Classical Fourier Analysis: Math 247A, Fall 99

Christoph Thiele

September 21, 2000

0.1 Schwartz functions

We plan to develop a theory of very general "functions" in $\mathbf{R}^n$ which handles problems of differentiability and similar questions in an elegant manner. We start by introducing a vector space of nice functions, on which all operations in question such as mainly differentiation and Fourier transform are defined. This is the class of Schwartz functions which we discuss in this section. This space of functions is ofcourse very limited, but the idea is to define the discussed operations on the dual space of the Schwarz space by transposition. The dual vector then contains a sufficiently large class of objects.

We discuss complex valued functions on $\mathbf{R}^n$. We define the partial differential operator

$$D_j = \frac{1}{2\pi i}\partial_j \quad ,$$

i.e.,

$$D_j f = \frac{1}{2\pi i}\partial_j f = \frac{1}{2\pi i}\frac{\partial f}{\partial x_j} \quad .$$

A multi-index is of the form

$$\alpha = (\alpha_1, \dots, \alpha_n)$$

with nonnegative integers α_j. The length of α is

$$|\alpha| = \sum_{j=1}^{n} \alpha_j \quad .$$

We define

$$D^\alpha = D_1^{\alpha_1} D_2^{\alpha_2} \dots D_n^{\alpha_n} \quad .$$

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n} \quad .$$

Let $C^\infty(\mathbf{R}^n)$ be the set of infinitely often differentiable functions, i.e. the functions f for which all partial derivatives $D^\alpha f$ exist in the classical sense (as limits of difference quotients) and are continuous.

Definition 1 (The N- norm) *We define for every $f \in C^\infty(\mathbf{R}^n)$*

$$\|f\|_{(N)} := \sup_{|\alpha|,|\beta|\le N} \sup_{x\in\mathbf{R}^n} \left|x^\alpha D^\beta f(x)\right| \quad .$$

Observe that $\|f\|_{(N)}$ may be infinite, and it is monotone increasing in N.

Definition 2 *The Schwartz space $\mathcal{S}(\mathbf{R}^n)$ is the set of all functions $f \in C^\infty(\mathbf{R}^n)$ such that $\|f\|_N < \infty$ for all $N \ge 0$.*

Observe that any linear combination and any product of Schwartz functions (Leibnitz rule) is a Schwartz function. The product of a Schwartz function and a polynomial is a Schwartz function. Any partial derivative of a Schwartz function is again a Schwartz function. The product of a Schwartz function with a polynomial is again a Schwartz function.

Examples

1) For any two (complex) polynomials P, Q with $\Re(Q(x)) > c|x|$ for some $c > 0$ and all $x \in \mathbf{R}^n$ with $|x| > c^{-1}$ the function

$$f(x) = P(x)e^{-Q(x)}$$

is in $\mathcal{S}(\mathbf{R}^n)$. Namely, we have

$$\sup_{x\in\mathbf{R}^n} |f(x)| \le \max\left\{ \max_{|x|\le c^{-1}} |f(x)|,\ \sup_{|x|>c^{-1}} \frac{|P(x)|(\deg(P))!)}{(c|x|)^{\deg(\mathrm{P})}} \right\} < \infty$$

and $x^\alpha D^\beta f$ is by an easy algebraic argument of the same form as f with a new polynomial P' and the same polynomial Q.

A frequent choice for Q is $Q(x) = \langle x, Ax\rangle$ for a positive definite real matrix A, in which case $e^{-Q(x)}$ is called a Gaussian, or $Q(x) = \langle x, Ax\rangle + \langle \xi, x\rangle$ for a positive definite real A and a complex valued B, in which case $e^{-Q(x)}$ is called a coherent state. Here we use the notation $\langle x, y\rangle$ for the standard scalar product of two vectors.

2) Every compactly supported function in $C^\infty(\mathbf{R}^n)$ is in $\mathcal{S}(\mathbf{R}^n)$. The space of such functions is denoted by $\mathcal{D}(\mathbf{R}^n)$. There are lots of such functions, e.g. $f \in \mathcal{D}(\mathbf{R}^n)$ where

$$f(x) = 0 \quad \text{if} |x| \ge 1 \quad .$$

$$f(x) = e^{\frac{-1}{1-|x|^2}} \quad \text{if} |x| < 1 \quad .$$

3) If $f \in S(\mathbf{R}^n)$, then $T_y f$, $M_\eta f$ and $D^2_\lambda f$ are in $\mathcal{S}(\mathbf{R})$ for all $y, \eta \in \mathbf{R}^n$ and $\lambda > 0$, where

$$T_y f(x) = f(x-y) \quad ,$$

$$M_\eta f(x) = f(x)e^{2\pi i\langle \eta, x\rangle} \quad ,$$

$$D^p_\lambda f(x) = \lambda^{-\frac{n}{p}} f(\lambda^{-1}x) \quad ,$$

4) Sometimes $\mathcal{S}(\mathbf{R}^n)$ is called the space of rapidly decaying function. But observe that the function $f(x) = e^{-x^2}e^{ie^x}$ is not in $\mathcal{S}(\mathbf{R})$. Hence rapid decay of the value of the function alone does not satisfy to assure membership in $\mathcal{S}(\mathbf{R})$.

0.2 The space $\mathcal{S}(\mathbf{R}^n)$ as Frechet space.

We first define a Frechet space. The notions in this definition will be explained gradually in this section.

Definition 3 *A Frechet space is a locally convex topological vector space which is induced by a complete invariant metric*

In order for $\mathcal{S}(\mathbf{R}^n)$ to qualify as Frechet space, we have to first define a topology for it. We shall do this by defining a metric on $\mathcal{S}(\mathbf{R}^n)$.

Definition 4 *For $f, g \in \mathcal{S}(\mathbf{R}^n)$ define*

$$\rho(f,g) := \sum_{N=0}^{\infty} 2^{-N} \frac{\|f-g\|_{(N)}}{1+\|f-g\|_{(N)}} \quad .$$

That ρ is a metric means that it satisfies positive definiteness, symmetry, and triangle inequality, which are easily verified.

It is also easy that ρ is invariant, which means

$$\rho(f-h, g-h) = \rho(f,g)$$

for all $f, g, h \in \mathcal{S}(\mathbf{R}^n)$.

Now we equip $\mathcal{S}(\mathbf{R}^n)$ with the topology $\mathcal{T}$ induced by the metric ρ, thus a set $U \subset \mathcal{S}(\mathbf{R})$ is open if and only if for all $f \in U$ there is an $\epsilon > 0$ such that the ϵ- ball about f in $\mathcal{S}(\mathbf{R})$ is contained in U:

$$\{g : \rho(f,g) \leq \epsilon\} \subset U \quad .$$

This is clearly a topology, i.e., $\emptyset$ and $\mathcal{S}(\mathbf{R})$ are open and finite intersections as well as arbitrary unions of open sets are open.

Moreover, $\mathcal{T}$ makes $\mathcal{S}(\mathbf{R})$ a topological vector space, which means that every point is a closed set and the vector operations

$$(f,g) \to f+g, (\alpha, f) \to \alpha f$$

are continuous (product topology on the product spaces). By invariance of ρ we only need to see that if $\rho(f,0) \to 0$ and $\rho(g,0) \to 0$ then $\rho(f+g,0) \to 0$ and similar for the multiplication with α. This is easy.

The following fundamental convergence lemma is useful

Lemma 1 *A sequence f_n in $\mathcal{S}(\mathbf{R}^n)$ converges to $f \in \mathcal{S}(\mathbf{R}^n)$ if and only if for all $N \geq 0$:*

(1) $$\lim_{n\to\infty} \|f_n - f\|_{(N)} = 0 \quad .$$

Proof

Assume that f_n converges to f. Since

$$(2) \qquad \sum_{N=0}^{\infty} 2^{-N} \frac{\|f_n - f\|_{(N)}}{1 + \|f_n - f\|_{(N)}}$$

converges to 0, so does

$$2^{-N} \frac{\|f_n - f\|_{(N)}}{1 + \|f_n - f\|_{(N)}}$$

for each N because all summands in (2) are positive, and therefore so does

$$\|f_n - f\|_{(N)}$$

This proves (1).

Now assume (1) holds and let $\epsilon > 0$. Pick N_0 such that $2^{-N_0} < \epsilon$. Then

$$\rho(f_n, f) \leq \epsilon + \sum_{N=0}^{N_0} 2^{-N} \frac{\|f_n - f\|_{(N)}}{1 + \|f_n - f\|_{(N)}} \quad .$$

By (1) each of the finitely many terms in the second summand converges to 0 for $n \to \infty$, hence this term is smaller than ϵ for large n. Thus f_n converges to f. ■

For $\epsilon > 0$ and $N \in \mathbf{N}$ define

$$V_{\epsilon,N} := \{f \in \mathcal{S}(\mathbf{R}^n) : \|f\|_{(N)} < \epsilon\} \quad .$$

The proof of Lemma 1 shows that each $V_{\epsilon,N}$ contains in a ball $B_{\epsilon'}(0)$ and vice versa each ball $B_\epsilon(0)$ contains a $V_{\epsilon',N}$. Thus the set of all $V_{\epsilon,N}$ is a neighborhood basis of the point 0 of the topology $\mathcal{T}$.

A topological vector space is locally convex, if 0 has a neighborhood basis consisting of convex sets. Here a set V in a vector space is convex if for all $x, y \in V$ we have $\lambda x + (1 - \lambda) y \in V$ for all $\lambda \in (0, 1)$. Since all $V_{\epsilon,N}$ are convex, $\mathcal{S}(\mathbf{R}^n)$ is locally convex. It is not true, that the balls $B_\epsilon(0)$ are convex (exercise).

Finally, we have to show that $\mathcal{S}(\mathbf{R}^n)$ is a complete topological vector space. A Cauchy sequence in a topological vector space is a sequence $(f_n)_{n \in \mathbf{N}}$ such that for each neighborhood of 0 there is an m such that for $n, n' > m$ we have $f_n - f_{n'} \in V$. If the topology is given by an invariant metric, this definition is easily seen to coincide with the definition of Cauchy sequences in metric spaces. We have to show that if f_n is a Cauchy sequence in $\mathcal{S}(\mathbf{R}^n)$, then f_n converges in $\mathcal{S}(\mathbf{R}^n)$. However, if f_n is such a sequence, then $x^\alpha D^\beta f_n$ is a Cauchy sequence in the sup -norm, hence converges to a bounded continuous function. In particular f_n converges to a function f and an standard lemma from calculus implies

$$\lim_{n \to \infty} x^\alpha D^\beta f_n = x^\alpha D^\beta f$$

in the sup norm. Now Lemma 1 shows that f_n converges to f in $\mathcal{S}(\mathbf{R}^n)$.

We close this section with two more easy but fundamental lemmata.

Lemma 2 *Let $\Lambda : \mathcal{S}(\mathbf{R}^n) \to \mathbf{C}$ be a linear functional. Then Λ is continuous if and only if there exists a $N \geq 0$ and a $C > 0$ such that*

$$|\Lambda(f)| \leq C\|f\|_{(N)} \tag{3}$$

for all $f \in \mathcal{S}(\mathbf{R})$.

Proof Assume Λ is continuous. Let $B_1(0)$ be the open 1 ball about 0 in $\mathbf{C}$. Then $\Lambda^{-1}(B_1(0))$ is an open neighborhood of 0, hence contains a $V_{\epsilon,N}$. Thus $\|f\|_{(N)} < \epsilon$ implies $\Lambda(f) < 1$, which implies (3) because

$$\Lambda(f) \leq 2\epsilon^{-1}\|f\|_{(N)}\Lambda\left(\frac{\epsilon}{2\|f\|_{(N)}}f\right) \leq 2\epsilon^{-1}\|f\|_{(N)} \quad .$$

Now assume that (3) holds. By linearity it suffices to show that Λ is continuous at 0, i.e., for each $\epsilon > 0$ there is a $V_{\epsilon',N}$ such that

$$|\Lambda(V_{\epsilon',N})| \leq \epsilon \quad .$$

However, this is immediate if we pick N as in (3) and $\epsilon' = \epsilon C^{-1}$.
■

Lemma 3 *Let $T : \mathcal{S}(\mathbf{R}^n) \to \mathcal{S}(\mathbf{R}^n)$ be a linear map. Then T is continuous if and only if for each N there is a N' and a $C > 0$ such that*

$$\|T(f)\|_{(N)} \leq C\,\|f\|_{(N')}$$

for all $f \in \mathcal{S}(\mathbf{R})$.

Proof This follows by the same ideas as the two preceeding lemmas. ■

We have

$$\|D_j f\|_{(N)} \leq c\|f\|_{(N+1)} \quad ,$$

hence the linear map D_j is continuous on $\mathcal{S}(\mathbf{R})$. Define the operator X_j by

$$X_j F(y) = y_j f(y) \quad ,$$

thus X_j is multiplication with the j-th coordinate. Again, we have

$$\|X_j f\|_{(N)} \leq c\|f\|_{(N+1)} \quad ,$$

hence X_j is a continuous linear operator on $\mathcal{S}(\mathbf{R})$.

The operations T_y, M_η, and D^p_λ are also continuous for all y, η, λ, p (exercise).

0.3 The Fourier transform

Define for $x, \xi \in \mathbf{R}^n$

$$\chi_\xi(x) := e^{2\pi i\langle x,\xi\rangle} \quad .$$

Hence $\chi_\xi(x) = \chi_x(\xi)$. For $f \in \mathcal{S}(\mathbf{R}^n)$ we define the Fourier transform

$$\mathcal{F}(f)(\xi) = \widehat{f}(\xi) = \int_{\mathbf{R}^n} f(x)\chi_{-\xi}(x)\, dx \tag{4}$$

Thus the Fourier transform is again a function on $\mathbf{R}^n$.

Observe that this integral is well defined in the (improper) Riemann or Lebesgue sense because the integrand is continuous and bounded by

$$C \min(1, |x|^{-n-1}) \quad .$$

Moreover we have for all ξ:

$$\widehat{f}(\xi) \le \|f\|_1 := \int_{\mathbf{R}^n} |f(x)|\, dx \quad .$$

The Fourier transform interchanges partial differentiation with multiplication by monomials in the following sense.

Observe that

$$D_j\chi_\xi(x) = \xi_j\, \chi_\xi(x) = X_j\chi_x(\xi) \quad .$$

Hence we have by partial integration

$$X_j\widehat{f}(\xi) = -\int_{\mathbf{R}^n} f(x)D_j\chi_{-\xi}(x)\, dx = \int_{\mathbf{R}^n} D_jf(x)\chi_{-\xi}(x)\, dx = \mathcal{F}(D_jf) \quad .$$

Moreover we have

$$\mathcal{F}(X_jf) = \int_{\mathbf{R}^n} X_jf(x)\chi_{-\xi}(x)\, dx$$

$$= -\int_{\mathbf{R}^n} f(x)D_j\chi_{-x}(\xi)\, dx = -D_j\int_{\mathbf{R}^n} f(x)\chi_{-\xi}(x)\, dx = -D_j\widehat{f}(\xi) \quad .$$

The interchange of differentiation and integration in the third equality ist easily justified with elementary calculus.

This shows that $\mathcal{F}f$ is infinitely often differentiable and we have

$$\|\mathcal{F}f\|_{(N)} = \sup_{|\alpha|,|\beta|\le N} \left\|X^\alpha D^\beta \mathcal{F}f\right\|_\infty$$

$$= \sup_{|\alpha|,|\beta|\le N} \left\|\mathcal{F}D^\alpha X^\beta f\right\|_\infty \le \sup_{|\alpha|,|\beta|\le N} \left\|D^\alpha X^\beta f\right\|_1$$

$$\le C \sup_{|\alpha|,|\beta|\le N} \left\|(1+|.|^2)^n D^\alpha X^\beta f\right\|_\infty \le C\|f\|_{N+2n}$$

Thus we have proved the following lemma:

Lemma 4 *The Fourier transform is a continuous linear map* $\mathcal{F} : \mathcal{S}(\mathbf{R}^n) \to \mathcal{S}(\mathbf{R}^n)$

Example We determine the Fourier transform of $f \in \mathcal{S}(\mathbf{R}^n)$ given by

$$f(x) = e^{-\pi|x|^2} \quad .$$

Observe that

$$D_j f = ix_j e^{-\pi|x|^2} = iX_j f \quad .$$

Taking the Fourier transform on both sides gives

$$X_j \widehat{f} = -iD_j \widehat{f} \quad .$$

Thus for $n = 1$, $\widehat{f}$ satisfies the same linear ordinary differential equation as f and therefore $\widehat{f} = cf$ because f does not vanish. If $n \neq 1$, then we have

$$f(x) = \prod_{j=1}^{n} e^{-\pi|x_j|^2}$$

and hence

$$\widehat{f}(\xi) = c^n \prod_{j=1}^{n} e^{-\pi|\xi_j|^2} = c^n f(\xi) \quad .$$

To calculate c we observe for $n = 2$

$$\widehat{f}(0) = \int_{\mathbf{R}^2} e^{-\pi|x|^2}\, dx = \int_0^\infty e^{-\pi r^2} 2\pi r\, dr = \int_0^\infty e^{-s}\, ds = 1 = f(0)$$

Hence $c \in \{-1, 1\}$. But c is positive so $c = 1$.

■

We observe the following easy formulas.

$$\mathcal{F}(T_y f) = M_{-y}(\mathcal{F} f) \tag{5}$$

$$\mathcal{F}(M_\eta f) = T_\eta(\mathcal{F} f) \tag{6}$$

$$\mathcal{F}(D_\lambda^p f) = D_{\lambda^{-1}}^{p'}(\mathcal{F} f) \tag{7}$$

where p' is defined by $1/p + 1/p' = 1$. They follow from the identities

$$\int_{\mathbf{R}^n} f(x-y) e^{-2\pi i<x,\xi>}\, dx = e^{2\pi i<-y,\xi>} \int_{\mathbf{R}^n} f(x-y) e^{-2\pi i<x-y,\xi>}\, dx$$

$$\int_{\mathbf{R}^n} e^{2\pi i<x,\eta>} f(x-y) e^{-2\pi i<x,\xi>}\, dx = \int_{\mathbf{R}^n} f(x-y) e^{-2\pi i<x,\xi-\eta>}\, dx$$

$$\int_{\mathbf{R}^n} \lambda^{-n/p} f(\lambda^{-1}x) e^{-2\pi i<x,\xi>}\, dx = \int_{\mathbf{R}^n} \lambda^n \lambda^{-n/p} f(x) e^{-2\pi i<x,\lambda\xi>}\, dx$$

Define

$$\langle f, g \rangle = \int_{\mathbf{R}^n} f(x)\overline{g(x)}\, dx$$

and

$$\overline{\mathcal{F}} f(\xi) = \int_{\mathbf{R}^n} f(x) e^{+2\pi i<x,\xi>}\, dx \quad .$$

For all $f, g \in \mathcal{S}(\mathbf{R}^n)$ we have

$$\langle \mathcal{F}f, g\rangle = \left\langle f, \overline{\mathcal{F}}g \right\rangle \quad .$$

This is the simply Fubini's theorem:

$$\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} f(x) e^{-2\pi i <x,\xi>} \overline{g(\xi)}\, dx\, d\xi = \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} f(x) \overline{g(\xi) e^{2\pi i <x,\xi>}}\, d\xi\, dx \quad .$$

The operator $\overline{\mathcal{F}}$ is also a continuous linear map on $\mathcal{S}(\mathbf{R}^n)$. In fact it is the inverse of $\mathcal{F}$:

Theorem 1 *We have for all $f \in \mathcal{S}(\mathbf{R})$:*

$$\overline{\mathcal{F}}\mathcal{F}f = f \quad , \quad \mathcal{F}\overline{\mathcal{F}}f = f \quad .$$

In particular $\mathcal{F}$ is an homeomorphic isomorphism and we have Plancherel's formula

$$\langle f, g\rangle = \langle \mathcal{F}f, \mathcal{F}g\rangle$$

for all $f, g \in \mathcal{S}(\mathbf{R}^n)$.

Proof It suffices to prove one of the inversion formulas, the other statements then are obvious. Observe that we have for all $f, g \in \mathcal{S}(\mathbf{R})$ and all $\lambda > 0$

$$\langle M_\lambda^\infty f, \mathcal{F}g\rangle = \left\langle f, M_{\lambda^{-1}}^1 \mathcal{F}g \right\rangle$$

$$= \langle f, \mathcal{F}M_\lambda^\infty g\rangle = \left\langle \overline{\mathcal{F}}f, M_\lambda^\infty g \right\rangle$$

Letting $\lambda \to \infty$, we obtain by Lebesgue's dominated convergence theorem

$$= f(0)\int_{\mathbf{R}^n} \overline{\mathcal{F}g(\xi)}\, d\xi = \overline{g(0)}\int_{\mathbf{R}^n} \overline{\mathcal{F}}(f)(\xi)\, d\xi \quad .$$

Letting $g(x) = e^{-\pi|x|^2}$ we obtain

$$f(0) = \mathcal{F}\overline{\mathcal{F}}f(0) \quad .$$

Thus we have for all x:

$$f(x) = T_{-x}f(0) = \mathcal{F}\overline{\mathcal{F}}T_{-x}f(0) = \mathcal{F}M_{-x}\overline{\mathcal{F}}f(0) = T_{-x}\mathcal{F}\overline{\mathcal{F}}f(0) = \mathcal{F}\overline{\mathcal{F}}f(x) \quad .$$

This proves one of the inversion inequalities. ∎

Observe that we have in particular for each $f \in \mathcal{S}(\mathbf{R}^n)$:

$$\|f\|_2^2 = \langle f, f\rangle = \langle \mathcal{F}f, \mathcal{F}f\rangle = \|\mathcal{F}f\|_2^2 \quad .$$

0.4 Tempered Distributions

Definition 5 *A continuous linear functional $\Lambda : \mathcal{S}(\mathbf{R}^n) \to \mathbf{C}$ is called a tempered distribution. The space of all tempered distributions with the weak star topology is denoted by $\mathcal{S}'(\mathbf{R}^n)$.*

Thus the topology on $\mathcal{S}'(\mathbf{R}^n)$ is the weakest topology such that for each $f \in \mathcal{S}(\mathbf{R}^n)$ the functional

$$f^* : \mathcal{S}'(\mathbf{R}^n) \to \mathbf{C}, \Lambda \mapsto \Lambda(f)$$

is continuous.

Each element $f \in \mathcal{S}(\mathbf{R}^n)$ can be identified as an element $\Lambda_f \in \mathcal{S}'(\mathbf{R}^n)$ via the linear functional

$$\mathcal{S}(\mathbf{R}^n) \to \mathbf{C} : g \to \int_{\mathbf{R}^n} f(x)g(x)\,dx \quad .$$

We have to verify that this is a continuous linear functional. However, we have

$$\left| \int_{\mathbf{R}^n} f(x)g(x)\,dx \right| \le \|f\|_2 \|g\|_2 \le C \|f\|_{(n)} \|g\|_{(n)}$$

This proves continuity by Lemma 2.

If $\Lambda \in \mathcal{S}'(\mathbf{R}^n)$ and $f \in \mathcal{S}(\mathbf{R}^n)$, then we also write (Λ, f) or (f, Λ) for $\Lambda(f)$. If Λ happens to be a Schwartz function, there is no ambiguity because of the symmetry $\Lambda_f(g) = \Lambda_g(f)$. The pairing (f, g) is a bilinear and not to be confused with the pairing $\langle f, g \rangle$, which is sesquilinear (antilinear in the second argument).

An example for a tempered distribution which is not a Schwartz function is the Dirac distribution $\delta_0 \in \mathcal{S}'(\mathbf{R}^n)$ given by

$$\delta_0(f) = f(0) \quad .$$

To verify continuity of δ_0 observe that $f(0) \le \|f\|_{(0)}$.

Lemma 5 *Let $\mathcal{V}$ be a topological vector space. A map $T : \mathcal{V} \to \mathcal{S}'(\mathbf{R}^m)$ is continuous if and only if for each $f \in \mathcal{S}(\mathbf{R}^m)$ the map*

$$f^* \circ T : \mathcal{V} \to \mathbf{C} : v \to T(v)(f)$$

is continuous.

Proof

If T is continuous, then $f^* \circ T$ is continuous because f^* is continuous for each f by definition of the topology of $\mathcal{S}'(\mathbf{R}^n)$.

Now assume $f^* \circ T$ is continuous for all $f \in \mathcal{S}(\mathbf{R}^n)$. By linearity it suffices to show that T is continuous at 0. Let V be an open neighborhood of 0 in $\mathcal{S}'(\mathbf{R}^n)$. Then there is $N > 0$ and $f_1, \dots, f_N$ and $\epsilon_1, \dots \epsilon_N$ such that the open set

$$U = \{\Lambda : |\Lambda(f_j)| < \epsilon_j \quad \text{for} \quad \text{all} \quad 1 \le j \le N\}$$

is contained in V. However, the pre-image of U under T is the set

$$\{v \in \mathcal{V} : |f_j^* \circ T(v)| < \epsilon_j \quad \text{for} \quad \text{all} \quad 1 \le j \le N\} \quad ,$$

which is open by the continuity of the maps $f_j^* \circ T$. This proves continuity of T at 0. ■

Let $T : \mathcal{S}(\mathbf{R}^n) \to \mathcal{S}'(\mathbf{R}^m)$ be a continuous map. Then there is a unique map $T^T : \mathcal{S}(\mathbf{R}^m) \to \mathcal{S}'(\mathbf{R}^n)$ such that

$$(Tf, g) = (f, T^T g)$$

for all $f, g \in \mathcal{S}(\mathbf{R}^n)$. Namely, $T^T g$ is the distribution $g^* \circ T$. We call T^T the transpose of T. Obviously T is the transpose of T^T.

Lemma 6 *Let $T : \mathcal{S}(\mathbf{R}^n) \to \mathcal{S}(\mathbf{R}^m)$ and $S : \mathcal{S}(\mathbf{R}^m) \to \mathcal{S}(\mathbf{R}^n)$ be two continuous linear maps which are transposes of each other (via the embedding of $\mathcal{S}$ into $\mathcal{S}'$). Then T extends to a continuous map*

$$T : \mathcal{S}'(\mathbf{R}^n) \to \mathcal{S}'(\mathbf{R}^m)$$

via the formula

$$(T\Lambda, f) := (\Lambda, Sf) \quad .$$

Proof Obviously $T\Lambda = \Lambda \circ S$ is a continuous linear functional on $\mathcal{S}(\mathbf{R}^m)$ and hence a tempered distribution. If Λ happens to be a Schwartz function, then the two possible definitions for $T\Lambda$ coincide because S is the transpose of T. It remains to prove that the extension of T to $\mathcal{S}'(\mathbf{R}^n)$ is continuous. By Lemma 5 it suffices to prove that $f^* \circ T$ is continuous for each $f \in \mathcal{S}(\mathbf{R}^m)$. However $f^* \circ T = (Sf)^*$, and the latter is continuous by definition of the topology of $\mathcal{S}'(\mathbf{R})$. ■

Examples: The operators D_j, X_j, $\mathcal{F}$, T_y, M_η, D^p_λ have the transpose operators $-D_j$, X_j, $\mathcal{F}$, T_{-y}, M_η, $D^{p'}_{\lambda^{-1}}$. Hence they all extend to bounded operators on $\mathcal{S}(\mathbf{R}^n)$.

Lemma 7 below will show that extension defined in Lemma 6 is the unique extension of T to a continuous linear operator on $\mathcal{S}'(\mathbf{R}^n)$.

For example we have

$$(D_j\delta_0, f) = (\delta_0, -D_j f) = -\frac{1}{2\pi i}\partial_j f(0) \quad ,$$

$$(X_j\delta_0, f) = (\delta_0, X_j f) = 0 f(0) = 0 \quad ,$$

$$(\mathcal{F}\delta_0, f) = (\delta_0, \mathcal{F}f) = \int_{\mathbf{R}^n} f(x) 1_{\mathbf{R}}(x)\, dx \quad ,$$

$$(T_y\delta_0, f) = (\delta_0, T_{-y} f) = f(y) \quad ,$$

$$(M_\eta\delta_0, f) = (\delta_0, M_\eta f) = f(0) = (\delta_0, f) \quad ,$$

$$(D^p_\lambda\delta_0, f) = (\delta_0, D^{p'}_{\lambda^{-1}} f) = \lambda^{\frac{n}{p'}} f(0) = (\lambda^{\frac{n}{p'}}\delta_0, f) \quad .$$

Lemma 7 *The space $\mathcal{S}(\mathbf{R}^n)$ is dense in $\mathcal{S}'(\mathbf{R}^n)$.*

Proof Let $\Lambda \in \mathcal{S}'(\mathbf{R}^n)$, we have to show that there exists a sequence $f_i \in \mathcal{S}(\mathbf{R}^n)$ converging to Λ. As tempered distribution Λ satisfies

$$|\Lambda(g)| \leq C\|g\|_{(N)}$$

for some $N \geq 0$. Observe that

$$\begin{aligned} \|X^\alpha D^\beta f\|_\infty &\leq \|D^\alpha X^\beta \mathcal{F} f\|_1 \\ \leq \|(1+|.|^2)^{-n}\|_2 &\|(1+|.|^2)^n D^\alpha X^\beta \mathcal{F}_f\|_2 \\ &\leq C\|(1+|D|^2)^n X^\alpha D^\beta f\|_2 \quad . \end{aligned}$$

Here we have use theorem 1. Hence

$$|\Lambda(g)| \leq C \left(\sum_{|\alpha|,|\beta| \leq N+2n} \|X^\alpha D^\beta g\|_2^2 \right)^{\frac{1}{2}} \quad .$$

Hence there is an element $(f_{\alpha,\beta})_{|\alpha|,|\beta|\leq N+2n}$ in the Hilbert space $L^2(\mathbf{R}^n)^{\{|\alpha|,|\beta|\leq N+2n\}}$ such that for every $g \in \mathcal{S}(\mathbf{R}^n)$

$$\Lambda(f) = \sum_{|\alpha|,|\beta|\leq N+2n} \int_{\mathbf{R}^n} X^\alpha D^\beta g(x)\overline{f(x)}\, dx \quad .$$

Let $(f_{\alpha,\beta}^{(k)})$ be a sequence of tuples in $\mathcal{S}(\mathbf{R}^n)^{\{|\alpha|,|\beta|\leq N+2n\}}$ converging to $(f_{\alpha,\beta})$. (Here we are using that $\mathcal{S}(\mathbf{R}^n)$ is dense in $L^2(\mathbf{R}^n)$. Then partial integration shows that

$$\Lambda(f) = \lim_{k\to\infty} \int_{\mathbf{R}^n} g(x) \left(\sum_{|\alpha|,|\beta|\leq N+n} \overline{D^\beta X^\alpha f_{\alpha,\beta}^{(k)}(x)} \right) dx \quad .$$

This provides the desired sequence of Schwartz functions. ∎

We observe that this lemma also proves that each distribution can be written as

$$\sum_{|\alpha|,|\beta|\leq N} D^\beta X^\alpha f_{\alpha,\beta}$$

for a certain $N \geq 0$ and functions $f_{\alpha,\beta}$ in $L^2(\mathbf{R}^n) \subset \mathcal{S}'(\mathbf{R}^n)$. Thus the space of tempered distributions is the most economic class of objects that contains $L^2(\mathbf{R}^n)$ and allows for application of the operators X_j and D_j for all $j = 1, \ldots n$.

0.5 Intermediate Banach Spaces, Regularity

Let $\|.\|_B$ be a norm on $\mathcal{S}(\mathbf{R}^n)$. As with the $\|.\|_{(N)}$ norms we do not require $\mathcal{S}(\mathbf{R}^n)$ to be complete with respect to this norm. A variant of Lemma 2 shows that $\|.\|_B$ is a continuous map from the Schwartz space to $\mathbf{C}$ if and only if for some $C > 0$ and $N \geq 0$

$$\|f\|_B \leq C\|f\|_{(N)}$$

for all $f \in \mathcal{S}(\mathbf{R}^n)$.

Lemma 8 *Let $\|.\|_B$ be a continuous norm on $\mathcal{S}(\mathbf{R}^n)$. Then the set of all linear functionals on $\mathcal{S}(\mathbf{R}^n)$ which are continuous with respect to $\|.\|_B$ is equal to the space B' of all distributions in $\mathcal{S}'(\mathbf{R}^n)$ for which the quantity*

$$\|\Lambda\|_{B'} := \sup_{f \in \mathcal{S}(\mathbf{R}^n):\|f\|_B \leq 1} |\Lambda(f)|$$

is finite. The space B' is a Banach space under the norm $\|.\|_{B'}$.

Proof Let Λ be a linear functional on $\mathcal{S}(\mathbf{R}^n)$ which is continuous with respect to $\|.\|_B$. Then

$$|\Lambda(f)| \leq C\|f\|_B \leq C'\|f\|_{(N)} \quad ,$$

hence Λ is a distribution with $\|\Lambda\|_{(B')} \leq C' < \infty$. Vice versa, if Λ is a distribution with $\|\Lambda\|_{(B')} < \infty$, then

$$\Lambda(f) = \|f\|_B \Lambda(f/\|f\|_B) \leq C\|f\|_B \quad ,$$

hence Λ is continuous with respect to $\|.\|_B$.

The space B' is clearly a normed space. To show that it is complete let Λ_k be a Cauchy sequence in B'. Then for each $f \in \mathcal{S}(\mathbf{R}^n)$ the sequence $\Lambda_k f$ is Cauchy in $\mathbf{C}$ because

$$\|\Lambda_k f - \Lambda_m f\| \leq \|\Lambda_k - \Lambda_m\|_{B'} \, \|f\|_B \quad .$$

Define $\Lambda f := \lim_{k\to\infty} \Lambda_k f$, then it is easy to see that $\Lambda \in B'$ and Λ_k converges to Λ strongly (in B'- norm). ■

The statement that a distribution Λ is in such a space B' is called a regularity statement for Λ. A regularity result usually is used to perform operations with distributions that are not possible for general distributions. E.g. we see immediately that if B is the completion of $\mathcal{S}(\mathbf{R}^n)$ under the norm $\|.\|_B$, then a distribution in B' acts as linear functional on all of B and B'- regularity is necessary for this. Frequently the space

Example Consider on $\mathcal{S}(\mathbf{R}^n)$ the continuous norm

$$\|f\|_{(0)} = \|f\|_\infty \quad .$$

Every complex Borel measure on $\mathbf{R}^n$ is a continuous linear functional on $\mathcal{S}(\mathbf{R}^n)$ with respect to this norm, and these are all continuous linear functionals because $\mathcal{S}(\mathbf{R}^n)$

is dense in $C_0(\mathbf{R}^n)$. The space $M_1(\mathbf{R}^n)$ of complex Borel measures is a Banch space under the total variation norm, which coincides with the norm of $M_1(\mathbf{R}^n)$ as dual space of $C_0(\mathbf{R}^n)$. Thus $M_1(\mathbf{R}^n)$ is a subset of $\mathcal{S}'(\mathbf{R}^n)$. ■

Example It is trivial that $\mathcal{S}(\mathbf{R}^n) \subset M_1(\mathbf{R}^n)$. The total variation norm is again a continuous norm on $\mathcal{S}(\mathbf{R}^n)$, which coincides with the L^1 norm. Moreover, $\mathcal{S}(\mathbf{R}^n)$ is dense in L^1 which is a closed subspace of M_1. The dual space of L^1 can be identified with $L^\infty(\mathbf{R}^n)$, thus $L^\infty(\mathbf{R}^n)$ is a space of the type B' as in the above theorem. To close the circle, $\mathcal{S}(\mathbf{R}^n)$ is a subset of $L^\infty(\mathbf{R}^n)$ whose closure is $C_0(\mathbf{R}^n)$. ■

Example Let $n = 1$. The space of distributions Λ such that $D_1(\Lambda)$ is L^1, M_1, C_0, L^∞, resp. consists exactly of the functions which are absolutely continuous, have bounded variation, are continuously differentiable (with certain behaviour at $\pm\infty$, are Lipschitz resp. ■

Example The L^p norms $\|.\|_p$ are continuous on $\mathcal{S}(\mathbf{R}^n)$ for $1 < p < \infty$. The dual space is then $L^{p'}$. The most important example is $p = 2$, the only example for which $\|.\|_B$ and $\|.\|_{B'}$ coincide on $\mathcal{S}(\mathbf{R}^n)$. Thus the Hilbert space $L^2(\mathbf{R}^n)$ in some sense is exactly in the middle between $\mathcal{S}(\mathbf{R}^n)$ and $\mathcal{S}'(\mathbf{R}^n)$. ■

Let $f \in L^1(\mathbf{R}^n) \subset \mathcal{S}'(\mathbf{R}^n)$. We can approximate f by a sequence of Schwartz functions f_n in the L^1 norm. Then f_n also converges to f in the distributional sense. The sequence $\mathcal{F}(f_n)(\xi)$ converges to

$$\int_{\mathbf{R}^n} f(x) e^{-2\pi i <x,\xi>} \, dx$$

uniformly (in the C_0 norm). On the other hand, $\mathcal{F}(f_n)$ converges to$\mathcal{F}(f)$ in the distributional sense, therefore $\mathcal{F}(f)$ is given by the Fourier integral.

0.6 The Poisson Summation Formula

We consider the following distribution on $\mathcal{S}(\mathbf{R}^n)$:

$$\Lambda(f) = \sum_{z \in \mathbf{Z}^n} f(z)$$

Using weights it is easy to see that $|\Lambda(f)| \leq C\|f\|_{(2n)}$, verifying that Λ is indeed a distribution.

Lemma 9 (Poisson Summation Formula) *We have $\mathcal{F}\Lambda = \Lambda$.*

Proof We first prove the lemma for $n = 1$. Observe that Λ satisfies $T_1\Lambda = M_1\Lambda = \Lambda$. We claim that these properties determine Λ up to a scalar multiple. Since $\mathcal{F}\Lambda$ obviously satisfies the same invariances, this proves $\mathcal{F}\Lambda = \Lambda$ (the scalar is verified by applying Λ to the Schwartz function $e^{-\pi|x|^2}$).

To see the claim, assume Λ' satisfies these invariances. We show that Λ' coincides with Λ on the set of compactly supported Schwartz functions, which proves the claim by density. Let g be a fixed Schwartz function supported in $B_{1/4}(0)$ with $g(0) = 1$. Then we have for every compactly supported $f \in \mathcal{S}(\mathbf{R}^n)$

$$\Lambda'(f) = \Lambda'\left(f - \sum_{z\in\mathbf{Z}} f(z)T_z g\right) + \Lambda'(g)\sum_{z\in\mathbf{Z}} f(z) \quad .$$

(The sums in here are really finite sums.) In the second summand we have used invariance under all T_z. Since the second summand is of the desired form, it remains to prove that the first summand is 0. However, define

$$h(x) = \frac{f(x) - \sum_{z\in\mathbf{Z}} f(z)T_z g(x)}{1 - e^{2\pi i x}}$$

The denominator vanishes of first order at all integers, but the enumerator also vanishes at these points and therefore h extends to a smooth compactly supported function. But we have

$$\Lambda'\left(f - \sum_{z\in\mathbf{Z}} f(z)T_z g\right) = \Lambda'(h - M_1 h) = 0 \quad ,$$

which completes the proof in the case $n = 1$.

For general $n > 1$ let $\mathcal{F}_j$ denote the partial Fourier transform in direction j:

$$\mathcal{F}_j f(x_1, \ldots, \xi_j, \ldots, x_n) = \int_{\mathbf{R}} f(x)e^{2\pi i x_j \xi + j}\, dx_j \quad .$$

Then Fubini's theorem implies

$$\mathcal{F} = \mathcal{F}_1\mathcal{F}_2\ldots\mathcal{F}_n \quad .$$

The statement of the lemma for $n = 1$ implies

$$\mathcal{F}_j\Lambda = \Lambda \quad .$$

Thus the Lemma follows by an iteration of this formula.
■

0.7 Homogeneous distributions in $\mathcal{S}'(\mathbf{R})$

We introduce a new symbol for the dilation operators by

$$S^z_\lambda f(x) = \lambda^z f(\lambda^{-1}x) \quad .$$

Thus $D^p_\lambda = S_\lambda^{-n/p}$. The point is that we want to allow z to be an aritrary complex number, so the association of p with an L^p space becomes obsolete. The formal transpose of S^z_λ is $S^{-n-z}_{\lambda^{-1}}$.

Clearly S_λ^z extends to a continuous operator on $\mathcal{S}'(\mathbf{R}^n)$. A distribution in $\mathcal{S}'(\mathbf{R}^n)$ is called homogeneous of degree z if

$$S_\lambda^z \Lambda = \Lambda \quad .$$

For example if z is a positive integer then homogeneous polynomials of degree z are homogeneous distributions of degree z

We shall restrict attention to the case $n = 1$ for now.

Let $\Omega = \{z \in \mathbf{C} : \Re(z) > 0\}$. For $z \in \Omega$ the function $|t|^z$ is continuous and grows polynomially, hence we can define homogeneous distributions by

$$\Lambda_z^+(f) = \int_{t>0} |t|^z f(t)\, dt$$

$$\Lambda_z^-(f) = \int_{t<0} |t|^z f(t)\, dt$$

Thus Λ_z^+ is homogeneous of degree z. It is easy to see that for fixed $f \in \mathcal{S}(\mathbf{R}^n)$ the function $\Lambda_z^+(f)$ is holomorphic in z in the region Ω. Namely, the function $|t|^z f(t)$ is pointwise holomorphic (complex differentiable) in z and both $|t|^z f(t)$ and its complex derivative w.r.t. z are uniformly in L^1 on Ω.

For all $z \in \Omega$ we have by partial integration

$$\Lambda_z^+(f) = \int_{t>0} |t|^z f(t)\, dt$$

$$= -\frac{1}{z+1} \int_{t>0} |t|^{z+1} f'(t)\, dt = -\frac{1}{z+1} \Lambda_{z+1}^+(f') \quad .$$

By iteration we obtain

$$\Lambda_z^+(f) = (-1)^n \frac{1}{(z+N)(z+N-1)\dots(z+1)} \Lambda_{z+N}^+(f^{(N)})$$

whenever $z \in \Omega$. The right hand side is a meromorphic function for $\Re(z) > -N$, thus we can use this formula to extend Λ_z^+ meromorphically to the region $\Re(z) > -N$. By uniqueness of meromorphic extensions these extensions coincide for different N on the intersections of their supports. Thus we can extend $\Lambda_z^+(f)$ meromorphically to all of $\mathbf{C}$ by taking the union of these extensions. Obviously $\Lambda_z^+(f)$ has poles at most at the negative integers and these poles are at most of order one.

Example We take $f(x) = 2e^{-\pi x^2}$. Then for $z \in \Omega$

$$\Lambda_z^+(f) = \Lambda_z^-(f) = \int_{\mathbf{R}} |t|^z e^{-\pi t^2}\, dt \quad .$$

We have

$$\Lambda_1^+(f) = \int_0^\infty 2te^{-\pi t^2}\, dt = \frac{1}{\pi} \int_0^\infty e^{-s}\, ds = \frac{1}{\pi} \quad .$$

We derive the following functional equation for $\lambda_z^+(f)$:

$$\Lambda_z^+(f) = -\frac{1}{z+1} \Lambda_{z+1}^+(f') = \frac{2\pi}{z+1} \int_0^\infty t^{z+1} 2e^{-\pi t^2}\, dt = \frac{2\pi}{z+1} \Lambda_{z+2}^+(f) \quad .$$

Thus $\Lambda_z^+(f)$ has simple poles at exactly the negative odd integers. Define

$$\Gamma(z) = \pi^z \Lambda_{2z-1}(f) \quad ,$$

then Γ has simple poles at exactly the non-positive integers and we have

$$\Gamma(1) = \pi \Lambda_1(f) = 1 \quad ,$$

and with the previous functional equation

$$\Gamma(z) = \pi^z \frac{2\pi}{2z} \Lambda_{2z+1}(f) = \frac{1}{z}\Gamma(z+1) \quad .$$

We would like to argue that Γ has no zeroes. Observe that Γ is bounded in the strip $1 \leq \Re(z) \leq 2$ and hence on the set $0 \leq \Re(z) \leq 1$, $|\Im(z)| > 1$. Thus $S(z) = \Gamma(z)\Gamma(1-z)$ is bounded in the latter set and we have

$$S(z+1) = \Gamma(z+1)\Gamma((1-z)-1) = z(-z)^{-1}\Gamma(z)\Gamma(1-z) = -S(z) \quad .$$

Thus $S(z)$ is (anti)- periodic or period 1. It is bounded on the set $|\Im(z)| > 1$ and has poles at all integers. Thus

$$I(z) = \sin(\pi z) S(z)$$

is periodic with period 1 and satisfies

$$|I(z)| \leq C e^{\pi |\Im(z)|} \quad .$$

Thus $J(z) = I(log(z)/2\pi i)$ is a holomorphic function in $\mathbf{C} \setminus \{0\}$ growing at most of order 1 and

$$|I(z)| \leq C \max(|z|^{\frac{1}{2}}, |z|^{-\frac{1}{2}}) \quad .$$

Thus $J(z)$ is constant. Evaluating J at 0 gives π for this constant. Hence

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi x)} \quad .$$

Therefore Γ has no zeroes. ■

Lemma 10 *For each $z \in \mathbf{C}$ the vector space of distributions in $\mathcal{S}'(\mathbf{R})$ which are homogeneous of degree z has dimension 2.*

Proof

We set $n = 1$. If Λ is homogeneous of degree z, then $\mathcal{F}\Lambda$ is homogeneous of degree $-n-z$. Since $\mathcal{F}$ is an isomorphism of $\mathcal{S}'(\mathbf{R}^n)$ (the inverse being $\overline{\mathcal{F}}$) it suffices to prove the statement for $z > -1$.

Since we have seen the (linearly independent) distribution Λ_z^+ and Λ_z^- for $z > -1$, the main point is to prove that the vector space of homogeneous distributions of degree z is at most 2 dimensional.

Let $z > -1$ and $\Lambda \in \mathcal{S}'(\mathbf{R})$ be a homogeneous distribution of degree z. Let $C_p(\mathbf{R})$ be the class of continuous functions on $\mathbf{R}$ which grow at most polynomially (i.e. are bounded by a polynomial). In particular the primitive of every element in $L^2(\mathbf{R})$ is in $C(\mathbf{R}) \cap \mathcal{S}'(\mathbf{R})$. We know that we can write Λ as

$$\sum_{\alpha,\beta \leq N} X^\alpha D^\beta f_{\alpha,\beta}$$

for some large $N > 1$ with $f_{\alpha,\beta} \in C(\mathbf{R}) \cap \mathcal{S}'(\mathbf{R})$. By Leibniz' rule we can write it as

$$\sum_{\alpha,\beta \leq N} D^\alpha X^\beta g_{\alpha,\beta}$$

with $g_{\alpha,\beta} \in C(\mathbf{R}) \cap \mathcal{S}'(\mathbf{R})$. Adding over β gives the expression

$$\sum_{\alpha \leq N} D^\alpha g_\alpha$$

Taking primitives and summing over α we can finally write this as

$$\Lambda = D^N g$$

for some $g \in C(\mathbf{R}) \cap \mathcal{S}'(\mathbf{R})$.

Now we have for each $\lambda > 0$

$$D^N g = S^z_\lambda D^N g = D^N S^{z+N}_\lambda g$$

Thus the distribution $g - S^{z+N}_\lambda g$ (which clearly is a continuous function) vanishes on all Schwartz functions of the form $D^N h$. This is a co-N- dimensional vector space of $\mathcal{S}(\mathbf{R})$. Thus we conclude $g - S^{z+N}_\lambda g$ is member of a N-dimensional vector space of tempered distributions which is easily identified as the space of polynomials of degree at most $N-1$. Thus $g - S^{z+N}_\lambda g = \sum_{k=0}^{N-1} a_k(\lambda) x^k$ in the sense of distributions and thus in the sense of continuous functions. It is easy to see that for each $\lambda > 0$, $\lambda \neq 1$ there is a unique continuous function g_λ differing from g by a polynomial of degree at most $N-1$ such that

$$g_\lambda - S^{z+N}_\lambda g_\lambda = 0 \quad ,$$

namely

$$g - g_\lambda = \sum_{k=0}^{N-1} \frac{a_k(\lambda)}{1 - \lambda^{N-k+z}} \quad .$$

By iteration we have

$$g_\lambda - S^{z+N}_{\lambda^k} g_\lambda = 0$$

for all integers k. By uniqueness we have $g_{\lambda^k} = g_\lambda$ for all integers k and thus $g_\lambda = g_{\lambda'}$ whenever λ' is a rational power of λ. By continuity of g_2 we have

$$g_2 - S^{z+N}_\lambda g_2 = 0$$

for all positive λ. Thus $g_2(x) = |x|^{z+N} g_2(x/|x|)$ and $g_2(0) = 0$ by continuity. Thus $D^{N-1}g_2(x)$ is continuous and member of the 2 dimensional vector space V of continuous functions homogeneous of degree $z+1$. Since $D^N g_2 = \Lambda$, we have the D maps V surjectively to the space of homogeneous distributions of degree z. This finishes the proof.

■

We now give various descriptions of the two dimensional space of homegenous distributions of degree z.

For z not a negative integer, we do have introduced two homogeneous distributions Λ^z_+ and Λ^z_-. For $\Re(z) > -1$ they are given by locally integrable functions.

Suppose $\Re(z) \leq -1$ and $N \geq -\Re(z)$. Partial integration shows that if f vanishes of order N at 0 then we have also the integral formulas

$$\Lambda^z_+(f) = \int_{t>0} t^z f(t)\, dt \quad ,$$

$$\Lambda^z_+(f) = \int_{t<0} |t|^z f(t)\, dt \quad .$$

This describes these distributions on a co-finite dimensional subspace. Obviously the two distributions are nonzero and linearly independent. A general Schwartz function $f(x)$ can be written as $g(x) + P(x)e^{-\pi x^2}$ with a polynomial of degree $N-1$ and g vanishing of order N at 0. Thus to extend Λ^+_z and Λ^-_z to all of $\mathcal{S}(\mathbf{R}^n)$ one only needs to know the value of $\Lambda^+_z(f_k)$ for the functions $f_k(x) = x^k e^{-\pi x^2}$ for $0 \leq k \leq N-1$.

For each z let $(t)^z_+$ be the holomorphic extension of the function t^z on $\mathbf{R}^+$ to the slit plane $\mathbf{C} \setminus \{is : s \leq 0\}$. Similarly let $(t)^z_-$ be the holomorphic extension of the function t^z on $\mathbf{R}^+$ to the slit plane $\mathbf{C} \setminus \{is : s \geq 0\}$. By partial intagration we see that

$$\Lambda^{(+)}_z f = \lim_{\epsilon \to +0} \int_{\mathbf{R}} (t + i\epsilon)^z_+ f(t)\, dt$$

and

$$\Lambda^{(-)}_z f = \lim_{\epsilon \to -0} \int_{\mathbf{R}} (t + i\epsilon)^z_- f(t)\, dt$$

exist and describe tempered distributions homogeneous of degree z. These linear functionals are obviously homogeneous of degree z.

Let $f(x) = e^{-\pi x^2}$. By partial integration one observes

$$\Lambda^{(+)}_z f = \int_{\mathbf{R}} (t + i\epsilon)^z_+ f(t + i\epsilon)\, dt$$

for $\epsilon > 0$ and

$$\Lambda^{(-)}_z f = \int_{\mathbf{R}} (t + i\epsilon)^z_- f(t + i\epsilon)\, dt$$

for $\epsilon < 0$, the right hand sides being independent of ϵ. This observation holds for any appropriate function instead of $e^{-\pi x^2}$ which extends analytically to the complex plane.

If z is not an integer, $\Lambda_z^{(+)}$ and $\Lambda_z^{(-)}$ are easily seen to be linearly independent. Let z be negative integer. Let again $f(x) = e^{-\pi x^2}$, then

$$\Lambda_z^{(-)} f - \Lambda_z^{(+)} f = \int_\Gamma y^z f(y)\, dy = 2\pi i(-1-z)!\, f^{(-1-z)}(0)$$

here Γ is any contour counterclockwise around 0 and the identity holds by Cauchy's integral formula. Observe that this embeds the derivatives of the Dirac distribution into an analytic family of distributions.

The distribution

$$\frac{1}{2}\left(\Lambda_z^{(-)} f + \Lambda_z^{(+)} f\right) =: p.v. \int_{\mathbf{R}} t^z f(t)\, dt$$

is called the principal value distribution. It is easy to verify that

$$p.v. \int t^{-1} f(t)\, dt = \lim_{\epsilon \to 0} \int_{\mathbf{R}\setminus[-\epsilon,\epsilon]} t^{-1} f(t)\, dt$$

for all $f \in \mathcal{S}(\mathbf{R})$.

0.8 The Riemann ζ function

Consider the homogeneous distribution

$$\Lambda_z = \Lambda_z^+ + \Lambda_z^-$$

For $\Re(z) > 0$ this is an even and real distribution,

$$\Lambda_z f = \Lambda_z f^- = \overline{\Lambda_z \overline{f}}$$

with $f^-(x) = f(-x)$ hence it is even and real for all z. Its Fourier transform is also real and even, so it follows easily that

$$\mathcal{F}\Lambda_z = c(z)\Lambda_{-1-z} \quad .$$

To obtain the (meromorphic) function $c(z)$ we apply this distribution to $g(t) = e^{-\pi t^2}$. Hence

(8) $$(\Lambda_{-1-z} g)\mathcal{F}\Lambda_z = (\Lambda_z g)\Lambda_{-1-z} \quad .$$

The idea of the Riemann zeta function is to apply Λ_{-z} formally to the distribution $\Lambda - \delta_0$ where Λ is the Poisson distribution defined by $\Lambda f = \sum_{z \in \mathbf{Z}} f(z)$ and δ_0 is the Dirac distribution:

$$\zeta(z) = \frac{1}{2}(\Lambda_{-z}, \Lambda - \delta_0) = \sum_{n \geq 1} \frac{1}{n^z} \quad .$$

The last equality gives sense to the a priori meaningless pairing of the two distributions in the case $\Re(z) > 1$. Since $\mathcal{F}\Lambda = \Lambda$, equation 8 suggests that if the function

$$(\Lambda_{z-1}, g)(\Lambda_{-z}, \Lambda)$$

extends to a meromorphic function on $\mathbf{C}$ we expect it to be invariant under $z \to 1-z$. Once we know what we are going for we can verify it as follows. We have for $\Re(z) > 1$

$$\begin{aligned}
&(\Lambda_{z-1}, g)(\Lambda_{-z}, \Lambda) \\
&= 2\sum_{n>0} \frac{1}{n^z} \int_{t>0} t^z e^{-\pi t^2}\, dt/t = 2\sum_{n>0} \int_{t>0} t^z e^{-\pi t^2 n^2}\, dt/t \\
&= 2\sum_{n>0} \int_1^\infty t^z e^{-\pi t^2 n^2}\, dt/t + \int_1^\infty t^{-z}((2\sum_{n>0} e^{-\pi t^{-2} n^2}) + 1 - 1)\, dt/t \\
&= 2\sum_{n>0} \int_1^\infty t^z e^{-\pi t^2 n^2}\, dt/t + \int_1^\infty t^{-z}(t(2\sum_{n>0} e^{-\pi t^2 n^2}) + t - 1)\, dt/t \\
&= 2\sum_{n>0} \int_1^\infty t^z e^{-\pi t^2 n^2}\, dt/t + 2\int_1^\infty t^{1-z} \sum_{n>0} e^{-\pi t^2 n^2}\, dt/t + \frac{1}{z} + \frac{1}{(1-z)} \quad .
\end{aligned}$$

This expression is obviously meromorphic on $\mathbf{C}$ and invariant under $z \to 1-z$. It follows that the Riemann ζ function extends to a meromorphic function on $\mathbf{C}$. The famous Riemann hypothesis is the conjecture that all zeroes ρ of ζ satisfy $\Re(\rho) = 1/2$.

0.9 Paley Wiener theorems

The theme of Paley Wiener theorems is to characterize the Fourier transform of compactly supported functions or distributions. To begin with, we would like to characterize the space $\mathcal{F}C_c^\infty(\mathbf{R}^n)$, i.e. the space of all Schwartz functions which are the Fourier transform of a compactly supported Schwartz functions.

Theorem 2 *The space $\mathcal{F}C_c^\infty(\mathbf{R}^n)$ consists of all functions $f \in \mathcal{S}(\mathbf{R}^n)$ which have a holomorphic extension to $\mathbf{C}^n$, also denoted by f, satisfying*

$$\|f(.+iy)\|_{(N)} \le C_N e^{C|y|} \tag{9}$$

for all $N \ge 0$ and appropriate constants C, C_N. The constant C can be chosen as 2π times the radius of a ball containing the support of the inverse Fourier transform of f.

Before we prove the theorem, we start with a few remarks:

A function f on $\mathbf{C}^n$ is called holomorphic if it is holomorphic in each argument. A basic theorem in the theory of holomorphic functions of several variables is that this definition is invariant under rotation of f, i.e., it is equivalent to requiring $f(x + tx')$ is holomorphic in t for all $x, x' \in \mathbf{C}^n$. We shall take this theorem for granted.

Under the assumption that f is holomorphic inequality (9) is equivalent to

$$\|X^N f(.+iy)\|_\infty \le C_N e^{C|y|} \tag{10}$$

for all $N \ge 0$, because one can calculate partial derivatives as

$$D_j f(z) = c\int_\Gamma \frac{f(\zeta)}{(\zeta - z_j)^2}\, d\zeta$$

where Γ is the unit circle centered at z in the plane through z which is parallel to the j-th coordinate plane. This is the dual phenomenon to the observation that for compactly supported Schwartz functions one only needs to bound the partial derivatives $D^\alpha f$ in order to control the (N)- norms.

We scetch a proof that (10) implies (9). We have to show

$$\left|X^\alpha D^\beta f(x+iy)\right| \le Ce^{c|y|}$$

for all multi indices α, β, where x and D act in the x variable. We may assume that whenever $|x_j| < 2$ we have $\alpha_j = 0$. Let $z = x + iy$, Z_j the multiplication operator by z_j, and Z^α as usual. It suffices to show

$$\left|Z^\alpha D^\beta f(z)\right| \le C\left(\prod_j |z_j/x_j|^{\alpha_j}\right) e^{c|y|} \quad .$$

Observe that applied to a holomorphic function D^β can also be viewed as complex partial differentiation. Using Leibniz' rule it suffices to show for every $\alpha' \le \alpha$ and $\beta' \le \beta$

$$\left|D^{\beta'} Z^{\alpha'} f(z)\right| \le C\left(\prod_j |z_j/x_j|^{\alpha}_j\right) e^{c|y|} \quad .$$

By the Cauchy integral formula there is a probability measure on the ball of radius 1 around z such that the left hand side is up to a universal scalar equal to

$$\int_{B_1(z)} Z^{\alpha'} f(z)\, d\mu \quad .$$

Let $z' \in B_1(z)$. Since $|x_j| \ge 2$ whenever $\alpha'_j \ne 0$, we have that z_j and z'_j as well as x_j and x'_j are comparable whenever $\alpha'_j \ne 0$. Apply (10) with α', then we have for all $z' \in B_1(z)$:

$$|z'|^{\alpha'} |f(z')| \le C\left(\prod_j |z_j/x_j|^{\alpha'_j}\right) e^{c|y|} \quad .$$

This implies the desired inequality.

We turn to the proof of the Paley Wiener theorem.

Proof

Let $f = \mathcal{F}g$ with $g \in C_c^\infty(\mathbf{R})$, we aim to show that f extends to an entire function and satisfies (10). For $z \in \mathbf{C}^n$ define

$$f(z) = \int_{\mathbf{R}^n} g(\zeta) e^{-2\pi i <z,\zeta>}\, d\zeta \quad .$$

This integral is defined in the Riemann sense since the integrand is continuous and compactly supported, and it is continuous in z since the support of the integrand is constant and the integrand varies continuously in z w.r.t. the sup norm. Obviously this definition of f coincides on $\mathbf{R}^n$ with $\mathcal{F}g$.

Let $\gamma : [0,1] \to \mathbf{C}^n$ be a closed differentiable curve in C^n such that all coordinates $\gamma_j(t)$ with $j \neq 1$ are constant. Then by Fubini

$$\int_0^1 f(z)\gamma'(t)\,dt = \int_{\mathbf{R}^n}\int_0^1 g(\zeta)e^{-2\pi i<z,\zeta>}\gamma'(t)\,dt\,d\zeta \quad .$$

This vanishes because $e^{-2\pi i<z,\zeta>}$ is holomorphic. Hence $f(z)$ is holomorphic by Morera's theorem.

Integrations by part give

$$z^\alpha f(z) = \int_{\mathbf{R}^n} D^\alpha g(\zeta)e^{-2\pi i<z,\zeta>}\,d\zeta \quad ,$$

hence

$$|x^\alpha||f(x+iy)| \leq \|D^\alpha g\|_1 e^{C|y|} \quad ,$$

which implies (9) by one of the previous remarks.

Now assume that $f : \mathbf{C}^n \to \mathbf{C}$ satisfies (9). Let g be the inverse Fourier transform of the restriction of f to $\mathbf{R}^n$, then obviously $g \in \mathcal{S}(\mathbf{R}^n)$. We have to show that g is compactly supported.

Let $x \in \mathbf{R}^n$. Then

$$g(x) = \int_{\mathbf{R}^n} f(\xi)e^{2\pi i<\xi,x>}\,d\xi \quad .$$

Let $\eta \in \mathbf{R}^n$. By a change of contour we can replace this integral by

$$g(x) = \int_{\mathbf{R}^n} f(\xi+i\eta)e^{2\pi i<\xi+i\eta,x>}\,d\xi \quad .$$

(We change the contour successively in each coordinate, using inequality (9)). Thus we have

$$|g(x)| \leq \int_{\mathbf{R}^n}(1+|\xi|^2)^{-n}e^{C|\eta|}e^{-2\pi<x,\eta>}\,d\xi \quad .$$

If we choose η in direction of x we obtain

$$|g(x)| \leq \int_{\mathbf{R}^n}(1+|\xi|^2)^{-n}e^{(C-2\pi|x|)|\eta|}d\xi \quad .$$

The right hand side tends to 0 as $|\eta|$ tends to ∞ provided $|x|$ is large enough. This proves compact support of g.

■

If $f \in C_c^\infty$, then the holomorphic extension of $\mathcal{F}f$ to $\mathbf{C}^n$ is called the Laplace transform of f. (In the literature the definition of Laplace transform may vary, e.g. by factors of i in the argument and other details)

Remark: Suppose $g \in \mathcal{S}(\mathbf{R}^n)$ is supported in a circle of radius r about the origin. Then the Laplace transform f of the translate $T_s g$, which is supported in a circle of radius r around s, satisfies

$$\|X^N f(.+iy)\|_\infty \leq C_N e^{2\pi(r|y|+<y,s>)}$$

because

$$\mathcal{F}(T_s g)(z) = e^{-2\pi i<z,s>}\mathcal{F}g(z) \quad .$$

A distribution $\Lambda \in \mathcal{S}'(\mathbf{R}^n)$ is said to vanish at a point $x \in \mathbf{R}^n$ if there is an open neighborhood U of x such that $\Lambda(f) = 0$ for all $f \in \mathcal{S}(\mathbf{R}^n)$ supported in U. The set U of all points where Λ vanishes is open. Its complement is called the support of Λ. If the support of a compactly supported Schwartz function is contained in the complement of the support of Λ, then $\Lambda(f) = 0$. This follows by covering the support of f by finitely many balls $B_{r_k}(x_k)$ such that $\Lambda(g) = 0$ for any g supported in any of the balls $R_{2r_k}(x_k)$, and then writing f as a sum of Schwartz functions each supported in one of the balls. E.g., this decomposition can be achieved by iteratively replacing f by $f - f\phi_k$ where ϕ_k is supported on $B_{2r_k}(x_k)$ and equal to 1 on $B_{r_k}(x_k)$.

Now let f be any Schwartz function supported in the complement of the support of Λ. Since we can approximate f by $f = \lim_{\lambda\to\infty} f D_\lambda^\infty \phi$ where ϕ is compactly supported and $\phi(0) = 1$ we obtain $\Lambda(f) = 0$.

The analogue of the previous theorem for distributions is given by

Theorem 3 *A distribution in $\mathcal{S}'(\mathbf{R}^n)$ is the Fourier transform of a compactly supported distribution Λ if and only if it is represented by a continuous function f which extends holomorphically to $\mathbf{C}^n$ and satisfies*

$$|f(x+iy)| \leq C(1+|x|^2)^N e^{C|y|} \quad , \tag{11}$$

for some $N > 0$, where C can be chosen to be any number larger than $2\pi r$ if Λ is supported in $B_r(0)$.

Proof Assume Λ is a tempered distribution supported in $B_r(0)$. We write it as

$$\Lambda = \sum_{|\beta|\leq N} D^\beta g_\beta$$

with continuous functions f_β. Pick $C > 2\pi r$ and $\phi \in \mathcal{S}(\mathbf{R}^n)$ supported in $B_C(0)$ such that ϕ is constant equal to 1 on an open neighborhood of the support of Λ. Define for $z \in \mathbf{C}^n$

$$f(z) = \Lambda(\phi e^{-2\pi i<z,.>})$$

This function is continuous and holomorphic in $\mathbf{C}^n$ as one can see by similar arguments as before from writing it as the integral

$$f(z) = \sum_{|\beta|\leq N} \int_{\mathbf{R}^n} g_\beta(x)(-1)^\beta D^\beta(\phi e^{-2\pi i<z,.>})(x)\, dx \quad .$$

By the Leibniz rule we can find certain ϕ_β with compact support such that

$$f(z) = \sum_{|\beta|\leq N} \sum_{\alpha\leq\beta} \int_{\mathbf{R}^n} g_\beta(x)\phi_{\alpha,\beta} z^\alpha e^{-2\pi i<z,.>}(x)\, dx \quad .$$

This easily proves the estimate (11) for f. We still have to see that f restricted to $\mathbf{R}^n$ is the Fourier transform of Λ. However, we have from the last expression

$$f = \sum_{|\beta|\le N} \sum_{\alpha\le\beta} (X^\alpha g_\beta \phi_{\alpha,\beta}) \quad ,$$

because for compactly supported continuous functions the Fourier transform (in the distributional sense) is given by the Fourier integral, as one can see from the continuity of the Fourier integral from $L^1(\mathbf{R}^n)$ into $L^\infty(\mathbf{R}^n)$. This implies $f = \mathcal{F}\Lambda$ by undoing the previous calculations in the distributional sense.

Now let conversely f be a holomorphic function on $\mathbf{C}^n$ satisfying (11) with certain C. Then f restricted to $\mathbf{R}^n$ is a distribution. Let $B_\epsilon(s)$ be any ball outside the circle of radius $C/2\pi$ about 0. We pair $\overline{\mathcal{F}}f$ with an arbitrary Schwartz function g supported in $B_\epsilon(s)$. Then we have

$$(\overline{\mathcal{F}}f, g) = \int_{\mathbf{R}^n} f(x)\mathcal{F}g(-x)\, dx \quad .$$

By a change of contour integrals (using decay estimates in x) we obtain

$$|(\overline{\mathcal{F}}f, g)| = |\int_{\mathbf{R}^n} f(x+iy)\mathcal{F}g(-x-iy)\, dx|$$

$$\le C|\int_{\mathbf{R}^n} (1+|x|^2)^N e^{2\pi r|y|}(1+|x|^2)^{-N-n} e^{2\pi\epsilon|y| - <s,y>}\, dx| \quad .$$

This tends to 0 if y pointss in direction of s and $|y|$ tends to ∞. This proves that the support of $\overline{\mathcal{F}}f$ is contained in $B_{C/2\pi}(0)$. Letting C tend to $2\pi r$ we conlude that the support of $\overline{\mathcal{F}}f$ is contained in $B_r(0)$.

■

0.10 Distributions versus tempered distributions

Often it is necessary to have a more localized concept than tempered distributions. E.g. any continuous function on $\mathbf{R}^n$ looks locally like a tempered distribution, but may fail to be a tempered distribution because of growth at infinity. Or one would like to introduce distributions on manifolds.

Let U be an open set in $\mathbf{R}^n$. Following the classical paper by Schwartz we write $\mathcal{D}(U)$ for the set of functions in $C_c^\infty(U)$.

Definition 6 *A linear functional Λ on the space $\mathcal{D}(U)$ is called distribution (on U), if for every $\phi \in \mathcal{D}(U)$ the functional*

$$f \mapsto \Lambda(\phi\, f|_U), \quad \mathcal{S}(\mathbf{R}^n) \to \mathbf{C}$$

is a tempered distribution.

As with $\mathcal{S}(\mathbf{R}^n)$ one can introduce a topology on $\mathcal{D}(U)$ which makes the space of distributions coincide with the space $\mathcal{D}'(U)$ of continuous linear functionals on $\mathcal{D}(U)$. We shall not discuss this topology in detail now, but adopt the notation $\mathcal{D}'(U)$.

Observe that $\mathcal{D}'(\mathbf{R})$ is strictly larger than $\mathcal{S}'(\mathbf{R}^n)$, e.g. it contains all continuous functions. Operations like D^α and X^α can easily be defined on $\mathcal{D}(U)$, also M_η, T_y, S^z_λ, the latter two mapping into $\mathcal{D}(U')$ with possibly different U'.

The Fourier transform of an element in $\mathcal{D}(\mathbf{R}^n)$ is in general not a distribution. The formal identity

$$(\mathcal{F}\Lambda, f) = (\Lambda, \mathcal{F}f)$$

suggests that for $\Lambda \in \mathcal{D}(\mathbf{R}^n)$, $\mathcal{F}\Lambda$ is a linear functional on the space $\mathcal{F}(\mathcal{D}(U))$, which we have identified in the theorem of Paley Wiener. By transport of structure we can also equip this space with a topology so that $\mathcal{F}(\mathcal{D}'(\mathbf{R}^n))$ becomes the space of continuous linear functionals.

Lemma 11 *A linear functional Λ on $\mathcal{D}(U)$ is in $\mathcal{D}'(U)$ iff for every compact set $K \subset U$ there is a N and a C such that*

$$|\Lambda(f)| \leq C \sup_{|\alpha|\leq N} \|D^\alpha f\|_\infty \tag{12}$$

for all $f \in \mathcal{D}(U)$ with support in K.

Proof

Suppose $\Lambda \in \mathcal{D}'(U)$. Let $K \subset U$ be compact. We can find a $\phi \in \mathcal{D}(U)$ which is constant equal to 1 on an open neighborhood of K. Then $\Lambda\phi$ is in $\mathcal{S}(\mathbf{R}^n)$ (in the sense of the above definition), hence for each $f \in \mathcal{D}(U)$ supported in K:

$$|\Lambda(f)| = |\Lambda(\phi f)| \leq C\|f\|_{(N)} \leq C \sup_{|\alpha|,|\beta|\leq N} \|X^\alpha D^\beta f\|_\infty$$

for some N. This implies (12) because x^α is bounded on K.

Vice versa assume that Λ satisfies (12). Let $\phi \in \mathcal{D}(U)$ and let K be the support of ϕ. Then (12) implies for every $f \in \mathcal{S}(\mathbf{R}^n)$:

$$|\Lambda(\phi f)| \leq C_N \|\phi f\|_{(N)} \quad .$$

Since multiplication with ϕ is continuous in $\mathcal{S}(\mathbf{R}^n)$, the right hand side is bounded by $C'\|f\|_{(N')}$ for some N', which implies that $\Lambda\phi$ is a tempered distribution

■

If N in the previous lemma can be chosen independently of K, then the minimal such N is called the (regularity) degree of the distribution. A distribution may not have finite degree, as the example

$$\sum_{n>0} T_n D^n \delta_0$$

in $\mathcal{D}'(\mathbf{R})$ shows.

Lemma 12 *Let U and V be two open sets in $\mathbf{R}^n$ and let $\Phi : U \to V$ be a diffeomorphism. Then there is an isomorphism of vector spaces*

$$\Phi^* : \mathcal{D}'(U) \to \mathcal{D}'(V)$$

defined by

$$\Phi_*(\Lambda)(\phi) = \Lambda(\phi \circ \Phi) \quad .$$

Proof All that needs to be shown is that $\Phi_*\Lambda$ is an element of $\mathcal{D}'(V)$ if $\Lambda \in \mathcal{D}'(U)$. Then the map is clearly linear and invertible with inverse given by the analogous construction with Φ^{-1}.

If ϕ is supported in a compactum, then $\phi \circ \Phi$ is supported in a compactum, hence

$$|\Phi_*(\Lambda)(\phi)| \leq C \sup_{|\alpha|<N} \|D^\alpha(\phi \circ \Phi)\|_\infty$$

for some $N > 0$. However, we can easily prove by induction on $|\alpha|$ that

$$D^\alpha(\phi \circ \Phi) = \sum_{|\beta| \leq |\alpha|} ((\psi_\beta D^\beta \phi) \circ \Phi)$$

for some smooth functions ψ_β on U. Namely,

$$D_j(\phi \circ \Phi) = \sum_k J(\Phi)_{j,k}((D_k\phi) \circ \Phi) \quad ,$$

$$\sum_k ((J(\Phi)_{j,k} \circ \Phi^{-1}) D_k \phi) \circ \Phi) \quad ,$$

where $J(\Phi)$ is the Jacobian of Φ. The statement for general α follows by induction and the Leibniz rule.

Since smooth functions on V are bounded on any compact subset of V, we obtain

$$\|D^\alpha(\phi \circ \Phi)\|_\infty \leq C \sum_{|\beta|<|\alpha|} \|D^\beta \phi\|_\infty \quad .$$

This proves that $\Phi_*\Lambda \in \mathcal{D}'(V)$. ∎

Let U_1 and U_2 be two open sets in $\mathbf{R}^n$ and let $V \subset U_1 \cap U_2$ be open. If $\Lambda_j \in \mathcal{D}'(U_j)$ for $j = 1, 2$, then then Λ_1 and Λ_2 are said to coincide on W if

$$\Lambda_1 \phi = \Lambda_2 \phi$$

for all $\phi \in \mathcal{D}(V) \subset \mathcal{D}(U_1) \cap D(U_2)$.

A n-dimensional smooth manifold M is a Hausdorff topological space with a covering by open sets U_j together with homeomorphisms $\Phi_j : U_j \to V_j \subset \mathbf{R}^n$ with open sets V_j such that whenever $U_j \cap U_k \neq \emptyset$ then

$$\Phi_k \circ \Phi_j^{-1} : \Phi(U_j \cap U_k) \to \Phi_k(U_k \cap U_j)$$

is a diffeomorphism.

A distribution Λ on M is defined to be a collection Λ_j of distributions in $\mathcal{D}(V_j)$ such that whenever $U_j \cap U_k \neq \emptyset$ we have that Λ_k and $(\Phi_k \circ \Phi_j^{-1})_* \Lambda_j$ coincide on $\Phi_k(U_k \cap U_j)$. If ϕ is a compactly supported smooth function on M, we apply Λ by writing $\phi = \sum \phi_j$ (the sum can be asusmed to be finite by compact support of ϕ) with ϕ_j supported in U_j and defining

$$\Lambda(\phi) = \sum_j \Lambda_j(\phi_j \circ \Phi_j^{-1}) \quad .$$

The concept of support of a tempered distribution translates literally to the space of distributions. In particular the set of compactly supported tempered distributions coincides with the set of compactly supported distributions.

A related and dual concept to that of distributions is that of compactly supported distributions. Let $\mathcal{E}(\mathbf{R}^n)$ be the set of all C^∞ functions on $\mathbf{R}^n$ if Λ is a compactly supported distribution, with support contained in a compact set K say, then we can pair this distribution with any element f in $\mathcal{E}(\mathbf{R}^n)$ by picking a function $\phi \in \mathcal{D}(\mathbf{R}^n)$ which is constant equal to 1 on a neighborhood of K and defining

$$\Lambda(f) = \Lambda(\phi f) \quad .$$

This definition is independent of the particular choice of ϕ, because if ϕ' is another choice, then

$$\Lambda(\phi f) - \Lambda(\phi' f) = \Lambda((\phi - \phi')f) = 0$$

because the support of $(\phi - \phi')f$ is disjoint from the support of Λ. Thus Λ defines a linear functional on $\mathcal{E}(\mathbf{R}^n)$. Again, one can define a topology on $\mathbf{R}^n$ which makes the set of compactly supported distributions be the dual space of $\mathcal{E}(\mathbf{R}^n)$, which we denote by $\mathcal{E}'(\mathbf{R}^n)$.

0.11 Convolution

We shall work with tempered distributions in this section.

The convolution of two functions $f, g \in \mathcal{S}(\mathbf{R}^n)$ is defined by the absolutely convergent integral

$$f * g(x) = \int_{\mathbf{R}^n} f(x-t)g(t)\,dt \quad .$$

It is easy to see by the Lebesgue dominated convergence theorem that $f * g$ is a continuous function. By Fubini we have

$$\|f * g\|_1 \leq \int_{\mathbf{R}^n} \int_{\mathbf{R}^n} |f(x-t)g(t)|\,dt\,dx \leq \|f\|_1 \|g\|_1 \quad ,$$

(with actual equality if f and g are nonnegative real) in particular $f * g \in L^1(\mathbf{R}^n)$ and $f * g$ is a distribution. Since it is in $L^1(\mathbf{R}^n)$, the Fourier transform of $f * g$ is given by

$$\mathcal{F}(f * g)(\xi) = \int_{\mathbf{R}^n} \int_{\mathbf{R}^n} f(x-t)g(t)e^{-2\pi i <\xi, x-t>} e^{-2\pi i <\xi, t>}\,dt\,dx = \mathcal{F}(f)(\xi)\mathcal{F}(g)(\xi) \quad .$$

Since the product of two Schwartz functions is a Schwartz function, we see that $\mathcal{F}(f * g)$ as well as $f * g$ are Schwartz functions. Moreover we immediately see the identities

$$f * g = g * f \quad ,$$

$$(f * g) * h = f * (g * h) \quad ,$$

$$D^\alpha (f * g) = (D^\alpha f) * g = f * D^\alpha g \quad .$$

Direct calculaion shows for Schwartz functions

$$\int_{\mathbf{R}^n} f * g(x) h(x) = \int g(x) \check{f} * h(x)\, dx$$

where $\check{f}(x) = f(-x)$. Now let $f \in \mathcal{S}(\mathbf{R}^n)$ and $\Lambda \in \mathcal{S}'(\mathbf{R}^n)$. We can formally define a distribution $f * \Lambda = \Lambda * f$

$$(f * \Lambda)(\phi) = \Lambda(\check{f} * \phi) = \mathcal{F}\Lambda(\mathcal{F} f \overline{\mathcal{F}} \phi)$$

where $\check{f}(x) = f(-x)$. If Λ is a Schwartz function, this definition ofcourse coincides with the previous one.

Lemma 13 *The distribution $f * \Lambda$ is a continuous function given by*

$$f * \Lambda(x) = \Lambda(T_x \check{f}) \tag{13}$$

Proof First we observe that the right hand side of (13) is a continuous function which describes a distribution. There is a N such that

$$\Lambda(T_x \check{f}) \leq C \|T_x \check{f}\|_{(N)}$$

$$\leq C \sup_{|\alpha|,|\beta| \leq N} \sup_{y \in \mathbf{R}^n} |(y + x)^\alpha D^\beta \check{f}(y)| \leq C \sup_{|\alpha'| \leq N} |x^{\alpha'}| \|\check{f}\|_{(N)} \quad .$$

Thus the right hand side of (13) grows at most polynomially in x. To show that it is continuous in x it suffices by an easy translation argument to show that it is continuous at $x = 0$. We have

$$|\Lambda(T_x \check{f}) - \Lambda(\check{f})| \leq C \sup_{|\alpha|,|\beta| \leq N} \|X^\alpha D^\beta (T_x \check{f} - \check{f})\|_{(N)}$$

Since $D^\beta \check{f} \in \mathcal{S}(\mathbf{R}^n)$ it suffice to prove for any Schwartz function f and any α that

$$\|X^\alpha (T_x f - f)\|_\infty$$

tends to 0 as $x \to 0$. Assume $|x| < 1$ we can estimate this by

$$|x| \sup_{y \in \mathbf{R}^n} \sup_{\xi \in B_1(y)} |y^\alpha f'(y)| \leq C|x|$$

This proves continuity of (13) at $x = 0$.

Let $f, g \in \mathcal{S}(\mathbf{R}^n)$ and assume for simplicity that f is compactly supported. We claim

$$f * g = \lim_{\epsilon \to 0} \epsilon^n \sum_{x \in \epsilon \mathbf{Z}^n} f(x) T_x g$$

where the limit is in the sense of $\mathcal{S}(\mathbf{R}^n)$ and the sum is actually a finite one. (This is a vector valued Riemann sum). This claim proves the lemma immediately because for all $\phi \in \mathcal{D}(\mathbf{R}^n)$ we have by continuity of Λ

$$f * \Lambda(\phi) = \Lambda(\phi * \check{f}) = \lim_{\epsilon \to 0} \epsilon^n \sum_{x \in \epsilon \mathbf{Z}^n} \phi(x) \Lambda(T_x \check{f})$$

$$= \int_{\mathbf{R}^n} \phi(x) \Lambda(T_x \check{f})) \, dx = (\phi, \Lambda(T_x \check{f}) \quad .$$

Here we have also used an ordinary Riemann sum. Since on both sides of this equation we have a distribution applied to ϕ, this equality holds for all $\phi \in \mathcal{S}(\mathbf{R}^n)$.

We have to see the above claim. We have to show

$$\|f * g - \sum_{x \in \epsilon \mathbf{Z}^n} f(x) T_x g\|_{(N) \to 0}$$

It suffices to show for every $g \in \mathcal{S}(\mathbf{R}^n)$ and α:

$$\|X^\alpha f * g - \sum_{x \in \epsilon \mathbf{Z}^n} f(x) X^\alpha T_x g\|_\infty \to 0 \quad .$$

We estimate this by

$$\|X^\alpha((f - f_\epsilon) * g)\| + \|X^\alpha(f_\epsilon * g - \sum_{x \in \epsilon \mathbf{Z}^n} f(x) X^\alpha T_x g)\|_\infty \to 0 \quad .$$

Here $f_\epsilon(x) = f(x')$ if x' is the nearest point to x on the lattice $\epsilon \mathbf{Z}^n$ and convolution with f_ϵ is defined in L^1 sense.

Since $\|(f - f_\epsilon)\| \leq C\epsilon$, $X^\alpha g$ is bounded, and f is compactly supported we see that

$$\|X^\alpha((f - f_\epsilon) * g) \to 0$$

as $\epsilon \to 0$. Finally we have

$$\|X^\alpha(f_\epsilon * g - \sum_{x \in \epsilon \mathbf{Z}^n} f(x) X^\alpha T_x g)\|_\infty$$

$$\leq \sup_y |y^\alpha| \sum_{x \in \epsilon Z} |f(x)| |g(y - x) - \sup_{|t| < \epsilon} g(x - y + t)|$$

Again we can estimate the difference of the two evaluations of g by ϵ times the derivative of g nearby. Thus this expression also tends to 0 as ϵ tends to 0. This finishes the proof of the claim and therefore the proof of the lemma. ■

Observe that the continuous function $f * \Lambda$ is in fact C^∞, because its distributional derivative is given by

$$D^\alpha(f * \Lambda) = (D^\alpha f * \Lambda)$$

and is again a continuous function by the Lemma. (The equation holds for $\Lambda \in \mathcal{S}'(\mathbf{R}^n)$ because it holds for $\Lambda \in \mathcal{S}(\mathbf{R}^n)$ by continuity).

Lemma 14 *If $\Lambda \in \mathcal{S}'(\mathbf{R}^n)$ is compactly supported, then $f \to f * \Lambda$ is a continuous map from $\mathcal{S}(\mathbf{R}^n)$ to $S(\mathbf{R}^n)$*

Proof We have to show that multiplication with the (continuous) function $\mathcal{F}(\Lambda)$ is continuous. However, we have by the Paley Wiener theorem

$$\|\mathcal{F}(\Lambda)f\|_{(N)} \leq C \sup_{|\alpha|,|\alpha'|\leq N} \|(1+|X|^2)^N D^\alpha(\mathcal{F}(\Lambda))D^{\alpha'}f\|_\infty$$

$$\leq C \sup_{|\alpha|,|\alpha'|\leq N} \|(1+|X|^2)^N(1+|X|^2)^{N'}D^{\alpha'}f\|_\infty \leq C\|f\|_{2N+N'} \quad .$$

■

Thus convolution with compactly supported distributions can be extended to a map $\mathcal{S}'(\mathbf{R}^n) \to \mathcal{S}'(\mathbf{R}^n)$.

Examples: Let $\delta_x(f) = f(x)$. Then

$$\delta_0 * f = \overline{F}(1\mathcal{F}f) = f$$

$$\delta_x * \delta_y = \overline{F}(\mathcal{F}\delta_x\mathcal{F}\delta_y) = \overline{F}(\mathcal{F}\delta_{x+y}) = \delta_{x+y} \quad .$$

$$(D^\alpha\delta_0) * f = D^\alpha(\delta_0 * f) = D^\alpha f$$

Rewmark, similarly to the last lemma we can define convolution of elements in $\mathcal{E}'(\mathbf{R})$ with elements in $\mathcal{D}'(\mathbf{R})$.

0.12 The Central Limit Theorem

A finite measure μ on $\mathbf{R}^n$ is a distribution in $\mathcal{S}'(\mathbf{R}^n)$ such that

$$\mu(f) \leq C\|f\|_\infty$$

for all $f \in \mathcal{S}(\mathbf{R}^n)$. The best constant C is called the total variation norm of μ. μ extends to a continuous linear functional on $C_0(\mathbf{R}^n)$.

We observe that $\mu * f$ is a bounded continuous function. Moreover, the map $f \to f * \mu$ is bounded in the L^∞ norm. Also $\mu * f$ is again a measure, because

$$\mu * f(g) = \mu(\check{f} * g) \leq \|\mu\|\|\check{f} * g\|_\infty \leq \|\mu\|\|f\|_1\|g\|_\infty \quad .$$

It follows that $\mu * f \in L^1(\mathbf{R}^n)$. Together with $D^\alpha(\mu * f) \in L^1(\mathbf{R}^n)$ this implies, $f \in C_0(\mathbf{R}^n)$. Thus convolution with μ extends to a bounded operator from $C_0(\mathbf{R}^n)$ to $C_0(\mathbf{R}^n)$. Thus (by a line as above) one can define a finite measure $\mu * \nu$ for two finite measures μ and ν with $\|\mu * \nu\| \leq \|\mu\|\|\nu\|$.

We claim that the Fourier transform of a measure is also given by a bounded continuous function. First we observe that

$$(\mathcal{F}\mu, f) \leq \|\mu\|\|\mathcal{F}f\|_\infty \leq \|\mu\|\|f\|_1 \quad ,$$

therefore $\mathcal{F}\mu$ is represented by an L^∞ function and we have $\|\mathcal{F}\mu\|_\infty \leq \|\mu\|$. If μ is actually in L^1, then we can approximate it by Schwartz functions in norm, so $\mathcal{F}\mu$ is approximated by Schwartz functions in L^∞ norm, therefore $\mathcal{F}\mu \in C_0(\mathbf{R}^n)$. If μ is merely a finite measure, then $\mu * f$ is in L^1 for all $f \in \mathcal{S}(\mathbf{R}^n)$, hence $\mathcal{F}\mu\mathcal{F}f$ is continuous for all f, which implies that $\mathcal{F}\mu$ is continuous. Observe that $\mathcal{F}\mu$ is not necessarily in $C_0(\mathbf{R}^n)$ as the example of the Dirac distribution shows.

A probability measure on $\mathbf{R}^n$ is a finite measure with total variation norm 1 and

$$\mathcal{F}(\mu)(0) = 1 \quad .$$

If μ and ν are probability measures, then so is $\mu * \nu$.

A probability measure is nonnegative, i.e., $\mu(f) \geq 0$ if $f \geq 0$. Namely, if μ is a probability measure, and $g = e^{-\pi x^2}$, then

$$\lim_{\lambda\to\infty} (\mu, D_\lambda^\infty g) = \lim_{\lambda\to\infty} (\mathcal{F}\mu, D_{\lambda^{-1}}^1 \mathcal{F}g) = 1 \quad .$$

Assume $f > 0, \|f\| = 1$ and $\mu(f)$ has nonzero imaginary part. Then

$$\|D_\lambda^\infty g \pm i\epsilon f\|_\infty \leq 1 + \epsilon^2$$

for sufficiently small ϵ, but

$$\Re(\mu(D_\lambda^\infty g \pm i\epsilon f)) = \Re(\mu(D_\lambda^\infty g)) \mp \epsilon \Im(\mu(f)) \quad ,$$

which leads to a contradiction to $\|\mu\| = 1$ for sufficiently small ϵ and large λ. Similarly one sees that $\mu(f)$ cannot be negative.

A probability measure is said to have finite second moment if $\sum_j X_j^2 \mu$ is a finite measure. If μ has finite second moment, then $\mathcal{F}\mu$ is twice continuously differentiable. Assume μ has finite second moment. We call the vector

$$m = (m_j) = (D_j \mathcal{F}\mu(0))$$

the mean of μ. The mean of $T_{-m}\mu$ is zero.

Assume the mean of μ is zero. Then $\sum_j D_j^2(\mathcal{F}\mu)(0)$ is a nonnegative real number, because it is the limit for $\lambda \to \infty$ of

$$(\mu, \sum_j X_j^2 D_\lambda^\infty g) > 0 \quad .$$

We call

$$\sigma = \sqrt{\sum_j D_j^2(\mathcal{F}\mu)(0)}$$

the standard deviation of μ.

If the standard deviation of μ is 0, then μ is supported in $\{0\}$. Namely, assume f with $\|f\| < 1$ is (compactly) supported outside 0 and $\mu(f) \neq 0$. we can assume f is real (by taking real and imaginary part separately). Then

$$\sum_j X_j^2 \mu(D_\lambda^\infty \pm f) = \pm \sum_j X_j^2 \mu(f)$$

is positive for sufficiently large λ by positivity of for both choices of $\pm$, which is clearly absurd. By direct inspection we see that the only probability measure supported on 0 is δ_0, thus $\mu = \delta_0$.

If μ has mean 0 and finite standard deviation σ, then

$$D^1_{\sigma^{-1}} T_{-m} \mu$$

is a probability measure with standard deviation 1 and mean 0.

Let μ be a probability measure, then it describes the probability "distribution" of a (vector valued) random variable Y. If we have another random variable Z with distribution ν, then the probability distribution of $Y + Z$ is given by $\mu * \nu$.

Suppose we have a sequence Y_k of identically distributed random variables in $\mathbf{R}$ with probability distribution μ, mean 0 and standard deviation 1, then the distribution of $k^{-\frac{1}{2}} \sum_{i=1}^{n} (Y_k)$ is given by $\mu_k = -D^1_{k^{-1/2}} \mu^{*k}$. The central limit theorem states that μ_k converges in measure to a Gaussian.

Theorem 4 *Central Limit Theorem Let μ be a probability measure on $\mathcal{S}(\mathbf{R})$ such that $\sum_j X_j^2 \mu$ is a measure,*

$$\mathcal{F}(X\mu)(0) = 0 \quad ,$$
$$\mathcal{F}(X^2\mu)(0) = 1 \quad .$$

Then the sequence of measures

$$\mu_k = -D^1_{k^{-1/2}} \mu^{*k}$$

converges in the weak star topology on M_1 to the measure $\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

Proof Consider the twice continuously differentiable function $f = \mathcal{F}\mu$. We have $f(0) = 1$, $Df(0) = 0$, $D^2 f(0) = 1$. Thus we have

$$f(x) = 1 - 4\pi^2 x^2 + h(x) x^2$$

for some continuous function h with $h(0) = 0$. We have

$$f_k(x) = f(k^{-1/2} x)^k = (1 - k^{-1} 2\pi^2 x^2 + h(k^{-1/2} x) k^{-1} x^2)^k \quad .$$

Given any small ϵ and fixed x, we conclude for sufficiently large k we conclude

$$(1 - k^{-1}(2\pi^2 + \epsilon) x^2)^k < f_k(x) < (1 - k^{-1}(2\pi^2 - \epsilon) x^2)^k \quad .$$

This implies for large k

$$e^{-(2\pi^2 - \epsilon) x^2} - \epsilon < f_k(x) < e^{-(2\pi^2 + \epsilon) x^2} + \epsilon \quad .$$

Thus $\mathcal{F}\mu$ tends to $e^{-2\pi^2 x^2}$ pointwise. This is the Fourier transform of

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{x^2}{2}} \quad .$$

By the Lebesgue dominated convergence theorem, (μ_k, f) tends to (g, f) for all $f \in \mathcal{S}(\mathbf{R})$. Since $\mathcal{S}(\mathbf{R})$ is dense in $C_0(\mathbf{R})$ and μ_k is a bounded family, (μ_k, f) tends to (g, f) for all $f \in C_0(\mathbf{R})$ because

$$|(\mu_k - g, f)| \leq \|\mu_k - g\| \|f - \tilde{f}\|_\infty + \|\mu_k - g, \tilde{f})| \quad ,$$

where $\tilde{f} \in \mathcal{S}(\mathbf{R})$ is an approximation to $f \in C_0(\mathbf{R})$, and both terms can be made smaller than any $\epsilon > 0$ by choice of $\tilde{f}$ and k. ■

0.13 Malgrange Ehrenpreis

We have seen that for each integer $k > 0$ there is a tempered distribution $\Lambda_k \in \mathcal{S}(\mathbf{R})$ such that

$$(\Lambda_k, f) = \int t^{-k} f(t)\, dt$$

whenever f vanishes of order k at 0. E.g., the derivatives of $p.v.(1/t)$ do this. It follows that

$$(X^k \Lambda_k, f) = (\Lambda, X^k f) = \int t^{k-k} f(t)\, dt = (1, f) \quad ,$$

where 1 denotes the distribution which is constant equal to 1.

A corollary of this is that for every polynomial P in one variable of degree $N \geq 0$ there is a tempered distribution Λ_P such that $P(X)\Lambda = 1$. Namely, we can write for x not a zero of P by partial fractions

$$\frac{1}{P(x)} = \sum_{k=1}^{N} \frac{b_k}{(x - a_k)^{n_k}} \quad ,$$

where n_k is less than the vanishing order of P at a_k. Then for every f which vanishes at least of the order of P everywhere we have

$$\Lambda_P(f) := \left(\sum_{k: a_k = \Re(a_k)} b_k T_{a_k} \Lambda_{-n_k} + \sum_{k: a_k \neq \Re(a_k)} \frac{b_k}{(. - a_k)^{n_k}} \right) f = \int \frac{1}{P(t)} f(t)\, dt$$

and therefore $P(X)\Lambda_P := 1$.

For polynomials in several variables this is more subtle. We have

Theorem 5 *Malgrange-Ehrenpreis For each polynomial P in n variables there is an element $\Lambda_P \in \mathcal{F}(\mathcal{D}'(\mathbf{R}^n)$ such that*

$$P(X)\Lambda_P(f) = (1, f)$$

for all $f \in \mathcal{F}(\mathcal{D}(\mathbf{R}^n))$.

Proof

A linear functional Λ on the space $\mathcal{F}(\mathcal{D}(\mathbf{R}^n))$ is continuous, if there are functions C_γ and M_γ such that for all $f \in \mathcal{F}(\mathcal{D}(\mathbf{R}^n))$ there is a $\gamma_f > 0$ such that for all $\gamma > \gamma_f$

$$|(\Lambda, f)| \leq C_\gamma \sup_{x,y} |f(x+iy)(1+|x|^2)^{M_\gamma} e^{-\gamma|y|}|.$$

In particular Λ satisfies this estimate if for some C, M

$$|(\Lambda, f)| \leq C \sup_{x,|y|<1} |f(x+iy)(1+|x|^2)^M| \quad .$$

(See the section on Paley Wiener.) We shall show that the functional

$$(\Lambda_P, f) = \int_{\mathbf{R}^n} \frac{1}{P(x)} f(x)\, dx \quad ,$$

defined on the space $P(X)\mathcal{F}(\mathcal{D}(\mathbf{R}^n))$ satisfies this estimate. By Hahn Banach the functional extends to a linear functional on all of $\mathcal{F}\mathcal{D}(\mathbf{R}n)$ satisfying this estimate. The extension is therefore is an element of $\mathcal{F}\mathcal{D}'(\mathbf{R}^n)$, which had to be shown.

Assume P has degree N. Let P_N be the homogeneous part of P of degree N. Pick $x \in \mathbf{R}^n$ so that $P_N(x) \neq 0$. Let V be the orthogonal complement of x in $\mathbf{R}^n$.

Let f be an element of $P(X)\mathcal{F}\mathcal{D}(\mathbf{R}^n)$. Then

$$\frac{1}{P(x)} f(x)$$

is an analytic function satisfying

$$\left| \frac{1}{P(x+iy)} f(x+iy) \right| \leq C(1+|x|^2)^{-2n}$$

for $x, y \in \mathbf{R}^n$ and $|y| \leq N+2$. Thus we can change contours to write for any choice of integers $0 < k(y) < N+1$

$$\int_{\mathbf{R}^n} \frac{1}{P(x)} f(x) = \int_V \int_{\mathbf{R}} \frac{1}{P(y+(t+ik(y))x)} f(y+(t+ik(y))x)\, dt\, dy$$

With appropriate choice of k this can be estimated by

$$C \int_V \int_R \frac{1}{P(y+(t+ik(y))x)} (1+|t|^2)^{-1}(1+|y|^2)^{-n}\, dt\, dy$$

$$\leq C \int_V \inf_{k \in 0,\ldots,N+1} \sup_{t \in \mathbf{R}} \frac{1}{P(y+(t+ik)x)} (1+|y|^2)^{-n}\, dy$$

Thus it suffices to estimate

$$\inf_{k \in 0,\ldots,N+1} \quad \sup_{t \in \mathbf{R}} \frac{1}{P(y+(t+ik)x)}$$

uniformly in y

However, we can write

$$P(y + (t + ik)x) = c \prod_{l=1}^{N} ((t + ik) - a_l)$$

with zeroes a_k depending on y but c independent of y (c depends only on the homogeneaous part of $P(y + tx)$ of degree N in t). By the pigeon hole there is at least one $k \in \{0, \dots, N+1\}$ such that the strip $|Im(z) - k| < 0$ does not contain any zero a_l. Thus, for this k,

$$|P(y + ty_0 + ky_0)| \geq c2^{-N} \quad .$$

This proves the theorem. ■

0.14 The wave equation

As an example of fundamental solutions we consider the partial differential operator

$$\Box f = -D_{n+1}^2 f + \sum_{j=1}^{n} D_j^2 f$$

For each $w \in \mathcal{S}(\mathbf{R}^{n+1})$ we would like to find a solution $f \in \mathcal{S}'(\mathbf{R}^{n+1})$ to the inhomogeneous wave equation

$$\Box f = w \quad .$$

(Eventually we hope to have better regularity of the solution that $\mathcal{S}'(\mathbf{R}^{n+1})$. Taking partial Fourier transform in the first n variables leads to

$$-D_{n+1}^2 \hat{f} + \sum_{j=1}^{n} X_j^2 \hat{f} = \hat{w}$$

where

$$\hat{f} = \mathcal{F}_1 \dots \mathcal{F}_n f$$

and similarly for w. We write t for the $n+1$-st variable and $\xi = (\xi_1, \dots, \xi_n)$ for first n variables. For fixed ξ we have the general solution to the homogeneous ODE ($\hat{w} = 0$):

$$\hat{f}(\xi, t) = Ae^{2\pi i|x|t} + Be^{-2\pi i|x|t} \quad .$$

To solve the inhomogeneous equation ($\hat{w} \neq 0$) we read this as an ansatz with varying coefficients $A = A(t)$, $B + B(t)$. We have

$$D_{n+1}\hat{f}(\xi, t) = \frac{1}{2\pi i}\left(A'(t)e^{2\pi i|\xi|t} + B'(t)e^{-2\pi i|\xi|t}\right) + \left(A|\xi|e^{2\pi i|\xi|t} - B|\xi|e^{-2\pi i|\xi|t}\right)$$

Setting the first summand equal to 0 (as part of the ansatz) we have

$$-D_{n+1}^2\hat{f}(\xi, t) + |\xi|^2 f(\xi, t) = \frac{-1}{2\pi i}\left(A'(t)|\xi|e^{2\pi i|\xi|t} - B'(t)|\xi|e^{-2\pi i|\xi|t}\right) = \hat{w}(\xi, t)$$

Solving for A' gives

$$A'(t) = -\pi i|\xi^{-1}|\hat{w}(\xi,t)e^{-2\pi i|\xi|t}$$

and thus

$$A(t) = -\int_{-\infty}^{t} \pi i|\xi|^{-1}\hat{w}(\xi,s)e^{-2\pi i|\xi|s}\,ds$$

is a particular solution. Similarly

$$B(t) = \int_{-\infty}^{t} \pi i|\xi|^{-1}\hat{w}(\xi,s)e^{2\pi i|\xi|s}\,ds$$

Thus we have as one particular solution to the inhomogeneous equation

$$\hat{f}(\xi,t) = \int_{-\infty}^{t} 2\pi|\xi|^{-1}\hat{w}(\xi,s)\sin(2\pi|\xi|(t-s))\,ds$$

$$= 2\pi\int |\xi^{-1}|\hat{w}(\xi,s)1_{+}(t-s)\sin(2\pi|\xi|(t-s))\,ds \quad .$$

Consider the following distribution in $\mathcal{S}(\mathbf{R}^n)$:

$$\Lambda(g) = \int_{S_1^{n-1}} f(x)\,d\mu \quad ,$$

where $S_{r,y}^{n-1}$ is the sphere of radius r about y and $d\mu$ is the standard surface measure on the sphere. Since $\mathcal{S}(\mathbf{R}^n)$ is weak star dense in $M_1(\mathbf{R}^n)$, by Lebesgue's dominated convergence theorem the Fourier transform of $\Lambda(g)$ is pointwise given by

$$\mathcal{F}\Lambda(\xi) = \int_{S_{0,1}^{n-1}} e^{-2\pi i<x,\xi>}\,d\mu(x) \quad .$$

By rotation symmetry this integral depends only on $|\xi|$, we can therefore have ξ point to the north pole. Let α be the angle between ξ and x, then

$$\mathcal{F}\Lambda(\xi) = c_n\int_0^{\pi} e^{-2\pi i|\xi|\cos(\alpha)}\sin(\alpha)^{n-2}\,d\alpha \quad ,$$

where c_n is a certain constant. For odd $n \geq 3$ this can be calculated in closed form by the following substitution:

$$\mathcal{F}\Lambda(\xi) = c_n\int_{-1}^{1} e^{-2\pi i|\xi|x}(1-x^2)^{\frac{n-3}{2}}\,dx \quad ,$$

In particular we have for $n = 3$:

$$\mathcal{F}\Lambda(\xi) = c_n\int_{-1}^{1} e^{-2\pi i|\xi|x}\,dx$$

$$= \frac{-c_n}{2\pi i}|\xi|^{-1}(e^{-2\pi i|\xi|} - e^{2\pi i|\xi|}) = \frac{c_n}{\pi}|\xi|^{-1}\sin(2\pi|\xi|) \quad .$$

Since we know that $\mathcal{F}\Lambda(0) = 4\pi$ we have

$$\mathcal{F}\Lambda(\xi) = 2|\xi|^{-1}2\pi|\xi| \quad .$$

Inserting this into our formula for $\hat{f}$ gives

$$\hat{f}(\xi, t) = \pi \int_{\mathbf{R}} \hat{w}(\xi, s) 1_+(t-s) D^{\infty}_{t-s^{-1}} \mathcal{F}\Lambda(\xi)\, ds$$

$$= \pi \int_{\mathbf{R}} (w * D^{\infty}_{t-s}\Lambda)\hat{\ }(\xi, s) 1_+(t-s)\, ds$$

$$= \pi \int_{\mathbf{R}} \left[\int_{\mathbf{R}^n} |t-s|^{-2} \left[\int_{S^1_{t-s,x}} w(\eta)\, d\mu \right] e^{-2\pi i <\xi, x>}\, dx \right] 1_+(t-s)\Lambda(\xi)\, ds \quad .$$

Here $*$ denotes convolution in the first n variables only. The order of integration can be interchanged by Fubini (actually not just Fubini, needs some extra justification), thus we obtain, this time $*$ denoting convolution in all variables:

$$f(x, t) = (w * \Lambda_{\text{cone}}) \quad ,$$

where

$$\Lambda_{\text{cone}}(g) = \frac{\pi}{\sqrt{2}} \int_{\{(x,t) \in \mathbf{R}^4 : t < 0, |x| = t\}} |t|^{-2} g\, d\mu$$

and μ is surface measure. From this formula we can conclude the strong Huygens' principle, wich states that light at time t and location x influences at time $t' > t$ only those locations which have distance $t' - t$ from x. The strong Huygens' principle holds only for n odd. For even n, one gets instead of Λ_c a distribution which is supported on the solid light cone, but does not vanish in the interior of the solid light cone.

As we have seen, Λ_{cone} is a fundamental solution of the wave equation, i.e., $\mathcal{F}\Lambda_{\text{cone}}$ is equal to of $(-t^2 + x^2)^{-1}$ outside the set where the latter is singular.

Convolution with Λ_{cone} actually maps $\mathcal{S}(\mathbf{R}^{n+1})$ into $C^\infty(\mathbf{R}^n) \cap L^\infty$. However, more can be said about the regularity of this operaton. E.g., as we quote here without proof, it maps $L^{4/3}$ to L^4 boundedly (for $n+1 = 4$). This may be used to solve the nonlinear wave equation

$$\Box f = f^3 + w \quad .$$

Namely, we obtain

$$f = \Lambda_{\text{cone}} * (f^3 + w) \quad .$$

We can use this to define an iteration scheme

$$f_{k+1} = \Lambda_{\text{cone}} * (f_k^3 + w) := A(f_k)$$

with $f_k = 0$. If $w \in L^{\frac{4}{3}}$, then $f_1 \in L^4$ by the mentioned regularity result. Thus f^3 is again in $L^{\frac{4}{7}3}$ and we can iterate. In fact the iteration map A is a contraction in a small ball around 0, because

$$\|Af - Ag\|_4 \leq C\|f^3 - g^3\|_{4/3}$$

$$\leq C\|f - g\|_4 \|f^2 + fg + g^2\|_2 \leq C\|f - g\|_4 (\|f\|_4^2 + \|g\|_4^2) \leq 1/2\|f - g\|_4 \quad ,$$

provided f and g have sufficiently small norm. Thus the iteration converges it w has sufficiently small norm (to ensure that the small ball is mapped to itself).

1 Hermite functions

Hermite functions arise naturally in a number of ways. We shall introduce them from the point of view of looking for eigenfunctions of the Fourier transform of the form $P(x)e^{-\pi|x|^2}$ with polynomials P. In fact we will obtain an entire orthonormal basis of $L^2(\mathbf{R}^n)$ consisting of such eigenfunctions.

We shall first discuss the case $n = 1$. Consider the operators $X + iD$ and $X - iD$. They satisfy the commutation identities

$$\mathcal{F}(X + iD) = i(X + iD)\mathcal{F}$$

$$\mathcal{F}(X - iD) = -i(X - iD)\mathcal{F}$$

Thus, if f is an eigenfunction of $\mathcal{F}$ with eigenvalue λ, then $(X+iD)f$ and $(X-iD)f$, unless they are 0, are eigenfunction of $\mathcal{F}$ with eigenvalues $i\lambda$ and $-i\lambda$.

Define $f_0(x) = e^{-\pi x^2}$. We easily see that

$$(X + iD)f_0 = 0 \quad .$$

It follows from linear ODE that the every function satisfying this identity is a multiple of f_0.

Define the n-th hermite function as

$$f_n = (X - iD)^n f_0 \quad .$$

Then f_n is of the form $P_n(x)e^{-\pi x^2}$ with some polynomial P_n which is easily seen to have real coefficients and has degree at most n. We shall see momentarily that the degree is actually equal to n.

We observe the identities

$$(X - iD)(X + iD) = X^2 + D^2 - i(DX - XD) = X^2 + D^2 - \frac{1}{2\pi}I \quad ,$$

$$(X + iD)(X - iD) = X^2 + D^2 + i(DX - XD) = X^2 + D^2 + \frac{1}{2\pi}I \quad ,$$

where I is the dientity operator. Thus, for $n > 0$,

$$(X + iD)f_n = (X + iD)(X - iD)f_{n-1} = (X - iD)(X + iD)f_{n-1} + \frac{1}{\pi}f_{n-1} \quad .$$

This shows eaily by induction that

$$(X + iD)f_n = \frac{n}{\pi}f_{n-1} \quad .$$

In particular, f_n is an eigenfunction of $(X - iD)(X + id)$ with eigenvalue $\frac{n}{\pi}$. This implies that all functions f_n are linearly independent and therefore $P_n(x)$ has necessarily exact degree n.

We observe that f_n is also an eigenfunction of the Schroedinger operator of the harmonic oscillator, X^2+D^2, with eigenvalue $\frac{2n+1}{2\pi}$. Since X^2+D^2 is self adjoint, the functions f_n are pairwise orthogonal.

We aim to show that the collection of Hermite functions is a complete orthogonal set. This means that if $f \in L^2(\mathbf{R})$ and $\langle f, f_n\rangle = 0$ for all n, then $f = 0$. Thus assume $f \in L^2(\mathbf{R})$ and $\langle f, f_n\rangle = 0$ for all n. Then ff_0 is an integrable functyion and we have

$$\mathcal{F}(ff_0)(\xi) = \int f(x)e^{-\pi x^2} e^{2\pi i x\xi}\, dx$$

$$= \int \sum_n \left[f(x)e^{-\pi x^2} \frac{(2\pi i x\xi)^n}{n!} \right] dx$$

The term in brackets is clearly bounded by

$$\frac{C}{n^2} e^{-\pi x^2} e^{|x\xi|}(1+|x\xi|)^2 \quad .$$

Thus we can interchange integration and summation and obtain

$$\mathcal{F}(ff_0)(\xi) = \sum_n \int f(x) \left[f(x)e^{-\pi x^2} \frac{(2\pi i x\xi)^n}{n!} \right] dx = 0$$

becaus integrating f against any function of the form $P(x)e^{-\pi x^2}$ gives zero by assumption.

Thus f_n is a complete orthogonal set. We can make it an orthonormal basis by normalizing. Define

$$g_0 = 2^{\frac{1}{4}} f_0$$

because

$$\|f_0\|_2^2 = \int e^{-2\pi x^2}\, dx = 2^{-\frac{1}{2}} \quad ,$$

and

$$g_n = \left(\frac{\pi}{n}\right)^{\frac{1}{2}} (X - iD) g_{n-1}$$

because

$$\|(X-iD)f_{n-1}\|_2^2 = \langle (X+iD)(X-iD)f_{n-1}, f_{n-1}\rangle = \frac{n}{\pi} \langle f_{n-1}, f_{n-1}\rangle \quad .$$

Then g_n is an orthonormal basis of $L^2(\mathbf{R})$.

We can easily obtain an orthonormal basis of $L^2(\mathbf{R}^n)$ by defining

$$g_{j_1,\ldots,j_n}(x) = \prod_{k=1}^{n} g_{j_k}(x_k) \quad .$$

Orthonormality is easily seen directly, and completeness can be proved in the same way as for $n = 1$. (It also follows from very general principles).

For each $N \geq 0$ define the norm

$$\|f\|_{[N]} = \left(\sum_{j_1,\ldots,j_n} |\langle f, g_{j_1,\ldots,j_N} \rangle|^2 \, (1 + j_1 + \ldots + j_n)^N \right)^{\frac{1}{2}} \quad .$$

This is defined (when allowed to be infinite) for all tempered distributions f. The space of all distributions for which this norm is finite we denote by $L^2_{[N]}(\mathbf{R}^n)$. It is easily seen that this space is a Hilbert space with scalar product

$$\langle f, g \rangle = \sum_{j_1,\ldots,j_n} \langle f, g \rangle \, (1 + j_1 + \ldots + j_n)^N \quad .$$

It has an orthonormal basis consisting of suitably normalized multiples of the functions $g_{j_1,\ldots,j_n}$. The norms $\|.\|_{[N]}$ are equivalent to the Schwartz norms in the following sense:

Lemma 15 *For each $N \geq 0$ there is a $N' \geq 0$ such that for all $f \in \mathcal{S}'(\mathbf{R}^n)$ we have*

$$\|f\|_{[N]} \leq C \|f\|_{(N)} \quad ,$$

$$\|f\|_{(N)} \leq C \|f\|_{[N]} \quad .$$

Proof We shall prove this lemma in the case $n = 1$, the general case being very similar. We have

$$\|f\|^2_{[N]} = \sum_j | \langle f, g_j \rangle |^2 (1 + j)^N$$

$$\leq C \sum_j | \left\langle f, (X^2 + D^2)^{N'} g_j \right\rangle |^2 = C \sum_j | \left\langle (X^2 + D^2)^{N'} f, g \right\rangle |^2$$

$$\leq C \|(X^2 + D^2)^{N'} f\|_2^2 \leq C \|(1 + |x|^2)(X^2 + D^2)^{N'} f\|^2_\infty$$

$$\leq C \|f\|_{(2N'+2)}$$

Vice versa, we have

$$\|D^\alpha X^\beta f\|_\infty \leq C \|(1 + |D|^2) X^\alpha D^\beta f\|_2$$

$$\leq \|(1 + (X + iD)(X - iD))((X + iD) + (X - iD))^\alpha (X + iD) - (X - iD))^\beta f\|_2$$

The latter is easily estimated by a constant time a sufficiently large $[N]$-norm.

This completes the proof.

■

2 The Schwartz kernel theorem

Lemma 16 *Let A be a bilinear form*

$$\mathcal{S}(\mathbf{R}^n) \times \mathcal{S}(\mathbf{R}^m) \to \mathbf{C}$$

such that

$$|A(f,g)| \leq C\|f\|_{[N]}\|g\|_{[N]}$$

for all $f \in \mathcal{S}(\mathbf{R}^n), g \in \mathcal{S}(\mathbf{R}^m)$. Then there exists a unique distribution $\Lambda \in \mathcal{S}(\mathbf{R}^{n+m})$ such that $A(f,g) = \Lambda(f \otimes g)$.

Proof Uniqueness: Λ is determined on all elements $g_{j_1,\ldots,j_{n+m}}$, because these are elementary tensor products of Schwartz functions in $\mathcal{S}(\mathbf{R}^n)$ and in $\mathcal{S}(\mathbf{R}^m)$. Thus it is determined on any finite linear combination of the elements $g_{j_1,\ldots,j_{n+m}}$.

Now let f be arbitrary in $\mathcal{S}(\mathbf{R}^{n+m})$. We can find finite linear combinations

$$f_k = \sum a^{(k)}_{j_1,\ldots,j_n} g_{j_1,\ldots,j_n}$$

such that $\|f - f_k\|_{[N]} \leq 1/k$, because the functions $g_{j_1,\ldots,j_n}$ form a complete orthogonal set in $L^2_{[N]}$.

Then the sequence f_k converges in $L^2_{[K]}$ for all K, hence in $\mathcal{S}(\mathbf{R}^{n+m})$. Since Λ is determined on all f_k, it is determined on f.

Existence: Define Λ in the natural way on the elements $g_{j_1,\ldots,j_{n+m}}$. We will show that there are constants C and K such that for all finite linear combinations $\sum a_{j_1,\ldots,j_{n+m}} g_{j_1,\ldots,j_{n+m}}$ we have

$$|\Lambda(\sum a_{j_1,\ldots,j_n} g_{j_1,\ldots,j_n})| \leq C\|\sum a_{j_1,\ldots,j_n} g_{j_1,\ldots,j_n})\|_{[K]} \quad .$$

Thus Λ extends to a bounded linear functional in $L^2_{[K]}$ and thus to a bounded linear functional on $\mathcal{S}(\mathbf{R}^{n+m})$.

But we have

$$|\Lambda(\sum a_{j_1,\ldots,j_n} g_{j_1,\ldots,j_n})| \leq \sum |a_{j_1,\ldots,j_n}||A(g_{j_1,\ldots,j_n} \otimes g_{j_{n+1}}, \ldots, g_{j_{n+m}})|$$

$$\leq C \sum |a_{j_1,\ldots,j_n}|(1 + j_1 + \ldots + j_n)^N (1 + j_{n+1} + \ldots + j_{n+m})^M$$

$$\leq C \left(\sum |a_{j_1,\ldots,j_n}|^2 (1 + j_1 + \ldots + j_n)^{2(N+n)} (1 + j_{n+1} + \ldots + j_{n+m})^{2(M+n)}\right)^{\frac{1}{2}}$$

$$\leq C \left(\sum |a_{j_1,\ldots,j_n}|^2 (1 + j_1 + \ldots + j_{n+m})^{2(N+M+2n)}\right)^{\frac{1}{2}} \quad .$$

This proves the desired estimate.

■

Theorem 6 *The Schwartz kernel theorem*

Let $T : \mathcal{S}(\mathbf{R}^n) \to \mathcal{S}'(\mathbf{R}^m)$ *be a continuous linear operator. Then there exists a unique* $\Lambda \in \mathcal{S}'(\mathbf{R}^{n+m})$ *such that*

$$(Tf, g) = \Lambda(f \otimes g)$$

for all $f \in \mathcal{D}(\mathbf{R}^m)$ *and* $g \in \mathcal{S}(\mathbf{R}^m)$.

Proof The bilinear map $f, g \to (T(f), g)$ is continuous in each argument. Since $\mathcal{S}(\mathbf{R}^n)$ and $\mathcal{S}(\mathbf{R}^m)$ are complete metric spaces, an application of Banach-Steinhaus shows that the bilinear map $(f, g) \to (T(f), g)$ is continuous in the product topology. Thus for each ϵ there is a δ such that

$$|(Tf, g)| \leq \epsilon$$

follows from

$$\rho(f, 0) + \rho(g, 0) \leq \delta \quad .$$

The latter follows from

$$\|f\|_{[N]} \leq \delta', \|g\|_{[N]} \leq \delta'$$

for certain N and δ'. By linearity in each argument we conclude

$$|(Tf, g)| \leq C\|f\|_{[N]}\|g\|_{[N]} \quad .$$

The Schwartz kernel theorem now follows from the previous lemma. ■

As an application of the Schwartz kernel theorem we prove the following little statement.

Lemma 17 *Let* $A : \mathcal{S}(\mathbf{R}) \to \mathcal{S}'(\mathbf{R})$ *be a continuosu linear operator such that*

$$A(T_x f) = T_x(Af)$$

and

$$A(D_\lambda^\infty f) = D_\lambda^\infty(Af)$$

for all $x \in \mathbf{R}$, $\lambda > 0$, *and* $f \in \mathcal{S}(\mathbf{R})$. *Then* A *extends to a bounded operator from* $L^2(\mathbf{R})$ *to itself and is a linear combination of the identity and the Hilbert transform (convolution with p.v. 1/t).*

Proof We pass to the Fourier transform: Define

$$M = \mathcal{F} A \mathcal{F}^{-1}$$

Then M is again a bounded operator $\mathcal{S}(\mathbf{R}) \to \mathcal{S}'(\mathbf{R})$. By the Schwartz kernel theorem there is a unique distribution $\Lambda \in \mathcal{S}(\mathbf{R}^2)$ such that $(Mf, g) = \Lambda(f \otimes g)$. We have

$$\Lambda(f \otimes g) = (Mf, g) = (M_x(Mf), M_{-x}g) = (M(M_x f), M_{-x}g) = \Lambda(M_x f, M_{-x}g)$$

By uniqueness of Λ we see

$$\Lambda(F) = \Lambda(M_{(x,-x)}F) \quad .$$

This implies that Λ vanishes on all functions F which vanish on the diagonal of $\mathbf{R}^2$. (If F is compactly supported, we can divide by $1 - e^{2\pi i(\epsilon(x_1-x_2))}$.)

Thus there is a linear functional Λ' on $\mathcal{S}(\mathbf{R})$ which satisfies

$$\Lambda(f) = \Lambda'(x \to f(x,x)) \quad ,$$

namely

$$\Lambda'(f) = \Lambda((x_1,x_2) \to f(x_1)e^{-\pi(x_1-x_2)^2}) \quad ,$$

and Λ' is easily seen to be a tempered distribution. Dilation invariance of A implies that Λ' is homogeneous of degree 0:

$$\Lambda(f \otimes g) = (Mf, g) = (D_\lambda^2 Mf, D_\lambda^2 g) =$$

$$(MD_\lambda^2 f, D_\lambda^2 g) = \Lambda(D_\lambda^2 f \otimes D_\lambda^2 g) = \Lambda(D^2(f \times g)) \quad ,$$

and

$$\begin{aligned}\Lambda'(D_\lambda^1 f) &= \Lambda((x_1,x_2) \to \lambda^{-1} f(\lambda^{-1}x_1)e^{-\pi(x_1-x_2)^2}) \\ &= \Lambda((x_1,x_2) \to \lambda^{-1} f(\lambda^{-1}x_1)e^{-\pi\lambda^{-2}(x_1-x_2)^2}) \\ &= \Lambda((x_1,x_2) \to f(x_1)e^{-\pi(x_1-x_2)^2} = \Lambda'(f) \quad .\end{aligned}$$

Thus Λ is the element of a two dimensional space, which is easily identified as the space of linear combinations of characteristic functions of $\mathbf{R}_+$ and $\mathbf{R}_-$. Thus we have

$$(Mf,g) = a\int_{-\infty^0} f(x)g(x)\,dx + b\int_0^\infty f(x)g(x)\,dx \quad ,$$

$$M(f,g) \le C\|f\|_2\|g\|_2 \quad .$$

The operator A itself is an element of a two dimensional space, which clearly contains the multiples of the identity and the Hilbert transform, as an easy verification shows.
■

3 The Heisenberg uncertainty relation

The Heisenberg uncertainty relation is a principle saying that a function and its Fourier transform cannot be simultaneoulsy well localized. It comes in many manifestations. E.g., if a function is compactly supported, then its Fourier transform cannot be compctly supported, because it is analytic, unless it is zero everywhere.

Another manifestation is given by the following inequality:

$$\|Xf\|_2\|Df\|_2 \ge \frac{1}{4\pi}\|f\|_2^2 \quad ,$$

which follows from the following easy calculation:

$$2\Im\langle Xf, Df\rangle = \Im(\langle Xf, Df\rangle - \langle DF, Xf\rangle)$$

$$= \Im\langle (DX - XD)f, f\rangle = \frac{-1}{2\pi}\langle f, f, \rangle \quad .$$

That this estimate is sharp follows from the example $f(x) = e^{-\pi x^2}$, because Xf and DX are multiplies of each other and their product is purely imaginary.

If we assume $\|f\|_2 = \|\mathcal{F}f\|_2 = 1$, then we can interpret both $|f|^2$ and $|\mathcal{F}f|^2$ as probability distribution. By applying translation and modulation operators we can assume that the mean of both distributions is equal to 1. Then the Heisenberg uncertainty relation gives a lower bound on the product of the standard deviations of $|f|^2$ and $|\mathcal{F}f|^2$.

4 Reproducing formulas, Gabor frames

The phase plane transform of a distribution Λ with respect to a fixed Schwartz function ϕ (normalized so that $\|\phi\|_2 = 1$ is obtained by testing Λ against all translates and modulates of ϕ. More precisely, assume $\|\phi\|_2 = 1$, then we define $V = V_\phi$ by

$$V\Lambda(x, \xi) = \Lambda(M_\xi T_x \phi) \quad .$$

We observe that this definition depends on the choice of order of translation and modulation. However, we shall mostly be concerned with the product

$$V\Lambda_1(x, \xi)\overline{V\Lambda_2(x, \xi)} \quad ,$$

which does not depend on that choice because

$$M_\xi T_x = e^{2\pi i <x,\xi>} T_x M_\xi \quad .$$

Lemma 18 *The transform V is an isometry from $L^2(\mathbf{R}^n)$ to $L^2(\mathbf{R}^{2n})$, i.e., we ahve for all $f_1, f_2 \in L^2(\mathbf{R}^N)$ that Vf_1 and Vf_2 are in $L^2(\mathbf{R}^{2n})$ and*

$$\langle Vf_1, Vf_2\rangle = \langle f_1, f_2\rangle \quad .$$

Proof

For Schwartz functions f we can write

$$Vf(x, \xi) = \int f(y) e^{2\pi i <\xi, y>} \phi(y - x)\, dy \quad ,$$

which is easily seen as the composition of the unitary map $F(x, y) \to F(y, y - x)$ and the Fourier transform in y direction, applied to the tensor product $f \otimes \phi$. Thus V is an isometry if restricted to $\mathcal{S}(\mathbf{R}^n)$.

It remains to show that V is continuous form $\mathcal{S}'(\mathbf{R}^n)$ tro $\mathcal{S}'(\mathbf{R}^{2n})$, which is an exercise.

■

As a result of this lemma we have for all $f \in \mathcal{S}(\mathbf{R}^n)$ the reproducing formula

$$f = \int_{\mathbf{R}^n \times \mathbf{R}^n} \langle f, M_\xi T_x \phi \rangle \, M_\xi T_x \phi \, dx \, d\xi \quad .$$

which holds in the sense of a Bochner integral (in particular pointwise of $f \in \mathcal{S}(\mathbf{R}^n)$).

In the following we restrict ourselves to the case $n = 1$, similar results for general n follow immediately.

We shall study whether there are Schwartz functions ϕ such that the distrcetized reproducing formula

$$\langle f, g \rangle = c \sum_{x \in a\mathbf{Z}, \xi \in \alpha \mathbf{Z}} \langle f, M_\xi T_x \phi \rangle \, \langle M_\xi T_x \phi, g \rangle \tag{14}$$

holds.

Suppose the discretized reproducing formula holds for certain c, a, α, ϕ, then

$$c^{-1} \langle f, g \rangle = \int_{[0,a] \times [0,\alpha]} \sum_{x \in a\mathbf{Z}, \xi \in \alpha \mathbf{Z}} \langle M_{\xi'} T_{x'} f, M_\xi T_x \phi \rangle \, \langle M_\xi T_x \phi, M_{\xi'} T_{x' g} \rangle \; dx' \, d\xi'$$

$$= \int_{[0,a] \times [0,\alpha]} \langle M_{\xi'} T_{x'} f, M_{\xi'} T_{x' g} \rangle = a\alpha \, \langle f, g \rangle \quad .$$

Thus we necessarily have $a\alpha = c^{-1}$, which we shall henceforth always assume. In particular if $c = 1$, we observe that the discrete reproducing formula with $c = 1$ implies that the set of all $M_\xi T_x \phi$ is an orthonormal basis in $L^2(\mathbf{R})$ for all $x \in a\mathbf{Z}$ and $\xi \in \alpha \mathbf{Z}$. For $c \neq 1$, the set of these functions is called an exact frame. (Gabor frame in the current setting of wave packets.)

We have the following theorem:

Theorem 7 (Balian Low) *For $a\alpha < 1$ there exist Schwartz functions ϕ such that the discrete reproducing formula holds. For $a\alpha = 1$ no such Schwartz function exists, but there exist $L^2(\mathbf{R})$ functions ϕ such that the disrete reproducing formula holds. Any such ϕ satisfies*

$$\max(\|D^2 \phi\|_2, \|X^2 \phi\|_2) = \infty \quad .$$

For $a\alpha > 1$ no $L^2(\mathbf{R})$ function ϕ exists such that the discrete reproducing formula holds.

Remark: The theorem does hold in the case $a\alpha = 1$ if the second order operators X^2 and D^2 are replaced by X and D, but for the sake of a nice proof we will prove this weaker statement.

As we have seen, the case $a\alpha = 1$ is particularly interesting, because (14) then means that the collection $M_\xi T_x \phi$ with $x \in a\mathbf{Z}$ and $\xi \in \alpha \mathbf{Z}$ is an orthonormal basis of $L^2(\mathbf{R}^n)$ (provided $\phi \in L^2(\mathbf{R}^n)$.). This is called a windowed Fourier transform with window ϕ. An example for such a basis is given by $a = 1$, $\alpha = 1$, ϕ the characteristic function of $[0, 1]$.

Proof We first study the case $a\alpha < 1$. The first step is to find a C_0^∞ function ϕ which is supported in $[0, \alpha^{-1}]$ and satisfies

$$\sum_{n \in aZ} |T_n \phi|^2 = 1 \quad .$$

We know that $C_0^\infty(\mathbf{R})$ is not empty. Pick any such function, by multipying it with its complex conjugate we can assume it is nonnegative real. By appropriate scaling, dilating, and translating, we can assume that it has a primitive which is constant $-1/2$ on $[-\infty, -1]$ and constant $1/2$ on $1, \infty$. By antisymmetrizing we can assume it is odd. By adding $1/2$ to the function we obtain a non-negative function, of which we can take the square root, which we will denote by ψ. Then we have

$$\psi^2(x) + \psi^2(-x) = 1$$

From here some algebra yields the desired function ϕ.

We claim that the function ϕ satisfies the discrete reproducing formula with a and α.

First we have

$$\sum_{\xi \in \alpha\mathbf{Z}} \langle f, M_\xi \phi \rangle \langle M_\xi \phi, f \rangle$$

$$= \sum_{\xi \in \alpha\mathbf{Z}} \left| \int_{[0,\alpha]} f(x) \overline{\phi}(x) e^{2\pi i \xi \alpha} \, dx \right|^2$$

$$= \left\langle f\overline{\phi}, f\overline{\phi} \right\rangle = \left\langle f|\phi|^2, f \right\rangle$$

by the Plancherel formula for the discrete Fourier transform. The same formula holds with $T_x\phi$ in place of ϕ.

Thus we have

$$\sum_{x \in a\mathbf{Z}, \xi \in \alpha\mathbf{Z}} \langle f, M_\xi T_x \phi \rangle \langle M_\xi T_x \phi, f \rangle$$

$$= \sum_{x \in a\mathbf{Z}} \left\langle f|T_x\phi|^2, f \right\rangle = \langle f, f \rangle$$

By polarization we obtain the cooresponding formula for $\langle f, g \rangle$ which had to be proved.

Now we consider the case $a\alpha = 1$. By dilation we can assume $a = \alpha = 1$.

Assume there exists a $\phi \in L^2(\mathbf{R})$ such that the discrete reproducing formula holds. This implies that the set of all $M_n T_m \phi$ with $n, m \in \mathbf{Z}$ is an orthonormal basis of $L^2(\mathbf{R})$.

We take what is called the Zak transform of ϕ. It is defined for compactly supported functions $f \in L^2(\mathbf{R})$ by

$$Zf(x, \xi) = \sum_{l \in Z} T_l M_\xi f(x) \quad .$$

For any two integers n, m it satisfies

$$Zf(x-n,\xi-m) = Zf(x,\xi)\sum_{l\in\mathbf{Z}} T_n T_l M_{-m} M_\xi f(x) = e^{-2\pi m x} e^{Zf(x,\xi)}$$

and is thus determined by its values on the fundamental domain $S = [0,1) \times [0,1)$. Moreover the Zak transform satisfies

$$Z(M_n T_m f)(x,\xi) = \sum_l T_l M_\xi M_n T_m f(x)$$

$$= e^{2\pi n x} \sum_l T_l M_\xi T_m f(x) = e^{2\pi n x} e^{2\pi m \xi} Zf(x) \quad .$$

This leads to the crucial observation that for a compactly supported $f \in L^2(\mathbf{R})$ the Zak transforms of the functions $M_n T_M f$ form an orthonormal basis of $L^2(\mathbf{R}^n)$ if and only if the Zak transform of f has modulus 1 almost everywhere. Namely, by Plancherel we have

$$\sum_{n,m} |\langle G, Z(M_n T_m f)\rangle|^2 = \|GZf\|_2^2 \quad .$$

In particular this holds for f the characteristic function of $[0,1)$, because

$$Zf(x,\xi) = \sum_{l\in\mathbf{Z}} T_l M_\xi f(x)$$

$$M_\xi \sum_{l\in\mathbf{Z}} e^{2\pi i l \xi} T_l f(x) = M_\xi e^{2\pi i [x] \xi} \quad .$$

Since $M_n T_M f$ actually forms an orthonormal basis of $L^2([0,1])$ we have that the Zak transform extends to an isomorphism of $L^2(\mathbf{R})$ to $L^2(S)$. All the previous observations continue to hold for all $f \in L^2(\mathbf{R})$ under this extension of the Zak transform.

Now we are looking for a function $f \in L^2(\mathbf{R})$ that satisfies $\|X^2 f\|_2 \|D^2 f\|_2 < \infty$ and $M_n T_m f$ forms an orthonormal basis, which is the same as to say that $|Zf|$ is one almost everywhere. We will assume such an f exists. Then the regularity of f will imply that Zf is continuous, but we will show that for Zf to be continuous it has to vanish somewhere (from its invariance properties), which gives a contradiction to Zf being 1 almost everywhere.

We fill in the details. First we observe that (for nice f),

$$Z(Xf)(x,\xi) = \sum_l T_l M_\xi X f(x)$$

$$= \sum_l T_l D_\xi M_\xi f(x) = D_\xi Zf(x,\xi)$$

and

$$Z(Df)(x,\xi) = \sum_l T_l M_\xi D f(x)$$

$$= \sum_l T_l D M_\xi f(x) - \sum_l T_l X_\xi M_\xi f(x) = D_x Zf(x,\xi) - X_\xi Zf(x,\xi) \quad .$$

Thus, if Xf and Df are in L^2, so are $D_\xi Zf$ and $D_x Zf$ (since $X_\xi Zf$ clearly in L^2), and similarly if $X^2 f$ and $D^2 f$ are in $L^2(\mathbf{R})$, then $D_x^2 Zf$ and $D_\xi^2 Zf$ are in L^2. Hence Zf is continuous.

However, consider the loop γ in $\mathbf{R}^2$ passing along straight lines from $(0,0)$ over $(0,1)$, $(1,1)$, $(1,0)$ back to $(0,0)$. Then $Z \circ \gamma$ has nonzero winding number around 0. Since the curve γ is contractible, the image of Zf must contain 0.

This completes the proof of the Balian Low theorem in the case $a\alpha = 1$.

The proof in the case $a\alpha > 1$ is postponed.

■