covered by finitely many neighborhoods $O_{x_1}, \dots, O_{x_N}$. We can choose $\chi_j \in C_0^\infty(O_{x_j})$, $j = 1, \dots, N$, so that $\sum_1^N \chi_j = 1$ in $\operatorname{supp}\chi$ and conclude using Proposition 2.3.7 that

$$\chi f = \sum_1^N \chi(\chi_j f) \in H_{(s)}(\mathbf{R}^n),$$

which means that $f \in H_{(s)}^{\mathrm{loc}}(\Omega)$. Thus $H_{(s)}^{\mathrm{loc}}(\Omega)$ is indeed the space of distributions in Ω which are locally in $H_{(s)}(\mathbf{R}^n)$.

Instead of using the order of a distribution f as a measure of its regularity, as in Theorem 2.3.4, it is usually better to describe it by regularity conditions of the form $f \in H_{(s)}^{\mathrm{loc}}$. This gives a continuous scale of regularity conditions with s ranging from $-\infty$ to $+\infty$, and the fact that L^2 conditions are exactly translated by the Fourier transformation leads to precise statements. As an example we reformulate Theorem 2.3.4 with our new notions.

THEOREM 2.3.8. *If $1 \le q < 2$ then the Fourier transform of $H_{(s)}(\mathbf{R}^n)$ is contained in $L^q(\mathbf{R}^n)$ if and only if $s > n(1/q - 1/2)$. If $2 < p \le \infty$ then the Fourier transform of $L^p(\mathbf{R}^n)$ is contained in $H_{(-s)}$ if and only if $s > n(1/2 - 1/p)$.*

PROOF. The Fourier transform of $H_{(s)}(\mathbf{R}^n)$ is $L^2(\mathbf{R}^n, (1+|\xi|^2)^s\, d\xi)$ by definition. By Hölder's inequality with exponents $2/q$ and $2/(2-q)$

$$\int_{\mathbf{R}^n} |f(\xi)|^q\, d\xi \le M\Big(\int_{\mathbf{R}^n} |f(\xi)|^2(1+|\xi|^2)^s\, d\xi\Big)^{\frac{q}{2}},\ M = \Big(\int_{\mathbf{R}^n}(1+|\xi|^2)^{-sq/(2-q)}\, d\xi\Big)^{\frac{2-q}{2}}.$$

M is the best possible constant in this estimate, and if $M = +\infty$ then there is some $f \in L^2(\mathbf{R}^n, (2\pi)^{-n}(1+|\xi|^2)^s\, d\xi)$ such that $f \notin L^q$. Now M is finite if and only if $sq/(2-q) > n/2$, that is, $s > n(1/q - 1/2)$, which proves the first statement. The second one is dual: If $u \in L^p(\mathbf{R}^n)$ and $\varphi \in \mathscr{S}(\mathbf{R}^n)$ we have with $1/p + 1/q = 1$

$$|\langle \hat{u}, \varphi\rangle| = |\langle u, \hat{\varphi}\rangle| \le \|u\|_p \|\hat{\varphi}\|_q \le \|u\|_p M^{1/q}(2\pi)^{n/2}\|\varphi\|_{(s)}.$$

This proves that $\hat{u} \in H_{(-s)}(\mathbf{R}^n)$ if $s > n(1/q - 1/2) = n(1/2 - 1/p)$.

Conversely, if $\hat{u} \in H_{(-s)}(\mathbf{R}^n)$ for all $u \in L^p(\mathbf{R}^n)$ it follows from Banach's theorem that $\|\hat{u}\|_{(-s)} \le C\|u\|_p$, $u \in L^p(\mathbf{R}^n)$, hence

$$|\langle u, \hat{\varphi}\rangle| = |\langle \hat{u}, \varphi\rangle| \le \|\hat{u}\|_{(-s)}\|\varphi\|_{(s)} \le C\|u\|_p\|\varphi\|_{(s)}, \quad \varphi \in \mathscr{S}(\mathbf{R}^n),$$

so $\|\hat{\varphi}\|_q \le C\|\varphi\|_{(s)}$. This implies that the Fourier transform of $H_{(s)}(\mathbf{R}^n)$ is contained in $L^q(\mathbf{R}^n)$, so the first part of the proof gives $s > n(1/q - 1/2) = n(1/2 - 1/p)$. The proof is complete.

In particular we note that $\hat{f} \in L^1(\mathbf{R}^n)$ if $f \in H_{(s)}(\mathbf{R}^n)$ for some $s > n/2$, so Fourier's inversion formula

$$f(x) = (2\pi)^{-n}\int_{\mathbf{R}^n} e^{i\langle x,\xi\rangle}\hat{f}(\xi)\, d\xi$$

is then absolutely convergent and proves that f is continuous, $f(x) \to 0$ as $x \to \infty$. Note that this improves Proposition 2.1.2 a great deal. There is a corresponding improvement of Proposition 2.2.2: *The Fourier series of a periodic function in $H^{\mathrm{loc}}_{(s)}(\mathbf{R}^n)$ converges absolutely if $s > n/2$.* This is essentially a theorem of S. Bernstein.

2.4. The method of stationary phase. The core of the proof of Theorem 2.3.4 was the explicit form of the Fourier transform of a Gaussian. We shall now discuss more systematically the role of Gaussians in the study of oscillatory integrals. First we shall motivate why they occur.

THEOREM 2.4.1. *Let $u \in \mathscr{E}'(K)$, where $K \subset \mathbf{R}^n$ is compact, and assume that $D^\alpha u \in L^1$ when $|\alpha| \le k$. If $\varphi \in C^{k+1}$ in a neighborhood of K and φ is real valued, $\varphi' \ne 0$ on K, then*

$$\Big|\int_K u e^{i\tau\varphi}\, dx\Big| \le C_{k,\varphi}\tau^{-k}\sum_{|\alpha|\le k}\int |D^\alpha u|\, dx, \quad \tau > 0. \tag{2.4.1}$$

PROOF. First assume that $\varphi(x) = x_1$. Then

$$\int (D_1^k u) e^{i\tau\varphi}\, dx = (-\tau)^k \int u e^{i\tau\varphi}\, dx,$$

which proves (2.4.1) in this case. (This is just the same argument as in (2.1.10) for example which we have used to prove that the Fourier transform of a smooth function is rapidly decreasing.) If $k \ge 1$ the implicit function theorem shows that for any $x^0 \in K$ one can find new local coordinates $y = \Phi(x)$ in a neighborhood ω such that $\Phi_1(x) = \varphi(x)$. If $\chi \in C_0^\infty(\omega)$ and $\Psi = \Phi^{-1}$ then

$$\int \chi u e^{i\tau\varphi}\, dx = \int (\chi u) \circ \Psi(y) e^{i\tau y_1} |\det \Psi'(y)|\, dy$$

has a bound of the form (2.4.1) by the first part of the proof. Using a partition of unity we complete the proof.

Occasionally it is useful to have control of how the constant $C_{k,\varphi}$ in (2.4.1) depends on φ, so we shall elaborate this point in the following theorem where we also allow φ not to be real:

THEOREM 2.4.1$'$. *Let $u \in \mathscr{E}'(K)$ where K is a compact subset of $\mathbf{R}^n$, and assume that $D^\alpha u \in L^1$ when $|\alpha| \le k$. If $\varphi \in C^{k+1}$ in a neighborhood of K and $\operatorname{Im}\varphi \ge 0$, $\varphi' \ne 0$ in K, then*

$$\Big|\int_K u e^{i\tau\varphi}\, dx\Big| \le C_k \tau^{-k} \sum_{|\alpha|\le k} \int |D^\alpha u| |\varphi'|^{|\alpha|-2k} N_{k-|\alpha|}\, dx, \quad \tau > 0, \tag{2.4.1$'$}$$

where

$$N_j = \sum_{|\alpha_1|+\cdots+|\alpha_j|=2j, 1\le|\alpha_1|,\ldots,1\le|\alpha_j|} |D^{\alpha_1}\varphi| \cdots |D^{\alpha_j}\varphi|. \tag{2.4.2}$$

PROOF. We shall prove (2.4.1)$'$ by induction with respect to k. When $k = 0$ the statement is trivial. Assume that $k > 0$ and that it has already been proved with k replaced by $k-1$, and write

$$u e^{i\tau\varphi} = u \sum_1^n |\partial_j \varphi|^2 |\varphi'|^{-2} e^{i\tau\varphi} = \sum_1^n u(\partial_j\bar\varphi)|\varphi'|^{-2} \partial_j e^{i\tau\varphi}/i\tau.$$

An integration by parts gives now

$$\int u e^{i\tau\varphi}\, dx = \frac{i}{\tau} \sum_1^n \int \partial_j\big(u(\partial_j\bar\varphi)|\varphi'|^{-2}\big) e^{i\tau\varphi}\, dx.$$

If we apply the inductive hypothesis, with k replaced by $k-1$ in (2.4.1)$'$, it follows that

$$\tau^k\Big|\int ue^{i\tau\varphi}\,dx\Big| \le C_{k-1}\sum_{j=1}^n\sum_{|\alpha|\le k-1}\int|\partial^\alpha\partial_j(u(\partial_j\bar\varphi)|\varphi'|^{-2})||\varphi'|^{|\alpha|-2k+2}N_{k-1-|\alpha|}\,dx.$$

Letting $\partial^\alpha\partial_j$ act here gives a sum of terms where ∂^β acts on u and ∂^γ acts on $\partial_j\bar\varphi/|\varphi'|^2$, where $|\beta|+|\gamma|=|\alpha|+1$. The induction will be successful if we can prove that

$$|\partial^\gamma((\partial_j\bar\varphi)|\varphi'|^{-2})||\varphi'|^{|\alpha|-2k+2}N_{k-1-|\alpha|}\le C|\varphi'|^{|\beta|-2k}N_{k-|\beta|},$$

that is, since $|\alpha|-2k+2-|\beta|+2k=|\gamma|+1$

$$|\partial^\gamma((\partial_j\bar\varphi)|\varphi'|^{-2})||\varphi'|^{|\gamma|+1}N_{k-|\gamma|-|\beta|}\le CN_{k-|\beta|},\quad |\beta|+|\gamma|\le k.$$

By the definition of N_l it is sufficient to prove this when $|\gamma|+|\beta|=k$, that is, prove that

$$|\partial^\gamma((\partial_j\bar\varphi)|\varphi'|^{-2})||\varphi'|^{|\gamma|+1}\le CN_{|\gamma|}.$$

Now it is clear that

$$|\varphi'|^{2+2|\gamma|}\partial^\gamma\big((\partial_j\bar\varphi)|\varphi'|^{-2}\big)$$

is a homogeneous polynomial of degree $1+2|\gamma|$ in $\psi=\varphi'$ and $\bar\psi$ and their derivatives, with the total order of differentiation in each term equal to $|\gamma|$. Hence $1+|\gamma|$ factors are not differentiated, and the product of the other $|\gamma|$ is bounded by $N_{|\gamma|}$. The proof is complete.

Remark. It is clear that one can say much more where $\operatorname{Im}\varphi>0$. We refer to Hörmander [1, Theorem 7.7.1] for a precise estimate which takes this into account.

When examining the asymptotic properties of an oscillatory integral $\int ue^{i\tau\varphi}$ with u of compact support, φ real valued, and very smooth u and φ, we know from Theorem 2.4.1 that the main contributions will come from points where (2.4.1) is not applicable, that is, where $\varphi'=0$. The method of stationary phase which is the subject of this section consists of a study of the contributions from such points. For the sake of simplicity we shall not insist on minimal regularity conditions in what follows. The first step is to examine how much it is possible to simplify a stationary point by a change of coordinates as we did in the proof of Theorem 2.4.1.

Lemma 2.4.2 (Morse). *If φ is a real valued C^∞ function in a neighborhood of $x^0\in\mathbf{R}^n$ such that $\varphi'(x^0)=0$ but $\varphi''(x^0)\ne 0$, then there is a diffeomorphism ψ of a neighborhood of the origin on a neighborhood of x^0 such that $\psi(0)=x^0$ and*

$$\varphi(\psi(y))=\varphi(x^0)\pm y_1^2+\varrho(y_2,\dots,y_n). \tag{2.4.3}$$

Proof. It is no restriction to assume that $\varphi(x^0)=0$, and by a preliminary affine transformation we can achieve that $x^0=0$ and that $\partial_1^2\varphi(0)\ne 0$. By the implicit function theorem the equation $\partial_1\varphi(x_1,x')=0$ has a unique C^∞ solution $x_1=\psi(x')$ with $\psi(0)=0$, when $x'=(x_2,\dots,x_n)$ is close to the origin in $\mathbf{R}^{n-1}$. Taking $x_1-\psi(x')$ as a new variable instead of x_1 we reduce the proof to the case where $\partial_1\varphi(0,x')=0$ in a neighborhood of the origin. By Taylor's formula

$$\varphi(x)=\varphi(0,x')+x_1^2q(x),\quad q(x)=\int_0^1(\partial_1^2\varphi)(tx_1,x')(1-t)\,dt.$$

Here $q\in C^\infty$ in a neighborhood of the origin and $q(0)=\frac{1}{2}\partial_1^2\varphi(0)$. Taking $y_1=x_1\sqrt{|q(x)|}$ and $y'=x'$ we obtain $\varphi(x)=\varphi(0,y')\pm y_1^2$ as claimed in (2.4.3).

The lemma can again be applied to the error term ϱ if $\varrho''(0)\neq 0$. In particular, when $\det\varphi''(x^0)\neq 0$ we can continue until we have obtained a change of variables making $\varphi(\psi(y))$ equal to a non-degenerate quadratic form A in a neighborhood of the origin. We shall then say that x^0 is a *non-degenerate critical point*. Since $A(y)=\frac{1}{2}\varphi''(\psi'(0)y)$, it has the same signature as φ'' but there is no other condition since all non-degenerate quadratic forms with the same signature are equivalent under linear transformations. Note that non-degenerate critical points are isolated by the implicit function theorem.

Next we shall formalize a part of Example 5 given after Theorem 2.1.5 and Example 4 given after Theorem 2.2.5.

LEMMA 2.4.3. *If A is a real symmetric $n\times n$ matrix with* $\det A\neq 0$, *then the Fourier transform of the Gaussian* $\mathbf{R}^n\ni x\mapsto\exp(i\langle Ax,x\rangle/2)$ *is*

$$\mathbf{R}^n\ni\xi\mapsto(2\pi)^{\frac{n}{2}}|\det A|^{-\frac{1}{2}}e^{\frac{\pi i}{4}\operatorname{sgn}A}\exp(-i\langle A^{-1}\xi,\xi\rangle/2),$$

where $\operatorname{sgn}A$ *is the number of positive eigenvalues minus the number of negative eigenvalues.*

PROOF. First assume that $n=1$. Then Example 5 after Theorem 2.1.5 gives for $\varepsilon>0$ that the Fourier transform of $x\mapsto\exp((iA-\varepsilon)x^2)/2)$ is

$$\xi\mapsto\sqrt{2\pi/(\varepsilon-iA)}\exp((iA-\varepsilon)^{-1}\xi^2/2)$$

with the square root in the right half plane in $\mathbf{C}$. When $\varepsilon\to 0$ the square root converges to $\sqrt{2\pi/|A|}\exp(\frac{\pi i}{4}\operatorname{sgn}A)$ and the exponential converges boundedly to $\exp(-iA^{-1}\xi^2/2)$, so the lemma is true when $n=1$. Hence it follows if A is a diagonal matrix. We can always choose a linear bijection $T:\mathbf{R}^n\to\mathbf{R}^n$ such that $x\mapsto\langle ATx,Tx\rangle$ has diagonal form. Then the Fourier transform of $x\mapsto\exp(i\langle ATx,Tx\rangle)$ is

$$\xi\mapsto(2\pi)^{\frac{n}{2}}|\det A|^{-\frac{1}{2}}|\det T|^{-1}e^{\frac{\pi i}{4}\operatorname{sgn}A}\exp(-i\langle A^{-1}{}^tT^{-1}\xi,{}^tT^{-1}\xi\rangle),$$

because $T^{-1}A^{-1}{}^tT^{-1}$ is the inverse of ${}^tTAT$, which proves the lemma in view of Theorem 2.2.5.

If $u\in\mathscr{S}$ it follows from Lemma 2.4.3 and Fourier's inversion formula that

(2.4.4)
$$\int u(x)e^{i\tau\langle Ax,x\rangle/2}=(2\pi\tau)^{-\frac{n}{2}}|\det A|^{-\frac{1}{2}}e^{\frac{\pi i}{4}\operatorname{sgn}A}\int\exp(-i\langle A^{-1}\xi,\xi\rangle/2\tau)\hat{u}(\xi)\,d\xi,\quad\tau>0.$$

The advantage of this formula is that for large τ we can make a Taylor expansion of the exponential in the right-hand side, using the fact that by Taylor's formula, for every positive integer k,

$$|e^{it} - \sum_{j<k} (it)^j/j!| \le |t|^k/k!, \quad t \in \mathbf{R}.$$

Hence we obtain using Fourier's inversion formula

$$\Big| \int u(x) e^{i\tau\langle Ax,x\rangle/2}\, dx - (2\pi/\tau)^{\frac{n}{2}} |\det A|^{-\frac12} e^{\frac{\pi i}{4}\operatorname{sgn} A} \sum_{j<k} (\langle A^{-1}D, D\rangle/2i\tau)^j u(0)/j! \Big|$$
$$\le (2\pi\tau)^{-\frac{n}{2}} |\det A|^{-\frac12} (2\tau)^{-k} \int |\langle A^{-1}\xi,\xi\rangle|^k |\hat u(\xi)|\, d\xi/k!.$$

We can estimate the right-hand side by means of Theorem 2.3.8 which gives:

Proposition 2.4.4. *If A is a non-singular symmetric $n \times n$ matrix and $u \in \mathscr{S}$, $\tau > 0$ and s is an integer $> n/2$, we have for every integer $k > 0$*

$$\Big| \int u(x) e^{i\tau\langle Ax,x\rangle/2}\, dx - (2\pi/\tau)^{\frac{n}{2}} |\det A|^{-\frac12} e^{\frac{\pi i}{4}\operatorname{sgn} A} \sum_{j<k} (\langle A^{-1}D, D\rangle/2i\tau)^j u(0)/j! \Big| \le C_{k,A} \tau^{-\frac{n}{2}-k} \sum_{|\alpha| \le s+2k} \|D^\alpha u\|_2. \tag{2.4.5}$$

The statement remains true for $s = n/2$ if n is even. For a proof and extensions of most results in this section we refer to Hörmander [1, Sections 7.6 and 7.7].

We can now state the main result of this section. For the sake of simplicity we do not make minimal smoothness assumptions.

Theorem 2.4.5. *Let $K \subset \mathbf{R}^n$ be a compact set and let φ be a real valued C^∞ function in a neighborhood of K. If every critical point of φ in K is non-degenerate, then they form a finite set C_φ and if $u \in C_0^\infty(K)$ then*

(2.4.6)
$$\Big| \int u(x) e^{i\tau\varphi(x)}\, dx - \sum_{x \in C_\varphi} e^{i\tau\varphi(x) + \frac{\pi i}{4}\operatorname{sgn}\varphi''(x)} (2\pi/\tau)^{\frac{n}{2}} |\det \varphi''(x)|^{-\frac12} \sum_{j<k} \tau^{-j} L_j u(x) \Big|$$
$$\le C\tau^{-\frac{n}{2}-k} \sum_{|\alpha| \le \frac{n}{2}+1+2k} \sup |D^\alpha u|, \quad \tau > 0.$$

Here L_j is a differential operator of order $2j$ depending on φ, and $L_0 = 1$.

Proof. As already pointed out, non-degenerate critical points are isolated so C_φ is finite. We can choose a finite partition of unity $1 = \sum \chi_\nu$ in a neighborhood of K such that there is at most one critical point in the support of each χ_ν, and if there is one then the support is so small that Lemma 2.4.2 can be used there to change the coordinates so that φ becomes a quadratic form in the new variables. Then the theorem follows from Theorem 2.4.1 and Proposition 2.4.4.

As an example we shall study the Fourier transform of a smooth density on a hypersurface in $\mathbf{R}^{n+1}$ which has total curvature $\neq 0$. This application will be important in Chapter V.

THEOREM 2.4.6. *Let $K \subset \mathbf{R}^n$ be a compact set and let ψ be a real valued C^∞ function in a neighborhood X of K such that $\det \psi''(x) \neq 0$ when $x \in K$. If $a \in C_0^\infty(K)$, then the asymptotics of the Fourier transform of the density $a(x)\,dx$ on the hypersurface $\{(x, \psi(x)); x \in X\} \subset \mathbf{R}^{n+1}$,*

$$F(\xi, \xi_{n+1}) = \int e^{-i(\langle x,\xi\rangle + \psi(x)\xi_{n+1})} a(x)\, dx,$$

is given by

$$(2.4.7) \quad \left| |\xi_{n+1}/2\pi|^{\frac{n}{2}} F(\xi, \xi_{n+1}) - \sum_x a(x) |\det \psi''(x)|^{-\frac{1}{2}} e^{\mp \frac{\pi i}{4} \operatorname{sgn} \psi''(x)} e^{-i(\langle x,\xi\rangle + \psi(x)\xi_{n+1})} \right| \\ \leq C/(|\xi| + |\xi_{n+1}|),$$

where the sum is taken over the finitely many $x \in K$ such that $\xi + \xi_{n+1}\psi'(x) = 0$, and $\pm$ is the sign of ξ_{n+1}.

PROOF. Note that there are no such points if $|\xi_{n+1}| \sup_K |\psi'| < |\xi|/2$, say. In that case the estimate follows from Theorem 2.4.1. Otherwise the estimate is for fixed $\xi/\xi_{n+1} = \eta$ a consequence of Theorem 2.4.5 with $\tau = |\xi_{n+1}|$ and $\varphi(x) = \mp(\langle x, \eta\rangle + \psi(x))$. The proof of Theorem 2.4.5 shows that the estimate obtained will be uniform in η when $|\eta| \leq 2 \sup_K |\psi'|$, which completes the proof.

The terms in (2.4.7) have a clear geometrical meaning. The equation $\xi + \xi_{n+1}\psi'(x) = 0$ means that (ξ, ξ_{n+1}) is in the direction of the normal at $(x, \psi(x))$. The total curvature $\mathscr{K}$ of the hypersurface is there equal to $(\det \psi''(x))/(1 + |\psi'(x)|^2)^{(n+2)/2}$, and the surface measure is $dx\sqrt{1 + |\psi'(x)|^2}$. The absolute value of the main term in (2.4.7) multiplied by $(1 + |\xi|^2/\xi_{n+1}^2)^{n/4} = (1 + |\psi'(x)|^2)^{n/4}$ is therefore $|\mathscr{K}|^{-\frac{1}{2}}$ multiplied by the density divided by the surface measure. The number of curvatures at $(x, \psi(x))$ pointing in the direction (ξ, ξ_{n+1}) minus the number pointing in the opposite direction is $\pm \operatorname{sgn} \psi''(x)$.

Another important example is the solution of the initial value problem for the Schrödinger equation in $\mathbf{R}^{1+n}$,

$$\partial u(t, x)/\partial t = \tfrac{i}{2}\Delta_x u(t, x);\ u(0, \cdot) = f \in \mathscr{S}(\mathbf{R}^n).$$

Using Fourier transforms we obtain the solution

$$u(t, x) = (2\pi)^{-n} \int \hat{f}(\xi) e^{i(\langle x,\xi\rangle - \frac{1}{2}t|\xi|^2)}\, d\xi = e^{i|x|^2/2t} (2\pi)^{-n} \int \hat{f}(\xi + x/t) e^{-it|\xi|^2/2}\, d\xi.$$

Choose $\chi \in C_0^\infty(\mathbf{R}^n)$ equal to 1 in the unit ball. If a factor $\chi(\xi)$ is inserted in the integrand we can use Proposition 2.4.4, and if we insert a factor $1 - \chi(\xi)$ then the proof of Theorem 2.4.1$'$ gives that the integral is $O(t^{-\nu})$ for every ν. Hence

$$u(t, x) = e^{i|x|^2/2t - \pi i n/4} (2\pi t)^{-n/2} (\hat{f}(x/t) + O(1/t)),$$

which can be refined to a complete asymptotic series. This reflects the quantum mechanical interpretation of the dual variable ξ as the velocity of the "particle".

CHAPTER III

WAVELETS

3.1. Multiresolution analysis. The key to the discussion of the fast Fourier transform in Section 1.3 was that the Fourier transform of a function defined in $\mathbf{Z}_{2^N}$ was successively reduced to the Fourier transform of functions in the subspaces V_j, $0 \le j \le N$ of functions which are lifted from $\mathbf{Z}_{2^{N-j}}$, that is, only depend on the residue class modulo 2^{N-j}. If $f \in V_j$ then $x \mapsto f(2x)$ is in V_{j+1}. The situation is similar to the following notion of a multiresolution analysis but the fact that $x \mapsto 2x$ is not bijective on $\mathbf{Z}_{2^N}$ makes an essential difference. In particular, the spaces V_j will then increase with j:

DEFINITION 3.1.1. An orthonormal multiresolution analysis of $L^2(\mathbf{R}^n)$ is a sequence of closed subspaces V_j, $j \in \mathbf{Z}$, such that

(i) $V_j \subset V_{j+1}$ for all $j \in \mathbf{Z}$;[1]
(ii) $\bigcap_{-\infty}^{\infty} V_j = \{0\}$, and $\bigcup_{-\infty}^{\infty} V_j$ is dense in $L^2(\mathbf{R}^n)$;
(iii) $f \in V_j$ if and only if $\gamma f \in V_{j+1}$ where $(\gamma f)(x) = f(2x)$;
(iv) $f \in V_0$ implies $f(\cdot - k) \in V_0$ if $k \in \mathbf{Z}^n$;
(v) There is a function $\varphi \in V_0$ such that the functions $x \mapsto \varphi(x-k)$, $k \in \mathbf{Z}^n$, are an orthonormal basis for V_0.

Since $2^{n/2}\gamma$ is a unitary map in $L^2(\mathbf{R}^n)$, it follows from (iii), (iv) and (v) that

(iii)′ $f \in V_j$ if and only if $\gamma^{-j} f \in V_0$;
(iv)′ $f \in V_j$ implies $f(\cdot - k/2^j) \in V_j$ if $k \in \mathbf{Z}^n$;
(v)′ The functions $x \mapsto 2^{nj/2}\varphi(2^j x - k)$ with $k \in \mathbf{Z}^n$ are an orthonormal basis for V_j.

To clarify the meaning of these conditions we shall first discuss the most classical case, the *Haar basis* in one dimension. Let V_0 be the set of functions in $L^2(\mathbf{R})$ which are constant a.e. (almost everywhere) in every interval $\{x \in \mathbf{R}; k \le x \le k+1\}$ bounded by consecutive integers, and define V_j so that (iii)′ is valid. This means that V_j consists of the functions which are constant a.e. in the dyadic intervals

$$I_{j,k} = \{x \in \mathbf{R}; k \le 2^j x \le k+1\}, \quad k \in \mathbf{Z}. \tag{3.1.1}$$

The intervals are divided in half when j is increased by 1, so it is clear that V_j increases with j. Since step functions of compact support are dense in L^2 and we can choose the points of discontinuity as rational numbers with a power of 2 as denominator, the union

[1] We follow the notation of Meyer [2]. In Daubechies [1] there is a change of sign for the indices so that the spaces V_j decrease instead.

of the spaces V_j is dense in L^2. The intersection of all V_j consists of L^2 functions which are constant on the positive and on the negative half axis, hence equal to 0. Thus (ii) is fulfilled, and (v) is valid with φ equal to the characteristic function of $[0,1]$.

For this example the orthogonal complement $W_0 = V_1 \ominus V_0$ of V_0 in V_1 consists of functions which are constant in the intervals $I_{1,k}$ and have integral 0 over the intervals $I_{0,k}$. Hence an orthonormal basis is given by the functions $x \mapsto \psi(x-k)$, $k \in \mathbf{Z}$, where

$$\psi(x) = \begin{cases} 1, & \text{if } 0 \le x < \frac{1}{2}, \\ -1, & \text{if } \frac{1}{2} \le x < 1, \\ 0, & \text{if } x < 0 \text{ or } x \ge 1. \end{cases}$$

By condition (iii)$'$ it follows that the functions $x \mapsto 2^{j/2}\psi(2^j x - k)$, $k \in \mathbf{Z}$, are an orthonormal basis in $W_j = V_{j+1} \ominus V_j$. Since $L^2(\mathbf{R}) = \bigoplus_{-\infty}^{\infty} W_j$ by condition (ii), it follows that the functions

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k), \quad j \in \mathbf{Z},\ k \in \mathbf{Z},$$

form an orthonormal basis for $L^2(\mathbf{R})$. It is called the *Haar basis*, and ψ is called the *Haar wavelet.*

Actually this is slightly incorrect historically, for the Haar basis is a basis for $L^2(0,1)$. However, we can regard $L^2(0,1)$ as the subspace of $L^2(\mathbf{R})$ consisting of functions vanishing outside $(0,1)$. Then $(f, \psi_{j,k}) = 0$ unless $0 \le k < 2^j$. If $j < 0$ and $k = 0$ then

$$(f, \psi_{j,0}) = 2^{j/2} \int_0^1 f(x)\, dx = 2^{(j+1)/2}(f, \psi_{-1,0}).$$

Since $\sum_{-\infty}^{-1} 2^{j+1} = 2$ we obtain

$$f = \sum_{0 \le j, 0 \le k < 2^j} \psi_{j,k}(f, \psi_{j,k}) + 2\psi_{-1,0}(f, \psi_{-1,0}).$$

Thus the functions $\psi_{j,k}$ in the sum and $\sqrt{2}\psi_{-1,0} \equiv 1$ are an orthonormal basis for $L^2(0,1)$, which is the original Haar basis. (See Haar [1].)

After this motivating example we shall now return to Definition 3.1.1 and examine the consequences of the conditions there, beginning with (v).

PROPOSITION 3.1.2. *If $\varphi \in L^2(\mathbf{R}^n)$ then the functions $x \mapsto \varphi(x-k)$, $k \in \mathbf{Z}^n$, are an orthonormal system in $L^2(\mathbf{R}^n)$ if and only if*

(3.1.2)
$$\sum_{k \in \mathbf{Z}^n} |\hat{\varphi}(\xi + 2\pi k)|^2 = 1 \quad a.e..$$

PROOF. The orthonormality means that

(3.1.3)
$$\int_{\mathbf{R}^n} \varphi(x-k)\overline{\varphi(x)}\, dx = \delta_{k,0}, \quad k \in \mathbf{Z}^n.$$

The Fourier transform of $x \mapsto \varphi(x-k)$ is $\xi \mapsto e^{-i\langle k,\xi\rangle}\hat\varphi(\xi)$, so (3.1.3) can be rewritten as follows using Parseval's formula

$$(3.1.3)' \qquad (2\pi)^{-n}\int_{\mathbf{R}^n} |\hat\varphi(\xi)|^2 e^{-i\langle k,\xi\rangle}\,d\xi = \delta_{k,0}.$$

The left-hand side is a Fourier coefficient of the function

$$(3.1.4) \qquad \Phi(\xi) = \sum_{j\in\mathbf{Z}^n} |\hat\varphi(\xi+2\pi j)|^2,$$

which is $2\pi\mathbf{Z}^n$ periodic. Hence $(3.1.3)'$ means precisely that $\Phi(\xi) = 1$ as a distribution, that is, almost everywhere as a function.

Given a function $\varphi \in L^2(\mathbf{R}^n)$ satisfying (3.1.2) we can define V_0 as the closed linear hull of the orthonormal functions $\varphi(\cdot - k)$, $k \in \mathbf{Z}^n$, and then define V_j by condition (iii)$'$. The condition (i) will then be fulfilled if and only if $V_{-1} \subset V_0$, that is, $x \mapsto \varphi(x/2)$ is in V_0.

PROPOSITION 3.1.3. *Let φ satisfy* (3.1.2)*. Then $x \mapsto \varphi(x/2)$ is in the closed linear hull V_0 in $L^2(\mathbf{R}^n)$ of the functions $\varphi(\cdot - k)$, $k \in \mathbf{Z}^n$, if and only if*

$$(3.1.5) \qquad \hat\varphi(2\xi) = m_0(\xi)\hat\varphi(\xi),$$

where m_0 is a $2\pi\mathbf{Z}^n$ periodic function in L^∞ and

$$(3.1.6) \qquad \sum_{k\in\{0,1\}^n} |m_0(\xi+\pi k)|^2 = 1 \quad a.e..$$

PROOF. The scalar product α_k of $\varphi(x/2)$ with the orthonormal functions $\varphi(x-k)$ which span V_0 can be calculated by Parseval's formula,

$$(3.1.7) \qquad \alpha_k = \int_{\mathbf{R}^n} \varphi(x/2)\overline{\varphi(x-k)}\,dx = (2\pi)^{-n}\int_{\mathbf{R}^n} 2^n\hat\varphi(2\xi)\overline{\hat\varphi(\xi)}e^{i\langle k,\xi\rangle}\,d\xi, \quad k \in \mathbf{Z}^n.$$

If $x \mapsto \varphi(x/2)$ is in V_0 then

$$\varphi(x/2) = \sum_{k\in\mathbf{Z}^n} \alpha_k\varphi(x-k)$$

with convergence in L^2. By Parseval's formula this is equivalent to

$$2^n\hat\varphi(2\xi) = \sum_{k\in\mathbf{Z}^n} \alpha_k e^{-i\langle k,\xi\rangle}\hat\varphi(\xi),$$

also with L^2 convergence. Since $\sum|\alpha_k|^2 < \infty$ the Fourier series

$$m_0(\xi) = 2^{-n}\sum_{k\in\mathbf{Z}^n} \alpha_k e^{-i\langle k,\xi\rangle}$$

converges in $L^2(\mathbf{R}^n/2\pi\mathbf{Z}^n)$, which proves (3.1.5). Replacing ξ by $(\xi+2\pi k)/2$ we obtain

$$|\hat\varphi(\xi+2\pi k)|^2 = |m_0((\xi+2\pi k)/2)|^2|\hat\varphi((\xi+2\pi k)/2)|^2, \quad k\in\mathbf{Z}^n.$$

We sum over k using (3.1.2) and the fact that each residue class of $(2\mathbf{Z}^n)/\mathbf{Z}^n$ contains precisely one element in $\{0,1\}^n$, which proves (3.1.6).

Assuming now that (3.1.5) and (3.1.6) are fulfilled we shall prove that $x\mapsto\varphi(x/2)$ is in V_0. To do so we observe that (3.1.7) and (3.1.5) give in view of (3.1.2) and the periodicity of m_0

$$\alpha_k = (2\pi)^{-n}\int_{\mathbf{R}^n} 2^n m_0(\xi)|\hat\varphi(\xi)|^2 e^{i\langle k,\xi\rangle}\,d\xi = (2\pi)^{-n}\int_{\mathbf{R}^n/2\pi\mathbf{Z}^n} 2^n m_0(\xi)e^{i\langle k,\xi\rangle}\,d\xi,$$

so it follows from Parseval's formula (for Fourier series now) that

$$(2\pi)^n\sum_{k\in\mathbf{Z}^n}|\alpha_k|^2 = \int_{\mathbf{R}^n/2\pi\mathbf{Z}^n}|2^n m_0(\xi)|^2\,d\xi.$$

By (3.1.6) we have

$$2^n\int_{\mathbf{R}^n/2\pi\mathbf{Z}^n}|m_0(\xi)|^2\,d\xi = (2\pi)^n,$$

so it follows that

$$\sum_{k\in\mathbf{Z}^n}|\alpha_k|^2 = 2^n = \int_{\mathbf{R}^n}|\varphi(x/2)|^2\,dx,$$

which proves that $x\mapsto\varphi(x/2)$ is in V_0.

We have now seen that the equations (3.1.2), (3.1.5), (3.1.6) express the conditions (i), (iii), (iv), (v) completely when V_0 is defined by (v) and V_j is then defined by (iii)$'$. It remains to examine the condition (ii).

PROPOSITION 3.1.4. *Let φ satisfy* (3.1.2), *define V_j by* (v) *and* (iii)$'$, *and denote the orthogonal projection $L^2(\mathbf{R}^n)\to V_j$ by P_j. Then $P_j\to 0$ strongly as $j\to-\infty$, hence $\bigcap_{-\infty}^{\infty}V_j=\{0\}$. When $j\to\infty$ we have $P_j\to\mathrm{Id}$ strongly if and only if one of the following equivalent conditions is fulfilled:*

(1) *$|\hat\varphi(\varepsilon\xi)|^2\to 1$ in $\mathscr{D}'(\mathbf{R}^n)$ as $\varepsilon\to 0$;*
(2) *$|\hat\varphi(\varepsilon\xi)|^2\to 1$ in $L^1_{\mathrm{loc}}(\mathbf{R}^n)$ as $\varepsilon\to 0$;*
(3) *If we define $|\hat\varphi(0)|=1$ then 0 is a Lebesgue point for $|\hat\varphi|^2$.*

They imply that $\bigcup_{-\infty}^{\infty}V_j$ is dense in $L^2(\mathbf{R}^n)$, and the converse is true if (i) *is fulfilled.*

PROOF. To prove the first statement we must verify that $\|P_jf\|_{L^2}\to 0$ as $j\to-\infty$ for all f in a dense subset of $L^2(\mathbf{R}^n)$, say $f\in C_0(\mathbf{R}^n)$. We have

$$\|P_jf\|^2 = \sum_{k\in\mathbf{Z}^n}|\alpha_{jk}|^2, \quad \alpha_{jk} = \int_{\mathbf{R}^n} f(x)2^{nj/2}\overline{\varphi(2^jx-k)}\,dx.$$

If $|f| \le M$ and B is a ball containing $\operatorname{supp} f$, then

$$|\alpha_{jk}|^2 \le M^2 m(B) \int_B 2^{nj} |\varphi(2^j x - k)|^2 \, dx = M^2 m(B) \int_{y+k \in 2^j B} |\varphi(y)|^2 \, dy,$$

so we have

$$\|P_j f\|^2 \le M^2 m(B) \int_{E_j} |\varphi(y)|^2 \, dy, \quad E_j = \bigcup_{k \in \mathbf{Z}^n} (\{k\} + 2^j B).$$

The sets E_j decrease to the null set $\mathbf{Z}^n$ as $j \to -\infty$ so the integral converges to 0 then.

Since $\|f\|_{L^2}^2 = \|P_j f\|_{L^2}^2 + \|f - P_j f\|_{L^2}^2$, we have $P_j \to \mathrm{Id}$ strongly as $j \to \infty$ if and only if $\|P_j f\|_{L^2}^2 \to \|f\|_{L^2}^2$ for f in a dense subset of $L^2(\mathbf{R}^n)$. This time we choose f with $\hat f \in C_0^\infty(\mathbf{R}^n)$ and rewrite α_{jk} using Parseval's formula

$$\alpha_{jk} = (2\pi)^{-n} \int_{\mathbf{R}^n} \hat f(\xi) 2^{-nj/2} \overline{\hat\varphi(2^{-j}\xi)} e^{i\langle k,\xi\rangle/2^j} \, d\xi = (2\pi)^{-n} \int_{\mathbf{R}^n} \hat f(2^j \xi) 2^{nj/2} \overline{\hat\varphi(\xi)} e^{i\langle k,\xi\rangle} \, d\xi,$$

which can be viewed as the Fourier coefficients of the $2\pi\mathbf{Z}^n$ periodic function

$$\sum_{l \in \mathbf{Z}^n} \hat f(2^j(\xi + 2\pi l)) 2^{nj/2} \overline{\hat\varphi(\xi + 2\pi l)}.$$

If $\operatorname{supp} \hat f \subset \{\xi; |\xi| \le R\}$ then the terms in the sum have disjoint supports if $2\pi 2^j > 2R$. Then the square of the absolute value of the sum is the sum of the squares of the absolute values of the terms, and we obtain using Parseval's formula for Fourier series

$$\begin{aligned}(2\pi)^n \|P_j f\|_{L^2}^2 = (2\pi)^n \sum_k |\alpha_{jk}|^2 &= \int_{\mathbf{R}^n} |\hat f(2^j\xi)|^2 2^{nj} |\hat\varphi(\xi)|^2 \, d\xi \\ &= \int_{\mathbf{R}^n} |\hat f(\xi)|^2 |\hat\varphi(2^{-j}\xi)|^2 \, d\xi.\end{aligned}$$

This converges to $(2\pi)^n \|f\|_{L^2}^2 = \int_{\mathbf{R}^n} |\hat f(\xi)|^2 \, d\xi$ as $j \to \infty$ if and only if

$$\int_{\mathbf{R}^n} |\hat f(\xi)|^2 (1 - |\hat\varphi(2^{-j}\xi)|^2) \, d\xi \to 0, \quad \text{as } j \to \infty. \tag{3.1.8}$$

Since the integrand is non-negative by (3.1.2) and we can choose $\hat f \in C_0^\infty(\mathbf{R}^n)$ equal to 1 on any given compact set, this implies that $|\hat\varphi(\varepsilon\xi)|^2 \to 1$ in L^1_{loc} as $\varepsilon \to 0$, for we can always choose j so that $2^{-j-1} \le \varepsilon \le 2^{-j}$, thus $j \to \infty$ when $\varepsilon \to 0$. The condition (2) above is therefore necessary, and it implies (3) which implies (1). Since (1) implies (3.1.8) when $\hat f \in C_0^\infty$, the proof is complete.

Summing up, the conditions (3.1.2), (3.1.5), (3.1.6) and the very mild equivalent conditions in Proposition 3.1.4 are necessary and sufficient for the *scale function* (or "father wavelet") φ to generate a multiresolution analysis.

If we know the projection in V_j of a function $f \in L^2(\mathbf{R}^n)$, it is clear that we can calculate the projection in $V_{j-1} \subset V_j$, for it can be obtained by first projecting f to V_j. To derive a formula for this passage to a coarser resolution we assume to simplify notation that $j = 0$ and consider a finite sum

$$f_0(x) = \sum c_k \varphi(x-k).$$

To compute the projection in V_{-1} we must calculate the scalar products with the basis functions $2^{-n/2}\varphi(x/2 - l)$, $l \in \mathbf{Z}^n$, in V_{-1}. We have by Parseval's formula

$$\begin{aligned}\int_{\mathbf{R}^n} \varphi(x-k) 2^{-n/2}\overline{\varphi(x/2-l)}\, dx &= (2\pi)^{-n} \int_{\mathbf{R}^n} \hat{\varphi}(\xi) e^{-i\langle k,\xi\rangle} 2^{n/2}\overline{\hat{\varphi}(2\xi)} e^{i\langle 2l,\xi\rangle}\, d\xi \\ &= (2\pi)^{-n} \int_{\mathbf{R}^n} 2^{n/2}\overline{m_0(\xi)} |\hat{\varphi}(\xi)|^2 e^{i\langle 2l-k,\xi\rangle}\, d\xi,\end{aligned}$$

where we have used (3.1.5). In view of (3.1.2) this is equal to $2^{n/2}\overline{\mu(2l-k)}$ where $\mu(k)$ denote the Fourier coefficients of the $2\pi\mathbf{Z}^n$ periodic function m_0. Thus the projection f_{-1} of f_0 in V_{-1} is

$$f_{-1}(x) = \sum_{l \in \mathbf{Z}^n} (Tc)_l 2^{-n/2} \varphi(x/2 - l), \quad (Tc)_l = 2^{n/2} \sum_{k \in \mathbf{Z}^n} \overline{\mu(2l-k)} c_k. \tag{3.1.9}$$

We have now proved:

Proposition 3.1.5. *If m_0 is the function in* (3.1.5) *then the components c_k and c'_k, $k \in \mathbf{Z}^n$, of the projection of a function $f \in L^2(\mathbf{R}^n)$ in V_0 resp. V_{-1} are related by $c' = Tc$ where T is the contraction operator defined by* (3.1.9) *when only finitely many c_k are different from* 0. *Here $\mu(k)$ are the Fourier coefficients of the $2\pi\mathbf{Z}^n$ periodic function m_0 in* (3.1.5).

3.2. The wavelets associated with a multiresolution analysis. Assume given an orthonormal multiresolution analysis of $L^2(\mathbf{R}^n)$ (see Definition 3.1.1). As in the example of the Haar basis we want now to examine the quotient spaces $W_j = V_{j+1} \ominus V_j$. Note that

$$V_{j+k} = V_j \oplus W_j \oplus \cdots \oplus W_{j+k-1},\ 0 < k \in \mathbf{Z}; \quad V_j = \bigoplus_{k<j} W_k, \quad L^2(\mathbf{R}^n) = \bigoplus_{-\infty}^{\infty} W_k.$$

From condition (iii)$'$ it follows that $W_j = \gamma^j W_0$, and we shall now discuss the properties of W_0, which will lead to the desired wavelets.

Proposition 3.2.1. *A function $f \in L^2(\mathbf{R}^n)$ is in W_0 if and only if*

$$\hat{f}(\xi) = A(\xi/2)\hat{\varphi}(\xi/2), \tag{3.2.1}$$

where $A(\xi)$ is a $2\pi\mathbf{Z}^n$ periodic function in L^2_{loc} with

$$\sum_{k \in \{0,1\}^n} A(\xi + \pi k)\overline{m_0(\xi+\pi k)} = 0. \tag{3.2.2}$$

Here m_0 is the function in (3.1.5)*. We have*

$$\|f\|_{L^2}^2 = \pi^{-n} \int_{\mathbf{R}^n/2\pi\mathbf{Z}^n} |A(\xi)|^2 \, d\xi, \quad f \in W_0. \tag{3.2.3}$$

PROOF. That $f \in V_1$ means precisely that $x \mapsto f(x/2)$ is in V_0, and we know from the beginning of the proof of Proposition 3.1.3 that this is equivalent to

$$\hat{f}(2\xi) = A(\xi)\hat{\varphi}(\xi),$$

where A is a $2\pi\mathbf{Z}^n$ periodic function. This proves (3.2.1). By Parseval's formula and (3.1.2)

$$(2\pi)^n \|f\|_{L^2}^2 = \int_{\mathbf{R}^n} |\hat{f}(\xi)|^2 \, d\xi = 2^n \int_{\mathbf{R}^n} |A(\xi)\hat{\varphi}(\xi)|^2 \, d\xi = 2^n \int_{\mathbf{R}^n/2\pi\mathbf{Z}^n} |A(\xi)|^2,$$

which proves (3.2.3). That $f \in W_0$ means that f is orthogonal to $\varphi(\cdot - l)$ for every $l \in \mathbf{Z}^n$, that is,

$$\int_{\mathbf{R}^n} A(\xi/2)\hat{\varphi}(\xi/2)\overline{\hat{\varphi}(\xi)} e^{i\langle l,\xi\rangle} \, d\xi = \int_{\mathbf{R}^n} A(\xi/2)\overline{m_0(\xi/2)}|\hat{\varphi}(\xi/2)|^2 e^{i\langle l,\xi\rangle} \, d\xi = 0, \quad l \in \mathbf{Z}^n,$$

where we have used (3.1.5). Interpreting these integrals as Fourier coefficients of a $2\pi\mathbf{Z}^n$ periodic function we conclude that this is equivalent to

$$\sum_{k \in \mathbf{Z}^n} A(\xi/2 + \pi k)\overline{m_0(\xi/2 + \pi k)}|\hat{\varphi}(\xi/2 + \pi k)|^2 = 0, \quad \xi \in \mathbf{R}^n.$$

We replace ξ by 2ξ and note that every residue class in $\mathbf{Z}^n/2\mathbf{Z}^n$ contains precisely one element in $\{0,1\}^n$. This gives

$$\sum_{k \in \{0,1\}^n} \sum_{l \in \mathbf{Z}^n} A(\xi + \pi k + 2\pi l)\overline{m_0(\xi + \pi k + 2\pi l)}|\hat{\varphi}(\xi + \pi k + 2\pi l)|^2 = 0.$$

Since A and m_0 are $2\pi\mathbf{Z}^n$ periodic we can use (3.1.2) to calculate the sum over l and are left with the equation (3.2.2). The proof is complete.

Recall that by (3.1.6) the equation (3.2.2) means that the 2^n vector $(A(\xi + \pi k))_{k \in \{0,1\}^n}$ must be orthogonal to the corresponding unit vector in $\mathbf{C}^{2^n}$ defined by m_0. For every ξ the equation (3.2.2) has therefore $2^n - 1$ linearly independent solutions. Before discussing the higher dimensional case we shall examine the much more elementary one-dimensional case. Then the equation (3.2.2) reads

$$A(\xi)\overline{m_0(\xi)} + A(\xi + \pi)\overline{m_0(\xi + \pi)} = 0,$$

which means that

$$(A(\xi), A(\xi+\pi)) = \lambda(\xi)(\overline{m_0(\xi+\pi)}, -\overline{m_0(\xi)}) \tag{3.2.2$'$}$$

for some complex valued function λ with period 2π. Replacing ξ by $\xi+\pi$ in the second component of this equation we obtain the equivalent equations

$$A(\xi) = \lambda(\xi)\overline{m_0(\xi+\pi)} = -\lambda(\xi+\pi)\overline{m_0(\xi+\pi)}.$$

This requires that $\lambda(\xi) = -\lambda(\xi+\pi)$, for $m_0(\xi)$ and $m_0(\xi+\pi)$ are not both equal to 0. Conversely, this condition implies that $A(\xi) = \lambda(\xi)\overline{m_0(\xi+\pi)}$ is a 2π periodic solution of (3.2.2). A particular solution is given by

$$m_1(\xi) = e^{i\xi}\overline{m_0(\xi+\pi)}, \tag{3.2.4}$$

and it is normalized so that

$$|m_1(\xi)|^2 + |m_1(\xi+\pi)|^2 = 1. \tag{3.2.5}$$

The general solution of (3.2.2)$'$ is now of the form $A(\xi) = B(\xi)m_1(\xi)$ where B has period π. Thus (3.2.1) can be written

$$\hat{f}(\xi) = B(\xi/2)m_1(\xi/2)\hat{\varphi}(\xi/2),$$

where $B(\xi/2)$ has period 2π. The equation (3.2.3) takes the form

$$\|f\|_{L^2}^2 = \pi^{-1}\int_0^{2\pi} |B(\xi)|^2|m_1(\xi)|^2\,d\xi = \pi^{-1}\int_0^{\pi} |B(\xi)|^2\,d\xi$$

where we have used (3.2.5) and that B has period π. We have now proved the following basic theorem on wavelets in one dimension:

THEOREM 3.2.2. *Given an orthonormal multiresolution analysis of $L^2(\mathbf{R})$, let $\psi \in L^2(\mathbf{R})$ be the* wavelet *defined by*

$$\hat{\psi}(\xi) = e^{i\xi/2}\overline{m_0(\xi/2+\pi)}\hat{\varphi}(\xi/2) \tag{3.2.6}$$

with m_0 as in Proposition 3.1.3. Then the functions $x \mapsto \psi(x-k)$ with $k \in \mathbf{Z}$ form an orthonormal basis for $W_0 = V_1 \ominus V_0$; the functions $x \mapsto 2^{j/2}\psi(2^j x - k)$, $k \in \mathbf{Z}$, form an orthonormal basis for $W_j = V_{j+1} \ominus V_j$ when $j \in \mathbf{Z}$ is fixed, and they form an orthonormal basis in $L^2(\mathbf{R})$ when j is also allowed to vary in $\mathbf{Z}$.

In Section 3.3 we shall give a detailed discussion of wavelets in $L^2(\mathbf{R})$ with compact support, in particular their regularity properties, but we shall now return to the higher dimensional case to complete the discussion of the equation (3.2.2). It is inevitably a more difficult problem for higher dimensions since the solution is much less determined then. We could argue quite brutally by extending the unit vector $(m_0(\xi+\pi k))_{k\in\{0,1\}^n}$ to an orthonormal basis for $\mathbf{C}^{2^n}$ by the Gram-Schmidt procedure for every ξ with $0 \le \xi_j < \pi$, $j = 1, \dots, n$, and extend the definition of the other vectors so obtained when $0 \le \xi_j < 2\pi$ to a $2\pi\mathbf{Z}^n$ periodic function. However, even if m_0 is a smooth function this would introduce very bad singularities, so we shall proceed more gently.

The set $\{0,1\}^n$ in (3.2.2) is clearly identified with the group $G = \mathbf{Z}_2^n$, identified in turn with the subgroup $\pi\mathbf{Z}^n/2\pi\mathbf{Z}^n$ of the torus $\mathbf{R}^n/2\pi\mathbf{Z}^n$, and (3.2.2) is a summation over a coset with respect to this subgroup. To obtain an orthonormal basis for the solutions of (3.2.2) we must find $2\pi\mathbf{Z}^n$ periodic functions $m_r(\xi)$ also for $r \in G \setminus \{0\}$ so that for every $\xi \in \mathbf{R}^n$ the vectors $(m_r(\xi + \pi k))_{k\in G}$ with $r \in G$ is a complete orthonormal system in $\mathbf{C}^G$. Taking the Fourier transform in G, normalized to be unitary, we introduce the functions on $\widehat{G} \cong G$ defined by

$$\widehat{m}_{r,\varrho}(\xi) = 2^{-n/2} \sum_{k\in G} m_r(\xi + \pi k)(-1)^{\langle k,\varrho\rangle}, \quad r, \varrho \in G,$$

where we have used the explicit form of the characters on $\mathbf{Z}_2^n$ given in a remark after Theorem 1.2.1.[2] We have

$$\widehat{m}_{r,\varrho}(\xi + \pi l) = 2^{-n/2} \sum_{k\in G} m_r(\xi + \pi(k + l))(-1)^{\langle k,\varrho\rangle}$$
$$= (-1)^{\langle l,\varrho\rangle} 2^{-n/2} \sum_{k\in G} m_r(\xi + \pi k)(-1)^{\langle k,\varrho\rangle} = (-1)^{\langle l,\varrho\rangle} \widehat{m}_{r,\varrho}(\xi).$$

Hence

$$M_{r,\varrho}(\xi) = e^{-i\langle \varrho,\xi\rangle} \widehat{m}_{r,\varrho}(\xi) \tag{3.2.7}$$

are $\pi\mathbf{Z}^n$ periodic functions. The matrix $(m_r(\xi + \pi k))_{r,k\in G}$ is unitary if and only if the matrix $(\widehat{m}_{r,\varrho}(\xi))_{r,\varrho\in G}$ is unitary, for the normalized Fourier transform in G is unitary. This is also equivalent to unitarity of the matrix $(M_{r,\varrho}(\xi))_{r,\varrho\in G}$, for it differs only by a factor of absolute value 1 in each column.

Conversely, if $(M_{r,\varrho}(\xi))_{r,\varrho\in G}$ is a $\pi\mathbf{Z}^n$ periodic unitary matrix then (3.2.7) defines a $2\pi\mathbf{Z}^n$ periodic unitary matrix, and inverting the Fourier transform in G we obtain the unitary matrix

$$m_{r,k}(\xi) = 2^{-n/2} \sum_{\varrho\in G} \widehat{m}_{r,\varrho}(\xi)(-1)^{\langle k,\varrho\rangle}, \quad r, k \in G.$$

Now

$$m_{r,0}(\xi + \pi k) = 2^{-n/2} \sum_{\varrho\in G} \widehat{m}_{r,\varrho}(\xi + \pi k) = 2^{-n/2} \sum_{\varrho\in G} \widehat{m}_{r,\varrho}(\xi)(-1)^{\langle k,\varrho\rangle} = m_{r,k}(\xi),$$

so the $2\pi\mathbf{Z}^n$ periodic functions $m_{r,0}(\xi)$ has the desired properties if $m_{0,k}(\xi) = m_0(\xi + \pi k)$.

The problem has now been reduced to finding a unitary matrix $(M_{r,\varrho}(\xi))_{r,\varrho\in G}$ when one row $M_0(\xi) = (M_{0,\varrho}(\xi))_{\varrho\in G}$ of unit length is given. When $n = 1$ so that $\mathbf{Z}_2^n$ has only two

[2] Note that introducing these functions is a case of the decomposition in (1.2.6), (1.2.7).

elements we used that for $z = (z_1, z_2)$ in the unit sphere in $\mathbf{C}^2$ the matrix $\begin{pmatrix} z_1 & z_2 \\ \bar{z}_2 & -\bar{z}_1 \end{pmatrix}$ is unitary. This is very exceptional, for if $z = (z_1, \dots, z_N)$ is the first row of a unitary matrix $U(z)$ depending continuously on z when $z \in \mathbf{C}^N$ has length 1, then iz together with the other rows and their products by i would be an orthonormal basis for the tangent space of the $2N-1$ sphere which is not possible unless $N = 2$ or $N = 4$.[3] It is therefore not possible to give a universal formula for m_j like that in (3.2.6). However, we recall that the reason for the present discussion was that we wanted to preserve regularity properties of m_0, and when m_0 has some regularity the following simple lemma will show that there is no difficulty in making a construction adapted to m_0.

LEMMA 3.2.3. *If f is a map from a cube $I \subset \mathbf{R}^n$ to $\mathbf{R}^N$, where $N > n$, then the range of f has measure 0 if f is Hölder continuous of some order $\alpha \in (n/N, 1)$, that is,*

$$|f(x) - f(y)| \le C|x - y|^\alpha, \quad x, y \in I.$$

PROOF. Dividing each side of I into ν equal pieces we decompose I into ν^n cubes, with diameter $\le C/\nu$. The range of f restricted to such a cube is contained in a cube in $\mathbf{R}^N$ with measure $\le C'(\nu^{-\alpha})^N$. The outer measure of the range of f is therefore $\le C'\nu^{n-\alpha N} \to 0$ as $\nu \to \infty$.

If M_0 is Hölder continuous of order $> n/(2^{2n} - 1)$, it follows from Lemma 3.2.3 that the range of M_0 is not the whole unit sphere in $\mathbf{C}^{2^n}$. Then there is no difficulty in finding a unitary matrix extension, for we have:

LEMMA 3.2.4. *If q is a given point on the unit sphere $S = \{z \in \mathbf{C}^N; \sum_1^N |z_j|^2 = 1\}$ then there is an orthonormal basis $V_0(z), \dots, V_{N-1}(z)$ for $\mathbf{C}^N$ depending real analytically on $z \in S \setminus \{q\}$, such that $V_0(z) = (z_1, \dots, z_N)$.*

PROOF. If U is a unitary mapping in $\mathbf{C}^N$ and the vector fields V_j satisfy the conditions listed in the theorem, then the vector fields $z \mapsto UV_j(U^{-1}z)$ also satisfy them if q is replaced by Uq. It is therefore sufficient to prove the lemma for a special choice such as $q = (0, \dots, 0, -1)$. Then we note that the differential of the "stereographic projection"

$$S \setminus \{q\} \ni (z_1, \dots, z_N) \mapsto (z_1, \dots, z_{N-1})/(z_N + 1) \in \mathbf{C}^{N-1}$$

gives an analytic bijection of the complex tangent plane of S at z defined by $\sum_1^N \bar{z}_j dz_j = 0$ on $\mathbf{C}^{N-1}$. Thus the inverse images $v_1(z), \dots, v_{N-1}(z)$ of the basis vectors in $\mathbf{C}^{N-1}$ form a complex basis in the complex tangent plane of S at z. Together with the vector $v_0(z) = (z_1, \dots, z_N)$ they form a basis for $\mathbf{C}^N$, depending analytically on $z \in S \setminus \{q\}$. If we orthonormalize using the Gram-Schmidt procedure, starting with V_0, we obtain an orthonormal basis $V_0(z), \dots, V_{N-1}(z)$ with $V_0(z) = v_0(z) = (z_1, \dots, z_N)$ which also depends real analytically on $z \in S \setminus \{q\}$, which proves the lemma.

---

[3]It is probably known but not to me if in the case $N = 4$ there is such a unitary matrix; that the tangent space is parallelizable is a weaker property.

If q is a point in the unit sphere of $\mathbf{C}^G$ which is not in the range of M_0, we can now choose

$$(M_{r,\varrho}(\xi))_{\varrho\in G} = V_r(M_0(\xi)),$$

for $V_0(M_0(\xi)) = M_0(\xi)$ then. Since V_r is real analytic this will preserve all reasonable regularity properties of W_0. The passage from m_0 to M_0 and from $(M_{r,\varrho})$ back to $(m_{r,\varrho})$ does not affect the regularity either, so we have now proved:

THEOREM 3.2.5. *If $n > 1$ and m_0 is Hölder continuous of order $> n/(2^{2n} - 1)$, then there exist $2\pi\mathbf{Z}^n$ periodic functions $m_r(\xi)$ in $\mathbf{R}^n$, $r \in \{0,1\}^n \setminus \{0\}$ which are real analytic functions of $m_0(\xi + \pi k)$, $k \in \{0,1\}^n$, such that*

$$\sum_{k\in\{0,1\}^n} m_r(\xi + \pi k)\overline{m_s(\xi + \pi k)} = \delta_{rs}, \quad r, s \in \{0,1\}^n. \tag{3.2.8}$$

In any case one can find bounded measurable m_r with these properties.

For an arbitrary $2\pi\mathbf{Z}^n$ periodic solution of (3.2.2) it follows from (3.2.8) that there are uniquely determined coefficients $B_r(\xi)$, $r \in \{0,1\}^n \setminus \{0\}$ such that

$$A(\xi + \pi k) = \sum_{r\in\{0,1\}^n\setminus\{0\}} B_r(\xi)m_r(\xi + \pi k), \quad k \in \{0,1\}^n.$$

In view of the $2\pi\mathbf{Z}^n$ periodicity this is then true for all $k \in \mathbf{Z}^n$, and since the coefficients $B_r(\xi)$ are unique it follows that they are $\pi\mathbf{Z}^n$ periodic. (If the functions involved are not continuous all statements should be understood to hold a.e..) Now it follows from (3.2.1) that

$$\hat{f}(\xi) = \sum_{r\in\{0,1\}^n\setminus\{0\}} B_r(\xi/2)m_r(\xi/2)\hat{\varphi}(\xi/2).$$

The equation (3.2.3) takes the form

$$\begin{aligned}\|f\|_{L^2}^2 = \pi^{-n}\int_{\mathbf{R}^n/2\pi\mathbf{Z}^n} \Big| \sum_{r\in\{0,1\}^n\setminus\{0\}} B_r(\xi)m_r(\xi)\Big|^2 d\xi \\ = \pi^{-n}\int_{\mathbf{R}^n/\pi\mathbf{Z}^n} \sum_{k\in\{0,1\}^n} \Big| \sum_{r\in\{0,1\}^n\setminus\{0\}} B_r(\xi)m_r(\xi + \pi k)\Big|^2 d\xi \\ = \pi^{-n}\int_{\mathbf{R}^n/\pi\mathbf{Z}^n} \sum_{r\in\{0,1\}^n\setminus\{0\}} |B_r(\xi)|^2\, d\xi.\end{aligned}$$

Here we have used that $B_r(\xi)$ is $\pi\mathbf{Z}^n$ periodic and that the vectors $(m_r(\xi+\pi k))_{k\in\{0,1\}^n}$ are orthonormal. Hence we have proved an analogue of Theorem 3.2.2 in higher dimensions:

THEOREM 3.2.6. *Given an orthonormal multiresolution analysis of $L^2(\mathbf{R}^n)$, let m_r be defined as in Theorem 3.2.5 for $r \in \{0,1\}^n \setminus \{0\}$, and define $\psi_r \in L^2(\mathbf{R}^n)$ by*

$$\hat{\psi}_r(\xi) = m_r(\xi/2)\hat{\varphi}(\xi/2). \tag{3.2.9}$$

Then the functions $x \mapsto \psi_r(x-k)$ *with* $k \in \mathbf{Z}^n$ *and* $r \in \{0,1\}^n \setminus \{0\}$ *form an orthonormal basis for* $W_0 = V_1 \ominus V_0$*; the functions* $x \mapsto 2^{nj/2}\psi_r(2^j x - k)$ *with* $k \in \mathbf{Z}^n$ *and* $r \in \{0,1\}^n \setminus \{0\}$ *form an orthonormal basis for* $W_j = V_{j+1} \ominus V_j$ *when* $j \in \mathbf{Z}$ *is fixed, and they form an orthonormal basis in* $L^2(\mathbf{R}^n)$ *when* j *is also allowed to vary in* $\mathbf{Z}$.

Remark. Taking $r = 0$ we obtain in view of (3.1.5) that $\psi_0 = \varphi$. With this definition the functions $2^{nj/2}\psi_r(2^j x - k)$ with $k \in \mathbf{Z}^n$ and arbitrary $r \in \{0,1\}^n$ form an orthonormal basis for V_{j+1}. The wavelets ψ_r with $r \neq 0$ provide the additional information which occurs in a refinement of the analysis.

The appearance of $2^n - 1$ wavelets in Theorem 3.2.6 is very natural for a multiresolution analysis in $\mathbf{R}^n$ which is constructed from one in $\mathbf{R}$ by tensor products as follows. For any multiresolution analysis of $L^2(\mathbf{R})$, with subspaces V_j and scaling function φ, a multiresolution analysis of $L^2(\mathbf{R}^n)$ is given by

$$V_j^{(n)} = V_j \otimes \cdots \otimes V_j, \tag{3.2.10}$$

that is, the closed linear hull of products $u(x) = \prod_1^n u_\nu(x_\nu)$ where $u_\nu \in V_j$. Since the functions $x \mapsto 2^{j/2}\varphi(2^j x - k)$, $k \in \mathbf{Z}$, are an orthonormal basis for V_j, the functions

$$\mathbf{R}^n \ni x \mapsto 2^{nj/2} \prod_{\nu=1}^n \varphi(2^j x_\nu - k_\nu), \quad k = (k_1, \dots, k_\nu) \in \mathbf{Z}^n, \tag{3.2.11}$$

are an orthonormal basis for $V_j^{(n)}$. In other words, $\varphi^{(n)}(x) = \prod_1^n \varphi(x_\nu)$, $x = (x_1, \dots, x_n)$, is a scaling function for this multiresolution analysis.

Since $V_{j+1} = V_j \oplus W_j$ it follows that $V_{j+1}^{(n)}$ is the orthogonal direct sum of tensor products $W_{j,r}^{(n)}$, where $r = (r_1, \dots, r_n) \in \{0,1\}^n$ and $W_{j,r}^{(n)}$ is the tensor product obtained when V_j is replaced by W_j at the ν^{th} position in (3.2.10) when $r_\nu = 1$. For example,

$$V_{j+1}^{(2)} = (V_j \otimes V_j) \oplus (V_j \otimes W_j) \oplus (W_j \otimes V_j) \oplus (W_j \otimes W_j),$$

where the spaces correspond in order to r equal to $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$. We have always $W_{j,0}^{(n)} = V_j^{(n)}$. If ψ is a wavelet such that $x \mapsto 2^{j/2}\psi(2^j x - k)$, $k \in \mathbf{Z}$, is an orthonormal basis for W_j, then an orthonormal basis for $W_{j,r}^{(n)}$ is given by $x \mapsto 2^{nj/2}\psi_r(2^j x - k)$, $k \in \mathbf{Z}^n$, where

$$\psi_r(x) = \prod_{r_\nu = 0} \varphi(x_\nu) \prod_{r_\nu = 1} \psi(x_\nu). \tag{3.2.12}$$

This gives very explicit wavelets with the properties in Theorem 3.2.6. Combined with the results of Section 3.3 we obtain multidimensional wavelets of compact support with as many derivatives and vanishing moments as we wish.

3.3. Compactly supported orthogonal wavelets in one dimension. For applications of wavelets a number of properties of the scaling function φ and the wavelets ψ are desirable such as fast decrease, vanishing moments for ψ and smoothness. Priorities depend on the applications. We shall here confine ourselves to a study of wavelets of *compact support* in one dimension. The interest of compactness is of course that it makes the coefficients in the wavelet expansion locally determined and the refinement algorithm (3.1.9) finite. The Haar basis already has this property, but it is not even continuous.

At first we assume that we are given a *multiresolution analysis of* $L^2(\mathbf{R})$ *with scaling function* φ *of compact support.* Using the results proved in Sections 3.1 and 3.2 we can immediately draw some useful conclusions:

1. The Fourier transform $\hat\varphi$ can be extended to an entire analytic function, and $|\hat\varphi(0)| = 1$ by Proposition 3.1.4. We can multiply φ by a constant of absolute value 1 so it is no restriction to assume in what follows that $\hat\varphi(0) = 1$.

2. By the proof of Proposition 3.1.3 the function m_0 in (3.1.5) is given by

$$m_0(\xi) = \tfrac{1}{2}\sum_{k\in\mathbf{Z}} \alpha_k e^{-ik\xi}, \quad \alpha_k = \int \varphi(x/2)\overline{\varphi(x-k)}\,dx,$$

and if $a \le x \le b$ when $x \in \operatorname{supp}\varphi$, then $\alpha_k = 0$ unless the intersection $[2a, 2b] \cap [a+k, b+k]$ has interior points, that is, $2a - b < k < 2b - a$. Hence m_0 is a trigonometric polynomial, $m_0(0) = 1$ by (3.1.5) so $m_0(\pi) = 0$ by (3.1.6), and $|m_0(\xi)|^2$ is a trigonometric polynomial of degree $< 3b - 3a$. If φ is real valued, then α_k are real and $|m_0(\xi)|^2 = m_0(\xi)m_0(-\xi)$ is even.

3. The wavelet ψ has also compact support; in fact, $\psi(x/2)$ is a finite linear combination of integer translates of φ in view of (3.2.6), which also gives $\hat\psi(0) = 0$, that is, $\int_{\mathbf{R}} \psi(x)\,dx = 0$. This statement is strengthened by regularity properties of ψ:

Proposition 3.3.1. *If* $\psi \in C_0^\mu(\mathbf{R})$ *where* μ *is an integer* > 0, *then*

$$\int \psi(x)x^\nu\,dx = 0, \quad \textit{that is,} \quad \hat\psi^{(\nu)}(0) = 0, \quad \textit{if} \ \ 0 \le \nu \le \mu, \tag{3.3.1}$$

which implies that m_0 *has a zero of order* $\mu + 1$ *at* π.

Proof. For arbitrary integers k and $j > 0$ we have

$$0 = 2^j \int_{\mathbf{R}} \psi(x)\overline{\psi(2^j x - k)}\,dx = \int_{\mathbf{R}} \psi(2^{-j}y + 2^{-j}k)\overline{\psi(y)}\,dy.$$

By Taylor's formula

$$\psi(2^{-j}y + 2^{-j}k) = \sum_{\nu\le\mu} \psi^{(\nu)}(2^{-j}k)(2^{-j}y)^\nu/\nu! + o(2^{-j\mu}),$$

uniformly when $y \in \operatorname{supp}\psi$. Hence

$$\sum_{\nu\le\mu} \psi^{(\nu)}(2^{-j}k)2^{-j\nu}\int_{\mathbf{R}} y^\nu\overline{\psi(y)}\,dy/\nu! = o(2^{-j\mu})$$

as $j \to +\infty$. Choose a point x_0 with $\psi(x_0) \neq 0$ and then a sequence $k_j \in \mathbf{Z}$ such that $2^{-j}k_j \to x_0$. Then it follows that $\psi(x_0)\int \overline{\psi(y)}\,dy = 0$, which of course we already knew without any regularity assumption. Assume that we have already proved (3.3.1) when $0 \le \nu < \sigma$ where $0 < \sigma \le \mu$. If we then choose x_0 so that $\psi^{(\sigma)}(x_0) \neq 0$, it follows in the same way that $\int x^\sigma \overline{\psi(y)}\,dy = 0$, which proves (3.3.1) inductively.

COROLLARY 3.3.2. *If the scaling function is compactly supported then it cannot be in* C^∞.

PROOF. If $\varphi \in C_0^\infty$ then $\psi \in C_0^\infty$ and the analytic function $\hat\psi$ has a zero of infinite order at the origin by Proposition 3.3.1. This implies $\hat\psi = 0$ which is a contradiction.

REMARK. It is easy to see that (3.3.1) follows from the somewhat weaker assumption that $\psi \in C^{\mu-1}$ and that $\psi^{(\mu-1)}$ is Lipschitz continuous. We leave the proof as an exercise.

Let N be the order of the zero of m_0 at π; by Proposition 3.3.1 and (3.2.6) we know that $N \ge \mu + 1$ if $\varphi \in C^\mu$, and even without any regularity assumptions we know that $N \ge 1$. Then it follows that

$$m_0(\xi)/(1+e^{-i\xi})^N$$

is a trigonometric polynomial. In fact, $m_0(\xi)$ can be written as a power of $e^{i\xi}$ times a polynomial in $e^{-i\xi}$ with a zero of order N at -1, so it is divisible by $(1+e^{-i\xi})^N$ as a polynomial in $e^{-i\xi}$. From now on we *assume that* φ *is real valued.* Then the Fourier coefficients α_k of m_0 are real, so $m_0(-\xi) = \overline{m_0(\xi)}$ and it follows that

$$M_0(\xi) = |m_0(\xi)|^2 = (\cos^2(\tfrac12\xi))^N L(\xi), \tag{3.3.2}$$

where L is a polynomial in $\cos\xi = 2\cos^2(\frac12\xi) - 1 = 1 - 2\sin^2(\frac12\xi)$, so we can write

$$M_0(\xi) = (1-y)^N P(y), \quad y = \sin^2(\tfrac12\xi). \tag{3.3.3}$$

Here P is a polynomial, of degree $< 3b - 3a - N$ if $\operatorname{supp}\varphi \subset [a,b]$. The condition (3.1.6) can be written

$$(1-y)^N P(y) + y^N P(1-y) = 1. \tag{3.3.4}$$

LEMMA 3.3.3. *There is a unique polynomial* P_N *of degree* $< N$ *satisfying* (3.3.4)*, and it is given by*

$$P_N(y) = \sum_{k=0}^{N-1} \binom{N-1+k}{k} y^k. \tag{3.3.5}$$

Every solution can be written $P(y) = P_N(y) + y^N R(\frac12 - y)$ *where* R *is an odd polynomial.*

PROOF. Since the polynomials y^N and $(1-y)^N$ have no common factor, the Euclidean algorithm gives that there exist polynomials Q_1 and Q_2 such that

$$(1-y)^N Q_1(y) + y^N Q_2(y) = 1.$$

We can write $Q_1(y) = q_1(y) + y^N r(y)$ with uniquely determined polynomials q_1 and r such that $q_1(y)$ is of degree $< N$. Writing $q_2(y) = Q_2(y) + (1-y)^N r(y)$ we obtain

$$(1-y)^N q_1(y) + y^N q_2(y) = 1,$$

which proves that q_2 is also of degree $< N$. There are no other polynomials of degree $< N$ with this property. Replacing y by $1-y$ gives

$$y^N q_1(1-y) + (1-y)^N q_2(1-y) = 1,$$

and we conclude that $q_1(y) = q_2(1-y)$ so $P_N(y) = q_1(y)$ is a solution of (3.3.4) of degree $< N$, and the only one. Since (3.3.4) implies

$$P_N(y) = (1-y)^{-N} + O(y^N), \quad \text{as } y \to 0,$$

it follows that P_N is the $(N-1)^{\text{st}}$ partial sum of the Taylor expansion of $(1-y)^{-N}$ which gives (3.3.5). The general solution can be written $P(y) = P_N(y) + y^N r(y)$ where $r(y) + r(1-y) = 0$, which means that $r(y) = R(y - \frac{1}{2})$ with an odd polynomial R. The proof is complete.

The polynomial P_N is obviously ≥ 1 in $[0,1]$, but for the general solution of (3.3.4) non-negativity in $[0,1]$ is a restriction which implies that R for a given degree μ must belong to a a convex compact neighborhood of the origin in the space of odd polynomials of degree μ.

To return to the function m_0 and ultimately to the scaling function φ we must first find a trigonometric polynomial with absolute value squared equal to $P(\sin^2(\frac{1}{2}\xi))$.

Lemma 3.3.4. *If P is a polynomial of degree μ which is non-negative in $[0,1]$, then there is a polynomial B of the same degree with real coefficients such that*

$$P(\sin^2(\tfrac{1}{2}\xi)) = |B(e^{i\xi})|^2. \tag{3.3.6}$$

Proof. Since $\sin^2(\frac{1}{2}\xi) = \frac{1}{2}(1-\cos\xi)$ we can write $P(\sin^2(\frac{1}{2}\xi)) = Q(\cos\xi)$ where Q is a polynomial of degree μ which is non-negative in $[-1,1]$. We can factor Q as a product of polynomials of the form $Q(x) = (x+\lambda)\operatorname{sgn}\lambda$ with $\lambda \in \mathbf{R}$ and $|\lambda| \geq 1$ or $(x-\zeta)(x-\bar\zeta)$ with $\zeta \in \mathbf{C}$, so it suffices to verify the lemma for these two cases.

a) In the first case we write

$$\begin{aligned}(\cos\xi+\lambda)\operatorname{sgn}\lambda &= \tfrac{1}{2}(e^{i\xi}+e^{-i\xi}+2\lambda)\operatorname{sgn}\lambda = \tfrac{1}{2}e^{-i\xi}(e^{2i\xi}+2\lambda e^{i\xi}+1)\operatorname{sgn}\lambda \\ &= \tfrac{1}{2}e^{-i\xi}(e^{i\xi}+a)(e^{i\xi}+1/a)\operatorname{sgn}\lambda = \tfrac{1}{2}(e^{i\xi}+a)(e^{-i\xi}+a)/|a|\end{aligned}$$

where $a = \lambda + \sqrt{\lambda^2-1}$ is real and has the same sign as λ. We can therefore take $B(x) = (x+a)/\sqrt{2|a|}$.

b) In the second case we write

$$\begin{aligned}(\cos\xi-\zeta)(\cos\xi-\bar\zeta) &= \tfrac{1}{4}e^{-2i\xi}(e^{2i\xi}-2\zeta e^{i\xi}+1)(e^{2i\xi}-2\bar\zeta e^{i\xi}+1) \\ &= \tfrac{1}{4}e^{-2i\xi}(e^{i\xi}-\zeta_1)(e^{i\xi}-\zeta_1^{-1})(e^{i\xi}-\bar\zeta_1)(e^{i\xi}-\bar\zeta_1^{-1}) \\ &= \tfrac{1}{4}(e^{i\xi}-\zeta_1)(e^{i\xi}-\bar\zeta_1)(e^{-i\xi}-\bar\zeta_1)(e^{-i\xi}-\zeta_1)/|\zeta_1|^2\end{aligned}$$

so we can take $B(x) = \frac{1}{2}(x-\zeta_1)(x-\bar{\zeta}_1)/|\zeta_1|$.

In the second case the factorisation is not unique: we can replace ζ_1 by $1/\zeta_1$, so each of the polynomials given by Lemma 3.3.3 which is non-negative in $[0,1]$ yields a finite number of candidates for the function m_0, in addition to a factor $e^{ik\xi}$ with $k \in \mathbf{Z}$. For each choice it remains to see if there is a corresponding scale function satisfying (3.1.5) and the other conditions established in Section 3.1. (Since (3.1.5) does not change if both $m_0(\xi)$ and $\hat{\varphi}(\xi)$ are multiplied by $e^{ik\xi}$, multiplication of $m_0(\xi)$ by $e^{ik\xi}$ just implies that $\hat{\varphi}(\xi)$ is multiplied by $e^{ik\xi}$, which means an integer translation of φ.) From (3.1.5) and the conditions $\hat{\varphi}(0) = m_0(0) = 1$ it follows that once m_0 has been chosen we must necessarily have $\hat{\varphi} = \Phi$ where

$$\Phi(\xi) = \prod_1^\infty m_0(\xi/2^k), \quad \xi \in \mathbf{R}. \tag{3.3.7}$$

The product is convergent even for all $\xi \in \mathbf{C}$ since $m_0(\xi/2^k) = 1 + O(2^{-k})$ for ξ in a bounded subset of $\mathbf{C}$, so Φ extends to an entire analytic function. Since

$$m_0(\xi) = \tfrac{1}{2} \sum_{k_- \le k \le k_+} \alpha_k e^{-ik\xi}$$

and $|m_0(\xi)| \le 1$ when $\xi \in \mathbf{R}$ (by the condition (3.1.6) which is fulfilled by our choice of m_0) it follows from the maximum principle applied to $\frac{1}{2}\sum \alpha_k z^{k-k_\pm}$ for $|z| > 1$ and $|z| < 1$ respectively that

$$|m_0(\zeta)| \le e^{k_\pm \operatorname{Im}\zeta}, \quad \zeta \in \mathbf{C},\ \pm\operatorname{Im}\zeta \ge 0,$$

which implies that

$$|\Phi(\zeta)| \le e^{k_\pm \operatorname{Im}\zeta}, \quad \zeta \in \mathbf{C},\ \pm\operatorname{Im}\zeta \ge 0.$$

By the Paley-Wiener-Schwartz theorem it follows that Φ is in fact the Fourier-Laplace transform of a distribution $\varphi \in \mathscr{E}'([k_-, k_+])$. (If $\alpha_{k_\pm} \ne 0$ it follows from the theorem of supports that $[k_-, k_+]$ is the smallest interval containing $\operatorname{supp}\varphi$.) What remains is to decide if φ satisfies (3.1.2) and to determine the regularity properties of φ.

To prove that $\Phi \in L^2(\mathbf{R})$ we denote the partial products (3.3.7) by Φ_k and note that

$$\begin{aligned}\int_{-2^k\pi}^{2^k\pi} |\Phi_k(\xi)|^2\,d\xi &= \int_0^{2^{k+1}\pi} |\Phi_k(\xi)|^2\,d\xi \\ &= \int_0^{2^k\pi} |\Phi_{k-1}(\xi)|^2(|m_0(2^{-k}\xi)|^2 + |m_0(2^{-k}\xi+\pi)|^2)\,d\xi \\ &= \int_0^{2^k\pi} |\Phi_{k-1}(\xi)|^2\,d\xi = \int_{-2^{k-1}\pi}^{2^{k-1}\pi} |\Phi_{k-1}(\xi)|^2\,d\xi = \dots = \int_{-\pi}^{\pi} |\Phi_0(\xi)|^2\,d\xi,\end{aligned}$$

where we have first used the periodicity of m_0 and then (3.1.6). By Fatou's lemma it follows that $\Phi \in L^2$ and that

$$\int_{\mathbf{R}} |\Phi(\xi)|^2\,d\xi \le \int_{-\pi}^{\pi} |\Phi_0(\xi)|^2\,d\xi = 2\pi,$$

for $\Phi_0 = 1$. (Since the convergence is locally uniform this is quite elementary; one just has to consider the integral over a compact interval first.) Note that (3.1.2) implies that $\|\hat{\varphi}\|_{L^2}^2 = 2\pi$. However, (3.1.2) is not always fulfilled.

Example. If

$$m_0(\xi) = \tfrac{1}{2}(1+e^{-i\xi})(1-e^{-i\xi}+e^{-2i\xi}) = \tfrac{1}{2}(1+e^{-3i\xi}) = e^{-3i\xi/2}\cos(\tfrac{3}{2}\xi)$$

then (3.1.6) is fulfilled and

$$\hat\varphi(\xi) = e^{-3i\xi/2}\sin(\tfrac{3}{2}\xi)/(\tfrac{3}{2}\xi) = -i(1-e^{-3i\xi})/3\xi$$

for the right-hand side equals 1 at 0 and (3.1.5) is satisfied since $(1-e^{-3i\xi})(1+e^{-3i\xi}) = 1-e^{-6i\xi}$. Hence we have $\varphi(x) = \frac{1}{3}$ in $[0,3]$ and $\varphi(x) = 0$ elsewhere. This resembles the Haar system but $\varphi(\cdot - k)$, $k \in \mathbf{Z}$, is not an orthonormal system because of its redundancy.

Before we tackle the problem to decide when (3.1.2) is valid, we note that in any case the corresponding wavelets will always be complete in a very strong sense:

Theorem 3.3.5. *With a trigonometric polynomial m_0 and a corresponding $\varphi \in L^2$ defined as above so that* (3.1.6) *and* (3.1.5) *are valid, with ψ defined by* (3.2.6) *and*

$$\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k), \quad \psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k),$$

we have for $f \in L^2(\mathbf{R})$

$$\sum_k |(f,\varphi_{\nu,k})|^2 + \sum_{\nu\le j\le\mu}\sum_k |(f,\psi_{j,k})|^2 = \sum_k |(f,\varphi_{\mu+1,k})|^2, \tag{3.3.8}$$

$$\sum_{j,k}|(f,\psi_{j,k})|^2 = \|f\|_{L^2}^2. \tag{3.3.9}$$

Moreover, $\|\psi\|_{L^2} = \|\varphi\|_{L^2} \le 1$, and

$$g(\xi) = \sum_{l\in\mathbf{Z}} |\hat\varphi(\xi + 2\pi l)|^2 \tag{3.3.10}$$

is a trigonometric polynomial such that

$$g(\xi) = |m_0(\tfrac{1}{2}\xi)|^2 g(\tfrac{1}{2}\xi) + |m_0(\tfrac{1}{2}\xi+\pi)|^2 g(\tfrac{1}{2}\xi+\pi). \tag{3.3.11}$$

Proof. It is sufficient to prove (3.3.8) when $\mu = \nu = 0$. Then

$$(\varphi_{0,n}, f) = \frac{1}{2\pi}\int \hat\varphi(\xi)\overline{\hat f(\xi)}e^{-in\xi}\,d\xi, \quad (\psi_{0,n}, f) = \frac{1}{2\pi}\int \hat\psi(\xi)\overline{\hat f(\xi)}e^{-in\xi}\,d\xi,$$

can be considered as Fourier coefficients of 2π periodic functions. Hence the left-hand side of (3.3.8) with $\nu = \mu = 0$ is equal to $\int_0^{2\pi}(|A(\xi)|^2 + |B(\xi)|^2)\,d\xi/2\pi$ where

$$A(\xi) = \sum_{l\in\mathbf{Z}} m_0(\tfrac{1}{2}\xi+\pi l)\hat\varphi(\tfrac{1}{2}\xi+\pi l)\overline{\hat f(\xi+2\pi l)},$$
$$B(\xi) = e^{i\xi/2}\sum_{l\in\mathbf{Z}}(-1)^l\overline{m_0(\tfrac{1}{2}\xi+\pi(l+1))}\hat\varphi(\tfrac{1}{2}\xi+\pi l)\overline{\hat f(\xi+2\pi l)},$$

where we have used (3.1.5) and (3.2.6) respectively. With the notation

$$C(\xi)=\sum_{l\in\mathbf{Z}}\hat\varphi(\tfrac12\xi+2\pi l)\overline{\hat f(\xi+4\pi l)}$$

we have since m_0 has period 2π

$$A(\xi)=m_0(\tfrac12\xi)C(\xi)+m_0(\tfrac12\xi+\pi)C(\xi+2\pi),$$
$$B(\xi)=e^{i\xi/2}(\overline{m_0(\tfrac12\xi+\pi)}C(\xi)-\overline{m_0(\tfrac12\xi)}C(\xi+2\pi)).$$

Hence it follows from (3.1.6) that

$$|A(\xi)|^2+|B(\xi)|^2=|C(\xi)|^2+|C(\xi+2\pi)|^2,$$

which gives

$$\frac{1}{2\pi}\int_0^{2\pi}(|A(\xi)|^2+|B(\xi)|^2)\,d\xi=\frac{1}{2\pi}\int_0^{4\pi}|C(\xi)|^2\,d\xi.$$

Now

$$(\varphi_{1,k},f)=\frac{1}{4\pi}\int\sqrt2\hat\varphi(\tfrac12\xi)\overline{\hat f(\xi)}e^{-ik\xi/2}\,d\xi$$

can be considered as the Fourier coefficients of the 4π periodic function $\sqrt2 C(\xi)$, so the right-hand side of (3.3.8) is equal to $\int_0^{4\pi}|C(\xi)|^2/2\pi$, which completes the proof of (3.3.8).

With the notation of the proof of Proposition 3.1.4 the right-hand side of (3.3.8) is $\sum_k|\alpha_{\mu+1,k}|^2$, and since $\hat\varphi(0)=1$ we proved then without using (3.1.2) that it converges to $\|f\|^2_{L^2}$ as $\mu\to\infty$ if $\hat f\in C_0^\infty(\mathbf{R})$. Hence it follows from (3.3.8) that $\sum_k|(f,\varphi_{\nu,k})|^2\le\|f\|^2_{L^2}$ for every ν and for every f in this dense subset of L^2. Restricting first to finite sums we conclude that this inequality is true for every $f\in L^2$ and every ν. Hence it follows for every $f\in L^2$ that the right-hand side of (3.3.8) converges to $\|f\|^2$ as $\mu\to+\infty$ and that the first sum on the left converges to 0 as $\nu\to-\infty$, for each of these statements is true in a dense subset of L^2 by the proof of Proposition 3.1.4. This proves (3.3.9).

We have already proved that $\|\varphi\|_{L^2}\le1$. Using (3.1.5), (3.1.6) and (3.2.6) we obtain

$$\int_{\mathbf{R}}(|\hat\varphi(\xi)|^2+|\hat\psi(\xi)|^2)\,d\xi=\int_{\mathbf{R}}(|m_0(\tfrac12\xi)|^2+|m_0(\tfrac12\xi+\pi)|^2)|\hat\varphi(\tfrac12\xi)|^2\,d\xi$$
$$=\int_{\mathbf{R}}|\hat\varphi(\tfrac12\xi)|^2\,d\xi=2\int_{\mathbf{R}}|\hat\varphi(\xi)|^2\,d\xi,$$

which proves that $\|\varphi\|_{L^2}=\|\psi\|_{L^2}$. Since the Fourier coefficients of g are the scalar products $(\varphi(\cdot-n),\varphi)$ it is clear that g is almost everywhere equal to a trigonometric polynomial. Now it follows from the arguments used to prove (2.1.25) that the series (3.3.10) is locally uniformly convergent, so g is continuous and everywhere equal to a trigonometric polynomial. When (3.3.10) is entered, the right-hand side of (3.3.11) becomes

$$|m_0(\tfrac12\xi)|^2\sum_{l\in\mathbf{Z}}|\hat\varphi(\tfrac12\xi+2\pi l)|^2+|m_0(\tfrac12\xi+\pi)|^2\sum_{l\in\mathbf{Z}}|\hat\varphi(\tfrac12\xi+2\pi l+\pi)|^2$$
$$=\sum_{l\in\mathbf{Z}}|m_0(\tfrac12\xi+\pi l)|^2|\hat\varphi(\tfrac12\xi+\pi l)|^2=\sum_{l\in\mathbf{Z}}|\hat\varphi(\xi+2\pi l)|^2,$$

by (3.1.5), and this is equal to the left-hand side. The proof is complete.

Many equivalent necessary and sufficient conditions for the validity of (3.1.2) are given by the following theorem:

THEOREM 3.3.6. *With a trigonometric polynomial m_0 and φ defined as above the following conditions are equivalent:*

(i) *The functions $x \mapsto \varphi(x-k)$ with $k \in \mathbf{Z}$ are orthonormal in $L^2(\mathbf{R})$;*
(ii) $\|\varphi\|_{L^2} = 1$;
(iii) $\Phi_k \chi(\cdot/2^k) \to \hat{\varphi}$ *in $L^2(\mathbf{R})$ as $k \to \infty$ if χ is the characteristic function of $(-\pi, \pi)$ and Φ_k are the partial products in* (3.3.7);
(iv) $\sum_{l\in\mathbf{Z}} |\hat{\varphi}(\xi + 2\pi l)|^2 = 1$ *for every $\xi \in \mathbf{R}$;*
(v) $\sum_{l\in\mathbf{Z}} |\hat{\varphi}(\xi + 2\pi l)|^2 > 0$ *for every $\xi \in \mathbf{R}$;*
(vi) *For every $\xi \in \mathbf{R}$ there is some $l \in \mathbf{Z}$ such that $\hat{\varphi}(\xi + 2\pi l) \neq 0$;*
(vii) *The projection of $\{\xi \in \mathbf{R}; \hat{\varphi}(\xi) \neq 0\}$ in $\mathbf{R}/2\pi\mathbf{Z}$ is surjective;*
(viii) *There is a function $\tilde{\chi} \in C_0^\infty(\mathbf{R})$ such that $\hat{\varphi}(\xi) \neq 0$ when $\xi \in \operatorname{supp}\tilde{\chi}$, $\tilde{\chi} = 1$ in a neighborhood of the origin, and $\sum_{l\in\mathbf{Z}} \tilde{\chi}(\xi + 2\pi l) \equiv 1$.*
(ix) *Every trigonometric polynomial with period 2π satisfying* (3.3.11) *is a constant.*
(x) *There is no trigonometrical polynomial g satisfying* (3.3.11) *with period 2π and* $\min g = 0$, $g(0) > 0$.
(xi) *There is no non-trival cycle in $\{\dot{\xi} \in \mathbf{R}/2\pi\mathbf{Z}; |m_0(\dot{\xi})| = 1\}$ for the doubling map $\dot{\xi} \mapsto 2\dot{\xi}$.*

PROOF. (i) $\implies$ (ii) is trivial. We proved above that the L^2 norm of $\Phi_k\chi(\cdot/2^k)$ is equal to $\sqrt{2\pi}$, and this sequence converges locally uniformly, hence weakly, to $\hat{\varphi}$. Norm convergence is then equivalent to convergence of the norms which proves the equivalence of (ii) and (iii). From Proposition 3.1.2 and its proof we recall that (i) is equivalent to (iv) and to the conditions on the Fourier coefficients

$$\int |\hat{\varphi}(\xi)|^2 e^{-in\xi}\, d\xi = 2\pi\delta_{n,0}, \quad n \in \mathbf{Z}. \tag{3.3.12}$$

To prove that (3.3.12) follows from (iii) we note that $\Phi_k(\xi) = m_0(\xi/2^k)\Phi_{k-1}(\xi)$ and use an argument similar to the proof that $\Phi \in L^2$,

$$\begin{aligned}\int |\Phi_k(\xi)|^2\chi(\xi/2^k)e^{-in\xi}\, d\xi &= 2^k \int |\Phi_k(2^k\xi)|^2\chi(\xi)e^{-in2^k\xi}\, d\xi \\ &= 2^k \int_{-\pi}^{\pi} |m_0(\xi)|^2|\Phi_{k-1}(2^k\xi)|^2 e^{-in2^k\xi}\, d\xi \\ &= 2^k \int_0^{\pi} (|m_0(\xi)|^2 + |m_0(\xi+\pi)|^2)|\Phi_{k-1}(2^k\xi)|^2 e^{-in2^k\xi}\, d\xi \\ &= 2^k \int_0^{\pi} |\Phi_{k-1}(2^k\xi)|^2 e^{-in2^k\xi}\, d\xi = \int_0^{2^k\pi} |\Phi_{k-1}(\xi)|^2 e^{-in\xi}\, d\xi \\ &= \int |\Phi_{k-1}(\xi)|^2\chi(\xi/2^{k-1})e^{-in\xi}\, d\xi = \cdots = \int \chi(\xi)e^{-in\xi}\, d\xi = 2\pi\delta_{n,0}.\end{aligned} \tag{3.3.13}$$

If $\Phi_k\chi(\cdot/2^k)\to\Phi$ in L^2 then $|\Phi_k|^2\chi(\cdot/2^k)\to|\Phi|^2$ in L^1 which proves (3.3.12), hence (iv). The equivalence of the first four conditions is now established. and it is trivial that (iv) $\Longrightarrow$ (v) $\Longrightarrow$ (vi) $\Longrightarrow$ (vii). If (vii) is fulfilled it follows from the Borel-Lebesgue lemma that we can find a finite number of open sets $O_j\subset\mathbf{R}$ where $\hat\varphi\neq 0$ such that the projection $q:\mathbf{R}\to\mathbf{R}/2\pi\mathbf{Z}$ is bijective in each O_j and $\cup qO_j=\mathbf{R}/2\pi\mathbf{Z}$. Choose a partition of unity $\chi_j\in C_0^\infty(qO_j)$ in the circle and let $\tilde\chi_j=\chi_j\circ q$ in O_j, $\chi_j=0$ in $\mathbf{R}\setminus O_j$. Then $\tilde\chi_j\in C_0^\infty(\mathbf{R})$, and if $\tilde\chi=\sum\tilde\chi_j$ then $\hat\varphi\neq 0$ in $\operatorname{supp}\tilde\chi$ and $\sum_{l\in\mathbf{Z}}\tilde\chi(\xi+2\pi l)=\sum\chi_j\circ q(\xi)\equiv 1$. Since $\hat\varphi(0)=1$ we could take one of the sets O_j as a neighborhood of the origin and the corresponding χ_j equal to 1 in a neighborhood of the origin to attain that $\tilde\chi=1$ in a neighborhood of the origin, which proves (viii).

Next we prove that (viii) implies (iv), that is, that (3.3.12) is valid. To do so we can essentially repeat the proof that (iii) $\Longrightarrow$ (iv). In fact, for any function f of period 2π we have $\int_{\mathbf{R}}\tilde\chi(\xi)f(\xi)\,d\xi=\int_0^{2\pi}f(\xi)\,d\xi$. Replacing χ by $\tilde\chi$ in (3.3.13) we thus obtain

$$\int|\Phi_k(\xi)|^2\tilde\chi(\xi/2^k)e^{-in\xi}\,d\xi=\int\tilde\chi(\xi)e^{-in\xi}\,d\xi=2\pi\delta_{n,0}.$$

If $C=\min_{\xi\in\operatorname{supp}\tilde\chi}|\hat\varphi(\xi)|$, which is >0 by condition (viii), we have

$$|\Phi_k(\xi)|=|\hat\varphi(\xi)|/|\hat\varphi(\xi/2^k)|\le|\hat\varphi(\xi)|/C,\quad \text{if } \xi/2^k\in\operatorname{supp}\tilde\chi.$$

Since $|\hat\varphi|^2\in L^1$ and $\tilde\chi(\xi/2^k)\to 1$ as $k\to\infty$, we conclude by dominated convergence that (3.3.12) holds.

Since the sum g defined in (3.3.10) satisfies (3.3.11) it follows from (ix) that this sum is a constant, and it is not equal to 0 since $\hat\varphi(0)=1$, so (v) follows. We also have (x) $\Longrightarrow$ (ix), for assume that g is a non-constant solution of (3.3.11). It is no restriction to assume that g is real valued, and since g may be replaced by $-g$ we may assume that 0 is not a minimum point. By (3.1.6) every constant satisfies (3.3.11), so subtracting the minimum of g from g we obtain a solution of (3.11) with minimum equal to 0 which is positive at the origin. This would contradict (x). The proof will be complete if we can establish that (v) $\Longrightarrow$ (xi) $\Longrightarrow$ (x).

Assume that (x) is false so that there is a trigonometric polynomial g with period 2π satisfying (3.3.11) with $\min g=0$ and $g(0)>0$. Let $N=\{\dot\xi\in\mathbf{R}/2\pi\mathbf{Z};g(\dot\xi)=0\}$ which is a finite set. If $\dot\xi\in N$ then $\dot\xi\neq 0$ and it follows from (3.3.11) that there is some $\dot\eta\in\mathbf{R}/2\pi\mathbf{Z}$ with $2\dot\eta=\dot\xi$ such that $g(\dot\eta)=0$, thus $\dot\eta\in N$. Repeating the argument we get a sequence $\dot\xi_1,\dot\xi_2,\dot\xi_3,\dots$ with $\dot\xi_1=\dot\xi$ and $2\dot\xi_{j+1}=\dot\xi_j$. Since N is finite there must be some repetition, that is, $\dot\xi_j=\dot\xi_{j+r}$ for some j and some $r>0$. But then we have $\dot\xi_1=2^{j-1}\dot\xi_j=2^{j-1}\dot\xi_{j+r}=\dot\xi_{r+1}$. If r is chosen minimal it follows that we have an invariant cycle $\dot\xi_1,\dot\xi_r,\dots,\dot\xi_2,\dot\xi_1$ for the map $\dot\xi\mapsto 2\dot\xi$. Different cycles are disjoint since a cycle is uniquely determined by any one of its elements, so N is a union of such cycles, and in particular invariant under multiplication by 2. If $\dot\xi\in N$ then $\dot\xi+\dot\pi\notin N$, for since $2\dot\xi=2(\dot\xi+\dot\pi)$ they would otherwise be in the same cycle, of period r, so $\dot\xi=2^r\dot\xi=2^r(\dot\xi+\dot\pi)=\dot\xi+\dot\pi$ which is impossible. Since

$$0=g(2\dot\xi)=|m_0(\dot\xi)|^2g(\dot\xi)+|m_0(\dot\xi+\dot\pi)|^2g(\dot\xi+\dot\pi)=|m_0(\dot\xi+\dot\pi)|^2g(\dot\xi+\dot\pi)$$

it follows that $m_0(\dot{\xi}+\dot{\pi})=0$, hence $|m_0(\dot{\xi})|=1$. Each of the cycles which constitute N is therefore one of the forbidden cycles in (xi), so (xi) $\implies$ (x).

Finally we prove that (v) $\implies$ (xi). Assume that (xi) is false, and let $\dot{\xi}_1, \dot{\xi}_2 = 2\dot{\xi}_1, \dots, \dot{\xi}_r = 2^{r-1}\dot{\xi}_1$ be a cycle of period precisely $r > 1$ on which $|m_0| = 1$. Since $r > 1$ these elements in $\mathbf{R}/2\pi\mathbf{Z}$ are not 0. Hence there is a unique $x_j \in (0,1)$ such that $2\pi x_j$ is in the class of $\dot{\xi}_j$. We have a periodic binary expansion $x_1 = .d_1 \dots d_r d_1 \dots d_r \dots$, hence

$$x_j = .d_j \dots d_r d_1 \dots d_r d_1 \dots d_r \dots .$$

Both digits 0 and 1 will occur in these periodic expansions. By (3.1.6) we have $m_0(\dot{\eta}_j) = 0$ if $\dot{\eta}_j = \dot{\xi}_j + \dot{\pi}$, which is the class of $2\pi y_j$ where

$$y_j = .d_j' d_{j+1} \dots d_r d_1 \dots d_r \dots$$

with the notation $d' = 1 - d$ for the complementary digit. Now we claim that $\hat{\varphi}(\xi) = 0$ for every ξ in the residue class $\dot{\xi}_1$. Assume to the contrary that $\hat{\varphi}(\xi) \neq 0$ for some such ξ, and write the binary expansion

$$\xi/2\pi = ...D_k D_{k-1} \dots D_1.d_1 d_2 \dots d_r d_1 \dots d_r \dots$$

which is finite to the left of the binary point and equal to x_1 after it. That $\hat{\varphi}(\xi) \neq 0$ means that $m_0(2^{-k}\xi) \neq 0$ for all $k \geq 1$. Now $2^{-k}\xi/2\pi$ is obtained by moving the binary point k steps to the left. The digits after the new position will determine the value of $m_0(2^{-k}\xi)$. When $k = 1$ we must not have the digits of y_r, so it follows that $D_1 = d_r$. Repeating the argument we then conclude when $k = 2$ that $D_2 = d_{r-1}$, and continuing in this way we see that the digits $d_1 \dots d_r$ must be indefinitely repeated to the left. This is a contradiction proving that (v) is not fulfilled. The proof is complete.

Corollary 3.3.7. *If $m_0 \neq 0$ in $[-\pi/3, \pi/3]$ then the conditions in Theorem 3.3.6 are satisfied.*

Proof. It suffices to verify condition (xi). Since $m_0 \neq 0$ in $[-\pi/3, \pi/3]$ we have $|m_0(\xi)| \neq 1$ when $\big||\xi| - \pi\big| \leq \pi/3$. Suppose that S is a cycle as in condition (xi) and identify S with a subset of $(-\pi, \pi]$. Then $S \subset (-2\pi/3, 2\pi/3)$, and since $2S \subset (-2\pi/3, 2\pi/3)$ we have $S \subset (-\pi/3, \pi/3)$. Iterating this argument we obtain $S \subset (-2^{-\nu}\pi/3, 2^{-\nu}\pi/3)$ for every integer $\nu > 0$, so $S = \{0\}$, which is a trivial cycle. Hence (xi) is fulfilled.

Note that the example given before Theorem 3.3.5 shows that $\pi/3$ cannot be replaced by any smaller number in Corollary 3.3.7. The result would be trivial with $\pi/3$ replaced by $\pi/2$. Even that condition is fulfilled when m_0 is obtained from the polynomials P_N in Lemma 3.3.3, so we have obtained a large family of compactly supported orthonormal wavelets. Our next goal is to examine their regularity.

Recall that with a positive integer N we have then

$$m_0(\xi) = ((1 + e^{-i\xi})/2)^N \mathscr{L}(\xi), \tag{3.3.14}$$

where $\mathscr{L}(\xi)$ is a trigonometric polynomial with period 2π and

$$|\mathscr{L}(\xi)|^2 = P_N(\sin^2(\tfrac{1}{2}\xi)), \tag{3.3.15}$$

where P_N is the unique polynomial of degree $N-1$ satisfying (3.3.4), given by (3.3.5). In particular, $\mathscr{L}(0) = P_N(0) = 1$. We have

$$\hat{\varphi}(\xi) = \prod_1^\infty m_0(2^{-j}\xi) = ((1-e^{-i\xi})/i\xi)^N \prod_1^\infty \mathscr{L}(2^{-j}\xi), \tag{3.3.16}$$

for $\prod_1^\infty ((1+e^{-i\xi/2^j})/2) = (1-e^{-i\xi})/i\xi$. An instructive proof of this classical product formula is obtained by noting that the product for j from 1 to k is the Fourier transform of the measure

$$2^{-k}(\delta_0 + \delta_{\frac{1}{2}}) * (\delta_0 + \delta_{\frac{1}{4}}) * \cdots * (\delta_0 + \delta_{2^{-k}}) = 2^{-k} \sum_{0 \le \nu < 2^k} \delta_{\nu/2^k}.$$

Acting on a test function it gives a Riemann sum for the integral from 0 to 1, and when $k \to \infty$ it follows that the measure converges as a distribution to the characteristic function of $(0,1)$, so the Fourier transform converges to $(1-e^{-i\xi})/i\xi$.

Using only (3.3.14) and (3.3.16) we shall now give a lower bound for the decay of $\hat{\varphi}$ at infinity. We do not assume (3.3.15) but just that $\mathscr{L}(0) = 1$.

Proposition 3.3.7. *If $\dot{\xi}_1, \dots, \dot{\xi}_r \in \mathbf{R}/2\pi\mathbf{Z}$ is a non-trivial cycle for the doubling map $\dot{\xi} \mapsto 2\dot{\xi}$, then there is a constant $C > 0$ such that for every $\xi \in \mathbf{R}$ with residue class $\dot{\xi}$ in the cycle and every integer $\nu > 0$*

$$|\hat{\varphi}(2^\nu \xi)| \ge C|\hat{\varphi}(\xi)| 2^{-\nu N} K^\nu, \quad K = \prod_1^r |\mathscr{L}(\dot{\xi}_j)|^{\frac{1}{r}}. \tag{3.3.17}$$

Proof. Since the cycle is non-trivial we have $\dot{\xi}_j \ne 0$ for $j = 1, \dots, r$, hence $|1-e^{i\dot{\xi}_j}| \ge C_1 > 0$, $j = 1, \dots, r$, which implies

$$|(1-e^{-i2^\nu \xi})/i2^\nu \xi|^N \ge C_1^N 2^{-\nu N} |\xi|^{-N}.$$

If μ is the largest integer $\le \nu/r$ then

$$|\prod_1^\infty \mathscr{L}(2^{\nu-j}\xi)| = K^{\mu r} \prod_{\mu r+1}^\infty |\mathscr{L}(2^{\nu-j}\xi)|.$$

There are at most $r-1$ factors with $\mu r < j \le \nu$, and since the residue class of $2^{\nu-j}\xi$ is then in the cycle, we have lower bounds for them. (We may assume that $K > 0$ for (3.3.17) is trivial when $K = 0$.) By (3.3.16)

$$|\hat{\varphi}(\xi)| \le 2^N |\xi|^{-N} \prod_{j > \nu} |\mathscr{L}(2^{\nu-j}\xi)|,$$

so we have proved that

$$|\hat{\varphi}(2^\nu \xi)| \geq C_1^N 2^{-\nu N} |\xi|^{-N} K^\nu C_2 \prod_{j>\nu} \mathscr{L}(2^{\nu-j}\xi) \geq C_1^N 2^{-N} 2^{-\nu N} C_2 |\hat{\varphi}(\xi)| K^\nu$$

where C_2 is the minimum of products of the values $|\mathscr{L}(\dot{\xi}_j)|/K$ for different $j \in [1, r]$. The proof is complete.

An important example of a cycle already encountered several times consists of the residue classes of $\pm 2\pi/3$. If $\mathscr{L}$ is defined by (3.3.15) we have $K = P_N(\frac{3}{4})^{\frac{1}{2}}$ then. The following proposition gives an upper bound for $\hat{\varphi}$ which is closely related to the lower bound in (3.3.17).

PROPOSITION 3.3.8. *Suppose that there is an integer r such that for every $\xi \in \mathbf{R}$*

$$\min_{1 \leq \varrho \leq r} |\mathscr{L}(\xi)|^{1/\varrho} \ldots |\mathscr{L}(2^{\varrho-1}\xi)|^{1/\varrho} \leq K. \tag{3.3.18}$$

Then there is a constant C such that

$$|\hat{\varphi}(\xi)| \leq C(1 + |\xi|)^{-N + \log_2 K}. \tag{3.3.19}$$

PROOF. We may assume that $|\xi| \geq 1$ and that $K \leq \sup |\mathscr{L}|$. By (3.3.16)

$$|\hat{\varphi}(\xi)| \leq 2^N |\xi|^{-N} \prod_1^\infty |\mathscr{L}(2^{-k}\xi)|, \quad |\xi| \geq 1.$$

Since $\mathscr{L}(0) = 1$ we have $|\mathscr{L}(\xi)| \leq 1 + C|\xi| \leq e^{C|\xi|}$ for some C. If j is the smallest integer such that $2^{-j}|\xi| \leq 1$ it follows that

$$\prod_1^\infty |\mathscr{L}(2^{-k}\xi)| \leq e^C \prod_1^j |\mathscr{L}(2^{-k}\xi)|.$$

The hypothesis (3.3.18) applied to $2^{-j}\xi$ proves that

$$|\mathscr{L}(2^{-j}\xi) \cdots \mathscr{L}(2^{\varrho-1-j}\xi)| \leq K^\varrho$$

for some integer $\varrho \in [1, r]$. If $j > r$ it follows that

$$\prod_1^j |\mathscr{L}(2^{-k}\xi)| \leq K^\varrho \prod_1^{j-\varrho} |\mathscr{L}(2^{-k}\xi)|.$$

We can repeat this argument as long as there are r factors left in the product, which gives

$$\prod_1^\infty |\mathscr{L}(2^{-k}\xi)| \leq K^j e^C \sup |\mathscr{L}/K|^r.$$

Since $2^{-j}|\xi| \geq \frac{1}{2}$ we have $K^j \leq |2\xi|^{\log_2 K}$, which completes the proof.

We shall now prove that Propositions 3.3.7 and 3.3.8 suffice to determine the decay of $\hat{\varphi}$ when $\mathscr{L}$ satisfies (3.3.15). This requires some preliminaries on the polynomials P_N.

LEMMA 3.3.9. *$P_N(x)$ and $((1-x)^{-N} - P_N(x))/x^N$ are increasing on $(0,1)$, and $P_N(x)x^{1-N}$ is decreasing. We have*

$$P_N(0)=1, \quad P_N(\tfrac12)=2^{N-1}, \quad P_N(1)=\tbinom{2N-1}{N}=\tfrac12\tbinom{2N}{N}\ge 4^{N-1}/\sqrt{N}, \tag{3.3.20}$$

$$(1-x)^{-N}-\tfrac12(4x)^N \le P_N(x)\le (1-x)^{-N}, \quad 0\le x\le \tfrac12, \tag{3.3.21}$$

$$(4x)^{N-1}/\sqrt{N}\le P_N(x)\le (4x)^{N-1}, \quad \tfrac12\le x\le 1, \tag{3.3.22}$$

$$\lim_{N\to\infty} P_N(x)^{1/N} = \begin{cases} (1-x)^{-1}, & \text{if } 0\le x\le \frac12 \\ 4x, & \text{if } \frac12 \le x\le 1, \end{cases} \tag{3.3.23}$$

$$P_N'(x) = N(P_N(x)-P_N(1)x^{N-1})/(1-x). \tag{3.3.24}$$

PROOF. The first statements follow since power series with positive coefficients are increasing for positive arguments. From (3.3.5) we know that $P_N(0)=1$, and when $y=\frac12$ it follows from (3.3.4) that $P_N(\frac12)=2^{N-1}$. Hence

$$((1-x)^{-N}-P_N(x))/x^N \le 2^{2N-1}, \quad 0<x\le \tfrac12,$$

for the left-hand side is increasing and there is equality when $x=\frac12$. This proves (3.3.21), which implies (3.3.23) when $0\le x<\frac12$. Since $P_N(x)/x^{N-1}\le P_N(\frac12)2^{N-1}=4^{N-1}$ the upper bound in (3.3.22) holds, and the lower bound follows in the same way when we have proved the last part of (3.3.20). To do so we use Pascal's triangle

$$\tbinom{2N-j-1}{N} = \tbinom{2N-j-2}{N-1} + \tbinom{2N-j-2}{N}, \quad j=0,\dots,N-1,$$

where the last term should be omitted when $j=N-1$. Summation gives

$$P_N(1)=\sum_0^{N-1}\tbinom{2N-j-2}{N-1} = \tbinom{2N-1}{N},$$

hence $P_1(1)=1$ and $P_{N+1}(1)/P_N(1)=(4N+2)/(N+1)\ge 4\sqrt{N/(N+1)}$ because $(2N+1)^2\ge 4N(N+1)$, so the last inequality in (3.3.20) follows by induction. (3.3.23) follows from (3.3.22) when $\frac12\le x\le 1$. To prove (3.3.24) finally we use that

$$P_N(x)(1-x)^N = 1+O(x^N) \quad\Longrightarrow\quad P_N'(x)(1-x)^N - NP_N(x)(1-x)^{N-1} = O(x^{N-1}).$$

Hence the difference between the two sides of (3.3.24) is a polynomial of degree $\le N-2$ which is $O(x^{N-1})$ as $x\to 0$, so it is equal to 0.

REMARK. From (3.3.21) it follows that $P_N(x)$ is asymptotic to $(1-x)^{-N}$ if $0\le x<\frac12$. Using Stirling's formula to deduce the asymptotics of the highest coefficients in P_N one obtains for $\frac12<x\le 1$ that $P_N(x)$ is asymptotic to $(\pi N)^{-\frac12}2x(2x-1)^{-1}(4x)^{N-1}$.

We can now prove the properties of P_N required to apply Proposition 3.3.9 with $r=2$. Note that if $\sin^2(\frac12\xi)=x$ then $\sin^2\xi = 4x(1-x)$.

LEMMA 3.3.10. *For every positive integer N we have*

$$0 \le P_N(x) \le P_N(\tfrac{3}{4}), \quad 0 \le x \le \tfrac{3}{4}, \tag{3.3.25}$$

$$P_N(x)P_N(4x(1-x)) \le P_N(\tfrac{3}{4})^2, \quad \tfrac{3}{4} \le x \le 1. \tag{3.3.26}$$

PROOF. (3.3.25) is obvious since P_N is increasing. To prove (3.3.26) we set

$$f(x) = P_N(x)P_N(y(x)), \quad y(x) = 4x(1-x).$$

Since $y'(x) = 4(1-2x)$ and $1 - y(x) = (1-2x)^2$, it follows from (3.3.24) that

$$\begin{aligned} &f'(x)(1-x)(2x-1)/N \\ &\quad = (P_N(x) - P_N(1)x^{N-1})(2x-1)P_N(y) - 4(1-x)P_N(x)(P_N(y) - P_N(1)y^{N-1}) \\ &\qquad = P_N(x)P_N(y)(6x-5) - P_N(1)((2x-1)P_N(y)x^{N-1} + 4(x-1)P_N(x)y^{N-1}). \end{aligned}$$

Now $P_N(x)y^{N-1} \le P_N(y)x^{N-1}$ since $y \le x$, so the right-hand side is bounded above by

$$(6x-5)P_N(y)(P_N(x) - x^{N-1}P_N(1)) \le 0, \quad \text{if } 6x - 5 \le 0.$$

This proves (3.3.26) when $\frac{3}{4} \le x \le \frac{5}{6}$.

Now assume that $\frac{5}{6} \le x \le 1$. Since $P_N(x) \le (4x/3)^{N-1}P_N(3/4)$ the inequality (3.3.26) follows if

$$(4x/3)^{N-1}P_N(y) \le P_N(3/4).$$

We have $P_N(y) \le (1-y)^{-N} = (2x-1)^{-2N}$, and $P_N(3/4) \ge (3/4)^{N-1}P_N(1) \ge 3^{N-1}/\sqrt{N}$ by (3.3.20). Hence (3.3.26) follows if

$$\big(x/(2x-1)^2\big)^N x^{-1} \le (9/4)^{N-1}/\sqrt{N}, \quad \tfrac{5}{6} \le x \le 1.$$

Since $x/(2x-1)^2$ is decreasing for $\frac{5}{6} \le x \le 1$ we only have to verify this inequality when $x = \frac{5}{6}$, and then it requires that

$$(5/6)^{N-1} \le 4/(9\sqrt{N}),$$

which is true when $N \ge 13$.

We may now also assume that $1 \le N \le 12$. At first we require that $\frac{5}{6} \le x \le x_0$ where $x_0 = (2+\sqrt{2})/4$, which means that $y \ge \frac{1}{2}$. Then we have by (3.3.22)

$$P_N(y) \le (4y)^{N-1} = (16x(1-x))^{N-1},$$

and since $P_N(x) \le (6x/5)^{N-1}P_N(5/6)$ it follows that for $\frac{5}{6} \le x \le x_0$

$$P_N(x)P_N(y) \le (6/5)^{N-1}P_N(5/6)(16x^2(1-x))^{N-1} \le (20/9)^{N-1}P_N(5/6).$$

Numerical calculation shows that the right-hand side is $\le P_N(3/4)^2$ for $1 \le N \le 12$.

If $x_0 \le x \le 1$ we have $y \le \frac{1}{2}$ and $P_N(y) \le (1-y)^{-N} = (2x-1)^{-2N}$ by (3.3.21). Since $P_N(x) \le (x/x_0)^{N-1}P_N(x_0)$ it follows that

$$P_N(x)P_N(y) \le x_0^{1-N}P_N(x_0)\big(x/(2x-1)^2\big)^N x^{-1} \le 2^N P_N(x_0),$$

for $x/(2x-1)^2$ is decreasing in $[x_0, 1]$ and equal to $2x_0$ when $x = x_0$. Numerical calculation shows that this is $\le P_N(3/4)^2$ for $5 \le N \le 12$. The cases where $1 \le N \le 4$ are settled by calculating the zeros of $(P_N(x)P_N(4x(1-x)) - P_N(3/4)^2)/(x - 3/4)$ numerically; the degree of this polynomial is $3(N-1) - 1 \le 8$. (See Daubechies [1, p. 225].)

THEOREM 3.3.11. *If the scale function* φ *is defined by* (3.3.16) *with* $\mathscr{L}$ *satisfying* (3.3.15), *then*

$$|\hat{\varphi}(\xi)| \le C(1+|\xi|)^{-N+\frac{1}{2}\log_2 P_N(\frac{3}{4})}, \quad \xi \in \mathbf{R}, \tag{3.3.27}$$

and there is no such estimate with a smaller exponent.

PROOF. The estimate (3.3.27) follows from Proposition 3.3.8 and Lemma 3.3.10, and Proposition 3.3.7 with the cycle consisting of the residue classes of $\pm 2\pi/3$ proves that the estimate is optimal.

By the remark after Lemma 3.3.9 the exponent of $(1+|\xi|)^{-1}$ in (3.3.27) is asymptotically

$$N(1-\tfrac{1}{2}\log_2 3) + \tfrac{1}{4}\log_2(N\pi) + 1 + o(1).$$

For $N = 2, \dots, 10$ the numerical values are 1.3390, 1.6360, 1.9125, 2.1766, 2.4322, 2.6817, 2.9265, 3.1676, 3.4057. These are upper bounds for the Hölder class of φ and ψ, and subtracting $1+\varepsilon$ with $\varepsilon > 0$ gives a lower bound. The precise determination of the Hölder class will not be discussed here. Note that asymptotically the exponent is only about $0.21N$ although the length of the support of φ is $2N-1$. We have N vanishing moments for ψ, many more than the regularity of φ implies by Proposition 3.3.1. Using polynomials $R \not\equiv 0$ in Lemma 3.3.3 one can increase the regularity while decreasing the number of vanishing moments, keeping the length of the support fixed. However, we shall not discuss these matters here.

CHAPTER IV

SINGULAR INTEGRAL OPERATORS

4.1. The conjugate function. A basic question in the study of Fourier series is to decide in what sense the partial sums converge. In Proposition 2.1.1 and Theorem 2.1.3 we saw that the Fourier series of a periodic C^∞ or L^2_{loc} function or distribution f converges to f in C^∞, L^2_{loc} or $\mathscr{D}'$ respectively. In this section we shall discuss the convergence problem when $f \in L^p_{\text{loc}}(\mathbf{R})$. However, we shall first discuss the analogue for non-periodic functions $f \in L^p(\mathbf{R})$ since the formulas are more transparent then. The question is thus if the inverse Fourier transform $f_{a,b}$ of the product of the Fourier transform of f by the characteristic function of $[a, b]$ converges to f in L^p as $a \to -\infty$ and $b \to +\infty$. The product is well defined at least if $1 \le p \le 2$, by Theorem 2.3.1, but even then it is not a priori clear that $f_{a,b}$ is in L^p. The proof of that is a major part of our task, which is not present in the case of Fourier series. The passage from f to $f_{a,b}$ can be made in two steps, multiplying the Fourier transform first by the characteristic function of $[a, \infty]$ and then by the characteristic function of $[-\infty, b]$. The first step is equivalent to multiplication of the Fourier transform of $x \mapsto f(x)e^{iax}$ by the Heaviside function H, the characteristic function of the positive real axis, followed by the inverse Fourier transformation and multiplication by e^{-iax}. The second step is similar. Thus we can reduce to the study of the operator consisting of multiplication of the Fourier transform by $H(\xi) = (1 + \operatorname{sgn}\xi)/2$; for reasons of symmetry and tradition we shall discuss $\operatorname{sgn}\xi$ instead of $H(\xi)$.

The inverse Fourier transform of $\xi \mapsto \operatorname{sgn}\xi$ is the distribution $(i/\pi)\operatorname{vp}(1/x)$ (see Example 2 after Theorem 2.1.5). If $f \in C_0^\infty(\mathbf{R})$ it follows that $\hat f(\xi)\operatorname{sgn}\xi$ is the Fourier transform of $i\tilde f(x)$ where $\tilde f$ is the *conjugate function* defined by the convolution

$$\tilde f(x) = \operatorname{vp}\frac{1}{\pi}\int \frac{f(t)}{x-t}\,dt = \lim_{\varepsilon\to+0}\frac{1}{\pi}\int_{|x-t|>\varepsilon}\frac{f(t)}{x-t}\,dt = \lim_{\varepsilon\to+0}\frac{1}{\pi}\int_{|t|>\varepsilon}\frac{f(x-t)}{t}\,dt. \tag{4.1.1}$$

This is a C^∞ function and

$$\tilde f(x) = \frac{1}{\pi x}\int_{\mathbf{R}} f(t)\,dt + O(x^{-2}) \quad \text{as } x \to \infty, \tag{4.1.2}$$

so $\tilde f \in L^p$ if $p > 1$ but not if $p = 1$ unless $\int_{\mathbf{R}} f(t)\,dt = 0$.

The Fourier transform of $F_\pm(x) = f(x) \pm i\tilde f(x)$ vanishes on the negative or positive real axis respectively, so $F_\pm$ can be extended to an analytic function in the half plane $\{z \in \mathbf{C}; \pm\operatorname{Im} z > 0\}$. This is made explicit by the formula

$$f(x) \pm i\tilde f(x) = \lim_{\varepsilon\to+0}\frac{\pm i}{\pi}\int_{\mathbf{R}}\frac{f(t)}{x \pm i\varepsilon - t}\,dt = \lim_{\varepsilon\to+0}\frac{\pm i}{\pi}\int_{\mathbf{R}}\frac{f(x-t)}{t \pm i\varepsilon}\,dt,$$

which follows since $\pm i/(t \pm i\varepsilon) = \pm it/(t^2+\varepsilon^2) + \varepsilon/(t^2+\varepsilon^2)$ and

$$\frac{1}{\pi}\int_{\mathbf{R}} \frac{f(x-t)\varepsilon}{t^2+\varepsilon^2}\,dt = \frac{1}{\pi}\int_{\mathbf{R}} \frac{f(x-\varepsilon t)}{t^2+1}\,dt \to f(x),$$

$$\int_{\mathbf{R}} \frac{tf(x-t)}{t^2+\varepsilon^2}\,dt = \int_0^\infty \frac{t(f(x-t)-f(x+t))}{t^2+\varepsilon^2}\,dt \to \int_0^\infty \frac{f(x-t)-f(x+t)}{t}\,dt$$
$$= \lim_{\varepsilon\to+0}\int_\varepsilon^\infty \frac{f(x-t)-f(x+t)}{t}\,dt = \pi\tilde f(x),$$

when $\varepsilon \to +0$. Thus

$$F_\pm(z) = \pm\frac{i}{\pi}\int_{\mathbf{R}} \frac{f(t)}{z-t}\,dt$$

is analytic when $\pm\operatorname{Im} z > 0$ and has boundary values $f \pm i\tilde f$ on the real axis. We assume now that f is real valued, which implies that f is the real part of the boundary values.

To estimate $\tilde f$ we shall use the fact that

$$z \mapsto p|\operatorname{Re} F_\pm(z)|^p - (p-1)|F_\pm(z)|^p \tag{4.1.3}$$

is subharmonic when $\pm\operatorname{Im} z > 0$ if $1 < p \le 2$. This follows since subharmonicity is invariant under composition with analytic maps and

$$\mathbf{C} \ni w \mapsto p|\operatorname{Re} w|^p - (p-1)|w|^p \tag{4.1.4}$$

is subharmonic, because the Laplacian in the sense of distribution theory is the function

$$w \mapsto p^2(p-1)(|\operatorname{Re} w|^{p-2} - |w|^{p-2}) \ge 0.$$

When $z = x+iy$, $\pm y > 0$, then the integral of (4.1.3) with respect to x exists by (4.1.2) and is a continuous function $I_\pm(y)$ which $\to 0$ at $\pm\infty$. As a limit of the Riemann sums at the points $z + \varepsilon\mathbf{Z}$ with $\varepsilon \to 0$ it is clear that $I_\pm$ is subharmonic as a function of z, hence a convex function of y. Since $I_\pm(y) \to 0$ at infinity it is a decreasing function of $|y|$, which proves that $I_\pm(0) \ge 0$, that is,

$$\int_{\mathbf{R}} (p|f(x)|^p - (p-1)|f(x) \pm i\tilde f(x)|^p)\,dx \ge 0.$$

This means that

$$\|f \pm i\tilde f\|_p \le (p/(p-1))^{1/p}\|f\|_p, \quad \|\tilde f\|_p \le (p/(p-1))^{1/p}\|f\|_p,$$

where the second inequality follows from the first and the triangle inequality.

Theorem 4.1.1. *If $f \in C_0^\infty(\mathbf{R})$ then the conjugate function $\tilde{f}$ satisfies* (4.1.2) *and*

$$\|\tilde{f}\|_p \le \begin{cases} p'^{1/p}\|f\|_p, & \text{if } 1 < p \le 2, \\ p^{1/p'}\|f\|_p, & \text{if } 2 \le p < \infty. \end{cases} \tag{4.1.5}$$

Here $1/p + 1/p' = 1$. The map $f \mapsto \tilde{f}$ extends to a continuous map $L^p(\mathbf{R}) \to L^p(\mathbf{R})$ for every $p \in (1,\infty)$.

Proof. We have already proved (4.1.5) when $1 < p \le 2$. The other case follows by duality, for if $g \in C_0^\infty(\mathbf{R})$ and $p > 2$ then

$$\int \tilde{f}(x)g(x)\,dx = \lim_{\varepsilon\to 0}\iint_{|x-t|>\varepsilon} \frac{f(t)g(x)}{x-t}\,dx\,dt = -\int f(x)\tilde{g}(x)\,dx,$$

hence

$$\Big|\int \tilde{f}(x)g(x)\,dx\Big| = \Big|\int f(x)\tilde{g}(x)\,dx\Big| \le \|f\|_p\|\tilde{g}\|_{p'} \le \|f\|_p\|g\|_{p'}p^{1/p'},$$

so the converse of Hölder's inequality gives the second part of (4.1.5).

Remark. Apart from the size of the constant the estimate (4.1.5) is due to M. Riesz [1]. The proof given here is due to P. Stein [1] and is a prototype for much later work, by E. M. Stein and others. Essén [1] has obtained optimal constants by modifying the function (4.1.4). The constant in (4.1.5) is reasonably good though. If we take for f the characteristic function of $(-1,1)$, then $\pi\tilde{f}(x) = \log((x+1)/(x-1)) > 2/x$ when $x > 1$, and we obtain $\|\tilde{f}\|_p^p/\|f\|_p^p \ge 2^p/(p-1)$. This proves that the best possible constant is at least $\frac{1}{2}$ times that in (4.1.5).

Next we shall prove that for the operator theoretically defined extension of $\tilde{f}$ in Theorem 4.1.1 the equation (4.1.1) remains valid almost everywhere. To do so we need the one dimensional case of the *Hardy-Littlewood maximal theorem*, which is particularly elementary to prove in that case. If $f \in L^1_{\text{loc}}(\mathbf{R})$ then the Hardy-Littlewood maximal function f^*_{HL} is defined by

$$f^*_{\text{HL}}(x) = \sup_{x\in I}\int_I |f(t)|\,dt/m(I) \tag{4.1.6}$$

where I is an interval with measure $m(I) > 0$. Since $\int_I |f(t)|\,dt$ is a continuous function of the end points of I, it is clear that $f^*_{\text{HL}}(x)$ does not change if we require I to be open, hence

$$f^*_{\text{HL}}(x) = \sup_{\varepsilon>0,\delta>0}\frac{1}{\varepsilon+\delta}\int_{-\varepsilon}^{\delta}|f(x+t)|\,dt$$

is a lower semi-continuous function. A useful version of (4.1.6) is that

$$\int_{\mathbf{R}} |f(x+t)|\varrho(t)\,dt \le f^*_{\text{HL}}(x)\int_{\mathbf{R}}\varrho(t)\,dt, \tag{4.1.6$'$}$$

if $\varrho \ge 0$ is increasing for $t < 0$ and decreasing for $t > 0$. In fact, (4.1.6) means precisely that $f^*_{\text{HL}}(x)$ is the smallest number such that this is true when ϱ is the characteristic function of an interval containing 0, and this implies (4.1.6)$'$ if ϱ is piecewise constant, hence in general.

THEOREM 4.1.2. *If* $\in L^1(\mathbf{R})$ *then*

$$m\{x; f^*_{\rm HL}(x) > \alpha\} \le 2\|f\|_1/\alpha, \quad \alpha > 0. \tag{4.1.7}$$

If $1 < p < \infty$ *then*

$$\|f^*_{\rm HL}\|_p \le (2^{p+1}p')^{1/p}\|f\|_p, \quad f \in L^p(\mathbf{R}). \tag{4.1.8}$$

PROOF. To prove (4.1.7) we introduce two additional maximal functions

$$f^*_\pm(x) = \sup_{\varepsilon>0}\int_0^\varepsilon |f(x\pm t)|\,dt/\varepsilon.$$

It is clear that $f^*_{\rm HL} = \max(f^*_+, f^*_-)$, so (4.1.7) follows if we prove the corresponding estimate for $f^*_\pm$ with half the constant. Of course it suffices to examine f^*_+. If we set

$$F_\alpha(x) = \int_{-\infty}^x |f(t)|\,dt - \alpha x, \quad O = \{x \in \mathbf{R}; \exists y > x,\ F_\alpha(y) > F_\alpha(x)\}$$

the statement is that $m(O) \le \|f\|_1/\alpha$. Let $I = (a, b)$ be a component of the open set O. It cannot be an infinite interval for if $x_0 \in I$ then a maximum point y for F_α in $[x_0, \infty)$ cannot be in O, and $F_\alpha(x) < F_\alpha(y)$ for $x \in I$. In fact, this is true if $x \in I$ and $x \ge x_0$ and the infimum of all $x \in I$ with $F_\alpha(x) < F_\alpha(y)$ cannot belong to I. It is now clear that b is the smallest maximum point for F_α in $[x_0, \infty)$, and that $F_\alpha(a) = F_\alpha(b)$. Hence $\int_I |f(t)|\,dt = \alpha(b-a)$ and $\alpha m(O) = \int_O |f(t)|\,dt \le \|f\|_1$, which proves (4.1.7).

To prove (4.1.8) we write for an arbitrary $s > 0$

$$f = g + h \quad \text{where } g = \begin{cases} f, & \text{if } |f| \le s/2,\\ 0, & \text{if } |f| > s/2.\end{cases}$$

It is obvious that $g^*_{\rm HL} \le s/2$, and since $f^*_{\rm HL} \le g^*_{\rm HL} + h^*_{\rm HL}$ it follows that $h^*_{\rm HL} \ge s/2$ in $E_s = \{x; f^*_{\rm HL}(x) > s\}$. Hence (4.1.7) gives $m(E_s) \le 4\|h\|_1/s$, and

$$\begin{aligned}\int_{\mathbf{R}} |f^*_{\rm HL}|^p\,dx &= \int_0^\infty s^p d(-m(E_s)) = \int_0^\infty m(E_s)d(s^p)\\ &= 4p\int_0^\infty s^{p-2}\int_{|f(x)|>s/2}|f(x)|\,dx = 2^{p+1}p/(p-1)\int_{\mathbf{R}}|f(x)|^p\,dx,\end{aligned}$$

where the last equality follows by changing the order of integration. This completes the proof.

REMARK. The proof of (4.1.7) can be interpreted as a determination of the points on the graph of $x \mapsto \int_{-\infty}^x |f(t)|\,dt$ where one cannot see the sun if it is at infinity to the right in the direction with slope α, so it is often called the rising sun lemma. The constant in (4.1.7) is optimal as follows immediately by letting f approach the Dirac measure at 0. Taking for f the characteristic function of $(-1, 1)$ it is easy to see that the constant in (4.1.8) cannot be improved by a factor greater than 3; in particular it has the right magnitude as $p \to 1$. The proof of (4.1.8) is our first encounter with the *Marcinkiewicz interpolation method* which will be presented in generality later on.

Example. The maximal theorem 4.1.2 also yields estimates for transformations where $(4.1.6)'$ is not directly applicable. An example is a classical *potential estimate of Hardy and Littlewood*: If $1 < p < q < \infty$ and $1/q = 1/p - \gamma$, then

$$f_\gamma(x) = \int_{\mathbf{R}} f(x-y)|y|^{\gamma-1}\,dy$$

exists for almost all $x \in \mathbf{R}$ if $f \in L^p(\mathbf{R})$, and $\|f_\gamma\|_q \le C_{p,q}\|f\|_p$. Since $y \mapsto |y|^{\gamma-1}$ is not integrable at infinity we apply $(4.1.6)'$ to the integral when $|y| < R$ for some R to be chosen later and apply Hölder's inequality when $|y| > R$. With $1/p + 1/p' = 1$ we obtain

$$\int_{\mathbf{R}} |f(x-y)||y|^{\gamma-1}\,dy \le f^*_{\text{HL}}(x)\int_{|y|<R}|y|^{\gamma-1}\,dy + \|f\|_p\Big(\int_{|y|>R}|y|^{p'(\gamma-1)}\,dy\Big)^{\frac{1}{p'}}.$$

Here $p'(\gamma-1) = -1-p'/q$ so the second integral converges, which proves that $f_\gamma(x)$ exists when $f^*_{\text{HL}}(x) < \infty$ and that

$$|f_\gamma(x)| \le C_{p,q}(f^*_{\text{HL}}(x)R^{\frac{1}{p}-\frac{1}{q}} + \|f\|_p R^{-\frac{1}{q}}).$$

When $R^{1/p} = \|f\|_p/f^*_{\text{HL}}(x)$ it follows that

$$|f_\gamma(x)| \le 2C_{p,q}f^*_{\text{HL}}(x)^{\frac{p}{q}}\|f\|_p^{1-\frac{p}{q}}.$$

Hence $\|f_\gamma\|_q \le 2C_{p,q}\|f^*_{\text{HL}}\|_p^{p/q}\|f\|_p^{1-p/q}$ and (4.1.8) gives $\|f_\gamma\|_q \le C'_{p,q}\|f\|_p$.

Later on in this section we shall need a variant of Theorem 4.1.2 with essentially the same proof, so we pause to prove it now. The issue is the maximal function

$$(4.1.6)'' \qquad f^{**}_{\text{HL}}(x,s) = \sup_{(x-s,x+s)\subset I}\int_I |f(t)|\,dt/m(I), \quad x \in \mathbf{R},\ s > 0,$$

which also takes the size of the interval I into account. The point of this function is that in analogy to $(4.1.6)'$

$$(4.1.6)''' \qquad \int_{\mathbf{R}} |f(x+t)|\varrho(t)\,dt \le f^{**}_{\text{HL}}(x,s)\int_{\mathbf{R}}\varrho(t)\,dt,$$

if $\varrho \ge 0$ is increasing for $t < 0$, decreasing for $t > 0$ and constant in $(-s,s)$. The following result is essentially due to L. Carleson:

Theorem 4.1.2′. *Let $d\nu$ be a positive measure in $\{(x,s) \in \mathbf{R}^2; s > 0\}$ and assume that $\nu(I \times (0,|I|)) \le |I|$ for every interval $I \subset \mathbf{R}$. Then it follows that*

$$(4.1.7)' \qquad \nu(\{(x,s); f^{**}_{\text{HL}}(x,s) > \alpha\}) \le 4\|f\|_1/\alpha, \quad \alpha > 0,\ f \in L^1(\mathbf{R}),$$

$$(4.1.8)' \qquad \Big(\iint_{s>0}|f^{**}_{\text{HL}}(x,s)|^p\,d\nu(x,s)\Big)^{1/p} \le (2^{p+2}p')^{1/p}\|f\|_p, \quad f \in L^p(\mathbf{R}).$$

PROOF. As in the proof of Theorem 4.1.2 we introduce corresponding left and right maximal functions

$$f_{\pm}^{**}(x,s) = \sup_{\varepsilon>s} \int_0^{\varepsilon} |f(x\pm t)|dt/\varepsilon.$$

If $f_{\rm HL}^{**}(x,s) > \alpha$ then $f_{+}^{**}(x,s) > \alpha$ or $f_{-}^{**}(x,s) > \alpha$, so to prove (4.1.7)′ it suffices to prove that

$$\nu(\{(x,s); f_{+}^{**}(x,s) > \alpha\}) \le 2\|f\|_1/\alpha, \tag{4.1.7$''$}$$

for this gives a similar bound for f_{-}^{**}. Now $f_{+}^{**}(x,s) > \alpha$ implies $f_{+}^{*}(x) > \alpha$, so x belongs to one of the intervals $I = (a,b)$ in the proof of Theorem 4.1.2; we keep the notation used there. Since $F_\alpha(x) \le F_\alpha(a)$ when $x \ge a$, we have $\int_x^{x+t} |f(y)|\,dy \le \int_a^{x+t} |f(y)|\,dy \le \alpha(x+t-a)$, hence

$$f_{+}^{**}(x,s) \le \sup_{t>s} \alpha(x-a+t)/t = \alpha(x-a+s)/s \le \alpha(m(I)+s)/s \le 2\alpha,$$

if $x \in I$ and $s \ge m(I)$. This means that $\{(x,s); f_{+}^{**}(x,s) > 2\alpha\} \subset \bigcup I \times (0,|I|)$, with the union taken over the disjoint components of the open set O, which proves (4.1.7)″ with α replaced by 2α. The proof that (4.1.7)′ implies (4.1.8)′ is a repetition of the proof that (4.1.7) implies (4.1.8) and is left for the reader; it is our second case of Marcinkiewicz' interpolation method. (See Theorem 4.2.4.)

We return now to the study of the principal value integral (4.1.1) and introduce another maximal function

$$f_{\rm CZ}^{*}(x) = \sup_{0<\varepsilon<\delta} \Big| \int_{\varepsilon<|t|<\delta} \frac{f(x-t)}{t}\,dt \Big|. \tag{4.1.9}$$

To estimate $f_{\rm CZ}^{*}$ we first assume that $f \in C_0^\infty(\mathbf{R})$. Choose a fixed $\chi \in C_0^\infty(\mathbf{R})$ with $\int_{\mathbf{R}} \chi(t)\,dt = 1$ such that $\chi(x)$ is a decreasing function of $|x|$. By (4.1.2)

$$|\tilde\chi(x) - 1/\pi x| \le C/x^2,$$

and $\tilde\chi \in C^\infty$. With the notation $\chi_\varepsilon(x) = \chi(x/\varepsilon)/\varepsilon$ we have $\tilde f * \chi_\varepsilon = f * \widetilde{\chi_\varepsilon} = f * \tilde\chi_\varepsilon$, hence

$$|f*(\tilde\chi_\varepsilon - \tilde\chi_\delta)(x)| \le |\tilde f * \chi_\varepsilon(x)| + |\tilde f * \chi_\delta(x)| \le 2\tilde f_{\rm HL}^{*}(x), \tag{4.1.10}$$

by (4.1.6)′. Let $K_{\varepsilon,\delta}(x) = 1/\pi x$ when $\varepsilon < |x| < \delta$ and $K_{\varepsilon,\delta}(x) = 0$ otherwise. Then

$$|\tilde\chi_\varepsilon(x) - \tilde\chi_\delta(x) - K_{\varepsilon,\delta}(x)| \le \begin{cases} C(\varepsilon+\delta)/|x|^2, & \text{if } |x|>\delta, \\ C\varepsilon/|x|^2 + |\tilde\chi_\delta(x)|, & \text{if } \varepsilon<|x|<\delta, \\ |\tilde\chi_\varepsilon(x)| + |\tilde\chi_\delta(x)|, & \text{if } |x|<\varepsilon, \end{cases}$$

because $|\tilde\chi_\varepsilon(x) - 1/\pi x| \le C\varepsilon/x^2$. By (4.1.6)′ again it follows that

$$|f*\tilde\chi_\varepsilon(x) - f*\tilde\chi_\delta(x) - f*K_{\varepsilon,\delta}(x)| \le 4(C+\max|\tilde\chi|)f_{\rm HL}^{*}(x). \tag{4.1.11}$$

Combining (4.1.11) with (4.1.10) and (4.1.5), (4.1.8) we obtain the estimate in the following theorem:

THEOREM 4.1.3. *When* $1 < p < \infty$ *there is a constant* C_p *such that the maximal function* (4.1.9) *has the bound*

$$\|f^*_{\mathrm{CZ}}\|_p \le C_p\|f\|_p, \quad f \in L^p(\mathbf{R}). \tag{4.1.12}$$

When $f \in L^p(\mathbf{R})$ *the limit* $\tilde f(x) = \lim_{\varepsilon\to 0,\delta\to\infty}\int_{\varepsilon<|t|<\delta} f(x-t)\,dt/\pi t$ *exists for almost every* $x \in \mathbf{R}$, *and is in* $L^p(\mathbf{R})$. *The map* $L^p(\mathbf{R}) \ni f \mapsto \tilde f \in L^p(\mathbf{R})$ *is continuous.*

PROOF. So far we have only proved (4.1.12) when $f \in C_0^\infty(\mathbf{R})$, but the general statement follows at once since this is a dense subset of $L^p(\mathbf{R})$. To prove that the principal values exist almost everywhere we introduce

$$F(x) = \varlimsup_{\varepsilon,\varepsilon'\to 0,\delta,\delta'\to\infty}\Big|\int_{\varepsilon<|t|<\delta}\frac{f(x-t)}{t}\,dt - \int_{\varepsilon'<|t|<\delta'}\frac{f(x-t)}{t}\,dt\Big|.$$

It follows from (4.1.12) that

$$\|F\|_p \le 2C_p\|f\|_p.$$

Now F_p does not change if we replace f by $f-g$ where $g \in C_0^\infty(\mathbf{R})$. Hence $\|F\|_p \le 2C_p\|f-g\|_p$, $g \in C_0^\infty(\mathbf{R})$, so we conclude that $F = 0$ almost everywhere which proves the pointwise existence of the principal value. The map $f \mapsto \tilde f$ with this pointwise definition of $\tilde f$ is a continuous linear map in L^p so it agrees with the extension defined after Theorem 4.1.1.

Thus we now have an unambigous definition of $\tilde f$ when $f \in L^p(\mathbf{R})$ and $1 < p < \infty$. Returning to the discussion at the beginning of the section we define $f_{a,b}$ as the function in $L^p(\mathbf{R})$ with Fourier transform equal to $\hat f(\xi)$ when $a \le \xi \le b$ and 0 elsewhere, when $f \in L^p(\mathbf{R}) \cap L^1(\mathbf{R})$, say. It follows from (4.1.5) that $f_{a,b} \in L^p(\mathbf{R})$ and that

$$\|f_{a,b}\|_p \le C_p\|f\|_p, \quad f \in L^p(\mathbf{R}).$$

Since $\|f_{a,b} - f\|_p \to 0$ as $a \to -\infty$ and $b \to +\infty$ provided that $\hat f \in C_0^\infty(\mathbf{R})$, and such functions are dense in $\mathscr{S}(\mathbf{R})$, hence in $L^p(\mathbf{R})$, we obtain:

COROLLARY 4.1.4. *If* $f \in L^p(\mathbf{R})$ *where* $1 < p < \infty$ *then the partial inverse Fourier transforms* $f_{a,b}$ *converge to* f *in* $L^p(\mathbf{R})$ *as* $a \to -\infty$ *and* $b \to +\infty$.

We started the section with discussing the Fourier series of a periodic function. If $f \in L^p_{\mathrm{loc}}(\mathbf{R})$ is periodic with period T we define the conjugate function $\tilde f$ as above by multiplying the Fourier transform by $-i\operatorname{sgn}\xi$ defined as 0 when $\xi = 0$, that is, multiplying the coefficient of $e^{2\pi ikx/T}$ in the Fourier series by $-i\operatorname{sgn}k$, thus dropping the constant term. It follows at once from Proposition 2.1.1 that $\tilde f \in C^\infty$ if $f \in C^\infty$. We could extend the estimate (4.1.5) by repeating the proof, but instead of doing that we shall prove that the estimate in the non-periodic case carries over to the periodic case. This is a principle which is useful in other contexts as well. Choose $\varphi \in \mathscr{S}\setminus\{0\}$ so that $\operatorname{supp}\hat\varphi \subset (-1,1)$, and form for a given $f \in C^\infty(\mathbf{R})$ with period T

$$f_\varepsilon(x) = \varphi(\varepsilon x)f(x) = \sum_{k\in\mathbf{Z}} c_k e^{2\pi ikx/T}\varphi(\varepsilon x)$$

where c_k are the Fourier coefficients of f. It is clear that $f_\varepsilon \in \mathscr{S}$, and the Fourier transform is

$$\xi \mapsto \sum_{k\in\mathbf{Z}} c_k \hat{\varphi}((\xi - 2\pi k/T)/\varepsilon)/\varepsilon.$$

The term with index k has support in $\{\xi; |\xi - 2\pi k/T| < \varepsilon\}$. If $T\varepsilon < \pi$ and $c_0 = 0$ it follows that the conjugate function of f_ε is

$$\tilde{f}_\varepsilon(x) = \varphi(\varepsilon x)\tilde{f}(x).$$

With the notation C_p for the constant in (4.1.5) we obtain

$$\int_0^T |\tilde{f}(x)|^p \sum_{j\in\mathbf{Z}} |\varphi(\varepsilon(x+jT))|^p\,dx \le C_p^p \int_0^T |f(x)|^p \sum_{j\in\mathbf{Z}} |\varphi(\varepsilon(x+jT))|^p\,dx.$$

After multiplication by εT the sums converge to $\int |\varphi(y)|^p\,dy$ as $\varepsilon \to 0$, so we obtain

$$\|\tilde{f}\|_{L^p(\mathbf{R}/T\mathbf{Z})} \le C_p \|f\|_{L^p(\mathbf{R}/T\mathbf{Z})},$$

provided that $c_0 = 0$. Now the L^p norm in $\mathbf{R}/T\mathbf{Z}$ of the mean value c_0 cannot exceed that of f, so it follows that

$$\|\tilde{f}\|_{L^p(\mathbf{R}/T\mathbf{Z})} \le 2C_p \|f\|_{L^p(\mathbf{R}/T\mathbf{Z})},$$

when f is periodic with period T, first when $f \in C^\infty$ and then by approximation when $f \in L^p(\mathbf{R}/T\mathbf{Z})$. Hence we have proved:

Corollary 4.1.5. *If $f \in L^p(\mathbf{R}/T\mathbf{Z})$ and $1 < p < \infty$, then the partial sums $s_n = \sum_{|k|\le n} c_k e^{2\pi i k x/T}$ of the Fourier series converge to f in $L^p(\mathbf{R}/T\mathbf{Z})$ as $n \to \infty$.*

Remark. Note that it is not claimed that there is convergence for an arbitrary order of summation of the terms. Nor is it stated in Corollaries 4.1.4 and 4.1.5 that there is pointwise convergence almost everywhere which is true but much more difficult to prove.

If $f \in L^1(\mathbf{R})$ we can still define a conjugate *distribution* $\tilde{f}$ as the inverse Fourier transform of the function $\xi \mapsto \hat{f}(\xi)\operatorname{sgn}\xi$, and $L^1(\mathbf{R}) \ni f \mapsto \tilde{f} \in \mathscr{S}'(\mathbf{R})$ is continuous. However, it was clear already from (4.1.2) that $\tilde{f}$ is usually not in L^1 even if $f \in C_0^\infty(\mathbf{R})$. The problem is not only that $\tilde{f}$ may be too large at infinity, as seen from (4.1.2), but there is a problem with local regularity too:

Proposition 4.1.6. *For all $f \in L^1(\mathbf{R})$ outside a certain set of first category the conjugate distribution $\tilde{f}$ is of positive order in every open subset of $\mathbf{R}$.*

Proof. Let $I = (x_0 - a, x_0 + a)$ be a bounded open non-empty interval $\subset \mathbf{R}$, and denote by B the set of all $f \in L^1(\mathbf{R})$ such that the restriction of $\tilde{f}$ to I is a bounded measure. B is a Banach space with $\|f\|_B$ equal to the sum of $\|f\|_1$ and the total mass of $\tilde{f}$ in I. If the range of the obvious injection $B \to L^1(\mathbf{R})$ is not of the first category, then it follows from Banach's theorem that the injection is an isomorphism, hence that there is a constant C such that

$$\int_I |\tilde{f}|\,dx \le C \int_{\mathbf{R}} |f|\,dx, \tag{4.1.13}$$

if $f \in L^1(\mathbf{R})$ and $\tilde f \in L^1(\mathbf{R})$. With $\varphi \in C_0^\infty(\mathbf{R})$, $\int \varphi\, dx = 1$, we apply (4.1.13) to $f_\varepsilon(x) = \varphi((x-x_0)/\varepsilon)/\varepsilon$. Then $\tilde f_\varepsilon(x) = \tilde\varphi((x-x_0)/\varepsilon)/\varepsilon$, and it follows from (4.1.2) that

$$\int_I |\tilde f_\varepsilon(x)|\, dx = \int_{-a}^{a} |\tilde\varphi(x/\varepsilon)|\, dx/\varepsilon = \int_{-a/\varepsilon}^{a/\varepsilon} |\tilde\varphi(x)|\, dx = -(2/\pi)\log\varepsilon + O(1), \quad \text{as } \varepsilon \to 0.$$

Since $\int_{\mathbf{R}} |f_\varepsilon(x)|\, dx = \|\varphi\|_1$ is independent of ε this contradicts (4.1.13), so the set of all $f \in L^1(\mathbf{R})$ such that $\tilde f$ is a bounded measure in I is of the first category. Taking the union of these sets for all I with x_0 and a rational proves the proposition.

A sufficient condition for $\tilde f$ to be in L^1_{loc} is given by the following proposition, where we use the standard notation $\log^+ x = \log x$ when $x > 1$, $\log^+ x = 0$ when $x \le 1$.

Proposition 4.1.7. *If $f \in L^1(\mathbf{R})$ and $|f|\log^+|f| \in L^1(\mathbf{R})$, then $\varrho\tilde f \in L^1(\mathbf{R})$ if $\varrho \in L^q(\mathbf{R}) \cap L^\infty(\mathbf{R})$ for some $q < \infty$.*

Proof. Let $E_0 = \{x \in \mathbf{R}; |f(x)| < 1\}$ and $E_k = \{x \in \mathbf{R}; 2^{k-1} \le |f(x)| < 2^k\}$ for $k = 1, 2, \dots$; define $f_k(x) = f(x)$ when $x \in E_k$ and $f_k(x) = 0$ if $x \notin E_k$. Then $f_k \in L^p$ for $p \in [1, \infty]$, and if $1 < p \le 2$ we have by (4.1.5)

$$\|\tilde f_k\|_p \le C(p-1)^{-1}\|f_k\|_p \le C(p-1)^{-1}2^k m(E_k)^{1/p}.$$

Without restriction we may assume that $q \ge 2$. If $1 < p \le q/(q-1)$ it follows from Hölder's inequality that

$$\int_{\mathbf{R}} |\tilde f_k(x)||\varrho(x)|\, dx \le C(\|\varrho\|_\infty + \|\varrho\|_q)(p-1)^{-1}2^k m(E_k)^{1/p}.$$

The hypothesis means that $\sum_0^\infty (k+1)2^k m(E_k) < \infty$, so to estimate the sum of the preceding integrals we must aim for this sum. Thus we choose $1/p = 1 - 1/(q(k+1))$ and write the preceding estimate in the form

$$\int_{\mathbf{R}} |\tilde f_k(x)||\varrho(x)|\, dx \le C(\|\varrho\|_\infty + \|\varrho\|_q)q2^{2/q}\Big((k+1)2^{-k}\Big)^{1-\frac{1}{p}}\Big((k+1)2^k m(E_k)\Big)^{1/p}$$

where we have used that $2^{k(1/p-1)} = 2^{-k/(q(k+1))} \ge 2^{-1/q}$. By the inequality between geometric and arithmetic means, with weights $1 - 1/p$ and $1/p$ it follows that

$$\int_{\mathbf{R}} |\tilde f_k(x)||\varrho(x)|\, dx \le C(\|\varrho\|_\infty + \|\varrho\|_q)q2^{2/q}\big(2^{-k}/q + (k+1)2^k m(E_k)\big).$$

Hence $\sum_0^\infty \tilde f_k$ converges in L^1_{loc}, so $\tilde f = \sum_0^\infty \tilde f_k \in L^1_{\text{loc}}$ and $\tilde f\varrho \in L^1$. The proof is complete.

Looking for substitutes for the estimate (4.1.5) when $p = 1$ may seem to be of marginal interest at first sight. However, it has played an important role in the development of harmonic analysis during the past 30 years and led to important techniques which cannot be ignored. We shall therefore pursue the matter further in the one dimensional case as an introduction to the case of higher dimension in the following sections.

DEFINITION 4.1.8. The *Hardy space* $\mathscr{H}^1(\mathbf{R})$ is the space of all $f \in L^1(\mathbf{R})$ with $\tilde{f} \in L^1(\mathbf{R})$.

PROPOSITION 4.1.9. *$\mathscr{H}^1(\mathbf{R})$ is a Banach space with the norm $\|f\|_{H^1} = \|f\|_{L^1} + \|\tilde{f}\|_{L^1}$, and it is invariant under the map $f \mapsto \tilde{f}$ and complex conjugation, both of which preserve the norm. When $f \in \mathscr{H}^1(\mathbf{R})$ the analytic functions*

$$F_\pm(z) = \pm\frac{i}{\pi}\int_{\mathbf{R}} \frac{f(t)}{z-t}\,dt, \quad \pm\operatorname{Im} z > 0, \tag{4.1.14}$$

have boundary values $f \pm i\tilde{f}$ in the L^1 sense,

$$\lim_{y\to\pm 0}\int_{\mathbf{R}} |F_\pm(x+iy) - f(x) \mp i\tilde{f}(x)|\,dx = 0, \quad \int_{\mathbf{R}} |F_\pm(x+iy)|\,dx \le \|f \pm i\tilde{f}\|, \quad \pm y > 0. \tag{4.1.15}$$

If $f \in L^1(\mathbf{R})$ and $\hat{f}$ has compact support not containing the origin, then $f \in \mathscr{H}^1(\mathbf{R})$. Such functions in $\mathscr{S}(\mathbf{R})$ are dense in $\mathscr{H}^1(\mathbf{R})$, and the closure of $\mathscr{H}^1(\mathbf{R})$ in $L^1(\mathbf{R})$ is the hyperplane $\{f \in L^1(\mathbf{R}); \hat{f}(0) = 0\}$.

PROOF. Since the map $L^1(\mathbf{R}) \ni f \mapsto \tilde{f} \in \mathscr{S}'(\mathbf{R})$ is continuous, it follows that $\mathscr{H}^1(\mathbf{R})$ is complete, for if $f_j \to f$ and $\tilde{f}_j \to g$ in $L^1(\mathbf{R})$, then $g = \tilde{f}$. If $f \in \mathscr{H}^1(\mathbf{R})$ then $\tilde{\tilde{f}} = -f$, so $\tilde{f} \in \mathscr{H}^1(\mathbf{R})$ and $\|\tilde{f}\|_{\mathscr{H}^1} = \|f\|_{\mathscr{H}^1}$. Since the map $f \mapsto \tilde{f}$ commutes with complex conjugation we have $\bar{f} \in \mathscr{H}^1$ and $\|\bar{f}\|_{\mathscr{H}^1} = \|f\|_{\mathscr{H}^1}$ if $f \in \mathscr{H}^1(\mathbf{R})$. If $f \in \mathscr{H}^1(\mathbf{R})$ then $\hat{f}(\xi)$ and $-i\hat{f}(\xi)\operatorname{sgn}\xi$ are continuous, so $\hat{f}(0) = 0$, that is, $\int_{\mathbf{R}} f\,dx = 0$. On the other hand, if $f \in L^1(\mathbf{R})$ and $\operatorname{supp}\hat{f}$ is a compact subset of $\mathbf{R}\setminus\{0\}$ then we can choose $\varphi \in \mathscr{S}$ so that $\hat{\varphi} \in C_0^\infty(\mathbf{R})$ and $\hat{\varphi}(\xi) = -i\operatorname{sgn}\xi$ in a neighborhood of $\operatorname{supp}\hat{f}$. Then $\tilde{f} = \varphi * f \in L^1(\mathbf{R})$, so $f \in \mathscr{H}^1(\mathbf{R})$. Thus the condition for a function in $L^1(\mathbf{R})$ to be in $\mathscr{H}^1(\mathbf{R})$ is only a restriction on the behavior of $\hat{f}$ at 0 and at ∞.

Choose $\chi \in \mathscr{S}(\mathbf{R})$ so that $\hat{\chi} \in C_0^\infty(\mathbf{R})$ and $\hat{\chi} = 1$ in a neighborhood of the origin, and set $\chi_\varepsilon(x) = \varepsilon\chi(\varepsilon x)$. Then

$$\lim_{\varepsilon\to 0}\|\chi_\varepsilon * f\|_{L^1} = |\hat{f}(0)|\|\chi\|_{L^1}, \quad \lim_{t\to\infty}\|\chi_t * f - f\|_{L^1} = 0, \quad f \in L^1(\mathbf{R}). \tag{4.1.16}$$

Since $\|\chi_t * f\|_{L^1} \le \|f\|_{L^1}\|\chi\|_{L^1}$, it is sufficient to prove (4.1.16) for all f in a dense subset of L^1 such as all $f \in \mathscr{S}$ with $\hat{f} \in C_0^\infty(\mathbf{R})$. The Fourier transform of $\chi_t * f$ is $\hat{\chi}(\xi/t)\hat{f}(\xi) = \hat{f}(\xi)$ for large t, so the second part of (4.1.16) is trivial then. To prove the first part we write the Fourier transform of $\chi_\varepsilon * f$ as

$$\hat{\chi}(\xi/\varepsilon)\hat{f}(\xi) = \hat{\chi}(\xi/\varepsilon)\hat{f}(0) + \hat{g}_\varepsilon(\xi), \quad \hat{g}_\varepsilon(\xi) = \hat{\chi}(\xi/\varepsilon)(\hat{f}(\xi) - \hat{f}(0)),$$

and conclude that $\int |\hat{g}_\varepsilon(\xi)|^2\,d\xi \le C\varepsilon^3$, $\int |d\hat{g}_\varepsilon(\xi)/d\xi|^2\,d\xi \le C\varepsilon$, which by Parseval's formula and Cauchy-Schwarz gives

$$\int (1 + \varepsilon^2 x^2)|g_\varepsilon(x)|^2\,dx \le C\varepsilon^3/\pi, \quad \int |g_\varepsilon(x)|\,dx \le \sqrt{C}\varepsilon,$$

and proves (4.1.16).

If $f \in L^1(\mathbf{R})$ and $\hat{f}(0) = 0$ it follows with $f_{t,\varepsilon} = \chi_t * (f - \chi_\varepsilon * f)$ that $\|f_{t,\varepsilon} - f\|_{L^1} \to 0$ as $t \to \infty$ and $\varepsilon \to 0$. The Fourier transform $\xi \mapsto \hat{\chi}(\xi/t)(1 - \hat{\chi}(\xi/\varepsilon))\hat{f}(\xi)$ of $f_{t,\varepsilon}$ has compact support in $\mathbf{R} \setminus \{0\}$, so $f_{t,\varepsilon} \in \mathscr{H}^1(\mathbf{R})$. This proves that the closure of $\mathscr{H}^1(\mathbf{R})$ in $L^1(\mathbf{R})$ consists of all $f \in L^1(\mathbf{R})$ with $\hat{f}(0) = 0$. (It is an easy exercise to prove this using the Hahn-Banach theorem.) If $\tilde{f} \in L^1(\mathbf{R})$ then $(\tilde{f})_{t,\varepsilon}$ is the conjugate function of $f_{t,\varepsilon}$ and $\|(\tilde{f})_{t,\varepsilon} - \tilde{f}\|_{L^1} \to 0$, hence $f_{t,\varepsilon} \to f$ in $\mathscr{H}^1$ as $\varepsilon \to 0$ and $t \to \infty$. To achieve fast decrease at infinity we choose $\varphi \in \mathscr{S}(\mathbf{R})$ with $\varphi(0) = 1$ and $\hat{\varphi} \in C_0^\infty$, and set $\varphi^\delta(x) = \varphi(\delta x)$. The Fourier transform of $\varphi^\delta f_{t,\varepsilon}$ is then the convolution of $\hat{\varphi}(\xi/\delta)/2\pi\delta$ and $\hat{f}_{t,\varepsilon}$, and since $\hat{\varphi} \in C_0^\infty(\mathbf{R})$ the support of the convolution does not contain the origin for small δ. To multiply the convolution by $\operatorname{sgn}\xi$ is then equivalent to multiplying $\hat{f}_{t,\varepsilon}(\xi)$ by $\operatorname{sgn}\xi$ before the convolution, that is, the conjugate function of $\varphi^\delta f_{t,\varepsilon}$ is $\varphi^\delta \tilde{f}_{t,\varepsilon}$, so $\varphi^\delta f_{t,\varepsilon} \to f_{t,\varepsilon}$ in $\mathscr{H}^1(\mathbf{R})$ as $\delta \to 0$. Since $\varphi^\delta f_{t,\varepsilon} \in \mathscr{S}(\mathbf{R})$ and the Fourier transform has compact support in $\mathbf{R} \setminus \{0\}$, this proves the density statement in the proposition.

When $\hat{f} \in C_0^\infty(\mathbf{R} \setminus \{0\})$ the Fourier transform of $x \mapsto F_\pm(x + iy)$ where $\pm y > 0$ is equal to

$$\hat{f}(\xi)(1 \pm \operatorname{sgn}\xi)e^{-y\xi} \to \hat{f}(\xi)(1 \pm \operatorname{sgn}\xi) = \hat{f}(\xi) \pm i(-i\hat{f}(\xi)\operatorname{sgn}\xi) \quad \text{in } \mathscr{S} \text{ as } y \to \pm 0.$$

This proves the first part of (4.1.15) for f in a dense subset of $\mathscr{H}^1(\mathbf{R})$, and the second part is then a consequence. In fact, if $\psi \in C_0^\infty(\mathbf{R})$, $|\psi| \le 1$, then $\int F_\pm(z + x)\psi(x)\,dx$ is analytic when $\pm \operatorname{Im} z > 0$, continuous in the closed half plane with boundary values bounded by $\|f \pm i\tilde{f}\|_{L^1}$, and tends to 0 at ∞. Hence it follows from the maximum principle that it is always bounded by $\|f \pm i\tilde{f}\|_{L^1}$, which proves the second part of (4.1.15) for all f in a dense subset of $\mathscr{H}^1(\mathbf{R})$. The bound follows by continuity for all $f \in \mathscr{H}^1(\mathbf{R})$, and the first part of (4.1.15) also follows then for all $f \in \mathscr{H}^1(\mathbf{R})$. The proof is complete.

The following lemma gives an important subset of $\mathscr{H}^1$. The simple proof was already used to prove (4.1.16).

Lemma 4.1.10. *If $f \in L^2(\mathbf{R})$ then*

(4.1.17)
$$\int_{\mathbf{R}} (1 + x^2/\delta^2)|f(x)|^2\,dx = M^2 < \infty, \ \int_{\mathbf{R}} f(x)\,dx = 0 \implies f \in \mathscr{H}^1, \ \|f\|_{\mathscr{H}^1} \le 2M\sqrt{\delta\pi}.$$

Proof. Cauchy-Schwarz' inequality gives $\|f\|_1 \le M\sqrt{\delta\pi}$, and by Parseval's formula $\hat{f}$ and $d\hat{f}/d\xi$ are in L^2 and

$$\frac{1}{2\pi}\int (|\hat{f}(\xi)|^2 + |d\hat{f}(\xi)/d\xi|^2/\delta^2)\,d\xi = M^2.$$

For $g = \tilde{f}$ we have $\hat{g}(\xi) = -i\operatorname{sgn}\xi\hat{f}(\xi)$, and since $\hat{f}(0) = 0$ it follows that $d\hat{g}(\xi)/d\xi = -i\operatorname{sgn}\xi d\hat{f}(\xi)/d\xi$. (Otherwise there would have been additional term $-2i\hat{f}(0)\delta_0$.) Hence Parseval's formula again gives $\int_{\mathbf{R}}(1 + x^2/\delta^2)|g(x)|^2\,dx = M^2$, so $\|\tilde{f}\|_1 = \|g\|_1 \le M\sqrt{\delta\pi}$, which proves (4.1.17).

If L is a continuous linear form on $\mathscr{H}^1(\mathbf{R})$, it follows from Lemma 4.1.10 that L restricts to a continuous linear form on $\{f \in L^2(\mathbf{R}, (1 + x^2/\delta^2)\,dx); \int_{\mathbf{R}} f\,dx = 0\}$ with norm $\le 2\sqrt{\delta\pi}\|L\|_{(\mathscr{H}^1)'}$. By Proposition 4.1.9 this is a dense subset of $\mathscr{H}^1(\mathbf{R})$. We extend the restriction uniquely to $L^2(\mathbf{R}, (1 + x^2/\delta^2)\,dx)$ so that the extension is equal to 0 for $x \mapsto (1 + x^2/\delta^2)^{-1}$, which does not increase the norm. In view of the translation invariance of $\mathscr{H}^1(\mathbf{R})$ it follows that for arbitrary $\delta > 0$ and $y \in \mathbf{R}$ there exists a function $\psi_{y,\delta} \in L^2_{\text{loc}}(\mathbf{R})$ such that

$$\delta \int_{\mathbf{R}} \frac{|\psi_{y,\delta}(x)|^2}{\delta^2 + |x-y|^2}\,dx \le 4\pi\|L\|^2_{(\mathscr{H}^1)'},$$
$$L(f) = \int_{\mathbf{R}} \psi_{y,\delta}(x) f(x)\,dx, \quad \text{if } f \in L^2(\mathbf{R}, (1+x^2)\,dx),\ \hat{f}(0) = 0. \tag{4.1.18}$$

Here $\psi_{y,\delta}$ was uniquely determined by the condition $\int_{\mathbf{R}} \psi_{y,\delta}(x)\,dx/(\delta^2 + (x-y)^2) = 0$. However, it is usually more convenient to use another normalisation which involves only the values of $\psi_{y,\delta}$ in the interval $(y-\delta, y+\delta)$. If $c_{y,\delta} = \int_{y-\delta}^{y+\delta} \psi_{y,\delta}(x)/2\delta$ then $\psi^0_{y,\delta} = \psi_{y,\delta} - c_{y,\delta}$ also defines L, the mean value over $(y-\delta, y+\delta)$ vanishes, and

$$\frac{1}{2\delta} \int_{y-\delta}^{y+\delta} |\psi^0_{y,\delta}(x)|^2\,dx \le 4\pi\|L\|^2_{(\mathscr{H}^1)'}. \tag{4.1.19}$$

Definition 4.1.11. A function $\varphi \in L^2_{\text{loc}}(\mathbf{R})$ is said to be in BMO($\mathbf{R}$) if there is a constant B such that for $y \in \mathbf{R}$ and $\delta > 0$

$$\frac{1}{2\delta} \int_{y-\delta}^{y+\delta} |\varphi(x) - \varphi_{y,\delta}|^2\,dx \le B^2, \quad \text{where } \varphi_{y,\delta} = \frac{1}{2\delta} \int_{y-\delta}^{y+\delta} \varphi(t)\,dt. \tag{4.1.20}$$

Thus we have proved that every continuous linear form on $\mathscr{H}^1(\mathbf{R})$ is defined by a function in BMO($\mathbf{R}$); the converse will be proved below. It may seem strange that in (4.1.20) we have dropped the global information contained in (4.1.18), but that is only apparent since it can be recovered from (4.1.20):

Proposition 4.1.12. *BMO($\mathbf{R}$)/$\mathbf{C}$ is a Banach space with norm equal to the smallest constant B such that* (4.1.20) *is valid, and* (4.1.20) *implies*

$$\delta \int_{\mathbf{R}} \frac{|\varphi(x) - \varphi_{y,\delta}|^2}{(x-y)^2 + \delta^2}\,dx \le 210B^2, \quad y \in \mathbf{R},\ \delta > 0. \tag{4.1.20$'$}$$

Proof. For fixed y and δ let $c_k = 2^{-k-1}\delta^{-1} \int_{|x-y|<2^k\delta} \varphi(x)\,dx$, $k = 0, 1, \dots,$ thus $c_0 = \varphi_{y,\delta}$. By (4.1.20) applied with δ replaced by $2^{k+1}\delta$

$$\frac{1}{2^{k+1}\delta} \int_{|x-y|<2^k\delta} |\varphi(x) - c_{k+1}|^2\,dx \le 2B^2, \quad \text{hence } |c_k - c_{k+1}| \le \sqrt{2}B,$$

which implies that $|c_k - c_0| \le \sqrt{2}kB$. By the triangle inequality it follows that

$$\int_{|x-y|<2^k\delta} |\varphi(x) - c_0|^2\,dx \le 2B^2(2^{k+1}\delta + 2k^2 2^{k+1}\delta) = B^2 2^{k+2}\delta(1+2k^2).$$

Summing up we obtain

$$\delta\int_{\mathbf{R}} \frac{|\varphi(x)-c_0|^2}{(x-y)^2+\delta^2}\,dx \le \frac{1}{\delta}\int_{|x-y|<\delta} |\varphi(x)-c_0|^2\,dx + \sum_1^\infty \frac{2^{2-2k}}{\delta}\int_{|x-y|<2^k\delta} |\varphi(x)-c_0|^2\,dx$$
$$\le 2B^2(1 + 8\sum_1^\infty 2^{-k}(1+2k^2)) = 210B^2.$$

This proves (4.1.20)$'$. It is obvious that the minimal B in (4.1.20) is a norm $\|\varphi\|_{\mathrm{BMO}}$ in $\mathrm{BMO}(\mathbf{R})/\mathbf{C}$. To prove completeness we consider a Cauchy sequence $\varphi_j \in \mathrm{BMO}(\mathbf{R})$. Without restriction we may assume that $\int_{-1}^1 \varphi_j(x)\,dx = 0$ for every j. Then it follows from (4.1.20)$'$ applied to $\varphi_j - \varphi_k$ that we have a Cauchy sequence in $L^2(\mathbf{R}, dx/(1+x^2))$. The limit φ there is obviously in $\mathrm{BMO}(\mathbf{R})$, and $\|\varphi - \varphi_j\|_{\mathrm{BMO}} \le \lim_{k\to\infty} \|\varphi_k - \varphi_j\|_{\mathrm{BMO}}$, which completes the proof.

To prove that conversely every $\varphi \in \mathrm{BMO}(\mathbf{R})$ defines a continuous linear form on $\mathscr{H}^1$ we shall study the Poisson integral

$$\Phi(t,x) = \frac{t}{\pi}\int_{\mathbf{R}} \frac{\varphi(y)}{(y-x)^2+t^2}\,dy = \frac{1}{\pi}\int_{\mathbf{R}} \frac{\varphi(x+ty)}{y^2+1}\,dy, \tag{4.1.21}$$

which is harmonic in $\{(t,x) \in \mathbf{R}^2; t > 0\}$ and continuous in the closed half space with boundary values φ when φ is continuous. When φ is a constant c then $\Phi = c$, and since we are interested in $\mathrm{BMO}/\mathbf{C}$ it is natural to focus attention on the derivatives of Φ rather than on Φ to remove constant terms.

Lemma 4.1.13. *If $\varphi \in L^2(\mathbf{R})$ and Φ is the Poisson integral* (4.1.21) *then*

$$2\iint_{t>0} t|\Phi'(t,x)|^2\,dx\,dt = \|\varphi\|_2^2, \quad |\Phi'(t,x)|^2 \le \|\varphi\|_2^2/\pi t^3. \tag{4.1.22}$$

Here $|\Phi'(t,x)|^2 = |\partial\Phi(t,x)/\partial t|^2 + |\partial\Phi(t,x)/\partial x|^2$. If $\varphi(x) = x\psi(x)$ where $\psi \in L^2$, then

$$|\Phi'(t,x)|^2 \le \|\psi\|_2^2(x^2+t^2)/\pi t^3. \tag{4.1.23}$$

If $\varphi \in \mathrm{BMO}(\mathbf{R})$ then for $y \in \mathbf{R}$ and $\delta > 0$

$$\iint_{T_{y,\delta}} t|\Phi'(t,x)|^2\,dx\,dt \le 603\delta\|\varphi\|_{\mathrm{BMO}}^2, \quad T_{y,\delta} = \{(t,x); 0<t<2\delta, |x-y|<\delta\}. \tag{4.1.24}$$

Proof. We may assume that φ is real valued. It suffices to prove (4.1.22) when $\varphi \in C_0^\infty$. Then it is obvious that $\Phi(t,x)$ is in C^∞ for $t \ge 0$ with $\Phi(0,x) = \varphi(x)$. At infinity

$\Phi(t,x)=O(t/(x^2+t^2))$ and the derivatives of $\Phi(t,x)$ are $O(1/(t^2+x^2))$. Since $\Delta\Phi=0$ we obtain by partial integration

$$\iint_{t>0} t|\Phi'(t,x)|^2\,dx\,dt = -\iint_{t>0}\partial\Phi(t,x)/\partial t\Phi(t,x)\,dx\,dt = \tfrac12\int_{\mathbf{R}}|\Phi(0,x)|^2\,dx = \tfrac12\|\varphi\|_2^2.$$

From the fact that

$$\Phi(t,x)=\operatorname{Im}\frac{1}{\pi}\int_{\mathbf{R}}\frac{\varphi(y)}{y-x-it}\,dy,$$

it follows that

$$\partial\Phi(t,x)/\partial t + i\partial\Phi(t,x)/\partial x = \frac{1}{\pi}\int_{\mathbf{R}}\frac{\varphi(y)}{(y-x-it)^2}\,dy,$$

so Cauchy-Schwarz' inequality gives

$$|\Phi'(t,x)|^2 \le \frac{\|\varphi\|_2^2}{\pi^2}\int_{\mathbf{R}}\frac{dy}{((y-x)^2+t^2)^2} = \frac{\|\varphi\|_2^2}{\pi^2 t^3}\int_{\mathbf{R}}\frac{dy}{(y^2+1)^2} \le \|\varphi\|_2^2/\pi t^3.$$

This proves (4.1.22). If $\varphi(x)=x\psi(x)$ with $\psi\in C_0^\infty$ we just replace $\varphi(y)$ by $y\psi(y)$ above before using the Cauchy-Schwarz inequality, and obtain

$$|\Phi'(t,x)|^2 \le \frac{\|\psi\|_2^2}{\pi^2}\int_{\mathbf{R}}\frac{y^2\,dy}{((y-x)^2+t^2)^2} = \frac{\|\psi\|_2^2}{\pi^2}\int_{\mathbf{R}}\frac{(y^2+x^2)\,dy}{(y^2+t^2)^2} \le \frac{\|\psi\|_2^2}{\pi}\Big(\frac{1}{t}+\frac{x^2}{t^3}\Big),$$

which proves (4.1.23). Note that when $t\to 0$ this bound is much better if $x=0$, and this is the only case that we shall use.

By the translation invariance of BMO($\mathbf{R}$) it suffices to prove (4.1.24) when $y=0$. Write $\varphi(x)=\varphi_{0,2\delta}+\varphi_0(x)+\varphi_1(x)$ where $\varphi_{0,2\delta}$ is the mean value of φ in $(-2\delta,2\delta)$ and

$$\varphi_0(x)=\begin{cases}\varphi(x)-\varphi_{0,2\delta}, & \text{when } |x|<2\delta,\\ 0, & \text{when } |x|\ge 2\delta\ ,\end{cases}\qquad \varphi_1(x)=\begin{cases}0, & \text{when } |x|<2\delta,\\ \varphi(x)-\varphi_{0,2\delta}, & \text{when } |x|\ge 2\delta.\end{cases}$$

By (4.1.20) and (4.1.20)$'$ we have, with $B=\|\varphi\|_{\mathrm{BMO}}$,

$$\frac{1}{4\delta}\int_{\mathbf{R}}|\varphi_0(x)|^2\,dx\le B^2,\quad \delta\int_{\mathbf{R}}|\varphi_1(x)|^2/x^2\,dx\le 105B^2.$$

The Poisson integral of $\varphi_{0,2\delta}$ is a constant which does not contribute to (4.1.24). Let Φ_0 and Φ_1 be the Poisson integrals of φ_0 and φ_1. By (4.1.22) we have

$$\iint_{T_{0,\delta}} t|\Phi_0'(t,x)|^2\,dx\,dt\le 2\delta B^2.$$

If $|y|<\delta$ then $\delta\int_{\mathbf{R}}|\varphi_1(x+y)|^2/x^2\,dx\le 420B^2$, and it follows from (4.1.23) that

$$t|\Phi_1'(t,y)|^2\le\frac{420B^2}{\pi\delta},\ |y|<\delta;\quad \text{hence}\ \iint_{T_{0,\delta}} t|\Phi'(t,x)|^2\,dx\,dt\le\frac{1680}{\pi}\delta B^2.$$

The triangle inequality completes the proof of (4.1.24).

We have now arrived at the final step in the proof that BMO($\mathbf{R}$) is the dual of $\mathscr{H}^1$.

PROPOSITION 4.1.14. *Let $\varphi \in L^2(\mathbf{R}, dx/(1+x^2))$, and assume that for the Poisson integral Φ defined by* (4.1.21) *we have*

$$\iint_{T_{y,\delta}} t|\Phi'(t,x)|^2\,dx\,dt \le 2A^2\delta, \quad y\in\mathbf{R},\ \delta>0, \tag{4.1.25}$$

where $T_{y,\delta}$ is defined by (4.1.24). *Then it follows that*

$$\Big|\int_{\mathbf{R}} \varphi f\,dx\Big| \le 16A\|f\|_{\mathscr{H}^1}, \quad \text{if } f\in\mathscr{S}, \hat f\in C_0^\infty(\mathbf{R}\setminus\{0\}), \tag{4.1.26}$$

so φ defines a continuous linear form on $\mathscr{H}^1$ with norm $\le 16A$.

PROOF. We may assume that φ and f are real valued. Polarization of (4.1.22) gives if $\varphi\in L^2(\mathbf{R})$

$$\int \varphi f\,dx = 2\iint_{t>0} t(\Phi'(t,x), F'(t,x))\,dx\,dt, \tag{4.1.27}$$

where F is the Poisson integral of f. When φ is bounded in $L^2(\mathbf{R}, dx/(1+x^2))$ we can write $\varphi(x) = \varphi_0(x) + x\varphi_1(x)$ with φ_0 and φ_1 bounded in L^2 and conclude from Lemma 4.1.13 that $|\Phi'(t,x)| = O((1+|x|+|t|)t^{-3/2})$. Since $|F'(t,x)| = O(e^{-ct}(1+x^2)^{-N})$ for some $c>0$ and all N, because $\hat f\in C_0^\infty(\mathbf{R}\setminus\{0\})$, it follows that both sides of (4.1.27) are continuous functions of $\varphi\in L^2(\mathbf{R}, dx/(1+x^2))$, so the formula follows for such φ from the case where $\varphi\in L^2(\mathbf{R})$.

The Poisson integral $F_+(t,x)$ of $f+i\tilde f$ is an analytic function of $x+it$ in the upper half plane with real part F, so Cauchy-Schwarz' inequality and (4.1.27) give

$$\begin{aligned}\Big|\int_{\mathbf{R}} \varphi f\,dx\Big| &\le 2\Big|\iint_{t>0} t(\Phi'(t,x), F_+'(t,x))\,dx\,dt\Big| \\ &\le 2\Big(\iint_{t>0} t|\Phi'(t,x)|^2|F_+(t,x)|\,dx\,dt\Big)^{\frac12}\Big(\iint_{t>0} t|F_+'(t,x)|^2|F_+(t,x)|^{-1}\,dx\,dt\Big)^{\frac12}.\end{aligned} \tag{4.1.28}$$

The second factor can be simplified since $|F_+'(t,x)|^2 = |F_+(t,x)|\Delta|F_+(t,x)|$, which follows from the analyticity of F_+ and fact that for the function $\mathbf{C}\ni w\mapsto |w|$ we have $|w|\Delta|w| = 1$ by the expression for Δ in polar coordinates. (It suffices to observe that this is true when $F_+(t,x)\ne 0$, for zeros are discrete. Note that $\Delta|F_+|$ is bounded except at simple zeros of F_+ where it may grow as the reciprocal of the distance to the zero.) Thus the second parenthesis in (4.1.28) is equal to

$$\iint_{t>0} t\Delta|F_+(t,x)|\,dx.$$

A formal integration by parts gives that this is equal to

$$\int_{\mathbf{R}} |F_+(0,x)|\,dx \le \|f\|_1 + \|\tilde f\|_1 = \|f\|_{\mathscr{H}^1}.$$

To argue rigorously we choose $\chi \in C_0^\infty(\mathbf{R}^2)$ so that $\chi \ge 0$, $\int \chi(t,x)\,dx\,dt = 1$ and $\chi(t,x)$ only depends on t^2+x^2. Set $\chi_\varepsilon(t,x) = \varepsilon^{-2}\chi(t/\varepsilon, x/\varepsilon)$. Since $F_+(t,x)$ is an entire function of $x+it$ which is rapidly decreasing when $t > -1$, say, it follows that $\chi_\varepsilon * |F_+|$ is in C^∞ and rapidly decreasing for $t \ge 0$ when ε is small enough. Hence

$$\iint_{t>0} t((\Delta|F_+|) * \chi_\varepsilon)\,dx\,dt = \iint_{t>0} t\Delta(|F_+| * \chi_\varepsilon)\,dx\,dt = \int_{\mathbf{R}} (|F_+| * \chi_\varepsilon)(0,x)\,dx.$$

The integrand in the right-hand side decreases to $|F_+(0,x)|$ as $\varepsilon \downarrow 0$, because $|F_+|$ is subharmonic. The integrand on the left is non-negative so Fatou's lemma gives

$$\iint_{t>0} t\Delta|F_+|\,dx\,dt \le \int_R |F_+(0,x)|\,dx \le \|f\|_{\mathscr{H}^1}. \tag{4.1.29}$$

To estimate the first factor in the right-hand side of (4.1.28) we note that

$$\sqrt{|F_+(t,x)|} = \exp(\tfrac{1}{2}\log|F_+(t,x)|)$$

is subharmonic and $\to 0$ at ∞. Hence $\sqrt{|F_+(t,x)|} \le G(t,x)$ where G is the Poisson integral of $g(x) = \sqrt{|F_+(0,x)|}$. If the Poisson kernel $y \mapsto t/(\pi(y^2+t^2))$ is replaced by the value $1/\pi t$ at the origin for $|y| < t$ then the integral is < 2, so it follows from $(4.1.6)'''$ that $G(t,x) \le 2g^{**}_{\mathrm{HL}}(x,t)$. By (4.1.25) the measure $t|\Phi'(t,x)|^2/A^2$ satisfies the hypotheses of Theorem 4.1.2$'$. With $p=2$ in $(4.1.8)'$ the first parenthesis in (4.1.28) can be estimated by $2^5\|Ag\|_2^2 = 2^5A^2\|F_+(0,\cdot)\|_1 \le 2^5A^2\|f\|_{\mathscr{H}^1}$. In view of (4.1.29) this completes the proof of (4.1.26).

We shall now sum up the results obtained on the duality of $\mathscr{H}^1(\mathbf{R})$ and $\mathrm{BMO}(\mathbf{R})$:

Theorem 4.1.15. *The restriction of a continuous linear form L on $\mathscr{H}^1(\mathbf{R})$ to the dense subset $\{f \in \mathscr{S}; \hat{f} \in C_0^\infty(\mathbf{R}\setminus\{0\})\}$ is of the form $L(f) = \int f\varphi\,dx$ where φ is uniquely determined in* $\mathrm{BMO}(\mathbf{R})/\mathbf{C}$ *and*

$$\|L\|_{(H^1)'}/278 \le \|\varphi\|_{\mathrm{BMO}} \le 4\|L\|_{(\mathscr{H}^1)'}. \tag{4.1.30}$$

Every $\varphi \in \mathrm{BMO}(\mathbf{R})$ defines a continuous linear form in $\mathscr{H}^1(\mathbf{R})$.

Proof. The upper bound in (4.1.30) follows from (4.1.19), and the lower bound is a combination of (4.1.24) and (4.1.26), for $16\sqrt{603/2} < 278$.

The duality established in Theorem 4.1.15 can be rephrased as the "atomic decomposition" of $\mathscr{H}^1(\mathbf{R})$. To state it we need a definition.

Definition 4.1.16. An atom in $\mathscr{H}^1(\mathbf{R})$ is a function $a \in L^2(\mathbf{R})$ with support in a compact interval I such that

$$m(I)\int_I |a(x)|^2\,dx \le 1, \qquad \int_I a(x)\,dx = 0. \tag{4.1.31}$$

By the Cauchy-Schwarz inequality (4.1.31) implies $\|a\|_1 \le 1$, and it follows from Lemma 4.1.10 that $\|a\|_{\mathscr{H}^1} \le 4$. In fact, by the translation invariance of $\mathscr{H}^1$ we may assume that $I = [-\delta,\delta]$, and then we have

$$\int (1+x^2/\delta^2)|a(x)|^2\,dx \le 1/\delta,$$

so $\|a\|_{\mathscr{H}^1} \le 2\sqrt{\pi} < 4$.

COROLLARY 4.1.17. *If B is the closed convex hull of the atoms in $\mathscr{H}^1(\mathbf{R})$ then*

$$\{f \in \mathscr{H}^1(\mathbf{R}); \|f\|_{\mathscr{H}^1} \le 1/278\} \subset B \subset \{f \in \mathscr{H}^1(\mathbf{R}); \|f\|_{\mathscr{H}^1} \le 4\}. \tag{4.1.32}$$

Every $f \in \mathscr{H}^1(\mathbf{R})$ has an atomic decomposition

$$f = \sum_1^\infty \lambda_j a_j, \quad \sum_1^\infty |\lambda_j| \le 279\|f\|_{\mathscr{H}^1}, \tag{4.1.33}$$

where a_j are atoms in $\mathscr{H}^1$; we have $\|f\|_{\mathscr{H}^1} \le 4\sum_1^\infty |\lambda_j|$.

PROOF. We have already proved that $\|f\|_{\mathscr{H}^1} \le 4$ if f is an atom, so this is also true in B. By the Hahn-Banach theorem we can describe B as the polar of the polar of the atoms. Thus let $L \in (\mathscr{H}^1)'$ and assume that $|L(a)| \le 1$ for all atoms a. If $\varphi \in \mathrm{BMO}(\mathbf{R})$ defines L this means that

$$\Big|\int_I a(x)\varphi(x)\,dx\Big| \le 1, \quad \text{if } \int_I a(x)\,dx = 0 \text{ and } m(I)\int_I |a(x)|^2\,dx \le 1, \quad \text{that is,}$$
$$\frac{1}{m(I)}\int_I |\varphi(x) - \varphi_I|^2\,dx \le 1, \quad \text{if } \varphi_I = \frac{1}{m(I)}\int_I \varphi(x)\,dx.$$

Thus $\|\varphi\|_{\mathrm{BMO}} \le 1$, and it follows from Theorem 4.1.15 that $\|L\|_{(\mathscr{H}^1)'} \le 278$, so $|L(f)| \le 1$ if $\|f\|_{\mathscr{H}^1} \le 1/278$. This proves (4.1.32).

Let $0 < \varepsilon < 1$. Given $f \in \mathscr{H}^1$ we can choose $\lambda_1, \dots, \lambda_j \in \mathbf{C}$ and atoms $a_1, \dots, a_j$ so that

$$\|f - \sum_1^j \lambda_\nu a_\nu\|_{\mathscr{H}^1} \le \varepsilon\|f\|_{\mathscr{H}^1}, \quad \sum_1^j |\lambda_\nu| \le 278\|f\|_{\mathscr{H}^1}.$$

We can repeat this argument with f replaced by the remainder $f - \sum_1^j \lambda_\nu a_\nu$. After an infinite number of iterations we obtain the decomposition (4.1.33) with

$$\sum_1^\infty |\lambda_\nu| \le 278\|f\|_{\mathscr{H}^1}(1 + \varepsilon + \varepsilon^2 + \dots) = 278\|f\|_{\mathscr{H}^1}/(1-\varepsilon).$$

With $\varepsilon = 1/279$ we obtain (4.1.33).

COROLLARY 4.1.18. *Let $\Phi = \{\varphi \in C^1(\mathbf{R}); (1+x^2)(|\varphi(x)| + |\varphi'(x)|) \le 1\}$. Set $\varphi_t(x) = \varphi(x/t)/t$ and $f^* = \sup_{\varphi\in\Phi}\sup_{t>0}|f * \varphi_t|$ for $f \in L^1(\mathbf{R})$. Then*

$$\|f^*\|_1 \le 6000\|f\|_{\mathscr{H}^1(\mathbf{R})}, \quad \textit{if } f \in \mathscr{H}^1(\mathbf{R}). \tag{4.1.34}$$

PROOF. First assume that f is an atom a; we may assume that $\operatorname{supp} a \subset [-\delta, \delta]$, that $\int a(x)\,dx = 0$ and that $2\delta\int |a(x)|^2\,dx \le 1$, thus $\int |a(x)|\,dx \le 1$. Since $|\varphi(x)| \le 1/(1+x^2)$

it follows from $(4.1.6)'$ that $a^*(x) \le \pi a^*_{\mathrm{HL}}(x)$, hence $\|a^*\|_2 \le 4\pi/\sqrt{2\delta}$ by (4.1.8). This gives an estimate for the L^1 norm on a finite interval, say

$$\int_{|x|<2\delta} a^*(x)\,dx \le 4\pi\sqrt{2}. \tag{4.1.35}$$

Now assume that $|x| > 2\delta$. To estimate

$$(a * \varphi_t)(x) = \frac{1}{t}\int_{-\delta}^{\delta} a(y)\varphi((x-y)/t)\,dy$$

we first use the bound $|\varphi((x-y)/t)| \le 1/(1+|x/2t|^2)$, $|y| < \delta$, and obtain

$$|(a * \varphi_t)(x)| \le \frac{4t}{4t^2+|x|^2} \le \frac{4\delta}{4\delta^2+|x|^2}, \quad t \le \delta.$$

For $t = |x|/2$ we would just get the bound $1/|x|$ which is not integrable at infinity. However, we can exploit the fact that $\int a(x)\,dx = 0$ by subtracting a term independent of y from $\varphi((x-y)/t)$ and using that when $|x| \ge 2\delta$ and $|y| < \delta$ then

$$|\varphi((x-y)/t) - \varphi(x/t)| \le \frac{|y|}{t}\frac{1}{1+|x/2t|^2}$$

by the hypothesis on φ'. This gives

$$|(a * \varphi_t(x)| \le \frac{4\delta}{4t^2+|x|^2} \le \frac{4\delta}{4\delta^2+|x|^2}, \quad t \ge \delta.$$

Since $\int_{|x|>2\delta} 4\delta\,dx/(4\delta^2+x^2) = \pi$, we have proved that $\|a^*\|_1 \le \pi(1+4\sqrt{2}) < 21$. For an atomic decomposition $f = \sum \lambda_j a_j$ we have $f^* \le \sum |\lambda_j| a_j^*$, so (4.1.34) follows from Corollary 4.1.17. The proof is complete.

The maximal function estimate in Corollary 4.1.18 is much more subtle than that in Theorem 4.1.2, for it takes cancellations into account whereas the Hardy-Littlewood maximal function only examines absolute values. There is an inverse of Corollary 4.1.18: If (4.1.34) is valid for a single fixed φ with $\int \varphi(x)\,dx \ne 0$, then $f \in \mathscr{H}^1$. This is interesting since it shows that the space $\mathscr{H}^1$ has a significance beyond the study of problems from analytic function theory. However, we shall not give the proof here and refer instead to Fefferman-Stein [1], Stein [2], and the references given there.

We introduced $\mathscr{H}^1$ as the largest subspace of $L^1(\mathbf{R})$ which is invariant under the map $f \mapsto \tilde{f}$ consisting of multiplying the Fourier transform by $-i\,\mathrm{sgn}\,\xi$. The space $\mathscr{H}^1$ obtained admits a much larger class of such *multipliers*. For a function $m \in L^\infty(\mathbf{R})$ and $f \in L^1(\mathbf{R})$ we shall denote by $m(D)f$ the inverse Fourier transform of $\xi \mapsto m(\xi)\hat{f}(\xi)$. It is clear that $m(D)$ is continuous from $L^1(\mathbf{R})$ to $\mathscr{S}'(\mathbf{R})$. Homogeneous functions m are linear combinations of the form $c_0 + c_1\,\mathrm{sgn}\,\xi$ with constants c_0, c_1, but we shall now discuss non-homogeneous functions with similar qualitative behavior.

Corollary 4.1.19. *If $m \in C^1(\mathbf{R}\setminus\{0\})$ and $|m(\xi)| \le 1$, $|\xi||m'(\xi)| \le 1$ when $\xi \ne 0$, then $m(D)$ is continuous from $\mathscr{H}^1(\mathbf{R})$ to $\mathscr{H}^1(\mathbf{R})$ and*

$$\|m(D)f\|_{\mathscr{H}^1} \le 3400\|f\|_{\mathscr{H}^1}, \quad f \in \mathscr{H}^1(\mathbf{R}). \tag{4.1.36}$$

Proof. If f is an atom corresponding to the interval $[-\delta,\delta]$ then (4.1.17) is valid with $M^2 \le 1/\delta$, hence

$$\frac{1}{2\pi}\int (|\hat f(\xi)|^2 + |d\hat f(\xi)/d\xi|^2/\delta^2)\,d\xi \le 1/\delta.$$

For $G(\xi) = m(\xi)\hat f(\xi)$ we have $|G(\xi)| \le |\hat f(\xi)|$ and $|G'(\xi)| \le |d\hat f(\xi)/d\xi| + |\hat f(\xi)|/|\xi|$ when $\xi \ne 0$. Since $\hat f(0) = 0$ it follows from Hardy's inequality that

$$\int_{\mathbf{R}} |\hat f(\xi)|^2/\xi^2\,d\xi \le 4\int_{\mathbf{R}} |d\hat f(\xi)/d\xi|^2\,d\xi.$$

The proof is obtained by an integration by parts,

$$\int_{-T}^{T} |\hat f(\xi)|^2/\xi^2\,d\xi \le -2\operatorname{Re}\int_{-T}^{T} \overline{\hat f(\xi)}/\xi\, d\hat f(\xi)/d\xi\,d\xi,$$

followed by an application of Cauchy-Schwarz' inequality, cancellation of one factor and letting $T \to \infty$. (One could also note that $\hat f(\xi)/\xi = \hat h(\xi)$ where $\operatorname{supp} h \subset [-\delta,\delta]$ and $-ih'(x) = f(x)$. Partial integration of $\int |h|^2\,dx$ gives the same conclusion.) By the triangle inequality $\|G'\|_2 \le 3\|\hat f'\|_2$, hence

$$\frac{1}{2\pi}\int (|G(\xi)|^2 + |G'(\xi)|^2/\delta^2)\,d\xi \le 9/\delta.$$

It follows from Lemma 4.1.10 that $G = \hat g$ where $\|g\|_{\mathscr{H}^1} \le 6\sqrt{\pi} < 12$. Thus $\|m(D)a\|_{\mathscr{H}^1} \le 12$ if a is an atom in $\mathscr{H}^1$. Since $m(D)$ is continuous from L^1 to $\mathscr{S}'$ it follows if $f = \sum \lambda_j a_j$, $\sum|\lambda_j| \le 279\|f\|_{\mathscr{H}^1}$, is an atomic decomposition of $f \in \mathscr{H}^1$ that $m(D)f = \sum \lambda_j m(D)a_j$ in $\mathscr{S}'$, and since the series also converges in $\mathscr{H}^1$ we have $m(D)f \in \mathscr{H}^1$ and

$$\|m(D)f\|_{\mathscr{H}^1} \le \sum |\lambda_j|\|m(D)a_j\|_{\mathscr{H}^1} \le 3400\|f\|_{\mathscr{H}^1}$$

by (4.1.33).

Parseval's formula proves that $m(D)$ maps $L^2(\mathbf{R})$ to $L^2(\mathbf{R})$ with norm ≤ 1 if m satisfies the hypotheses in Corollary 4.1.19. From an interpolation theorem which will be proved in Section 4.4 it follows that $m(D)$ is a continuous map from $L^p(\mathbf{R})$ to $L^p(\mathbf{R})$ for $1 < p \le 2$; by duality the same result follows for $2 \le p < \infty$ and we also get a continuous map in $\mathrm{BMO}(\mathbf{R})/\mathbf{C}$. The result, for $1 < p < \infty$, is known as *Mihlin's theorem*. It can be proved directly with arguments which we have to use anyway in the proof of the interpolation theorem just quoted. However, it is appealing to have an end point result with such an easy proof as Corollary 4.1.19 and move the rest of the argument to a general theorem with no relation to the specific situation.

4.2. Singular integrals in higher dimensions. In Section 4.1 we motivated the study of the conjugate function by questions on L^p convergence of partial sums of Fourier series and the analogue for Fourier integrals. There are many higher dimensional analogues of this, depending on how one groups terms into partial sums. However, we shall postpone discussion of such questions to Chapter V and instead study the formal analogue of the conjugate function, homogeneous multipliers on Fourier transforms of L^p functions. Such operators occur naturally in the study of elliptic differential operators $P(D)$ where P is a homogeneous polynomial of degree μ in $D = -i\partial/\partial x$ and $P(\xi) \neq 0$ for $0 \neq \xi \in \mathbf{R}^n$: If $P(D)u = f$ and $u \in \mathscr{S}$ then the Fourier transform of $D^\alpha u$ is $(\xi^\alpha/P(\xi))\hat{f}(\xi)$. If $|\alpha| = \mu$ then $\xi^\alpha/P(\xi)$ is a homogeneous function of degree 0 in $C^\infty(\mathbf{R}^n \setminus \{0\})$.

Let $m \in C^\infty(\mathbf{R}^n \setminus \{0\})$ be positively homogeneous of degree 0 in the sense that

$$m(t\xi) = m(\xi), \quad \xi \in \mathbf{R}^n \setminus \{0\},\ t > 0. \tag{4.2.1}$$

We regard m as an element in $L^\infty(\mathbf{R}^n)$. When $n = 1$ all such functions are linear combinations of the constant function and the sign function, but when $n > 1$ they form an infinite dimensional space as shown by functions such as $\xi \mapsto \xi^\alpha/|\xi|^\alpha$ with any multiindex α. We shall pay particular attention to those with $|\alpha| = 0$ or $|\alpha| = 1$.

Lemma 4.2.1. *If M is the inverse Fourier transform of a function $m \in C^\infty(\mathbf{R}^n \setminus \{0\})$ which is positively homogeneous of degree 0 then the restriction to $\mathbf{R}^n \setminus \{0\}$ is in C^∞, and it is positively homogeneous of degree $-n$,*

$$M(t\xi) = t^{-n}M(\xi), \quad \xi \in \mathbf{R}^n \setminus \{0\},\ t > 0. \tag{4.2.2}$$

For every bounded neighborhood Ω of 0 there is a constant a_Ω such that

$$M(\varphi) = a_\Omega\varphi(0) + \lim_{\varepsilon\to 0}\int_{\complement\varepsilon\Omega} M(x)\varphi(x)\,dx, \quad \varphi \in \mathscr{S}(\mathbf{R}^n). \tag{4.2.3}$$

Proof. $\xi^\beta D_\xi^\alpha m(\xi)$ is in L^1 outside a compact set if $|\alpha| > n + |\beta|$, hence $D^\beta x^\alpha M$ is then a continuous function. This proves that M is a C^∞ function outside the origin. If $\varphi \in \mathscr{S}$ then the Fourier transform of $\varphi_t(x) = t^n\varphi(tx)$ is $\xi \mapsto \hat{\varphi}(\xi/t)$ when $t > 0$, hence

$$M(\hat{\varphi}) = \int m(\xi)\varphi(\xi)\,d\xi = \int m(\xi)\varphi_t(\xi)\,d\xi = M(\hat{\varphi}(\cdot/t)), \quad \varphi \in \mathscr{S},\ t > 0, \tag{4.2.4}$$

which proves (4.2.2) when $\varphi \in C_0^\infty(\mathbf{R}^n \setminus \{0\})$. Choosing $\hat{\varphi}$ as a decreasing function of $|x|$ which is equal to 1 in the unit ball B we also conclude that

$$\int_{|x|=1} M(x)\,dS(x) = 0, \tag{4.2.5}$$

where dS is the surface measure on the unit sphere. If Ω is the unit ball, it follows that the limit in (4.2.3) exists, for the integral is independent of ε for small ε if $\varphi = 1$ in a neighborhood of the origin, and the integral over $\mathbf{R}^n$ is absolutely convergent if $\varphi(0) = 0$.

The difference between $M(\varphi)$ and this limit is thus a distribution with support at the origin satisfying (4.2.4) so it is a multiple of the Dirac measure at the origin which proves (4.2.3) when $\Omega = B$. The formula follows in general with

$$a_\Omega = a_B + \int_{\Omega \setminus rB} M(x)\,dx$$

when $r > 0$ is so small that $rB \subset \Omega$.

If M is given in $C^\infty(\mathbf{R}^n \setminus \{0\})$ and satisfies (4.2.2) and (4.2.5), then the proof of Lemma 4.2.1 shows that the limit in (4.2.3) exists and defines a distribution with $M(\hat\varphi) = M(\hat\varphi(\cdot/t))$ when $\varphi \in \mathscr{S}$. The Fourier transform m is homogeneous of degree 0 and C^∞ in $\mathbf{R}^n \setminus \{0\}$ so all such functions M can occur in Lemma 4.2.1. The constant a_Ω in (4.2.3) gives rise to a constant term in m. If Ω is the unit ball then a_Ω is the mean value of m on a sphere $\{\xi; |\xi| = r\}$.

The representation (4.2.3) of the kernel of the convolution operator $m(D)$ is the reason why $m(D)$ is called a *singular integral operator*. The homogeneity is on the borderline where the kernel just fails to be integrable both at 0 and at ∞.

The proof of Lemma 4.2.1 gives with no change that if $m \in C^{n+2}(\mathbf{R}^n \setminus \{0\})$ and

$$|\xi|^{|\alpha|}|D^\alpha m(\xi)| \le 1, \quad \xi \in \mathbf{R}^n \setminus \{0\},\ |\alpha| \le n+2, \tag{4.2.6}$$

then the inverse Fourier transform M is in $C^1(\mathbf{R}^n \setminus \{0\})$ and $|M(x)| \le C$, $|M'(x)| \le C$ when $|x| = 1$ with C independent of m. If $t > 0$ then $\xi \mapsto m(\xi/t)$ also satisfies (4.2.6), and the inverse Fourier transform is $x \mapsto t^n M(tx)$ in $\mathbf{R}^n \setminus \{0\}$, so we conclude that

$$|M(x)| \le C|x|^{-n}, \quad |M'(x)| \le C|x|^{-n-1}, \quad x \in \mathbf{R}^n \setminus \{0\}.$$

In the following extension of Theorem 4.1.1 we define $m(D)f$ when $f \in L^1(\mathbf{R}^n)$ as the inverse Fourier transform of $m\hat f$. If $f \in L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n)$ it is clear that $m(D)f \in L^2(\mathbf{R}^n)$ and that $\|m(D)f\|_2 \le \|f\|_2$ if $|m| \le 1$.

THEOREM 4.2.2. *There is a constant C depending only on n such that for every $m \in C^{n+2}(\mathbf{R}^n \setminus 0)$ satisfying* (4.2.6) *we have*

$$m_L(\{x; |m(D)f(x)| > \alpha\}) \le C\|f\|_1/\alpha, \quad \textit{for } f \in L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n), \tag{4.2.7}$$

$$\|m(D)f\|_p \le \begin{cases} Cp'^{1/p}\|f\|_p, & \textit{if } 1 < p \le 2, \\ Cp^{1/p'}\|f\|_p, & \textit{if } 2 \le p < \infty, \end{cases} \quad \textit{for } f \in L^1(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n). \tag{4.2.8}$$

Here $1/p + 1/p' = 1$, and m_L denotes the Lebesgue measure.

For the proof we need the Calderón-Zygmund decomposition lemma:

LEMMA 4.2.3. *Let $f \in L^1(I)$ where I is a cube in $\mathbf{R}^n$, and let $s > \int_I |f(x)|\,dx/m(I)$, where m is the Lebesgue measure. Then we can write*

$$f = v + \sum_1^\infty w_k, \tag{4.2.9}$$

where all terms are in $L^1(I)$, $w_k(x) = 0$ when $x \in I \setminus I_k$ for certain cubes $I_k \subset I$ with disjoint interiors, and

$$\int_I (|v(x)| + \sum_1^\infty |w_k(x)|)\,dx \le 3\int_I |f(x)|\,dx, \tag{4.2.10}$$

(4.2.11)
$$v(x) = \frac{1}{m(I_k)}\int_{I_k} f(y)\,dy \text{ and } w_k(x) = f(x) - v(x), \text{ if } x \in I_k, \text{ thus } \int_{I_k} w_k(x)\,dx = 0,$$

$$|v(x)| \le s \quad \text{almost everywhere in } I \setminus \cup I_k, \tag{4.2.12}$$

$$s \le \frac{1}{m(I_k)}\int_{I_k} |f(x)|\,dx < 2^n s, \quad \text{which implies} \tag{4.2.13}$$

$$|v(x)| < 2^n s \text{ in } \cup I_k, \quad s\sum_1^\infty m(I_k) \le \int_I |f(x)|\,dx. \tag{4.2.13$'$}$$

The lemma remains valid when $I = \mathbf{R}^n$.

Proof. We divide I into 2^n cubes J_ν, $1 \le \nu \le 2^n$, by halving each side. For each of these cubes we have

$$\frac{1}{m(J_\nu)}\int_{J_\nu} |f(x)|\,dx \le \frac{2^n}{m(I)}\int_I |f(x)|\,dx < 2^n s.$$

If the mean value on the left is $\ge s$, then J_ν is included among the cubes I_k, and we define w_k and v by (4.2.11) in I_k. It is clear that (4.2.10) is then valid for the integrals over I_k. For the other cubes J_ν, for which $\int_{J_\nu} |f(x)|\,dx/m(J_\nu) < s$, the same procedure is again applied and so on. This gives a possibly finite sequence of cubes I_k. If $x \in I \setminus \cup I_k$ then the mean value of $|f|$ is $< s$ over arbitrarily small cubes containing x, so $|f(x)| \le s$ if x is a Lebesgue point. With $v = f$ in $I \setminus \cup I_k$, the lemma is proved when I is a finite cube. If $I = \mathbf{R}^n$ we first divide $\mathbf{R}^n$ into a mesh of cubes of measure $2\|f\|_1/s$ and can then apply the result already proved to each of them.

Proof of Theorem 4.2.2. We decompose $f \in L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n)$ using Lemma 4.2.3 with $I = \mathbf{R}^n$ and $s = \alpha$. By (4.2.12), (4.2.13)$'$ and (4.2.10)

$$\|v\|_2^2 \le 2^n\alpha\|v\|_1 \le 3\cdot 2^n\alpha\|f\|_1,$$

which implies $\|m(D)v\|_2^2 \le 3\cdot 2^n\alpha\|f\|_1$, hence

$$(\tfrac12\alpha)^2 m_L(\{x; |m(D)v(x)| > \tfrac12\alpha\}) \le 3\cdot 2^n\alpha\|f\|_1.$$

Since the terms w_k have supports in the cubes I_k with disjoint interiors, it follows that $\sum_1^\infty w_k$ converges in L^2, so $m(D)\sum_1^\infty w_k = \sum_1^\infty m(D)w_k$ with convergence in L^2, hence in L^1_{loc}. Let y_k and s_k be the center and the side of I_k, and let $2I_k$ be the cube with center y_k and side $2s_k$. If $x \notin 2I_k$ then

$$(m(D)w_k)(x) = \int_{I_k} M(x-y)w_k(y)\,dy = \int_{I_k} (M(x-y) - M(x-y_k))w_k(y)\,dy.$$

Since $|M'(x)| \le C|x|^{-n-1}$ by the remarks before the statement of Theorem 4.2.2 we have for $y \in I_k$

$$\int_{\complement 2I_k} |M(x-y) - M(x-y_k)|\,dx \le C' s_k \int_{\complement 2I_k} |x-y_k|^{-n-1}\,dx = C'', \tag{4.2.14}$$

where C'' is independent of the cube I_k by homogeneity and translation invariance. Hence it follows that

$$\int_{\complement 2I_k} |m(D)w_k(x)|\,dx \le C'' \int_{I_k} |w_k(y)|\,dy. \tag{4.2.14$'$}$$

The measure of $E = \cup(2I_k)$ is at most $2^n\|f\|_1/\alpha$, by (4.2.13)$'$, and

$$\int_{\complement E} \sum |m(D)w_k(x)|\,dx \le 3C''\|f\|_1$$

by (4.2.14)$'$ and (4.2.10), so it follows that

$$m_L(\{x; \sum |m(D)w_k(x)| > \tfrac{1}{2}\alpha\}) \le m(E) + 6C''\|f\|_1/\alpha \le 2^n\|f\|_1/\alpha + 6C''\|f\|_1/\alpha.$$

Summing up, we have proved that

$$m_L(\{x; |m(D)f(x)| > \alpha\}) \le (3\cdot 2^{n+2} + 2^n + 6C'')\|f\|_1/\alpha,$$

which proves (4.2.7).

The estimate (4.2.8) for $2 < p < \infty$ follows by duality from the estimate for $1 < p < 2$. In that case it is a consequence of (4.2.7) and the estimate $\|m(D)f\|_2 \le \|f\|_2$, which implies

$$\alpha^2 m_L(\{x; |m(D)f(x)| > \alpha\}) \le \|f\|_2^2,$$

for we can apply the *Marcinkiewicz interpolation theorem*:

Theorem 4.2.4. *Let X and Y be two locally compact spaces with positive Radon measures $d\mu$ and $d\nu$, and let T be a sublinear map from $L^1(X,d\mu)\cap L^\infty(X,d\mu)$ to $L^1_{\text{loc}}(Y,d\nu)$, that is,*

$$|T(g+h)| \le |T(g)| + |T(h)|, \quad g,h \in L^1(X,d\mu)\cap L^\infty(X,d\mu).$$

If $p_1, p_2 \in [1,\infty)$ and T is of weak type (p_j,p_j) for $j = 1,2$, that is,

$$s^{p_j}\nu(\{y\in Y; |Tf(y)| > s\}) \le C_j \int_X |f(x)|^{p_j}\,d\mu(x), \quad f \in L^1(X,d\mu)\cap L^\infty(X,d\mu), \tag{4.2.15}$$

for $j = 1,2$, then it follows that for $p_2 < p < p_1$

$$\Big(\int_Y |Tf(y)|^p\,d\nu(y)\Big)^{\frac{1}{p}} \le C_p\Big(\int_X |f(x)|^p\,d\mu(x)\Big)^{\frac{1}{p}}, \quad f \in L^1(X,d\mu)\cap L^\infty(X,d\mu).$$

If $p_1 = \infty$ and the hypothesis (4.2.15) is replaced by $\|Tf\|_\infty \le C_\infty \|f\|_\infty$ when $j = 1$, then the same conclusion holds for $p_2 < p < \infty$.

PROOF. Let $f = g_s + h_s$ where

$$g_s(x) = \begin{cases} f(x), & \text{if } |f(x)| < s \\ 0, & \text{if } |f(x)| \ge s \end{cases}, \quad h_s(x) = \begin{cases} 0, & \text{if } |f(x)| < s \\ f(x), & \text{if } |f(x)| \ge s \end{cases}.$$

Since $|Tf| \le |Tg_s| + |Th_s|$ we have

$$\begin{aligned} N(s) = \nu(\{y; |Tf(y)| > s\}) &\le \nu(\{y; |Tg_s(y)| > \tfrac12 s\}) + \nu(\{y; |Th_s(y)| > \tfrac12 s\}) \\ &\le (2/s)^{p_1} C_1 \int_{|f(x)|<s} |f(x)|^{p_1}\, d\mu(x) + (2/s)^{p_2} C_2 \int_{|f(x)|\ge s} |f(x)|^{p_2}\, d\mu(x), \end{aligned}$$

by (4.2.15) if $p_1 < \infty$. Hence

$$\begin{aligned} \|Tf\|_p^p = p\int_0^\infty s^{p-1} N(s)\, ds &\le p\Big(2^{p_1} C_1 \iint_{|f(x)|<s} s^{p-1-p_1} |f(x)|^{p_1}\, d\mu(x)\, ds \\ &\quad + 2^{p_2} C_2 \iint_{|f(x)|\ge s} s^{p-1-p_2} |f(x)|^{p_2}\, d\mu(x)\, ds\Big) \\ &= p(2^{p_1}(p_1-p)^{-1} C_1 + 2^{p_2}(p-p_2)^{-1} C_2) \int_X |f(x)|^p\, d\mu(x). \end{aligned}$$

This proves the theorem for $p_1 < \infty$. When $p_1 = \infty$ we can reduce the proof to the case where $\|Tf\|_\infty < \frac12 \|f\|_\infty$. Then $|Tg_s| < \frac12 s$ above so $N(s) \le \nu\{y; |Th_s(y)| > \frac12 s\}$, and the proof proceeds as before with one term less.

REMARK. There is a more general version of Marcinkiewicz' interpolation theorem which deals with maps from L^p to L^q when $p \ne q$. The proof is similar but somewhat more complicated and can be found in Zygmund [1].

Next we extend the Hardy-Littlewood maximal theorem to n dimensions. If $f \in L^1_{\text{loc}}(\mathbf{R}^n)$ we define the maximal function by

$$f^*_{\text{HL}}(x) = \sup_{x\in B} \frac{1}{m(B)} \int_B |f(y)|\, dy, \tag{4.2.16}$$

where B is a ball with respect to a fixed norm in $\mathbf{R}^n$. It does not matter which norm is chosen, so we can use cubes as well as Euclidean balls, but it is important that the shape is fixed. If $\varrho(x)$ is a positive decreasing function of $|x|$, then the proof of (4.1.6)′ gives

$$\int_{\mathbf{R}^n} |f(x+t)| \varrho(t)\, dt \le f^*_{\text{HL}}(x) \int_{\mathbf{R}^n} \varrho(t)\, dt. \tag{4.2.16$'$}$$

The maximal theorem takes the form:

THEOREM 4.2.5. *If $f \in L^1(\mathbf{R}^n)$ then*

$$m(\{x; f^*_{\mathrm{HL}}(x) > s\}) \le \frac{3^n}{s}\int_{\mathbf{R}^n} |f(x)|\,dx, \quad s > 0, \tag{4.2.17}$$

where m is the Lebesgue measure. If $1 < p < \infty$ then

$$\|f^*_{\mathrm{HL}}\|_p \le (3^n 2^p p')^{\frac{1}{p}}\|f\|_p, \quad f \in L^p(\mathbf{R}^n). \tag{4.2.18}$$

PROOF. The set $E_s = \{x; f^*_{\mathrm{HL}}(x) > s\}$ is the union of all open balls B such that the mean value of $|f|$ over B exceeds s. For every compact set $K \subset E_s$ there is a finite family $\mathscr{F}_K$ of such balls which contains K. We shall prove that there is a subset $\mathscr{F}'_K \subset \mathscr{F}_K$ consisting of disjoint balls such that

$$\bigcup_{B\in\mathscr{F}_K} B \subset \bigcup_{B\in\mathscr{F}'_K} 3B \tag{4.2.19}$$

where $3B$ is the ball with the same center as B but three times larger radius. This implies that

$$m(K) \le \sum_{B\in\mathscr{F}'_K} m(3B) = 3^n \sum_{B\in\mathscr{F}'_K} m(B) \le 3^n s^{-1} \sum_{B\in\mathscr{F}'_K}\int_B |f(x)|\,dx,$$

which implies (4.2.17) since the balls in $\mathscr{F}'_K$ are disjoint and K is an arbitrary compact subset of E_s.

To select the balls $\mathscr{F}'_K$ we first choose a ball $B_1 \in \mathscr{F}_K$ with maximal radius. Among the balls in $\mathscr{F}_K$ which do not intersect B_1 we then choose a ball B_2 with maximal radius and continue so that B_j is always a ball in $\mathscr{F}_K$ with maximal radius not intersecting $B_1, \dots, B_{j-1}$. The selection breaks off since $\mathscr{F}_K$ is finite. If a ball $B \in \mathscr{F}_K$ has not been chosen it must intersect one of the chosen ones. If $B_j \cap B \ne \emptyset$ and j is minimal, then the radius of B_j is at least as large as that of B since B should otherwise have been chosen instead of B_j. By the triangle inequality it follows that $B \subset 3B_j$, which proves (4.2.19) and (4.2.17). The estimate (4.2.18) is now a consequence of the Marcinkiewicz interpolation theorem, with some constant. The constant in (4.2.18) is obtained if one repeats the proof of (4.1.8) which is left for the reader to do.

EXERCISE 4.2.1. a) Prove for the Hardy-Littlewood maximal function (4.2.16) that

$$m(\{x; f^*_{\mathrm{HL}}(x) > s\}) \le 2\cdot 3^n s^{-1}\int_{|f(x)|>s/2} |f(x)|\,dx, \quad f \in L^1_{\mathrm{loc}}(\mathbf{R}^n), \quad s > 0.$$

b) Prove that if $\varphi \in C^1(\mathbf{R})$, $\varphi(0) = 0$ and $\varphi' \ge 0$ then

$$\int \varphi(f^*_{\mathrm{HL}}(x))\,dx \le 2\cdot 3^n \int |f(x)| \int_{0<t<2|f(x)|} \frac{\varphi'(t)}{t}\,dt.$$

c) Prove that when $\varepsilon > 0$ then

$$\int f^*_{\mathrm{HL}}(x)^{\varepsilon+1}(1 + f^*_{\mathrm{HL}}(x))^{-\varepsilon}\,dx$$

$$\le 2\cdot 3^n\Big((1+\varepsilon^{-1})\int_{2|f(x)|<1} |f(x)|\,dx + \int_{2|f(x)|\ge 1} |f(x)|(1+\varepsilon+\varepsilon^{-1}+\log|2f(x)|)\,dx\Big).$$

EXERCISE 4.2.2. Prove with the notation in Lemma 4.2.3, $I = \mathbf{R}^n$, that if J is an axis parallel cube with $J \cap I_j \neq \emptyset$ but $J \not\subset 2I_j$ then $I_j \subset 5J$, and deduce that

$$\frac{1}{m(J)} \int_J |f(y)|\, dy \le s(1 + 10^n) \quad \text{if } J \not\subset \cup(2I_j),\ s > 0.$$

Use this to give another proof of (4.2.17) (with 3^n replaced by a larger constant).

EXERCISE 4.2.3. With f^*_{HL} defined using the norm $|x| = \max_{1\le j\le n} |x_j|$, $x \in \mathbf{R}^n$,
a) prove that

$$m(\{x; f^*_{\mathrm{HL}}(x) \ge s\}) \ge 2^{-n} s^{-1} \int_{|f(x)|>s} |f(x)|\, dx;$$

b) prove that if $\varphi \in C^1(\mathbf{R})$, $\varphi(0) = 0$ and $\varphi' \ge 0$ then

$$\varphi(f^*_{\mathrm{HL}}(x))\, dx \ge 2^{-n} \int |f(x)| \int_{t<|f(x)|} \frac{\varphi'(t)}{t}\, dt;$$

c) prove that when $0 < \varepsilon < 1$ then

$$\int f^*_{\mathrm{HL}}(x)^{\varepsilon+1}(1 + f^*_{\mathrm{HL}}(x))^{-\varepsilon}\, dx$$
$$\ge 2^{-n-1}\Big((\tfrac12 + \varepsilon^{-1}) \int_{|f(x)<1} |f(x)|^{1+\varepsilon}\, dx + \int_{|f(x)|>1} |f(x)|(\tfrac12 + \varepsilon^{-1} + \log|f(x)|)\, dx\Big).$$

Later on we shall also need an analogue of Theorem 4.1.2′ for the maximal function

$$f^{**}_{HL}(x,t) = \sup_{x\in B, r(B)>t} \frac{1}{m(B)} \int_B |f(y)|\, dy, \quad x \in \mathbf{R}^n,\ t > 0, \tag{4.2.16$''$}$$

which also takes the radius $r(B)$ of the ball B into account. We can replace $f^*_{\mathrm{HL}}(x)$ by $f^{**}_{\mathrm{HL}}(x,t)$ in (4.2.16)′ if $\varrho(y)$ is constant when $|y| < t$.

THEOREM 4.2.5′. *Let ν be a positive measure in $\mathbf{R}^n \times \mathbf{R}_+$ such that for every ball $B \subset \mathbf{R}^n$*

$$\nu(B \times (0, r(B)) \le m(B), \tag{4.2.20}$$

where m is the Lebesgue measure in $\mathbf{R}^n$. Then it follows that

$$\nu(\{(x,t); f^{**}_{\mathrm{HL}}(x,t) > s\}) \le \frac{3^n}{s} \int |f(x)|\, dx, \quad f \in L^1(\mathbf{R}^n), \quad s > 0. \tag{4.2.17$'$}$$

If $1 < p < \infty$ then

$$\Big(\int |f^{**}_{\mathrm{HL}}(x,t)|^p d\nu(x,t)\Big)^{1/p} \le (3^n 2^p p')^{\frac1p} \|f\|_p, \quad f \in L^p(\mathbf{R}^n). \tag{4.2.18$'$}$$

Proof. For any compact subset K of the open set where $f^{**}_{\mathrm{HL}}(x,t) > s$ we can find a finite family $\mathscr{F}_K$ of balls B such that $\int_B |f(y)|\,dy > sm(B)$ and

$$K \subset \bigcup_{B\subset\mathscr{F}_K} B\times(0,r(B)).$$

As in the proof of Theorem 4.2.5 we choose a finite disjoint sequence $B_1, B_2, \dots \in \mathscr{F}_K$ with $r(B_j)$ decreasing such that for every $B\in\mathscr{F}_K$ we have $r(B)\le r(B_j)$ if $B\cap B_k=\emptyset$ for $k<j$ and $B\cap B_j\ne\emptyset$ for some j. When $B\cap B_j\ne\emptyset$ and j is minimal it follows that

$$B\times(0,r(B))\subset(3B_j)\times(0,r(B_j))\subset(3B_j)\times(0,r(3B_j))$$

which gives $K\subset\bigcup(3B_j)\times(0,r(3B_j))$, hence by (4.2.20)

$$\nu(K)\le\sum m(3B_j)=3^n\sum m(B_j)\le 3^n\sum\int_{B_j}|f(y)|\,dy/s\le 3^n\|f\|_1/s.$$

This proves the weak type estimate (4.2.17)$'$, and as before it implies (4.2.18)$'$ by the Marcinkiewicz interpolation theorem.

With m and M as in Lemma 4.2.1 we can form a maximal function

$$f^*_M(x)=\sup_{0<\varepsilon<\delta}\Big|\int_{\varepsilon<|x|<\delta} f(x-y)M(y)\,dy\Big|.$$

The proof of (4.1.12) gives with no essential change apart from substitution of Theorems 4.2.2 and 4.2.5 for Theorems 4.1.1 and 4.1.2

$$\|f^*_M\|_p\le C_p\|f\|_p,\quad f\in L^p(\mathbf{R}),\ 1<p<\infty.$$

This implies that

$$\lim_{\varepsilon\to0,\delta\to\infty}\int_{\varepsilon<|x|<\delta} f(x-y)M(y)\,dy=(m(D)f)(x)-a_Bf(x)\quad\text{for almost all } x\in\mathbf{R}^n$$

where $m(D)f$ is defined by continuous extension of $m(D)$ from $\mathscr{S}$ to $L^p(\mathbf{R}^n)$ and B is the unit ball. We leave the repetition of the details for the reader. The analogue of Proposition 4.1.6 for functions in $\mathbf{R}^n$ is obvious, for it suffices to consider products of functions of one variable, and Proposition 4.1.7 also carries over to the n-dimensional case with the same proof.

Thus we arrive at the discussion of Hardy spaces, which will elucidate the failure of Theorem 4.2.2 for $p=1$. When $n>1$ we have an infinite dimensional supply of operators $m(D)$ to choose from, but at first we shall only consider n of them, chosen so that the proof of Theorem 4.1.15 can be extended. The others will be controlled afterwards.

Let $|\xi|=(\xi_1^2+\dots+\xi_n^2)^{\frac12}$ be the Euclidean norm now, and set

$$R_j(\xi)=-i\xi_j/|\xi|,\quad j=1,\dots,n. \tag{4.2.21}$$

The corresponding convolution operators $R_j(D)$ are called the Riesz operators. Since the inverse Fourier transform of $\xi\mapsto1/|\xi|$ is a constant times the function $x\mapsto|x|^{1-n}$, for reasons of homogeneity and orthogonal invariance, it follows by differentiation that the kernel of the convolution operator $R_j(D)$ is $x\mapsto cx_j|x|^{-n-1}$ for some constant c which will be determined in a moment. Since R_j is odd the constant a_B in the representation (4.2.3) of the inverse Fourier transform of R_j must vanish. Sometimes we shall use the notation $R_0=1$, so that $R_0(D)=\mathrm{Id}$, the identity operator.

Definition 4.2.6. The *Hardy space* $\mathscr{H}^1(\mathbf{R}^n)$ is the space of all $f \in L^1(\mathbf{R}^n)$ such that $R_j(D)f \in L^1(\mathbf{R}^n)$ for $j = 1, \dots, n$.

Before stating an analogue of Proposition 4.1.9 we must make some preliminary remarks on the Poisson kernel P_0 in $\mathbf{R}^{1+n}$. It is the kernel giving the unique bounded solution of the Laplace equation in $\mathbf{R}^{1+n}_+ = \{(t,x); t > 0, x \in \mathbf{R}^n\}$ with given boundary values $f \in L^\infty(\mathbf{R}^n) \cap C^0(\mathbf{R}^n)$, say, when $t = 0$. If E is the fundamental solution of the Laplacian in $\mathbf{R}^{1+n}$, then

$$P_0(t,x) = 2\partial E(t,x)/\partial t = 2t(|x|^2 + t^2)^{-\frac{1}{2}(n+1)}/c_{n+1}, \quad (t,x) \in \mathbf{R}^{1+n}_+, \tag{4.2.22}$$

where c_{n+1} is the area of the unit sphere $S^n \subset \mathbf{R}^{n+1}$. The Fourier transform of $P_0(t,x)$ with respect to x is $\xi \mapsto e^{-t|\xi|}$, for it is continuous and uniformly bounded, $\to 1$ as $t \to 0$ and is annihilated by $\partial^2/\partial t^2 - |\xi|^2$ since P_0 is harmonic. The kernel

$$P_j(t,x) = 2\partial E(t,x)/\partial x_j = 2x_j(|x|^2 + t^2)^{-\frac{1}{2}(n+1)}/c_{n+1}, \quad (t,x) \in \mathbf{R}^{1+n}_+, \tag{4.2.23}$$

is also harmonic, and the distribution limit when $t \to 0$ is the inverse Fourier transform of R_j, which is thus equal to $\operatorname{vp} 2x_j|x|^{-n-1}/c_{n+1}$. In fact, the Fourier transform of $\partial P_j(t,x)/\partial t = \partial P_0(t,x)/\partial x_j$ with respect to x is $\xi \mapsto i\xi_j e^{-t|\xi|}$, and the Fourier transform of P_j tends to 0 in $\mathscr{S}'$ as $t \to +\infty$, so it is the integral $\xi \mapsto -i\xi_j|\xi|^{-1}e^{-t|\xi|} = R_j(\xi)e^{-t|\xi|}$ vanishing when $t \to +\infty$. Thus the Fourier transform of $P_j(t,x)$ with respect to x is $R_j(\xi)e^{-t|\xi|}$ for $j = 0, \dots, n$.

We can now state and prove an analogue of Proposition 4.1.9. To simplify notation we shall sometimes use the notation $x_0 = t$, $\partial_0 = \partial/\partial t$, $\partial_j = \partial/\partial x_j$ when $j = 1, \dots, n$.

Proposition 4.2.7. *$\mathscr{H}^1(\mathbf{R}^n)$ is a Banach space with the norm $\|f\|_{\mathscr{H}^1} = \sum_0^n \|R_j f\|_{L^1}$, and it is invariant under complex conjugation which even preserves the norm. When $f \in \mathscr{H}^1(\mathbf{R}^n)$ then the functions*

$$(P_j f)(t,x) = \int P_j(t, x-y) f(y)\, dy, \quad j = 0, \dots, n, \tag{4.2.24}$$

are conjugate harmonic in $\mathbf{R}^{1+n}_+$, in the sense that $\partial_k P_j f = \partial_j P_k f$, $j,k = 0, \dots, n$, and $\sum_0^n \partial_j P_j f = 0$. They have boundary values $R_j f$ in the L^1 sense,

$$\begin{aligned} &\int_{\mathbf{R}^n} |(P_j f)(t,x)|\, dx \le \|R_j f\|_{L^1}, \ t > 0, \\ &\lim_{t \to +0} \int_{\mathbf{R}^n} |(P_j f)(t,x) - R_j(D)f(x)|\, dx = 0, \quad f \in \mathscr{H}^1(\mathbf{R}^n). \end{aligned} \tag{4.2.25}$$

If $f \in L^1(\mathbf{R}^n)$ and $\hat f$ has compact support not containing the origin, then $f \in \mathscr{H}^1(\mathbf{R}^n)$. Such functions in $\mathscr{S}(\mathbf{R}^n)$ are dense in $\mathscr{H}^1(\mathbf{R}^n)$, and the closure of $\mathscr{H}^1(\mathbf{R}^n)$ in $L^1(\mathbf{R}^n)$ is $\{f \in L^1(\mathbf{R}^n); \hat f(0) = 0\}$.

Proof. Since $R_j(D)$ is continuous from $L^1(\mathbf{R}^n)$ to $\mathscr{S}'(\mathbf{R}^n)$ it follows that $\mathscr{H}^1(\mathbf{R}^n)$ is complete, hence a Banach space, and since the kernel of $R_j(D)$ is real valued it is clear that $\mathscr{H}^1(\mathbf{R}^n)$ is invariant under complex conjugation. If $f \in \mathscr{H}^1(\mathbf{R}^n)$ then $R_j(\xi)\hat{f}(\xi)$ is continuous for $j = 0, \dots, n$, which implies $\hat{f}(0) = 0$, that is, $\int_{\mathbf{R}^n} f(x)\,dx = 0$. If $f \in L^1$ and $\operatorname{supp}\hat{f}$ is a compact subset of $\mathbf{R}^n \setminus 0$ it follows as in the one dimensional case that $f \in \mathscr{H}^1(\mathbf{R}^n)$.

Choose $\chi \in \mathscr{S}(\mathbf{R}^n)$ so that $\hat{\chi} \in C_0^\infty(\mathbf{R}^n)$ and $\hat{\chi} = 1$ in a neighborhood of the origin. With $\chi_\varepsilon(x) = \varepsilon^n \chi(\varepsilon x)$ we claim that

$$\lim_{\varepsilon\to 0} \|\chi_\varepsilon * f\|_{L^1} = |\hat{f}(0)|\|\chi\|_{L^1}, \quad \lim_{t\to\infty} \|\chi_t * f - f\|_{L^1} = 0, \quad f \in L^1(\mathbf{R}^n). \tag{4.2.26}$$

As for (4.1.16) it suffices to prove this when $\hat{f} \in C_0^\infty(\mathbf{R}^n)$, and the second part is trivial then. To prove the first part we write the Fourier transform of $g_\varepsilon = \chi_\varepsilon * f - \chi_\varepsilon \hat{f}(0)$ as

$$\hat{g}_\varepsilon(\xi) = \hat{\chi}(\xi/\varepsilon)(\hat{f}(\xi) - \hat{f}(0)),$$

and conclude that

$$\varepsilon^{2|\alpha|} \int_{\mathbf{R}^n} |D^\alpha \hat{g}_\varepsilon(\xi)|^2\,d\xi \le C_\alpha \varepsilon^{n+2}$$

for every α. By Parseval's formula it follows for every positive integer N that

$$\int_{\mathbf{R}^n} (1 + \varepsilon^2|x|^2)^N |g_\varepsilon(x)|^2\,dx \le C_N \varepsilon^{n+2},$$

so Cauchy-Schwarz' inequality gives $\|g_\varepsilon\|_{L^1} \le C'\varepsilon$, which proves (4.2.26).

If $f \in L^1(\mathbf{R}^n)$ and $\hat{f}(0) = 0$ it follows that $f_{t,\varepsilon} = \chi_t * (f - \chi_\varepsilon * f) \to f$ in L^1 as $t \to \infty$ and $\varepsilon \to 0$, and the Fourier transform of $f_{t,\varepsilon}$ has compact support in $\mathbf{R}^n \setminus \{0\}$ so $f_{t,\varepsilon} \in \mathscr{H}^1(\mathbf{R}^n)$. This proves that the closure of $\mathscr{H}^1(\mathbf{R}^n)$ in $L^1(\mathbf{R}^n)$ consists of all $f \in L^1(\mathbf{R}^n)$ with $\hat{f}(0) = 0$. Since $R_j(D)f_{t,\varepsilon} = (R_j(D)f)_{t,\varepsilon}$ we have $f_{t,\varepsilon} \to f$ in $\mathscr{H}^1(\mathbf{R}^n)$ if $f \in \mathscr{H}^1(\mathbf{R}^n)$. If we regularize $f_{t,\varepsilon}$ to $\varphi^\delta f_{t,\varepsilon} \in \mathscr{S}(\mathbf{R}^n)$ as in the proof of Proposition 4.1.9, then $\varphi^\delta f_{t,\varepsilon} \to f_{t,\varepsilon}$ in L^1 as $\delta \to 0$, and the support of the Fourier transform is contained in a fixed compact set $K \subset \mathbf{R}^n \setminus \{0\}$ for small δ. We can choose $r_j \in \mathscr{S}$ so that $\hat{r}_j = R_j$ in a neighborhood of K, and then $R_j(D)(\varphi^\delta f_{t,\varepsilon}) = r_j * (\varphi^\delta f_{t,\varepsilon})$ for small δ. This converges in L^1 to $r_j * f_{t,\varepsilon} = R_j(D)f_{t,\varepsilon}$ as $\delta \to 0$, which completes the proof of the density statement in the proposition.

When $\hat{f} \in C_0^\infty(\mathbf{R}^n \setminus \{0\})$ then the Fourier transform of $(P_j f)(t,x)$ with respect to x is $\xi \mapsto e^{-t|\xi|} R_j(\xi)\hat{f}(\xi)$, so $(P_j f)(t,x)$ is the Poisson integral of $R_j f$. This proves (4.2.25) for f in a dense subset of $\mathscr{H}^1(\mathbf{R}^n)$, and by continuity it follows for all $f \in \mathscr{H}^1(\mathbf{R}^n)$.

The following lemma is analogous to Lemma 4.1.10 but it is slightly harder to prove when n is even. We state it in somewhat greater generality than needed right now to prepare for an analogue of Corollary 4.1.19, but to simplify the proof we restrict ourselves to functions which will become atoms in $\mathscr{H}^1(\mathbf{R}^n)$. (See also Lemma 4.4.3.)

LEMMA 4.2.8. *Let $f \in L^2(\mathbf{R}^n)$ have support in a ball B and assume that $\int f(x)\,dx = 0$. Let $m \in C^\nu(\mathbf{R}^n \setminus \{0\})$ for some $\nu > n/2$, and assume that*

$$|\xi|^{|\alpha|}|D^\alpha m(\xi)| \le 1, \quad \text{if } 0 \ne \xi \in \mathbf{R}^n,\ |\alpha| \le \nu. \tag{4.2.27}$$

Then it follows that $\|m(D)f\|_1 \le C_n\sqrt{|B|}\|f\|_2$. *In particular,* $f \in \mathscr{H}^1(\mathbf{R}^n)$ *and* $\|f\|_{\mathscr{H}^1} \le (n+1)C_n\sqrt{|B|}\|f\|_2$. *Here* $|B|$ *is the Lebesgue measure of* B.

Proof. Without restriction we may assume that the center of B is at the origin. If the radius is δ then $g(x) = \delta^n f(\delta x)$ has support in the unit ball, $\hat{g}(\xi) = \hat{f}(\xi/\delta)$ and $m(\xi)\hat{f}(\xi) = m(\xi)\hat{g}(\delta\xi) = m_\delta(\delta\xi)\hat{g}(\delta\xi)$ where $m_\delta(\xi) = m(\xi/\delta)$ also satisfies (4.2.27). Thus $m(D)f = \delta^{-n}(m_\delta(D)g)(\cdot/\delta)$, so $\|m(D)f\|_1 = \|m_\delta(D)g\|_1$. Since $\delta^{n/2}\|f\|_2 = \|g\|_2$ the proof has now been reduced to the case where B is the unit ball, which we assume from now on.

Cauchy-Schwarz' inequality gives at once that $\|f\|_1 \le \sqrt{|B|}\|f\|_2$, and by Parseval's formula

$$(2\pi)^{-n}\int_{\mathbf{R}^n} |D^\alpha \hat{f}(\xi)|^2\,d\xi = \int_{\mathbf{R}^n} |x^\alpha f(x)|^2\,dx \le \|f\|_2^2,$$

for arbitrary α. Choose $\psi \in C_0^\infty(\mathbf{R}^n)$ so that $\psi(\xi) = 1$ when $|\xi| \le 1$ and $\psi(\xi) = 0$ when $|\xi| \ge 2$, and set $m = m_1 + m_2$ where $m_1 = \psi m$ and $m_2 = (1-\psi)m$. Since $|\xi| \ge 1$ when $\xi \in \operatorname{supp} m_2$, the derivatives of m_2 of order $\le \nu$ are bounded, and we obtain

$$\int_{\mathbf{R}^n} |x^\alpha m_2(D)f(x)|^2\,dx = (2\pi)^{-n}\int_{\mathbf{R}^n} |D^\alpha(m_2(\xi)\hat{f}(\xi))|^2\,d\xi \le C_\alpha\|f\|_2^2,$$

when $|\alpha| \le \nu$. Hence

$$\int (1+|x|^2)^\nu |m_2(D)f(x)|^2\,dx \le C\|f\|^2,$$

and by Cauchy-Schwarz' inequality this implies $\|m_2(D)f\|_1 \le C\|f\|_2$, with another constant C, because $2\nu > n$.

When $\xi \in \operatorname{supp} m_1$ we have $|\xi| \le 2$, and $|D^\alpha \hat{f}(\xi)| \le C_\alpha\|f\|_2$ for every α when $|\xi| \le 2$, but we only have bounds for $|\xi|^\alpha D^\alpha m_1(\xi)$. Let ν be the smallest integer $> n/2$, thus $(n+1)/2 \le \nu \le (n+2)/2$. Since

$$D^\alpha(m_1(\xi)\hat{f}(\xi)) - (D^\alpha m_1(\xi))\hat{f}(\xi)$$

only contains derivatives of m_1 of order $|\alpha|-1$ and $|\hat{f}(\xi)| = |\hat{f}(\xi) - \hat{f}(0)| \le C\|f\|_2|\xi|$, we have

$$|D^\alpha(m_1(\xi)\hat{f}(\xi))| \le C_\alpha|\xi|^{1-\nu}\|f\|_2, \quad |\alpha| \le \nu.$$

If n is even then $\nu - 1 = n/2$ so this bound is not in L^2 when $|\xi| < 2$. However, if $1 < p < 2$ we have $\|D^\alpha(m_1\hat{f})\|_p \le C_p\|f\|_2$ when $|\alpha| \le \nu$, and it follows from the Hausdorff-Young inequality that with $1/p + 1/p' = 1$

$$\Big(\int_{\mathbf{R}^n} |x^\alpha m_1(D)f(x)|^{p'}\,dx\Big)^{\frac{1}{p'}} \le C_p\|f\|_2.$$

Thus we have a bound for the norm of $(1+|x|)^\nu|m_1(D)f(x)|$ in $L^{p'}(\mathbf{R}^n)$. When $(1+|x|)^{-\nu}$ is in L^p, that is, $n/\nu < p < 2$, it follows from Hölder's inequality that $\|m_1(D)f\|_1 \le C\|f\|_2$; we can for example take $p = (2n+1)/(n+1)$. This completes the proof of the lemma.

If L is a continuous linear form on $\mathscr{H}^1(\mathbf{R}^n)$ it follows from Lemma 4.2.8 that L for every ball B restricts to a continuous linear form on $\{f \in L^2(B); \int_B f\,dx = 0\}$. Hence there is a unique function $\Phi_B \in L^2(B)$ with $\int_B \Phi_B\,dx = 0$ such that

$$L(f) = \int f(x)\Phi_B(x)\,dx, \quad \text{if } f \in L^2,\ \operatorname{supp} f \subset B,\ \int f(x)\,dx = 0,$$

and $\int_B |\Phi_B(x)|^2\,dx \le C'_n m(B)\|L\|^2_{(\mathscr{H}^1)'}$. If $B_1 \subset B_2$ then $\Phi_{B_2} - \Phi_{B_1}$ is a constant $c_{B_2B_1}$ in B_1, and $\Phi_{B_2} - c_{B_2B_1}$ extends Φ_{B_1} to B_2. From the sequence Φ_{B_j} where $B_j = \{x \in \mathbf{R}^n; |x| < j\}$ we obtain a function $\varphi \in L^2_{\text{loc}}(\mathbf{R}^n)$ equal to $\Phi_{B_j} - c_{B_jB_1}$ in B_j. For every ball B we have $\Phi_B = \varphi - \varphi_B$ where $\varphi_B = \int_B \varphi\,dx/m(B)$, and

$$L(f) = \int f(x)\varphi(x)\,dx \tag{4.2.28}$$

for every $f \in L^2(\mathbf{R}^n)$ with compact support and $\int_{\mathbf{R}^n} f(x)\,dx = 0$. This is a dense subset of $\mathscr{H}^1(\mathbf{R}^n)$. In fact, by Proposition 4.2.7 functions $f \in \mathscr{S}$ with $\hat{f} \in C_0^\infty(\mathbf{R}^n \setminus \{0\})$ are dense. If f is such a function and $0 \le \psi \in C_0^\infty$, $\psi = 1$ in a neighborhood of 0 and $\int \psi\,dx = 1$, it follows from Lemma 4.2.8 that

$$f_j(x) = \psi(2^{-j}x)f(x) - c_j\psi(2^{-j}x)$$

is in $\mathscr{H}^1(\mathbf{R}^n)$ if the integral is 0, that is,

$$c_j = 2^{-nj}\int \psi(2^{-j}x)f(x)\,dx = 2^{-nj}\int (\psi(2^{-j}x) - 1)f(x)\,dx.$$

Thus $c_j = O(2^{-\nu j})$ for every ν, and it follows from Lemma 4.2.8 that $\|f_j - f_{j+1}\|_{\mathscr{H}^1} = O(2^{-\nu j})$ for every ν, since this is true for the L^2 norm. Since $f_j \to f$ it follows that $\|f - f_j\|_{\mathscr{H}^1} = O(2^{-\nu j})$ for every ν, so we have proved (4.2.28) for a dense subset of $\mathscr{H}^1$. In a moment we shall see that $\varphi \in \mathscr{S}'$, and since $f_j \to f$ in $\mathscr{S}$ this will prove that (4.2.28) is valid when $\hat{f} \in C_0^\infty(\mathbf{R}^n \setminus \{0\})$.

We have now been led to introduce an analogue of Definition 4.1.11:

Definition 4.2.9. A function $f \in L^2_{\text{loc}}(\mathbf{R}^n)$ is said to be in $\text{BMO}(\mathbf{R}^n)$ if there is a constant K such that for every ball $B \subset \mathbf{R}^n$

$$\frac{1}{m(B)}\int_B |f(x) - f_B|^2\,dx \le K^2, \quad \text{if } f_B = \frac{1}{m(B)}\int_B f(y)\,dy. \tag{4.2.29}$$

Example. $f(x) = \log|x|$ is in $\text{BMO}(\mathbf{R}^n)$ but not even locally bounded. Since $f(tx) = \log t + \log|x|$ when $t > 0$ it suffices to verify (4.2.29) for balls B of radius 1. If $|x| \le 3$ when $x \in B$ then

$$\int_B |f(x) - f_B|^2\,dx \le \int_B |f(x)|^2\,dx \le \int_{|x|\le 3} |f(x)|^2 dx < \infty.$$

If $|x| \ge 1$ when $x \in B$ then $|f'(x)| \le 1$ when $x \in B$, hence $|f(x) - f_B| \le 2$ if $x \in B$, and $\int_B |f(x) - f_B|^2\,dx \le 4m(B)$ which proves that $f \in \text{BMO}(\mathbf{R}^n)$. More generally it follows that $f(x) = \int \log|x - y|d\mu(y)$ is in $\text{BMO}(\mathbf{R}^n)$ if $d\mu$ is a measure in $\mathbf{R}^n$ with finite total mass $\int |d\mu|$.

PROPOSITION 4.2.10. BMO($\mathbf{R}^n$)/$\mathbf{C}$ *is a Banach space with norm equal to the smallest constant* K *such that* (4.2.29) *is valid. With* $B = \{x \in \mathbf{R}^n; |y - x| < \delta\}$ *it follows from* (4.2.29) *that*

$$(4.2.29)' \qquad \delta \int_{\mathbf{R}^n} \frac{|\varphi(x) - \varphi_B|^2}{(|x-y| + \delta|)^{n+1}}\, dx \le C_n \|\varphi\|^2_{\mathrm{BMO}}.$$

PROOF. Let $B_k = \{x \in \mathbf{R}^n; |x - y| < 2^k\delta\}$ for $k = 0, 1, \dots,$ and set $c_k = \varphi_{B_k} = \int_{B_k} \varphi\, dx/m(B_k)$. If $\|\varphi\|_{\mathrm{BMO}} = 1$ then

$$\frac{1}{m(B_{k+1})} \int_{B_{k+1}} |\varphi(x) - c_{k+1}|^2\, dx \le 1, \quad \text{hence } |c_k - c_{k+1}|^2 \le 2^n,$$

because $m(B_{k+1})/m(B_k) = 2^n$. By the triangle inequality $|c_k - c_0| \le 2^{n/2}k$ and

$$\int_{B_k} |\varphi(x) - c_0|^2\, dx \le 2 \int_{B_k} (|\varphi(x) - c_k|^2 + |c_k - c_0|^2)\, dx$$
$$\le 2(1 + 2^n k^2) m(B_k) = 2(1 + 2^n k^2) 2^{kn} m(B_0).$$

Hence it follows that

$$\sum_0^\infty 2^{-k(n+1)} \int_{B_k} |\varphi(x) - c_0|^2\, dx \le C m(B_0),$$

which gives

$$2^{-n-1} \int_{\mathbf{R}^n} \frac{|\varphi(x) - c_0|^2}{(|x-y|/\delta + 1)^{n+1}}\, dx \le C m(B)$$

and proves (4.2.29)′ for another constant C_n. The rest of the proof is exactly the same as that of Proposition 4.1.12.

As in Section 4.1 we must now study the Poisson integral of a function $\varphi \in \mathrm{BMO}(\mathbf{R}^n)$, defined by (4.2.22), and prove an analogue of Lemma 4.1.13.

LEMMA 4.2.11. *If* $\varphi \in L^2(\mathbf{R}^n)$ *and* Φ *is the Poisson integral* $P_0\varphi$ *of* φ, *then*

$$(4.2.30) \qquad 2 \iint_{t>0} t |\Phi'(t,x)|^2\, dx\, dt = \|\varphi\|_2^2,$$

where $|\Phi'(t,x)|^2 = |\partial\Phi(t,x)/\partial t|^2 + \sum_1^n |\partial\Phi(t,x)/\partial x_j|^2$. *If* $\varphi(x) = 0$ *when* $|x| < \delta$ *and* $\varphi(x)/|x|^{(n+1)/2} \in L^2$, *then*

$$(4.2.31) \qquad |\Phi'(t,0)|^2 \le C_n (t+\delta)^{-1} \int |\varphi(x)|^2 |x|^{-n-1}\, dx.$$

If $\varphi \in \mathrm{BMO}(\mathbf{R}^n)$ *then we have for* $y \in \mathbf{R}^n$ *and* $\delta > 0$

$$\iint_{T_{y,\delta}} t|\Phi'(t,x)|^2\,dx\,dt \le C_n \delta^n \|\varphi\|^2_{\mathrm{BMO}}, \quad T_{y,\delta} = \{(t,x); |x-y| < \delta, 0 < t < \delta\}. \tag{4.2.32}$$

PROOF. The proof of (4.2.30) is the same as that of the first part of (4.1.22), which is independent of the dimension. To prove (4.2.31) we note that $|P_0'(t,x)| \le C(t+|x|)^{-n-1}$ for reasons of homogeneity. Since $\Phi'(t,x) = \int P_0'(t,x-y)\varphi(y)\,dy$ we obtain using Cauchy-Schwarz' inequality

$$|\Phi'(t,0)|^2 \le C^2 \int_{\mathbf{R}^n} |\varphi(y)|^2 |y|^{-n-1}\,dy \int_{|y|>\delta} |y|^{n+1}/(t+|y|)^{2n+2}\,dy.$$

In the last integral we estimate $t+|y|$ below by $\frac{1}{2}(t+\delta+|y|)$. The integral over the whole of $\mathbf{R}^n$ is then convergent and equal to a constant times $(t+\delta)^{-1}$, which proves (4.2.31). The estimate (4.2.32) follows then as in the proof of Lemma 4.1.13 by writing φ as the sum of the mean value over the ball $\{x; |y-x| < 2\delta\}$, a function supported by this ball and one which vanishes in it. The repetition is left as an exercise.

In the following analogue of Proposition 4.1.14 special properties of the Riesz operators will be essential.

PROPOSITION 4.2.12. *Let* $\varphi \in L^2(\mathbf{R}^n, dx/(1+|x|)^{n+1})$ *and assume that for the Poisson integral* Φ *of* φ *we have*

$$\int_{T_{y,\delta}} t|\Phi'(t,x)|^2\,dx\,dt \le A^2\delta^n, \quad y \in \mathbf{R}^n,\ \delta > 0, \tag{4.2.33}$$

where $T_{y,\delta}$ *is defined by* (4.2.32). *Then it follows that*

$$\Big|\int_{\mathbf{R}^n} \varphi f\,dx\Big| \le C_n A \|f\|_{\mathscr{H}^1}, \quad \text{if } f \in \mathscr{S},\ \hat{f} \in C_0^\infty(\mathbf{R}^n \setminus \{0\}), \tag{4.2.34}$$

so φ *defines a continuous linear form on* $\mathscr{H}^1(\mathbf{R}^n)$.

PROOF. We may assume that φ and f are real valued. As in the proof of Proposition 4.1.14 it follows from (4.2.30) that

$$\int_{\mathbf{R}^n} \varphi f\,dx = 2\iint_{t>0} t(\Phi'(t,x), F_0'(t,x))\,dx, \tag{4.2.35}$$

where $F_0 = P_0 f$ is the Poisson integral of f. As in the proof of (4.1.28) we shall also consider the harmonic functions $F_j = P_j f$ and the vector $\vec{F} = (F_0, \dots, F_n)$ which is the gradient of a harmonic function. The reason is that as will be verified in a moment

$$|\vec{F}'|^2 = \sum_{j,k=0}^n |\partial_j F_k|^2 \le (n+1)|\vec{F}|\Delta|\vec{F}|, \quad \text{when } \vec{F} \ne 0. \tag{4.2.36}$$

Here $\partial_j = \partial/\partial x_j$ when $j = 1, \dots, n$ and $\partial_0 = \partial/\partial t$. Since $|F_0'| \le |\vec{F}'|$ it follows from (4.2.35) that

$$\text{(4.2.37)} \quad \Big|\int \varphi f\, dx\Big| \le 2\iint_{t>0} t|\Phi'(t,x)||\vec{F}'(t,x)|\, dx\, dt \\ \le 2\Big(\iint_{t>0} t|\Phi'(t,x)|^2|\vec{F}(t,x)|\, dx\, dt\Big)^{\frac12}\Big(\iint_{t>0} t|\vec{F}'(t,x)|^2|\vec{F}(t,x)|^{-1}\, dx\, dt\Big)^{\frac12}.$$

Using (4.2.36) we obtain

$$\iint_{t>0} t|\vec{F}'(t,x)|^2|\vec{F}(t,x)|^{-1}\, dx\, dt \le (n+1)\iint_{t>0} t\Delta|\vec{F}|\, dx\, dt \\ \le (n+1)\int |\vec{F}(0,x)|\, dx \le (n+1)\|f\|_{\mathscr{H}^1},$$

exactly as in the proof of (4.1.26). In a moment we shall prove that (4.2.36) means that $|\vec{F}|^q$ is subharmonic when $q = n/(n+1)$. Accepting this for a moment we have $|\vec{F}|^q \le G$ where G is the Poisson integral of $g(x) = |\vec{F}(0,x)|^q$. Thus $G(t,x) \le Cg^{**}_{\mathrm{HL}}(x,t)$, so we have $|\vec{F}| \le C^p g^{**}(x,t)^p$ if $p = 1/q$. Hence it follows from (4.2.33) and Theorem 4.2.5′, with $p = 1/q$, that the first parenthesis in (4.2.37) can be estimated by $CA^2\|g\|_p^p = C\|\vec{F}(0,\cdot)\|_1$, which completes the proof of (4.2.34) apart from the verification of (4.2.36) and the subharmonicity of $|\vec{F}|^q$.

Differentiation of the equation $|\vec{F}|^2 = \sum_0^n F_\nu^2$ gives

$$|\vec{F}|\partial_k|\vec{F}| = \sum_{\nu=0}^n F_\nu F_{\nu k}, \quad k = 0, \dots, n,$$

where $F_{\nu k} = \partial_k F_\nu$ is *symmetric* in ν and k. Since $\Delta F_\nu = 0$ another differentiation gives

$$|\vec{F}|\Delta|\vec{F}| + \sum_0^n (\partial_k|\vec{F}|)^2 = \sum_{k,\nu=0}^n F_{\nu k}^2, \quad \text{hence}$$

$$|\vec{F}|\Delta|\vec{F}| = \sum_{k,\nu=0}^n F_{\nu k}^2 - \sum_{k=0}^n\Big(\sum_{\nu=0}^n F_{\nu k}F_\nu\Big)^2/|\vec{F}|^2.$$

We claim that the right-hand side is $\ge \sum_{k,\nu=0}^n F_{\nu k}^2/(n+1)$. By the orthogonal invariance of the Hilbert-Schmidt norm it suffices to prove this when $\vec{F}$ is along a coordinate axis, say the x_n axis. Then the inequality (4.2.36) becomes

$$\sum_{k,\nu=0}^n F_{\nu k}^2 \le (n+1)\sum_{\nu=0}^{n-1}\sum_{k=0}^n F_{\nu k}^2.$$

This is obvious for the off diagonal terms because of the symmetry, since $n+1 \ge 2$, and for the diagonal terms this means that $F_{nn}^2 \le n\sum_0^{n-1} F_{\nu\nu}^2$, which follows from the fact that the trace is equal to 0, that is, $F_{nn} = -\sum_0^{n-1} F_{\nu\nu}$. This equation which is also orthogonally invariant comes from the fact that $\vec{F}$ is the gradient of a harmonic function, the Newton potential of $2f \otimes \delta(t)$.

When $\vec{F} \neq 0$ we have

$$\Delta|\vec{F}|^q = q|\vec{F}|^{q-1}\Delta|\vec{F}| + q(q-1)|\vec{F}|^{q-2}||\vec{F}|'|^2 \geq q|\vec{F}|^{q-2}(|\vec{F}'|^2/(n+1) + (q-1)||\vec{F}|'|^2),$$

by (4.2.36). Since $|\vec{F}|\partial_k|\vec{F}| = \sum_\nu F_{k\nu}F_\nu$ we have $||\vec{F}|'| \leq |\vec{F}'|$ and it follows that $\Delta|\vec{F}|^q \geq 0$ if $\vec{F} \neq 0$ and $q - 1 + 1/(n+1) \geq 0$, that is, $q \geq n/(n+1)$. This implies that $|\vec{F}|^q$ is subharmonic and completes the proof.

Summing up, we have now extended Theorem 4.1.15 to several variables:

THEOREM 4.2.13. *The restriction of a continuous linear form L on $\mathscr{H}^1(\mathbf{R}^n)$ to the dense subset $\{f \in \mathscr{S}; \hat{f} \in C_0^\infty(\mathbf{R}^n)\setminus\{0\}\}$ is of the form $L(f) = \int f\varphi\, dx$ where φ is uniquely determined in* $\mathrm{BMO}(\mathbf{R}^n)/\mathbf{C}$. *The norm of L in the dual space of $\mathscr{H}^1(\mathbf{R}^n)$ is equivalent to the norm of φ in* $\mathrm{BMO}(\mathbf{R}^n)/\mathbf{C}$, *and every $\varphi \in \mathrm{BMO}(\mathbf{R}^n)/\mathbf{C}$ defines a continuous linear form in $\mathscr{H}^1(\mathbf{R}^n)$.*

We leave for the reader to assemble the proof using the preceding results, and pass to introducing n-dimensional atoms:

DEFINITION 4.2.14. An atom in $\mathscr{H}^1(\mathbf{R}^n)$ is a function $a \in L^2(\mathbf{R}^n)$ with support in a ball B such that

$$m(B)\int_B |a(x)|^2\, dx \leq 1, \qquad \int_B a(x)\, dx = 0. \tag{4.2.38}$$

From Lemma 4.2.8 it follows that the atoms form a bounded subset of $\mathscr{H}^1(\mathbf{R}^n)$, and Theorem 4.2.13 gives the other half of the following corollary:

COROLLARY 4.2.15. *If A is the closed convex hull of the atoms in $\mathscr{H}^1(\mathbf{R}^n)$ then there are positive constants C_n', C_n'' such that*

$$\{f \in \mathscr{H}^1(\mathbf{R}^n); \|f\|_{\mathscr{H}^1} \leq C_n'\} \subset A \subset \{f \in \mathscr{H}^1(\mathbf{R}^n); \|f\|_{\mathscr{H}^1} \leq C_n''\}. \tag{4.2.39}$$

If $C_n''' > 1/C_n'$ then every $f \in \mathscr{H}^1(\mathbf{R}^n)$ has an atomic decomposition

$$f = \sum_1^\infty \lambda_j a_j, \quad \sum_1^\infty |\lambda_j| \leq C_n'''\|f\|_{\mathscr{H}^1}, \tag{4.2.40}$$

where a_j are atoms in $\mathscr{H}^1(\mathbf{R}^n)$; we have $\|f\|_{\mathscr{H}^1} \leq C_n'' \sum_1^\infty |\lambda_j|$.

The proof is left as an exercise since it is a repetition of that of Corollary 4.1.17. We also leave as an exercise to prove the following extension of Corollary 4.1.18 to the n-dimensional case:

COROLLARY 4.2.16. *Let* $\Phi = \{\varphi \in C^1(\mathbf{R}^n); (1+|x|)^{n+1}(|\varphi(x)|+|\varphi'(x)|) \le 1\}$. *Set* $\varphi_t(x) = \varphi(x/t)/t^n$ *and* $f^* = \sup_{\varphi\in\Phi}\sup_{t>0}|f*\varphi_t|$ *for* $f \in L^1(\mathbf{R}^n)$. *Then*

$$\|f^*\|_1 \le C_n\|f\|_{\mathscr{H}^1(\mathbf{R}^n)}, \quad \text{if } f \in \mathscr{H}^1(\mathbf{R}^n). \tag{4.2.41}$$

However, we shall prove a consequence of this result and Exercise 4.2.3, due to Stein [3]:

COROLLARY 4.2.17. *If* $f \in \mathscr{H}^1(\mathbf{R}^n)$ *and* $f \ge 0$ *in an open set* $\Omega \subset \mathbf{R}^n$, *then* $|f|\log^+|f| \in L^1_{\text{loc}}(\Omega)$.

PROOF. Let $\varrho \in C_0^\infty(\Omega)$ and $0 \le \varrho \le 1$, and choose $\varphi \in \Phi$ with support in the unit ball, $\varphi \ge 0$, and $\varphi(0) > 0$. (We keep the notation in Corollary 4.2.16.) Then $0 \le g = \varrho f$, and $g^*_{\text{HL}} \le Cf^*$ in a neighborhood $K \Subset \Omega$ of $\operatorname{supp}\varrho$. In fact, $\varphi(x) > \frac{1}{2}\varphi(0) > 0$ when $|x| < c$, hence

$$0 \le \int_{|y|<\delta c} g(x-y)\,dy \le \int_{|y|<\delta c} f(x-y)\,dy \le 2\delta^n\varphi(0)^{-1}(f*\varphi_\delta)(x) \le 2\delta^n\varphi(0)^{-1}f^*(x),$$

if $x \in K$ and $\max(\delta c, \delta)$ is smaller than the distance from K to $\complement\Omega$. This implies $g^*_{\text{HL}} \le Cf^*$ in K for some C, and it is clear that $g^*_{\text{HL}}(x) \le C/(1+|x|)^n$, $x \in \complement K$, for some C. Thus

$$\int g^*_{\text{HL}}(x)^{\varepsilon+1}/(1+g^*_{\text{HL}}(x))^\varepsilon\,dx < \infty,$$

if $\varepsilon > 0$, and it follows from Exercise 4.2.3 that $|g|\log^+|g| \in L^1$.

The result should be compared to the n dimensional version of Proposition 4.1.7. Next we prove an extension of Corollary 4.1.19:

COROLLARY 4.2.18. *If* $m \in C^\nu(\mathbf{R}^n \setminus \{0\})$ *satisfies* (4.2.27) *for some* $\nu > n/2$ *then* $m(D)$ *is continuous from* $\mathscr{H}^1(\mathbf{R}^n)$ *to* $\mathscr{H}^1(\mathbf{R}^n)$ *and*

$$\|m(D)f\|_{\mathscr{H}^1} \le C_n\|f\|_{\mathscr{H}^1}, \quad f \in \mathscr{H}^1(\mathbf{R}^n). \tag{4.2.42}$$

PROOF. First we prove the weaker result

$$\|m(D)f\|_{L^1} \le C_n\|f\|_{\mathscr{H}^1}, \quad f \in \mathscr{H}^1(\mathbf{R}^n). \tag{4.2.43}$$

It is sufficient to verify this when f is an atom, and then it was proved in Lemma 4.2.8. In particular, (4.2.43) is true when $\hat{f} \in C_0^\infty(\mathbf{R}^n\setminus\{0\})$, and then it is clear that $R_j(D)m(D)f = m_j(D)f$ where $m_j = R_jm$ satisfies (4.2.27) after division by a suitable constant. Hence

$$\|m(D)f\|_{\mathscr{H}^1} = \sum_0^n \|R_j(D)m(D)f\|_{L^1} = \sum_0^n \|m_j(D)f\|_{L^1} \le C'\|f\|_{\mathscr{H}^1},$$

for all f in a dense subset of $\mathscr{H}^1(\mathbf{R}^n)$, which completes the proof.

Corollary 4.2.16 justifies the claim made earlier that nothing was lost by just including the Riesz operators R_j in the definition of $\mathscr{H}^1(\mathbf{R}^n)$.

4.3. Wavelets as bases in L^p and in $\mathscr{H}^1$. Let us first recall basic definitions and facts concerning bases in separable Banach spaces.

DEFINITION 4.3.1. A *Schauder basis* in a Banach space B is a sequence $e_j \in B$, $j = 1, 2, \dots$ such that every $x \in B$ has a unique representation as a sum

$$x = \sum_1^\infty x_j e_j, \quad \text{that is, } \|x - \sum_1^n x_j e_j\| \to 0, \text{ as } n \to \infty.$$

Here x_j are real or complex depending on whether B is a real or a complex Banach space.

PROPOSITION 4.3.2. *If $(e_j)_1^\infty$ is a Schauder basis in B then there is a constant C such that*

$$\sup_n \|\sum_1^n x_j e_j\| \le C\|x\|, \quad \text{if } x = \sum_1^\infty x_j e_j. \tag{4.3.1}$$

In particular, $|x_j|\|e_j\| \le 2C\|x\|$, so the linear forms $\sum_1^\infty x_j e_j \mapsto x_j\|e_j\|$ in B are uniformly bounded. Conversely, if $e_j \in B \setminus \{0\}$ and the finite linear combinations of $e_1, e_2, \dots$ are dense in B, then $(e_j)_1^\infty$ is a Schauder basis in B if there is a constant C such that

$$\|\sum_1^n x_j e_j\| \le C\|\sum_1^N x_j e_j\|, \quad \text{when } n \le N. \tag{4.3.2}$$

Here x_j are arbitrary scalars.

PROOF. Assume first that $(e_j)_1^\infty$ is a Schauder basis, and set

$$|||x||| = \sup_n \|\sum_1^n x_j e_j\|, \quad \text{if } x = \sum_1^\infty x_j e_j.$$

This is a new norm with $\|x\| \le |||x|||$, $x \in B$, and we shall prove that B is complete also with this norm. By Banach's theorem this will imply that $|||x||| \le C\|x\|$, which is the inequality (4.3.1). To prove the completeness we first observe that

$$|x_j|\|e_j\| \le 2|||x|||, \quad \text{if } x = \sum_1^\infty x_j e_j,$$

so the map $x \mapsto |x_j|$ is continuous for this larger norm. If $x^\nu = \sum_1^\infty x_k^\nu e_k$ and $|||x^\nu - x^\mu||| \to 0$ as $\nu, \mu \to \infty$, it follows that $x_k^\nu - x_k^\mu \to 0$, hence $\lim_{\nu\to\infty} x_k^\nu = x_k$ exists for every k. Since

$$\|\sum_1^n (x_k^\nu - x_k^\mu) e_k\| \le |||x^\nu - x^\mu|||,$$

we have for every $\varepsilon > 0$

$$\|\sum_1^n (x_k^\nu - x_k)e_k\| \le \lim_{\mu\to\infty} |||x^\nu - x^\mu||| < \varepsilon$$

if $\nu > \nu_\varepsilon$. Hence

$$\|\sum_{n+1}^m x_k e_k\| \le \|\sum_{n+1}^m x_k^\nu e_k\| + 2\varepsilon < 3\varepsilon, \quad n < m,$$

if n and m are large enough, so $\sum_1^\infty x_k e_k = x$ exists and $|||x^\nu - x||| \le \varepsilon$ when $\nu > \nu_\varepsilon$. This completes the proof of the first part.

Now assume that (4.3.2) is valid. Then

$$\max_{j\le N} |x_j|\|e_j\| \le 2C\|\sum_1^N x_j e_j\|,$$

which proves that the elements e_j are linearly independent and that the map $\sum_1^N x_j e_j \mapsto x_j\|e_j\|$ extends from the set E of finite linear combinations of the elements e_j to a linear form L_j on B with norm $\le 2C$. We have

$$\sup_n \|\sum_1^n L_j(x)e_j/\|e_j\|\| \le C\|x\|, \quad x \in E.$$

Since the maps $B \ni x \mapsto \sum_1^n L_j(x)e_j/\|e_j\| \in B$ are uniformly bounded and converge to the identity when $x \in E$, it follows that they converge strongly to the identity, so $x = \sum_1^\infty L_j(x)e_j/\|e_j\|$ for every $x \in B$. If $\sum_1^\infty x_j e_j = 0$ then

$$\|\sum_1^n x_j e_j\| \le \lim_{N\to\infty} \|\sum_1^N x_j e_j\| = 0$$

which proves that $x_j = 0$ for every j. Thus $(e_j)_1^\infty$ is a Schauder basis.

The useful property of a Schauder basis is that with the notation in the proof the projections $x \mapsto \sum_1^n L_j(x)e_j/\|e_j\|$ on the linear span of the first n elements have a uniformly bounded norm. This allows approximation of arbitrary bounded operators by operators of finite rank with uniformly bounded norm.

A series $\sum_1^\infty y_j$ with $y_j \in B$ is said to converge *unconditionally* to $y \in B$ if for every $\varepsilon > 0$ there is a finite subset I of $\mathbf{N} = \{1, 2, \dots\}$ such that $\|\sum_{j\in J} y_j - y\| < \varepsilon$ for all finite subsets J of $\mathbf{N}$ containing I. This implies that $\sum_1^\infty y_j = y$ and that this remains true for an arbitrary rearrangement of the terms; conversely, if the convergence is independent of the ordering then the series converges unconditionally.

Definition 4.3.3. A Schauder basis $(e_j)_1^\infty$ in B is called *unconditional* if every $x \in B$ has a unique representation $x = \sum_1^\infty x_j e_j$ and the series is unconditionally convergent.

Proposition 4.3.4. *If* $(e_j)_1^\infty$ *is an unconditional Schauder basis in* B *then there is a constant* C *such that*

$$\sup_J \|\sum_{j\in J} x_j e_j\| \le C\|x\|, \quad \text{if } x = \sum_1^\infty x_j e_j. \tag{4.3.3}$$

Here J *is an arbitrary finite subset of* $\mathbf{N}$. *Conversely, if* $e_j \in B \setminus \{0\}$ *and the finite linear combinations are dense in* B, *then* $(e_j)_1^\infty$ *is an unconditional Schauder basis if there is a constant* C *such that*

$$\|\sum_{j\in J} x_j e_j\| \le C\|\sum_{i\in I} x_i e_i\|, \quad \text{if } J \subset I, \tag{4.3.4}$$

where I *is any finite subset of* $\mathbf{N}$ *and* x_i *are scalars.*

The proof is essentially a repetition of the proof of Proposition 4.3.2, now with

$$|||x||| = \sup_J \|\sum_{j\in J} x_j e_j\|, \quad \text{if } x = \sum_1^\infty x_j e_j,$$

so we leave it as an exercise.

The inequality (4.3.4) is equivalent to

$$\|\sum_{i\in I} \lambda_i x_i e_i\| \le C\|\sum_{i\in I} x_i e_i\|, \quad \text{if } 0 \le \lambda_i \le 1,\ i \in I. \tag{4.3.4$'$}$$

In fact, (4.3.4) means that this is true for all $\lambda \in \{0,1\}^I$, while (4.3.4)$'$ states that this is true for all λ in the cube with these vertices, which follows from the convexity of the norm. From (4.3.4)$'$ it follows that

$$\|\sum_{i\in I} \lambda_i x_i e_i\| \le 2C\|\sum_{i\in I} x_i e_i\|, \quad \text{if } -1 \le \lambda_i \le 1,\ i \in I, \tag{4.3.4$''$}$$

and even for complex scalars λ_i we obtain

$$\|\sum_{i\in I} \lambda_i x_i e_i\| \le 4C\|\sum_{i\in I} x_i e_i\|, \quad \text{if } |\lambda_i| \le 1,\ i \in I. \tag{4.3.4$'''$}$$

We can take any one of these conditions as a criterion for unconditional Schauder bases. Roughly speaking they mean that the norm of a linear combination of the basis elements is essentially determined by the absolute values of the coefficients, independently of the arguments (or signs).

It is known that there is no unconditional Schauder basis in $L^1(\mathbf{R}^n)$, but wavelets provide such bases in the L^p spaces which have one:

THEOREM 4.3.5. *Let ψ_r, $r \in \{0,1\}^n \setminus \{0\}$ be orthonormal wavelets in $C_0^1(\mathbf{R}^n)$, as in Theorem 3.2.6. Then the orthonormal basis*

$$\psi_{r,j,k}(x) = 2^{nj/2}\psi_r(2^j x - k), \quad 0 \neq r \in \{0,1\}^n,\ j \in \mathbf{Z},\ k \in \mathbf{Z}^n,$$

in $L^2(\mathbf{R}^n)$ is an unconditional basis in $L^p(\mathbf{R}^n)$ when $1 < p < \infty$. If $\psi_r \in C_0^2(\mathbf{R}^n)$ it is also an unconditional basis in $\mathscr{H}^1(\mathbf{R}^n)$.

The theorem remains true with much weaker assumptions on smoothness and for wavelets which just decay sufficiently fast at infinity, but the hypotheses made here are convenient in the proof and we know from Chapter III that such wavelets exist. Since $\int \psi_r(x)\,dx = 0$ (see Proposition 3.3.1 for the one-dimensional case) we know that $2^{nj/2}\psi_{r,j,k}/C$ is an atom in $\mathscr{H}^1(\mathbf{R}^n)$ for some constant C, and it is clear that $\psi_{r,j,k} \in L^p(\mathbf{R}^n)$ for $1 \le p \le \infty$. The first step in the proof of Theorem 4.3.5 is to establish completeness of the wavelets.

LEMMA 4.3.6. *If $f \in C_0^\infty(\mathbf{R}^n)$ then the wavelet expansion $\sum \psi_{r,j,k}(f, \psi_{r,j,k})$ converges to f in L^p for $1 < p \le \infty$. If $\int_{\mathbf{R}^n} f(x)\,dx = 0$ it converges to f in $\mathscr{H}^1(\mathbf{R}^n)$.*

PROOF. We know already that the series converges to f in $L^2(\mathbf{R}^n)$. It is therefore sufficient to prove that it converges in L^p (resp. $\mathscr{H}^1$), for the sum must then be equal to f. To estimate the coefficients

$$f_{r,j,k} = \int f(x)\psi_{r,j,k}(x)\,dx = 2^{nj/2}\int f(x)\psi_r(2^j x - k)\,dx$$

we first assume $j \ge 0$. Since $\int \psi_r(x)\,dx = 0$ we have

$$f_{r,j,k} = 2^{nj/2}\int (f(x) - f(k/2^j))\psi_r(2^j x - k)\,dx,$$

and since $|f(x) - f(k/2^j)| \le C|x - k/2^j|$, we obtain

$$|f_{r,j,k}| \le 2^{nj/2}C\int |x - k/2^j||\psi_r(2^j x - k)|\,dx \le C' 2^{-nj/2-j}.$$

Hence $\|\psi_{r,j,k}f_{r,j,k}\|_\infty \le C2^{-j}$, and since only a bounded number of supports can overlap for fixed j this proves uniform convergence of the sum for $j \ge 0$. We have $\|\psi_{r,j,k}f_{r,j,k}\|_{\mathscr{H}^1} \le C2^{-j-nj}$, and since $f_{r,j,k} = 0$ unless $|2^j x - k| \le C$ for some $x \in \operatorname{supp} f$, the number of terms is $O(2^{nj})$ which proves $\mathscr{H}^1$ convergence too. Thus there is convergence in L^p for $1 \le p \le \infty$.

Now assume that $j < 0$. Then there is a fixed bound for the number of non-zero coefficients $f_{r,j,k}$ with the same j, and $|f_{r,j,k}| \le C2^{nj/2}$. Since $\|\psi_{r,j,k}\|_p = 2^{nj(1/2-1/p)}\|\psi_r\|_p$ we have $\|\psi_{r,j,k}f_{r,j,k}\|_p \le C2^{nj(1-1/p)}$ which proves L^p convergence of the sum for $j < 0$ when $p > 1$. When $\int f(x)\,dx = 0$ and $f(y)\psi_r(2^j y - k) \neq 0$ for some y we have

$$|f_{r,j,k}| = |2^{nj/2}\int f(x)(\psi_r(2^j x - k) - \psi_r(2^j y - k))\,dx|$$
$$\le C2^{nj/2}\int |f(x)|2^j|x - y|\,dx \le C'2^{nj/2+j}.$$

Since $2^{nj/2}\psi_{r,j,k}$ has bounded norm in $\mathscr{H}^1$ it follows that

$$\|\psi_{r,j,k}f_{r,j,k}\|_{\mathscr{H}^1} \le C2^j,$$

and we get convergence of the sum for $j<0$ in $\mathscr{H}^1$ too.

The second part of the proof of Theorem 4.3.5 is to verify an estimate of the form (4.3.4). Thus we must prove for any finite subset J of $\{(r,j,k); 0\ne r\in\{0,1\}^n, j\in\mathbf{Z}, k\in\mathbf{Z}^n\}$ that the norm of the operator

$$\mathscr{K}_J : f\mapsto \sum_{(r,j,k)\in J}\psi_{r,j,k}(f,\psi_{r,j,k}) \tag{4.3.5}$$

in $L^p(\mathbf{R}^n)$, $1<p<\infty$, and in $\mathscr{H}^1(\mathbf{R}^n)$ has a bound independent of J. The kernel of $\mathscr{K}_J$ is

$$K_J(x,y) = \sum_{(r,j,k)\in J} 2^{nj}\psi_r(2^jx-k)\psi_r(2^jy-k). \tag{4.3.6}$$

The norm of the orthogonal projection operator $\mathscr{K}_J$ in L^2 is ≤ 1, and we shall estimate the norm in the other spaces by modifying the study of singular integral operators in Section 4.2. First we prove a substitute for the regularity property of the convolution kernel M given by (4.2.14).

Lemma 4.3.7. *If K_J is defined by (4.3.6) with a finite subset J of $(\{0,1\}^n\setminus\{0\})\times\mathbf{Z}\times\mathbf{Z}^n$ then*

$$|K_J(x,y)|\le C|x-y|^{-n},\quad |\partial K_J(x,y)/\partial(x,y)|\le C|x-y|^{-n-1}, \tag{4.3.7}$$

where C is independent of J. For the operator $\mathscr{K}_J$ with kernel K_J we have

$$m(\{x; |\mathscr{K}_Jf(x)|>\alpha\})\le C\|f\|_1/\alpha,\quad f\in L^1(\mathbf{R}^n)\cap L^2(\mathbf{R}^n), \tag{4.3.8}$$

$$\|\mathscr{K}_Jf\|_p \le \begin{cases} Cp'^{1/p}\|f\|_p, & \text{if } 1<p\le 2,\\ Cp^{1/p'}\|f\|_p, & \text{if } 2\le p<\infty,\end{cases}\quad \text{for } f\in L^1(\mathbf{R}^n)\cap L^\infty(\mathbf{R}^n). \tag{4.3.9}$$

Proof. If $|x|\le R$ when $x\in\operatorname{supp}\psi_r$ then $|2^jx-k|\le R$ and $|2^jy-k|\le R$ in the support of the terms in (4.3.6). This implies $2^j|x-y|\le 2R$. For fixed (x,y) and r,j the number of such terms is at most equal to the largest number N of lattice points in a ball of radius R. Hence

$$|K_J(x,y)|\le C\sum_r\sum_{2^j|x-y|\le 2R}2^{nj}\sup\psi_r^2 \le 2C\sum_r \sup\psi_r^2(2R/|x-y|)^n$$

which proves the first estimate in (4.3.7). Differentiation with respect to x or y contributes another factor 2^j which gives the second estimate (4.3.7).

If $I \subset \mathbf{R}^n$ is a cube with center z and side s, and $2I$ is the cube with center z and side $2s$, it follows from the second part of (4.3.7) that

$$\int_{\complement 2I} |K_J(x,y) - K_J(x,z)|\,dx \le C's \int_{\complement 2I} |x-z|^{-n-1}\,dx = C'', \quad y \in I,$$

where C'' is independent of I. Hence

$$\int_{x \notin 2I} |\mathscr{K}_J w(x)|\,dx \le C'' \int_I |w(y)|\,dy, \quad \text{if } w \in L^1, \int_I w\,dy = 0, \ \operatorname{supp} w \subset I, \tag{4.3.10}$$

for we have

$$\mathscr{K}_J w(x) = \int_I (K_J(x,y) - K_J(x,z))w(y)\,dy.$$

Now the estimates (4.3.8) and (4.3.9) follow from the proof of Theorem 4.2.2, for (4.3.10) is a substitute for (4.2.14)$'$ there, and K_J has norm ≤ 1 in $L^2(\mathbf{R}^n)$.

With this lemma we have completed the proof of the statement on L^p spaces in Theorem 4.3.5. It is also very easy to see that

$$\|\mathscr{K}_J f\|_1 \le C\|f\|_{\mathscr{H}^1}, \tag{4.3.11}$$

for by Corollary 4.2.15 it suffices to prove this for an atom f. If f has support in a cube I and $\sqrt{m(I)}\|f\|_2 \le 1$, $\int f\,dx = 0$, then $\|f\|_1 \le 1$ and it follows from (4.3.10) that

$$\int_{x \notin 2I} |\mathscr{K}_J f(x)|\,dx \le C.$$

Since $\sqrt{m(I)}\|\mathscr{K}_J f\|_2 \le 1$ we have

$$\int_{2I} |\mathscr{K}_J f(x)|\,dx \le 2^n,$$

which gives $\|\mathscr{K}_J f\|_1 \le C + 2^n$ and proves (4.3.11) with another C.

To estimate $\|\mathscr{K}_J f\|_{\mathscr{H}^1}$ we must also estimate $\|R_\nu \mathscr{K}_J f\|_1$ where R_ν is one of the Riesz operators. The kernel of $R_\nu \mathscr{K}_J$ is

$$K_J^\nu(x,y) = \sum_{(r,j,k)\in J} 2^{nj}\psi_r^\nu(2^j x - k)\psi_r(2^j y - k), \tag{4.3.6$'$}$$

where $\psi_r^\nu = R_\nu \psi_r$, for the Riesz operators commute with translations and scale changes. Since $\int \psi_r(x)\,dx = 0$ we have $\psi_r^\nu(x) = O(|x|^{-n-1})$ as $x \to \infty$, and the derivatives decrease even faster. We have with a constant c

$$\psi_r^\nu(x) = c\int (\psi_r(x-y) - \psi_r(x))y_\nu |y|^{-n-1}\,dy$$

with absolute and uniform convergence since $\psi_r \in C_0^1$. Assuming that $\psi_r \in C_0^2$ we can differentiate under the integral sign and conclude that $\psi_r^\nu \in C^1$.

We are now ready to prove that

$$(4.3.7)' \qquad |K_J^\nu(x,y)| \le C|x-y|^{-n}, \quad |\partial K_J^\nu(x,y)/\partial(x,y)| \le C|x-y|^{-n-1},$$

where C is independent of J. The proof does not differ much from the proof of (4.3.7) and we keep the notation used there. For fixed j, r we only have N non-zero terms to consider, since $|2^j y - k| \le R$, and we have

$$2^j|x-y| \le |2^j x - k| + |2^j y - k| \le |2^j x - k| + R$$

then. Since $\psi_r^\nu(2^j x - k)(1 + |2^j x - k|)^{n+1}$ is bounded, it follows that

$$|K_J^\nu(x,y)| \le C \sum 2^{nj}(1 + 2^j|x-y|)^{-n-1}.$$

The sum when $2^j|x-y| \le 1$ was estimated in the proof of (4.3.7), and the sum when $2^j|x-y| > 1$ is at most $|x-y|^{-n-1}(2|x-y|) = 2|x-y|^{-n}$. This proves the first estimate $(4.3.7)'$, and the second follows in the same way since differentiation of $K_J^\nu(x,y)$ only contributes a factor 2^j to the estimates. However, now we need the estimate

$$(1+|x|)^{n+2}(|\psi_r^\nu(x)| + |\partial\psi_r^\nu(x)/\partial x|) \le C,$$

which follows since $\int x^\alpha \psi_r(x)\,dx = 0$ when $|\alpha| \le 1$. (See Proposition 3.3.1.)

From $(4.3.7)'$ it follows by repetition of the proof of (4.3.11) that

$$(4.3.12) \qquad \|\mathscr{K}_J f\|_{\mathscr{H}^1} = \sum_{\nu=0}^{n} \|\mathscr{K}_J^\nu f\|_1 \le C\|f\|_{\mathscr{H}^1},$$

which completes the proof of Theorem 4.3.5.

4.4. More on $\mathscr{H}^1$ atoms and on BMO. The use of L^2 norms in the defining property (4.2.29) of BMO($\mathbf{R}^n$) may seem surprising since after all BMO is a space closely related to L^∞. By duality we obtained the defining property (4.2.38) of atoms in $\mathscr{H}^1(\mathbf{R}^n)$ which also involved L^2 norms although $\mathscr{H}^1$ is closely related to L^1. We shall now prove that L^2 actually has no special role in these contexts although it was convenient in the developments in Section 4.2.

If I is a cube in $\mathbf{R}^n$ we shall denote by $\widetilde{I}$ the set of all cubes $\subset I$ obtained when I is first divided into 2^n equal cubes, and the process is repeated indefinitely for the cubes so obtained.

Theorem 4.4.1 (John-Nirenberg). *Let $f \in L^1(I_0)$ where I_0 is a cube in $\mathbf{R}^n$, and assume that there is a constant K such that for every cube $I \in \widetilde{I}_0$*

$$(4.4.1) \qquad \frac{1}{m(I)}\int_I |f(x) - f_I|\,dx \le K, \quad f_I = \frac{1}{m(I)}\int_I f(y)\,dy.$$

Then it follows that

$$m(\{x \in I_0; |f(x) - f_{I_0}| > \sigma\}) \le e e^{-a\sigma/K} m(I_0), \quad \sigma > 0, \tag{4.4.2}$$

$$\Big(\frac{1}{m(I_0)} \int_{I_0} |f(x) - f_{I_0}|^p \, dx\Big)^{\frac{1}{p}} \le a^{-1} epK, \tag{4.4.3}$$

where $a = 2^{-n}e^{-1}$.

Before the proof we make a few observations.

1. If instead of (4.4.1) we had only assumed that

$$\frac{1}{m(I)} \int_I |f(x) - c| \, dx \le K \quad \text{for some } c, \tag{4.4.1$'$}$$

then (4.4.1) would have been fulfilled with K replaced by $2K$. In fact, (4.4.1)$'$ implies that $|f_I - c| \le K$, hence $\int_I |f(x) - f_I| \, dx/m(I) \le 2K$. Thus $c = f_I$ is always a good choice although it is not always exactly minimizing except in the L^2 norm.

2. Since $(\int_I |f(x)-f_I|^p \, dx/m(I))^{1/p}$ is an increasing function of $p \in [1,\infty)$, the condition (4.4.1) is for a fixed I weaker than the corresponding L^p condition. However, (4.4.3) shows that apart from the size of the constants such L^p conditions posed for *all* $I \in \widetilde{I}$ are in fact independent of p.

Proof of Theorem 4.4.1. Replacing f by f/K we may assume that $K = 1$, which simplifies the notation. In particular, $\int_{I_0} |f(x) - f_{I_0}| \, dx \le m(I_0)$. Denote by $F(\sigma)$ the smallest constant such that

$$m(\{x \in I_0; |f(x) - f_{I_0}| > \sigma\}) \le F(\sigma) m(I_0) \tag{4.4.4}$$

for all f satisfying the hypotheses of Theorem 4.4.1 with $K = 1$. Note that $F(\sigma) \le 1$ and that $F(\sigma)$ is obviously invariant under translation and scale changes so it is independent of I_0. Using the Calderón-Zygmund decomposition (Lemma 4.2.3) we shall prove that

$$F(\sigma + 2^n s) \le F(\sigma)/s, \quad \text{if } \sigma > 0, \ s > 1. \tag{4.4.5}$$

We apply the lemma to $f(x) - f_{I_0}$ with $s > \int_{I_0} |f(x) - f_{I_0}| \, dx/m(I_0)$, in particular any $s > 1$. In the decomposition $f - f_{I_0} = v + \sum_1^\infty w_k$ we have $|v| \le 2^n s$ almost everywhere, and the cubes I_k are in $\widetilde{I}_0$ by the proof of Lemma 4.2.3. If $|f(x) - f_{I_0}| > \sigma + 2^n s$ it follows that $x \in I_k$ and that $|w_k(x)| > \sigma$, for some k. (We ignore null sets throughout.) Hence

$$\begin{aligned} m(\{x \in I_0; |f(x) - f_{I_0}| > \sigma + 2^n s\} &\le \sum_k m(\{x \in I_k; w_k(x) > \sigma\}) \\ &\le F(\sigma) \sum_k m(I_k) \le F(\sigma) \int_{I_0} |f(x) - f_{I_0}| \, dx/s \le F(\sigma) s^{-1} m(I_0), \end{aligned}$$

where the second inequality follows from the definition of F, applied to I_k, and the third follows from (4.2.13)$'$.

When $s = e$ it follows from (4.4.5) that

$$F(\sigma + 2^n e) \le F(\sigma)e^{-1}, \quad \text{if } \sigma > 0. \tag{4.4.6}$$

Since $F \le 1$ we have $F(\sigma) \le ee^{-a\sigma}$ when $0 < \sigma \le 2^n e$ if $2^n ea = 1$, and then it follows inductively from (4.4.6) that

$$F(\sigma) \le ee^{-a\sigma}, \quad \sigma > 0, \tag{4.4.7}$$

which proves (4.4.2).

Since

$$\begin{aligned}\int_{I_0} |f(x) - f_{I_0}|^p\, dx &= \int_0^\infty p\sigma^{p-1} m(\{x \in I_0; |f(x) - f_{I_0}| > \sigma\}\, d\sigma \\ &\le em(I_0)\int_0^\infty p\sigma^{p-1}e^{-a\sigma}\, d\sigma = a^{-p}em(I_0)\Gamma(p+1),\end{aligned} \tag{4.4.8}$$

and $e\Gamma(p+1) \le (ep)^p$ when $p \ge 1$, the estimate (4.4.3) follows.

Corollary 4.4.2. *If $f \in L^1_{\text{loc}}(\mathbf{R}^n)$ and for every axis parallel cube $I \subset \mathbf{R}^n$*

$$\frac{1}{m(I)}\int_I |f(x) - f_I|\, dx \le K, \quad f_I = \frac{1}{m(I)}\int_I f(y)\, dy, \tag{4.4.9}$$

then $f \in \text{BMO}(\mathbf{R}^n)$ and $\|f\|_{\text{BMO}} \le CK$. We have $f \in L^p_{\text{loc}}(\mathbf{R}^n)$ for every $p \in [1, \infty)$.

Proof. By Theorem 4.4.1 it follows from (4.4.9) that

$$\Big(\frac{1}{m(I)}\int_I |f(x) - f_I|^2\, dx\Big)^{1/2} \le CK. \tag{4.4.10}$$

Every ball B is contained in a cube I such that $m(I) \le C_n m(B)$, so we may replace the cube in (4.4.10) by a ball if the constant is replaced by $\sqrt{C_n}C$. This means that our definition (4.2.29) is fulfilled. From Theorem 4.4.1 it follows also that $f \in L^p$ in every cube, which completes the proof.

Remark. We can of course replace cubes by balls also in the hypothesis (4.4.9).

To prepare for the next corollary we prove a lemma:

Lemma 4.4.3. *Let $1 < p \le \infty$. If $f \in L^p(\mathbf{R}^n)$ and* $\operatorname{supp} f \subset B$ *where B is a ball in $\mathbf{R}^n$, and if $\int f\, dx = 0$, then $f \in \mathscr{H}^1(\mathbf{R}^n)$ and*

$$\|f\|_{\mathscr{H}^1} \le Cp(p-1)^{-1} m(B)^{1-1/p}\|f\|_p, \tag{4.4.11}$$

where C only depends on n.

Proof. Since $m(B)^{-1/p}\|f\|_p \ge m(B)^{-1/2}\|f\|_2$ when $p \ge 2$, the statement follows from Lemma 4.2.8 then, so we assume that $1 < p \le 2$. As in the proof of Lemma 4.2.8 we may also assume that B is the unit ball. By Theorem 4.2.2 we have for the Riesz transforms

$$\|R_j f\|_p \le C\|f\|_p/(p-1),$$

hence by Hölder's inequality

$$\int_{|x|\le 2} |R_j f(x)|\,dx \le C(2^n m(B))^{1-1/p}\|f\|_p/(p-1).$$

Since

$$|R_j f(x)| = c\Big|\int_{|y|<1} f(y)((x_j-y_j)|x-y|^{-n-1} - x_j|x|^{-n-1})\,dy\Big|$$
$$\le C|x|^{-n-1}\int_{|y|<1}|f(y)|\,dy, \quad \text{if } |x|\ge 2,$$

it follows that

$$\int_{|x|\ge 2}|R_j f(x)|\,dx \le C'\|f\|_1 \le C''\|f\|_p,$$

which proves the lemma.

REMARK. The same proof gives Lemma 4.2.8 at once if we assume that m satisfies (4.2.6).

In the following corollary of Theorem 4.4.1 we shall call a function $f\in L^p(\mathbf{R}^n)$ an $\mathscr{H}^1$ atom of type p, $p>1$, if there is a ball $B\subset\mathbf{R}^n$ such that $\operatorname{supp} f\subset B$, $\int_B f(x)\,dx = 0$ and $m(B)^{1-1/p}\|f\|_p\le 1$. Thus the atoms in Definition 4.2.14 are of type 2.

COROLLARY 4.4.4. *For every $p\in(1,\infty]$ there are positive constants C'_{np} and C''_{np} such that*

$$\{f\in\mathscr{H}^1(\mathbf{R}^n);\|f\|_{\mathscr{H}^1}\le C'_{np}\}\subset A_p\subset\{f\in\mathscr{H}^1(\mathbf{R}^n);\|f\|_{\mathscr{H}^1}\le C''_{np}\}, \tag{4.4.12}$$

where A_p is the closed convex hull in $\mathscr{H}^1(\mathbf{R}^n)$ of the atoms of type p. If $C'''_{n,p} > 1/C'_{np}$ then every $f\in\mathscr{H}^1(\mathbf{R}^n)$ has an atomic decomposition

$$f=\sum_1^\infty \lambda_j a_j,\quad \sum_1^\infty|\lambda_j|\le C'''_{np}\|f\|_{\mathscr{H}^1}, \tag{4.4.13}$$

where a_j are atoms of type p in $\mathscr{H}^1(\mathbf{R}^n)$; we have $\|f\|_{\mathscr{H}^1}\le C''_{np}\sum_1^\infty|\lambda_j|$.

The proof is again a repetition of that of Corollary 4.1.17 and we leave it as an exercise. The most precise atomic decomposition is of course the one with $p=\infty$.

We shall now prove a result closely related to Theorem 4.4.1 which proves that also L^p spaces have a characterization similar to that of BMO. This will be useful in the proof of interpolation theorems. We keep the notation $\widetilde{I}_0$ in Theorem 4.4.1.

THEOREM 4.4.5. *Let $f\in L^1(I_0)$ where I_0 is a cube in $\mathbf{R}^n$, and set*

$$f^\sharp(x)=\sup_{x\in I\in\widetilde{I}_0}\frac{1}{m(I)}\int_I|f(y)-f_I|\,dy;\quad f_I=\frac{1}{m(I)}\int_I f(y)\,dy. \tag{4.4.14}$$

If $f^\sharp \in L^p(I_0)$ *and* $1 < p < \infty$ *it follows that* $f \in L^p(I_0)$ *and that*

$$\|f - f_{I_0}\|_{L^p(I_0)} \le C\|f^\sharp\|_{L^p(I_0)}, \tag{4.4.15}$$

where C *only depends on* p *and* n.

Proof. To simplify notation we assume that $f_{I_0} = 0$ and that

$$\frac{1}{m(I_0)} \int_{I_0} |f^\sharp(y)|^p\, dy = 1. \tag{4.4.16}$$

Since $f^\sharp(x) \ge \int_{I_0} |f(y)|\, dy/m(I_0)$ this implies that

$$\frac{1}{m(I_0)} \int_{I_0} |f(y)|\, dy \le 1, \tag{4.4.17}$$

which allows us to make a Calderón-Zygmund decomposition of f according to Lemma 4.2.3, for every $s > 1$. Let I_k^s, v^s and w_k^s denote the cubes and functions in this decomposition, and set

$$\mu(s) = \sum_k m(I_k^s). \tag{4.4.18}$$

The proof of Lemma 4.2.3 shows that if $s_1 < s_2$ then each cube $I_k^{s_2}$ is contained in a cube $I_j^{s_1}$, so $\mu(s)$ is decreasing. We claim that

$$\mu(s) \le m(\{x \in I_0; f^\sharp(x) > As\}) + 2A\mu(2^{-n-1}s), \quad s > 2^{n+1},\ A > 0. \tag{4.4.19}$$

Let $s' = 2^{-n-1}s$, thus $s' \ge 1$. For every cube I_k^s we can find a cube $I_j^{s'}$ with $I_k^s \subset I_j^{s'}$. If $f^\sharp(x) \le As$ for some $x \in I_j^{s'} = I$ then

$$\frac{1}{m(I)} \int_I |f(x) - f_I|\, dx \le As.$$

Since $|f_I| \le 2^n s' = s/2$, by (4.2.13) with s replaced by s', we obtain using (4.2.13) once more

$$\int_{I_k^s} |f(x) - f_I|\, dx \ge \int_{I_k^s} |f(x)|\, dx - |f_I| m(I_k^s) \ge \tfrac{1}{2} s m(I_k^s).$$

Hence

$$\tfrac{1}{2} s \sum_{I_k^s \subset I} m(I_k^s) \le \int_I |f(x) - f_I|\, dx \le m(I)As,$$

which proves that

$$\sum_{I_k^s \subset I_j^{s'}} m(I_k^s) \le 2Am(I_j^{s'}). \tag{4.4.20}$$

Since all cubes I_k^s contained in a cube $I_j^{s'}$ where $f^\sharp > As$ are contained in $\{x; f^\sharp(x) > As\}$, we have proved (4.4.19).

From (4.4.19) it follows for $S > 2^{n+1}$ that

$$p\int_{2^{n+1}}^{S} s^{p-1}\mu(s)\,ds \le p\int_0^S s^{p-1}m(\{x\in I_0; f^\sharp(x) > As\})ds + 2Ap\int_{2^{n+1}}^{S} s^{p-1}\mu(2^{-n-1}s)\,ds.$$

The first term in the right-hand side is bounded by

$$pA^{-p}\int_0^\infty m(\{x\in I_0; f^\sharp(x) > s\})\,ds = A^{-p}\int_{I_0} |f^\sharp(x)|^p\,dx.$$

The second term in the right-hand side is equal to

$$2Ap2^{(n+1)p}\int_1^{2^{-n-1}S} s^{p-1}\mu(s)\,ds.$$

If we choose $1/A = 4\cdot 2^{(n+1)p}$ it follows that

$$p\int_{2^{n+1}}^{S} s^{p-1}\mu(s)\,ds \le 2A^{-p}\int_{I_0}|f^\sharp(x)|^p\,dx + p\int_1^{2^{n+1}} s^{p-1}\mu(s)\,ds,$$

and when $S\to+\infty$ we obtain using (4.4.16)

$$p\int_0^\infty s^{p-1}\mu(s)\,ds \le Cm(I_0), \tag{4.4.21}$$

because $\mu(s)\le m(I_0)$.

It remains to connect the integral on the left to the L^p norm of f. To do so we shall examine the maximal function

$$f^*(x) = \sup_{x\in I\in\widetilde{I}_0}\frac{1}{m(I)}\int_I |f(y)|\,dy,$$

which is $\ge |f(x)|$ almost everywhere. We shall estimate the measure of

$$E(s) = \{x\in I_0; f^*(x) > s\},$$

which is the union of the intervals $I\in\widetilde{I}_0$ such that $\int_I |f(y)|\,dy/m(I) > s$.

For every $I\in\widetilde{I}_0$ and $s>1$ we have

$$\int_I |f(y)|\,dy = \sum_k\int_{I_k^s\cap I}|f(y)|\,dy + \int_{I\setminus\cup I_k^s}|f(y)|\,dy.$$

In the last integral we have $|f(y)| \le s$ by (4.2.12), and when $I_k^s \subset I$ we have $\int_{I_k^s} |f(y)|\, dy \le 2^n s m(I_k^s)$ by (4.2.13), hence

$$\int_I |f(y)|\, dy \le 2^n s m(I)$$

unless $I \subset I_k^s$ for some k. In fact, two cubes in $\widetilde{I}_0$ have either disjoint interiors or else one is contained in the other. Thus we conclude that $E(2^n s) \subset \cup I_k^s$, hence

$$m(E(2^n s)) \le \mu(s), \quad \text{if } s \ge 1.$$

This proves that

$$\int_{I_0} |f(x)|^p\, dx \le \int_{I_0} |f^*(x)|^p\, dx = p\int_0^\infty s^{p-1} m(E(s))\, ds$$
$$= 2^{np} p \int_0^\infty s^{p-1} m(E(2^n s))\, ds \le 2^{np} p \int_1^\infty s^{p-1}\mu(s)\, ds + 2^{np} m(I),$$

and by (4.4.21) this gives $\int_{I_0} |f(x)|^p\, dx \le Cm(I)$ and completes the proof.

Corollary 4.4.6. *Let $f \in L^1_{\text{loc}}(\mathbf{R}^n)$ and set*

$$f^\sharp(x) = \sup_{x\in I} \frac{1}{m(I)} \int_I |f(x) - f_I|\, dx; \quad f_I = \frac{1}{m(I)} \int_I f(y)\, dy, \tag{4.4.22}$$

where I runs over dyadic cubes, defined by $0 \le 2^j x_\nu - k_\nu \le 1$, $1 \le \nu \le n$, where $k \in \mathbf{Z}^n$ and $j \in \mathbf{Z}$. If $f^\sharp \in L^p(\mathbf{R}^n)$ and $1 < p < \infty$ it follows that $f - c \in L^p(\mathbf{R}^n)$ for some constant c, and that

$$\|f - c\|_{L^p(\mathbf{R}^n)} \le C_p \|f^\sharp\|_{L^p(\mathbf{R}^n)}. \tag{4.4.23}$$

We have $c = \lim_{m(E)\to\infty} \int_E f(x)\, dx/m(E)$ for arbitrary measurable sets E, and $c = 0$ if $f \in L^q$ for some $q \in [1, \infty)$.

Proof. If we apply Theorem 4.4.5 to the cube I_N defined by $|x_\nu| \le N$ for $\nu = 1, \dots, n$, it follows that

$$\|f - c_N\|_{L^p(I_N)} \le C\|f^\sharp\|_{L^p(I_N)}, \quad c_N = f_{I_N}.$$

This implies that $|c_{N+1} - c_N| m(I_N)^{1/p} \le 2C\|f^\sharp\|_{L^p(\mathbf{R}^n)}$, so c_N has a limit c as $N \to \infty$ which has the desired property. The last statement follows since

$$\int_E |f(x) - c|\, dx/m(E) \le \Big(\int_E |f(x) - c|^p\, dx/m(E)\Big)^{\frac{1}{p}} \to 0, \quad \text{when } m(E) \to \infty,$$

which also proves that $c = 0$ if $f \in L^q$ for some $q \in [1, \infty)$.

An inequality in the sense opposite to (4.4.23) is an obvious consequence of Theorem 4.2.5, for $f^\sharp \le 2(f - c)^*_{\text{HL}}$ for every c. The space $\text{BMO}(\mathbf{R}^n)$ has a similar characterization with $p=\infty$. However, the essential difference is that while the constant term is uniquely determined in $L^p(\mathbf{R}^n)+\mathbf{C}$ when $1<p<\infty$ there is no natural way to factor out $\mathbf{C}$ from BMO.

In what follows we define $f^\sharp(x)$ by (4.4.22) with arbitrary cubes I. The definition is not changed if we require I to be open, and (4.4.23) remains valid since $f^\sharp$ can only be increased by this change. It is then clear that $f^\sharp(x)$ is a lower semicontinuous function. If $\psi\in C_0(I)$ and $\int_I \psi\,dx=0$, $|\psi|\le 1$, then

$$\frac{1}{m(I)}\int f(x)\psi(x)\,dx \le \frac{1}{m(I)}\int_I |f(x)-f_I|\,dx.$$

If the left-hand side is $\le M$ for all such ψ, it follows from the Hahn-Banach theorem that for some measure $d\mu$ with support in $\overline{I}$ and total mass $\le M$ we have

$$\frac{1}{m(I)}\int f(x)\psi(x)\,dx = \int \psi(x)d\mu(x), \quad \text{if } \psi\in C_0(I),\ \int \psi\,dx=0,$$

which means that $d\mu(x)=(f(x)-c)/m(I)$ for some constant c. As already observed this implies $\int_I |f(x)-c|\,dx/m(I)\le M$, hence $|f_I-c|\le M$, so $\int_I |f(x)-f_I|\,dx/m(I)\le 2M$. Hence

$$\tfrac12 f^\sharp(x)\le \sup_{\psi,I}\frac{1}{m(I)}\int_I f(x)\psi(x)\,dx \le f^\sharp(x), \tag{4.4.24}$$

where the supremum is taken over all open cubes I with $x\in I$ and all $\psi\in C_0(I)$ with $\int\psi(y)\,dy=0$ and $|\psi|\le 1$.

Assume that $f\in L^1_{\text{loc}}$ is not a constant. Then $f^\sharp(x)>0$ for every x. If $0\le\varphi<\frac12 f^\sharp$ and $\varphi\in C_0(\mathbf{R}^n)$, it follows from (4.4.24) and the Borel-Lebesgue lemma that we can choose finitely many open cubes I_j, functions $\psi_j\in C_0(I_j)$ with $\int\psi_j(y)\,dy=0$ and $|\psi_j|\le 1$, and functions $\chi_j\in C_0(I_j)$ with $\chi_j\ge 0$ and $\sum\chi_j\le 1$ such that

$$\varphi(x)\le \sum \chi_j(x)\int_{I_j} f(y)\psi_j(y)\,dy/m(I_j), \quad x\in\mathbf{R}^n.$$

The right-hand side is bounded by $f^\sharp(x)$. Hence

$$\tfrac12\|f^\sharp\|_p\le \sup_\Psi \|\Psi f\|_p\le \|f^\sharp\|_p, \quad f\in L^1_{\text{loc}}, \quad \text{where} \tag{4.4.25}$$

$$\Psi f(x)=\int \Psi(x,y)f(y)\,dy, \quad \Psi(x,y)=\sum\chi_j(x)\psi_j(y)/m(I_j),\ x,y\in\mathbf{R}^n. \tag{4.4.26}$$

Here the sum is finite, I_j denotes open cubes, and $\chi_j,\psi_j\in C_0(I_j)$ are as described above. If $f\in L^q$ for some $q\in(1,\infty)$ then it follows from Corollary 4.4.6 that $\|f\|_p$ is equivalent to $\sup_\Psi\|\Psi f\|_p$. This gives the following interpolation theorem already announced above:

THEOREM 4.4.7. *Let $1 < p < \infty$ and let T be a linear operator from $L^p(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n)$ to $L^p(\mathbf{R}^n) \cap \mathrm{BMO}(\mathbf{R}^n)$ such that*

$$(4.4.27) \qquad \begin{aligned} \|Tf\|_p &\le C\|f\|_p, \quad f \in L^p(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n), \\ \|Tf\|_{\mathrm{BMO}} &\le C\|f\|_\infty, \quad f \in L^p(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n). \end{aligned}$$

Then it follows that

$$(4.4.28) \qquad \|Tf\|_q \le CC_q\|f\|_q, \quad f \in L^p(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n),\ p < q < \infty,$$

so the closure of T in $L^q(\mathbf{R}^n)$ is a bounded operator.

PROOF. With Ψ as in (4.4.25), (4.4.26) we have by (4.4.27)

$$\|\Psi Tf\|_p \le C\|f\|_p, \quad \|\Psi Tf\|_\infty \le C\|f\|_\infty, \quad f \in L^p(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n),$$

for $\|\Psi Tf\|_p \le \|(Tf)^\sharp\|_p \le C\|Tf\|_p$, and $\|\Psi Tf\|_\infty \le C\|Tf\|_{\mathrm{BMO}}$. Hence it follows from the Riesz-Thorin convexity theorem (Theorem 2.3.2) applied to ΨT that

$$\|\Psi Tf\|_q \le C\|f\|_q, \quad f \in L^p \cap L^\infty,\ p < q < \infty,$$

and we conclude using (4.4.26) and Corollary 4.4.6 that (4.4.28) is valid. The proof is complete.

COROLLARY 4.4.8. *Let $1 < p < \infty$ and let T be a linear operator from $L^p(\mathbf{R}^n) \cap \mathscr{H}^1(\mathbf{R}^n)$ to $L^p(\mathbf{R}^n) \cap L^1(\mathbf{R}^n)$ such that*

$$(4.4.29) \qquad \begin{aligned} \|Tf\|_p &\le C\|f\|_p, \quad f \in L^p(\mathbf{R}^n) \cap \mathscr{H}^1(\mathbf{R}^n), \\ \|Tf\|_1 &\le C\|f\|_{\mathscr{H}^1}, \quad f \in L^p(\mathbf{R}^n) \cap \mathscr{H}^1(\mathbf{R}^n). \end{aligned}$$

Then it follows that

$$(4.4.30) \qquad \|Tf\|_q \le CC_q\|f\|_q, \quad f \in L^p(\mathbf{R}^n) \cap \mathscr{H}^1(\mathbf{R}^n), \quad 1 < q < p,$$

so the closure of T in $L^q(\mathbf{R}^n)$ is a bounded operator (defined in $L^q(\mathbf{R}^n)$).

PROOF. T is densely defined in $L^p(\mathbf{R}^n)$ and has a bounded adjoint $T^* : L^{p'}(\mathbf{R}^n) \to L^{p'}(\mathbf{R}^n)$ where $1/p + 1/p' = 1$. If $g \in L^{p'}(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n)$ and $f \in L^p(\mathbf{R}^n) \cap \mathscr{H}^1(\mathbf{R}^n)$ then

$$|\langle T^*g, f\rangle| = |\langle g, Tf\rangle| \le \|g\|_\infty \|Tf\|_1 \le C\|g\|_\infty \|f\|_{\mathscr{H}^1}.$$

By Theorem 4.2.13 this proves that $\|T^*g\|_{\mathrm{BMO}} \le CC'\|g\|_\infty$, for $L^p(\mathbf{R}^n) \cap \mathscr{H}^1(\mathbf{R}^n)$ is a dense subset of $\mathscr{H}^1(\mathbf{R}^n)$. Now it follows from Theorem 4.4.7 that

$$\|T^*g\|_{q'} \le CC_q\|g\|_{q'}, \quad g \in L^{q'}(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n).$$

Hence

$$|\langle g, Tf\rangle| = |\langle T^*g, f\rangle| \le CC_q\|g\|_{q'}\|f\|_q,$$

and since $L^{p'}(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n)$ is dense in $L^{q'}(\mathbf{R}^n)$ the estimate (4.4.30) follows and the corollary is proved.

CHAPTER V

CONVERGENCE AND SUMMABILITY OF THE FOURIER EXPANSION

5.1. The role of multipliers. In Section 4.1 our motivation was the question on L^p convergence of the Fourier series of a function in $L^p(\mathbf{R})$ and the analogous question for Fourier transforms. However, the n-dimensional results of Section 4.2 were instead related to estimates of derivatives of potentials. We shall now return to the convergence problem. For the case of Fourier integrals the question is whether the inverse Fourier transform $\chi_\nu(D)f$ of $\chi_\nu \hat{f}$ converges to f in $L^p(\mathbf{R}^n)$ when $f \in L^p(\mathbf{R}^n)$ and χ_ν is the characteristic function of a subset of $\mathbf{R}^n$ increasing to $\mathbf{R}^n$. This formulation is not quite adequate unless $1 \le p \le 2$ so this will be assumed for a moment.

Definition 5.1.1. A measurable function χ in $\mathbf{R}^n$ is called a multiplier on the Fourier transform of $L^p(\mathbf{R}^n)$, where $1 \le p \le 2$, if $\chi \hat{f}$ is the Fourier transform of a function $\chi(D)f \in L^p(\mathbf{R}^n)$ for every $f \in L^p(\mathbf{R}^n)$. The set of such multipliers is denoted by $\mathscr{M}_p(\mathbf{R}^n)$.

If $\chi \in \mathscr{M}_p(\mathbf{R}^n)$ then the map

$$L^p(\mathbf{R}^n) \ni f \mapsto \chi(D)f \in L^p(\mathbf{R}^n) \tag{5.1.1}$$

is closed, for if $f_\nu \to 0$ in $L^p(\mathbf{R}^n)$ and $\chi(D)f_\nu \to g$ in $L^p(\mathbf{R}^n)$, then $\hat{f}_\nu \to 0$ in $L^{p'}(\mathbf{R}^n)$ and $\chi \hat{f}_\nu \to \hat{g}$ in $L^{p'}(\mathbf{R}^n)$ if $1/p + 1/p' = 1$. This implies convergence almost everywhere for a suitable subsequence, so it follows that $\hat{g} = 0$ almost everywhere. Thus (5.1.1) is closed and it follows from the closed graph theorem that for some constant C

$$\|\chi(D)f\|_p \le C\|f\|_p, \quad f \in L^p(\mathbf{R}^n). \tag{5.1.2}$$

In particular, (5.1.2) is valid for all $f \in \mathscr{S}(\mathbf{R}^n)$. Conversely, if (5.1.2) is true when $f \in \mathscr{S}(\mathbf{R}^n)$, then $\chi \in L^{p'}_{\text{loc}}(\mathbf{R}^n)$ and $\chi \in \mathscr{M}_p(\mathbf{R}^n)$, for if $\mathscr{S}(\mathbf{R}^n) \ni f_\nu \to f$ in $L^p(\mathbf{R}^n)$ then $\chi(D)f_\nu$ converges to a limit $g \in L^p(\mathbf{R}^n)$, so $\hat{f}_\nu \to \hat{f}$ and $\chi \hat{f}_\nu \to \hat{g}$ in $L^{p'}(\mathbf{R}^n)$ which implies that $\hat{g} = \chi \hat{f}$. We can therefore extend the definition of multipliers to all $p \in [1, \infty]$ as follows:

Definition 5.1.1′. A function $\chi \in L^1_{\text{loc}}(\mathbf{R}^n)$ is called a multiplier on the Fourier transform of $L^p(\mathbf{R}^n)$, where $1 \le p \le \infty$, if $\chi \hat{f}$ is the Fourier transform of a function $\chi(D)f \in L^p(\mathbf{R}^n)$ for every $f \in \mathscr{S}(\mathbf{R}^n)$ and (5.1.2) is valid when $f \in \mathscr{S}(\mathbf{R}^n)$. The set of such multipliers is denoted by $\mathscr{M}_p(\mathbf{R}^n)$, and $\|\chi\|_{\mathscr{M}_p}$ is defined to be the best constant C for which (5.1.2) is valid when $f \in \mathscr{S}(\mathbf{R}^n)$.

PROPOSITION 5.1.2. *$\mathscr{M}_p(\mathbf{R}^n)$ is a Banach algebra contained in $\mathscr{M}_2(\mathbf{R}^n) = L^\infty(\mathbf{R}^n)$. If $\chi \in \mathscr{M}_p(\mathbf{R}^n)$ then $\chi \in \mathscr{M}_q(\mathbf{R}^n)$ when $|1/q - 1/2| \le |1/p - 1/2|$, and*

$$\|\chi\|_{\mathscr{M}_q} \le \|\chi\|_{\mathscr{M}_p} \quad \text{if } \chi \in \mathscr{M}_p(\mathbf{R}^n) \text{ and } |\tfrac{1}{q} - \tfrac{1}{2}| \le |\tfrac{1}{p} - \tfrac{1}{2}|. \tag{5.1.3}$$

If $\chi \in \mathscr{M}_p(\mathbf{R}^n)$ and $\varphi \in \mathscr{S}(\mathbf{R}^n)$ then $\widehat{\varphi\chi} \in L^q(\mathbf{R}^n)$ when $1/q \le |1/p - 1/2| + 1/2$.

PROOF. By Hölder's inequality (5.1.2) implies

$$|\langle \chi(D)f, g\rangle| \le C\|f\|_p \|g\|_{p'}, \quad f, g \in \mathscr{S}(\mathbf{R}^n). \tag{5.1.2$'$}$$

Since $\langle \chi(D)f, g\rangle = \langle f, \chi(-D)g\rangle$ it follows that

$$\|\chi(-D)g\|_{p'} \le C\|g\|_{p'}, \quad g \in \mathscr{S}(\mathbf{R}^n).$$

There is a slight complication when $p = \infty$ for then we only obtain that $\chi(-D)g$ is a measure with total mass $\le C\|g\|_1$. However, if $\hat{g} \in C_0^\infty$ then $\chi(-D)g$ is a continuous function so it is in L^1 and (5.1.2)$'$ holds. Every $g \in \mathscr{S}$ is the limit in $\mathscr{S}$ of a sequence in C_0^∞, which proves (5.1.2)$'$ in general when $p' = 1$ also. If we replace g by $\check{g}$ where $\check{g}(x) = g(-x)$ it follows that

$$\|\chi(D)g\|_{p'} \le C\|g\|_{p'}, \quad g \in \mathscr{S}(\mathbf{R}^n),$$

so $\chi \in \mathscr{M}_{p'}$ and $\|\chi\|_{\mathscr{M}_{p'}} \le \|\chi\|_{\mathscr{M}_p}$. Replacing p by p' we conclude that there is equality.

Thus the map

$$\mathscr{S} \ni f \mapsto \chi(D)f$$

is continuous in the L^p norm and in the $L^{p'}$ norm. If $1 < p < \infty$ the closure is therefore a continuous map in $L^p(\mathbf{R}^n) \cap L^{p'}(\mathbf{R}^n)$ satisfying the hypotheses of the Riesz-Thorin convexity theorem (Theorem 2.3.2), so it follows that

$$\|\chi(D)f\|_q \le C\|f\|_q, \quad \text{if } |\tfrac{1}{q} - \tfrac{1}{2}| \le |\tfrac{1}{p} - \tfrac{1}{2}|, \quad f \in \mathscr{S}(\mathbf{R}^n).$$

If $p = 1$ or $p = \infty$ there is again a slight complication since the closure of $\mathscr{S}(\mathbf{R}^n)$ in $L^1(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n)$ consists of continuous integrable functions converging to 0 at ∞. However, the proof of Theorem 2.3.2 shows that this is sufficient for the conclusion.

When $p = 2$ then the estimate (5.1.2) is equivalent to

$$\|\chi F\|_2 \le C\|F\|_2, \quad F \in L^2(\mathbf{R}^n),$$

and this is true if and only if $|\chi| \le C$ almost everywhere. Thus $\mathscr{M}_2(\mathbf{R}^n) = L^\infty(\mathbf{R}^n)$, and $\mathscr{M}_p(\mathbf{R}^n) \subset L^\infty(\mathbf{R}^n)$ for $1 \le p \le \infty$.

If $1 \le p \le 2$ and $\chi \in \mathscr{M}_p(\mathbf{R}^n)$ then (5.1.2) is well defined and valid for all $f \in L^p(\mathbf{R}^n)$, so it follows at once that $\chi_1\chi_2 \in \mathscr{M}_p$ and that $\|\chi_1\chi_2\|_{\mathscr{M}_p} \le \|\chi_1\|_{\mathscr{M}_p}\|\chi_2\|_{\mathscr{M}_p}$ if $\chi_1, \chi_2 \in \mathscr{M}_p$. If χ_ν, $\nu = 1, 2, \dots$ is a Cauchy sequence in $\mathscr{M}_p(\mathbf{R}^n)$ then $\chi_\nu \hat{f}$ converges in $L^{p'}(\mathbf{R}^n)$ for every $\hat{f} \in \mathscr{S}(\mathbf{R}^n)$, so χ_ν converges to a limit $\chi \in L^{p'}_{\text{loc}}(\mathbf{R}^n)$ which is in $\mathscr{M}_p$ with norm $\le \lim_{\nu\to\infty} \|\chi_\nu\|_{\mathscr{M}_p}$. Hence $\|\chi - \chi_\mu\|_{\mathscr{M}_p} \le \lim_{\nu\to\infty} \|\chi_\nu - \chi_\mu\|_{\mathscr{M}_p}$, and when $\mu \to \infty$ we conclude that $\chi_\mu \to \chi$ in $\mathscr{M}_p$, so $\mathscr{M}_p$ is complete.

When proving the last statement we may assume that $p \le 2$. Since $\chi \in L^\infty(\mathbf{R}^n)$ we have $\varphi\chi \in L^1(\mathbf{R}^n)$ so $\widehat{\varphi\chi} \in L^\infty(\mathbf{R}^n)$, and since $\varphi = \hat{\psi}$ where $\psi \in \mathscr{S}(\mathbf{R}^n) \subset L^p(\mathbf{R}^n)$ we have $\chi\varphi = \widehat{\Psi}$ where $\Psi \in L^p(\mathbf{R}^n)$, which proves that $\widehat{\chi\varphi} \in L^p(\mathbf{R}^n)$. Hence $\widehat{\chi\varphi} \in L^q(\mathbf{R}^n)$ when $p \le q \le \infty$ as claimed.

Multipliers are invariant under linear operations:

Proposition 5.1.3. *If $T : \mathbf{R}^n \to \mathbf{R}^m$ is an affine surjective map and $\chi \in \mathscr{M}_p(\mathbf{R}^m)$, then $\chi \circ T \in \mathscr{M}_p(\mathbf{R}^n)$ and $\|\chi \circ T\|_{\mathscr{M}_p(\mathbf{R}^n)} = \|\chi\|_{\mathscr{M}_p(\mathbf{R}^m)}$.*

Proof. First assume that $n = m$. If (5.1.2) is valid then replacing f by $fe^{i\langle \cdot,\theta\rangle}$ where $\theta \in \mathbf{R}^n$ gives $\|\chi(D+\theta)f\|_p \le C\|f\|_p$, for $\chi(D)(fe^{i\langle\cdot,\theta\rangle}) = e^{i\langle\cdot,\theta\rangle}\chi(D+\theta)f$. This proves the statement when T is a translation. If A is a linear bijection in $\mathbf{R}^n$ and $f_A = f \circ A$ then $\widehat{f_A}(\xi) = |\det A|^{-1}\hat{f}({}^tA^{-1}\xi)$, so

$$\chi(\xi)\widehat{f_A}(\xi) = |\det A|^{-1}\chi({}^tA\eta)\hat{f}(\eta), \quad \eta = {}^tA^{-1}\xi.$$

Thus $\chi(D)f_A = (\chi({}^tAD)f)_A$, which proves the statement when $T = {}^tA^{-1}$.

What remains is to prove the statement when $m < n$ and

$$T(\xi_1, \dots, \xi_n) = (\xi_1, \dots, \xi_m).$$

Then $\chi_T(\xi) = (\chi\circ T)(\xi) = \chi(\xi_1, \dots, \xi_m)$, so $\chi_T(D)f(x) = \chi(D')f(x)$ where $\chi(D')$ operates on f as a function of $x' = (x_1, \dots, x_m)$ when $x'' = (x_{m+1}, \dots, x_n)$ is fixed. If (5.1.2) is valid for $\chi(D')$ in $\mathbf{R}^m$ and $1 \le p \le 2$, it follows that

$$\int |\chi_T(D)f(x', x'')|^p\, dx' \le C^p \int |f(x', x'')|^p\, dx'$$

for fixed x'', and integration with respect to x'' gives $\|\chi_T(D)f\|_p \le C\|f\|_p$. Conversely, if this is true we can choose $f(x) = g(x')h(x'')$ with a fixed $h \in \mathscr{S}(\mathbf{R}^{n-m})\setminus\{0\}$ and conclude that (5.1.2) is valid for $\chi(D)$. This completes the proof.

The hypothesis that T is surjective made in Proposition 5.1.3 is essential, for $\chi \circ T$ may not even be a measurable function otherwise since it only depends on the values of χ in a null set. However, under conditions which make $\chi \circ T$ meaningful there is a valid version of Proposition 5.1.3:

Proposition 5.1.4. *If $T : \mathbf{R}^n \to \mathbf{R}^m$ is an affine map and $\chi \in \mathscr{M}_p(\mathbf{R}^m)$ then $\chi \circ T \in \mathscr{M}_p(\mathbf{R}^n)$ and $\|\chi \circ T\|_{\mathscr{M}_p(\mathbf{R}^n)} \le \|\chi\|_{\mathscr{M}_p(\mathbf{R}^m)}$ if the points in $T\mathbf{R}^n$ which are not Lebesgue points for χ have Lebesgue measure 0 in $T\mathbf{R}^n$.*

Proof. We may assume that $1 \le p \le 2$. By Proposition 5.1.3 we may also assume that $n < m$ and that

$$T\xi = (\xi_1, \dots, \xi_n, 0, \dots, 0), \quad \xi \in \mathbf{R}^n.$$

We shall denote the variable in $\mathbf{R}^m$ by (x, y) or (ξ, η) where $x, \xi \in \mathbf{R}^n$ and $y, \eta \in \mathbf{R}^{m-n}$. Since $(\xi, \eta) \mapsto \chi(\xi, \varepsilon\eta)$ is a multiplier with the same norm M for every $\varepsilon > 0$, an application of (5.1.2) to $(x, y) \mapsto f(x)g(y)$ gives for $\varepsilon > 0$

$$\Big(\iint |h_\varepsilon(x,y)|^p\, dx\, dy\Big)^{\frac{1}{p}} \le M\|f\|_p\|g\|_p, \quad f \in \mathscr{S}(\mathbf{R}^n),\ g \in \mathscr{S}(\mathbf{R}^{m-n}),$$

$$h_\varepsilon(x,y) = (2\pi)^{-m} \iint e^{i\langle x,\xi\rangle + i\langle y,\eta\rangle}\hat{f}(\xi)\hat{g}(\eta)\chi(\xi, \varepsilon\eta)\, d\xi\, d\eta.$$

Since $|\chi| \le M$ almost everywhere it follows that $|\chi(\xi, 0)| \le M$ when $(\xi, 0)$ is a Lebesgue point for χ, hence for almost every $\xi \in \mathbf{R}^n$ by hypothesis, and since

$$\chi(\xi, 0) = \lim_{\varepsilon \to 0} \iint_{|\xi'|<\varepsilon, |\eta'|<\varepsilon} \chi(\xi + \xi', \eta')\, d\xi' d\eta' \Big/ \iint_{|\xi'|<\varepsilon, |\eta'|<\varepsilon} d\xi'\, d\eta'$$

for almost all $\xi \in \mathbf{R}^n$ and the integral here is a continuous function of ξ, it follows that $\xi \mapsto \chi(\xi, 0)$ is a measurable function. Let $\hat{f} \in C_0^\infty(\mathbf{R}^n)$, $\hat{g} \in C_0^\infty(\mathbf{R}^{m-n})$. We have

$$\begin{aligned} |h_\varepsilon(x, y) - h_0(x, y)| &\le (2\pi)^{-m} \iint |\hat{f}(\xi)\hat{g}(\eta)||\chi(\xi, \varepsilon\eta) - \chi(\xi, 0)|\, d\xi\, d\eta \\ &= C \iiint_{|\xi'|<1} |\hat{f}(\xi + \varepsilon\xi')\hat{g}(\eta)||\chi(\xi + \varepsilon\xi', \varepsilon\eta) - \chi(\xi + \varepsilon\xi', 0)|\, d\xi\, d\eta\, d\xi'. \end{aligned}$$

The integral with respect to (ξ', η) is bounded by a constant and tends to 0 when $\varepsilon \to 0$ if $(\xi, 0)$ is a Lebesgue point for χ and ξ is a Lebesgue point for $\chi(\cdot, 0)$, hence $h_\varepsilon(x, y) \to h_0(x, y)$ as $\varepsilon \to 0$. By Fatou's lemma we obtain

$$\Big(\iint |h_0(x, y)|^p\, dx\, dy\Big)^{\frac{1}{p}} \le M\|f\|_p\|g\|_p,$$

and since $h_0(x, y) = (\chi(D, 0)f)g$ we conclude that $\|\chi(D, 0)f\|_p \le M\|f\|_p$, which proves the proposition.

By Theorem 4.1.1 we know that the characteristic function of $\{t \in \mathbf{R}; t > 0\}$ is in $\mathscr{M}_p(\mathbf{R})$ for $1 < p < \infty$. Hence it follows from Proposition 5.1.3 that the characteristic function of any half space in $\mathbf{R}^n$ is in $\mathscr{M}_p(\mathbf{R}^n)$ for $1 < p < \infty$, and since $\mathscr{M}_p(\mathbf{R}^n)$ is a Banach algebra we conclude that the characteristic function of any polyhedron in $\mathbf{R}^n$ is in $\mathscr{M}_p(\mathbf{R}^n)$.

THEOREM 5.1.5. *If $\chi \in \mathscr{M}_p(\mathbf{R}^n)$ then*

$$\|\chi(D/t)f - f\|_q \to 0 \quad \text{as } t \to \infty, \tag{5.1.4}$$

if $f \in L^q(\mathbf{R}^n)$ and $|\frac{1}{q} - \frac{1}{2}| < |\frac{1}{p} - \frac{1}{2}|$ or $2 \le q = p < \infty$, provided that

$$\int_{|\xi|<\varepsilon} |\chi(\xi) - 1|\, d\xi/\varepsilon^n \to 0, \quad \varepsilon \to 0, \tag{5.1.5}$$

that is, 0 is a Lebesgue point for χ and $\chi(0) = 1$.

PROOF. Since $\|\chi(\cdot/t)\|_{\mathscr{M}_p}$ is independent of t it suffices to prove that (5.1.4) follows from (5.1.5) when $\hat{f} \in C_0^\infty$, for such functions are dense in $\mathscr{S}(\mathbf{R}^n)$, hence in L^q for $1 \le q < \infty$. Now

$$\chi(\xi/t)\hat{f}(\xi) - \hat{f}(\xi) = (\chi(\xi/t) - 1)\hat{f}(\xi)$$

is uniformly bounded, and if $|\xi| < R$ when $\hat{f}(\xi) \neq 0$ we have by (5.1.5)

$$\int |(\chi(\xi/t) - 1)\hat{f}(\xi)|\, d\xi \leq C \int_{|\xi|<R} |\chi(\xi/t) - 1|\, d\xi = Ct^n \int_{|\xi|<R/t} |\chi(\xi) - 1|\, d\xi \to 0,\ t \to \infty.$$

Hence the L^2 norm converges to 0 so $\|\chi(D/t)f - f\|_2 \to 0$ as $t \to \infty$. Since the L^q norm is bounded when $|\frac{1}{q} - \frac{1}{2}| \leq |\frac{1}{p} - \frac{1}{2}|$ it follows from Hölder's inequality that it converges to 0 when there is strict inequality. By the Hausdorff-Young inequality this is also true when there is equality and $q \geq 2$, which proves the statement.

REMARK. If $\chi \in H^{\mathrm{loc}}_{(s)}$ in a neighborhood of the origin for some $s > n|\frac{1}{p} - \frac{1}{2}|$, then it follows from Theorem 2.3.8 that (5.1.4) is also valid when $|\frac{1}{q} - \frac{1}{2}| = |\frac{1}{p} - \frac{1}{2}|$ and $q < 2$.

In particular it follows from Theorem 5.1.5 that $\chi(D)f \to f$ in $L^p(\mathbf{R}^n)$ if $1 < p < \infty$ and χ is the characteristic function of a polyhedron with the origin in its interior. We shall now show that the situation is quite different for characteristic functions of sets with smooth boundaries.

THEOREM 5.1.6. *Let χ be the characteristic function of an open set $\Omega \subset \mathbf{R}^n$, and assume that $\chi \in \mathscr{M}_p(\mathbf{R}^n)$ for some $p \neq 2$. If $\partial\Omega$ is a C^2 hypersurface in a neighborhood of a point $\xi^0 \in \partial\Omega$ and $\Omega \cup \partial\Omega$ is not a neighborhood of ξ^0, then $\partial\Omega$ is a subset of a hyperplane in a neighborhood of ξ^0.*

PROOF. By Proposition 5.1.3 we may assume that $\xi^0 = 0$ and that $\Omega \cap U$ is defined by $\xi_n > \psi(\xi')$ where $\xi' = (\xi_1, \dots, \xi_{n-1})$, $\psi \in C^2$ and $\psi(0) = 0$, $\psi'(0) = 0$. We must prove that $\psi'' = 0$ in a neighborhood of the origin. Let $\varphi \in C_0^\infty(U)$ be equal to 1 in a neighborhood U_0 of the origin. Then $\varphi\chi \in \mathscr{M}_p(\mathbf{R}^n)$ for some $p \neq 2$, and all points where $\xi_n \neq \psi(\xi')$ are Lebesgue points. Hence it follows from Proposition 5.1.4 that $(\xi_1, \xi_2) \mapsto (\varphi\chi)(\xi_1 a + b, \xi_2)$ is in $\mathscr{M}_p(\mathbf{R}^2)$ if $0 \neq a \in \mathbf{R}^{n-1}$ and $b \in \mathbf{R}^{n-1}$. If the theorem has been proved in the two-dimensional case it follows that $\psi''(b)$ vanishes in the direction a if $(b, \psi(b)) \in U_0$, so $\psi'' = 0$ in a neighborhood of the origin.

Thus it suffices to prove Theorem 5.1.6 when $n = 2$, that is, to prove that *if $\psi \in C^2(\mathbf{R})$ and $\varphi\chi_\psi \in \mathscr{M}_p(\mathbf{R}^2)$ for some $p \neq 2$, where $\varphi \in C_0^\infty(\mathbf{R}^2)$ and χ_ψ is the characteristic function of $\{x \in \mathbf{R}^2; \xi_2 > \psi(\xi_1\}$, then $\psi''(\xi_1) = 0$ if $\varphi(\xi_1, \psi(\xi_1)) \neq 0$.* This will require some preparations. The first is a lemma explaining how $(\varphi\chi_\psi)(D)$ operates on functions with support oblong in the direction of a normal of the curve $\{(\xi_1, \psi(\xi_1)); \xi_1 \in \mathbf{R}\}$ and with the corresponding frequency. The second is a famous construction in measure theory, the Besicovitch solution of the so-called Kakeya needle problem. The proof of the two-dimensional case of Theorem 5.1.6 will be given after that. (Theorem 5.1.6 was proved by Fefferman [1] for the unit ball in $\mathbf{R}^n$. The general case is not essentially different, but the proof is more transparent then.)

LEMMA 5.1.7. *Let $u \in C_0^\infty(\mathbf{R}^2)$ and set for $t \in \mathbf{R}$ and $\varepsilon > 0$*

$$u_{t,\varepsilon}(x_1, x_2) = u(\varepsilon(x_1 + \psi'(t)x_2), \varepsilon^2 x_2)e^{i(tx_1 + \psi(t)x_2)}. \tag{5.1.6}$$

Then $(\varphi\chi_\psi)(D)u_{t,\varepsilon} = v_{t,\varepsilon}$ where

$$v_{t,\varepsilon}(x_1, x_2) = V_{t,\varepsilon}(\varepsilon(x_1 + \psi'(t)x_2), \varepsilon^2 x_2)e^{i(tx_1 + \psi(t)x_2)}, \tag{5.1.7}$$

$V_{t,\varepsilon}(x) \to V_t(x)$ locally uniformly in (t,x) when $\varepsilon \to 0$,

$$\widehat{V}_t(\xi) = \varphi(t,\psi(t))\chi_t(\xi)\hat{u}(\xi). \tag{5.1.8}$$

Here χ_t is the characteristic function of $\{\xi \in \mathbf{R}^2; \xi_2 > \psi''(t)\xi_1^2/2\}$. If $a \le x_2 \le b$ when $(x_1,x_2) \in \operatorname{supp} u$, then $V_t(x)$ is real analytic for $x_2 < a$ and for $x_2 > b$, and if $\int_{\mathbf{R}^2} u(x)\,dx \ne 0$ then $V_t \not\equiv 0$ in these half planes when $\varphi(t,\psi(t)) \ne 0$.

Proof. Set $A_\varepsilon(x_1,x_2) = (\varepsilon(x_1+\psi'(t)x_2), \varepsilon^2 x_2)$ and $\theta = (t,\psi(t))$. Then

$$u_{t,\varepsilon} = (u\circ A_\varepsilon)e^{i\langle\cdot,\theta\rangle}, \quad v_{t,\varepsilon} = (V_{t,\varepsilon}\circ A_\varepsilon)e^{i\langle\cdot,\theta\rangle},$$
$$(\varphi\chi_\psi)(D)u_{t,\varepsilon} = e^{i\langle\cdot,\theta\rangle}(\varphi\chi_\psi)(D+\theta)(u\circ A_\varepsilon) = e^{i\langle\cdot,\theta\rangle}V_{t,\varepsilon}\circ A_\varepsilon.$$

The Fourier transform of $u\circ A_\varepsilon$ is $\xi \mapsto |\det A_\varepsilon|^{-1}\hat{u}({}^tA_\varepsilon^{-1}\xi)$, so the Fourier transform of $(\varphi\chi_\psi)(D+\theta)(u\circ A_\varepsilon)$ is

$$\xi \mapsto |\det A_\varepsilon|^{-1}(\varphi\chi_\psi)(\xi+\theta)\hat{u}({}^tA_\varepsilon^{-1}\xi) = |\det A_\varepsilon|^{-1}(\varphi\chi_\psi)({}^tA_\varepsilon\,{}^tA_\varepsilon^{-1}\xi+\theta)\hat{u}({}^tA_\varepsilon^{-1}\xi),$$

which means that

$$\widehat{V_{t,\varepsilon}}(\eta) = (\varphi\chi_\psi)({}^tA_\varepsilon\eta+\theta)\hat{u}(\eta).$$

The right-hand side is uniformly bounded by $|\hat{u}|\sup|\varphi|$, and $\varphi({}^tA_\varepsilon\eta+\theta) \to \varphi(\theta)$ as $\varepsilon \to 0$. Since ${}^tA_\varepsilon\eta = (\varepsilon\eta_1, \varepsilon\psi'(t)\eta_1+\varepsilon^2\eta_2)$ the factor $\chi_\psi({}^tA_\varepsilon\eta+\theta)$ is the characteristic function of

$$\{\eta \in \mathbf{R}^2; \varepsilon\psi'(t)\eta_1+\varepsilon^2\eta_2+\psi(t) > \psi(\varepsilon\eta_1+t)\}$$

which converges to $\{\eta \in \mathbf{R}^2; \eta_2 > \psi''(t)\eta_1^2/2\}$ when $\varepsilon \to 0$. This proves that $V_{t,\varepsilon}$ converges to V_t as defined by (5.1.8).

The inverse Fourier transform of χ_t is

$$x \mapsto (2\pi)^{-2}\iint_{\xi_2>\psi''(t)\xi_1^2/2} e^{i(x_1\xi_1+x_2\xi_2)}\,d\xi_1\,d\xi_2$$
$$= (2\pi)^{-2}\int e^{i(x_1\xi_1+\psi''(t)x_2\xi_1^2/2)}\,d\xi_1\int_{\xi_2>0} e^{ix_2\xi_2}\,d\xi_2$$

with the integrals taken in the sense of distribution theory. This follows at once if we first introduce a convergence factor $e^{-\delta|\xi|^2}$ and then let $\delta \to 0$. The integral with respect to ξ_2 is $i/(x_2+i0)$, and the integral with respect to ξ_1 is $2\pi\delta_0$ if $\psi''(t)=0$ and otherwise it is the Gaussian $\sqrt{2\pi i/\psi''(t)x_2}\exp(-ix_1^2/2\psi''(t)x_2)$. Thus it is analytic when $x_2 \ne 0$ which proves the stated analyticity. If $\psi''(t)=0$ then the convolution with u is asymptotic to $i/2\pi x_2\int u(x_1,y_2)\,dy_2$ when $x_2 \to \infty$, and when $\psi''(t) \ne 0$ it is asymptotic to $c_\pm|x_2|^{-3/2}\int u(y)\,dy$ when $x_2 \to \pm\infty$, where $c_\pm \ne 0$. This completes the proof.

We shall now discuss the construction of a Besicovitch set in $\mathbf{R}^2$. For a triangle ABC with base AB, of length l, and height h from the vertex C we form two new triangles ADA' and BDB' where D is the midpoint on AB and C divides AA' (resp. BB') in the ratio h to d for some $d > 0$. Thus the bases AD and BD of the triangles have length $\frac{1}{2}l$, and the heights from A' and B' are $h + d$. The union of the triangles ADA' and BDB' consists of ABC and two additional triangles $A'CA''$ and $B'CB''$. (Draw a figure!) The triangle $A'CA''$ is homothetic with respect to A' to $A'A'''D$ where A''' is the midpoint of AC, and the ratio is d to $d + \frac{1}{2}h$. Thus

$$m(A'CA'') = m(A'A'''D)d^2/(d + \tfrac{1}{2}h)^2,$$
$$m(A'A'''D) = m(A'AD) - m(AA'''D) = l(h + d)/4 - lh/8 = l(h + 2d)/8.$$

Hence $m(A'CA'') = ld^2/(4d + 2h)$, and $m(ADA' \cup BDB') = lh/2 + ld^2/(2d + h)$. Now we fix $d = 1$, and starting with $h = 1$ we repeat the construction k times ending up with 2^k triangles R_j, $j = 1, \dots, 2^k$, with bases of length $2^{-k}l$ and height $k + 1$,

$$m\Big(\bigcup_{j=1}^{2^k} R_j\Big) \le \tfrac{1}{2}l + l\sum_1^k \frac{1}{j+2} < l\log(k + 1). \tag{5.1.9}$$

If $\widetilde{CAB}$ denotes the union of the half lines with one end point at C intersecting AB, with CAB removed, then $\widetilde{A'AD}$ and $\widetilde{B'BD}$ are disjoint subsets of $\widetilde{CAB}$. If we define $\widetilde{R}_j$ similarly it follows that all $\widetilde{R}_j$ are disjoint. This means that if $\widehat{R}_j$ is the part of $\widetilde{R}_j$ at distance $\le k + 1$ from the line through A and B, then

$$m\Big(\bigcup_1^{2^k} \widehat{R}_j\Big) = \sum_1^{2^k} m(\widehat{R}_j) = 3\sum_1^{2^k} m(R_j) = 3l(k + 1)/2$$

which according to (5.1.9) is larger than the measure of $\cup R_j$ by a factor $(k+1)/\log(k+1)$. This is the essence of the "sprouting" construction.

Before using the construction to finish the proof of Theorem 5.1.6 we shall digress to discuss its consequences for the definition of maximal functions. Suppose that for $f \in L^1_{\text{loc}}(\mathbf{R}^2)$ we define

$$f^*(x) = \sup_{x\in K} \int_K |f(y)|\, dy/m(K)$$

where K is an arbitrary bounded open convex set $\subset \mathbf{R}^2$. (It would make no essential difference if we required K to be a rectangle or to be the interior of an ellipse.) If there is an estimate $\|f^*\|_q \le C\|f\|_p$ then it follows by a scale change that $q = p$. Now take for f the characteristic function of $\cup R_j$. In the sets $\widehat{R}_j$ we find by taking $K = R_j \cup \widehat{R}_j$ that $f^* \ge 1/4$. Hence

$$\|f\|_p \le (l\log(k + 1))^{\frac{1}{p}}, \quad 4\|f^*\|_p \ge (3l(k + 1)/2)^{\frac{1}{p}},$$

and since k can be chosen arbitrarily large it follows that no L^p estimate is possible.

End of proof of Theorem 5.1.6. It remains to prove the two dimensional statement in the form given in slanted text at the end of the first part of the proof. Assuming that $\psi''(0) \neq 0$ we shall derive a contradiction. To do so we start from the Besicovitch construction above with a triangle ABC whose base AB is on the x-axis and top C is on the line $x_2 = 1$, chosen so that the range of $(-\psi'(t), 1)$ for t so small that $\varphi(t, \psi(t)) = 1$ covers all vectors $\overrightarrow{DC}$ with D between A and B. In the construction we choose a large number k of iterations. Let

$$\varepsilon = 2^{-k} l/(k+1) \tag{5.1.10}$$

be the ratio between the base and the height of the triangles R_j, and let $(-\psi'(t_j), 1)$ be the direction of the median in R_j, the line from the midpoint m_j of the base to the top. Choose

$$u \in C_0^\infty(\{x \in \mathbf{R}^2; -1 < x_2 < 0, |x_1| < \tfrac{1}{2}(1 - x_2)\}), \tag{5.1.11}$$

and consider $u_{t_j,\varepsilon}$ as defined by (5.1.6). The support lies in the triangle $\widehat{R}$ if R is the triangle with vertices $\varepsilon^{-1}(\pm\frac{1}{2}, 0)$ and $\varepsilon^{-2}(-\psi'(t_j), 1)$, which is similar to R_j with the ratio $2^{-k} l/\varepsilon^{-1} = \varepsilon^2(k+1)$. Thus the support of $u_{t_j,\varepsilon}(\cdot - \kappa m_j)$ is in $\kappa \widehat{R}_j$ where $\kappa = 1/(\varepsilon^2(k+1))$. Let

$$f_\theta = \sum_{j=1}^{2^k} e^{i\theta_j} u_{t_j,\varepsilon}(\cdot - \kappa m_j), \quad \theta_j \in \mathbf{R},$$

and write $T = (\varphi\chi_\psi)(D)$, $E = \cup(\kappa R_j)$. Then

$$\int_E |T f_\theta|^2 \, dx \geq \sum_j \int_E |T u_{t_j,\varepsilon}(\cdot - \kappa m_j)|^2 \, dx$$

for a suitable choice of θ, for if we expand the left-hand side and take the mean value over θ, the cross products drop out so the mean value is equal to the right-hand side. Since $E \supset \kappa R_j$ it follows from Lemma 5.1.7 when ε is small enough that for some $c > 0$

$$c 2^k \varepsilon^{-3} \leq \int_E |T f_\theta|^2 \, dx \leq \Big(\int_E |T f_\theta|^p \, dx \Big)^{\frac{2}{p}} (m(E))^{1 - \frac{2}{p}}$$

where we have also used Hölder's inequality, assuming that $2 < p < \infty$. If T is bounded in L^p norm then

$$\|T f_\theta\|_p^p \leq C \|f_\theta\|_p^p = C \sum_j \|u_{t_j,\varepsilon}(\cdot - \kappa m_j)\|_p^p \leq C' 2^k \varepsilon^{-3}$$

for the sets $\widehat{R}_j$ are disjoint. Hence

$$2^k \varepsilon^{-3} \leq C'' m(E) \leq C'' \kappa^2 l \log(k+1)$$

and since $2^k \varepsilon^{-3} \kappa^{-2} = 2^k \varepsilon (k+1)^2 = l(k+1)$ we get a contradiction when k is so large that $k + 1 > C'' \log(k+1)$. The proof is complete.

Note that the point in the proof was that the L^p norm of f_θ is bounded by the norm of a term times the number of terms raised to the power $1/p$, since the supports are disjoint, whereas the L^2 norm of Tf_θ in E involved the number of terms raised to the power $1/2$ only. The discrepancy between L^2 and L^p norm could be overcome by Hölder's inequality since the terms in Tf_θ were all large in the *same set* E of not too large measure. It is this focussing effect which is the main point in the proof.

Theorem 5.1.6 proves in particular that for $p \neq 2$ and $n > 1$ the characteristic function χ of the unit ball in $\mathbf{R}^n$ is not in $\mathscr{M}_p$, so spherical summation of the Fourier expansion is not possible. The difficulty stems from the discontinuity of χ at the unit sphere and disappears if χ is replaced by a cutoff function in $C_0^\infty(\mathbf{R}^n)$. We shall now study how smooth it has to be, in particular examine when

$$R_\alpha(\xi) = \begin{cases} (1-|\xi|^2)^\alpha, & \text{if } |\xi| < 1 \\ 0, & \text{if } |\xi| \geq 1 \end{cases}, \quad \xi \in \mathbf{R}^n, \tag{5.1.12}$$

is in $\mathscr{M}_p(\mathbf{R}^n)$. When $R_\alpha \in \mathscr{M}_p(\mathbf{R}^n)$ it follows from Theorem 5.1.5 that the Riesz-Bochner means $R_\alpha(D/t)f$ converge to f in $L^p(\mathbf{R}^n)$ when $t \to \infty$, for all $f \in L^p(\mathbf{R}^n)$. The following theorem gives a necessary condition which is actually much more elementary than that in Theorem 5.1.6.

Theorem 5.1.8. *Let $\varrho \in C^\infty(\mathbf{R}^n)$ be real valued, $n > 1$, and define $\varrho_+ = \max(\varrho, 0)$. Assume that ϱ has a zero $\xi^0 \in \mathbf{R}^n$ such that $\varrho'(\xi^0) \neq 0$ and $\varrho''(\xi^0)t \neq 0$ if $0 \neq t \in \mathbf{R}^n$ and $\varrho'(\xi^0)t = 0$, that is, the zeros of ϱ form a hypersurface with total curvature $\neq 0$ at ξ^0. Then*

$$\left|\tfrac{1}{p} - \tfrac{1}{2}\right| < \tfrac{1}{2n} + \tfrac{\alpha}{n}, \quad \textit{if } \chi\varrho_+^\alpha \in \mathscr{M}_p, \tag{5.1.13}$$

for some $\chi \in C_0^\infty$ with $\chi(\xi^0) \neq 0$.

Proof. In view of Proposition 5.1.2 it suffices to prove the statement when $p < 2$, that is, prove that $1/p - 1/2 < 1/2n + \alpha/n$. By Proposition 5.1.3 it is no restriction to assume that $\xi^0 = 0$ and that $\varrho'(\xi^0) = (0, \dots, 0, 1)$. Write $\xi = (\xi', \xi_n)$ where $\xi' = (\xi_1, \dots, \xi_{n-1})$. Then the equation $\varrho(\xi', \xi_n) = 0$ has a unique solution $\xi_n = \psi(\xi')$ in a neighborhood of 0, and $\psi(0) = 0$, $\psi'(0) = 0$ and $\det \psi''(\xi') \neq 0$. The quotient $\varrho(\xi)/(\xi_n - \psi(\xi'))$ is in C^∞ in a neighborhood of 0 and equal to 1 at 0. If $a \in C_0^\infty(\mathbf{R}^{n-1})$ and $b \in C_0^\infty(\mathbf{R})$ have support sufficiently close to the origin, it follows that

$$m(\xi) = a(\xi')b(\xi_n - \psi(\xi'))(\xi_n - \psi(\xi'))_+^\alpha = \varphi(\xi)\chi(\xi)\varrho_+(\xi)^\alpha,$$

where $\varphi(\xi) = a(\xi')b(\xi_n - \psi(\xi'))\big((\xi_n - \psi(\xi'))/\varrho(\xi)\big)^\alpha/\chi(\xi)$ is in C_0^∞. If $\chi\varrho_+^\alpha \in \mathscr{M}_p(\mathbf{R}^n)$ it follows that $m \in \mathscr{M}_p(\mathbf{R}^n)$ and by Proposition 5.1.2 that $\widehat{m} \in L^p(\mathbf{R}^n)$. A change of variables gives

(5.1.14)

$$\widehat{m}(x) = A(x)B(x_n), \quad A(x) = \int e^{-i(\langle x', \xi'\rangle + x_n\psi(\xi'))}a(\xi')\,d\xi', \;\; B(x_n) = \int_0^\infty e^{-ix_n t}b(t)t^\alpha\,dt.$$

The method of stationary phase (see Theorem 2.4.6) gives that if $a(0) \neq 0$ then $|A(x)| > C|x_n|^{-(n-1)/2}$ at infinity in a conic neighborhood of the x_n axis, for some $C > 0$, for the equation $x' + x_n\psi'(\xi') = 0$ for the critical point has a unique solution ξ' near 0 when $|x'/x_n|$ is small, since $\psi'(0) = 0$ and $\det \psi''(0) \neq 0$. The Fourier transform of t_+^α is

$$\tau \mapsto \Gamma(\alpha + 1)(i\tau + 0)^{-\alpha-1} = \Gamma(\alpha + 1)|\tau|^{-\alpha-1}e^{-\frac{\pi i}{2}(\alpha+1)\operatorname{sgn}\tau},$$

for it is the limit as $\varepsilon \to +0$ of the Fourier transform

$$\int_0^\infty t^\alpha e^{-\varepsilon t - i\tau t}\, dt = \int_0^\infty t^\alpha e^{-t(\varepsilon+i\tau)}\, dt = (\varepsilon + i\tau)^{-\alpha-1} \int_0^\infty t^\alpha e^{-t}\, dt.$$

The last integral is $\Gamma(\alpha + 1)$, and the second equality follows from Cauchy's integral formula. The Fourier transform B is the convolution with $\hat{b}/2\pi$, which is in $\mathscr{S}(\mathbf{R})$ and has the integral $b(0)$. Hence

$$|B(x_n)||x_n|^{\alpha+1} \to \Gamma(\alpha + 1)|b(0)|, \quad \text{as } x_n \to \infty.$$

If $b(0) \neq 0$ it follows that for some $C' > 0$

$$|\widehat{m}(x)| \geq C'|x_n|^{-\frac{1}{2}(n-1)-\alpha-1}$$

at infinity in a conic neighborhood of the x_n axis. Since $\widehat{m} \in L^p(\mathbf{R}^n)$ it follows that

$$p(\tfrac{1}{2}(n-1) + \alpha + 1) > n, \quad \text{that is, } \tfrac{1}{p} < \tfrac{1}{2} + \tfrac{1}{2n} + \tfrac{\alpha}{n}$$

which proves (5.1.13).

It is not known whether in general the necessary conditions $\alpha > 0$ in Theorem 5.1.6 and (5.1.13) in Theorem 5.1.8 are sufficient to guarantee that say $R_\alpha \in \mathscr{M}_p$. An exception is the two-dimensional case which will be studied in Section 5.2. When $n > 2$ the sufficiency will be proved in Section 5.3 when $|1/p - 1/2| > 1/(n + 1)$. It is actually known in a slightly wider range, but the proofs are then much more difficult and we shall only give some references to the literature for such results.

5.2. The two-dimensional case. The proof of Theorem 5.1.8 showed that after appropriate localization the operator $\varrho_+^\alpha(D)$ in $\mathbf{R}^n$ is essentially a convolution with the product of a function which is homogeneous of degree $-\frac{1}{2}(n + 1) - \alpha$ at infinity and an oscillatory factor $e^{i\Phi}$ with Φ homogeneous of degree 1, which comes from the phase factor in (2.4.7). To study such operators we shall begin with a modification of the Hausdorff-Young inequality (Theorem 2.3.1) which is valid in any dimension.

THEOREM 5.2.1. *Let $a \in C_0^\infty(\mathbf{R}^{2n})$, let $\varphi \in C^\infty(\mathbf{R}^{2n})$ be real valued, and assume that* $\det(\partial^2\varphi(x, y)/\partial x\partial y) \neq 0$ *when* $(x, y) \in \operatorname{supp} a$. *Set*

$$T_N u(x) = \int e^{iN\varphi(x,y)} a(x, y)u(y)\, dy, \quad u \in L^1_{\text{loc}}(\mathbf{R}^n). \tag{5.2.1}$$

If $1 \le p \le 2$ *and* $1/p + 1/p' = 1$, *then*

$$N^{n/p'} \|T_N u\|_{p'} \le C\|u\|_p, \quad u \in L^p(\mathbf{R}^n),\ N \ge 0. \tag{5.2.2}$$

Note that if $1/p + 1/p' = 1$ then $p \le 2$ is equivalent to $p' \ge p$.

PROOF. By the Riesz-Thorin convexity theorem (Theorem 2.3.2) we only have to prove the estimate for $p = 1$ and for $p = 2$, and it is trivial for $p = 1$. To prove it when $p = 2$ we write

$$\|T_N u\|^2 = \iint a_N(y,z) u(y)\overline{u(z)}\, dy\, dz,$$
$$a_N(y,z) = \int e^{iN(\varphi(x,y)-\varphi(x,z))} a(x,y)\overline{a(x,z)}\, dx.$$

If $\det \varphi''_{xy} \ne 0$ at (x^0, y^0) then it follows from Taylor's formula that

$$|\varphi'_x(x,y) - \varphi'_x(x,z)| = |\varphi''_{xy}(x,y)(y-z)| + O(|y-z|^2) \ge c|y-z|$$

for some $c > 0$ if $|x - x^0| + |y - y^0| + |z - y^0|$ is sufficiently small. If this is true when $(x,y) \in \operatorname{supp} a$ and $(x,z) \in \operatorname{supp} a$ we can apply Theorem 2.4.1$'$ with the phase function $x \mapsto (\varphi(x,y) - \varphi(x,z))/|y-z|$ and $\tau = N|y-z|$ and obtain

$$|a_N(y,z)| \le C_k (1 + N|y-z|)^{-k}$$

for any positive integer k. When $k = n+1$ it follows that

$$\|T_N u\|^2 \le C_{n+1} \iint (1 + N|y-z|)^{-n-1} |u(y)||u(z)|\, dy\, dz$$
$$\le C_{n+1} \|u\|^2 \int (1 + N|y|)^{-n-1}\, dy = CN^{-n}\|u\|^2,$$

which proves (5.2.2) when $\operatorname{supp} a$ is sufficiently close to (x^0, y^0). By hypothesis we can always choose a partition of unity $1 = \sum_1^J \chi_j$ in a neighborhood of $\operatorname{supp} a$ so that (5.2.2) is valid with a replaced by $\chi_j a$, $j = 1, \dots, J$. Hence the estimate follows for $a = \sum_1^J \chi_j a_j$.

The hypothesis $\det \varphi''_{xy} \ne 0$ in Theorem 5.2.1 is quite essential. In fact, assume for example that $a(0,0) \ne 0$ and that (5.2.2) is valid. With $A = \varphi''_{xy}(0,0)$ we have by Taylor's formula

$$\varphi(x,y) = \varphi(x,0) + \varphi(0,y) - \varphi(0,0) + \langle x, Ay\rangle + O(|x|^3 + |y|^3).$$

If we choose $v \in C_0^\infty(\mathbf{R}^n)$ and set $u_\varepsilon(y) = v(y/\varepsilon) e^{-iN\varphi(0,y)}$, then

$$(T_N u_\varepsilon)(\varepsilon x) = \int e^{iN(\varphi(\varepsilon x, \varepsilon y) - \varphi(0,\varepsilon y))} a(\varepsilon x, \varepsilon y) v(y) \varepsilon^n\, dy$$
$$= e^{iN(\varphi(\varepsilon x,0)-\varphi(0,0))} \varepsilon^n \int e^{iN\varepsilon^2\langle x,Ay\rangle + O(N\varepsilon^3)} (a(0,0) + O(\varepsilon)) v(y)\, dy.$$

With $\varepsilon = 1/\sqrt{N}$ the integral converges to $a(0,0)\hat{v}(-\,{}^tAx)$ uniformly on compact sets. Since

$$\|T_N u_\varepsilon\|_{p'}^{p'} = \int |(T_N u)(\varepsilon x)|^{p'} \varepsilon^n \, dx,$$

it follows that

$$\varliminf_{N\to\infty} \|T_N u_\varepsilon\|_{p'} \varepsilon^{-n(1+1/p')} \geq |a(0,0)| \|\hat{v}(-\,{}^tA\cdot)\|_{p'} = |a(0,0)| |\det A|^{-1/p'} \|v\|_{p'}.$$

We have $\|u_\varepsilon\|_p = \|v\|_p \varepsilon^{n/p}$ and $\varepsilon^{n(1+1/p')-n/p} = \varepsilon^{2n/p'} = N^{-n/p'}$, so the estimate (5.2.2) implies that $|a(0,0)| \|\hat{v}\|_{p'} \leq C |\det A|^{1/p'} \|v\|_p$. When v is fixed this gives a positive lower bound for $|\det A|$ where $|a|$ has a positive lower bound. In addition we see that the Hausdorff-Young inequality is a consequence of Theorem 5.2.2.

However, a weakened version of (5.2.2) is always valid:

THEOREM 5.2.1′. *Let $a \in C_0^\infty(\mathbf{R}^n \times \mathbf{R}^m)$, let $\varphi \in C^\infty(\mathbf{R}^n \times \mathbf{R}^m)$ be real valued, and let ν be the minimum rank of $\partial^2\varphi(x,y)/\partial x \partial y$ when $(x,y) \in \operatorname{supp} a$. Set*

$$T_N u(x) = \int e^{iN\varphi(x,y)} a(x,y) u(y) \, dy, \quad u \in L^1_{\text{loc}}(\mathbf{R}^m). \tag{5.2.1′}$$

If $1 \leq p \leq 2$ and $1/p + 1/p' = 1$, then

$$N^{\nu/p'} \|T_N u\|_{p'} \leq C \|u\|_p, \quad u \in L^p(\mathbf{R}^m), \; N \geq 0. \tag{5.2.2′}$$

PROOF. It is sufficient to prove this when $p = 2$. For arbitrary $(x^0, y^0) \in \operatorname{supp} a$ we can label the coordinates so that $\det(\partial^2\varphi(x,y)/\partial x_j \partial y_k)_{j,k=1}^\nu \neq 0$ at (x^0, y^0). Set

$$x' = (x_1, \dots, x_\nu), \; x'' = (x_{\nu+1}, \dots, x_n), \; y' = (y_1, \dots, y_\nu), \; y'' = (y_{\nu+1}, \dots, y_m),$$
$$S_N u(x, y'') = \int e^{iN\varphi(x,y)} a(x,y) u(y) \, dy', \quad x = (x', x''), \; y = (y', y'').$$

It supp a is sufficiently close to (x^0, y^0) it follows from Theorem 5.2.1 that

$$N^\nu \int |S_N u(x, y'')|^2 \, dx' \leq C^2 \int |u(y', y'')|^2 \, dy'.$$

Integration with respect to x'' and y'' gives (5.2.2)′, for $T_N u(x) = \int S_N(x, y'') \, dy''$ implies $|T_N u(x)|^2 \leq C \int |S_N u(x, y'')|^2 \, dy''$ since y'' is bounded in the support, and x'' is also bounded there.

We shall actually need an estimate similar to (5.2.2), but with different L^q norms and other powers of N when φ does not satisfy the assumption in Theorem 5.2.1. We shall begin with a quite degenerate case where $n = 2$ and φ is independent of one of the y variables. A statement which is closer to Theorem 5.2.1 will be given afterwards.

THEOREM 5.2.2. *Let $a \in C_0^\infty(\mathbf{R}^2 \times \mathbf{R})$, let $\varphi \in C^\infty(\mathbf{R}^2 \times \mathbf{R})$ be real valued, and assume that*

$$\det \begin{pmatrix} \partial^2\varphi(x,t)/\partial t\partial x_1 & \partial^2\varphi(x,t)/\partial t\partial x_2 \\ \partial^3\varphi(x,t)/\partial t^2\partial x_1 & \partial^3\varphi(x,t)/\partial t^2\partial x_2 \end{pmatrix} \neq 0, \quad \text{if } (x,t) \in \operatorname{supp} a. \tag{5.2.3}$$

Here $x = (x_1, x_2) \in \mathbf{R}^2$ and $t \in \mathbf{R}$. Set

$$T_N f(x) = \int e^{iN\varphi(x,t)} a(x,t) f(t)\, dt, \quad f \in L^1_{\text{loc}}(\mathbf{R}),\ x \in \mathbf{R}^2. \tag{5.2.4}$$

Then it follows that

$$\|T_N f\|_q \leq C N^{-2/q} (q/(q-4))^{\frac14} \|f\|_r, \quad f \in L^r(\mathbf{R}), \quad \text{if } q > 4 \text{ and } \tfrac{3}{q} + \tfrac{1}{r} = 1. \tag{5.2.5}$$

Note that if $3/q + 1/r = 1$ then $q > 4$ is equivalent to $q > r$.

PROOF. Attempting to apply Theorem 5.2.1 we form

$$F_N(x) = (T_N f(x))^2 = \iint e^{iN(\varphi(x,t)+\varphi(x,s))} a(x,t) a(x,s) f(t) f(s)\, ds\, dt.$$

However, the hypotheses on the phase function are not fulfilled since the determinant

$$\det \begin{pmatrix} \partial^2\varphi(x,t)/\partial x_1\partial t & \partial^2\varphi(x,s)/\partial x_1\partial s \\ \partial^2\varphi(x,t)/\partial x_2\partial t & \partial^2\varphi(x,s)/\partial x_2\partial s \end{pmatrix}$$

vanishes when $t = s$. Subtracting the first column from the second we get by Taylor's formula that

$$\det \begin{pmatrix} \partial^2\varphi(x,t)/\partial x_1\partial t & \partial^2\varphi(x,s)/\partial x_1\partial s \\ \partial^2\varphi(x,t)/\partial x_2\partial t & \partial^2\varphi(x,s)/\partial x_2\partial s \end{pmatrix}$$
$$= (s-t) \det \begin{pmatrix} \partial^2\varphi(x,t)/\partial x_1\partial t & \partial^3\varphi(x,t)/\partial x_1\partial t^2 \\ \partial^2\varphi(x,t)/\partial x_2\partial t & \partial^3\varphi(x,t)/\partial x_2\partial t^2 \end{pmatrix} + O(s-t)^2,$$

so the absolute value is bounded below by $c|t-s|$ for a positive constant c in the support of $a(x,t)a(x,s)$ if $\operatorname{supp} a$ is sufficiently close to a point (x^0, t^0) where (5.2.3) is valid. Now $\varphi(x,t) + \varphi(x,s)$ is a symmetric function in s, t so it can be regarded as a function of (x, u, v) near $(x^0, 2t^0, 0)$, where $u = t + s$ and $v = t - s$, which is even in v. Hence there is a C^∞ function $\Phi(x, u, w)$ in a neighborhood of $(x^0, 2t^0, 0)$ such that $\varphi(x,t) + \varphi(x,s) = \Phi(x,u,w)$ if $u = t + s$ and $w = (t-s)^2$. Similarly $a(x,t)a(x,s) = A(x,u,w)$ where $A \in C^\infty$ and $\operatorname{supp} A$ is close to $(x^0, 2t^0, 0)$. The Jacobian $D(u,w)/D(t,s)$ is equal to $4(s-t)$, and the map $(s,t) \mapsto (u,w)$ is a double cover of the half plane where $w > 0$, so we obtain

$$F_N(x) = \tfrac12 \int_{w>0} e^{iN\Phi(x,u,w)} A(x,u,w) f(t) f(s) |t-s|^{-1}\, du\, dw$$

where $t = \frac{1}{2}(u \pm \sqrt{w})$ and $s = \frac{1}{2}(u \mp \sqrt{w})$. Now Φ satisfies the hypotheses of Theorem 5.2.1 in a neighborhood of $(x^0, 2t^0, 0)$, so Theorem 5.2.1 gives for $1 \le p \le 2$

$$\|T_N f\|_{2p'}^2 = \|F_N\|_{p'} \le CN^{-2/p'}\Big(\int_{w>0} |f(t)f(s)|^p |t-s|^{-p}\,du\,dw\Big)^{\frac{1}{p}}$$
$$= CN^{-2/p'}\Big(2\iint |f(t)|^p|f(s)|^p|t-s|^{1-p}\,ds\,dt\Big)^{\frac{1}{p}}.$$

We can estimate the right-hand side using the classical Hardy-Littlewood inequality proved in an example following Theorem 4.1.2 which states that

$$\iint |s-t|^{\gamma-1}|f(s)||g(t)|\,ds\,dt \le C_{p_1,p_2}\|f\|_{p_1}\|g\|_{p_2} \tag{5.2.6}$$

if $1/p_1 + 1/p_2 = 1+\gamma > 1$ and $1 < p_j < \infty$. With $\gamma = 2-p$ and $1/p_1 = 1/p_2 = (3-p)/2$ we obtain

$$\Big(\iint |f(t)f(s)|^p|s-t|^{1-p}\,ds\,dt\Big)^{\frac{1}{p}} \le C(2-p)^{-1/p}\|f\|_{2p/(3-p)}^2, \quad 1 \le p < 2,$$

for inspection of the proof of (5.2.6) gives $C_{p_1,p_2} \le C/\gamma = C/(2-p)$. We leave the verification as an exercise. Hence

$$\|T_N f\|_{2p'} \le CN^{-1/p'}(2-p)^{-1/2p}\|f\|_{2p/(3-p)}.$$

With the notation $2p' = q$ and $2p/(3-p) = r$ we have $1/r + 3/q = 3/2p - 1/2 + 3/2p' = 1$, and $p < 2$ means $q > 4$, $2-p = (q-4)/(q-2)$. Since $1/2p < 1/4$ the inequality (5.2.5) is now proved if $\operatorname{supp} a$ is sufficiently close to a point where (5.2.3) is valid. As in the proof of Theorem 5.2.1 a partition of unity completes the proof.

The condition (5.2.3) means that the first and second derivatives of $\partial\varphi(x,t)/\partial x$ with respect to t are linearly independent. Thus $t \mapsto \partial\varphi(x,t)/\partial x \in \mathbf{R}^2$ defines a smooth immersed curve with curvature different from 0. In the special case where φ is linear in x, the curve is independent of x and we are led to the following:

Corollary 5.2.3. *Let I be an open interval on $\mathbf{R}$, and let $I \ni t \mapsto \Phi(t) \in \mathbf{R}^2$ be an immersion of I as a curve Γ with curvature $\neq 0$. Set*

$$Sf(x) = \int e^{i\langle x,\Phi(t)\rangle}a(t)f(t)\,dt, \quad f \in L^1_{\text{loc}}(\mathbf{R}),\ x \in \mathbf{R}^2, \tag{5.2.7}$$

where $a \in C_0^\infty(I)$. Then it follows that

$$\|Sf\|_q \le C(q/(q-4))^{\frac{1}{4}}\|f\|_r, \quad f \in L^r(\mathbf{R}), \quad \text{if } q > 4,\ \tfrac{3}{q} + \tfrac{1}{r} = 1. \tag{5.2.8}$$

With $\hat{g}$ denoting the Fourier transform of $g \in L^1(\mathbf{R}^2) \cap L^q(\mathbf{R}^2)$ we have

$$\|a(\hat{g} \circ \Phi)\|_{L^r(I)} \le C(4-3q)^{-\frac{1}{4}}\|g\|_q, \quad \text{if } 1 \le q < \tfrac{4}{3},\ \tfrac{3}{q} + \tfrac{1}{r} = 3. \tag{5.2.9}$$

Proof. The function $\varphi(x,t) = \langle x, \Phi(t)\rangle$ satisfies (5.2.3) in every compact subset of $\mathbf{R}^2 \times I$. Choose $b \in C_0^\infty(\mathbf{R}^2)$ with $b(0) = 1$ and apply Theorem 5.2.2 with a replaced by $b(x)a(t)$. This gives

$$N^{2/q}\Big(\int |b(x)(Sf)(Nx)|^q\,dx\Big)^{\frac{1}{q}} \le C(q/(q-4))^{\frac14}\|f\|_r.$$

If we introduce Nx as a new integration variable in the left-hand side and let $N \to \infty$, the estimate (5.2.8) follows. The estimate (5.2.9) is dual: We have

$$\langle a(\hat g \circ \Phi), f\rangle = \iint e^{-i\langle x,\Phi(t)\rangle} g(x)a(t)f(t)\,dx\,dt = \int (Sf)(-x)g(x)\,dx,$$

so Hölder's inequality and (5.2.8) gives

$$|\langle a(\hat g \circ \Phi), f\rangle| \le \|Sf\|_q\|g\|_{q'} \le C(q/(q-4))^{\frac14}\|f\|_r\|g\|_{q'}.$$

Since $3/q' + 1/r' = 3$ and $q/(q-4) = q'/(4-3q')$ the inequality (5.2.9), with q and r replaced by q' and r', follows from the converse of Hölder's inequality.

We shall now rephrase Theorem 5.2.2 in closer analogy to Theorem 5.2.1.

Theorem 5.2.4. *Let $a \in C_0^\infty(\mathbf{R}^2 \times \mathbf{R}^2)$, let $\varphi \in C^\infty(\mathbf{R}^2 \times \mathbf{R}^2)$ be real valued, and suppose that when $(x,y) \in \operatorname{supp} a$ we have $\partial^2\varphi(x,y)/\partial x\partial y \ne 0$ and*

(5.2.10) $$\partial^2\langle t, \partial\varphi(x,y)/\partial x\rangle/\partial y^2 \ne 0, \quad \text{if } 0 \ne t \in \mathbf{R}^2,\ \partial\langle t, \partial\varphi(x,y)/\partial x\rangle/\partial y = 0.$$

If T_N is defined by (5.2.1) *with $n = 2$ then*

(5.2.11) $$\|T_N u\|_q \le CN^{-2/q}(q/(q-4))^{\frac14}\|u\|_r, \quad u \in L^r(\mathbf{R}^2), \quad \text{if } q > 4 \text{ and } \tfrac{3}{q} + \tfrac{1}{r} = 1.$$

Proof. Since $r' = q/3 < q$, thus $r > q'$, the estimate (5.2.10) follows from (5.2.2), with $n = 2$ and $p = q'$, if $\det \partial^2\varphi/\partial x\partial y \ne 0$ in $\operatorname{supp} a$. It is therefore sufficient to prove (5.2.11) when $\operatorname{supp} a$ is in a small neigbhorhood of a point (x^0, y^0) where $\partial^2\varphi(x,y)/\partial x\partial y$ has rank 1 and (5.2.10) is valid. After an affine change of x variables we may assume that $x^0 = 0$ and that $\partial^2\varphi/\partial x_1\partial y = 0$ at (x^0,y^0). Then $\partial^2\varphi/\partial x_2\partial y \ne 0$ and $\partial^3\varphi/\partial x_1\partial y^2 \ne 0$ at (x^0,y^0). After an affine change of y variables we may therefore assume that $y^0 = 0$ and that $\partial^2\varphi/\partial x_2\partial y_1 \ne 0$, $\partial^3\varphi/\partial x_1\partial y_1^2 \ne 0$ at $(0,0)$. Since $\partial^2\varphi/\partial x_1\partial y_1 = 0$ at $(0,0)$ it follows that $(x,t) \mapsto \varphi(x,t,y_2)$ satisfies (5.2.3) in a neighborhood of the origin when y_2 is fixed and small, for the determinant is equal to $-\partial^2\varphi/\partial x_2\partial y_1\partial^3\varphi/\partial x_1\partial y_1^2$ at the origin. Writing

$$\mathscr{T}_N u(x,y_2) = \int e^{iN\varphi(x,y)}a(x,y)u(y)\,dy_1$$

we have by (5.2.5) since $T_N u(x) = \int \mathscr{T}_N u(x,y_2)\,dy_2$

$$\|T_N u\|_q \le \int_K dy_2\|\mathscr{T}_N u(\cdot,y_2)\|_{L^q(\mathbf{R}^2)} \le CN^{-2/q}(q/(q-4))^{\frac14}\int_K \|u(\cdot,y_2)\|_{L^r(\mathbf{R})}\,dy_2,$$

where K is a compact set such that $a(x,y) = 0$ if $y_2 \notin K$. By Hölder's inequality the integral in the right-hand side is $\le m(K)^{1/r'}\|u\|_r$, which completes the proof.

Corollary 5.2.5. *Let $\Phi \in C^\infty(\mathbf{R}^2 \setminus \{0\})$ be real valued and positively homogeneous of degree 1, and let $A \in C_0^\infty(\mathbf{R}^2 \setminus \{0\})$. Set*

$$S_t f(x) = \int e^{it\Phi(x-y)} A(x-y) f(y)\, dy, \quad f \in L^1_{\mathrm{loc}}(\mathbf{R}^2).$$

If $\Phi'' \neq 0$ in $\operatorname{supp} A$, it follows that

$$\|S_t f\|_p \le C_p(t) \|f\|_p, \quad f \in L^p(\mathbf{R}^2),\ t > 2,\ p \ge 2, \quad \textit{where} \tag{5.2.12}$$

$$C_p(t) = \begin{cases} Ct^{-2/p}(p/(p-4))^{\frac14}, & \textit{if } p > 4, \\ Ct^{-\frac12}(\log t)^{\frac12 - \frac1p}, & \textit{if } 2 \le p \le 4. \end{cases} \tag{5.2.13}$$

Proof. The phase function $\varphi(x,y) = \Phi(x-y)$ satisfies the hypotheses of Theorem 5.2.4 when $x - y \in \operatorname{supp} A$. In fact, $\Phi''(z)z = 0$ since Φ' is homogeneous of degree 0, and since $\Phi''(z) \neq 0$ it follows that $\Phi''(z)t = 0$ implies that t is proportional to z, and $\Phi'''(z)z = -\Phi''(z)$ since Φ'' is homogeneous of degree -1.

Let $0 \le \chi \in C_0^\infty(\mathbf{R}^2)$, $\int \chi^2\, dx = 1$. Then the hypotheses of Theorem 5.2.4 are fulfilled if $a(x,y) = A(x-y)\chi(y)$. If $p > 4$ it follows that

$$\|S_t \chi^2 f\|_p \le C_p(t) \|\chi f\|_p, \quad f \in L^p(\mathbf{R}^2), \tag{5.2.14}$$

for $r = p/(p-3) < p$. Theorem 5.2.1$'$ gives the estimate (5.2.14) when $p = 2$. The Riesz-Thorin interpolation theorem (Theorem 2.3.2) gives (5.2.14) for $p < 2 \le 4$ if we apply it between $p_1 = 2$ and $p_2 = 4 + 1/\log t$.

By the translation invariance of S_t it follows that for $z \in \mathbf{R}^2$

$$\|S_{t,z} f\|_p \le C_p(t) \|\chi(\cdot - z) f\|_p, \quad \text{where } S_{t,z} f = S_t(\chi(\cdot - z)^2 f). \tag{5.2.15}$$

Since

$$S_{t,z} f(x) = \int e^{it\Phi(x-y)} A(x-y) \chi(y-z)^2 f(y)\, dy,$$

the integral with respect to z is equal to $S_t f(x)$, and since $A(x-y)\chi(y-z) \neq 0$ implies a bound for $x - z$, we have by Hölder's inequality

$$|S_t f(x)|^p \le C^p \int |S_{t,z} f(x)|^p\, dz.$$

If we raise the estimate (5.2.15) to the power p and integrate with respect to z it follows that

$$\|S_t f\|_p \le C C_p(t) \|\chi\|_p \|f\|_p$$

which completes the proof.

Since Φ is homogeneous we have

$$(S_t f)(x/t) = \int e^{i\Phi(x-ty)} A(x/t - y) f(y)\, dy = \int e^{i\Phi(x-y)} A((x-y)/t) g(y)\, dy \tag{5.2.16}$$

where $g(y) = f(y/t)/t^2$, and (5.2.12) gives

$$\|(S_t f)(\cdot/t)\|_p = t^{2/p}\|S_t f\|_p \le t^{2/p}C_p(t)\|f\|_p = t^2 C_p(t)\|g\|_p$$

If we multiply (5.2.16) by $t^{-1-\lambda}$ and integrate with respect to t from 2 to ∞, it follows that the operator

$$K_\lambda g(x) = \int e^{i\Phi(x-y)} b_\lambda(x-y)g(y)\,dy, \tag{5.2.17}$$

$$b_\lambda(z) = \int_2^\infty A(z/t)t^{-1-\lambda}dt, \tag{5.2.18}$$

is bounded in L^p provided that $p \ge 2$ and that

$$\int_2^\infty C_p(t)t^{1-\lambda}\,dt < \infty. \tag{5.2.19}$$

By (5.2.13) this condition is equivalent to

$$\lambda > \begin{cases} 2(1-\frac{1}{p}), & \text{if } p \ge 4 \\ \frac{3}{2}, & \text{if } 2 \le p \le 4. \end{cases}$$

The function b defined by (5.2.18) is positively homogeneous of degree $-\lambda$ when $|z|$ is so large that $z/t \in \operatorname{supp} A$ implies $t \ge 2$, for then we have $b(z) = b_0(z)$ where

$$b_0(z) = \int_0^\infty A(z/t)t^{-1-\lambda}\,dt. \tag{5.2.20}$$

Every b_0 which is positively homogeneous of degree $-\lambda$ is of the form (5.2.20) with $A(z) = b_0(z)\varrho(z)$ where $\varrho \in C_0^\infty(\mathbf{R}^2 \setminus \{0\})$ is chosen so that $\int_0^\infty \varrho(z/t)\,dt/t = 1$ for $z \ne 0$. This is true for any non-negative C^∞ function of $|z|$ with support in $(1,2)$ after multiplication by a suitable normalizing constant. This gives the following theorem when $p \ge 2$, and when $p \le 2$ it follows by duality.

Theorem 5.2.6. *Let $\Phi \in C^\infty(\mathbf{R}^2 \setminus \{0\})$ be real valued and positively homogeneous of degree 1, and let $a_0 \in C^\infty(\mathbf{R}^2 \setminus \{0\})$ be positively homogeneous of degree $-\lambda$. Assume that $\Phi''(x) \ne 0$ when $a_0(x) \ne 0$, $x \ne 0$, and that*

$$\lambda > \max(\tfrac{3}{2}, 2|\tfrac{1}{p} - \tfrac{1}{2}| + 1). \tag{5.2.21}$$

If $a \in L^1_{\mathrm{loc}}(\mathbf{R}^2)$ is equal to a_0 outside a compact set, then the operator $f \mapsto (e^{i\Phi}a) * f$ extends from $L^1(\mathbf{R}^2) \cap L^\infty(\mathbf{R}^2)$ to a continuous operator in $L^p(\mathbf{R}^2)$.

In fact, we have just proved the continuity of this operator for some $b \in C^\infty$ equal to a_0 outside a compact set. Hence $b - a \in L^1$, and $f \mapsto (b-a) * f$ is continuous in L^p for every p.

Theorem 5.2.6 gives the sufficiency of the necessary conditions in Theorem 5.1.8:

THEOREM 5.2.7. *Let $\varrho \in C^\infty(\mathbf{R}^2)$ be real valued and negative outside a compact set, and assume that $\varrho'(\xi) \neq 0$ and that $\varrho''(\xi)t \neq 0$ if $\xi \in \mathbf{R}^2$, $\varrho(\xi) = 0$ and $0 \neq t \in \mathbf{R}^2$, $\varrho'(\xi)t = 0$. Then $\varrho_+^\alpha \in \mathscr{M}_p(\mathbf{R}^2)$ if*

$$\alpha > \max\left(0, 2\left|\tfrac{1}{p} - \tfrac{1}{2}\right| - \tfrac{1}{2}\right). \tag{5.2.22}$$

PROOF. The zeros of ϱ form a compact set, so a partition of unity shows that it is sufficient to prove that for every ξ^0 with $\varrho(\xi^0) \neq 0$ there is some $\varphi \in C_0^\infty$ with $\varphi(\xi^0) \neq 0$ such that $\varphi\varrho_+^\alpha \in \mathscr{M}_p(\mathbf{R}^2)$. If we choose φ as in the proof of Theorem 5.1.8, keeping the notation there, we find that it is sufficient to prove L^p continuity of convolution with a C^∞ kernel K such that outside a compact set $K(x) - e^{i\Phi(x)}a_0(x) = O(|x|^{-1-\lambda})$ where a_0 is homogeneous of degree $-\lambda$ and $\lambda = 3/2 + \alpha$. Here Φ is defined in $\operatorname{supp} a_0$ by $\Phi(x) = x_1 t + x_2\psi(t)$ where t is a C^∞ function of x, homogeneous of degree 0, determined by the equation $x_1 + x_2\psi'(t) = 0$; $\psi''(t) \neq 0$. Thus $\Phi'(x) = (t, \psi'(t))$ and $\partial^2\Phi(x)/\partial x_1^2 = \partial t/\partial x_1 = -1/(x_2\psi''(t)) \neq 0$. Since (5.2.22) is identical to (5.2.21), the theorem is proved.

The hypothesis in Theorem 5.2.7 that the curvature of $\{\xi; \varrho(\xi) = 0\}$ is not equal to 0 is superfluous. A fairly simple localization argument proves that the result remains valid if there are no points which are flat of infinite order, and Sjölin [1] has proved it for arbitrary smooth curves. (However, the necessity proved in Theorem 5.1.8 requires a non-zero curvature at some point.) This suggests that there should exist a more natural proof which does not aim so directly at the kernel of the convolution operator corresponding to the multiplier, but none seems to be known.

5.3. The higher dimensional case. Our first goal is to prove an analogue of Theorem 5.2.2 which was the key to the results in Section 5.2. Thus let $\varphi \in C^\infty(\mathbf{R}^{2n-1})$ be real valued, let $a \in C_0^\infty(\mathbf{R}^{2n-1})$, and consider the operator analogous to (5.2.4) defined by

$$T_N f(x) = \int e^{iN\varphi(x,y)}a(x,y)f(y)\,dy, \quad f \in L^1_{\text{loc}}(\mathbf{R}^{n-1}). \tag{5.3.1}$$

Here the variables in $\mathbf{R}^{2n-1}$ are denoted by (x,y) where $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^{n-1}$. We shall assume that

$$\operatorname{rank}(\partial^2\varphi(x,y)/\partial x\partial y) = n-1 \quad \text{when } (x,y) \in \operatorname{supp} a. \tag{5.3.2}$$

Then there is for every $(x,y) \in \operatorname{supp} a$ a vector $t \in \mathbf{R}^n \setminus \{0\}$, uniquely determined up to a constant factor, such that $(\partial/\partial y)\langle\partial\varphi(x,y)/\partial x, t\rangle = 0$. The analogue of the hypothesis (5.2.3) is that for $(x,y) \in \operatorname{supp} a$

(5.3.3)

$$(\partial/\partial y)\langle\partial\varphi(x,y)/\partial x, t\rangle = 0 \implies \det(\partial^2/\partial y^2)\langle\partial\varphi(x,y)/\partial x, t\rangle \neq 0, \quad \text{if } 0 \neq t \in \mathbf{R}^n.$$

The conditions (5.3.2) and (5.3.3) do not change if we add a function of x or a function of y to φ or change the x or the y coordinates, and this does not affect L^rL^q estimates of the form (5.2.5) either, apart from the size of the constants. One can therefore simplify φ using the following lemma.

LEMMA 5.3.1. *If φ satisfies* (5.3.2) *at the origin, then there are new x and y coordinates such that at the origin*

$$\varphi(x,y)-\varphi(x,0)-\varphi(0,y)+\varphi(0,0)=\sum_{j=1}^{n-1}x_jy_j+\tfrac{1}{2}x_n\langle A(y)y,y\rangle+O(|x|^2|y|^2), \tag{5.3.4}$$

where $A(y)$ is a real symmetric $n\times n$ matrix which is a C^∞ function of y.

PROOF. By Taylor's formula applied first in the x variables and then in the y variables we can write

$$\varphi(x,y)-\varphi(x,0)-\varphi(0,y)+\varphi(0,0)=\sum_{j=1}^{n}\sum_{k=1}^{n-1}c_{jk}(x,y)x_jy_k.$$

The condition (5.3.2) at the origin means that $\sum c_{jk}(0,0)x_jy_k=\sum_1^{n-1}L_k(x)y_k$ where the linear forms L_k are linearly independent. By a linear change of the x variables we can achieve that $L_k(x)=x_k$ which we assume now. Again by Taylor's formula we can write

$$c_{jk}(x,y)=c_{jk}(0,0)+\sum_{l=1}^{n}d_{jkl}(x)x_l+\sum_{l=1}^{n-1}e_{jkl}(y)y_l+R_{jk}(x,y)$$

where $R_{jk}(x,y)=O(|x||y|)$. This gives

$$\varphi(x,y)=\sum_{j=1}^{n-1}\Big(x_j+\sum_{k,l=1}^{n}d_{kjl}(x)x_kx_l\Big)\Big(y_j+\sum_{k,l=1}^{n-1}e_{jkl}(y)y_ky_l\Big)+x_n\sum_{k,l=1}^{n-1}e_{nkl}(y)y_ky_l+R(x,y)$$

where $R(x,y)=O(|x|^2|y|^2)$. This proves the lemma.

The condition (5.3.3) with $(x,y)=(0,0)$ means that the matrix $A(0)$ is non-singular, so this is an invariant condition. When examining the conditions for the validity of an L^rL^q estimate of the form (5.2.5) for T_N we may assume that $\varphi(x,0)\equiv 0$, $\varphi(0,y)\equiv 0$, for otherwise φ may be replaced by the left-hand side of (5.3.4), and we assume that $a=1$ in a neighborhood of the origin. Let $f\in C_0^\infty(\mathbf{R}^{n-1})$ have so small support that $a=1$ in a neighborhood of $\{0\}\times\operatorname{supp} f$. To examine T_Nf in a conic neighborhood of the x_n axis close to the origin we set $x'=(x_1,\dots,x_{n-1})=x_nz$ and note that with some $\psi\in C^\infty$

$$\varphi(x,y)=\varphi((x_nz,x_n),y)=x_n(\langle z,y\rangle+\tfrac{1}{2}\langle A(y)y,y\rangle+x_n\psi(z,x_n,y)).$$

Hence $y\mapsto\varphi(x,y)/x_n$ has a unique non-degenerate critical point close to the origin if x_n and z are sufficiently small. If $\operatorname{supp} f$ is sufficiently small it follows from the method of stationary phase, Theorem 2.4.5, that there are positive constants $c_1,\dots,c_4$ such that

$$|T_Nf(x_nz,x_n)|\ge c_4(Nx_n)^{\frac{1}{2}(1-n)},\quad \text{if } |z|<c_3,\ c_1/N<x_n<c_2.$$

Hence, with another positive constant c_5,

$$\int |T_N f(x)|^q \, dx \geq c_5 N^{\frac{1}{2}(1-n)q} \int_{c_1/N}^{c_2} x_n^{(n-1)(1-\frac{1}{2}q)} dx_n.$$

Since $(n-1)(1-\frac{1}{2}q)+1 = nq(\frac{1}{q}-\frac{1}{2}+\frac{1}{2n})$ and $\frac{1}{2}(1-n)q-(n-1)(1-\frac{1}{2}q)-1=-n$, it follows if we distinguish the cases where the integral on the right converges or diverges at 0 that

$$\varliminf_{N\to\infty} \|T_N f\|_q N^{\frac{1}{2}(n-1)} \geq C\big(\tfrac{1}{q}-\tfrac{1}{2}+\tfrac{1}{2n}\big)^{-\frac{1}{q}}, \quad \text{if } \tfrac{1}{2}-\tfrac{1}{2n} < \tfrac{1}{q} \leq \tfrac{1}{2}, \tag{5.3.5}$$

$$\varliminf_{N\to\infty} \|T_N f\|_q N^{\frac{1}{2}(n-1)} (\log N)^{-\frac{1}{q}} \geq C, \quad \text{if } \tfrac{1}{q} = \tfrac{1}{2}-\tfrac{1}{2n}, \tag{5.3.6}$$

$$\varliminf_{N\to\infty} \|T_N f\|_q N^{n/q} \geq C\big(\tfrac{1}{2}-\tfrac{1}{q}-\tfrac{1}{2n}\big)^{-\frac{1}{q}}, \quad \text{if } \tfrac{1}{q} < \tfrac{1}{2}-\tfrac{1}{2n}. \tag{5.3.7}$$

The exponent $-1/q$ in (5.3.5) and (5.3.7) may of course be replaced by $1/2n - 1/2$. When $n=2$ we have therefore proved that the growth of the constant in (5.2.5) as $q \to 4$ cannot be improved and also that no estimate of the form (5.2.5) is valid with $q \leq 4$ no matter how r is chosen. Also in the higher dimensional case we conclude that an estimate of the form

$$\|T_N f\|_q \leq C N^{-n/q} \|f\|_r \tag{5.3.8}$$

cannot be valid unless $q > 2n/(n-1)$. To prove a necessary condition on r also we first observe that since

$$(T_N f)(x/N) = \int e^{iN\varphi(x/N,y)} f(y) dy \to \int e^{i\langle x, \Phi(y)\rangle} f(y)\, dy, \quad N \to \infty,$$

where $\Phi(y) = (y, \frac{1}{2}\langle A(y)y, y\rangle)$, it follows from (5.3.8) that

$$\|Sf\|_q \leq C\|f\|_r, \quad \text{where } Sf(x) = \int e^{i\langle x,\Phi(y)\rangle} f(y)\, dy.$$

Now we use a scaling argument. Let $f \in C_0^\infty(\mathbf{R}^{n-1}) \setminus \{0\}$ and set $f_\varepsilon(y) = f(y/\varepsilon)\varepsilon^{(1-n)/r}$ with a small $\varepsilon > 0$. Then $\|f_\varepsilon\|_r$ is independent of ε and

$$(Sf_\varepsilon)(x'/\varepsilon, x_n/\varepsilon^2)\varepsilon^{(1-n)/r'} = S_\varepsilon f(x) \to S_0 f(x)$$

where $1/r + 1/r' = 1$ and S_ε is defined as S but with $A(y)$ replaced by $A(\varepsilon y)$. Hence

$$\varliminf_{\varepsilon\to 0} \varepsilon^{(n+1)/q-(n-1)/r'} \|Sf_\varepsilon\|_q \geq \|S_0 f\|_q,$$

which proves that (5.3.8) implies $(n+1)/q - (n-1)/r' \leq 0$. Summing up, (5.3.8) requires that

$$\tfrac{1}{q} < \tfrac{1}{2} - \tfrac{1}{2n}, \qquad \tfrac{n+1}{(n-1)}\tfrac{1}{q} + \tfrac{1}{r} \leq 1. \tag{5.3.9}$$

When $n=2$ these are precisely the conditions in (5.2.5).

However, when $n > 2$ the conditions (5.3.9) are not sufficient to guarantee that (5.3.8) is valid. The following striking example is essentially due to Bourgain [1].

Example. Let $n = 2k+1$ be odd and set

$$\varphi(x,y) = \sum_1^k (\tfrac12 x_j + y_j + x_n y_{k+j})^2 + \sum_1^k (x_{j+k} - x_n x_j) y_{k+j} \tag{5.3.10}$$
$$= \sum_1^k (\tfrac14 x_j^2 + y_j^2) + \sum_1^{2k} x_j y_j + 2x_n \sum_1^k y_j y_{k+j} + x_n^2 \sum_{k+1}^{2k} y_j^2.$$

Note that this corresponds to $\langle A(y)y, y\rangle = 4\sum_1^k y_j y_{k+j}$ with the notation in Lemma 5.3.1, which means that $A(y)$ is independent of y and has k positive and k negative eigenvalues. If $a \in C_0^\infty(\mathbf{R}^{2n-1})$ and $f \in C_0^\infty(\mathbf{R}^{2n-1})$, then the stationary phase method applied to the integral (5.3.1) in the variables $y_1, \dots, y_k$ gives

$$T_N f(x) = cN^{-\frac12 k} \int \exp\big(iN \sum_1^k (x_{j+k} - x_n x_j) y_{k+j}\big) A(x, y'')\, dy'' + O(N^{-\frac12 k - 1})$$

where $y'' = (y_{k+1}, \dots, y_{2k})$ and $A(x,y'') = a(x,y)f(y)$ when $y_j = -x_n y_{k+j} - \frac12 x_j$ for $j = 1, \dots, k$. We can choose a and f so that $A(0,0) \neq 0$. Then we obtain

$$T_N f(x) = cN^{-\frac12 k} \widehat{A}(x, N\eta) + O(N^{-\frac12 k-1}), \quad \eta_j = x_n x_j - x_{j+k},\ j = 1, \dots, k,$$

where $\hat{A}$ denotes the Fourier transform of $A(x,y'')$ in y''. Hence

$$\varliminf_{N\to\infty} \|T_N f\|_q N^{k(\frac12 + \frac1q)} \geq |c| \Big(\int |\widehat{A}(x', x_n x', x_n, x'')|^q\, dx \Big)^{\frac1q}$$

where $x' = (x_1, \dots, x_k)$ and $x'' = (x_{k+1}, \dots, x_{2k})$. The right-hand side is not 0, so an estimate of the form (5.3.8) cannot hold even with $r = \infty$ unless $k(\frac12 + \frac1q) \geq \frac{n}{q}$, that is,

$$q \geq \tfrac{2n+2}{n-1}. \tag{5.3.11}$$

The preceding example combined with (5.3.9) shows that if n is odd and no hypothesis is made about φ beyond the conditions (5.3.2) and (5.3.3), then the best result which can be proved is the following theorem of Stein [4]:

Theorem 5.3.2. *Let $\varphi \in C^\infty(\mathbf{R}^{2n-1})$ be real valued, let $a \in C_0^\infty(\mathbf{R}^{2n-1})$, and assume that* (5.3.2) *and* (5.3.3) *are valid when $(x,y) \in \operatorname{supp} a$. Then there is a constant C such that with T_N defined by* (5.3.1)

$$\|T_N f\|_q \leq CN^{-n/q} \|f\|_r,\ f \in L^r(\mathbf{R}^{n-1}), \quad \textit{if } q \geq \tfrac{2n+2}{n-1},\ \tfrac{n+1}{n-1}\tfrac1q + \tfrac1r \leq 1. \tag{5.3.12}$$

Proof. Since f may be assumed to have support in a fixed compact set, the statement is strongest when $\frac{n+1}{n-1}\frac1q + \frac1r = 1$. Since (5.3.12) is trivial when $r = 1$ and $q = \infty$, it follows from the Riesz-Thorin interpolation theorem (Theorem 2.3.2) that it suffices to prove (5.3.12) when $r = 2$ and $q = (2n+2)/(n-1)$. It is then more convenient to discuss the equivalent dual estimate

$$\|T_N^* u\|_2 \le CN^{-n(n-1)/(2n+2)}\|u\|_{(2n+2)/(n+3)}, \quad u \in L^{(2n+2)/(n+3)}(\mathbf{R}^n). \tag{5.3.13}$$

The square of the left-hand side is equal to $(T_N T_N^* u, u)$ which by Hölder's inequality can be estimated by $\|T_N T_N^* u\|_{(2n+2)/(n-1)}\|u\|_{(2n+2)/(n+3)}$. Hence (5.3.13) will follow if we can prove that

$$\|T_N T_N^* u\|_{(2n+2)/(n-1)} \le C^2 N^{-n(n-1)/(n+1)}\|u\|_{(2n+2)/(n+3)}, \quad u \in L^{(2n+2)/(n+3)}(\mathbf{R}^n). \tag{5.3.14}$$

It is of course sufficient to prove (5.3.14) when φ has the form of the right-hand side in (5.3.4) and the support of a is very close to the origin.

In the proof of (5.3.14) we shall single out the x_n variable and write

$$(T_N(x_n)f)(x') = \int e^{iN\varphi(x',x_n,y)}a(x',x_n,y)f(y)\,dy, \quad f \in L^1_{\text{loc}}(\mathbf{R}^{n-1}).$$

Here $x' = (x_1, \dots, x_{n-1})$. Then

$$T_N^* u = \int T_N(s)^* u(\cdot, s)\,ds, \quad u \in L^1_{\text{loc}}(\mathbf{R}^n),$$

$$(T_N T_N^* u)(\cdot, t) = \int T_N(t)T_N(s)^* u(\cdot, s)\,ds. \tag{5.3.15}$$

To complete the proof we need an estimate for $T_N(t)T_N(s)^*$:

LEMMA 5.3.3. *When $1 \le p \le 2$ and $1/p + 1/p' = 1$ then*

$$\|T_N(t)T_N(s)^* f\|_{p'} \le C|t-s|^{-(n-1)(\frac{1}{2}-\frac{1}{p'})}N^{-(n-1)(\frac{1}{2}+\frac{1}{p'})}\|f\|_p, \quad f \in L^p(\mathbf{R}^{n-1}),\ s \ne t. \tag{5.3.16}$$

PROOF. By Theorem 5.2.1 we have with a constant C independent of t

$$\|T_N(t)f\|_2 \le CN^{-\frac{1}{2}(n-1)}\|f\|_2.$$

Since the same estimate is valid for the adjoint it follows that

$$\|T_N(t)T_N(s)^* f\|_2 \le CN^{-(n-1)}\|f\|_2,$$

which is the estimate (5.3.16) for $p = 2$. By the Riesz-Thorin interpolation theorem (Theorem 2.3.2) the estimate will follow for $1 \le p \le 2$ if we can prove it when $p = 1$, that is, prove that

$$\|T_N(t)T_N(s)^* f\|_\infty \le C|t-s|^{-\frac{1}{2}(n-1)}N^{-\frac{1}{2}(n-1)}\|f\|_1.$$

Such a bound is equivalent to an estimate

$$\|K_{t,s}\|_\infty \le C(N|t-s|)^{-\frac{1}{2}(n-1)}, \tag{5.3.17}$$

where $K_{t,s}$ is the kernel of $T_N(t)T_N(s)^*$, that is,

$$K_{t,s}(x',z) = \int_{\mathbf{R}^{n-1}} e^{iN(\varphi(x',t,y)-\varphi(z,s,y))} a(x',t,y)\overline{a(z,s,y)}\, dy, \quad x',z \in \mathbf{R}^{n-1}.$$

First assume that $z=0$, $s=0$, and that the coordinates have been chosen according to Lemma 5.3.1. Then

$$\frac{\partial}{\partial y}\big(\varphi(x',t,y)-\varphi(z,s,y)\big) = x' + t(A(y)y + O(|y|^2)) + O(|x'|^2+|t|^2)|y|.$$

If $|t|\le |x'|$ and $|y|$, $|x'|$ are sufficiently small, the norm is $\ge |x'|/2$, and all derivatives with respect to y are $O(|x'|)$. Hence it follows from Theorem 2.4.1 that

$$|K_{t,0}(x',0)| \le C_k(N|x'|)^{-k} \le C_k(N|t|)^{-k}, \quad k\ge 0,$$

if the support of a is sufficiently close to the origin. On the other hand, if $|x'|<t$ then

$$\frac{\partial^2}{\partial y^2}\big(\varphi(x',t,y)-\varphi(z,s,y)\big) = t(A(y)+O(|y|+|t|)),$$

so the absolute value of the determinant of the quotient by t has a positive lower bound if t and y are small enough, and all derivatives of $\varphi(x',t,y)/t$ with respect to y are also uniformly bounded then. Hence it follows from Theorem 2.4.3 that

$$|K_{t,0}(x',0)| \le C(N|t|)^{-\frac{1}{2}(n-1)},$$

if the support of a is sufficiently close to the origin. For any other z and s close to the origin the same conclusions are obtained after the change of coordinates achieved in Lemma 5.3.1, which completes the proof.

End of the proof of Theorem 5.3.2. When $p=(2n+2)/(n+3)$ and $p'=(2n+2)/(n-1)$ we obtain from (5.3.16) using (5.3.15) and Minkowski's inequality

$$V(t) \le C\int |t-s|^{-(n-1)/(n+1)} N^{-n(n-1)/(n+1)} U(s)\, ds, \quad \text{where}$$
$$V(t) = \|(T_N T_N^* u)(\cdot,t)\|_{p'}, \quad U(s) = \|u(\cdot,s)\|_p.$$

Since $1/p-1/p' = 2/(n+1) = 1-(n-1)/(n+1)$, it follows from the Hardy-Littlewood potential estimate (see the example following Theorem 4.1.2) that

$$\Big(\int |V(t)|^{p'}\, dt\Big)^{\frac{1}{p'}} \le CN^{-n(n-1)/(n+1)}\Big(\int |U(s)|^p\, ds\Big)^{\frac{1}{p}},$$

that is,

$$\|T_N T_N^* u\|_{p'} \le CN^{-n(n-1)/(n+1)}\|u\|_p,$$

which is the estimate (5.3.14). The proof is complete.

From this point on we can to a large extent repeat the arguments in Section 5.2. The special case of Theorem 5.3.2 where φ is linear in x merits a special emphasis, as in Corollary 5.2.3:

Corollary 5.3.4. *Let Y be an open subset of $\mathbf{R}^{n-1}$, $n \geq 3$, and let $Y \ni y \mapsto \Phi(y) \in \mathbf{R}^n$ be an immersed hypersurface with total curvature $\neq 0$. Set*

$$Sf(x) = \int e^{i\langle x, \Phi(y)\rangle} a(y) f(y)\, dy, \quad f \in L^1_{\text{loc}}(\mathbf{R}^{n-1}),\ x \in \mathbf{R}^n, \tag{5.3.18}$$

where $a \in C_0^\infty(Y)$. Then it follows that

$$\|Sf\|_q \leq C\|f\|_r, \quad f \in L^r(\mathbf{R}^{n-1}), \quad \text{if } q \geq \tfrac{2n+2}{n-1} \text{ and } \tfrac{n+1}{n-1}\tfrac{1}{q} + \tfrac{1}{r} \leq 1. \tag{5.3.19}$$

With $\hat{g}$ denoting the Fourier transform of $g \in L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n)$ we have

$$\|a(\hat{g} \circ \Phi)\|_r \leq C\|g\|_q, \quad \text{if } 1 \leq q \leq \tfrac{2n+2}{n+3} \text{ and } \tfrac{n+1}{n-1}\tfrac{1}{q} + \tfrac{1}{r} \geq \tfrac{n+1}{n-1}. \tag{5.3.20}$$

Proof. The function $\varphi(x, y) = \langle x, \Phi(y)\rangle$ satisfies (5.3.2), (5.3.3) when $y \in Y$. In fact, since $\Phi'(y)$ is injective, the rank of $\partial^2\varphi(x,y)/\partial x\partial y = \partial\Phi(y)/\partial y$ is $n-1$. The condition $(\partial/\partial y)\langle\Phi(y), t\rangle = 0$ means that t is orthogonal to the tangent plane at $\Phi(y)$. Thus (5.3.2) means that $\det(\partial^2/\partial y^2)\langle\Phi(y), t\rangle \neq 0$ when t is a normal $\neq 0$, that is, that the total curvature is not 0. Choose $b \in C_0^\infty(\mathbf{R}^n)$ with $b(0) = 1$. If we apply Theorem 5.3.2 with $a(x, y)$ replaced by $b(x)a(y)$ it follows that

$$N^{n/q}\Big(\int |b(x)(Sf)(Nx)|^q\, dx\Big)^{\frac{1}{q}} \leq C\|f\|_r, \quad \text{if } q \geq \tfrac{2n+2}{n-1} \text{ and } \tfrac{n+1}{n-1}\tfrac{1}{q} + \tfrac{1}{r} \leq 1.$$

The estimate (5.3.19) follows if we let $N \to \infty$ after introducing Nx as a new integration variable. The estimate (5.3.20) is dual as in the proof of Corollary 5.2.3, and we do not repeat the proof.

The estimate (5.3.20) is known as the restriction theorem. A weaker form was first proved by Tomas [1]. It has been proved by Bourgain [1] that it is valid for a wider range of q but the precise range is not known. One should note that the motivation we gave for the condition $q \geq (2n+2)/(n-1)$ in Theorem 5.3.2 assumed that no information was given on φ beyond conditions (5.3.2) and (5.3.3). Linearity in x is a strong additional hypothesis which extends the permissible range.

Next we prove an analogue of Theorem 5.2.4:

Theorem 5.3.5. *Let $a \in C_0^\infty(\mathbf{R}^n \times \mathbf{R}^n)$, $n \geq 3$, let $\varphi \in C^\infty(\mathbf{R}^n \times \mathbf{R}^n)$ be real valued, and assume that when $(x, y) \in \operatorname{supp} a$ the rank of $\partial^2\varphi(x,y)/\partial x\partial y$ is at least $n-1$ and that*

$$\operatorname{rank} \partial^2\langle t, \partial\varphi(x,y)/\partial x\rangle/\partial y^2 \geq n-1, \quad \text{if } 0 \neq t \in \mathbf{R}^n,\ \partial\langle t, \partial\varphi(x,y)/\partial x\rangle/\partial y = 0. \tag{5.3.21}$$

If T_N is defined by (5.2.1) then

$$\|T_N u\|_q \leq CN^{-n/q}\|u\|_r, \quad u \in L^r(\mathbf{R}^n), \quad \text{if } q \geq \tfrac{2n+2}{n-1} \text{ and } \tfrac{n+1}{n-1}\tfrac{1}{q} + \tfrac{1}{r} \leq 1. \tag{5.3.22}$$

Proof. Since $r' \leq q(n-1)/(n+1) < q$, thus $r > q'$, the estimate (5.3.22) follows from (5.2.2) with $p = q'$ if $\det \partial^2\varphi(x,y)/\partial x\partial y \neq 0$ in $\operatorname{supp} a$. It is therefore sufficient to prove (5.3.22) when $\operatorname{supp} a$ is contained in a small neighborhood of a point (x^0, y^0) where $\partial^2\varphi(x,y)/\partial x\partial y$ has rank $n-1$ and (5.3.21) is applicable. Choose $t \in \mathbf{R}^n \setminus \{0\}$ so that $(\partial/\partial y)\langle t, \partial\varphi/\partial x\rangle = 0$ at (x^0, y^0). Since the rank of $(\partial^2/\partial y^2)\langle t, \partial\varphi/\partial x\rangle$ is $\geq n-1$ at (x^0, y^0) by (5.3.21), we can make an affine change of y coordinates preserving y^0 such that in the new coordinates the rank of $(\partial^2\varphi(x,y)/\partial x_j\partial y_k)_{j=1,\dots,n}^{k=1,\dots,n-1}$ is equal to $n-1$ and $\det(\partial^2\langle t, \partial\varphi/\partial x\rangle/\partial y_j\partial y_k)_{j,k=1}^{n-1} \neq 0$ at (x^0, y^0). In fact, both conditions are fulfilled for a generic direction of the coordinate plane where $y_n = y_n^0$. Then Theorem 5.3.2 can be applied for fixed y_n to

$$\mathscr{T}_N u(x, y_n) = \int e^{iN\varphi(x,y)} a(x,y) u(y)\, dy_1 \dots dy_{n-1}.$$

Since $T_N u(x) = \int \mathscr{T}_N u(x, y_n)\, dy_n$ we obtain

$$\|T_N u\|_q \leq \int_K \|\mathscr{T}_N u(\cdot, y_n)\|_q\, dy_n \leq C N^{-n/q} \int_K \|u(\cdot, y_n)\|_{L^r(\mathbf{R}^{n-1})}\, dy_n,$$

where K is a compact set $\subset \mathbf{R}$ such that $a(x,y) = 0$ when $y_n \notin K$. By Hölder's inequality the right-hand side is $\leq m(K)^{1/r'}\|u\|_r$, which completes the proof.

Corollary 5.3.6. *Let $\Phi \in C^\infty(\mathbf{R}^n \setminus \{0\})$ be real valued and positively homogeneous of degree 1, $n \geq 3$, and let $A \in C_0^\infty(\mathbf{R}^n \setminus \{0\})$. Set*

$$S_t f(x) = \int e^{it\Phi(x-y)} A(x-y) f(y)\, dy, \quad f \in L^1_{\text{loc}}(\mathbf{R}^n).$$

If $\Phi''(x)$ has rank $n-1$ for every $x \in \operatorname{supp} A$ and $n \geq 3$, it follows that

$$\|S_t f\|_p \leq C_p(t)\|f\|_p, \quad f \in L^p(\mathbf{R}^n),\ t > 2,\ p \geq 2, \quad \textit{where} \tag{5.3.23}$$

$$C_p(t) = \begin{cases} Ct^{-n/p}, & \textit{if } p \geq \frac{2n+2}{n-1}, \\ Ct^{-(n-1)(\frac{1}{4}+\frac{1}{2p})}, & \textit{if } 2 \leq p \leq \frac{2n+2}{n-1}. \end{cases} \tag{5.3.24}$$

Proof. The phase function $\varphi(x,y) = \Phi(x-y)$ satisfies the hypotheses of Theorem 5.3.5 when $x - y \in \operatorname{supp} A$. In fact, $\Phi''(z)z = 0$ since Φ' is homogeneous of degree 0, and since $\Phi''(z)$ is of rank $n-1$ it follows that $\Phi''(z)t = 0$ implies that $t = cz$, hence $\Phi'''(z)t = -c\Phi''(z)$ since Φ'' is homogeneous of degree -1. This is of rank $n-1$ when $c \neq 0$.

Let $0 \leq \chi \in C_0^\infty(\mathbf{R}^n)$, $\int \chi^2\, dx = 1$. Then the hypotheses of Theorem 5.3.5 are fulfilled if $a(x,y) = A(x-y)\chi(y)$. If $p \geq (2n+2)/(n-1)$ it follows that

$$\|S_t \chi^2 f\|_p \leq C_p(t)\|\chi f\|_p, \quad f \in L^p(\mathbf{R}^n), \tag{5.3.25}$$

for $2n/(p(n-1)) \leq n/(n+1) < 1$. Theorem 5.2.1′ gives the estimate (5.3.25) when $p = 2$, and then it follows from the Riesz-Thorin interpolation theorem (Theorem 2.3.2) for $2 \leq p \leq (2n+2)/(n-1)$.

By the translation invariance of S_t it follows that for $z \in \mathbf{R}^n$

$$\|S_{t,z}f\|_p \le C_p(t)\|\chi(\cdot - z)f\|_p, \quad \text{where } S_{t,z}f = S_t(\chi(\cdot - z)^2 f). \tag{5.3.26}$$

Since

$$S_{t,z}f(x) = \int e^{it\Phi(x-y)}a(x-y)\chi(y-z)^2 f(y)\,dy,$$

the integral with respect to z is equal to $S_tf(x)$, and since $a(x-y)\chi(y-z) \ne 0$ implies a bound for $x-z$, we have by Hölder's inequality

$$|S_tf(x)|^p \le C^p \int |S_{t,z}f(x)|^p\,dz.$$

If we raise the estimate (5.3.26) to the power p and integrate with respect to z it follows that

$$\|S_tf\|_p \le CC_p(t)\|\chi\|_p\|f\|_p$$

which completes the proof.

Since Φ is homogeneous we have

$$(S_tf)(x/t) = \int e^{i\Phi(x-ty)}A(x/t-y)f(y)\,dy = \int e^{i\Phi(x-y)}A((x-y)/t)g(y)\,dy \tag{5.3.27}$$

where $g(y) = f(y/t)/t^n$, and (5.3.23) gives

$$\|(S_tf)(\cdot/t)\|_p = t^{n/p}\|S_tf\|_p \le t^{n/p}C_p(t)\|f\|_p = t^nC_p(t)\|g\|_p$$

If we multiply (5.3.27) by $t^{-1-\lambda}$ and integrate with respect to t from 2 to ∞, it follows that the operator

$$K_\lambda g(x) = \int e^{i\Phi(x-y)}b_\lambda(x-y)f(y)\,dy, \tag{5.3.28}$$

$$b_\lambda(z) = \int_2^\infty A(z/t)t^{-1-\lambda}dt, \tag{5.3.29}$$

is bounded in $L^p(\mathbf{R}^n)$ provided that $p \ge 2$ and that

$$\int_2^\infty C_p(t)t^{n-1-\lambda}\,dt < \infty. \tag{5.3.30}$$

By (5.3.24) this condition is equivalent to

$$\lambda > \begin{cases} n(1-\frac{1}{p}), & \text{if } p \ge \frac{2n+2}{n-1} \\ \frac{3n+1}{4} - \frac{n-1}{2p}, & \text{if } 2 \le p \le \frac{2n+2}{n-1}. \end{cases}$$

The function b defined by (5.3.29) is positively homogeneous of degree $-\lambda$ when $|z|$ is so large that $z/t \in \operatorname{supp} A$ implies $t \ge 2$, for then we have $b(z) = b_0(z)$ where

$$b_0(z) = \int_0^\infty A(z/t)t^{-1-\lambda}\,dt. \tag{5.3.31}$$

For every b_0 which is positively homogeneous of degree $-\lambda$ we have (5.3.31) if $A(z) = b_0(z)\varrho(z)$ where $\varrho \in C_0^\infty(\mathbf{R}^n \setminus \{0\})$ is chosen so that $\int_0^\infty \varrho(z/t)\,dt/t = 1$ for $z \ne 0$. This is true for any non-negative C^∞ function of $|z|$ with support in $(1,2)$ after multiplication by a suitable normalizing constant. This gives the following theorem when $p \ge 2$, and when $p \le 2$ it follows by duality.

THEOREM 5.3.7. *Let $\Phi \in C^\infty(\mathbf{R}^n \setminus \{0\})$ be real valued and positively homogeneous of degree 1, $n \geq 3$, and let $a_0 \in C^\infty(\mathbf{R}^n \setminus \{0\})$ be positively homogeneous of degree $-\lambda$. If $\Phi''(x)$ has rank $n-1$ when $a_0(x) \neq 0$ and*

$$\lambda > \max(\tfrac{n-1}{2}\left|\tfrac{1}{p} - \tfrac{1}{2}\right| + \tfrac{n+1}{2}, n\left|\tfrac{1}{p} - \tfrac{1}{2}\right| + \tfrac{n}{2}), \tag{5.3.32}$$

and $a \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$ is equal to a_0 outside a compact set, then the operator $f \mapsto (e^{i\Phi}a) * f$ extends from $L^1(\mathbf{R}^2) \cap L^\infty(\mathbf{R}^n)$ to a continuous operator in $L^p(\mathbf{R}^n)$.

In fact, we have just proved the continuity of this operator for some $b \in C^\infty$ equal to a_0 outside a compact set. Hence $b - a \in L^1$, and $f \mapsto (b-a) * f$ is continuous in L^p for every p.

Theorem 5.3.7 gives the sufficiency of the necessary conditions in Theorem 5.1.8 when $p \geq (2n+2)/(n-1)$ but a weaker result otherwise:

THEOREM 5.3.8. *Let $\varrho \in C^\infty(\mathbf{R}^n)$ be real valued and negative outside a compact set, $n \geq 3$, and assume that $\varrho'(\xi) \neq 0$ and that $\varrho''(\xi)t \neq 0$ if $\xi \in \mathbf{R}^n$, $\varrho(\xi) = 0$ and $0 \neq t \in \mathbf{R}^n$, $\varrho'(\xi)t = 0$. Then $\varrho_+^\alpha \in \mathscr{M}_p(\mathbf{R}^n)$ if*

$$\alpha > \max\left(\tfrac{n-1}{2}\left|\tfrac{1}{p} - \tfrac{1}{2}\right|, n\left|\tfrac{1}{p} - \tfrac{1}{2}\right| - \tfrac{1}{2}\right). \tag{5.3.33}$$

PROOF. The zeros of ϱ form a compact set, so a partition of unity shows that it is sufficient to prove that for every ξ^0 with $\varrho(\xi^0) \neq 0$ there is some $\varphi \in C_0^\infty$ with $\varphi(\xi^0) \neq 0$ such that $\varphi\varrho_+^\alpha \in \mathscr{M}_p(\mathbf{R}^n)$. If we choose φ as in the proof of Theorem 5.1.8, keeping the notation there, we find that it is sufficient to prove L^p continuity of convolution with a C^∞ kernel K such that outside a compact set $K(x) - e^{i\Phi(x)}a_0(x) = O(|x|^{-1-\lambda})$ where a_0 is homogeneous of degree $-\lambda$ and $\lambda = (n+1)/2 + \alpha$. Here Φ is defined in $\operatorname{supp} a_0$ by $\Phi(x) = \langle x', t\rangle + x_n\psi(t)$ where t is a C^∞ function of x, homogeneous of degree 0, determined by the equation $x' + x_n\psi'(t) = 0$; $\det \psi''(t) \neq 0$. Thus $\Phi'(x) = (t, \psi'(t))$ and $\partial^2\Phi(x)/\partial x'\partial x' = \partial t/\partial x' = -(x_n\psi''(t))^{-1}$ is non-singular. Since (5.3.33) is identical to (5.3.32), the theorem is proved.

NOTES

Chapter I. The discussion of general finite commutative groups in Sections 1.1 and 1.2 is only intended to give the algebraic side of the motivation for Fourier analysis. Apart from the simple explicit formulas (1.2.4) for $\mathbf{Z}_n$ it can be bypassed with no loss of continuity. Alternatively one can find further results in a textbook on algebra such as Lang [1]. The discussion of the fast Fourier transform follows Auslander and Tolmieri [1] to a large extent. In this reference one can also find an interesting discussion of the eigenvalues of the finite Fourier transform with applications to the quadratic reciprocity theorem. See also Strang [1] for a discussion of the virtues of the fast Fourier transform in applications.

Chapter II. Most of the material in Section 2.1 is by now so classical that we shall only give references to the origin of two of them. The Bernstein theorem (Th. 2.1.8) has been treated here following Achieser [1]. The theorem of supports (Th.2.1.11) was first proved by Titchmarsh [1].

The Hausdorff-Young theorem (Th. 2.3.1) was actually proved by these authors for Fourier series while the extension to Fourier integrals is due to Titchmarsh. The proof given here is that of M. Riesz [2] who proved a somewhat restricted version of Theorem 2.3.2 using real variable methods. The proof given here is due to Thorin [1]. The classical background of Theorem 2.3.8 is another theorem of Bernstein stating that the Fourier series of a function which is Hölder continuous of order $> \frac{1}{2}$ is absolutely convergent. We shall not try to trace the deep historical roots of the method of stationary phase. Instead we would like to refer to Hörmander [1, Section 7.7] for a much more extensive study.

Chapter III. We have here followed Daubechies [1] to a very large extent, first in discussing multiresolution analyses with scale functions which are only in L^2. (Proposition 3.1.4 rounds off her results which only give sufficient conditions at that point.) The construction of wavelets in several dimensions in Section 3.2 is mainly taken from Meyer [2] though. Both Daubechies [1], Meyer [1], [2] and Meyer-Coifman [1] should be consulted for further results on wavelets and their applications. Only the mathematical framework is presented here.

Chapter IV. The estimate of the conjugate function (Th. 4.1.1) is due to M. Riesz [1], but the proof we have chosen is due to P. Stein [1]. In the n-dimensional case studied in Section 4.2 we follow the methods of Calderón and Zygmund [1]. The Hardy-Littlewood maximal theorem is due to Hardy and Littlewood [1] but the proof of Theorem 4.1.2 is due to F. Riesz [1]. In the n-dimensional case in Section 4.2 we have instead used covering theorems which can be found for example in Aronszajn and Smith [1]. The potential estimate in the following example comes from Hardy and Littlewood [2]. The refined maximal theorem of Carleson in Theorem 4.1.2′ originates from Carleson [1], [2]. A simplification and generalisation given in Hörmander [2] is used in Section 4.2 here and has been adapted to the method of F. Riesz in the one-dimensional case. The proof of Theorem 4.1.3 comes from Calderón and Zygmund [1]. Proposition 4.1.7 has mainly been taken from Stein [4]. The discussion of the Hardy space $\mathscr{H}^1$ and the duality with BMO in Sections 4.1 and 4.2 follows Fefferman and Stein [1]; see also Stein [2]. In both these references there is also a discussion of the Hardy space $\mathscr{H}^p$ with $0 < p < 1$. The Mihlin theorem giving the L^p analogue of Corollary 4.2.18 is due to Mihlin [1] under somewhat more restrictive hypotheses and Hörmander [3] in a somewhat stronger form. Actually the result goes back to Marcinkiewicz [1]. His interpolation theorem first appeared in a somewhat special form in Marcinkiewicz [2]. The general statement, containing that given here, was published much later by Zygmund [1]. For examples of applications of the Hardy space in the theory of non-linear differential equations one can consult Coifman, Lions, Meyer and Semmes [1].

The proof in Section 4.3 that compactly supported wavelets give bases in L^p and in $\mathscr{H}^1$ follows Daubechies [1] and Meyer [1,2]. They use weaker hypotheses which make the proofs technically harder but the main points are the same as here.

The John-Nirenberg theorem (Th. 4.4.1) is the simplest of a number of related results proved by John and Nirenberg [1]. Spanne [1] has proved the interpolation Theorem 4.4.7 using the more refined results from that paper. (At that time the duality between $\mathscr{H}^1$ and BMO was not known.) Here we have instead used the properties of the function $f^\sharp$ due to Fefferman and Stein [1].

Chapter V. The basic facts on multipliers belong to the folklore which is hard to trace back. The striking Theorem 5.1.6 is due to Fefferman [1], while the necessary condition in Theorem 5.1.8 undoubtedly has many discoverers, for it is an immediate consequence of the stationary phase theorem. The important results in Section 5.2 are due to Carleson and Sjölin [1]. The proof given here is a simplification due to Hörmander [4], where Theorem 5.2.1 has been taken. Questions concerning analogues of the crucial Theorem 5.2.2 for higher dimensions were also raised there, but the example (5.3.10) due to Bourgain [1] proved that further conditions than anticipated in Hörmander [4] will be required then. (See Bourgain [1], [2] for further improvements of the results in Section 5.3.) Theorem 5.3.2 is due to Stein [5]; a weaker form of the restriction theorem (Th. 5.3.4) was proved before by Thomas [1]. The proof of Theorem 5.3.2 here follows Sogge [1] which is highly recommended for further study of the topics in Chapter V. Estimates of the type discussed in Section 5.3 are essential for the study of low regularity to non-linear hyperbolic differential equations.

References

N. I. Achieser, *Vorlesungen über Approximationstheorie*, Akademie-Verlag, Berlin, 1953.

N. Aronszajn and K. T. Smith, *Functional spaces and functional completion*, Ann. Inst. Fourier Grenoble **6** (1955–56), 125–185.

l. Auslander and R. Tolmieri [1], *Is computing with the finite Fourier transform pure or applied mathematics?*, Bull. Amer. Math. Soc. **1** (1979), 847–897.

W. Beckner [1], *Inequalities in Fourier analysis*, Ann. Math. **102** (1975), 159–182.

J. Bourgain [1], *L^p estimates for oscillatory integrals in several variables*, Geom. and Funct. Analysis **1** (1991), 321–374.

J. Bourgain [2], *Besicovitch type maximal operators and applications to Fourier analysis*, Geom. and Funct. Analysis **1** (1991), 147–187.

A. P. Calderon and A. Zygmund [1], *On the existence of certain singular integrals*, Acta Math. **88** (1952), 85–139.

L. Carleson [1], *An interpolation problem for bounded analytic functions*, Amer. J. Math. **80** (1958), 921–930.

L. Carleson [2], *Interpolation by bounded analytic functions and the corona problem*, Ann. of Math. **76** (1962), 547–559.

L. Carleson and P. Sjölin [1], *Oscillatory integrals and a multiplier problem for the disc*, Studia Math. **44** (1972), 287–299.

R. Coifman, P. L. Lions, Y. Meyer and S. Semmes, *Compensated compactness and Hardy spaces*, J. Math. Pures Appl. **72** (1993), 247–286.

I. Daubechies [1], *Ten lectures on wavelets*, SIAM, Philadelphia, PA, 1992, CBMS-NSF Regional conference series in applied mathematics.

M. Essén [1], *A superharmonic proof of the M. Riesz conjugate function theorem*, Ark. Mat. **22** (1984), 241–249.

C. Fefferman [1], *The multiplier problem for the ball*, Ann. of Math. **94** (1971), 330–336.

C. Fefferman and E. M. Stein [1], *H^p spaces of several variables*, Acta Math. **129** (1972), 137–193.

A. Haar [1], *Zur Theorie der orthogonalen Funktionensysteme*, Math. Ann. **69** (1910), 331–371.

G. H. Hardy and J. E. Littlewood [1], *A maximal theorem with function-theoretic applications*, Acta Math. **54** (1930), 81–116.

G. H. Hardy and J. E. Littlewood [2], *Some properties of fractional integrals (I)*, Math. Z. **27** (1928), 565–606; (II), Math. Z. **34** (1931–32), 403–439.

L.-I. Hedberg, *On certain convolution inequalities*, Proc. Amer. Math. Soc. **36** (1972), 505–510.

L. Hörmander [1], *The analysis of linear partial differential operators I*, Springer Verlag, 1983.

L. Hörmander [2], *L^p estimates for (pluri-)subharmonic functions*, Math. Scand. **20** (1967), 65–78.

L. Hörmander [3], *Estimates for translation invariant operators in L^p spaces*, Acta Math. **104** (1960), 93–140.

L. Hörmander [4], *Oscillatory integrals and multipliers on FL^p*, Ark. för Matematik **11** (1973), 1–11.

F. John and L. Nirenberg [1], *On functions of bounded mean oscillation*, Comm. Pure Appl. Math. **14** (1961), 391–413.

S. Lang [1], *Algebra*, Addison-Wesley, 1965.

E. H. Lieb [1], *Gaussian kernels have only Gaussian maximizers*, Inv. Math. **102** (1990), 179–208.

J. Marcinkiewicz [1], *Sur les multiplicateurs des séries de Fourier*, Studia Math. **8** (1939), 78–91.

J. Marcinkiewicz [2], *Sur l'interpolation d'opérations*, C. R. Acad. Sci. Paris **208** (1939), 1272–1273.

Y. Meyer [1], *Wavelets, Algorithms and applications*, SIAM, Philadelphia, PA, 1993.

Y. Meyer [2], *Ondelettes et opérateurs I, II*, Hermann, Paris, 1990.

Y. Meyer and R. R. Coifman [1], *Ondelettes et opérateurs III*, Hermann, Paris, 1991.

S. G. Mihlin [1], *On the multipliers of Fourier integrals*, Dokl. Akad. Nauk SSSR (N.S.) **109** (1956), 701–703. (Russian)

F. Riesz [1], *Sur l'existence de la dérivée des functions d'une variable réelle et des functions d'intervalle*, Verh. Int. Kongr. Zürich I, 1932, pp. 258–269.

M. Riesz [1], *Sur les fonctions conjuguées*, Math. Z. **27** (1927), 218–244.

M. Riesz [2], *Sur les maxima des formes bilinéaires et sur les functionelles linéaires*, Acta Math. **49** (1926), 465–497.

P. Sjölin [1], *Fourier multipliers and estimates of the Fourier transform of measures carried by smooth curves in* $\mathbf{R}^2$, Studia Math. **51** (1974), 169–182.

C. Sogge [1], *Fourier integrals in classical analysis*, Cambridge University Press, 1993.

S. Spanne [1], *Sur l'interpolation entre les espaces* $\mathscr{L}_k^{p\Phi}$, Ann. Scuola Norm. Sup. Pisa **20(3)** (1966), 625–648.

E. M. Stein [1], *Singular integrals and differentiability properties of functions*, Princeton University Press, Princeton, N.J., 1970.

E. M. Stein [2], *Harmonic analysis*, Princeton University Press, Princeton. N.J., 1993.

E. M. Stein [3], *Note on the class* $L \log L$, Studia Math. **32** (1969), 305–310.

E. M. Stein [4], *Localization and summability of multiple Fourier series*, Acta Math. **100** (1958), 93–147.

E. M. Stein [5], *Oscillatory integrals in Fourier analysis*, Beijing lectures in harmonic analysis, Princeton University Press, Princeton, N.J., 1971, pp. 307–356.

E. M. Stein and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton University Press, Princeton, N.J., 1971.

P. Stein [1], *On a theorem of M. Riesz*, J. London Math. Soc. **8** (1933), 242–247.

G. Strang [1], *Wavelet transfosrms versus Fourier transforms*, Bull. Amer. Math. Soc. **28** (1993), 288–305.

O. Thorin [1], *An extension of a convexity theorem due to M. Riesz*, Kungl. Fys. Sällsk. Förh. **8:14** (1938), 1–5.

P. Tomas [1], *Restriction theorems for the Fourier transform*, Proc. Symp. Pure Math. **35** (1979), 111–114.

E. C. Titchmarsh, *The zeros of certain integral functions*, Proc. London Math. Soc. **25** (1926), 283–302.

A. Zygmund [1], *On a theorem of Marcinkiewicz concerning interpolation of operators*, J. Math. Pures Appl. **35** (1956), 223–248.

Index of notation

General notation

L^p	space of measurable functions with integrable pth power
$\\| \\|_p$	the norm $\\| \\|_{L^p}$
$C^k(X)$	functions in X with continuous derivatives of order $\leq k$
$C_0^k(X)$	functions in $C^k(X)$ with compact support
$\mathscr{D}'(X)$	Schwartz distributions in X
$\mathscr{E}'(X)$	Schwartz distributions in X with compact support
$\mathscr{S}$	Schwartz space of rapidly decreasing C^∞ functions
$\mathscr{S}'$	temperate distributions
$f * g$	convolution of f and g
$\operatorname{supp} u$	support of u
$\operatorname{sing\,supp} u$	singular support of u
$X \Subset Y$	closure of X is a compact subset of Y
$\complement X$	complement of X (in some larger set)
∂X	boundary of X
α	usually a multiindex $\alpha = (\alpha_1, \dots, \alpha_n)$
$\lvert\alpha\rvert$	length $\alpha_1 + \cdots + \alpha_n$ of α
$\alpha!$	multifactorial $\alpha_1! \dots \alpha_n!$
x^α	monomial $x_1^{\alpha_1} \dots x_n^{\alpha_n}$ in $\mathbf{R}^n$
∂^α	partial derivative, $\partial_j = \partial/\partial x_j$
D^α	partial derivative, $D_j = -i\partial/\partial x_j$
$\hat{f}$	Fourier(-Laplace) transform of f
$\mathbf{Z}_\nu$	$\mathbf{Z}/\nu\mathbf{Z}$ where $\mathbf{Z}$ denotes the integers
$H_{(s)}$	Sobolev space of order s
$\\| \\|_{(s)}$	norm in $H_{(s)}$

Section 1.2

$\widehat{G}$	dual group

Section 4.1

$\tilde{f}$	conjugate function of f
f_{HL}^*	Hardy-Littlewood maximal function (4.1.6)
f_{HL}^{**}	the refined maximal function (4.1.6)$''$
f_{CZ}^*	Calderón-Zygmund maximal function (4.1.9)
$\mathscr{H}^1(\mathbf{R})$	Hardy space in $\mathbf{R}$
$\mathrm{BMO}(\mathbf{R})$	functions of bounded mean oscillation in $\mathbf{R}$

Section 4.2

f_{HL}^*	Hardy-Littlewood maximal function (4.2.16)
f_{HL}^{**}	the refined maximal function (4.2.16)$''$
f_{M}^*	maximal function for a singular integral operator with kernel M
R_j	Riesz kernels (4.2.21)
$\mathscr{H}^1(\mathbf{R}^n)$	Hardy space in $\mathbf{R}^n$
$\mathrm{BMO}(\mathbf{R}^n)$	functions of bounded mean oscillation in $\mathbf{R}^n$

P_0	Poisson kernel (4.2.22)
P_j	conjugates (4.2.23) of Poisson kernel

Section 4.4

$f^\sharp$	maximal function defined by (4.4.14)

Section 5.1

$\mathscr{M}_p$	multipliers on the Fourier transform of L^p

Index