LECTURES IN HARMONIC ANALYSIS

Thomas H. Wolff

Revised version, March 2002

1. L^1 Fourier transform

If $f \in L^1(\mathbb{R}^n)$ then its Fourier transform is $\hat{f} : \mathbb{R}^n \to \mathbb{C}$ defined by

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

More generally, let $M(\mathbb{R}^n)$ be the space of finite complex-valued measures on $\mathbb{R}^n$ with the norm

$$\|\mu\| = |\mu|(\mathbb{R}^n),$$

where $|\mu|$ is the total variation. Thus $L^1(\mathbb{R}^n)$ is contained in $M(\mathbb{R}^n)$ via the identification $f \to \mu$, $d\mu = f dx$. We can generalize the definition of Fourier transform via

$$\hat{\mu}(\xi) = \int e^{-2\pi i x\cdot\xi} d\mu(x).$$

Example 1 Let $a \in \mathbb{R}^n$ and let δ_a be the Dirac measure at a, $\delta_a(E) = 1$ if $a \in E$ and $\delta_a(E) = 0$ if $a \notin E$. Then $\widehat{\delta_a}(\xi) = e^{-2\pi i a\cdot\xi}$.

Example 2 Let $\Gamma(x) = e^{-\pi|x|^2}$. Then

$$\hat{\Gamma}(\xi) = e^{-\pi|\xi|^2}. \tag{1}$$

Proof The integral in question is

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

Notice that this factors as a product of one variable integrals. So it suffices to prove (1) when $n = 1$. For this we use the formula for the integral of a Gaussian: $\int_{-\infty}^{\infty} e^{-\pi x^2} dx = 1$. It follows that

$$\begin{aligned}
\int_{-\infty}^{\infty} e^{-2\pi i x\xi} e^{-\pi x^2} dx &= \int_{-\infty}^{\infty} e^{-\pi(x-i\xi)^2} dx \cdot e^{-\pi\xi^2} \\
&= \int_{-\infty-i\xi}^{\infty-i\xi} e^{-\pi x^2} dx \cdot e^{-\pi\xi^2} \\
&= \int_{-\infty}^{\infty} e^{-\pi x^2} dx \cdot e^{-\pi\xi^2} \\
&= e^{-\pi\xi^2},
\end{aligned}$$

where we used contour integration at the next to last line. $\square$

There are some basic estimates for the L^1 Fourier transform, which we state as Propositions 1 and 2 below. Consideration of Example 1 above shows that in complete generality not that much more can be said.

Proposition 1.1 If $\mu \in M(\mathbb{R}^n)$ then $\hat{\mu}$ is a bounded function, indeed

$$\|\hat{\mu}\|_\infty \leq \|\mu\|_{M(\mathbb{R}^n)}. \tag{2}$$

Proof For any ξ,

$$\begin{aligned} |\hat{\mu}(\xi)| &= |\int e^{-2\pi i x\cdot\xi} d\mu(x)| \\ &\leq \int |e^{-2\pi i x\cdot\xi}|\, d|\mu|(x) \\ &= \|\mu\|. \end{aligned}$$

$\square$

Proposition 1.2 If $\mu \in M(\mathbb{R}^n)$ then $\hat{\mu}$ is a continuous function.

Proof Fix ξ and consider

$$\hat{\mu}(\xi + h) = \int e^{-2\pi i x\cdot(\xi+h)} d\mu(x).$$

As $h \to 0$ the integrands converge pointwise to $e^{-2\pi i x\cdot\xi}$. Since all the integrands have absolute value 1 and $|\mu|(\mathbb{R}^n) < \infty$, the result follows from the dominated convergence theorem. $\square$

We now list some basic formulas for the Fourier transform; the ones listed here are roughly speaking those that do not involve any differentiations. They can all be proved by using the formula $e^{a+b} = e^a e^b$ and appropriate changes of variables. Let $f \in L^1$, $\tau \in \mathbb{R}^n$, and let T be an invertible linear map from $\mathbb{R}^n$ to $\mathbb{R}^n$.

1. Let $f_\tau(x) = f(x - \tau)$. Then

$$\widehat{f_\tau}(\xi) = e^{-2\pi i \tau\cdot\xi} \hat{f}(\xi). \tag{3}$$

2. Let $e_\tau(x) = e^{2\pi i x\cdot\tau}$. Then

$$\widehat{e_\tau f}(\xi) = \hat{f}(\xi - \tau). \tag{4}$$

3. Let T^{-t} be the inverse transpose of T. Then

$$\widehat{f \circ T} = |\det(T)|^{-1} \hat{f} \circ T^{-t}. \tag{5}$$

4. Define $\tilde{f}(x) = \overline{f(-x)}$. Then

$$\hat{\tilde{f}} = \overline{\hat{f}}. \tag{6}$$

We note some special cases of 3. If T is an orthogonal transformation (i.e. TT^t is the identity map) then $\widehat{f \circ T} = \hat{f} \circ T$, since $\det(T) = \pm 1$. In particular, this implies that if f is radial then so is $\hat{f}$, since orthogonal transformations act transitively on spheres. If T is a dilation, i.e. $Tx = r \cdot x$ for some $r > 0$, then 3. says that the Fourier transform of the function $f(rx)$ is the function $r^{-n}\hat{f}(r^{-1}\xi)$. Replacing r with r^{-1} and multiplying through by r^{-n}, we see that the reverse formula also holds: the Fourier transform of the function $r^{-n}f(r^{-1}x)$ is the function $\hat{f}(r\xi)$.

There is a general principle that if f is localized in space, then $\hat{f}$ should be smooth, and conversely if f is smooth then $\hat{f}$ should be localized. We now discuss some simple manifestations of this. Let $D(x, r) = \{y \in \mathbb{R}^n : |y - x| < r\}$.

<u>Proposition 1.3</u> Suppose that $\mu \in M(\mathbb{R}^n)$ and $\operatorname{supp} \mu$ is compact. Then $\hat{\mu}$ is C^∞ and

$$D^\alpha \hat{\mu} = ((-2\pi i x)^\alpha \mu)\widehat{\ }. \tag{7}$$

Further, if $\operatorname{supp} \mu \subset D(0, R)$ then

$$\|D^\alpha \hat{\mu}\|_\infty \leq (2\pi R)^{|\alpha|} \|\mu\|. \tag{8}$$

We are using multiindex notation here and will do so below as well. Namely, a *multiindex* is a vector $\alpha \in \mathbb{R}^n$ whose components are nonnegative integers. If α is a multiindex then by definition

$$D^\alpha = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}},$$

$$x^\alpha = \Pi_{j=1}^n x_j^{\alpha_j}.$$

The *length* of α, denoted $|\alpha|$, is $\sum_j \alpha_j$. One defines a partial order on multiindices via

$$\alpha \leq \beta \Leftrightarrow \alpha_i \leq \beta_i \text{ for each } i,$$

$$\alpha < \beta \Leftrightarrow \alpha \leq \beta \text{ and } \alpha \neq \beta.$$

<u>Proof of Proposition 1.3</u> Notice that (8) follows from (7) and Proposition 1 since the norm of the measure $(2\pi i x)^\alpha \mu$ is $\leq (2\pi R)^{|\alpha|} \|\mu\|$.

Furthermore, for any α the measure $(2\pi i x)^\alpha \mu$ is again a finite measure with compact support. Accordingly, if we can prove that $\hat{\mu}$ is C^1 and that (7) holds when $|\alpha| = 1$, then the lemma will follow by a straightforward induction.

Fix then a value $j \in \{1, \ldots, n\}$, and let e_j be the jth standard basis vector. Also fix $\xi \in \mathbb{R}^n$, and consider the difference quotient

$$\Delta(h) = \frac{\hat{\mu}(\xi + he_j) - \hat{\mu}(\xi)}{h}. \tag{9}$$

This is equal to

$$\int \frac{e^{-2\pi i h x_j} - 1}{h} e^{-2\pi i \xi \cdot x} d\mu(x). \tag{10}$$

As $h \to 0$, the quantity

$$\frac{e^{-2\pi i h x_j} - 1}{h}$$

converges pointwise to $-2\pi i x_j$. Furthermore, $|\frac{e^{-2\pi i h x_j} - 1}{h}| \leq 2\pi |x_j|$ for each h. Accordingly, the integrands in (10) are dominated by $|2\pi x_j|$, which is a bounded function on the support of μ. It follows by the dominated convergence theorem that

$$\lim_{h \to 0} \Delta(h) = \int \lim_{h \to 0} \frac{e^{-2\pi i h x_j} - 1}{h} e^{-2\pi i \xi \cdot x} d\mu(x),$$

which is equal to

$$\int -2\pi i x_j e^{-2\pi i \xi \cdot x} d\mu(x).$$

This proves the formula (7) when $|\alpha| = 1$. Formula (7) and Proposition 2 imply that $\hat{\mu}$ is C^1. $\square$

Remark The estimate (8) is tied to the support of μ. However, the fact that $\hat{\mu}$ is C^∞ and the formula (7) are still valid whenever μ has enough decay to justify the differentiations under the integral sign. For example, they are valid if μ has moments of all orders, i.e. $\int |x|^N d|\mu|(x) < \infty$ for all N.

The estimate (2) can be seen as justification of the idea that if μ is localized then $\hat{\mu}$ should be smooth. We now consider the converse statement, μ smooth implies $\hat{\mu}$ localized.

Proposition 1.4 Suppose that f is C^N and that $D^\alpha f \in L^1$ for all α with $0 \leq |\alpha| \leq N$. Then

$$\widehat{D^\alpha f}(\xi) = (2\pi i \xi)^\alpha \hat{f}(\xi) \tag{11}$$

when $|\alpha| \leq N$ and furthermore

$$|\hat{f}(\xi)| \leq C(1 + |\xi|)^{-N} \tag{12}$$

for a suitable constant C.

The proof is based on an integration by parts which is most easily justified when f has compact support. Accordingly, we include the following lemma before giving the proof.

Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a C^∞ function with the following properties (4. is actually irrelevant for present purposes):

1. $\phi(x) = 1$ if $|x| \leq 1$
2. $\phi(x) = 0$ if $|x| \geq 2$
3. $0 \leq \phi \leq 1$.
4. ϕ is radial.

Define $\phi_k(x) = \phi(\frac{x}{k})$; thus ϕ_k is similar to ϕ but lives on scale k instead of 1. If α is a multiindex, then there is a constant C_α such that $|D^\alpha \phi_k| \leq \frac{C_\alpha}{k^{|\alpha|}}$ uniformly in k. Furthermore, if $\alpha \neq 0$ then the support of $D^\alpha \phi$ is contained in the region $k \leq |x| \leq 2k$.

<u>Lemma 1.5</u> If f is C^N, $D^\alpha f \in L^1$ for all α with $|\alpha| \leq N$ and if we let $f_k = \phi_k f$ then $\lim_{k\to\infty} \|D^\alpha f_k - D^\alpha f\|_1 = 0$ for all α with $|\alpha| \leq N$.

<u>Proof</u> It is obvious that

$$\lim_{k\to\infty} \|\phi_k D^\alpha f - D^\alpha f\|_1 = 0,$$

so it suffices to show that

$$\lim_{k\to\infty} \|D^\alpha(\phi_k f) - \phi_k D^\alpha f\|_1 = 0. \tag{13}$$

However, by the Leibniz rule

$$D^\alpha(\phi_k f) - \phi_k D^\alpha f = \sum_{0<\beta\leq\alpha} c_\beta D^{\alpha-\beta} f D^\beta \phi_k,$$

where the c_β's are certain constants. Thus

$$\begin{aligned} \|D^\alpha(\phi_k f) - \phi_k D^\alpha f\|_1 &\leq C \sum_{0<\beta\leq\alpha} \|D^\beta \phi_k\|_\infty \|D^{\alpha-\beta} f\|_{L^1(\{x:|x|\geq k\})} \\ &\leq Ck^{-1} \sum_{0<\beta\leq\alpha} \|D^{\alpha-\beta} f\|_{L^1(\{x:|x|\geq k\})} \end{aligned}$$

The last line clearly goes to zero as $k \to \infty$. There are two reasons for this (either would suffice): the factor k^{-1}, and the fact that the L^1 norms are taken only over the region $|x| \geq k$. $\square$

Proof of Proposition 1.4 If f is C^1 with compact support, then by integration by parts we have

$$\int \frac{\partial f}{\partial x_j}(x)e^{-2\pi i x\cdot\xi}dx = 2\pi i\xi_j \int e^{-2\pi i x\cdot\xi} f(x)dx,$$

i.e. (11) holds when $|\alpha| = 1$. An easy induction then proves (11) for all α provided that f is C^N with compact support.

To remove the compact support assumption, let f_k be as in Lemma 1.5. Then (11) holds for f_k. Now we pass to the limit as $k \to \infty$. On the one hand $\widehat{D^\alpha f_k}$ converges uniformly to $\widehat{D^\alpha f}$ as $k \to \infty$ by Lemma 1.5 and Proposition 1.1. On the other hand $\widehat{f_k}$ converges uniformly to $\hat{f}$, so $(2\pi i\xi)^\alpha \widehat{f_k}$ converges to $(2\pi i\xi)^\alpha \hat{f}$ pointwise. This proves (11) in general.

To prove (12), observe that (11) and Proposition 1 imply that $\xi^\alpha \hat{f} \in L^\infty$ if $|\alpha| \le N$. On the other hand, it is easy to estimate

$$C_N^{-1}(1+|\xi|)^N \le \sum_{|\alpha|\le N} |\xi^\alpha| \le C_N(1+|\xi|)^N, \tag{14}$$

so (12) follows. $\square$

Together with (14), let us note the inequality

$$1+|x| \le (1+|y|)(1+|x-y|),\ x,y \in \mathbb{R}^n \tag{15}$$

which will be used several times below.

2. Schwartz space

The Schwartz space $\mathcal{S}$ is the space of functions $f : \mathbb{R}^n \to \mathbb{C}$ such that:

1. f is C^∞,
2. $x^\alpha D^\beta f$ is a bounded function for each pair of multiindices α and β.

For $f \in \mathcal{S}$ we define

$$\|f\|_{\alpha\beta} = \|x^\alpha D^\beta f\|_\infty.$$

It is possible to see that $\mathcal{S}$ with the family of norms $\|\cdot\|_{\alpha\beta}$ is a Frechet space, but we don't discuss such questions here (see [27]). However, we define a notion of sequential convergence in $\mathcal{S}$:

A sequence $\{f_k\} \subset \mathcal{S}$ *converges in* $\mathcal{S}$ to $f \in \mathcal{S}$ if $\lim_{k\to\infty} \|f_k - f\|_{\alpha\beta} = 0$ for each pair of multiindices α and β.

Examples 1. Let C_0^∞ be the C^∞ functions with compact support. Then $C_0^\infty \subset \mathcal{S}$.

Namely, to prove that $x^\alpha D^\beta f$ is bounded, just note that if $f \in C_0^\infty$ then $D^\beta f$ is a continuous function with compact support, hence bounded, and that x^α is a bounded function on the support of $D^\beta f$.

2. Let $f(x) = e^{-\pi|x|^2}$. Then $f \in \mathcal{S}$.

For the proof, notice that if $p(x)$ is a polynomial, then any first partial derivative $\frac{\partial}{\partial x_j}(p(x)e^{-\pi|x|^2})$ is again of the form $q(x)e^{-\pi|x|^2}$ for some polynomial q. It follows by induction that each $D^\beta f$ is a polynomial times f for each β. Hence $x^\alpha D^\beta f$ is a polynomial times f for each α and β. This implies using L'Hospital's rule that $x^\alpha D^\beta f$ is bounded for each α and β.

3. The following functions are not in $\mathcal{S}$: $f_N(x) = (1 + |x|^2)^{-N}$ for any given N, and $g(x) = e^{-\pi|x|^2}\sin(e^{\pi|x|^2})$. Roughly, although f_N decays rapidly at ∞, it does not decay rapidly enough, whereas g decays rapidly enough but its derivatives do not decay. Detailed verification is left to reader.

We now discuss some simple properties of $\mathcal{S}$, then some which are slightly less simple.

I. $\mathcal{S}$ is closed under differentiations and under multiplication by polynomials. Furthermore, these operations are continuous on $\mathcal{S}$ in the sense that they preserve sequential convergence. Also $f, g \in \mathcal{S}$ implies $fg \in \mathcal{S}$.

Proof Let $f \in \mathcal{S}$. If γ is a multiindex, then $x^\alpha D^\beta(D^\gamma)f = x^\alpha D^{\beta+\gamma}f$, which is bounded since $f \in \mathcal{S}$. So $D^\gamma f \in \mathcal{S}$.

By the Leibniz rule, $x^\alpha D^\beta(x^\gamma f)$ is a finite sum of terms each of which is a constant multiple of

$$x^\alpha D^\delta(x^\gamma)D^{\beta-\delta}f$$

for some $\delta \leq \beta$. Furthermore, $D^\delta(x^\gamma)$ is a constant multiple of $x^{\gamma-\delta}$ if $\delta \leq \gamma$, and otherwise is zero. Thus $x^\alpha D^\beta(x^\gamma f)$ is a linear combination of monomials times derivatives of f, and is therefore bounded. So $x^\gamma f \in \mathcal{S}$.

The continuity statements follow from the proofs of the closure statements; we will normally omit these arguments. As an indication of how they are done, let us show that if γ is a multiindex then $f \to D^\gamma f$ is continuous. Suppose that $f_n \to f$ in $\mathcal{S}$. Fix a pair of multiindices α and β. Applying the definition of convergence with the multiindices α and $\beta + \gamma$, we have

$$\lim_{n\to\infty} \|x^\alpha D^{\beta+\gamma}(f_n - f)\|_\infty = 0.$$

Equivalently,

$$\lim_{n\to\infty} \|x^\alpha D^\beta(D^\gamma f_n - D^\gamma f)\|_\infty = 0$$

which says that $D^\gamma f_n$ converges to $D^\gamma f$.

The last statement (that $\mathcal{S}$ is an algebra) follows readily from the product rule and the definitions. □

II. The following alternate definitions of $\mathcal{S}$ are often useful.

$$f \in \mathcal{S} \Leftrightarrow (1+|x|)^N D^\beta f \text{ is bounded for each } N \text{ and } \beta, \tag{16}$$

$$f \in \mathcal{S} \Leftrightarrow \lim_{x\to\infty} x^\alpha D^\beta f = 0 \text{ for each } \alpha \text{ and } \beta. \tag{17}$$

Indeed, (16) follows from the definition and (14). The backward implication in (17) is trivial, while the forward implication follows by applying the definition with α replaced by appropriate larger multiindices, e.g. $\alpha + e_j$ for arbitrary $j \in \{1, \ldots, n\}$.

Proposition 2.1 C_0^∞ is dense in $\mathcal{S}$, i.e. for any $f \in \mathcal{S}$ there is a sequence $\{f_k\} \subset C_0^\infty$ with $f_k \to f$ in $\mathcal{S}$.

Proof This is almost the same as the proof of Lemma 1.5. Namely, define ϕ_k as there and consider $f_k = \phi_k f$, which is evidently in C_0^∞. We must show that

$$x^\alpha D^\beta(\phi_k f) \to x^\alpha D^\beta f$$

uniformly as $k \to \infty$. For this, we estimate

$$\|x^\alpha D^\beta(\phi_k f) - x^\alpha D^\beta f\|_\infty \le \|\phi_k x^\alpha D^\beta f - x^\alpha D^\beta f\|_\infty + \|x^\alpha(D^\beta(\phi_k f) - \phi_k D^\beta f)\|_\infty.$$

The first term is bounded by $\sup_{|x|\ge k} |x^\alpha D^\beta f|$ and therefore goes to zero as $k \to \infty$ by (17). The second term is estimated using the Leibniz rule by

$$C \sum_{\gamma<\beta} \|x^\alpha D^\gamma f\|_\infty \|D^{\beta-\gamma}\phi_k\|_\infty. \tag{18}$$

Since $f \in \mathcal{S}$ and $\|D^{\beta-\gamma}\phi_k\| \le \frac{C}{k}$, the expression (18) goes to zero as $k \to \infty$. □

There is a stronger density statement which is sometimes needed. Define a C_0^∞ *tensor function* to be a function $f : \mathbb{R}^n \to \mathbb{C}$ of the form

$$f(x) = \prod_j \phi_j(x_j),$$

where each $\phi_j \in C_0^\infty(\mathbb{R})$.

Proposition 2.1′ Linear combinations of C_0^∞ tensor functions are dense in $\mathcal{S}$.

Proof In view of Proposition 2.1 it suffices to show that if $f \in C_0^\infty$ then there is a sequence $\{g_k\}$ such that:

1. Each g_k is a C_0^∞ tensor function.

2. The supports of the g_k are contained in a fixed compact set E which is independent of k.

3. $D^\alpha g_k$ converges uniformly to $D^\alpha f$ for each α.

To construct $\{g_k\}$, we use the fact (a basic fact about Fourier series) that if f is a C^∞ function in $\mathbb{R}^n$ which is 2π-periodic in each variable then f can be expanded in a series

$$f(\theta) = \sum_{\nu \in \mathbb{Z}^n} a_\nu e^{i\nu \cdot \theta}$$

where the $\{a_\nu\}$ satisfy

$$\sum_\nu (1 + |\nu|)^N |a_\nu| < \infty$$

for each N. Considering partial sums of the Fourier series, we therefore obtain a sequence of trigonometric polynomials p_k such that $D^\alpha p_k$ converges uniformly to $D^\alpha f$ for each α.

In constructing $\{g_k\}$ we can assume that $x \in \mathrm{supp} f$ implies $|x_j| \leq 1$, say, for each j - otherwise we work with $f(Rx)$ for suitable fixed R instead and undo the rescaling at the end. Let ϕ be a C_0^∞ function of one variable which is equal to 1 on $[-1, 1]$ and vanishes outside $[-2, 2]$. Let $\tilde{f}$ be the function which is equal to f on $[-\pi, \pi] \times \ldots \times [-\pi, \pi]$ and is 2π-periodic in each variable. Then we have a sequence of trigonometric polynomials p_k such that $D^\alpha p_k$ converges uniformly to $D^\alpha \tilde{f}$ for each α. Let $g_k(x) = \Pi_{j=1}^n \phi(x_j) \cdot p_k(x)$. Then g_k clearly satisfies 1. and 2, and an argument with the product rule as in Lemma 1.5 and Proposition 2.2 will show that $\{g_k\}$ satisfies 3. The proof is complete. $\square$

The next proposition is an alternate definition of $\mathcal{S}$ using L^1 instead of L^∞ norms.

Proposition 2.2 A C^∞ function f is in $\mathcal{S}$ iff the norms

$$\|x^\alpha D^\beta f\|_1$$

are finite for each α and β. Furthermore, a sequence $\{f_k\} \subset \mathcal{S}$ converges in $\mathcal{S}$ to $f \in \mathcal{S}$ iff

$$\lim_{k \to \infty} \|x^\alpha D^\beta (f_k - f)\|_1 = 0$$

for each α and β.

Proof We only prove the first part; the equivalence of the two notions of convergence follows from the proof and is left to the reader.

First suppose that $f \in \mathcal{S}$. Fix α and β. Let $N = |\alpha| + n + 1$. Then we know that $(1 + |x|)^N D^\beta f$ is bounded. Accordingly,

$$\begin{aligned} \|x^\alpha D^\beta f\|_1 &\leq \|(1 + |x|)^N D^\beta f\|_\infty \|x^\alpha (1 + |x|)^{-N}\|_1 \\ &< \infty \end{aligned}$$

using that the function $(1+|x|)^{-n-1}$ is integrable.

For the converse, we first make a definition and state a lemma. If $f:\mathbb{R}^n \to \infty$ is C^k and if $x \in \mathbb{R}^n$ then

$$\Delta_k^f(x) \stackrel{def}{=} \sum_{|\alpha|=k} |D^\alpha f(x)|.$$

We denote $D(x,r) = \{y : \ |x-y| \le r\}$. We also now start to use the notation $x \lesssim y$ to mean that $x \le Cy$ where C is a fixed but unspecified constant.

<u>Lemma 2.2'</u> Suppose f is a C^∞ function. Then for any x

$$|f(x)| \lesssim \sum_{0\le j\le n+1} \|\Delta_j^f\|_{L^1(D(x,1))}.$$

This is contained in Lemma A2 which is stated and proved at the end of the section.

To finish the proof of Proposition 2.2, we apply the preceding lemma to $D^\beta f$. This gives

$$|D^\beta f(x)| \lesssim \sum_{|\gamma|\le|\beta|+n+1} \int_{D(x,1)} |D^\gamma f(y)| dy,$$

therefore

$$\begin{aligned}(1+|x|)^N |D^\beta f(x)| &\lesssim (1+|x|)^N \sum_{|\gamma|\le|\beta|+n+1} \int_{D(x,1)} |D^\gamma f(y)| dy \\ &\lesssim \sum_{|\gamma|\le|\beta|+n+1} \int_{D(x,1)} (1+|y|)^N |D^\gamma f(y)| dy,\end{aligned}$$

where we used the elementary inequality

$$1+|x| \le 2 \min_{y\in D(x,1)} (1+|y|).$$

It follows that

$$\|(1+|x|)^N |D^\beta f\|_\infty \lesssim \sum_{|\gamma|\le|\beta|+n+1} \|(1+|x|)^N D^\gamma f\|_1,$$

and then Proposition 2.2 follows from (14). $\square$

<u>Theorem 2.3</u> If $f \in \mathcal{S}$ then $\hat{f} \in \mathcal{S}$. Furthermore, the map $f \to \hat{f}$ is continuous from $\mathcal{S}$ to $\mathcal{S}$.

<u>Proof</u> As usual we explicitly prove only the first statement.

If $f \in \mathcal{S}$ then $f \in L^1$, so $\hat{f}$ is bounded. Thus if $f \in \mathcal{S}$, then $\widehat{D^\alpha x^\beta f}$ is bounded for any given α and β, since $D^\alpha x^\beta f$ is again in $\mathcal{S}$. However, Propositions 1.3 and 1.4 imply that

$$\widehat{D^\alpha x^\beta f}(\xi) = (2\pi i)^{|\alpha|}(-2\pi i)^{-|\beta|} \xi^\alpha D^\beta \hat{f}(\xi)$$

so $\xi^\alpha D^\beta \hat{f}$ is again bounded, which means that $\hat{f} \in \mathcal{S}$. $\square$

Appendix: pointwise Poincare inequalities

This is a little more technical than the preceding and we will omit some details. We prove a frequently used pointwise estimate for a function in terms of integrals of its gradient, which plays a similar role that the mean value inequality plays in calculus. Then we prove a generalization involving higher derivatives which includes Lemma 2.3. Let ω be the volume of the unit ball.

Lemma A1 Suppose that f is C^1. Then

$$|f(x) - \frac{1}{\omega}\int_{D(x,1)} f(y)dy| \lesssim \int_{D(x,1)} \frac{|\nabla f(y)|}{|x-y|^{n-1}}dy.$$

Proof Applying the fundamental theorem of calculus to the function

$$t \to f(x + t(y-x))$$

shows that

$$|f(y) - f(x)| \leq |x-y| \int_0^1 |\nabla f(x + t(y-x))|dt.$$

Integrate this with respect to y over $D(x,1)$ and divide by ω. Thus

$$\begin{aligned} |f(x) - \frac{1}{\omega}\int_{D(x,1)} f(y)dy| &\leq \frac{1}{\omega}\int_{D(x,1)} |f(x) - f(y)|dy \\ &\lesssim \int_{D(x,1)} |x-y| \int_0^1 |\nabla f(x+t(y-x))|dtdy \\ &= \int_0^1 \int_{D(x,1)} |x-y||\nabla f(x+t(y-x))|dydt. \qquad (19) \end{aligned}$$

Make the change of variables $z = x + t(y-x)$, and then reverse the order of integration again. This leads to

$$\begin{aligned} (19) &= \int_{t=0}^1 \int_{D(x,t)} t^{-1}|z-x||\nabla f(z)|\frac{dz}{t^n}dt \\ &= \int_{D(x,1)} |z-x||\nabla f(z)| \int_{t=|z-x|}^1 t^{-(n+1)}dtdz \\ &\lesssim \int_{D(x,1)} |x-z|^{-(n-1)}|\nabla f(z)|dz \end{aligned}$$

as claimed.

Lemma A.2 Suppose that f is C^k. Then

$$|f(x)| \lesssim \sum_{0\le j<k} \|\Delta_j^f\|_{L^1(D(x,1))} + \begin{cases} \int_{D(x,1)} |x-y|^{-(n-k)}\Delta_k^f(y)dy & \text{if } 1\le k\le n-1, \\ \int_{D(x,1)} \log\frac{1}{|x-y|}\Delta_n^f(y)dy & \text{if } k=n, \\ \|\Delta_{n+1}^f\|_{L^1(D(x,1))} & \text{if } k=n+1. \end{cases} \tag{20}$$

The case $k = n+1$ is Lemma 2.2'.

Proof The case $k = 1$ follows immediately from Lemma A.1, so it has already been proved. To pass to general k we use induction based on the inequalities (here $a > 0, b > 0, |z-x| \le \text{constant}$)

$$\int_{y\in D(x,C)} |x-y|^{-(n-a)}|z-y|^{-(n-b)}dy \le \begin{cases} C|x-z|^{-(n-a-b)} & \text{if } a+b<n, \\ \log\frac{C}{|z-x|} & \text{if } a+b=n, \\ C & \text{if } a+b>n, \end{cases} \tag{21}$$

and

$$\int_{y\in D(x,C)} |x-y|^{-(n-1)}\log\frac{1}{|z-y|}dy \le C \tag{22}$$

In fact, (21) may be proved by subdividing the region of integration in the three regimes $|y-x| \le \frac{1}{2}|z-x|$, $|y-z| \le \frac{1}{2}|z-x|$ and "the rest", and noting that on the third regime the integrand is comparable to $|y-x|^{-(2n-a-b)}$. (22) may be proved similarly.

We now prove (20) by induction on k. We have done the case $k = 1$. Suppose that $2 \le k \le n-1$ and that the cases up to and including $k-1$ have been proved. Then

$$\begin{aligned} |f(x)| &\lesssim \sum_{j\le k-2} \|\Delta_j^f\|_{L^1(D(x,1))} + \int_{D(x,1)} |x-y|^{-(n-k+1)}\Delta_{k-1}^f(y)dy \\ &\lesssim \sum_{j\le k-2} \|\Delta_j^f\|_{L^1(D(x,1))} + \int_{D(x,1)} |x-y|^{-(n-k+1)}\int_{D(y,1)} \Delta_{k-1}^f(z)dzdy \\ &\quad + \int_{D(x,1)} |x-y|^{-(n-k+1)}\int_{D(y,1)} |y-z|^{-(n-1)}\Delta_k^f(z)dzdy \\ &\lesssim \sum_{j\le k-1} \|\Delta_j^f\|_{L^1(D(x,2))} + \int_{D(x,2)} |x-z|^{-(n-k)}\Delta_k^f(z)dz. \end{aligned}$$

For the first two inequalities we used (20) with k replaced by $k-1$ and 1 respectively, and for the last inequality we reversed the order of integration and used (21). The disc $D(x,2)$ can be replaced by $D(x,1)$ using rescaling, so we have proved (20) for $k \le n-1$.

To pass from $k = n-1$ to $k = n$ we argue similarly using the second case of (21), and to pass from $k = n$ to $k = n+1$ we argue similarly using (22). □

3. Fourier inversion and Plancherel

Convolution of ϕ and f is defined as follows:

$$\phi * f(x) = \int \phi(y) f(x-y) dy \tag{23}$$

We assume the reader has seen this definition before but will summarize some facts, mostly without giving the proofs. There is an issue of the appropriate conditions on ϕ and f under which the integral (23) makes sense. We recall the following.

1. If $\phi \in L^1$ and $f \in L^p$, $1 \leq p \leq \infty$, then the integral (23) is an absolutely convergent Lebesgue integral for a.e. x and

$$\|\phi * f\|_p \leq \|\phi\|_1 \|f\|_p. \tag{24}$$

2. If ϕ is a continuous function with compact support and $f \in L^1_{loc}$, then the integral (23) is an absolutely convergent Lebesgue integral for every x and $\phi * f$ is continuous.

3. If $\phi \in L^p$ and $f \in L^{p'}$, $\frac{1}{p} + \frac{1}{p'} = 1$, then the integral (23) is an absolutely convergent Lebesgue integral for every x, and $\phi * f$ is continuous. Furthermore,

$$\|\phi * f\|_\infty \leq \|\phi\|_p \|f\|_{p'}. \tag{25}$$

The absolute convergence of (23) in 2. is trivial, and (25) follows from Cauchy-Schwarz. Continuity then follows from the dominated convergence theorem. 1. is obvious when $p = \infty$. It is also true for $p = 1$, by Fubini's theorem and a change of variables. The general case follows by interpolation, see the Riesz-Thorin theorem in Section 4.

In any of the above situations convolution is commutative: the integral defining $f * \phi$ is again convergent for the same values of x, and $f * \phi = \phi * f$. This follows by making the change of variables $y \to x - y$. Notice also that

$$\mathrm{supp}(\phi * f) \subset \mathrm{supp}\phi + \mathrm{supp}f,$$

where the sum $E + F$ means $\{x + y : x \in E, y \in F\}$.

In many applications the function ϕ is fixed and very "nice", and one considers convolution as an operator

$$f \to \phi * f.$$

<u>Lemma 3.1</u> If $\phi \in C_0^\infty$ and $f \in L^1_{loc}$ then $\phi * f$ is C^∞ and

$$D^\alpha(\phi * f) = (D^\alpha \phi) * f. \tag{26}$$

<u>Proof</u> It is enough to prove that $\phi * f$ is C^1 and (26) holds for multiindices of length 1, since one can then use induction. Fix j and consider difference quotients

$$d(h) = \frac{1}{h}\Big((\phi * f)(x + he_j) - (\phi * f)(x)\Big).$$

Using (23) and commutativity of convolution, we can rewrite this as

$$d(h) = \int \frac{1}{h}(\phi(x + he_j - y) - \phi(x - y))f(y)dy.$$

The quotients

$$A_h(y) = \frac{1}{h}(\phi(x + he_j - y) - \phi(x - y))$$

are bounded by $\|\frac{\partial \phi}{\partial x_j}\|_\infty$ by the mean value theorem. For fixed x and for $|h| \le 1$, the support of A_h is contained in the fixed compact set $E = \overline{D(x,1)} + \operatorname{supp}\phi$. Thus the integrands $A_h f$ are dominated by the L^1 function $\|\frac{\partial \phi}{\partial x_j}\|_\infty \chi_E |f|$. The dominated convergence theorem implies

$$\lim_{h \to 0} d(h) = \int f(y) \lim_{h \to 0} A_h(y) dy = \int f(y) \frac{\partial \phi}{\partial x_j}(x - y) dy.$$

This proves (26) (when $|\alpha| = 1$). The continuity of the partials then follows from 2. above. □

Corollary 3.1' If $f, g \in \mathcal{S}$ then $f * g \in \mathcal{S}$.

Proof By Lemma 3.1 it suffices to show that if $(1 + |x|)^N f(x)$ and $(1 + |x|)^N g(x)$ are bounded for every N then so is $(1 + |x|)^N f * g(x)$. This follows by writing out the definitions and using (15); the details are left to the reader. □

Convolution interacts with the Fourier transform as follows: Fourier transform converts convolution to ordinary pointwise multiplication. Thus we have the following formulas:

$$\widehat{f * g} = \hat{f}\hat{g}, \;\; f, g \in L^1, \tag{27}$$

$$\widehat{fg} = \hat{f} * \hat{g}, \;\; f, g \in \mathcal{S}. \tag{28}$$

(27) follows from Fubini's theorem and is in many textbooks; the proof is left to the reader. (28) then follows easily from the inversion theorem, so we defer the proof until after Theorem 3.4.

Let $\phi \in \mathcal{S}$, and assume that $\int \phi = 1$. Define $\phi^\epsilon(x) = \epsilon^{-n}\phi(\epsilon^{-1}x)$. The family of functions $\{\phi^\epsilon\}$ is called an *approximate identity*. Notice that $\int \phi^\epsilon = 1$ for all ϵ. Thus one can regard the ϕ^ϵ as roughly convergent to the Dirac mass δ_0 as $\epsilon \to 0$. Indeed, the following fact is basic but quite standard; see any reasonable book on real analysis for the proof.

Lemma 3.2 Let $\phi \in \mathcal{S}$ and $\int \phi = 1$. Then:

1. If f is a continuous function which goes to zero at ∞ then $\phi^\epsilon * f \to f$ uniformly as $\epsilon \to 0$.

2. If $f \in L^p$, $1 \le p < \infty$ then $\phi^\epsilon * f \to f$ in L^p as $\epsilon \to 0$.

Let us note the following corollary:

Lemma 3.3 Suppose $f \in L^1_{loc}$. Then there is a fixed sequence $\{g_k\} \subset C_0^\infty$ such that if $p \in [1, \infty)$ and $f \in L^p$, then $g_k \to f$ in L^p. If f is continuous and goes to zero at ∞, then $g_k \to f$ uniformly.

The reason for stating the lemma in this way is that one sometimes has to deal with several notions of convergence simultaneously, e.g., L^1 and L^2 convergence, and it is convenient to be able to approximate f in both norms simultaneously.

Proof Let $\psi \in C_0^\infty$, $\int \psi = 1$, $\psi \geq 0$, and let ϕ be as in Lemma 1.5. Fix a sequence $\epsilon_k \downarrow 0$. Let $g_k(x) = \phi(\frac{x}{k}) \cdot (\psi^{\epsilon_k} * f)$.

If $f \in L^p$, then for large k the quantity $\|\psi^{\epsilon_k} * f\|_{L^p(|x| \geq k)}$ is bounded by $\|f\|_{L^p(|x| \geq k-1)}$ using (24) and that suppψ^{ϵ_k} is contained in $D(0,1)$. Accordingly, $\|g_k - \psi^{\epsilon_k} * f\|_p \to 0$ as $k \to \infty$. On the other hand, $\|\psi^{\epsilon_k} * f - f\|_p \to 0$ by Lemma 3.2. If f is continuous and goes to zero at ∞, then one can argue the same way using the first part of Lemma 3.2. Smoothness of g_k follows from Lemma 3.1, so the proof is complete. $\square$

Theorem 3.4 (Fourier inversion) Suppose that $f \in L^1$, and assume that $\hat{f}$ is also in L^1. Then for a.e. x,

$$f(x) = \int \hat{f}(\xi) e^{2\pi i \xi \cdot x} d\xi. \tag{29}$$

Equivalently,

$$\hat{\hat{f}}(x) = f(-x) \text{ for a.e. } x. \tag{30}$$

The proof uses Lemma 3.2 and also the following facts:

A. The gaussian $\Gamma(x) = e^{-\pi|x|^2}$ satisfies $\hat{\Gamma} = \Gamma$, and therefore also satisfies (30). So at any rate there is one function f for which Theorem 3.4 is true. In fact this implies that there are many such functions. Indeed, if we form the functions

$$\Gamma_\epsilon(x) = e^{-\pi\epsilon^2|x|^2},$$

then we have

$$\widehat{\Gamma_\epsilon}(\xi) = \epsilon^{-n} e^{-\pi\frac{|\xi|^2}{\epsilon^2}}. \tag{31}$$

Applying this again with ϵ replaced by ϵ^{-1}, one can verify that Γ_ϵ satisfies (30). See the discussion after formula (5).

B. The *duality relation* for the Fourier transform, i.e., the following lemma.

Lemma 3.5 Suppose that $\mu \in M(\mathbb{R}^n)$ and $\nu \in M(\mathbb{R}^n)$. Then

$$\int \hat{\mu} d\nu = \int \hat{\nu} d\mu. \tag{32}$$

In particular, if $f, g \in L^1$, then

$$\int \hat{f}(x) g(x) dx = \int f(x) \hat{g}(x) dx.$$

Proof This follows from Fubini's theorem:

$$\begin{aligned} \int \hat{\mu} d\nu &= \int \int e^{-2\pi i \xi \cdot x} d\mu(x) d\nu(\xi) \\ &= \int \int e^{-2\pi i \xi \cdot x} d\nu(\xi) d\mu(x) \\ &= \int \hat{\nu} d\mu. \end{aligned}$$

$\square$

Proof of Theorem 3.4 Consider the integral in (29) with a damping factor included:

$$I_\epsilon(x) = \int \hat{f}(\xi) e^{-\pi \epsilon^2 |\xi|^2} e^{2\pi i \xi \cdot x} d\xi. \tag{33}$$

We evaluate the limit as $\epsilon \to 0$ in two different ways.

1. As $\epsilon \to 0$, $I_\epsilon(x) \to \int \hat{f}(\xi) e^{2\pi i \xi \cdot x} d\xi$ for each fixed x. This follows from the dominated convergence theorem, since $\hat{f} \in L^1$.

2. With x and ϵ fixed, define $g(\xi) = e^{-\pi \epsilon^2 |\xi|^2} e^{2\pi i \xi \cdot x}$. Thus

$$I_\epsilon(x) = \int f(y) \hat{g}(y) dy$$

by Lemma 3.5. On the other hand, we can evaluate $\hat{g}$ using the fact that $g(\xi) = e_x(\xi) \Gamma_\epsilon(\xi)$ and (4), (31). Thus

$$\hat{g}(y) = \widehat{\Gamma_\epsilon}(y - x) = \Gamma^\epsilon(x - y),$$

where $\Gamma^\epsilon(y) = \epsilon^{-n} \Gamma(\frac{y}{\epsilon})$ is an approximate identity as in Lemma 3.2, and we have used that Γ is even.

Accordingly,

$$I_\epsilon = f * \Gamma^\epsilon,$$

and we conclude by Lemma 3.2 that

$$I_\epsilon \to f$$

in L^1 as $\epsilon \to 0$.

Summing up, we have seen that the functions I_ϵ converge pointwise to $\int \hat{f}(\xi) e^{2\pi i x\cdot\xi} d\xi$, and converge in L^1 to f. This is only possible when (29) holds. $\square$

Corollaries of the inversion theorem

The first corollary below is not really a corollary, but a reformulation of the proof without the assumption that $\hat{f} \in L^1$. This is the form the inversion theorem takes for general f. Notice that the integrals I_ϵ are well defined for any $f \in L^1$, since the Gaussian $e^{-\pi\epsilon^2|x|^2}$ is integrable for each fixed ϵ. Corollary 3.6 is often stated as "the Fourier transform of f is Gauss-Weierstrass summable to f", and can be compared to the theorem on Cesaro summability for Fourier series.

Corollary 3.6 1. Suppose $f \in L^1$ and define $I_\epsilon(x)$ via (33). Then $I_\epsilon \to f$ in L^1 as $\epsilon \to 0$.

2. If $1 < p < \infty$ and additionally $f \in L^p$, then $I_\epsilon \to f$ in L^p as $\epsilon \to 0$. If instead f is continuous and goes to zero at ∞, then $I_\epsilon \to f$ uniformly.

Proof This follows from the preceding argument showing that $I_\epsilon = \Gamma^\epsilon * f$, together with Proposition 3.2. $\square$

Corollary 3.7 If $f \in L^1$ and $\hat{f} = 0$, then $f = 0$.

This is immediate from Theorem 3.4. $\square$

Theorem 3.8 The Fourier transform maps $\mathcal{S}$ onto $\mathcal{S}$.

Proof Given $f \in \mathcal{S}$, let $F(x) = f(-x)$ and let $g = \hat{F}$. Then $g \in \mathcal{S}$ by Theorem 2.3, and (30) implies

$$\hat{g}(x) = \hat{\hat{F}}(x) = F(-x) = f(x)$$

$\square$

Let us also prove formula (28). Let $f, g \in \mathcal{S}$. Then

$$\begin{aligned} \widehat{\hat{f} * \hat{g}}(-x) &= \hat{\hat{f}}(-x)\hat{\hat{g}}(-x) \\ &= f(x)g(x) \end{aligned}$$

by (27) and Theorem 3.4. Using Theorem 3.4 again, it follows that $\hat{f} * \hat{g} = \widehat{fg}$. $\square$

Theorem 3.9 (Plancherel theorem, first version) If $u, v \in \mathcal{S}$ then

$$\int \hat{u}\overline{\hat{v}} = \int u\overline{v}. \tag{34}$$

Proof By the inversion theorem,

$$\int u(x)\overline{v}(x)dx = \int \hat{\hat{u}}(-x)\overline{v}(x)dx = \int \hat{\hat{u}}(x)\overline{v(-x)}dx$$

i.e.,

$$\int \hat{u}\overline{\hat{v}} = \int \hat{\hat{u}}\tilde{v}.$$

Applying the duality relation to the right-hand side we obtain

$$\int u\overline{v} = \int \hat{u}\hat{\tilde{v}},$$

and now (34) follows from (6). □

Theorem 3.9 says that the Fourier transform restricted to Schwartz functions is an isometry in the L^2 norm. Since $\mathcal{S}$ is dense in L^2 (e.g., by Lemma 3.3) this suggests a way of extending the Fourier transform to L^2.

Theorem 3.10 (Plancherel theorem, second version) There is a unique bounded operator $\mathcal{F} : L^2 \to L^2$ such that $\mathcal{F}f = \hat{f}$ when $f \in \mathcal{S}$. $\mathcal{F}$ has the following additional properties:

1. $\mathcal{F}$ is a unitary operator.
2. $\mathcal{F}f = \hat{f}$ if $f \in L^1 \cap L^2$.

Proof The existence and uniqueness statement is immediate from Theorem 3.9, as is the fact that $\|\mathcal{F}f\|_2 = \|f\|_2$. In view of this isometry property, the range of $\mathcal{F}$ must be closed, and unitarity of $\mathcal{F}$ will follow if we show that the range is dense. However, the latter statement is immediate from Theorem 3.8 and Lemma 3.3.

It remains to prove 2. For $f \in \mathcal{S}$, 2. is true by definition. Suppose now that $f \in L^1 \cap L^2$. By Lemma 3.3, there is a sequence $\{g_k\} \subset \mathcal{S}$ which converges to f both in L^1 and in L^2. By Proposition 1.1, $\widehat{g_k}$ converges to $\hat{f}$ uniformly. On the other hand, $\widehat{g_k}$ converges to $\mathcal{F}f$ in L^2 by boundedness of the operator $\mathcal{F}$. It follows that $\mathcal{F}f = \hat{f}$. □

Statement 2. allows us to use the notation $\hat{f}$ for $\mathcal{F}f$ if $f \in L^2$ without any possible ambiguity. We may therefore extend the definition of the Fourier transform to $L^1 + L^2$ (in fact to σ-finite measures of the form $\mu + f dx$, $\mu \in M(\mathbb{R}^n)$, $f \in L^2$) via $\widehat{f+g} = \hat{f} + \hat{g}$.

Corollary 3.12 The following form of the duality relation is valid:

$$\int \hat{\nu}\psi = \int \hat{\psi} d\nu, \ \psi \in \mathcal{S},$$

if $\nu = \mu + f dx$, $\mu \in M(\mathbb{R}^n)$, $f \in L^2$.

Proof We have already proved this in Lemma 3.5 if $f = 0$, so it suffices to prove it when $\mu = 0$, i.e., to show that

$$\int \hat{f}\psi = \int f\hat{\psi}$$

if $f \in L^2$, $\psi \in \mathcal{S}$. This is true by Lemma 3.5 if $f \in L^1 \cap L^2$. Therefore it is also true for $f \in L^2$, since for fixed ψ both sides depend continuously on f (in the case of the left-hand side this follows from the Plancherel theorem). $\square$

Theorem 3.13 If $\mu \in M(\mathbb{R}^n)$, $f \in L^2$ and

$$\hat{f} + \hat{\mu} = 0,$$

then $\mu = -f dx$. In particular, if $\mu \in M(\mathbb{R}^n)$ and $\hat{\mu} \in L^2$, then μ is absolutely continuous with respect to the Lebesgue measure with an L^2 density.

Proof By the Riesz representation theorem for measures on compact sets, the measure $\mu + f dx$ will be zero provided

$$\int \phi d\mu + \phi f dx = 0 \tag{35}$$

for continuous ϕ with compact support.

If $\phi \in C_0^\infty$ then (35) follows from Corollary 3.12. In general, we choose (e.g., by Proposition 3.2) a sequence ϕ_k in C_0^∞ which converges to ϕ uniformly and in L^2. We write down (35) for the ϕ_k's and pass to the limit. This proves (35).

To prove the last statement, suppose that $\hat{\mu} \in L^2$, and choose (by Theorem 3.10) a function $g \in L^2$ with $\hat{g} = \hat{\mu}$. Then $d\mu - g dx$ has Fourier transform zero, so by the first part of the proof $d\mu = g dx$. $\square$

All the basic formulas for the L^1 Fourier transform extend to the $L^1 + L^2$-Fourier transform by approximation arguments. This was done above in the case of the duality relation. Let us note in particular that the transformation formulas in Section 1 extend. For example, in the case of (5), one has

$$\widehat{f \circ T} = |\det(T)|^{-1} \hat{f} \circ T^{-t} \tag{36}$$

if $f \in L^1 + L^2$. Since we already know this when $f \in L^1$, it suffices to prove it when $f \in L^2$. Choose $\{f_k\} \subset L^1 \cap L^2$, $f_k \to f$ in L^2. Composition with T is continuous on L^2, as is Fourier transform, so we can write down (36) for the $\{f_k\}$ and pass to the limit.

The $L^1 + L^2$ domain for the Fourier transform is wide enough to include many natural examples. Note in particular that $L^p \subset L^1 + L^2$ if $p \in (1, 2)$. Furthermore, certain homogeneous functions belong to $L^1 + L^2$ although none of them can belong to L^p for any fixed p. For example, $|x|^{-a}$ belongs to $L^1 + L^2$ if $\frac{n}{2} < a < n$, since it belongs to L^1 at the origin and to L^2 at infinity. However, the $L^1 + L^2$ domain is not always sufficient. The most natural way to proceed would be to develop the idea of tempered distributions, but we don't want to do this explicitly. Instead, we further broaden the definition of Fourier transform as follows: a *tempered function* is a function $f \in L^1_{loc}(\mathbb{R}^n)$ such that

$$\int (1+|x|)^{-N}|f(x)|dx < \infty$$

for some constant N. Roughly, f has at most polynomial growth in the sense of L^1 averages.

It is clear that if f is tempered and $\phi \in \mathcal{S}$, then $\int |\phi f| < \infty$. Furthermore, the map $\phi \to \int \phi f$ is continuous on $\mathcal{S}$. It follows that $\phi * f$ is well defined if $\phi \in \mathcal{S}$ and f is tempered, and a simple estimation shows that $\phi * f$ is again tempered.

If f and g are tempered functions, we say that *g is the distributional Fourier transform of f* if

$$\int g\phi = \int f\hat{\phi} \tag{37}$$

for all $\phi \in \mathcal{S}$. For given f, such a function g is unique using the density properties of $\mathcal{S}$ as in several previous arguments. We denote g by $\hat{f}$. Notice also that if $f \in L^1 + L^2$, then its $L^1 + L^2$-Fourier transform coincides with its distributional Fourier transform by Proposition 3.12.

All the basic formulas, in particular (3), (4), (5), (6), extend to the case of distributional Fourier transforms, e.g., if g is the distributional Fourier transform of f, then $|\det T|^{-1}\hat{f} \circ T^{-t}$ is the distributional Fourier transform of $f \circ T$. This may be seen by making appropriate changes of variable in the integrals in (37). We indicate how these arguments are carried out by proving the extended version of formula (27). Namely, if f is tempered, $\psi \in \mathcal{S}$, and f has a distributional Fourier transform, then so does $\psi * f$ and

$$\widehat{\psi * f} = \hat{\psi}\hat{f} \tag{38}$$

The proof is as follows. Let ϕ be another Schwartz function. Then

$$\begin{aligned}
\int (\hat{\psi}\hat{f})\phi &= \int \hat{f}(\hat{\psi}\phi) \\
&= \int f\widehat{\hat{\psi}\phi} \\
&= \int f(\hat{\hat{\psi}} * \hat{\phi}) \\
&= \int f(x)\int \psi(-(x-y))\hat{\phi}(y)dydx \\
&= \int \psi * f(y)\hat{\phi}(y).
\end{aligned}$$

The second line followed from the definition of distributional Fourier transform, the third line from (28), and the next to last line used the inversion theorem for ψ. Comparing with the definition (37), we see that we have proved (38).

Let us note also that the inversion theorem is true for distributional Fourier transforms: if f is tempered and has a distributional Fourier transform $\hat{f}$, then $\hat{f}$ has the distributional Fourier transform $f(-x)$. Here is the proof. If $\phi \in \mathcal{S}$, then

$$\begin{aligned}
\int f(-x)\phi(x)dx &= \int f(x)\phi(-x)dx \\
&= \int f(x)\hat{\hat{\phi}}(x)dx \\
&= \int \hat{f}(x)\hat{\phi}(x)dx.
\end{aligned}$$

We used a change of variables, the inversion theorem for ϕ, and the definition (37) of $\hat{f}$. Comparing again with (37), we have the stated result. $\square$

4. Some specifics, and L^p for $p < 2$

We first discuss a couple of basic examples where the Fourier transform can be calculated, namely powers of the distance to the origin and complex Gaussians.

Proposition 4.1 Let $h_a(x) = \frac{\gamma(a/2)}{\pi^{a/2}}|x|^{-a}$. Then $\widehat{h_a} = h_{n-a}$ in the sense of $L^1 + L^2$ Fourier transforms if $\frac{n}{2} < \mathrm{Re}\,(a) < n$, and in the sense of distributional Fourier transforms if $0 < \mathrm{Re}\,(a) < n$.

Here γ is the gamma function, i.e.,

$$\gamma(s) = \int_0^\infty e^{-t} t^{s-1} dt.$$

Proof Suppose that a is real and $\frac{n}{2} < a < n$. Then $h_a \in L^1 + L^2$. The functions of the form $f(x) = c|x|^{-a}$ with c constant may be characterized by the following two transformation properties:

1. f is radial, i.e., $f \circ \rho = f$ for all linear $\rho : \mathbb{R}^n \to \mathbb{R}^n$ with $\rho^t \rho =$identity.
2. f is homogeneous of degree $-a$, i.e.,

$$f(\epsilon x) = \epsilon^{-a} f(x) \tag{39}$$

for each $\epsilon > 0$.

We will use the notation

$$f_\epsilon(x) = f(\epsilon x), \tag{40}$$

$$f^\epsilon(x) = \epsilon^{-n} f(\frac{x}{\epsilon}). \tag{41}$$

Let $f(x) = |x|^{-a}$, $\frac{n}{2} < a < n$. Taking Fourier transforms we obtain from 1. and 2. the following (see the discussion in Section 1 regarding special cases of (5)): $\hat{f}$ is radial, and $\hat{f}^{\epsilon} = \epsilon^{-a}\hat{f}$, which is equivalent to $\hat{f}_{\epsilon} = \epsilon^{-(n-a)}\hat{f}$. Hence $\hat{f} = c|x|^{-(n-a)}$, and it remains to evaluate the constant c. For this we use the duality relation, taking the Schwartz function ψ to be the Gaussian Γ. Thus

$$\int |x|^{-a} e^{-\pi|x|^2} dx = c \int |x|^{-(n-a)} e^{-\pi|x|^2} dx. \tag{42}$$

To evaluate the left hand side, change to polar coordinates and then make the change of variable $t = \pi r^2$. Thus, if σ is the area of the unit sphere, we get

$$\begin{aligned} \int |x|^{-a} e^{-\pi|x|^2} dx &= \sigma \int_0^\infty e^{-\pi r^2} r^{n-a} \frac{dr}{r} \\ &= \sigma \int_0^\infty e^{-t} (\frac{t}{\pi})^{\frac{n-a}{2}} \frac{dt}{2t} \\ &= \frac{\sigma}{2} \pi^{-(\frac{n-a}{2})} \gamma(\frac{n-a}{2}), \end{aligned}$$

and similarly the right hand side of (42) is $c\frac{\sigma}{2}\pi^{-(\frac{a}{2})}\gamma(\frac{a}{2})$. Hence

$$c = \frac{\pi^{\frac{a}{2}}\gamma(\frac{n-a}{2})}{\pi^{\frac{n-a}{2}}\gamma(\frac{a}{2})},$$

and the proposition is proved in the case $\frac{n}{2} < a < n$.

For the general case, fix $\phi \in \mathcal{S}$ and consider the two integrals

$$A(z) = \int h_z \hat{\phi},$$

$$B(z) = \int h_{n-z} \phi.$$

Both A and B may be seen to be analytic in z in the indicated regime: since γ is analytic, this reduces to showing that

$$\int |x|^{-z} \phi(x) dx$$

is analytic when $\phi \in \mathcal{S}$, which may be done by using the dominated convergence theorem to justify complex differentiation under the integral sign.

By Proposition 3.14, A and B agree for z in $(\frac{n}{2}, n)$. So they agree everywhere by the uniqueness theorem. This proves that the distributional Fourier transform of h_a exists and is h_{n-a}. If $\operatorname{Re} a > \frac{n}{2}$, then $h_a \in L^1 + L^2$, so that its $L^1 + L^2$ and distributional Fourier transforms coincide. $\square$

Let T be an invertible $n\times n$ real symmetric matrix. The *signature* of T is the quantity $k_+ - k_-$ where k_+ and k_- are the numbers of positive and negative eigenvalues of T, counted with multiplicity. We also define

$$G_T(x) = e^{-\pi i\langle Tx,x\rangle},$$

and observe that G_T has absolute value 1 and is therefore tempered.

Proposition 4.2 Let T be an invertible $n \times n$ real symmetric matrix with signature σ. Then G_T has a distributional Fourier transform, which is equal to

$$e^{-\pi i\frac{\sigma}{4}}|\det T|^{-\frac{1}{2}}G_{-T^{-1}}$$

Remark This can easily be generalized to complex symmetric T with nonnegative imaginary part (the latter condition is needed, else G_T is not tempered). See [17], Theorem 7.6.1. If $n = 1$, we do this case in the course of the proof.

Proof We need to show that

$$\int e^{-\pi i\langle Tx,x\rangle}\hat{\phi}(x)dx = e^{\frac{-\pi i}{4}\sigma}|\det T|^{-\frac{1}{2}}\int e^{\pi i\langle T^{-1}x,x\rangle}\phi(x)dx \tag{43}$$

if $\phi \in \mathcal{S}$ and T is invertible real symmetric.

First consider the $n = 1$ case. Let $\sqrt{z}$ be the branch of the square root defined on the complement of the nonpositive real numbers and positive on the positive real axis. Thus $\sqrt{\pm i} = e^{\pm\frac{\pi i}{4}}$. Accordingly, (43) with $n = 1$ is equivalent to

$$\int e^{-\pi zx^2}\hat{\phi}(x)dx = (\sqrt{z})^{-1}\int e^{-\pi\frac{x^2}{z}}\phi(x)dx \tag{44}$$

if $\phi \in \mathcal{S}$ and z is pure imaginary and not equal to zero. We prove this formula by analytic continuation from the real case.

Namely, if $z = 1$ then (44) is Example 2 in Section 1, and the case of z real and positive then follows from scaling, i.e., the fact that the Fourier transform of f_ϵ is $\hat{f}^\epsilon$, see (5). Both sides of (44) are easily seen to be analytic in z when $\mathrm{Re}\, z > 0$ and continuous in z when $\mathrm{Re}\, z \geq 0, z \neq 0$, so (44) is proved.

Now consider the $n \geq 2$ case. Observe that if (43) is true for a given T (and all ϕ), it is true also when T is replaced by UTU^{-1} for any $U \in SO(n)$. This follows from the fact that $\widehat{f\circ U} = \hat{f}\circ U$. However, since we did not give an explicit proof of the latter fact for distributional Fourier transforms, we will now exhibit the necessary calculations. Let $S = UTU^{-1}$. Thus S and T have the same determinant and the same signature. Accordingly, if (43) holds for T then

$$\int e^{-\pi i\langle Sx,x\rangle}\hat{\phi}(x)dx \quad = \quad \int e^{-\pi i\langle TU^{-1}x,U^{-1}x\rangle}\hat{\phi}(x)dx$$

$$\begin{aligned}
&= \int e^{-\pi i\langle Tx,x\rangle}\hat{\phi}(Ux)dx \\
&= \int e^{-\pi i\langle Tx,x\rangle}\widehat{\phi\circ U}(x)dx \\
&= e^{-\frac{\pi i}{4}\sigma}|\det T|^{-\frac{1}{2}}\int e^{\pi i\langle T^{-1}x,x\rangle}\phi\circ U(x)dx \\
&= e^{-\frac{\pi i}{4}\sigma}|\det T|^{-\frac{1}{2}}\int e^{\pi i\langle T^{-1}U^{-1}x,U^{-1}x\rangle}\phi(x)dx \\
&= e^{-\frac{\pi i}{4}\sigma}|\det S|^{-\frac{1}{2}}\int e^{\pi i\langle S^{-1}x,x\rangle}\phi(x)dx.
\end{aligned}$$

We used that $\widehat{\phi\circ U}=\hat{\phi}\circ U$ for Schwartz functions ϕ, see the comments after formula (5) in Section 1.

It therefore suffices to prove (43) when T is diagonal. If T is diagonal and ϕ is a tensor function, then the integrals in (43) factor as products of one variable integrals and (43) follows immediately from (44). The general case then follows from Proposition 2.1' and the fact that integration against a tempered function defines a continuous linear functional on $\mathcal{S}$. □

We now briefly discuss the L^p Fourier transform, $1<p<2$. The most basic result is the Hausdorff-Young theorem, which is a formal consequence of the Plancherel theorem and Proposition 1.1 via the following.

Riesz-Thorin interpolation theorem. Let T be a linear operator with domain $L^{p_0}+L^{p_1}$, $1\le p_0<p_1\le\infty$. Assume that $f\in L^{p_1}$ implies

$$\|Tf\|_{q_0}\le A_0\|f\|_{p_0} \tag{45}$$

and $f\in L^{p_1}$ implies

$$\|Tf\|_{q_1}\le A_1\|f\|_{p_1} \tag{46}$$

for some $1\le q_0,q_1\le\infty$. Suppose that for a certain $\theta\in(0,1)$,

$$\frac{1}{p}=\frac{1-\theta}{p_0}+\frac{\theta}{p_1} \tag{47}$$

and

$$\frac{1}{q}=\frac{1-\theta}{q_0}+\frac{\theta}{q_1}. \tag{48}$$

Then $f\in L^p$ implies

$$\|Tf\|_q\le A_0^{1-\theta}A_1^{\theta}\|f\|_p.$$

For the proof see [20], [34], or numerous other textbooks.

We will adopt the convention that when indices p and p' are used we must have

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

Proposition 4.3 (Hausdorff-Young) If $1 \le p \le 2$ then

$$\|\hat{f}\|_{p'} \le \|f\|_p. \tag{49}$$

Proof We interpolate between the cases $p=1$ and 2, which we already know. Namely, apply the Riesz-Thorin theorem with $p_0=1, q_0=\infty$, $p_1=q_1=2$, $A_0=A_1=1$. The hypotheses (45) and (46) follow from Proposition 1.1 and Theorem 3.10 respectively. For given p,q, existence of $\theta \in (0,1)$ for which (47) and (48) hold is equivalent to $1<p<2$ and $q=p'$. The result follows. $\square$

For later reference we insert here another basic result which follows from Riesz-Thorin, although this one (in contrast to Hausdorff-Young) could also be proved by elementary manipulation of inequalities.

Proposition 4.4(Young's inequality) Let $\phi \in L^p$, $\psi \in L^r$, where $1 \le p,r \le \infty$ and $\frac{1}{p}+\frac{1}{r} \ge 1$. Let $\frac{1}{q}=\frac{1}{p}-\frac{1}{r'}$. Then the integral defining $\phi * \psi$ is absolutely convergent for a.e. x and

$$\|\phi * \psi\|_q \le \|\phi\|_p \|\psi\|_r.$$

Proof View ϕ as fixed, i.e., define

$$T\psi = \phi * \psi.$$

Inequalities (24) and (25) imply that

$$T : L^1 + L^{p'} \to L^p + L^\infty$$

with

$$\|T\psi\|_p \le \|\phi\|_p \|\psi\|_1$$
$$\|T\psi\|_\infty \le \|\phi\|_p \|\psi\|_{p'}.$$

If $\frac{1}{q}=\frac{1}{p}-\frac{1}{r'}$ then there is $\theta \in [0,1]$ with

$$\frac{1}{r}=\frac{1-\theta}{1}+\frac{\theta}{p'}$$
$$\frac{1}{q}=\frac{1-\theta}{p}+\frac{\theta}{\infty}.$$

The result now follows from Riesz-Thorin. $\square$

Remarks 1. Unless $p = 1$ or 2, the constant 1 in the Hausdorff-Young inequality is not the best possible; indeed the best constant is found by testing the Gaussian function Γ. This is much deeper and is due to Babenko when p' is an even integer and to Beckner [1] in general. There are some related considerations in connection with Proposition 4.4, due also to Beckner.

2. Except in the case $p = 2$ the inequality (49) is not reversible, in the sense that there is no constant C such that $\|f\|_{p'} \geq \|f\|_p$ when $f \in \mathcal{S}$. Equivalently (in view of the inversion theorem) the result does not extend to the case $p > 2$. This is not at all difficult to show, but we discuss it at some length in order to illustrate a few different techniques used for constructing examples in connection with the L^p Fourier transform. Here is the most elementary argument.

Exercise Using translation and multiplication by characters, construct a sequence of Schwartz functions $\{\phi_n\}$ so that

1. Each ϕ_n has the same L^p norm.
2. Each $\widehat{\phi_n}$ has the same $L^{p'}$ norm.
3. The supports of the $\widehat{\phi_n}$ are disjoint.
4. The supports of the ϕ_n are "essentially disjoint" meaning that

$$\|\sum_{n=1}^{N} \phi_n\|_p^p \approx \sum_{n=1}^{N} \|\phi_n\|_p^p \ \ (\approx N)$$

uniformly in N.

Use this to disprove the converse of Hausdorff-Young.

Here is a second argument based on Proposition 4.2. This argument can readily be adapted to show that there are functions $f \in L^p$ for any $p > 2$ which do not have a distributional Fourier transform in our sense. See [17], Theorem 7.6.6.

Take $n = 1$ and $f_\lambda(x) = \phi(x)e^{-\pi i\lambda x^2}$, where $\phi \in C_0^\infty$ is fixed. Here λ is a large positive number. Then $\|f_\lambda\|_p$ is independent of λ for any p. By the Plancherel theorem, $\|\widehat{f_\lambda}\|_2$ is also independent of λ. On the other hand, $\widehat{f_\lambda}$ is the convolution of $\hat{\phi}$, which is in L^1, with $(\sqrt{i\lambda})^{-1}e^{\pi i\lambda^{-1}x^2}$, which has L^∞ norm $\lambda^{-\frac{1}{2}}$. Accordingly, if $p < 2$ then

$$\begin{aligned} \|\widehat{f_\lambda}\|_{p'} &\leq \|\widehat{f_\lambda}\|_2^{\frac{2}{p'}} \|\widehat{f_\lambda}\|_\infty^{1-\frac{2}{p'}} \\ &\lesssim \lambda^{-(\frac{1}{2}-\frac{1}{p'})}. \end{aligned}$$

Since $\|f_\lambda\|_p$ is independent of λ, this shows that when $p < 2$ there is no constant C such that $C\|f\|_{p'} \geq \|f\|_p$ for all $f \in \mathcal{S}$. $\square$

Here now is another important technique ("randomization") and a third disproof of the converse of Hausdorff-Young.

Let $\{\omega_n\}_{n=1}^N$ be independent random variables taking values ± 1 with equal probability. Denote expectation (a.k.a. integral over the probability space in question) by $\mathbb{E}$, and probability (a.k.a. measure) by Prob. Let $\{a_n\}_{n=1}^N$ be complex numbers.

<u>Proposition 4.5</u> (Khinchin's inequality)

$$\mathbb{E}(|\sum_{n=1}^N a_n\omega_n|^p) \approx (\sum_{n=1}^N |a_n|^2)^{\frac{p}{2}} \tag{50}$$

for any $0 < p < \infty$, where the implicit constants depend on p only.

Most books on probability and many analysis books give proofs. Here is the proof in the case $p > 1$. There are three steps.

(i) When $p = 2$ it is simple to see from independence that (50) is true with equality: expand out the left side and observe that the cross terms cancel.

(ii) The upper bound. This is best obtained as a consequence of a stronger ("subgaussian") estimate. One can clearly assume the $\{a_n\}$ are real and (52) below is for real $\{a_n\}$. Let $t > 0$. We have

$$\mathbb{E}(e^{t\sum_n a_n\omega_n}) = \prod_n \mathbb{E}(e^{ta_n\omega_n}) = \prod_n \frac{1}{2}(e^{ta_n} + e^{-ta_n}),$$

where the first equality follows from independence and the fact that $e^{x+y} = e^x e^y$. Use the numerical inequality

$$\frac{1}{2}(e^x + e^{-x}) \le e^{\frac{x^2}{2}} \tag{51}$$

to conclude that

$$\mathbb{E}(e^{t\sum_n a_n\omega_n}) \le e^{\frac{t^2}{2}\sum_n a_n^2},$$

therefore

$$\mathrm{Prob}(\sum_n a_n\omega_n \ge \lambda) \le e^{-t\lambda + \frac{t^2}{2}\sum_n a_n^2}$$

for any $t > 0$ and $\lambda > 0$. Taking $t = \frac{\lambda}{\sum_n a_n^2}$ gives

$$\mathrm{Prob}(\sum_n a_n\omega_n \ge \lambda) \le e^{-\frac{\lambda^2}{2\sum_n a_n^2}}, \tag{52}$$

hence

$$\mathrm{Prob}(|\sum_n a_n\omega_n| \ge \lambda) \le 2e^{-\frac{\lambda^2}{2\sum_n a_n^2}}.$$

From this and the formula for the L^p norm in terms of the distribution function,

$$\mathbb{E}(|f|^p) = p \int \lambda^{p-1} \mathrm{Prob}(|f| \geq \lambda) d\lambda$$

one gets

$$\mathbb{E}(|\sum_n a_n \omega_n|^p) \leq 2p \int \lambda^{p-1} e^{-\frac{\lambda^2}{2\sum_n a_n^2}} d\lambda = 2^{2+\frac{p}{2}} p \gamma(\frac{p}{2})(\sum_n a_n^2)^{\frac{p}{2}}.$$

This proves the upper bound.

(iii) The lower bound. This follows from (i) and (ii) by duality. Namely

$$\begin{aligned} \sum_n |a_n|^2 &= \mathbb{E}(|\sum_n a_n \omega_n|^2) \\ &\leq \mathbb{E}(|\sum_n a_n \omega_n|^p)^{\frac{1}{p}} \mathbb{E}(|\sum_n a_n \omega_n|^{p'})^{\frac{1}{p'}} \\ &\lesssim (\sum_n |a_n|^2)^{\frac{1}{2}} \mathbb{E}(|\sum_n a_n \omega_n|^p)^{\frac{1}{p}}, \end{aligned}$$

so that

$$\mathbb{E}(|\sum_n a_n \omega_n|^p)^{\frac{1}{p}} \gtrsim (\sum_n |a_n|^2)^{\frac{1}{2}}$$

as claimed. $\square$.

To apply this in connection with the converse of Hausdorff-Young, let ϕ be a C_0^∞ function, and let $\{k_j\}_{j=1}^N$ be such that the functions $\phi_j \overset{def}{=} \phi(\cdot - k_j)$ have disjoint support. Thus $\widehat{\phi_n}(\xi) = e^{2\pi i \xi \cdot k_n} \hat{\phi}(\xi)$. The L^p norm of $\sum_{n \leq N} \omega_n \phi_n$ is independent of ω in view of the disjoint support, indeed

$$\| \sum_{n \leq N} \omega_n \phi_n \|_p = C N^{\frac{1}{p}}, \tag{53}$$

where $C = \|\phi\|_p$.

Now consider the corresponding Fourier side norms, more precisely the expectation of their p' powers:

$$\mathbb{E}(\| \sum_{n \leq N} \omega_n \widehat{\phi_n} \|_{p'}^{p'}. \tag{54}$$

We have by Fubini's theorem

$$\begin{aligned} (54) &= \mathbb{E}(\| \sum_{n \leq N} \omega_n e^{2\pi i \xi \cdot k_n} \hat{\phi}(\xi)) \|_{L^{p'}(d\xi)}^{p'} \\ &= \int_{\mathbb{R}^n} |\hat{\phi}(\xi)|^{p'} \mathbb{E}(|\sum_{n \leq N} \omega_n e^{2\pi i \xi \cdot k_n}|^{p'}) \\ &\approx N^{\frac{p'}{2}}, \end{aligned}$$

where at the last step we used Khinchin.

It follows that we can make a choice of ω so that $\| \sum_{n\le N} \omega_n \widehat{\phi_n}\|_{p'} \lesssim N^{\frac{1}{2}}$. If $p<2$ and if N is large, this is much smaller than the right hand side of (53), so we are done.

5. Uncertainty Principle

The uncertainty principle is[1] the heuristic statement that if a measure μ is supported on R, then for many purposes $\hat{\mu}$ may be regarded as being constant on any dual ellipsoid R^*.

The simplest rigorous statement is

Proposition 5.1 (L^2 Bernstein inequality) Assume that $f \in L^2$ and $\hat{f}$ is supported in $D(0,R)$. Then f is C^∞ and there is an estimate

$$\|D^\alpha f\|_2 \le (2\pi R)^{|\alpha|}\|f\|_2. \tag{55}$$

Proof Essentially this is an immediate consequence of the Plancherel theorem, see the calculation at the end of the proof. However, there are some details to take care of if one wants to be rigorous.

The Fourier inversion formula

$$f(x) = \int \hat{f}(\xi) e^{2\pi i x\cdot\xi} d\xi \tag{56}$$

is valid (in the naive sense). Namely, note that the support assumption implies $\hat{f} \in L^1$, and choose a sequence of Schwartz functions ψ_k which converges to $\hat{f}$ both in L^1 and in L^2. Let $\phi_k \in \mathcal{S}$ satisfy $\widehat{\phi_k} = \psi_k$. The formula (56) is valid for ϕ_k. As $k \to \infty$, the left sides converge in L^2 to f by Theorem 3.10 and the right sides converge uniformly to $\int \hat{f}(\xi) e^{2\pi i x\cdot\xi} d\xi$, which proves (56) for f.

Proposition 1.3 applied to $\hat{f}$ now implies that f is C^∞ and that $D^\alpha f$ is obtained by differentiation under the integral sign in (56). The estimate (55) holds since

$$\|D^\alpha f\|_2 = \|\widehat{D^\alpha f}\|_2 = \|(2\pi i\xi)^\alpha \hat{f}\|_2 \le (2\pi R)^{|\alpha|}\|\hat{f}\|_2 = (2\pi R)^{|\alpha|}\|f\|_2. \ \square$$

A corresponding statement is also true in L^p norms, but proving this and other related results needs a different argument since there is no Plancherel theorem.

Lemma 5.2 There is a fixed Schwartz function ϕ such that if $f \in L^1 + L^2$ and $\widehat{f}$ is supported in $D(0,R)$, then

$$f = \phi^{R^{-1}} * f.$$

[1] This should be qualified by adding "as far as we are concerned". There are various more sophisticated related statements which are also called uncertainty principle; see for example [14], [15] and references there.

Proof Take $\phi \in \mathcal{S}$ so that $\widehat{\phi}$ is equal to 1 on $D(0,1)$. Thus $\widehat{\phi^{R^{-1}}}(\xi) = \widehat{\phi}(R^{-1}\xi)$ is equal to 1 on $D(0,R)$, so $(\phi^{R^{-1}} * f - f)\widehat{}$ vanishes identically. Hence $\phi^{R^{-1}} * f = f$. $\square$

Proposition 5.3 (Bernstein's inequality for a disc) Suppose that $f \in L^1 + L^2$ and $\hat{f}$ is supported in $D(0,R)$. Then

1. For any α and $p \in [1,\infty]$,

$$\|D^\alpha f\|_p \leq (CR)^{|\alpha|}\|f\|_p.$$

2. For any $1 \leq p \leq q \leq \infty$

$$\|f\|_q \leq CR^{n(\frac{1}{p}-\frac{1}{q})}\|f\|_p.$$

Proof The function $\psi = \phi^{R^{-1}}$ satisfies

$$\|\psi\|_r = CR^{\frac{n}{r'}} \tag{57}$$

for any $r \in [1,\infty]$, where $C = \|\phi\|_r$. Also, using the chain rule

$$\|\nabla\psi\|_1 = R\|\phi\|_1. \tag{58}$$

We know that $f = \psi * f$. In the case of first derivatives, 1. therefore follows from (57) and (24). The general case of 1. then follows by induction.

For 2., let r satisfy $\frac{1}{q} = \frac{1}{p} - \frac{1}{r'}$. Apply Young's inequality obtaining

$$\begin{aligned} \|f\|_q &= \|\psi * f\|_q \\ &\leq \|\psi\|_r\|f\|_p \\ &\lesssim R^{\frac{n}{r'}}\|f\|_p \\ &= R^{n(\frac{1}{p}-\frac{1}{q})}\|f\|_p. \end{aligned}$$

$\square$

We now extend the $L^p \to L^q$ bound to ellipsoids instead of balls, using change of variable. An *ellipsoid* in $\mathbb{R}^n$ is a set of the form

$$E = \{x \in \mathbb{R}^n : \sum_j \frac{|(x-a)\cdot e_j|^2}{r_j^2} \leq 1\} \tag{59}$$

for some $a \in \mathbb{R}^n$ (called the center of E), some choice of orthonormal basis $\{e_j\}$ (the axes) and some choice of positive numbers r_j (the axis lengths). If E and E^* are two ellipsoids, then we say that E^* is *dual* to E if E^* has the same axes as E and reciprocal axis lengths, i.e., if E is given by (59) then E^* should be of the form

$$\{x \in \mathbb{R}^n : \sum_j r_j^2 |(x-b)\cdot e_j|^2 \le 1$$

for some choice of the center point b.

Proposition 5.4 (Bernstein's inequality for an ellipsoid) Suppose that $f \in L^1 + L^2$ and $\hat{f}$ is supported in an ellipsoid E. Then

$$\|f\|_q \lesssim |E|^{(\frac{1}{p}-\frac{1}{q})}\|f\|_p$$

if $1 \le p \le q \le \infty$.

One could similarly extend the first part of Proposition 5.3 to ellipsoids centered at the origin, but the statement is awkward since one has to weight different directions differently, so we ignore this.

Proof Let k be the center of E. Let T be a linear map taking the unit ball onto $E-k$. Let $S = T^{-t}$; thus $T = S^{-t}$ also. Let $f_1(x) = e^{-2\pi i k\cdot x}f(x)$ and $g = f_1 \circ S$, so that

$$\begin{aligned}\hat{g}(\xi) &= |\det S|^{-1}\widehat{f_1}(S^{-t}(\xi)) \\ &= |\det S|^{-1}\hat{f}(S^{-t}(\xi+k)) \\ &= |\det T|\hat{f}(T(\xi)+k)).\end{aligned}$$

Thus $\hat{g}$ is supported in the unit ball, so by Proposition 5.3

$$\|g\|_q \lesssim \|g\|_p.$$

On the other hand,

$$\|g\|_q = |\det S|^{-\frac{1}{q}}\|f\|_q = |\det T|^{\frac{1}{q}}\|f\|_q = |E|^{\frac{1}{q}}\|f\|_q$$

and likewise with q replaced by p. So

$$|E|^{\frac{1}{q}}\|f\|_q \lesssim |E|^{\frac{1}{p}}\|f\|_p$$

as claimed. □

For some purposes one needs a related "pointwise statement", roughly that if supp$\hat{f} \subset E$, then for any dual ellipsoid E^* the values on E^* are controlled by the average over E^*.

To formulate this precisely, let N be a large number and let $\phi(x) = (1+|x|^2)^{-N}$. Suppose an ellipsoid R^* is given. Define $\phi_{E^*}(x) = \phi(T(x-k))$, where k is the center of E^* and T is a selfadjoint linear map taking $E^* - k$ onto the unit ball. If T_1 and T_2

are two such maps, then $T_1 \circ T_2^{-1}$ is an orthogonal transformation, so ϕ_{E^*} is well defined. Essentially, ϕ_{E^*} is roughly equal to 1 on E^* and decays rapidly as one moves away from E^*. We could also write more explicitly

$$\phi_{E^*}(x) = \Big(1 + \sum_j \frac{|(x-k)\cdot e_j|^2}{r_j^2}\Big)^{-N}.$$

Proposition 5.5 Suppose that $f \in L^1 + L^2$ and $\hat{f}$ is supported in an ellipsoid E. Then for any dual ellipsoid E^* and any $z \in E^*$,

$$|f(z)| \leq C_N \frac{1}{|E^*|} \int |f(x)| \phi_{E^*} dx. \tag{60}$$

Proof Assume first that E is the unit ball, and E^* is also the unit ball. Then f is the convolution of itself with a fixed Schwartz function ψ. Accordingly

$$\begin{aligned} |f(z)| &\leq \int |f(x)|\,|\psi(z-x)| dx \\ &\leq C_N \int |f(x)|(1+|z-x|^2)^{-N} \\ &\leq C_N \int |f(x)|(1+|x|^2)^{-N}. \end{aligned}$$

We used the Schwartz space bounds for ψ and that $1+|z-x|^2 \gtrsim 1+|x|^2$ uniformly in x when $|z| \leq 1$. This proves (60) when $E = E^*$ =unit ball.

Suppose next that E is centered at zero but E and E^* are otherwise arbitrary. Let k and T be as above, and consider

$$g(x) = f(T^{-1}x + k)).$$

Its Fourier transform is supported on $T^{-1}E$, and if T maps E^* onto the unit ball, then T^{-1} maps E onto the unit ball. Accordingly,

$$|g(y)| \leq \int \phi(x)|g(x)| dx$$

if $y \in D(0,1)$, so that

$$f(T^{-1}z + k) \leq \int \phi(x)|f(T^{-1}x+k)| dx = |\det T| \int \phi_{E^*}(x)|f(x)| dx$$

by changing variables. Since $|\det T| = \frac{1}{|E^*|}$, we get (60).

If E isn't centered at zero, then we can apply the preceding with f replaced by $e^{-2\pi i k\cdot x} f(x)$ where k is the center of E. $\square$

Remarks 1. Proposition 5.5 is an example of an estimate "with Schwartz tails". It is not possible to make the stronger conclusion that, say, $|f(x)|$ is bounded by the average of f over the double of E^* when $x \in E^*$, even in the one dimensional case with $E = E^* =$ unit interval. For this, consider a fixed Schwartz function g whose Fourier transform is supported in the unit interval $[-1, 1]$. Consider also the functions

$$f_N(x) = (1 - \frac{x^2}{4})^N g(x).$$

Since $\hat{f}_N$ are linear combinations of $\hat{g}$ and its derivatives, they have the same support as $\hat{g}$. Moreover, they converge pointwise boundedly to zero on $[-2, 2]$, except at the origin. It follows that there can be no estimate of the value of f_N at the origin by its average over $[-2, 2]$.

2. All the estimates related to Bernstein's inequality are sharp except for the values of the constants. For example, if E is an ellipsoid, E^* a dual ellipsoid, $N < \infty$, then there is a function f with $\text{supp}\hat{f} \subset E^*$ and with

$$\|f\|_1 \geq |E|, \tag{61}$$

$$|f(x)| \leq C\phi_E(x), \tag{62}$$

where $\phi_E = \phi_E^{(N)}$ was defined above. In the case $E = E^*$ =unit ball this is obvious: take f to be any Schwartz function with Fourier support in the unit ball and with the appropriate L^1 norm. The general case then follows as above by making changes of variable.

The estimates (61) and (62) imply that $\|f\|_p \approx |E|^{\frac{1}{p}}$ for any p, so it follows that Proposition 5.4 is also sharp.

6. Stationary phase

Let ϕ be a real valued C^∞ function, let a be a C_0^∞ function, and define

$$I(\lambda) = \int e^{-\pi i \lambda \phi(x)} a(x) dx.$$

Here λ is a parameter, which we always assume to be positive. The issue is the behavior of the integral $I(\lambda)$ as $\lambda \to +\infty$.

Some general remarks 1. $|I(\lambda)|$ is clearly bounded by a constant depending on a only. One may expect decay as $\lambda \to \infty$, since when λ is large the integral will involve a lot of cancellation.

2. On the other hand, if ϕ is constant then $|I(\lambda)|$ is independent of λ. So one needs to put nondegeneracy hypotheses on ϕ. As it turns out, properties of a are less important.

Note also that one can always cut up a with a partition of unity, which means that the question of how fast $I(\lambda)$ decays can be "localized" to a small neighborhood of a point.

3. Suppose that $\phi_1 = \phi_2 \circ G$ where G is a smooth diffeomorphism. Then

$$\begin{aligned} \int e^{-\pi i\lambda\phi_2(x)}a(x)dx &= \int e^{-\pi i\lambda\phi_1(G^{-1}x)}a(x)dx \\ &= \int e^{-\pi i\lambda\phi_1(y)}a(Gy)d(Gy) \\ &= \int e^{-\pi i\lambda\phi_1(y)}a(Gy)|J_G(y)|dy \end{aligned}$$

where J_G is the Jacobian determinant. The function $y \to a(Gy)|J_G(y)|$ is again C_0^∞, so we see that any bound for $I(\lambda)$ which is independent of choice of a will be "diffeomorphism invariant".

4. Recall from advanced calculus [23] the normal forms for a function near a regular point or a nondegenerate critical point:

<u>Straightening Lemma</u> Suppose $\Omega \subset \mathbb{R}^n$ is open, $f : \Omega \to \mathbb{R}$ is C^∞, $p \in \Omega$ and $\nabla f(p) \neq 0$. Then there are neighborhoods U and V of 0 and p respectively and a C^∞ diffeomorphism $G : U \to V$ with $G(0) = p$ and

$$f \circ G(x) = f(p) + x_n.$$

<u>Morse Lemma</u> Suppose $\Omega \subset \mathbb{R}^n$ is open, $f : \Omega \to \mathbb{R}$ is C^∞, $p \in \Omega, \nabla f(p) = 0$, and suppose that the Hessian matrix $H_f(p) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(p)\right)$ is invertible. Then, for a unique k (= number of positive eigenvalues of H_f; see Lemma 6.3 below) there are neighborhoods U and V of 0 and p respectively and a C^∞ diffeomorphism $G : U \to V$ with $G(0) = p$ and

$$f \circ G(x) = f(p) + \sum_{j=1}^{k} x_j^2 - \sum_{j=k+1}^{n} x_j^2.$$

We consider now $I(\lambda)$ first when a is supported near a regular point, and then when a is supported near a nondegenerate critical point. Degenerate critical points are easy to deal with if $n = 1$, see [33], chapter 8, but in higher dimensions they are much more complicated and only the two-dimensional case has been worked out, see [36].

<u>Proposition 6.1</u> (Nonstationary phase) Suppose $\Omega \subset \mathbb{R}^n$ is open, $\phi : \Omega \to \mathbb{R}$ is C^∞, $p \in \Omega$ and $\nabla\phi(p) \neq 0$. Suppose $a \in C_0^\infty$ has its support in a sufficiently small neighborhood of p. Then

$$\forall N\ \exists C_N : |I(\lambda)| \leq C_N \lambda^{-N},$$

and furthermore C_N depends only on bounds for finitely many derivatives of ϕ and a and a lower bound for $|\nabla\phi(p)|$ (and on N).

Proof The straightening lemma and the calculation in 3. above reduce this to the case $\phi(x) = x_n + c$. In this case, letting $e_n = (0, \ldots, 0, 1)$ we have

$$I(\lambda) = e^{-\pi i \lambda c}\hat{a}(\frac{\lambda}{2}e_n),$$

and this has the requisite decay by Proposition 2.3. $\square$

Now we consider the nondegenerate critical point case, and as in the preceding proof we first consider the normal form.

Proposition 6.2 Let T be a real symmetric invertible matrix with signature σ, let a be C_0^∞ (or just in $\mathcal{S}$), and define

$$I(\lambda) = \int e^{-\pi i \lambda \langle Tx, x\rangle} a(x) dx.$$

Then, for any N,

$$I(\lambda) = e^{-\pi i \frac{\sigma}{4}} |\det T|^{-\frac{1}{2}} \lambda^{-\frac{n}{2}} \left(a(0) + \sum_{j=1}^{N} \lambda^{-j} \mathcal{D}_j a(0) + \mathcal{O}(\lambda^{-(N+1)}) \right).$$

Here $\mathcal{D}_j$ are certain explicit homogeneous constant coefficient differential operators of order $2j$, depending on T only, and the implicit constant depends only on T and on bounds for finitely many Schwartz space seminorms of a.

Proof Essentially this is just another way of looking at the formula for the Fourier transform of an imaginary Gaussian. By Proposition 4.2, the definition of distributional Fourier transform, and the Fourier inversion theorem for a we have

$$I(\lambda) = e^{-\pi i \frac{\sigma}{4}} \lambda^{-\frac{n}{2}} |\det T|^{-\frac{1}{2}} \int \hat{a}(-\xi) e^{\pi i \lambda^{-1} \langle T^{-1}\xi, \xi\rangle} d\xi.$$

We can replace $\hat{a}(-\xi)$ with $\hat{a}(-\xi)$ by making a change of variables, since the Gaussian is even. To understand the resulting integral, use that $\lambda^{-1} \to 0$ as $\lambda \to \infty$, so the Gaussian term is approaching 1. To make this quantitative, use Taylor's theorem for e^{ix}:

$$e^{\pi i \lambda^{-1} \langle T^{-1}\xi, \xi\rangle} = \sum_{j=0}^{N} \frac{(\pi i \lambda^{-1} \langle T^{-1}\xi, \xi\rangle)^j}{j!} + \mathcal{O}(\frac{|\xi|^{2N+2}}{\lambda^{N+1}})$$

uniformly in ξ and λ. Accordingly,

$$\begin{aligned} \int \hat{a}(\xi) e^{\pi i \lambda^{-1} \langle T^{-1}\xi, \xi\rangle} d\xi &= \int \hat{a}(\xi)(1 + \sum_{j=1}^{N} \frac{(\pi i \lambda^{-1} \langle T^{-1}\xi, \xi\rangle)^j}{j!}) d\xi \\ &+ \mathcal{O}\Big(\int |\hat{a}(\xi)| \frac{|\xi|^{2N+2}}{\lambda^{N+1}} \Big). \end{aligned}$$

Now observe that $\int \hat{a}(\xi)d\xi = a(0)$ by the inversion theorem, and similarly

$$\int \hat{a}(\xi)\frac{(\pi i\langle T^{-1}\xi,\xi\rangle)^j}{j!}d\xi$$

is the value at zero of $\mathcal{D}_j a$ for an appropriate differential operator $\mathcal{D}_j$. This gives the result, since

$$\int |\hat{a}(\xi)||\xi|^{2N+2}d\xi$$

is bounded in terms of Schwartz space seminorms of $\hat{a}$, and therefore in terms of derivatives of a. □

Now we consider the case of a general phase function with a nondegenerate critical point. It is clear that this should be reducible to the Gaussian case using the Morse lemma and remark 3. above. However, there is more calculation involved than in the proof of Proposition 5.1, since we need to obtain the correct form for the asymptotic expansion. We recall the following formula which follows from the chain rule:

Lemma 6.3 Suppose that ϕ is smooth, $\nabla\phi(p) = 0$ and G is a smooth diffeomorphism, $G(0) = p$. Then

$$H_{\phi\circ G}(0) = DG(0)^t H_\phi(p)DG(0).$$

Thus $H_\phi(p)$ and $H_{\phi\circ G}(0)$ have the same signature and

$$\det(H_{\phi\circ G}(0)) = J_G(0)^2 \det(H_\phi(p)).$$

Proposition 6.4 Let ϕ be C^∞ and assume that $\nabla\phi(p) = 0$ and $H_\phi(p)$ is invertible. Let σ be the signature of $H_\phi(p)$, and let $\Delta = 2^{-n}|\det(H_\phi(p)|$. Let a be C_0^∞ and supported in a sufficiently small neighborhood of p. Define

$$I(\lambda) = \int e^{-\pi i\lambda\phi(x)dx}a(x).$$

Then, for any N,

$$I(\lambda) = e^{-\pi i\lambda\phi(p)}e^{-\pi i\frac{\sigma}{4}}\Delta^{-\frac{1}{2}}\lambda^{-\frac{n}{2}}\left(a(p) + \sum_{j=1}^{N}\lambda^{-j}\mathcal{D}_j a(p) + \mathcal{O}(\lambda^{-(N+1)})\right).$$

Here $\mathcal{D}_j$ are certain differential operators of order[2] $\leq 2j$, with coefficients depending on ϕ, and the implicit constant depends on ϕ and on bounds for finitely many derivatives of a.

[2]Actually the order is exactly $2j$ but we have no need to know that.

Proof We can assume that $\phi(p) = 0$; else we replace ϕ with $\phi - \phi(p)$. Choose a C^∞ diffeomorphism G by the Morse lemma and apply remark 3. Thus

$$I_\lambda = \int e^{-\pi i \lambda \langle Ty,y\rangle} a(Gy)|J_G(y)|dy,$$

where T is a diagonal matrix with diagonal entries ± 1 and with signature σ. Also $|J_G(0)| = \Delta^{-\frac{1}{2}}$ by Lemma 6.3 and an obvious calculation of the Hessian determinant of the function $y \to \langle Ty, y\rangle$. Let $\mathcal{D}_j$ be associated to this T as in Proposition 5.2 and let $b(y) = a(Gy)|J_G(y)|$. Then

$$I(\lambda) = e^{-\pi i \frac{\sigma}{4}} \lambda^{-\frac{n}{2}} \left(b(0) + \sum_{j=1}^{N} \lambda^{-j} \mathcal{D}_j b(0) + \mathcal{O}(\lambda^{-(N+1)}) \right)$$

by Proposition 6.2. Now $b(0) = |J_G(0)|a(p) = \Delta^{-\frac{1}{2}} a(p)$, so we can write this as

$$I(\lambda) = e^{-\pi i \frac{\sigma}{4}} \Delta^{-\frac{1}{2}} \lambda^{-\frac{n}{2}} \left(a(p) + \sum_{j=1}^{N} \lambda^{-j} \Delta^{\frac{1}{2}} \mathcal{D}_j b(0) + \mathcal{O}(\lambda^{-(N+1)}) \right).$$

Further, it is clear from the chain rule and product rule that any $2j$-th order derivative of b at the origin can be expressed as a linear combination of derivatives of a at p of order $\leq 2j$ with coefficients depending on G, i.e., on ϕ. Otherwise stated, the term $\Delta^{\frac{1}{2}} \mathcal{D}_j b(0)$ can be expressed in the form $\tilde{\mathcal{D}}_j a(p)$, where $\tilde{\mathcal{D}}_j$ is a new differential operator of order $\leq 2j$ with coefficients depending on ϕ. This gives the result. $\square$

In practice, it is often more useful to have estimates for $I(\lambda)$ instead of an asymptotic expansion. Clearly an estimate $|I(\lambda)| \lesssim \lambda^{-\frac{n}{2}}$ could be derived from Proposition 6.4, but one also sometimes needs estimates for the derivatives of $I(\lambda)$ with respect to suitable parameters. For now we just consider the technically easiest case where the parameter is λ itself.

Proposition 6.5 (i) Assume that $\nabla\phi(p) \neq 0$. Then for a supported in a small neighborhood of p, $\left|\frac{d^j I(\lambda)}{d\lambda^j}\right| \leq C_{jN} \lambda^{-N}$ for any N.

(ii)Assume that $\nabla\phi(p) = 0$, and $H_\phi(p)$ is invertible. Then, for a supported in a small neighborhood of p,

$$\left| \frac{d^k}{d\lambda^k} (e^{\pi i \lambda \phi(p)} I(\lambda)) \right| \leq C_k \lambda^{-(\frac{n}{2}+k)}.$$

Proof We only prove (ii), since (i) follows easily from Proposition 6.1 after differentiating under the integral sign as in the proof below. For (ii) we need the following.

Claim. Let $\{\phi_i\}_{i=1}^M$ be real valued smooth functions and assume that $\phi_i(p) = 0$, $\nabla\phi_i(p) = 0$. Let $\Phi = \Pi_{i=1}^M \phi_i$. Then all partial derivatives of Φ of order less than $2M$ also vanish at p.

Proof By the product rule any partial $D^\alpha\Phi$ is a linear combination of terms of the form
$$\prod_{i=1}^{M} D^{\beta_i}\phi_i$$
with $\sum_i \beta_i = \alpha$. If $|\alpha| < 2M$, then some β_i must be less than 2, so by hypothesis all such terms vanish at p.

To prove the proposition, differentiate $I(\lambda)$ under the integral sign obtaining
$$\frac{d^k(e^{\pi i\lambda\phi(p)}I(\lambda))}{d\lambda^k} = (-\pi i)^k \int (\phi(x)-\phi(p))^k a(x) e^{-\pi i\lambda(\phi(x)-\phi(p))}dx.$$
Let $b(x) = (\phi(x)-\phi(p))^k a(x)$. By the above claim all partials of b of order less than $2k$ vanish at p. Now look at the expansion in Proposition 6.4 replacing a with b and setting $N = k-1$. By the claim the terms $\mathcal{D}_j b(p)$ must vanish when $j < k$, as well as $b(p)$ itself. Hence Proposition 5.4 shows that $\frac{d^k}{d\lambda^k}(e^{\pi i\lambda\phi(p)}I(\lambda)) = \mathcal{O}(\lambda^{-(\frac{n}{2}+k)})$ as claimed. $\square$

As an application we estimate the Fourier transform of the surface measure σ on the sphere $S^{n-1} \subset \mathbb{R}^n$. For this and for other similar calculations one wants to work with an integral over a submanifold instead of over $\mathbb{R}^n$. This is not significantly different since it is always possible to work in local coordinates. However, things are easier if one uses the local coordinates as economically as possible. Recall then that if $\phi : \mathbb{R}^n \to \mathbb{R}$ is smooth and if M is a k-dimensional submanifold, $p \in M$, and if $F : U \to M$ is a local coordinate (more precisely the inverse map to one) near p, then $\phi \circ F$ will have a critical point at $F^{-1}p$ iff $\nabla\phi(p)$ is orthogonal to the tangent space to M at p; in particular this is independent of the choice of F.

Notice that $\hat{\sigma}$ is a radial function, because the surface measure is rotation invariant (exercise: prove this rigorously), and is smooth by Proposition 1.3. It therefore suffices to consider $\hat{\sigma}(\lambda e_n)$ where $e_n = (0, \ldots, 0, 1)$ and $\lambda > 0$.

Put local coordinates on the sphere as follows: the first "local coordinate" is the map
$$x \to (x, \sqrt{1-|x|^2}),$$
$$\mathbb{R}^{n-1} \supset D(0, \frac{1}{2}) \to S^{n-1}.$$
The second is the map
$$x \to (x, -\sqrt{1-|x|^2}),$$
$$\mathbb{R}^{n-1} \supset D(0, \frac{1}{2}) \to S^{n-1},$$
and the remaining ones map onto sets whose closures do not contain $\{\pm e_n\}$. Let $\{q_k\}$ be a suitable partition of unity subordinate to this covering by charts. Define $\phi(x) = e_n \cdot x$, $\phi : \mathbb{R}^n \to \mathbb{R}$. Thus the gradient of ϕ is e_n and is normal to the sphere at $\pm e_n$ only.

Now

$$\begin{aligned}
\hat{\sigma}(\lambda e_n) &= \int e^{-2\pi i\lambda e_n\cdot x}d\sigma(x) \\
&= \sum_{j=1}^{k}\int e^{-2\pi i\lambda e_n\cdot x}q_k(x)d\sigma(x) \\
&= \int_{D(0,\frac{1}{2})} e^{-2\pi i\lambda\sqrt{1-|x|^2}}\frac{q_1(x)}{\sqrt{1-|x|^2}}dx + \int_{D(0,\frac{1}{2})} e^{2\pi i\lambda\sqrt{1-|x|^2}}\frac{q_2(x)}{\sqrt{1-|x|^2}}dx \\
&\quad + \sum_{k\geq 3}\int e^{-2\pi i\lambda\phi_k(x)}a_k(x)dx, \qquad (63)
\end{aligned}$$

where the dx integrals are in $\mathbb{R}^{n-1}$, and the phase functions ϕ_k for $k \geq 3$ have no critical points in the support of a_k. The Hessian of $2\sqrt{1-|x|^2}$ at the origin is -2 times the identity matrix, and in particular is invertible. It is also clear that the first and second terms are complex conjugates. We conclude from Proposition 6.5 that

$$\hat{\sigma}(\lambda e_n) = \mathrm{Re}(a(\lambda)e^{2\pi i\lambda}) + y(\lambda)$$

with

$$\frac{d^j a(\lambda)}{d\lambda^j} \lesssim \lambda^{-\frac{n-1}{2}-j}, \qquad (64)$$

$$\frac{d^j y(\lambda)}{d\lambda^j} \lesssim \lambda^{-N} \qquad (65)$$

for any N. In fact σ is real and even and therefore $\hat{\sigma}$ must be real valued. Multiplying y by $e^{-2\pi i\lambda}$ does not affect the estimate (65), so we can absorb y into a and rewrite this as

$$\hat{\sigma}(\lambda e_n) = \mathrm{Re}(a(\lambda)e^{2\pi i\lambda}),$$

where a satisfies (64). Since $\hat{\sigma}$ is radial, we have proved the following.

Corollary 6.6 The function $\hat{\sigma}$ (is a C^∞ function and) satisfies

$$\hat{\sigma}(x) = \mathrm{Re}(a(|x|)e^{2\pi i|x|}), \qquad (66)$$

where for large r

$$\left|\frac{d^j a}{dr^j}\right| \leq C_j r^{-(\frac{n-1}{2}+j)}. \qquad (67)$$

Furthermore, looking at the first term in (63), and using the expansion of Proposition 6.4, with $N = 0$, we can obtain the leading behavior at ∞. Namely, for the first term in (63) at its critical point $x = 0$ we have a phase function $2\sqrt{1-|x|^2}$ with $\phi(0) = 2$, $\Delta = 1$

and signature $-(n-1)$, and an amplitude $q_2(x)(1-|x|^2)^{-1/2}$ which is 1 at the critical point. By Proposition 6.4 the integral is

$$e^{-2\pi i\lambda}e^{\frac{\pi i}{4}(n-1)}\lambda^{-\frac{n-1}{2}}+\mathcal{O}(\lambda^{-\frac{n+1}{2}}).$$

The second term is the complex conjugate and the others are $\mathcal{O}(\lambda^{-N})$ for any N. Hence the quantity (63) is

$$2\lambda^{-\frac{n-1}{2}}\cos(2\pi(\lambda-\frac{n-1}{8}))+\mathcal{O}(\lambda^{-\frac{n+1}{2}})$$

and we have proved

<u>Corollary 6.7</u> For large $|x|$

$$\hat{\sigma}(x)=2|x|^{-\frac{n-1}{2}}\cos(2\pi(|x|-\frac{n-1}{8}))+\mathcal{O}(|x|^{-\frac{n+1}{2}}).$$

<u>Remarks</u> Of course it is possible to consider surfaces other than the sphere, see for example [17], Theorem 7.7.14. The main point in regard to the latter is that the nondegeneracy of the critical points of the phase function which arises when calculating the Fourier transform is equivalent to nonzero Gaussian curvature, so a hypersurface with nonzero Gaussian curvature everywhere behaves essentially the same as the sphere, whereas if there are flat directions the decay becomes weaker. Obtaining derivative bounds like Corollary 6.6 in the above manner requires a somewhat more complicated version of Proposition 6.5 with ϕ and a depending on an auxiliary parameter z, which we now explain without giving the proofs.

Suppose that $\phi(x,z)$ is a C^∞ function of x and z, where $x\in\mathbb{R}^n$, and $z\in\mathbb{R}^k$ should be regarded a parameter. Assume that for a certain p and z_0 we have $\nabla_x\phi(p,z_0)=0$ and that the matrix of second x-partials of ϕ at (p,z_0) is invertible.

1. Prove that there are neighborhoods U of z_0 and V of p and a smooth function $\kappa: U\to V$ with the following property: if $z\in V$, then $\nabla_x\phi(x,z)=0$ if and only if $x=\kappa(z)$.

2. Let $a(x,z)$ be C_0^∞ and supported in a small enough neighborhood of (p,z_0). Define

$$I(\lambda,z)=\int e^{-\pi i\lambda\phi(x,z)}a(x,z)dx.$$

Prove the following:

$$\left|\frac{d^{j+k}(e^{\pi i\lambda\phi(\kappa(z),z)}I(\lambda,z))}{d\lambda^j dz^k}\right|\le C_{jk}\lambda^{-(\frac{n}{2}+j)}.$$

This is the analogue of Proposition 6.5 for general parameters.

7. Restriction problem

We now suppose given a function $f : \mathcal{S}^{n-1} \to \mathbb{C}$ and consider the Fourier transform

$$\widehat{fd\sigma}(\xi) = \int_{S^{n-1}} f(x) e^{-2\pi i x\cdot\xi} d\sigma(x). \tag{68}$$

If f is smooth, then one can use stationary phase to evaluate $\widehat{fd\sigma}$ to any desired degree of precision, just as with Corollary 6.6. In particular this leads to the bound

$$|\widehat{fd\sigma}(\xi)| \leq C\|f\|_{C^2}(1+|\xi|)^{-\frac{n-1}{2}}$$

say, where $\|f\|_{C^2} = \sum_{0\leq|\alpha|\leq 2} \|D^\alpha f\|_{L^\infty}$.

On the other hand there can be no similar decay estimate for functions f which are just bounded. The reason for this is that then there is no distinguished reference point in Fourier space. Thus, if we let $f_k(x) = e^{2\pi i k\cdot x}$ and set $\xi = k$, we have

$$|\widehat{f_k d\sigma}(\xi)| = \sigma(S^{n-1}) \approx 1.$$

Taking a sum of the form $f = \sum_j j^{-2} f_{k_j}$, where $|k_j| \to \infty$ sufficiently rapidly, we obtain a continuous function f such that there is no estimate

$$|\widehat{fd\sigma}(\xi)| \leq C(1+|\xi|)^{-\epsilon}$$

for any $\epsilon > 0$.

On the other hand, if we consider instead L^q norms then the issue of a distinguished origin is no longer relevant. The following is a long standing open problem in the area.

Restriction conjecture (Stein) Prove that if $f \in L^\infty(S^{n-1})$ then

$$\|\widehat{fd\sigma}\|_q \leq C_q\|f\|_\infty \tag{69}$$

for all $q > \frac{2n}{n-1}$.

The example of a constant function shows that the regime $q > \frac{2n}{n-1}$ would be best possible. Namely, Corollary 6.7 implies that $\hat\sigma \in L^q$ if and only if $q \cdot \frac{n-1}{2} > n$.

The corresponding problem for L^2 densities f was solved in the 1970's:

Theorem 7.1 (P. Tomas-Stein) If $f \in L^2(S^{n-1})$ then

$$\|\widehat{fd\sigma}\|_q \leq C\|f\|_{L^2(S^{n-1})} \tag{70}$$

for $q \geq \frac{2n+2}{n-1}$, and this range of q is best possible.

Remarks 1. Notice that the assumptions on q in (70) and (69) are of the form $q > q_0$ or $q \geq q_0$. The reason for this is that there is an obvious estimate

$$\|\widehat{fd\sigma}\|_\infty \leq \|f\|_1$$

by Proposition 1.1, and it follows by the Riesz-Thorin theorem that if (69) or (70) holds for a given q, then it also holds for any larger q.

2. The restriction conjecture (69) is known to be true when $n = 2$; this is due to C. Fefferman and Stein, early 1970's. See [12] and [33].

3. Of course there is a difference in the L^q exponent in (70) and the one which is conjectured for L^∞ densities. Until fairly recently it was unknown (in three or more dimensions) whether the estimate (69) was true even for some q less than the Stein-Tomas exponent $\frac{2n+2}{n-1}$. This was first shown by Bourgain [3], a paper which has been the starting point for a lot of recent work.

4. The fact that $q \geq \frac{2n+2}{n-1}$ is best possible for (70) is due I believe to A. Knapp. We now discuss the construction. Notice that in order to distinguish between L^2 and L^∞ norms, one should use a function f which is highly localized. Next, in view of the nice behavior of rectangles under the Fourier transform discussed e.g. in our section 5, it is natural to take the support of f to be the intersection of S^{n-1} with a small rectangle. Now we set up the proof.

Let

$$C_\delta = \{x \in S^{n-1} : 1 - x \cdot e_n \leq \delta^2\},$$

where $e_n = (0, \ldots, 0, 1)$. Since $|x - e_n|^2 = 2(1 - x \cdot e_n)$, it is easy to show that

$$|x - e_n| \leq C^{-1}\delta \Rightarrow x \in C_\delta \Rightarrow |x - e_n| \leq C\delta \tag{71}$$

for an appropriate constant C. Now let $f = f_\delta$ be the indicator function of C_δ. We calculate $\|f\|_{L^2(S^{n-1})}$ and $\|\widehat{fd\sigma}\|_q$. All constants are of course independent of δ.

In the first place, $\|f\|_2$ is the square root of the measure of C_δ, so by (71) and the dimensionality of the sphere we have

$$\|f\|_2 \approx \delta^{\frac{n-1}{2}}. \tag{72}$$

The support of $fd\sigma$ is contained in the rectangle centered at e_n with length about δ^2 in the e_n direction and length about δ in the orthogonal directions. We look at $\widehat{fd\sigma}$ on the dual rectangle centered at 0. Suppose then that $|\xi_n| \leq C_1^{-1}\delta^{-2}$ and that $|\xi_j| \leq C_1^{-1}\delta^{-1}$ when $j < n$; here C_1 is a large constant. Then

$$|\widehat{fd\sigma}(\xi)| \quad = \quad \left| \int_{C_\delta} e^{-2\pi i x \cdot \xi} d\sigma(x) \right|$$

$$
\begin{aligned}
&= \left| \int_{C_\delta} e^{-2\pi i(x-e_n)\cdot\xi} d\sigma(x) \right| \\
&\geq \int_{C_\delta} \cos(2\pi(x-e_n)\cdot\xi) d\sigma(x).
\end{aligned}
$$

Our conditions on ξ imply if C_1 is large enough that $|(x-e_n)\cdot\xi| \leq \frac{\pi}{3}$, say, for all $x \in C_\delta$. Accordingly,

$$
|\widehat{fd\sigma}(\xi)| \geq \frac{1}{2}|C_\delta| \approx \delta^{n-1}.
$$

Our set of ξ has volume about $\delta^{-(n+1)}$, so we conclude that

$$
\|\widehat{fd\sigma}\|_q \gtrsim \delta^{n-1-\frac{n+1}{q}}.
$$

Comparing this estimate with (72) we find that if (70) holds then

$$
\delta^{n-1-\frac{n+1}{q}} \lesssim \delta^{\frac{n-1}{2}}
$$

uniformly in $\delta \in (0,1]$. Hence $n-1-\frac{n+1}{q} \geq \frac{n-1}{2}$, i.e. $q \geq \frac{2n+2}{n-1}$.

For future reference we record the following variant on the above example: if f is as above and $g = e^{2\pi i x\cdot\eta}Tf$, where $\eta \in \mathbb{R}^n$ and T is a rotation mapping e_n to $v \in S^{n-1}$, then g is supported on

$$
\{x \in S^{n-1} : 1 - x\cdot v \leq \delta^2\},
$$

and

$$
|\widehat{gd\sigma}| \gtrsim \delta^{n-1}
$$

on a cylinder of length $C_1^{-1}\delta^{-2}$ and cross-section radius $C_1^{-1}\delta^{-1}$, centered at η and with the axis parallel to v.

Before giving the proof of Theorem 7.1 we need to discuss convolution of a Schwartz function with a measure, since this wasn't previously considered. Let $\mu \in M(\mathbb{R}^n)$; assume μ has compact support for simplicity, although this assumption is not really needed. Define

$$
\phi * \mu(x) = \int \phi(x-y) d\mu(y).
$$

Observe that $\phi * \mu$ is C^∞, since differentiation under the integral sign is justified as in Lemma 3.1.

It is convenient to use the notation $\check{\mu}$ for $\hat{\mu}(-x)$. We need to extend some of our formulas to the present context. In particular the following extends (28) since if $\mu \in \mathcal{S}$ then the Fourier transform of $\check{\mu}$ is μ by Theorem 3.4:

$$
\widehat{\phi\check{\mu}} = \hat{\phi} * \mu \text{ when } \phi \in \mathcal{S}, \tag{73}
$$

$$
\widehat{\phi\hat{\mu}} = \hat{\phi} * \hat{\mu} \text{ when } \phi \in \mathcal{S}. \tag{74}
$$

Notice that (73) can be interpreted naively: Proposition 1.3 and the product rule imply that $\phi\check{\mu}$ is a Schwartz function. To prove (73), by uniqueness of distributional Fourier transforms it suffices to show that

$$\int \hat{\psi}\phi\check{\mu} = \int \psi(\hat{\phi} * \mu)$$

if ψ is another Schwartz function. This is done as follows. Denote $Tx = -x$, then

$$\begin{aligned}
\int \hat{\psi}(x)\phi(x)\hat{\mu}(-x)dx &= \int \hat{\psi}(-x)\phi(-x)\hat{\mu}(x)dx \\
&= \int ((\hat{\psi}\phi) \circ T)\hat{\mu} \\
&= \int ((\hat{\psi}\phi) \circ T)\hat{}\, d\mu \text{ by the duality relation} \\
&= \int (\widehat{\hat{\psi} \circ T}) * (\widehat{\phi \circ T}) d\mu \\
&= \int \psi * (\hat{\phi} \circ T) d\mu \\
&= \int \int \psi(y)\hat{\phi}(-x+y)dyd\mu(x) \\
&= \int \psi(y)\hat{\phi} * \mu(y)dy.
\end{aligned}$$

For (74), again let ψ be another Schwartz function. Then

$$\begin{aligned}
\int \widehat{\phi\mu}\psi dx &= \int \hat{\psi}\phi d\mu \text{ by the duality relation} \\
&= \int \widehat{\check{\phi} * \psi} d\mu \\
&= \int (\check{\phi} * \psi)\hat{\mu} dx \text{ by the duality relation} \\
&= \int (\hat{\phi} * \hat{\mu})\psi dx.
\end{aligned}$$

The last line may be seen by writing out the definition of the convolution and using Fubini's theorem. Since this worked for all $\psi \in \mathcal{S}$, we get (74).

<u>Lemma 7.2</u> Let $f, g \in \mathcal{S}$, and let μ be a (say) compactly supported measure. Then

$$\int \hat{f}\overline{\hat{g}} d\mu = \int (\hat{\mu} * \overline{g}) \cdot f dx. \tag{75}$$

<u>Proof</u> Recall that

$$\hat{\hat{g}} = \overline{\hat{g}},$$

so that

$$\hat{\overline{\hat{g}}} = \hat{\hat{\hat{g}}} = \overline{g}$$

by the inversion theorem. Now apply the duality relation and (74), obtaining

$$\begin{aligned} \int \hat{f}\overline{\hat{g}} d\mu &= \int f \cdot (\overline{\hat{g}}\mu)\hat{}\, dx \\ &= \int f \cdot (\overline{g} * \hat{\mu}) dx \end{aligned}$$

as claimed. □

Lemma 7.3 Let μ be a finite positive measure. The following are equivalent for any q and any C.

1. $\|\widehat{fd\mu}\|_q \le C\|f\|_2$, $f \in L^2(d\mu)$.
2. $\|\hat{g}\|_{L^2(d\mu)} \le C\|g\|_{q'}$, $g \in \mathcal{S}$.
3. $\|\hat{\mu} * f\|_q \le C^2\|f\|_{q'}$, $f \in \mathcal{S}$.

Proof Let $g \in \mathcal{S}$, $f \in L^2(d\mu)$. By the duality relation

$$\int \hat{g} f d\mu = \int \widehat{fd\mu} \cdot g dx. \tag{76}$$

If 1. holds, then the right side of (76) is $\le \|g\|_{q'}\|\widehat{fd\mu}\|_q \le C\|g\|_{q'}\|f\|_{L^2(d\mu)}$ for any $f \in L^2(d\mu)$. Hence so is the left side. This proves 2. by duality. If 2. holds then the left side is $\le \|\hat{g}\|_{L^2(d\mu)}\|f\|_{L^2(d\mu)} \le C\|g\|_{q'}\|f\|_{L^2(d\mu)}$ for $g \in \mathcal{S}$. Hence so is the right side. Since $\mathcal{S}$ is dense in $L^{q'}$, this proves 1. by duality.

If 3. holds, then the right side of (75) is $\le C^2\|f\|_{q'}^2$ when $f = g \in \mathcal{S}$. Hence so is the left side, which proves 2. If 2. holds then, for any $f, g \in \mathcal{S}$, using also the Schwartz inequality the left side of (75) is $\le C^2\|f\|_{q'}\|g\|_{q'}$. Hence the right side of (75) is also $\le C^2\|f\|_{q'}\|g\|_{q'}$, which proves 3. by duality. □

Remark One can fit lemma 7.3 into the abstract setup

$$T : L^2 \to L^q \Leftrightarrow T^* : L^{q'} \to L^2 \Leftrightarrow TT^* : L^{q'} \to L^q.$$

This is the standard way to think about the lemma, although it is technically a bit easier to present the proof in the above ad hoc manner. Namely, if T is the operator $f \to \widehat{fd\sigma}$ then T^* is the operator $f \to \hat{f}$, where we regard $\hat{f}$ as being defined on the measure space associated to μ, and TT^* is convolution with $\hat{\mu}$.

Proof of Theorem 7.1 We will not give a complete proof; we only prove (70) when $q > \frac{2n+2}{n-1}$ instead of $\ge$. For the endpoint, see for example [35], [9], [32], [33].

We will show that if $q > \frac{2n+2}{n-1}$, then

$$\|\hat{\sigma} * f\|_q \le C_q\|f\|_{q'}, \tag{77}$$

which suffices by Lemma 7.3.

The relevant properties of σ will be

$$\sigma(D(x,r)) \lesssim r^{n-1}, \tag{78}$$

which reflects the $n-1$-dimensionality of the sphere, and the bound

$$|\hat{\sigma}(\xi)| \lesssim (1+|\xi|)^{-\frac{n-1}{2}} \tag{79}$$

from Corollary 6.6.

Let ϕ be a C^∞ function with the following properties:

$$\operatorname{supp}\phi \subset \{x : \frac{1}{4} \leq x \leq 1\},$$

$$\text{if } |x| \geq 1 \text{ then } \sum_{j\geq 0} \phi(2^{-j}x) = 1.$$

Such a function may be obtained as follows: let χ be a C^∞ function which is equal to 1 when $|x| \geq 1$ and to 0 when $|x| \leq \frac{1}{2}$, and let $\phi(x) = \chi(2x) - \chi(x)$.

We now cut up $\hat{\sigma}$ as follows:

$$\hat{\sigma} = K_{-\infty} + \sum_{j=0}^{\infty} K_j,$$

where

$$K_j(x) = \phi(2^{-j}x)\hat{\sigma}(x),$$

$$K_{-\infty}(x) = (1 - \sum_{j=0}^{\infty} \phi(2^{-j}x))\hat{\sigma}.$$

Then $K_{-\infty}$ is a C_0^∞ function, so

$$\|K_{-\infty} * f\|_q \lesssim \|f\|_p$$

by Young's inequality, provided $q \geq p$. In particular, since $q > 2$ we may take $p = q'$.

We now consider the terms in the sum. The logic will be that we estimate convolution with K_j as an operator from L^1 to L^∞ and from L^2 to L^2, and then we use Riesz-Thorin. We have

$$\|K_j\|_\infty \lesssim 2^{-j\frac{n-1}{2}}$$

by (79). Using the trivial bound $\|K_j * f\|_\infty \leq \|K_j\|_\infty \|f\|_1$ we conclude our $L^1 \to L^\infty$ bound,

$$\|K_j * f\|_\infty \lesssim 2^{-j\frac{n-1}{2}} \|f\|_1. \tag{80}$$

On the other hand, we can use (78) to estimate $\widehat{K_j}$. Namely, let $\psi = \hat{\phi}$. Note also that $\hat{\sigma} = \check{\sigma}$, since σ and therefore $\hat{\sigma}$ are invariant under the reflection $x \to -x$. Accordingly, we have

$$\widehat{K_j} = \psi^{2^{-j}} * \sigma,$$

using (73) and the fact that $\widehat{\phi_\epsilon} = \hat{\phi}^\epsilon$. Since $\psi \in \mathcal{S}$, it follows that

$$|\widehat{K_j}(\xi)| \leq C_N 2^{jn} \int (1 + 2^j|\xi - \eta|)^{-N} d\sigma(\eta)$$

for any fixed $N < \infty$. Therefore

$$\begin{aligned}
|\widehat{K_j}(\xi)| &\leq C_N 2^{jn} \Bigg(\int_{D(\xi, 2^{-j})} (1 + 2^j|\xi - \eta|)^{-N} d\sigma(\eta) \\
&\quad + \sum_{k \geq 0} \int_{D(\xi, 2^{k+1-j}) \setminus D(\xi, 2^{k-j})} (1 + 2^j|\xi - \eta|)^{-N} d\sigma(\eta) \Bigg) \\
&\leq C_N 2^{jn} \left(\sigma(D(\xi, 2^{-j})) + \sum_{k \geq 0} 2^{-Nk} \sigma(D(\xi, 2^{k+1-j}) \setminus D(\xi, 2^{k-j})) \right) \\
&\lesssim 2^{jn} \left(2^{-j(n-1)} + \sum_{k \geq 0} 2^{-Nk} 2^{(n-1)(k-j)} \right) \\
&\lesssim 2^j,
\end{aligned}$$

where we used (78) at the next to last line, and at the last line we fixed N to be equal to n and summed a geometric series. Thus

$$\|\widehat{K_j}\|_\infty \lesssim 2^j. \tag{81}$$

Now we mention the trivial but important fact that

$$\|K * f\|_2 \leq \|\hat{K}\|_\infty \|f\|_2 \tag{82}$$

if say K and f are in $\mathcal{S}$. This follows since

$$\begin{aligned}
\|K * f\|_2 &= \|\widehat{K * f}\|_2 \\
&= \|\hat{K}\hat{f}\|_2 \\
&\leq \|\hat{K}\|_\infty \|\hat{f}\|_2 \\
&= \|\hat{K}\|_\infty \|f\|_2.
\end{aligned}$$

Combining (81) and (82) we conclude

$$\|K_j * f\|_2 \lesssim 2^j \|f\|_2. \tag{83}$$

Accordingly, by (80), (83) and Riesz-Thorin we have

$$\|K_j * f\|_q \lesssim 2^{j\theta} 2^{-j\frac{n-1}{2}(1-\theta)} \|f\|_{q'}$$

if $\frac{\theta}{2} + \frac{1-\theta}{\infty} = \frac{1}{q}$. This works out to

$$\|K_j * f\|_q \lesssim 2^{j(\frac{n+1}{q} - \frac{n-1}{2})} \|f\|_{q'} \tag{84}$$

for any $q \in [2, \infty]$. If $q > \frac{2n+2}{n-1}$ then the exponent $\frac{n+1}{q} - \frac{n-1}{2}$ is negative, so we conclude that

$$\sum_j \|K_j * f\|_q \lesssim \|f\|_{q'}.$$

Since f is a Schwartz function, the sum

$$K_{-\infty} * f + \sum_j K_j * f$$

is easily seen to converge pointwise to $\hat{\sigma} * f$. We conclude using Fatou's lemma that $\|\hat{\sigma} * f\|_q \lesssim \|f\|_{q'}$, as claimed. □

Further remarks 1. Notice that the L^2 estimate in the preceding argument was based only on dimensionality considerations. This suggests that there should be an L^2 bound for $\widehat{f d\sigma}$ valid under very general conditions.

Theorem 7.4 Let ν be a positive finite measure satisfying the estimate

$$\nu(D(x,r)) \leq Cr^{\alpha}. \tag{85}$$

Then there is a bound

$$\|\widehat{f d\nu}\|_{L^2(D(0,R))} \leq CR^{\frac{n-\alpha}{2}} \|f\|_{L^2(d\nu)}. \tag{86}$$

The proof uses the following "generic" test for L^2 boundedness.

Lemma 7.5 (Schur's test) Let (X, μ) and (Y, ν) be measure spaces, and let $K(x,y)$ be a measurable function on $X \times Y$ with

$$\int_X |K(x,y)| d\mu(x) \leq A \text{ for each } y, \tag{87}$$

$$\int_Y |K(x,y)| d\nu(y) \leq B \text{ for each } x. \tag{88}$$

Define $T_K f(x) = \int K(x,y) f(y) d\nu(y)$. Then for $f \in L^2(d\nu)$ the integral defining $T_K f$ converges a.e. $(d\mu(x))$ and there is an estimate

$$\|T_K f\|_{L^2(d\mu)} \leq \sqrt{AB} \|f\|_{L^2(d\nu)}. \tag{89}$$

Proof It is possible to use Riesz-Thorin here, since (88) implies $\|T_K f\|_\infty \leq B\|f\|_\infty$ and (87) implies $\|T_K f\|_1 \leq A\|f\|_1$.

A more "elementary" argument goes as follows. If a and b are positive numbers then we have

$$\sqrt{ab} = \min_{\epsilon \in (0,\infty)} \frac{1}{2}(\epsilon a + \epsilon^{-1} b), \tag{90}$$

since $\leq$ is the arithmetic-geometric mean inequality and $\geq$ follows by taking $\epsilon = \sqrt{\frac{b}{a}}$.

To prove (89) it suffices to show that if $\|f\|_{L^2(d\mu)} \leq 1$, $\|g\|_{L^2(d\nu)} \leq 1$, then

$$\int\int |K(x,y)||f(x)||g(y)|d\mu(x)d\nu(y) \leq \sqrt{AB}. \tag{91}$$

To show (91), we estimate

$$\begin{aligned}
\int\int\ & |K(x,y)||f(x)||g(y)|d\mu(x)d\nu(y) \\
& = \frac{1}{2}\min_{\epsilon}\Big(\epsilon\int\int |K(x,y)||g(y)|^2 d\mu(x)d\nu(y) \\
& + \epsilon^{-1}\int\int |K(x,y)||f(x)|^2 d\nu(y)d\mu(x)\Big) \\
& \leq \frac{1}{2}\min_{\epsilon}\Big(\epsilon A\int |g(y)|^2 d\nu(y) + \epsilon^{-1}B\int |f(x)|^2 d\mu(x)\Big) \\
& \leq \frac{1}{2}\min_{\epsilon}(\epsilon A + \epsilon^{-1}B) \\
& = \sqrt{AB}.
\end{aligned}$$

□

To prove Theorem 7.4, let ϕ be an even Schwartz function which is ≥ 1 on the unit disc and whose Fourier transform has compact support. (Exercise: show that such a function exists.) In the usual way define $\phi_{R^{-1}}(x) = \phi(R^{-1}x)$. Then

$$\begin{aligned}
\|\widehat{fd\nu}\|_{L^2(D(0,R))} & \leq \|\phi_{R^{-1}}(x)\widehat{fd\nu}(-x)\|_{L^2(dx)} \\
& = \|\widehat{\phi_{R^{-1}}} * (fd\nu)\|_2
\end{aligned}$$

by (73) and Plancherel.

The last line is the $L^2(dx)$ norm of the function

$$\int R^n\hat{\phi}(R(x-y))f(y)d\nu(y).$$

We have estimates

$$\int |R^n\hat{\phi}(R(x-y))|dx = \|\hat{\phi}\|_1 < \infty$$

for each fixed y, by change of variables, and

$$\int |R^n\hat{\phi}(R(x-y))|d\nu(y) \lesssim R^{n-\alpha}$$

for each fixed x, by (85) and the compact support of $\hat{\phi}$. By Lemma 7.5

$$\|\int R^n\hat{\phi}(R(x-y))f(y)d\nu(y)\|_{L^2(dx)} \lesssim R^{\frac{n-\alpha}{2}}\|f\|_{L^2(d\nu)},$$

and the proof is complete. $\square$

2. Another remark is that it is possible to base the whole proof of Theorem 7.1 on the stationary phase asymptotics in section 6, instead of explicitly using the dimensionality of σ. This sort of argument has the obvious advantage that it is more flexible, since it works also in other situations where the "convolution kernel" $\hat{\sigma}(x-y)$ is replaced by a kernel $K(x,y)$ satisfing appropriate conditions. See for example [29], [32], [33]. On the other hand, it is more complicated and is not as relevant in connection with more delicate questions such as the restriction conjecture, which is known to be false in most of the more general situations (see [5], [26], [30]). We give a brief sketch omitting details. The basic result is the so-called variable coefficient Plancherel theorem, due to Hörmander [16].

Let ϕ be a real valued C^∞ function defined on $\mathbb{R}^n \times \mathbb{R}^n$, let $a \in C_0^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ and consider the "oscillatory integral operators"

$$T_\lambda f(x) = \int e^{-\pi i \lambda \phi(x,y)} a(x,y) f(y) dy. \tag{92}$$

Since a has compact support, it is obvious that these map $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n)$ with a norm bound independent of λ, but we want to show that the norm decays in a suitable way as $\lambda \to \infty$. As with the oscillatory integrals of section 6, this will not be the case if ϕ is too degenerate. In the present situation, note that if ϕ depends on x only, then the factor $e^{-\pi i \lambda \phi}$ in (92) may be taken outside the integral sign, so the norm is independent of λ. Similarly, if ϕ depends on y only, then the factor $e^{-\pi i \lambda \phi}$ may be incorporated into f. We conclude in fact that if $\phi(x,y) = a(x) + b(y)$, then $\|T_\lambda\|_{L^2 \to L^2}$ is independent of λ. This strongly suggests that the appropriate nondegeneracy condition should involve the "mixed Hessian"

$$\tilde{H}_\phi = \left(\frac{\partial^2 \phi}{\partial x_i \partial y_j} \right)_{i,j=1}^n$$

since the mixed Hessian vanishes identically if $\phi(x,y) = a(x) + b(y)$.

Theorem A (Hörmander) Assume that

$$\det(\tilde{H}_\phi(x,y)) \neq 0$$

at all points $(x,y) \in \operatorname{supp} a$. Then

$$\|T_\lambda\|_{L^2 \to L^2} \leq C \lambda^{-\frac{n}{2}}.$$

Sketch of proof This is evidently related to stationary phase, but one cannot apply stationary phase directly to the integral (92), since f isn't smooth. Instead, one looks at $T_\lambda T_\lambda^*$ which is an integral operator T_K with kernel

$$K(x,y) = \int e^{-\pi i \lambda (\phi(x,z) - \phi(y,z))} a(x,z) \overline{a(y,z)} dz. \tag{93}$$

The assumption about the mixed Hessian guarantees that the phase function in (93) has no critical points if x and y are close together. Using a version[3] of "nonstationary phase" one can obtain the estimate

$$\forall N\ \exists C_N : |K(x,y)| \leq C_N(1+\lambda|x-y|)^{-N},$$

provided $|x-y|$ is less than a suitable constant. It follows that if a has small support then

$$\int |K(x,y)|dy \lesssim \lambda^{-n}$$

for each fixed x, and similarly

$$\int |K(x,y)|dx \lesssim \lambda^{-n}$$

for each y. Then Schur's test shows that $\|T_\lambda T_\lambda^*\|_{L^2\to L^2} \lesssim \lambda^{-n}$, so $\|T_\lambda\|_{L^2\to L^2} \lesssim \lambda^{-\frac{n}{2}}$. The small support assumption on a can then be removed using a partition of unity. $\Box$.

It is possible to generalize this to the case where the rank of $\tilde{H}_\phi$ is $\geq k$, where $k \in \{1,\ldots,n\}$; just replace the exponent $\frac{n}{2}$ by $\frac{k}{2}$. Furthermore, the compact support assumption on a may be replaced by "proper support" (see the statement below), and finally one can obtain $L^{q'} \to L^q$ estimates by interpolating with the trivial $\|T_\lambda\|_{L^1\to L^\infty} \leq 1$. Here then is the variable coefficient Plancherel, souped up in a manner which makes it applicable in connection with Stein-Tomas. See the references mentioned above.

Theorem B (Hörmander) Assume that a is a C^∞ function supported on the set $\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : |x-y| \leq C\}$ whose all partial derivatives are bounded. Let ϕ be a real valued C^∞ function defined on a neighborhood of $\operatorname{supp} a$, all of whose partial derivatives are bounded, and such that the rank of $\tilde{H}_\phi(x,y)$ is at least k everywhere. Assume furthermore that the sum of the absolute values of the determinants of the k by k minors of $\tilde{H}_\phi$ is bounded away from zero. Then there is a bound

$$\|T_\lambda\|_{L^{q'}\to L^q} \leq C\lambda^{-\frac{k}{q}}$$

when $2 \leq q \leq \infty$.

Now look back at the proof we gave for Theorem 7.1. The main point was to obtain the bound (84). Now, $K_j(x)$ is the real part of

$$\tilde{K}_j(x) \stackrel{def}{=} \phi(2^{-j}x)a(|x|)e^{-2\pi i|x|},$$

where a satisfies the estimates in Corollary 6.6. Accordingly, it suffices to prove (84) with K_j replaced by $\tilde{K}_j$. Let T_j be convolution with $\tilde{K}_j$, and rescale by 2^j; thus we consider the operator

$$f \to T_j(f_{2^j})_{2^{-j}}.$$

[3]One needs something a bit more quantitative than our Proposition 6.1; the necessary lemma is best proved by integration by parts. See for example [32].

This is an integral operator S_j whose kernel is

$$2^{nj}\phi(x-y)a(2^j|x-y|)e^{-2\pi i 2^j|x-y|}.$$

We want to apply Theorem B to S_j; toward this end we make the following observations.

(i) From the estimates in Corollary 6.6, we see that the functions

$$2^{\frac{n-1}{2}j}\phi(x-y)a(2^j|x-y|)$$

have derivative bounds which are independent of j, and clearly they are supported in $\frac{1}{4} \leq |x-y| \leq 1$.

(ii) The mixed Hessian of the function $|x-y|$ has rank $n-1$. This is a calculation which we leave to the reader, just noting that the exceptional direction corresponds to the direction along the line segment $\overline{xy}$.

It follows that the operators $2^{-\frac{n+1}{2}j}S_j$ satisfy the hypotheses of Theorem B with $k = n-1$, uniformly in j, if we take $\lambda = 2^{j+1}$. Accordingly,

$$\|S_j f\|_q \lesssim 2^{(\frac{n+1}{2}-\frac{n-1}{q})j}\|f\|_{q'},$$

and therefore using change of variables

$$2^{-\frac{n}{q}j}\|T_j f\|_q \lesssim 2^{(\frac{n+1}{2}-\frac{n-1}{q})j}2^{-\frac{n}{q'}j}\|f\|_{q'},$$

i.e.

$$\|T_j f\|_q \lesssim 2^{(\frac{n+1}{q}-\frac{n-1}{2})j}\|f\|_{q'},$$

which is (84). $\square$

Exercise: Use Theorem A for an appropriate phase function, and a rescaling argument of the preceding type, to prove the bound

$$\|\hat{f}\|_2 \leq C\|f\|_2.$$

This explains the name "variable coefficient Plancherel theorem".

8. Hausdorff measures

Fix $\alpha > 0$, and let $E \subset \mathbb{R}^n$. For $\epsilon > 0$, one defines

$$H_\alpha^\epsilon(E) = \inf(\sum_{j=1}^{\infty} r_j^\alpha),$$

where the infimum is taken over all countable coverings of E by discs $D(x_j, r_j)$ with $r_j < \epsilon$. It is clear that $H^{\epsilon}_{\alpha}(E)$ increases as ϵ decreases, and we define

$$H_{\alpha}(E) = \lim_{\epsilon \to 0} H^{\epsilon}_{\alpha}(E).$$

It is also clear that $H^{\epsilon}_{\alpha}(E) \leq H^{\epsilon}_{\beta}(E)$ if $\alpha > \beta$ and $\epsilon \leq 1$; thus

$$H_{\alpha}(E) \text{ is a nonincreasing function of } \alpha. \tag{94}$$

<u>Remarks</u> 1. If $H^1_{\alpha}(E) = 0$, then $H_{\alpha}(E) = 0$. This follows readily from the definition, since a covering showing that $H^1_{\alpha}(E) < \delta$ will necessarily consist of discs of radius $< \delta^{\frac{1}{\alpha}}$.

2. It is also clear that $H_{\alpha}(E) = 0$ for all E if $\alpha > n$, since one can then cover $\mathbb{R}^n$ by discs $D(x_j, r_j)$ with $\sum_j r_j^{\alpha}$ arbitrarily small.

<u>Lemma 8.1</u> There is a unique number α_0, called the Hausdorff dimension of E or dim$E$, such that $H_{\alpha}(E) = \infty$ if $\alpha < \alpha_0$ and $H_{\alpha}(E) = 0$ if $\alpha > \alpha_0$.

<u>Proof</u> Define α_0 to be the supremum of all α such that $H_{\alpha}(E) = \infty$. Thus $H_{\alpha}(E) = \infty$ if $\alpha < \alpha_0$, by (94). Suppose $\alpha > \alpha_0$. Let $\beta \in (\alpha_0, \alpha)$. Define $M = 1 + H_{\beta}(E) < \infty$. If $\epsilon > 0$, then we have a covering by discs with $\sum_j r_j^{\beta} \leq M$ and $r_j < \epsilon$. So

$$\sum_j r_j^{\alpha} \leq \epsilon^{\alpha - \beta} \sum_j r_j^{\beta} \leq \epsilon^{\alpha - \beta} M$$

which goes to 0 as $\epsilon \to 0$. Thus $H_{\alpha}(E) = 0$. $\square$.

<u>Further remarks</u> 1. The set function H_{α} may be seen to be countably additive on Borel sets, i.e. defines a Borel measure. See standard references in the area like [6], [10], [25]. This is part of the reason one considers H_{α} instead of, say, H^1_{α}. Notice in this connection that if E and F are disjoint compact sets, then evidently $H_{\alpha}(E \cup F) = H_{\alpha}(E) + H_{\alpha}(F)$. This statement is already false for H^1_{α}.

2. The Borel measure H_n coincides with $\frac{1}{\omega}$ times Lebesgue measure, where ω is the volume of the unit ball. If $\alpha < n$, then H_{α} is non-sigma finite; this follows e.g. by Lemma 8.1, which implies that any set with nonzero Lebesgue measure will have infinite H_{α}-measure.

<u>Examples</u> The canonical example is the usual $\frac{1}{3}$-Cantor set on the line. This has a covering by 2^n intervals of length 3^{-n}, so it has finite $H_{\frac{\log 2}{\log 3}}$-measure. It is not difficult to show that in fact its $H_{\frac{\log 2}{\log 3}}$-measure is nonzero; this can be done geometrically, or one can apply Proposition 8.2 below to the Cantor measure. In particular, the dimension of the Cantor set is $\frac{\log 2}{\log 3}$.

Now consider instead a Cantor set with variable dissection ratios $\{\epsilon_n\}$, i.e. one starts with the interval $[0,1]$, removes the middle ϵ_1 proportion, then removes the middle ϵ_2 proportion of each of the resulting intervals and so forth. If we assume that $\epsilon_{n+1} \leq \epsilon_n$, and let $\epsilon = \lim_{n\to\infty} \epsilon_n$, then it is not hard to show that the dimension of the resulting set E will be $\frac{\log 2}{\log(\frac{2}{1-\epsilon})}$. In particular, if $\epsilon_n \to 0$ then $\dim E = 1$.

On the other hand, $H_1(E)$ will be positive only if $\sum_n \epsilon_n < \infty$, so this gives examples of sets with zero Lebesgue measure but "full" Hausdorff dimension.

There are numerous other notions of dimension. We mention only one of them, the Minkowski dimension, which is only defined for compact sets. Namely, if E is compact then let $E^\delta = \{x \in \mathbb{R}^n : \operatorname{dist}(x,E) < \delta\}$. Let α_0 be the supremum of all numbers α such that, for some constant C,

$$|E^\delta| \geq C\delta^{n-\alpha}$$

for all $\delta \in (0,1]$. Then α_0 is called the lower Minkowski dimension and denoted $d_L(E)$. Let α_1 be the supremum of all numbers α such that, for some constant C,

$$|E^\delta| \geq C\delta^{n-\alpha}$$

for a sequence of δ's which converges to zero. Then α_1 is called the upper Minkowski dimension and denoted $d_U(E)$.

It would also be possible to define these like Hausdorff dimension but restricting to coverings by discs all the same size, namely: define a set S to be *δ-separated* if any two distict points $x, y \in S$ satisfy $|x-y| > \delta$. Let $\mathcal{E}_\delta(E)$ ("*δ-entropy on E*") be the maximal possible cardinality for a δ-separated subset[4] of E. Then it is not hard to show that

$$d_L(E) = \liminf_{\delta\to 0} \frac{\log \mathcal{E}_\delta(E)}{\log \frac{1}{\delta}},$$

$$d_U(E) = \limsup_{\delta\to 0} \frac{\log \mathcal{E}_\delta(E)}{\log \frac{1}{\delta}}.$$

Notice that a countable set may have positive lower Minkowski dimension; for example, the set $\{\frac{1}{n}\}_{n=1}^\infty \cup \{0\}$ has upper and lower Minkowski dimension $\frac{1}{2}$.

If E is a compact set, then let $P(E)$ be the space of the probability measures supported on E.

<u>Proposition 8.2</u> Suppose $E \subset \mathbb{R}^n$ is compact. Assume that there is a $\mu \in P(E)$ with

$$\mu(D(x,r)) \leq Cr^\alpha \tag{95}$$

for a suitable constant C and all $x \in \mathbb{R}^n$, $r > 0$. Then $H_\alpha(E) > 0$. Conversely, if $H_\alpha(E) > 0$, there is a $\mu \in P(E)$ such that (95) holds.

[4] Exercise: show that $\mathcal{E}_\delta(E)$ is comparable to the minimum number of δ-discs required to cover E

Proof The first part is easy: let $\{D(x_j, r_j)\}$ be any covering of E by discs. Then

$$1 = \mu(E) \leq \sum_j \mu(D(x_j, r_j)) \leq C \sum_j r_j^\alpha,$$

which shows that $H_\alpha(E) \geq C^{-1}$.

The proof of the converse involves constructing a suitable measure, which is most easily done using dyadic cubes. Thus we let $\mathcal{Q}_k$ be all cubes of side length $\ell(Q) = 2^{-k}$ whose vertices are at points of $2^{-k}\mathbb{Z}^n$. We can take these to be closed cubes, for definiteness. It is standard to work with these in such contexts because of their nice combinatorics: if $Q \in \mathcal{Q}_k$, then there is a unique $\tilde{Q} \in \mathcal{Q}_{k-1}$ with $Q \subset \tilde{Q}$; furthermore if we fix $Q_1 \in \mathcal{Q}_{k-1}$, then Q_1 is the union of those $Q \in \mathcal{Q}_k$ with $\tilde{Q} = Q_1$, and the union is disjoint except for edges. A *dyadic cube* is a cube which is in $\mathcal{Q}_k$ for some k.

If Q is a dyadic cube, then clearly there is a disc $D(x, r)$ with $Q \subset D(x, r)$ and $r \leq C\ell(Q)$. Likewise, if we fix $D(x, r)$, then there are a bounded number of dyadic cubes $Q_1 \dots Q_C$ with $\ell(Q_j) \leq Cr$ and whose union contains $D(x, r)$. From these properties, it is easy to see that the definition of Hausdorff measure and also the property (95) could equally well be given in terms of dyadic cubes. Thus, except for the values of the constants,

$$\mu \text{ satisfies (95)} \Leftrightarrow \mu(Q) \leq C\ell(Q)^\alpha \text{ for all dyadic cubes } Q.$$

Furthermore, if we define

$$h_\alpha^\epsilon(E) = \inf\Big(\sum_{Q \in \mathcal{F}} \ell(Q)^\alpha : E \subset \bigcup_{Q \in \mathcal{F}} Q\Big),$$

where $\mathcal{F}$ runs over all coverings of E by dyadic cubes of side length $\ell(Q) < \epsilon$, and

$$h_\alpha(E) = \lim_{\epsilon \to 0} h_\alpha^\epsilon(E),$$

then we have

$$C^{-1} H_\alpha^\epsilon(E) \leq h_\alpha^\epsilon(E) \leq C H_\alpha^\epsilon(E),$$

therefore

$$h_\alpha(E) > 0 \Leftrightarrow H_\alpha(E) > 0.$$

We return now to the proof of Proposition 8.2. We may assume that E is contained in the unit cube $[0, 1] \times \ldots \times [0, 1]$. By the preceding remarks and Remark 1. above we may assume that $h_\alpha^1(E) > 0$, and it suffices to find $\mu \in P(E)$ so that $\mu(Q) \leq C\ell(Q)^\alpha$ for all dyadic cubes Q with $\ell(Q) \leq 1$. We now make a further reduction.

Claim It suffices to find, for each fixed $m \in \mathbb{Z}^+$, a positive measure μ with the following properties:

$$\mu \text{ is supported on the union of the cubes } Q \in \mathcal{Q}_m \text{ which intersect } E; \tag{96}$$

$$\|\mu\| \geq C^{-1}; \tag{97}$$

$$\mu(Q) \leq \ell(Q)^{\alpha} \text{ for all dyadic cubes with } \ell(Q) \geq 2^{-m}. \tag{98}$$

Here C is independent of m.

Namely, if this can be done, then denote the measures satisfying (96), (97), (98) by μ_m. (98) implies a bound on $\|\mu_m\|$, so there is a weak* limit point μ. (96) then shows that μ is supported on E, (98) shows that $\mu(Q) \leq \ell(Q)^{\alpha}$ for all dyadic cubes, and (97) shows that $\|\mu\| \geq C^{-1}$. Accordingly, a suitable scalar multiple of μ gives us the necessary probability measure.

There are a number of ways of constructing the measures satisfying (96), (97), (98). Roughly, the issue is that (97) and (98) are competing conditions, and one has to find a measure μ with the appropriate support and with total mass roughly as large as possible given that (97) holds. This can be done for example by using finite dimensional convexity theory (exercise!). We present a different (more constructive) argument taken from [6], Chapter 2.

We fix m, and will construct a finite sequence of measures $\nu_m, \ldots, \nu_0$, in that order; ν_0 will be the measure we want.

Start by defining ν_m to be the unique measure with the following properties.

1. On each $Q \in \mathcal{Q}_m$, ν_m is a scalar multiple of Lebesgue measure.
2. If $Q \in \mathcal{Q}_m$ and $Q \cap E = \emptyset$, then $\nu_m(Q) = 0$.
3. If $Q \in \mathcal{Q}_m$ and $Q \cap E \neq \emptyset$, then $\nu_m(Q) = 2^{-m\alpha}$.

If we set $k = m$, then ν_k has the following properties: it is absolutely continuous to Lebesgue measure, and

$$\nu_k(Q) \leq \ell(Q)^{\alpha} \text{ if } Q \text{ is a dyadic cube with side } 2^{-j},\ k \leq j \leq m; \tag{99}$$

if Q_1 is a dyadic cube of side 2^{-k}, then there is a covering $\mathcal{F}_{Q_1}$ of $Q_1 \cap E$ by dyadic cubes contained in Q_1 such that $\nu_k(Q_1) \geq \sum_{Q \in \mathcal{F}_{Q_1}} \ell(Q)^{\alpha}$. (100)

Assume now that $1 \leq k \leq m$ and we have constructed an absolutely continuous measure ν_k with properties (96), (99) and (100). We will construct ν_{k-1} having these same properties, where in (99) and (100) k is replaced by $k-1$. Namely, to define ν_{k-1} it suffices to define $\nu_{k-1}(Y)$ when Y is contained in a cube $Q \in \mathcal{Q}_{k-1}$. Fix $Q \in \mathcal{Q}_{k-1}$. Consider two cases.

(i) $\nu_k(Q) \leq \ell(Q)^{\alpha}$. In this case we let ν_{k-1} agree with ν_k on subsets of Q.

(ii) $\nu_k(Q) \geq \ell(Q)^{\alpha}$. In this case we let ν_{k-1} agree with $c\nu_k$ on subsets of Q, where c is the scalar $\frac{2^{-(k-1)\alpha}}{\nu_k(Q)}$.

Notice that $\nu_{k-1}(Y) \leq \nu_k(Y)$ for any set Y, and furthermore $\nu_{k-1}(Q) \leq \ell(Q)^\alpha$ if $Q \in \mathcal{Q}_{k-1}$. These properties and (99) for ν_k give (99) for ν_{k-1}, and (96) for ν_{k-1} follows trivially from (96) for ν_k. To see (100) for ν_{k-1}, fix $Q \in \mathcal{Q}_{k-1}$. If Q is as in case (ii), then $\nu_{k-1}(Q) = \ell(Q)^\alpha$, so we can use the covering by the singleton $\{Q\}$. If Q is as in case (i), then for each of the cubes $Q_j \in \mathcal{Q}_k$ whose union is Q we have the covering of $Q_j \cap E$ associated with (100) for ν_k. Since ν_k and ν_{k-1} agree on subsets of Q, we can simply put these coverings together to obtain a suitable covering of $Q \cap E$. This concludes the inductive step from ν_k to ν_{k-1}.

We therefore have constructed ν_0. It has properties (96), (98) (since for ν_0 this is equivalent to (99)), and by (100) and the definition of h^1_α it has property (99). $\square$

Let us now define the *α-dimensional energy* of a (positive) measure μ with compact support[5] by the formula

$$I_\alpha(\mu) = \int\int |x-y|^{-\alpha} d\mu(x) d\mu(y).$$

We always assume that $0 < \alpha < n$. We also define

$$V^\alpha_\mu(y) = \int |x-y|^{-\alpha} d\mu(x).$$

Thus

$$I_\alpha(\mu) = \int V^\alpha_\mu d\mu. \tag{101}$$

The "potential" V^α_μ is very important in other contexts (namely elliptic theory, since it is harmonic away from $\operatorname{supp}\mu$ when $\alpha = n-2$) but less important than the energy here. Nevertheless we will use it in a technical way below. Notice that it is actually the convolution of the function $|x|^{-\alpha}$ with the measure μ.

Roughly, one expects a measure to have $I_\alpha(\mu) < \infty$ if and only if it satisfies (95); this precise statement is false, but we see below that nevertheless the Hausdorff dimension of a compact set can be defined in terms of the energies of measures in $P(E)$.

Lemma 8.3 (i) If μ is a probability measure with compact support satisfying (95), then $I_\beta(\mu) < \infty$ for all $\beta < \alpha$.

(ii) Conversely, if μ is a probability measure with compact support and with $I_\alpha(\mu) < \infty$, then there is another probability measure ν such that $\nu(X) \leq 2\mu(X)$ for all sets X and such that ν satisfies (95).

Proof (i) We can assume that the diameter of the support of μ is ≤ 1. Then

$$V_\mu{}^\beta(x) \lesssim \sum_{j=0}^{\infty} 2^{j\beta} \mu(D(x, 2^{-j})).$$

[5] The compact support assumption is not needed; it is included to simplify the presentation.

Accordingly, if μ satisfies (95), and $\beta < \alpha$, then

$$\begin{aligned} V_\mu{}^\beta(x) &\lesssim \sum_{j=0}^{\infty} 2^{j\beta} 2^{-j\alpha} \\ &\lesssim 1. \end{aligned}$$

It follows by (101) that $I_\alpha(\mu) < \infty$.

(ii) Let F be the set of points x such that $V_\mu^\alpha(x) \leq 2I_\alpha(\mu)$. Then $\mu(F) \geq \frac{1}{2}$ by (101). Let χ_F be the indicator function of F and let $\nu(X) = \mu(X \cap F)/\mu(F)$. We need to show that ν satisfies (95). Suppose first that $x \in F$. If $r > 0$ then

$$r^{-\alpha}\nu(D(x,r)) \leq V_\nu^\alpha(x) \leq 2V_\mu^\alpha(x) \leq 4I_\alpha(\mu).$$

This verifies (95) when $x \in F$. For general x, consider two cases. If r is such that $D(x,r) \cap F = \emptyset$ then evidently $\nu(D(x,r)) = 0$. If $D(x,r) \cap F \neq \emptyset$, let $y \in D(x,r) \cap F$. Then $\nu(D(x,r)) \leq \nu(D(y,2r)) \lesssim r^\alpha$ by the first part of the proof. $\Box$

Proposition 8.4 If E is compact then the Hausdorff dimension of E coincides with the number

$$\sup\{\alpha : \exists \mu \in P(E) \text{ with } I_\alpha(\mu) < \infty\}.$$

Proof Denote the above supremum by s. If $\beta < s$ then by (ii) of Lemma 8.3 E supports a measure with $\mu(D(x,r)) \leq Cr^\beta$. Then by Proposition 8.2 $H_\beta(E) > 0$, so $\beta \leq \dim E$. So $s \leq \dim E$. Conversely, if $\beta < \dim E$ then by Proposition 8.2 E supports a measure with $\mu(D(x,r)) \leq Cr^{\beta+\epsilon}$ for $\epsilon > 0$ small enough. Then $I_\beta(\mu) < \infty$, so $\beta \leq s$, which shows that $\dim E \leq s$. $\Box$

The energy is a quadratic expression in μ and is therefore susceptible to Fourier transform arguments. Indeed, the following formula is essentially just Lemma 7.2 combined with the formula for the Fourier transform of $|x|^{-\alpha}$.

Proposition 8.5 Let μ be a positive measure with compact support and $0 < \alpha < n$. Then

$$\int\int |x-y|^{-\alpha} d\mu(x) d\mu(y) = c_\alpha \int |\hat{\mu}(\xi)|^2 |\xi|^{-(n-\alpha)} d\xi, \tag{102}$$

where $c_\alpha = \frac{\gamma(\frac{n-a}{2})\pi^{a-\frac{n}{2}}}{\gamma(\frac{a}{2})}$.

Proof Suppose first that $f \in L^1$ is real and even, and that $d\mu(x) = \phi(x)dx$ with $\phi \in \mathcal{S}$. Then we have

$$\int f(x-y) d\mu(x) d\mu(y) = \int |\hat{\mu}(\xi)|^2 \hat{f}(\xi) d\xi \tag{103}$$

This is proved like Lemma 7.2 using (73) instead of (74). Now fix ϕ. Then both sides of (103) are easily seen to define continuous linear maps from $f \in L^2$ to $\mathbb{R}$. Accordingly,

(103) remains valid when $f \in L^1 + L^2, \phi \in \mathcal{S}$. Applying Proposition 4.1, we conclude (102) if $d\mu(x) = \phi(x)dx$, $\phi \in \mathcal{S}$. To pass to general measures, we use the following fact.

Lemma 8.6 Let ϕ be any radial decreasing Schwartz function with L^1 norm 1, and let $0 < \alpha < n$. Then

$$\int |x-y|^{-\alpha}\phi(y)dy \lesssim |x|^{-\alpha},$$

where the implicit constant depends only on α, not on the choice of ϕ.

We sketch the proof as follows: one can easily reduce to the case where $\phi = \frac{1}{|D(0,R)|}\chi_{D(0,R)}$, and this case can be done by explicit calculation. $\square$

Now let $\phi(x) = e^{-\pi|x|^2}$. We have then $\phi^\epsilon * \mu \in \mathcal{S}$, so

$$\begin{aligned} &\int\int \left(\int\int |x-y|^{-\alpha}\phi^\epsilon(x-z)\phi^\epsilon(y-w)dxdy\right) d\mu(z)d\mu(w) \\ &= c_\alpha \int |\hat{\mu}(\xi)|^2|\hat{\phi}(\epsilon\xi)|^2|\xi|^{-(n-\alpha)}d\xi. \end{aligned} \tag{104}$$

Now let $\epsilon \to 0$. On the left side of (104), the expression inside the parentheses converges pointwise to $|z-w|^{-\alpha}$ using a minor variant on Lemma 3.2. If $I_\alpha(\mu) < \infty$ then the convergence is dominated in view of the preceding lemma, so the integrals on the left side converge to $I_\alpha(\mu)$. If $I_\alpha(\mu) = \infty$, then this remains true using Fatou's lemma. On the right hand side of (104) we can argue similarly: the integrands converge pointwise to $|\hat{\mu}(\xi)|^2|\xi|^{-(n-\alpha)}$. If $\int |\hat{\mu}(\xi)|^2|\xi|^{-(n-\alpha)}d\xi < \infty$ then the convergence is dominated since the factors $\hat{\phi}(\epsilon\xi)$ are bounded by 1, so the integrals converge to $\int |\hat{\mu}(\xi)|^2|\xi|^{-(n-\alpha)}d\xi$. If $\int |\hat{\mu}(\xi)|^2|\xi|^{-(n-\alpha)}d\xi = \infty$ then this remains true by Fatou's lemma. Accordingly, we can pass to the limit from (104) to obtain the proposition. $\square$

Corollary 8.7 Suppose μ is a compactly supported probability measure on $\mathbb{R}^n$ with

$$|\hat{\mu}(\xi)| \le C|\xi|^{-\beta} \tag{105}$$

for some $0 < \beta < n/2$, or more generally that (105) is true in the sense of L^2 means:

$$\int_{D(0,N)} |\hat{\mu}(\xi)|^2 d\xi \le CN^{n-2\beta}. \tag{106}$$

Then the dimension of the support of μ is at least 2β.

Proof It suffices by Proposition 8.4 to show that if (106) holds then $I_\alpha(\mu) < \infty$ for all $\alpha < 2\beta$. However,

$$\begin{aligned} \int_{|\xi|\ge 1} |\hat{\mu}(\xi)|^2|\xi|^{-(n-\alpha)}d\xi &\lesssim \sum_{j=0}^{\infty} 2^{-j(n-\alpha)} \int_{2^j \le |\xi| \le 2^{j+1}} |\hat{\mu}(\xi)|^2 d\xi \\ &\lesssim \sum_{j=0}^{\infty} 2^{-j(n-\alpha)}2^{j(n-2\beta)} \\ &< \infty \end{aligned}$$

if $\alpha < 2\beta$ and (106) holds. Observe also that the integral over $|\xi| \leq 1$ is finite since $|\hat{\mu}(\xi)| \leq \|\mu\| = 1$. This completes the proof in view of Proposition 8.5. $\square$

One can ask the converse question, whether a compact set with dimension α must support a measure μ with

$$|\hat{\mu}(\xi)| \leq C_\epsilon (1 + |\xi|)^{-\frac{\alpha}{2}+\epsilon} \tag{107}$$

for all $\epsilon > 0$. The answer is (emphatically) no[6]. Indeed, there are many sets with positive dimension which do not support any measure whose Fourier transform goes to zero as $|\xi| \to \infty$. The easiest way to see this is to consider the line segment $E = [0,1] \times \{0\} \subset \mathbb{R}^2$. E has dimension 1, but if μ is a measure supported on E then $\hat{\mu}(\xi)$ depends on ξ_1 only, so it cannot go to zero at ∞. If one considers only the case $n = 1$, this question is related to the classical question of "sets of uniqueness". See e.g. [28], [40]. One can show for example that the standard $\frac{1}{3}$ Cantor set does not support any measure such that $\hat{\mu}$ vanishes at ∞.

Indeed, it is nontrivial to show that a "noncounterexample" exists, i.e. a set E with given dimension α which supports a measure satisfying (107). We describe a construction of such a set due to R. Kaufman in the next section.

As a typical application (which is also important in its own right) we now discuss a special case of Marstrand's projection theorem. Let e be a unit vector in $\mathbb{R}^n$ and $E \subset \mathbb{R}^n$ a compact set. The projection $P_e(E)$ is the set $\{x \cdot e : x \in E\}$. We want to relate the dimensions of E and of its projections. Notice first of all that $\dim P_e E \leq \dim E$; this follows from the definition of dimension and the fact that the projection P_e is a Lipschitz function.

A reasonable example, although not very typical, is a smooth curve in $\mathbb{R}^2$. This is one-dimensional, and most of its projection will be also one-dimensional. However, if the curve is a line, then one of its projections will be just a point.

Theorem 8.8 (Marstrand's projection theorem for one-dimensional projections) Assume that $E \subset \mathbb{R}^n$ is compact and $\dim E = \alpha$. Then

(i) If $\alpha \leq 1$ then for a.e. $e \in S^{n-1}$ we have $\dim P_e E = \alpha$.

(ii) If $\alpha > 1$ then for a.e. $e \in S^{n-1}$ the projection $P_e E$ has positive one-dimensional Lebesgue measure.

Proof If μ is a measure supported on E, $e \in S^{n-1}$, then the projected measure μ_e is the measure on $\mathbb{R}$ defined by

$$\int f d\mu_e = \int f(x \cdot e) d\mu(x)$$

for continuous f. Notice that $\widehat{\mu_e}$ may readily be calculated from this definition:

$$\widehat{\mu_e}(k) = \int e^{-2\pi i k x \cdot e} d\mu(x)$$

[6]On the other hand, if one interprets decay in an L^2 averaged sense the answer becomes yes, because the calculation in the proof of the above corollary is reversible.

$$= \hat{\mu}(ke).$$

Let $\alpha < \dim E$, and let μ be a measure supported on E with $I_\alpha(\mu) < \infty$. We have then

$$\int |\hat{\mu}(ke)|^2 |k|^{-1+\alpha} dk d\sigma(e) < \infty \tag{108}$$

by Proposition 8.5 and polar coordinates. Thus, for a.e. e we have

$$\int |\hat{\mu}(ke)|^2 |k|^{-1+\alpha} dk < \infty.$$

It follows by Proposition 8.5 with $n = 1$ that for a.e. e the projected measure μ_e has finite α-dimensional energy. This and Proposition 8.4 give part (i), since μ_e is supported on the projected set $P_e E$. For part (ii), we note that if $\dim E > 1$ we can take $\alpha = 1$ in (108). Thus $\widehat{\mu_e}$ is in L^2 for almost all e. By Theorem 3.13, this condition implies that μ_e has an L^2 density, and in particular is absolutely continuous with respect to Lebesgue measure. Accordingly $P_e E$ must have positive Lebesgue measure. □

Remark Theorem 8.8 has a natural generalization to k-dimensional instead of 1-dimensional projections, which is proved in the same way. See [10].

9. Sets with maximal Fourier dimension, and distance sets

A. Sets with maximal Fourier dimension

Jarnik's theorem is the following Proposition 9A1. Fix a number $\alpha > 0$, and let

$$E_\alpha = \{x \in \mathbb{R} : \exists \text{ infinitely many rationals } \tfrac{a}{q} \text{ such that } |x - \frac{a}{q}| \le q^{-(2+\alpha)}\}.$$

Proposition 9A1 The Hausdorff dimension of E_α is equal to $\frac{2}{2+\alpha}$.

Proof We show only that $\dim E_\alpha \le \frac{2}{2+\alpha}$. The converse inequality is not much harder (see [10]) but we have no need to give a proof of it since it follows from Theorem 9A2 below using Corollary 8.7.

It suffices to prove the upper bound for $E_\alpha \cap [-N, N]$. Consider the set of intervals $I_{aq} = (\frac{a}{q} - q^{-(2+\alpha)}, \frac{a}{q} + q^{-(2+\alpha)})$, where $0 \le a \le Nq$ are integers. Then

$$\sum_{q > q_0} \sum_a |I_{aq}|^\beta \approx \sum_{q > q_0} q \cdot q^{-\beta(2+\alpha)},$$

which is finite and goes to 0 as $q_0 \to \infty$ if $\beta > \frac{2}{2+\alpha}$. For any given q_0 the set $\{I_{aq} : q > q_0\}$ covers $E_\alpha \cap [-N, N]$, which therefore has $H_\beta(E_\alpha \cap [-N, N]) = 0$ when $\beta > \frac{2}{2+\alpha}$, as claimed. □.

Theorem 9A2 (Kaufman [21]). For any $\alpha > 0$ there is a positive measure μ supported on a subset of E_α such that

$$|\hat{\mu}(\xi)| \le C_\epsilon |\xi|^{-\frac{1}{2+\alpha}+\epsilon} \tag{109}$$

for all $\epsilon > 0$.

This shows then that Corollary 8.7 is best possible of its type.

The proof is most naturally done using periodic functions, so we start with the following general remarks concerning "periodization" and "deperiodization". Let $\mathbb{T}^n$ be the n-torus which we regard as $[0,1] \times \ldots \times [0,1]$ with edges identified; thus a function on $\mathbb{T}^n$ is the same as a function on $\mathbb{R}^n$ periodic for the lattice $\mathbb{Z}^n$.

If $f \in L^1(\mathbb{T}^n)$ then one defines its Fourier coefficients by

$$\hat{f}(k) = \int_{\mathbb{T}^n} f(x) e^{-2\pi i k\cdot x} dx, \quad k \in \mathbb{Z}^n$$

and one also makes the analogous definition for measures. If f is smooth then one has $|\hat{f}(k)| \le C_N |k|^{-N}$ for all N and $\sum_{k\in\mathbb{Z}^n} \hat{f}(k) e^{2\pi i k\cdot x} = f(x)$.

Also, if $f \in L^1(\mathbb{R}^n)$ one defines its periodization by

$$f_{per}(x) = \sum_{\nu\in\mathbb{Z}^n} f(x-\nu).$$

Then $f_{per} \in L^1(\mathbb{T}^n)$, and we have

Lemma 9A3 If $k \in \mathbb{Z}^n$ then $\widehat{f_{per}}(k) = \hat{f}(k)$.

Proof

$$\begin{aligned}
\hat{f}(k) &= \int_{\mathbb{R}^n} f(x) e^{-2\pi i k\cdot x} dx \\
&= \sum_{\nu\in\mathbb{Z}^n} \int_{[0,1]\times\ldots\times[0,1]+\nu} f(x) e^{-2\pi i k\cdot x} dx \\
&= \sum_{\nu\in\mathbb{Z}^n} \int_{[0,1]\times\ldots\times[0,1]} f(x-\nu) e^{-2\pi i k\cdot(x-\nu)} dx \\
&= \int_{[0,1]\times\ldots\times[0,1]} \sum_{\nu\in\mathbb{Z}^n} f(x-\nu) e^{-2\pi i k\cdot(x-\nu)} dx \\
&= \int_{[0,1]\times\ldots\times[0,1]} f_{per}(x) e^{-2\pi i k\cdot x} dx.
\end{aligned}$$

At the last line we used that $e^{-2\pi i k\cdot\nu} = 1$. □

Suppose now that f is a smooth function on $\mathbb{T}^n$; regard it as a periodic function on $\mathbb{R}^n$. Let $\phi \in \mathcal{S}$ and consider the function $F(x) = \phi(x)f(x)$. We have then

$$\begin{aligned}\hat{F}(\xi) &= \sum_{\nu \in \mathbb{Z}^n} \hat{f}(\nu) \int e^{-2\pi i(\xi-\nu)\cdot x} \phi(x) dx \\ &= \sum_{\nu \in \mathbb{Z}^n} \hat{f}(\nu)\hat{\phi}(\xi - \nu).\end{aligned}$$

This formula extends by a limiting argument to the case where the smooth function f is replaced by a measure; we omit this argument[7]. Thus we have the following: let μ be a measure on $\mathbb{T}^n$, let $\phi \in \mathcal{S}$, and define a measure ν on $\mathbb{R}^n$ by

$$d\nu(x) = \phi(x) d\mu(\{x\}), \tag{110}$$

where $\{x\}$ is the fractional part of x. Then for $\xi \in \mathbb{R}^n$

$$\hat{\nu}(\xi) = \sum_{k \in \mathbb{Z}^n} \hat{\mu}(k)\hat{\phi}(\xi - k). \tag{111}$$

A corollary of this formula by simple estimates with absolutely convergent sums, using the Schwartz decay of $\hat{\phi}$, is the following.

<u>Lemma 9A4</u> If μ is a measure on $\mathbb{T}^n$, satisfying

$$|\hat{\mu}(k)| \le C(1 + |k|)^{-\alpha}$$

for a certain $\alpha > 0$, and if $\nu \in M(\mathbb{R}^n)$ is defined by (110), then

$$|\hat{\nu}(\xi)| \le C'(1 + |\xi|)^{-\alpha}.$$

This can be proved by using (111) and considering the range $|\xi - k| \le |\xi|/2$ and its complement separately. The details are left to the reader.

We now start to construct a measure on the 1-torus $\mathbb{T}$, which will be used to prove Theorem 9A2.

Let ϕ be a nonnegative C_0^∞ function on $\mathbb{R}$ supported in $[-1, 1]$ and with $\int \phi = 1$. Define $\phi^\epsilon(x) = \epsilon^{-1}\phi(\epsilon^{-1}x)$ and let Φ^ϵ be the periodization of ϕ^ϵ.

Let $\mathcal{P}(M)$ be the set of prime numbers which lie in the interval $(\frac{M}{2}, M]$. By the prime number theorem, $|\mathcal{P}(M)| \approx \frac{M}{\log M}$ for large M. If $p \in \mathcal{P}(M)$ then the function $\Phi_p^\epsilon(x) \overset{def}{=} \Phi^\epsilon(px)$ is again 1-periodic[8] and we have

$$\widehat{\Phi_p^\epsilon}(k) = \begin{cases} \hat{\phi}(\epsilon\frac{k}{p}) & \text{if } p \mid k, \\ 0 & \text{otherwise.} \end{cases} \tag{112}$$

[7] It is based on the fact that every measure on $\mathbb{T}^n$ is the weak$*$ limit of a sequence of absolutely continuous measures with smooth densities, which is a corollary e.g. of the Stone-Weierstrass theorem.

[8] Of course for fixed p it is p^{-1}-periodic

To see this, start from the formula

$$\widehat{\Phi^\epsilon}(k) = \widehat{\phi^\epsilon}(k) = \widehat{\phi}(\epsilon k)$$

which follows from Lemma 9A3. Thus

$$\Phi_\epsilon(x) = \sum_k \widehat{\phi}(\epsilon k) e^{2\pi i k\cdot x},$$

$$\Phi_p^\epsilon(x) = \sum_k \widehat{\phi}(\epsilon k) e^{2\pi i k p\cdot x},$$

which is equivalent to (112).

Now define

$$F = \frac{1}{|\mathcal{P}(M)|} \sum_{p\in\mathcal{P}(M)} \Phi_p^\epsilon.$$

Then F is smooth, 1-periodic, and $\int_0^1 F = 1$ (cf. (113)). Of course F depends on ϵ and M but we suppress this dependence.

<u>Lemma</u> The Fourier coefficients of F behave as follows:

$$\hat{F}(0) = 1, \tag{113}$$

$$\hat{F}(k) = 0 \text{ if } 0 < |k| \leq \tfrac{M}{2}, \tag{114}$$

for any N there is C_N such that

$$|\hat{F}(k)| \leq C_N \frac{\log|k|}{M}\Big(1 + \frac{\epsilon|k|}{M}\Big)^{-N} \text{ for all } k \neq 0. \tag{115}$$

<u>Proof</u> Both (113) and (114) are selfevident from (112). For (115) we use that a given integer $k > 0$ has at most $C\frac{\log k}{\log M}$ different prime divisors in the interval $(M/2, M]$. Hence, by (112) and the Schwartz decay of $\hat{\phi}$,

$$|\hat{F}(k)| \leq \frac{\log M}{M} \cdot \frac{\log|k|}{\log M} \cdot C_N\Big(1 + \frac{\epsilon|k|}{M}\Big)^{-N}$$

as claimed. □

We now make up our mind to choose $\epsilon = M^{-(1+\alpha)}$, and denote the function F by F_M. Thus we have the following

$$\text{supp} F_M \subset \{x : \; |x - \frac{a}{p}| \leq p^{-(2+\alpha)} \text{ for some } p \in \mathcal{P}_M, a \in [0,p]\}, \tag{116}$$

$$|\widehat{F_M}(k)| \leq C_N \frac{\log|k|}{M}\Big(1 + \frac{|k|}{M^{2+\alpha}}\Big)^{-N}, \tag{117}$$

and furthermore F_M is nonnegative and satisfies (113) and (114).

It is easy to deduce from (117) with $N = 1$ that

$$|\widehat{F_M}(k)| \lesssim |k|^{-\frac{1}{2+|\alpha|}} \log |k| \tag{118}$$

uniformly in M. In view of (116) we could now try to prove the theorem by taking a weak limit of the measures $F_M dx$ and then using Lemma 9A4. However this would not be correct since the set E_α is not closed, and ((116) notwithstanding) there is no reason why the weak limit should be supported on E_α. Indeed, (114) and (113) imply easily that the weak limit is the Lebesgue measure. The following is the standard way of getting around this kind of problem. It has something in common with the classical "Riesz product" construction; see [25].

I. First consider a fixed smooth function ψ on $\mathbb{T}$. We claim that if M has been chosen large enough then

$$|\widehat{\psi F_M}(k) - \widehat{\psi}(k)| \lesssim \begin{cases} \frac{\log|k|}{M}(1 + \frac{|k|}{M^{2+\alpha}})^{-100} & \text{when } |k| \geq \frac{M}{4} \\ M^{-100} & \text{when } |k| \leq \frac{M}{4}. \end{cases} \tag{119}$$

To see this, we write using (111) and (114), (113)

$$\begin{aligned} \widehat{\psi F_M}(k) - \widehat{\psi}(k) &= \textstyle\sum_{l \in \mathbb{Z}} \widehat{\psi}(l) \widehat{F_M}(k-l) - \widehat{\psi}(k) \\ &= \textstyle\sum_{l:|k-l| \geq M/2} \widehat{\psi}(l) \widehat{F_M}(k-l). \end{aligned} \tag{120}$$

Since ψ is smooth, $|\widehat{\psi}(l)| \leq c_N |l|^{-N}$ for any N. Hence the right side of (120) is bounded by

$$C \max_{l:|k-l| \geq M/2} |\widehat{F_M}(k-l)|.$$

Using (117), we can estimate this by

$$C_N \max_{l:|k-l| \geq M/2} \frac{\log|k-l|}{M} \Big(1 + \frac{|k-l|}{M^{2+\alpha}}\Big)^{-N}. \tag{121}$$

For any fixed N the function

$$f(t) = \frac{\log t}{M} \Big(1 + \frac{t}{M^{2+\alpha}}\Big)^{-N}$$

is decreasing for $t \geq M/10$, provided that M is large enough. Thus (121) is bounded by

$$\begin{gathered} C_N \tfrac{\log(M/2)}{M} \Big(1 + \tfrac{M/2}{M^{2+\alpha}}\Big)^{-N} \\ \lesssim M^{-100}, \end{gathered}$$

which proves the second part of (119). To prove the first part, we again use (120) and consider separately the range $|k-l| \geq |k|/2$ and its complement. For $|k-l| \geq |k|/2$ we argue as above, replacing $|k-l|$ by $|k|/2$ in (121). For $|k-l| \le |k|/2$ we have $|l| \ge |k|/2$, hence the estimate follows easily using the decay of the Fourier coefficients of ψ and the fact that $|\widehat{F_M}(k)| \le 1$. The details are left to the reader.

II. Let

$$g(r) = \begin{cases} r^{-\frac{1}{2+\alpha}} \log r & \text{when } r \ge r_0, \\ r_0^{-\frac{1}{2+\alpha}} \log r_0 & \text{when } r \le r_0, \end{cases}$$

where $r_0 > 1$ is chosen so that $g(r) \le 1$ and $g(r)$ is nonincreasing for all r. Then for any $\psi \in C^\infty(\mathbb{T})$, $\epsilon > 0$, and $M_0 > 10r_0$ we can choose N large enough and a rapidly increasing sequence $M_0 < M_1 < M_2 < \ldots < M_N$ so that

$$|\widehat{\psi G}(k) - \widehat{\psi}(k)| \le \epsilon g(|k|), \tag{122}$$

where $G = N^{-1}(F_{M_1} + \ldots + F_{M_N})$. This can be done as follows. Denote $E_M(k) = |\widehat{\psi F_M}(k) - \widehat{\psi}(k)|$ for $M \ge M_1$. We will need the following consequences of (119):

$$E_M(k) \le Cg(|k|) \text{ if } |k| \ge M/4, \tag{123}$$

$$\lim_{|k|\to\infty} \frac{E_M(k)}{g(|k|)} = 0 \text{ for any fixed } M, \tag{124}$$

$$E_M(k) \le CM^{-100} \text{ if } |k| \le M/4, \tag{125}$$

with the constant in (123), (125) independent of k, M.

Fix N so that $\frac{C}{N} < \frac{\epsilon}{100}$. Observe that for all M large enough

$$CM^{-100} < \frac{\epsilon}{100} g(M). \tag{126}$$

We now choose $M_1, M_2, \ldots, M_N$ inductively so that (126) holds, $M_{j+1} > 4M_j$ and

$$\frac{1}{N} \sum_{i=1}^{j} E_i(k) \le \frac{\epsilon}{100} g(|k|) \text{ if } |k| > M_{j+1}, \tag{127}$$

which is possible by (124). We claim that (122) holds for this choice of M_j. To show this, we start with

$$|\widehat{\psi G}(k) - \widehat{\psi}(k)| \le \frac{1}{N} \sum_{i=1}^{N} E_i(k). \tag{128}$$

Assume that $M_j \le |k| \le M_{j+1}$ (the cases $|k| \le M_1$ and $|k| \ge M_N$ are similar and are left to the reader). By (127) we have

$$\frac{1}{N} \sum_{i=1}^{j-1} E_i(k) \le \frac{\epsilon}{100} g(|k|).$$

We also see from (123), (125) that

$$\frac{1}{N}E_j(k) \leq \frac{C}{N}g(|k|) \leq \frac{\epsilon}{100}g(|k|),$$

$$\frac{1}{N}E_{j+1}(k) \leq \frac{C}{N}g(|k|) + \frac{C}{N}M_{j+1}^{-100} \leq \frac{\epsilon}{100}g(|k|) + \frac{C}{N}M_{j+1}^{-100},$$

$$\frac{1}{N}E_i(k) \leq \frac{C}{N}M_i^{-100} \text{ for } j+2 \leq i \leq N.$$

Thus the right side of (128) is bounded by

$$\frac{3\epsilon}{100}g(|k|) + \frac{C}{N}\sum_{i=j+1}^{N} M_i^{-100}.$$

By (126), we can estimate the last sum by

$$\frac{\epsilon}{100N}\sum_{i=j+1}^{N} g(M_i) \leq \frac{\epsilon}{100}g(M_{j+1}) \leq \frac{\epsilon}{100}g(|k|),$$

which proves (122).

We note that the support properties of G are similar to those of the F's. Namely, it follows from (116) that

$$\operatorname{supp} G \subset \{x: \ |x - \frac{a}{p}| \leq p^{-(2+\alpha)} \text{ for some } p \in (\frac{M_1}{2}, M_N), \ a \in [0,p]\}. \tag{129}$$

III. We now construct inductively the functions G_m and H_m, $m = 1, 2, \ldots$, as follows. Let $G_0 \equiv 1$. If G_m has been constructed, we let G_{m+1} be as in step II with $\psi = G_0 G_1 \ldots G_m$, $M_0 \geq 10r_0 + m$, and $\epsilon = 2^{-m-2}$. Then the functions $H_m = G_1 \ldots G_m$ satisfy

$$\frac{1}{2} \leq \widehat{H_m}(0) \leq \frac{3}{2}$$

for each m and moreover the estimate (118) holds also for the H's, i.e.

$$|\widehat{H_m}(k)| \lesssim |k|^{-\frac{1}{2+\alpha}} \log |k|$$

IV. Now let μ be a weak$*$ limit point of the sequence $\{H_m dx\}$. The support of μ will be contained in the intersection of the supports of the $\{G_m\}$, hence by (129) it will be a compact subset of E_α. From step III, we will have $|\hat{\mu}(k)| \lesssim |k|^{-\frac{1}{2+\alpha}} \log |k|$. The theorem now follows by Lemma 9A4. $\square$

B. Distance sets

If E is a compact set in $\mathbb{R}^2$ (or in $\mathbb{R}^n$), the distance set $\Delta(E)$ is defined as

$$\Delta(E) = \{|x-y| : x, y \in E\}.$$

One version of *Falconer's distance set problem* is the conjecture that

$$E \subset \mathbb{R}^2,\ \dim E > 1 \Rightarrow |\Delta(E)| > 0.$$

One can think of this as a version of the Marstrand projection theorem where the nonlinear projection $(x, y) \to |x-y|$ replaces the linear ones. In fact, it is also possible to make the stronger conjecture that the "pinned" distance sets

$$\{|x-y| : y \in E\}$$

should have positive measure for some $x \in E$, or for a set of $x \in E$ with large Hausdorff dimension. This would be analogous to Theorem 8.8 with the nonlinear maps $y \to |x-y|$ replacing the projections P_e.

Alternately, one can consider this problem as a continuous analogue of a well known open problem in discrete geometry (Erdős' distance set problem): prove that for finite sets $F \subset \mathbb{R}^2$ there is a bound $|\Delta(F)| \geq C_\epsilon^{-1}|F|^{1-\epsilon}$, $\epsilon > 0$. The example $F = \mathbb{Z}^2 \cap D(0, N)$, $N \to \infty$ can be used to show that in Erdős' problem one cannot take $\epsilon = 0$, and a related example [11] shows that in Falconer's problem it does not suffice to assume that $H_1(E) > 0$. The current best result on Erdős' problem is $\epsilon = \frac{1}{7}$ due to Solymosi and Tóth [31] (there were many previous contributions), and on Falconer's problem the current best result is $\dim E > \frac{4}{3}$ due to myself [37] using previous work of Mattila [24] and Bourgain [4].

The strongest results on Falconer's problem have been proved using Fourier transforms in a manner analogous to the proof of Theorem 8.8. We describe the basic strategy, which is due to Mattila [24]. Given a measure μ on E, there is a natural way to put a measure on $\Delta(E)$, namely push forward the measure $\mu \times \mu$ by the map $\Delta : (x, y) \to |x-y|$. If one can show that the pushforward measure has an L^2 Fourier transform, then $\Delta(E)$ must have positive measure by Theorem 3.13.

In fact one proceeds slightly differently for technical reasons. Let μ be a measure in $\mathbb{R}^2$, then [24] one associates to it the measure ν defined as follows. Let $\nu_0 = \Delta(\mu \times \mu)$, i.e.

$$\int f d\nu_0 = \int f(|x-y|) d\mu(x) d\mu(y).$$

Observe that

$$\int t^{-\frac{1}{2}} d\nu(t) = I_{\frac{1}{2}}(\mu).$$

Thus if $I_{\frac{1}{2}}(\mu) < \infty$, as we will always assume, then the measure we now define will be in $M(\mathbb{R})$. Namely, let

$$d\nu(t) = e^{i\frac{\pi}{4}} t^{-\frac{1}{2}} d\nu_0(t) + e^{-i\frac{\pi}{4}} |t|^{-\frac{1}{2}} d\nu_0(-t). \tag{130}$$

Since ν_0 is supported on $\Delta(E)$, ν is supported on $\Delta(E) \cup -\Delta(E)$.

Proposition 9B1 (Mattila [24]) Assume that $I_\alpha(\mu) < \infty$ for some $\alpha > 1$. Then the following are equivalent:
(i) $\hat{\nu} \in L^2(\mathbb{R})$,
(ii) the estimate

$$\int_{R=1}^{\infty} (\int |\hat{\mu}(Re^{i\theta})|^2 d\theta)^2 R dR < \infty. \tag{131}$$

Corollary 9B2 [24] Suppose that $\alpha > 1$ is a number with the following property: if μ is a positive compactly supported measure with $I_\alpha(\mu) < \infty$ then

$$\int |\hat{\mu}(Re^{i\theta})|^2 d\theta \leq C_\mu R^{-(2-\alpha)}. \tag{132}$$

Then any compact subset of $\mathbb{R}^2$ with dimension $> \alpha$ must have a positive measure distance set.

Here and below we identify $\mathbb{R}^2$ with $\mathbb{C}$ in the obvious way.

Proof of the corollary Assume $\dim E > \alpha$. Then E supports a measure with $I_\alpha(\mu) < \infty$. We have

$$\begin{aligned} \int_{R=1}^{\infty} (\int |\hat{\mu}(Re^{i\theta})|^2 d\theta)^2 R dR &\lesssim \int_{R=1}^{\infty} (\int |\hat{\mu}(Re^{i\theta})|^2 d\theta) R^{-(2-\alpha)} R dR \\ &< \infty. \end{aligned}$$

On the first line we used (132) to estimate one of the two angular integrals, and the last line then follows by recognizing that the resulting expression corresponds to the Fourier representation of the energy in Proposition 8.5. By Proposition 9B1 $\Delta(E) \cup -\Delta(E)$ supports a measure whose Fourier transform is in L^2, which suffices by Theorem 3.13.$\square$.

At the end of the section we will prove (132) in the easy case $\alpha = \frac{3}{2}$ where it follows from the uncertainty principle; we believe this is due to P. Sjölin. It is known [37] that (132) holds when $\alpha > \frac{4}{3}$, and this is essentially sharp since (132) fails when $\alpha < \frac{4}{3}$. The negative result follows from a variant on the Knapp argument (Remark 4. in section 7); this is due to [24], and is presented also in several other places, e.g. [37]. The positive result requires more sophisticated L^p type arguments.

Before proving the proposition we record a few more formulas. Let σ_R be the angular measure on the circle of radius R centered at zero; thus we are normalizing the arclength measure on this circle to have total mass 2π. Let μ be any measure with compact support. We then have

$$\int |\hat{\mu}(Re^{i\theta})|^2 d\theta = \int \widehat{\sigma_R} * \mu d\mu. \tag{133}$$

This is just one more instance of the formula which first appeared in Lemma 7.2 and was used in the proof of Proposition 8.5. This version is contained in Lemma 7.2 if μ has a Schwartz space density, and a limiting argument like the one in the proof of Proposition 8.4 shows that it holds for general μ. We also record the asymptotics for $\widehat{\sigma_R}$ which of course follow from those for $\widehat{\sigma_1}$ (Corollary 6.7) using dilations. Notice that the passage from σ_1 to σ_R preserves the total mass, i.e. essentially $\sigma_R = (\sigma_1)^R$. We conclude that

$$\widehat{\sigma_R}(x) = 2(R|x|)^{-\frac{1}{2}} \cos(2\pi(R|x| - \frac{1}{8})) + \mathcal{O}((R|x|)^{-\frac{3}{2}}) \tag{134}$$

when $R|x| \geq 1$, and $|\widehat{\sigma_R}|$ is clearly also bounded independently of R.

Proof of the proposition From the definition of ν we have

$$\begin{aligned}
\hat{\nu}(k) &= e^{i\frac{\pi}{4}} \int |x-y|^{-\frac{1}{2}} e^{-2\pi ik|x-y|} d\mu(x)d\mu(y) \\
&+ e^{-i\frac{\pi}{4}} \int |x-y|^{-\frac{1}{2}} e^{2\pi ik|x-y|} d\mu(x)d\mu(y) \\
&= 2\int |x-y|^{-\frac{1}{2}} \cos(2\pi(|k||x-y| - \frac{1}{8}))d\mu(x)d\mu(y).
\end{aligned}$$

On the other hand, by (133) and (134) we have

$$\begin{aligned}
\int |\hat{\mu}(ke^{i\theta})|^2 d\theta &= |k|^{-\frac{1}{2}} \int 2|x-y|^{-\frac{1}{2}} \cos(2\pi(|k||x-y| - \frac{1}{8}))d\mu(x)d\mu(y) \\
&+\mathcal{O}\Big(\int_{|x-y|\geq|k|^{-1}} (|k||x-y|)^{-\frac{3}{2}} d\mu(x)d\mu(y)\Big) \\
&+\mathcal{O}\Big(\int_{|x-y|\leq|k|^{-1}} (|k||x-y|)^{-\frac{1}{2}} d\mu(x)d\mu(y)\Big).
\end{aligned}$$

The last error term arises by comparing $\widehat{\sigma_R}$, which is bounded, to the main term on the right side of (134), which is $\mathcal{O}((R|x|)^{-\frac{1}{2}})$, in the regime $R|x| < 1$. We may combine the two error terms to obtain

$$\begin{aligned}
\int |\hat{\mu}(ke^{i\theta})|^2 d\theta &= |k|^{-\frac{1}{2}} \int 2|x-y|^{-\frac{1}{2}} \cos(2\pi(|k||x-y| - \frac{1}{8}))d\mu(x)d\mu(y) \\
&+\mathcal{O}(\int (|k||x-y|)^{-\alpha} d\mu(x)d\mu(y))
\end{aligned}$$

for any $\alpha \in [\frac{1}{2}, \frac{3}{2}]$. Therefore

$$\hat{\nu}(k) = |k|^{\frac{1}{2}} \int |\hat{\mu}(ke^{i\theta})|^2 d\theta + \mathcal{O}(|k|^{\frac{1}{2}-\alpha} I_\alpha(\mu)).$$

The error term here is evidently bounded by $|k|^{\frac{1}{2}-\alpha} I_\alpha(\mu)$ for any $\alpha \in (1, \frac{3}{2})$, and therefore belongs to $L^2(|k| \geq 1)$. We conclude then that $\hat{\nu}$ belongs to L^2 on $|k| \geq 1$ if and only if $|k|^{\frac{1}{2}} \int |\hat{\mu}(ke^{i\theta})|^2 d\theta$ does. This gives the proposition, since $\hat{\nu}$ (being the Fourier transform of a measure) clearly belongs to L^2 on $[-1, 1]$. $\square$

Proposition 9B3 If $\alpha \geq 1$ and if μ is a positive measure with compact support[9] then

$$\int |\hat{\mu}(Re^{i\theta})|^2 d\theta \lesssim I_\alpha(\mu) R^{-(\alpha-1)},$$

where the implicit constant depends on a bound for the radius of a disc centered at 0 which contains the support of μ. In particular, (132) holds if $\alpha = \frac{3}{2}$.

Corollary 9B4 (originally due to Falconer [11] with a different proof) If $\dim E > \frac{3}{2}$ then the distance set of E has positive measure.

Proofs The corollary is immediate from the proposition and Corollary 9B2. The proof of the proposition is very similar to the proofs of Bernstein's inequality and of Theorem 7.4. We can evidently assume that R is large. Let ϕ be a radial C_0^∞ function whose Fourier transform is ≥ 1 on the support of μ. Let $d\nu(x) = (\widehat{\phi}(x))^{-1} d\mu(x)$. Then it is obvious (from the definition, not the Fourier representation) that $I_\alpha(\nu) \leq I_\alpha(\mu)$. Also $\hat{\mu} = \phi * \hat{\nu}$. Accordingly

$$\begin{aligned}
\int |\hat{\mu}(Re^{i\theta})|^2 d\theta &= \int |\phi * \hat{\nu}(Re^{i\theta})|^2 d\theta \\
&\lesssim \int |\phi(Re^{i\theta} - x)|\, |\hat{\nu}(x)|^2 dx d\theta \\
&= \int |\hat{\nu}(x)|^2 \int |\phi(x - Re^{i\theta})| d\theta dx \\
&\lesssim R^{-1} \int_{|\,|x|-R| \leq C} |\hat{\nu}(x)|^2 dx \\
&\lesssim R^{-1+2-\alpha} \int |x|^{-(2-\alpha)} |\hat{\nu}(x)|^2 dx \\
&\approx R^{1-\alpha} I_\alpha(\mu).
\end{aligned}$$

Here the second line follows by writing

$$|\phi * \hat{\nu}(Re^{i\theta})| \leq \int \sqrt{|\phi(Re^{i\theta} - x)|} \cdot \sqrt{|\phi(Re^{i\theta} - x)|} |\hat{\nu}(x)| dx$$

and applying the Schwartz inequality. The fourth line follows since for fixed x the set of θ where $\phi(x - Re^{i\theta}) \neq 0$ has measure $\lesssim R^{-1}$, and is empty if $|x| - R$ is large. The proof is complete. $\square$

[9] Here, as opposed to in some previous situations, the compact support is important.

Remark The exponent $\alpha - 1$ is of course far from sharp; the sharp exponent is $\frac{\alpha}{2}$ if $\alpha > 1$, $\frac{1}{2}$ if $\alpha \in [\frac{1}{2}, 1]$ and α if $\alpha < \frac{1}{2}$.

Exercise Prove this in the case $\alpha \leq 1$. (This is a fairly hard exercise.)

Exercise Carry out Mattila's construction (formula (130) and the preceding discussion) in the case where μ is a measure in $\mathbb{R}^n$ instead of $\mathbb{R}^2$, and prove analogues of Proposition 9A1, Corollary 9A2, Proposition 9A3. Conclude that a set in $\mathbb{R}^n$ with dimension greater than $\frac{n+1}{2}$ has a positive measure distance set. (See [24]. The dimension result is also due originally to Falconer. The conjectured sharp exponent is $\frac{n}{2}$.)

10. The Kakeya Problem

A Besicovitch set, or a Kakeya set, is a compact set $E \subset \mathbb{R}^n$ which contains a unit line segment in every direction, i.e.

$$\forall e \in S^{n-1}\ \exists x \in \mathbb{R}^n : \ x + te \in E\ \forall t \in [-\frac{1}{2}, \frac{1}{2}]. \tag{135}$$

Theorem 10.1 (Besicovitch, 1920) If $n \geq 2$, then there are Kakeya sets in $\mathbb{R}^n$ with measure zero.

There are many variants on Besicovitch's construction in the literature, cf. [10], Chapter 7, or [39], Section 1.

There is a basic open question about Besicovitch sets which can be stated vaguely as "How small can this really be?" This can be formulated more precisely in terms of fractal dimension. If one uses the Hausdorff dimension, then the main question is the following.

Open question (the Kakeya conjecture) If $E \subset \mathbb{R}^n$ is a Kakeya set, does E necessarily have Hausdorff dimension n?

If $n = 2$ then the answer is yes; this was proved by Davies [8] in 1971. For general n, what is known at present is that $\dim(E) \geq \min(\frac{n+2}{2}, (2-\sqrt{2})(n-4)+3)$; the first bound which is better for $n = 3$ is due to myself [38], and the second one is due to Katz and Tao [19]. Instead of the Hausdorff dimension one can use other notions of dimension, for example the Minkowski dimension defined in Section 8. The current best results for the upper Minkowski dimension are due to Katz, Łaba and Tao [18], [22], [19].

There is also a more quantitative formulation of the problem in terms of the *Kakeya maximal functions*, which are defined as follows. For any $\delta > 0$, $e \in S^{n-1}$ and $a \in \mathbb{R}^n$, let

$$T_e^\delta(a) = \{x \in \mathbb{R}^n : |(x-a) \cdot e| \leq \frac{1}{2}, |(x-a)^\perp| \leq \delta\},$$

where $x^\perp = x - (x \cdot e)e$. Thus $T_e^\delta(a)$ is essentially the δ-neighborhood of the unit line segment in the e direction centered at a. Then the Kakeya maximal function of $f \in$

$L^1_{loc}(\mathbb{R}^n)$ is the function $f^*_\delta : S^{n-1} \to \mathbb{R}$ defined by

$$f^*_\delta(e) = \sup_{a\in\mathbb{R}^n} \frac{1}{|T^\delta_e(a)|} \int_{T^\delta_e(a)} |f|. \tag{136}$$

The issue is to prove a "$\delta^{-\epsilon}$" estimate for f^*_δ, i.e. an estimate of the form

$$\forall\varepsilon\exists C_\varepsilon : \|f^*_\delta\|_{L^p(S^{n-1})} \leq C_\varepsilon \delta^{-\varepsilon}\|f\|_p \tag{137}$$

for some $p < \infty$.

Remarks 1. It is clear from the definition that

$$\|f^*_\delta\|_\infty \leq \|f\|_\infty, \tag{138}$$

$$\|f^*_\delta\|_\infty \leq \delta^{-(n-1)}\|f\|_1. \tag{139}$$

2. If $n \geq 2$ and $p < \infty$, there can be no bound of the form

$$\|f^*_\delta\|_q \leq C\|f\|_p, \tag{140}$$

with C independent of δ. This can be seen as follows. Consider a zero measure Kakeya set E. Let E_δ be the δ-neighborhood of E, and let $f = \chi_{E_\delta}$. Then $f^*_\delta(e) = 1$ for all $e \in S^{n-1}$, so that $\|f^*_\delta\|_q \approx 1$. On the other hand, $\lim_{\delta\to 0} |E_\delta| = 0$, hence $\lim_{\delta\to 0} \|f_\delta\|_p = 0$ for any $p < \infty$.

3. Let $f = \chi_{D(0,\delta)}$. Then for all $e \in S^{n-1}$ the tube $T^\delta_e(0)$ contains $D(0,\delta)$, so that $f^*_\delta(e) = \frac{|D(0,\delta)|}{|T^\delta_e(0)|} \gtrsim \delta$. Hence $\|f^*_\delta\|_p \approx \delta$. However, $\|f\|_p \approx \delta^{n/p}$. This shows that (137) cannot hold for any $p < n$.

Open problem (the Kakeya maximal function conjecture): prove that (137) holds with $p = n$, i.e.

$$\forall\varepsilon\ \exists C_\varepsilon :\ \|f^*_\delta\|_{L^n(S^{n-1})} \leq C_\varepsilon\delta^{-\varepsilon}\|f\|_n. \tag{141}$$

When $n = 2$, this was proved by Córdoba [7] in a somewhat different formulation and by Bourgain [3] as stated. These results are relatively easy; from one point of view, this is because (141) is then an L^2 estimate. In higher dimensions the problem remains open. There are partial results on (141) which can be understood as follows. Interpolating between (139), which is the best possible bound on L^1, and (141) gives a family of conjectured inequalities

$$\|f^*_\delta\|_q \lesssim C_\varepsilon\delta^{-\frac{n}{p}+1-\varepsilon}\|f\|_p, \qquad q = q(p). \tag{142}$$

Note that if (142) holds for some $p_0 > 1$, it also holds for all $1 \leq p \leq p_0$ (again by interpolating with (135)). The current best results in this direction are that (142) holds with $p = \min((n+2)/2, (4n+3)/7)$ and a suitable q [38], [19].

Proposition 10.2 If (137) holds for some $p < \infty$, then Besicovitch sets in $\mathbb{R}^n$ have Hausdorff dimension n.

Remark The inequality

$$|E_\delta| \geq C_\varepsilon^{-1}\delta^\varepsilon \tag{143}$$

for any Kakeya set E follows immediately from (137) by the same argument that showed (140). (143) says that Besicovitch sets in $\mathbb{R}^n$ have lower Minkowski dimension n.

Proof of the proposition Let E be a Besicovitch set. Fix a covering of E by discs $D_j = D(x_j, r_j)$; we can assume that all r_j's are $\leq 1/100$. Let

$$J_k = \{j : \ 2^{-k} \leq r_j \leq 2^{-(k-1)}\}.$$

For every $e \in S^{n-1}$, E contains a unit line segment I_e parallel to e. Let

$$S_k = \{e \in S^{n-1} : \ |I_e \cap \bigcup_{j\in J_k} D_j| \geq \frac{1}{100k^2}\}.$$

Since $\sum_k \frac{1}{100k^2} < 1$ and $\sum_k |I_e \cap \bigcup_{j\in J_k} D_j| \geq |I_e| = 1$, it follows that $\bigcup_{k=1}^\infty S_k = S^{n-1}$.

Let

$$f = \chi_{F_k}, \ F_k = \bigcup_{j\in J_k} D(x_j, 10r_j).$$

Then for $e \in S_k$ we have

$$|T_e^{2^{-k}}(a_e) \cap F_k| \gtrsim \frac{1}{100k^2}|T_e^{2^{-k}}(a_e)|,$$

where a_e is the midpoint of I_e so that $T_e^{2^{-k}}(a_e)$ is a tube of radius 2^{-k} around I_e. Hence

$$\|f^*_{2^{-k}}\|_p \gtrsim k^{-2}\sigma(S_k)^{1/p}. \tag{144}$$

On the other hand, (137) implies that

$$\|f^*_{2^{-k}}\|_p \leq C_\varepsilon 2^{k\varepsilon}\|f\|_p \leq C_\varepsilon 2^{k\varepsilon}(|J_k| \cdot 2^{-(k-1)np})^{1/p}. \tag{145}$$

Comparing (144) and (145), we see that

$$\sigma(S_k) \lesssim 2^{kp\varepsilon - kn}k^{2p}|J_k| \lesssim 2^{-k(n-2p\varepsilon)}|J_k|.$$

Therefore

$$\sum_j r_j^{n-2p\varepsilon} \geq \sum_k 2^{-k(n-2p\varepsilon)}|J_k| \gtrsim \sum_k \sigma(S_k) \gtrsim 1.$$

We have shown that $\sum r_j^\alpha \gtrsim 1$ for any $\alpha < n$, which implies the claimed Hausdorff dimension bound. $\square$

Remark By the same argument as in the proof of Proposition 10.2, (142) implies that the dimension of a Kakeya set in $\mathbb{R}^n$ is at least p.

A. The $\mathbf{n = 2}$ case

Theorem 10.3 If $n = 2$, then there is a bound

$$\|f_\delta^*\|_2 \leq C(\log \frac{1}{\delta})^{1/2} \|f\|_2.$$

We give two different proofs of the theorem. The first one is due to Bourgain [3] and uses Fourier analysis. The second one is due to Córdoba [7] and is based on geometric arguments.

Proof 1 (Bourgain) We can assume that f is nonnegative. Let $\rho_\delta^e(x) = (2\delta)^{-1}\chi_{T_e^\delta(0)}$, then

$$f_\delta^*(e) = \sup_{a \in \mathbb{R}^2} (\rho_\delta^e * f)(a).$$

Let ψ be a nonnegative Schwartz function on $\mathbb{R}$ such that $\widehat{\phi}$ has compact support and $\phi(x) \geq 1$ when $|x| \leq 1$. Define $\psi : \mathbb{R}^2 \to \mathbb{R}$ by

$$\psi(x) = \phi(x_1)\delta^{-1}\phi(\delta^{-1}x_2).$$

Note that $\psi \geq \rho_\delta^e$ when $e = e_1$, so that $f_\delta^*(e_1) \leq \sup_a (\psi * f)(a)$. Similarly

$$f_\delta^*(e) \leq \sup_a (\psi_e * f)(a),$$

where $\psi_e = \psi \circ p_e$ for an appropriate rotation p_e. Hence

$$f_\delta^*(e) \leq \|\psi_e * f\|_\infty \leq \|\widehat{\psi_e}\widehat{f}\|_1 = \int |\widehat{\psi_e}(\xi)| \cdot |\widehat{f}(\xi)| d\xi. \tag{146}$$

By Hölder's inequality,

$$\begin{aligned} &\textstyle\int |\widehat{\psi_e}(\xi))| \, |\widehat{f}(\xi)| d\xi \\ &\leq \Big(\textstyle\int |\widehat{\psi_e}(\xi)| \, |\widehat{f}(\xi)|^2 (1 + |\xi|) d\xi \Big)^{1/2} \Big(\textstyle\int \frac{|\widehat{\psi_e}(\xi)|}{1+|\xi|} d\xi \Big)^{1/2}. \end{aligned} \tag{147}$$

Note that $\widehat{\psi_e} = \widehat{\psi} \circ p_e$ and $\widehat{\psi} = \widehat{\phi}(x_1)\widehat{\phi}(\delta x_2)$, so that $|\widehat{\psi_e}| \lesssim 1$ and $\widehat{\psi}$ is supported on a rectangle R_e of size about $1 \times 1/\delta$. Accordingly,

$$\int \frac{|\widehat{\psi_e}(\xi)|}{1 + |\xi|} d\xi \lesssim \int_{R_e} \frac{d\xi}{1 + |\xi|} \approx \int_1^{1/\delta} \frac{ds}{s} = \log(\frac{1}{\delta}). \tag{148}$$

Using (146), (147) and (148) we obtain

$$\begin{aligned}
\|f_\delta^*\|_2^2 &\lesssim \log(\tfrac{1}{\delta}) \int |\widehat{\psi_e}(\xi)|\, |\widehat{f}(\xi)|^2(1+|\xi|)d\xi\Big)^{1/2} \\
&\lesssim \log(\tfrac{1}{\delta}) \int_{\mathbb{R}^2} |\widehat{f}(\xi)|^2(1+|\xi|) \left(\int_{S^1} |\widehat{\psi_e}(\xi)|de\right) d\xi \\
&\lesssim \log\tfrac{1}{\delta} \int_{\mathbb{R}^2} |\hat{f}(\xi)|^2 d\xi \\
&= \log\tfrac{1}{\delta}\|f\|_2^2.
\end{aligned}$$

Here the third line follows since for fixed ξ the set of $e \in S^{n-1}$ where $\widehat{\psi_e}(\xi) \neq 0$ has measure $\lesssim 1/(1+|\xi|)$. The proof is complete. □

Remark For $n \geq 3$, the same argument shows that

$$\|f_\delta^*\|_2 \lesssim \delta^{-(n-2)/2}\|f\|_2, \tag{149}$$

which is the best possible L^2 bound.

Proof 2 (Córdoba) The proof uses the following duality argument.

Lemma 10.4 Let $1 < p < \infty$, and let p' be the dual exponent of p: $\frac{1}{p} + \frac{1}{p'} = 1$. Suppose that p has the following property: if $\{e_k\} \subset S^{n-1}$ is a maximal δ-separated set, and if $\delta^{n-1}\sum_k y_k^{p'} \leq 1$, then for any choice of points $a_k \in \mathbb{R}^n$ we have

$$\|\sum_k y_k \chi_{T_{e_k}^\delta(a_k)}\|_{p'} \leq A.$$

Then there is a bound

$$\|f_\delta^*\|_{L^p(S^{n-1})} \lesssim A\|f\|_p.$$

Proof. Let $\{e_k\}$ be a maximal δ-separated subset of S^{n-1}. Observe that if $|e-e'| < \delta$ then $f_\delta^*(e) \leq Cf_\delta^*(e')$; this is because any $T_e^\delta(a)$ can be covered by a bounded number of tubes $T_{e'}^\delta(a')$. Therefore

$$\begin{aligned}
\|f_\delta^*\|_p^p &\leq \sum_k \int_{D(e_k,\delta)} |f_\delta^*(e)|^p de \\
&\lesssim (\delta^{n-1}\sum_k |f_\delta^*(e_k)|^p)^{1/p} \\
&= \delta^{n-1}\sum_k y_k |f_\delta^*(e_k)|
\end{aligned}$$

for some sequence y_k with $\sum_k y_k^{p'}\delta^{n-1} = 1$. On the last line we used the duality between l_p and $l_{p'}$. Hence

$$\|f_\delta^*\|_p^p \lesssim \delta^{n-1}\sum_k y_k \frac{1}{|T_{e_k}^\delta(a_k)|}\int_{T_{e_k}^\delta(a_k)} |f|$$

for some choice of $\{a_k\}$. Since $|T^\delta_{e_k}(a_k)| \approx \delta^{n-1}$, it follows that

$$\begin{aligned} \|f^*_\delta\|_p^p &\lesssim \int \left(\sum_k y_k \chi_{T^\delta_{e_k}(a_k)}\right)|f| \\ &\leq \|\sum_k y_k \chi_{T^\delta_{e_k}(a_k)}\|_{p'} \cdot \|f\|_p \text{ (Hölder's inequality)} \\ &\leq A\|f\|_p \end{aligned}$$

as claimed. □

We continue with Córdoba's proof. In view of Lemma 10.4, it suffices to prove that for any sequence $\{y_k\}$ with $\delta \sum y_k^2 = 1$ and any maximal δ-separated subset $\{e_k\}$ of S^1 we have

$$\Big\|\sum_k y_k \chi_{T^\delta_{e_k}(a_k)}\Big\|_2 \lesssim \sqrt{\log\frac{1}{\delta}}. \tag{150}$$

The relevant geometric fact is

$$|T^\delta_{e_k}(a) \cap T^\delta_{e_l}(b)| \lesssim \frac{\delta^2}{|e_k - e_l| + \delta}. \tag{151}$$

Using (151) we estimate

$$\begin{aligned} \|\sum_k y_k \chi_{T^\delta_{e_k}(a_k)}\|_2^2 &= \sum_{k,l} y_k y_l |T^\delta_{e_k}(a_k) \cap T^\delta_{e_l}(a_l)| \\ &\lesssim \sum_{k,l} y_k y_l \frac{\delta^2}{|e_k - e_l| + \delta}. \\ &\lesssim \sum_{k,l} \sqrt{\delta} y_k \sqrt{\delta} y_l \frac{\delta}{|e_k - e_l| + \delta}. \end{aligned} \tag{152}$$

Observe that for fixed k

$$\sum_l \frac{\delta}{|e_l - e_k| + \delta} \lesssim \sum_{l \leq \frac{1}{\delta}} \frac{\delta}{l\delta + \delta} = \sum_{l \leq \frac{1}{\delta}} \frac{1}{l+1} \approx \log\frac{1}{\delta},$$

and similarly for fixed l

$$\sum_k \frac{\delta}{|e_k - e_l| + \delta} \lesssim \log\frac{1}{\delta}.$$

Applying Schur's test (Lemma 7.5) to the kernel $\delta/(|e_k - e_l| + \delta)$ we obtain that

$$\|\sum_k y_k \chi_{T^\delta_{e_k}(a_k)}\|_2^2 \lesssim \log\frac{1}{\delta} \sum_k (\sqrt{\delta} y_k)^2 \lesssim \log\frac{1}{\delta}, \tag{153}$$

which proves (150). $\square$

B. Kakeya Problem vs. Restriction Problem

Recall that the restriction conjecture states that

$$\|\widehat{fd\sigma}\|_p \leq C_p \|f\|_{L^\infty(S^{n-1})} \text{ if } p > \frac{2n}{n-1}.$$

In fact, the stronger estimate

$$\|\widehat{fd\sigma}\|_p \leq C_p \|f\|_{L^p(S^{n-1})} \text{ if } p > \frac{2n}{n-1} \tag{154}$$

can be proved to be formally equivalent, see e.g. [3].

It is known that the restriction conjecture implies the Kakeya conjecture. This is due to Bourgain [3], although a related construction had appeared earlier in [2]; both constructions are variants on the argument in [13].

Proposition 10.5. (Fefferman, Bourgain) If (154) is true then the conjectured bound

$$\|f_\delta^*\|_n \leq C_\varepsilon \delta^{-\varepsilon} \|f\|_n$$

is also true.

Proof We will use Lemma 10.4. Accordingly, we choose a maximal δ-separated set $\{e_j\}$ on S^{n-1}; observe that such a set has cardinality $\approx \delta^{-(n-1)}$. For each j pick a tube $T_{e_j}^\delta(a_j)$, and let τ_j be the cylinder obtained by dilating $T_{e_j}^\delta(a_j)$ by a factor of δ^{-2} around the origin. Thus τ_j has length δ^{-2}, cross-section radius δ^{-1}, and axis in the e_j direction. Also let

$$S_j = \{e \in S^{n-1} : 1 - e.e_j \leq C^{-1}\delta^2\}.$$

Then S_j is a spherical cap of radius approximately $C^{-1}\delta$, centered at e_j. We choose the constant C large enough so that the S_j's are disjoint. Knapp's construction (see chapter 7) gives a smooth function f_j on S^{n-1} such that f_j is supported on S_j and

$$\|f_j\|_{L^\infty(S^{n-1})} = 1,$$

$$|\widehat{f_j d\sigma}| \gtrsim \delta^{n-1} \text{ on } \tau_j.$$

We consider functions of the form

$$f_\omega = \sum_j \omega_j y_j f_j,$$

where y_j are nonnegative coefficients and ω_j are independent random variables taking values ± 1 with equal probability. Since f_j have disjoint supports, we have

$$\begin{aligned} \|f_\omega\|_{L^q(S^{n-1})}^q &= \sum_j \|y_j f_j\|_{L^q(S^{n-1})}^q \\ &\approx \sum_j y_j^q \delta^{n-1}. \end{aligned} \tag{155}$$

On the other hand,

$$\begin{aligned} \mathbb{E}(\|\widehat{f_\omega d\sigma}\|_{L^q(\mathbb{R}^n)}^q) &= \int_{\mathbb{R}^n} \mathbb{E}(|\widehat{f_\omega d\sigma}(x)|^q) dx \\ &\approx \int_{\mathbb{R}^n} (\sum_j y_j^2 |\widehat{f_\omega d\sigma}(x)|^2)^{q/2} dx \quad \text{(Khinchin's inequality)} \\ &\gtrsim \delta^{q(n-1)} \int_{\mathbb{R}^n} |\sum_j y_j^2 \chi_{\tau_j}(x)|^{q/2} dx. \end{aligned} \tag{156}$$

Assume now that (154) is true. Then for any $q > \frac{2n}{n-1}$ it follows from (155) and (156) that

$$\delta^{q(n-1)} \int_{\mathbb{R}^n} |\sum_j y_j^2 \chi_{\tau_j}(x)|^{q/2} dx \lesssim \sum_j y_j^q \delta^{n-1}.$$

Let $z_j = y_j^2$ and $p' = q/2$, then the above inequality is equivalent to the statement

$$\text{if } \delta^{n-1} \sum_j z_j^{q/2} \le 1, \text{ then } \|\sum_j z_j \chi_{\tau_j}\|_{q/2} \lesssim \delta^{-2(n-1)} \tag{157}$$

for any $p' \ge \frac{n}{n-1}$. We now rescale this by δ^2 to obtain

$$\text{if } \delta^{n-1} \sum_j z_j^{p'} \le 1, \text{ then } \|\sum_j z_j \chi_{T_j}\|_{p'} \lesssim \delta^{2(\frac{n}{p'} - (n-1))}.$$

Observe that $\frac{n}{p'} - (n-1) \searrow 0$ as $p' \searrow \frac{n}{n-1}$. Thus for any $\varepsilon > 0$ we have

$$\text{if } \delta^{n-1} \sum_j z_j^{p'} \le 1, \text{ then } \|\sum_j z_j \chi_{T_j}\|_{p'} \lesssim \delta^{-\varepsilon} \tag{158}$$

if p' is close enough to $\frac{n}{n-1}$. By Lemma 10.4, this implies that for any $\varepsilon > 0$

$$\|f_\delta^*\|_p \lesssim \delta^{-\varepsilon} \|f\|_p$$

provided that $p < n$ is close enough to n. Interpolating this with the trivial L^∞ bound, we conclude that

$$\|f_\delta^*\|_n \lesssim \delta^{-\varepsilon} \|f\|_n$$

as claimed. $\square$

We proved that the restriction conjecture is stronger than the Kakeya conjecture. Bourgain [3] partially reversed this and obtained a restriction theorem beyond Stein-Tomas by using a Kakeya set estimate that is stronger than the L^2 bound (153) used in the proof of (149). It is not known whether (either version of) the Kakeya conjecture implies the full restriction conjecture.

Theorem 10.6 (Bourgain [3]) Suppose that we have an estimate

$$\|\sum_j \chi_{T^\delta_{e_j}(a_j)}\|_{q'} \leq C_\varepsilon \delta^{-(\frac{n}{q}-1+\varepsilon)} \tag{159}$$

for any given $\varepsilon > 0$ and for some fixed $q > 2$. Then

$$\|\widehat{fd\sigma}\|_p \leq C_p \|f\|_{L^\infty(S^{n-1})} \tag{160}$$

for some $p < \frac{2n+2}{n-1}$.

Remark The geometrical statement corresponding to (159) is that Kakeya sets in $\mathbb{R}^n$ have dimension at least q.

We will sketch the proof only for $n = 3$. Recall that in $\mathbb{R}^3$ we have the estimates

$$\|\widehat{fd\sigma}\|_4 \lesssim \|f\|_{L^2(S^2)} \tag{161}$$

from the Stein-Tomas theorem, and

$$\|\widehat{fd\sigma}\|_{L^2(D(0,R))} \lesssim R^{1/2}\|f\|_{L^2(S^2)} \tag{162}$$

from Theorem 7.4 with $\alpha = n - 1 = 2$. Interpolating (161) and (162) yields a family of estimates

$$\|f\hat{d}\sigma\|_{L^p(D(0,R))} \lesssim R^{\frac{2}{p}-\frac{1}{2}}\|f\|_{L^2(S^2)} \tag{163}$$

for $2 \leq p \leq 4$. Below we sketch an argument showing that the exponent of R in (163) can be lowered if the L^2 norm on the right side is replaced by the L^∞ norm.

Proposition 10.7 Let $n = 3$, $2 < p < 4$, and assume that (159) holds for some $q > 2$. Then

$$\|\widehat{fd\sigma}\|_{L^p(D(0,R))} \lesssim R^{\alpha(p)}\|f\|_{L^\infty(S^{n-1})},$$

where $\alpha(p) < \frac{2}{p} - \frac{1}{2}$.

This of course implies (160) for all p such that $\alpha(p) \leq 0$; in particular, there are $p < 4$ for which (160) holds.

Heuristic proof of the proposition Assume that $\|f\|_{L^\infty(S^{n-1})} = 1$, and let $\delta = R^{-1}$. We cover S^2 by spherical caps

$$S_j = \{e \in S^2 : 1 - e.e_j \leq \delta\},$$

where $\{e_j\}$ is a maximal $\sqrt{\delta}$-separated set on S^2. Then

$$f = \sum_j f_j,$$

where each f_j is supported on S_j. Let $G = f\hat{d\sigma}$ and $G_j = f_j\hat{d\sigma}$, so that $G = \sum_j G_j$. By the uncertainty principle $|G_j|$ is roughly constant on cylinders of length R and diameter $\sqrt{R}$ pointing in the e_j direction. To simplify the presentation, we now make the assumption[10] that G_j is supported on only one such cylinder τ_j.

Next, we cover $D(0, R)$ with disjoint cubes Q of side $\sqrt{R}$. For each Q we denote by $N(Q)$ the number of cylinders τ_j which intersect it. Note that $G|_Q = \sum_j G_j|_Q$, where we sum only over those j's for which τ_j intersects Q. Using this and (163), we can estimate $\|G\|_{L^p(Q)}$ for $2 \leq p \leq 4$:

$$\begin{aligned} \|G\|_{L^p(Q)} &\lesssim \sqrt{R}^{\frac{2}{p}-\frac{1}{2}} \Big\| \sum_{j:\tau_j \cap Q \neq \emptyset} f_j \Big\|_{L^2(S^2)} \\ &\lesssim \sqrt{R}^{\frac{2}{p}-\frac{1}{2}} (N(Q) \cdot |S_i|)^{1/2} \\ &\approx \delta^{\frac{3}{4}-\frac{1}{p}} N(Q)^{1/2}. \end{aligned} \tag{164}$$

Summing over Q, we obtain

$$\begin{aligned} \|G\|^p_{L^p(D(0,R))} &\lesssim \delta^{\frac{3p}{4}-1} \sum_Q N(Q)^{p/2} \\ &\approx \delta^{\frac{3p}{4}p+\frac{1}{2}} \| \sum_j \chi_{\tau_j} \|^{p/2}_{p/2}. \end{aligned} \tag{165}$$

On the last line we used that

$$\| \sum_j \chi_{\tau_j} \|^{p/2}_{p/2} = \sum_Q N(Q)^{p/2} \cdot |Q| = \delta^{-3/2} \sum_Q N(Q)^{p/2}.$$

We now let $p = 2q'$, where q' is the exponent in (159), and assume that p is sufficiently close to 4 (interpolate (159) with (149) if necessary). We have from (159)

$$\| \sum_j \chi_{T^{\sqrt{\delta}}_{e_j}(a_j)} \|_{q'} \leq C_\varepsilon \sqrt{\delta}^{-(\frac{3}{q}-1+\varepsilon)}.$$

[10] It is because of this assumption that our proof is merely heuristic. Of course the Fourier transform of a compactly supported measure cannot be compactly supported; the rigorous proof uses the Schwartz decay of G_j instead.

Rescaling this inequality by δ^{-1} we obtain

$$\|\sum_j \chi_{\tau_j}\|_{q'} \lesssim \sqrt{\delta}^{-(\frac{3}{q}-1+\varepsilon)}\delta^{-3/q'} = \delta^{-1-\frac{3}{p}-\varepsilon}.$$

We combine this with (164) and conclude that

$$\|G\|^p_{L^p(Q)} \lesssim \delta^{\frac{p}{4}-1}\delta^{-\varepsilon},$$

i.e.,

$$\|f\hat{d}\sigma\|_{L^p(D(0,R))} \lesssim \delta^{\frac{1}{4}-\frac{1}{p}-\varepsilon} = R^{\frac{1}{p}-\frac{1}{4}+\varepsilon},$$

which proves the proposition since $\frac{1}{p}-\frac{1}{4} < \frac{2}{p}-\frac{1}{2}$ if $p < 4$. $\square$

References

[1] W. Beckner, *Inequalities in Fourier analysis*, Ann. of Math. 102 (1975), 159-182.

[2] W. Beckner, A. Carbery, S. Semmes, F. Soria: *A note on restriction of the Fourier transform to spheres*, Bull. London Math. Soc. 21 (1989), 394-398.

[3] J. Bourgain, *Besicovitch type maximal operators and applications to Fourier analysis*, Geometric and Functional Analysis 1 (1991), 147-187.

[4] J. Bourgain, *Hausdorff dimension and distance sets*, Israel J. Math 87(1994), 193-201.

[5] J. Bourgain, *L^p estimates for oscillatory integrals in several variables*, Geometric and Functional Analysis 1(1991), 321-374.

[6] L. Carleson, *Selected Problems on Exceptional Sets*, Van Nostrand Mathematical Studies. No. 13, Van Nostrand Co., Inc., Princeton-Toronto-London 1967.

[7] A. Córdoba, *The Kakeya maximal function and spherical summation multipliers*, Amer. J. Math. 99 (1977), 1-22.

[8] R. O. Davies, *Some remarks on the Kakeya problem*, Proc. Cambridge Phil. Soc. 69(1971), 417-421.

[9] K.M. Davis, Y.C. Chang, *Lectures on Bochner-Riesz Means* London Mathematical Society Lecture Note Series, 114, Cambridge University Press, Cambridge, 1987.

[10] K. J. Falconer, *The geometry of fractal sets*, Cambridge University Press, 1985.

[11] K. J. Falconer, *On the Hausdorff dimension of distance sets*, Mathematika 32(1985), 206-212.

[12] C. Fefferman, *Inequalities for strongly singular convolution operators*, Acta Math. 124(1970), 9-36.

[13] C. Fefferman, *The multiplier problem for the ball*, Ann. Math. 94(1971), 330-336.

[14] G. Folland, *Harmonic Analysis in Phase Space*, Annals of Mathematics Studies, 122. Princeton University Press, Princeton, NJ, 1989.

[15] V. Havin, B. Joricke, *The Uncertainty Principle in Harmonic Analysis*, Springer-Verlag, 1994.

[16] L. Hörmander, *Oscillatory integrals and multipliers on FL^p*, Ark. Mat. 11 (1973), 1-11.

[17] L. Hörmander, *The Analysis of Linear Partial Differential Operators*, volume 1, 2nd edition, Springer Verlag 1990.

[18] N. Katz, I. Łaba, T. Tao, *An improved bound on the Minkowski dimension of Besicovitch sets in* $\mathbb{R}^3$, Annals of Math. 152(2000), 383-446.

[19] N. Katz, T. Tao, *New bounds for Kakeya sets*, preprint, 2000.

[20] Y. Katznelson, *An Introduction to Harmonic Analysis* 2nd edition. Dover Publications, Inc., New York, 1976.

[21] R. Kaufman, *On the theorem of Jarnik and Besicovitch*, Acta Arithmetica 39(1981), 265-267.

[22] I. Łaba, T. Tao, *An improved bound for the Minkowski dimension of Besicovitch sets in medium dimension*, Geom. Funct. Anal. 11(2001), 773-806.

[23] J. E. Marsden, *Elementary Classical Analysis*, W. H. Freeman and Co., San Francisco, 1974

[24] P. Mattila, *Spherical averages of Fourier transforms of measures with finite energy; dimension of intersections and distance sets*, Mathematika 34 (1987), 207-228.

[25] P. Mattila, *Geometry of sets and measures in Euclidean spaces*, Cambridge University Press, 1995.

[26] W. Minicozzi, C. Sogge, *Negative results for Nikodym maximal functions and related oscillatory integrals in curved space*, . Math. Res. Lett. 4 (1997), 221–237.

[27] W. Rudin, *Functional Analysis*, McGraw-Hill, Inc., New York, 1991.

[28] R. Salem, *Algebraic Numbers and Fourier Analysis*, D.C. Heath and Co., Boston, Mass. 1963 .

[29] C. Sogge, *Fourier integrals in Classical Analysis*, Cambridge University Press, Cambridge, 1993.

[30] C. Sogge, *Concerning Nikodym-type sets in 3-dimensional curved space*, Journal Amer. Math. Soc., 1999.

[31] J. Solymosi, C. Tóth, *Distinct distances in the plane*, Discrete Comput. Geometry 25(2001), 629-634.

[32] E.M. Stein, in *Beijing Lectures in Harmonic Analysis*, edited by E. M. Stein.

[33] E. M. Stein, *Harmonic Analysis*, Princeton University Press 1993.

[34] E. M. Stein, G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, Princeton, N.J. 1990

[35] P. A. Tomas, *A restriction theorem for the Fourier transform*, Bull. Amer. Math. Soc. 81(1975), 477-478.

[36] A. Varchenko, *Newton polyhedra and estimations of oscillatory integrals*, Funct. Anal. Appl. 18(1976), 175-196.

[37] T. Wolff, *Decay of circular means of Fourier transforms of measures*, International Mathematics Research Notices 10(1999), 547-567.

[38] T. Wolff, *An improved bound for Kakeya type maximal functions*, Revista Mat. Iberoamericana 11(1995), 651–674.

[39] T. Wolff, *Recent work connected with the Kakeya problem*, in *Prospects in Mathematics*, H. Rossi, ed., Amer. Math. Soc., Providence, R.I. (1999), 129-162.

[40] A. Zygmund, *Trigonometric Series*, Cambridge University Press, Cambridge, 1968.