B4.4 Fourier Analysis

Jan Kristensen

Hilary Term 2021

Contents

1 The Fourier transform

1.1 The Fourier transform on L^1

Definition 1.1. Let $f \in \mathrm{L}^1(\mathbb{R}^n)$. Then the *Fourier transform* of f is

$$\mathcal{F}(f)(\xi) = \widehat{f}(\xi) := \int_{\mathbb{R}^n} f(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\,\mathrm{d}x, \quad \xi \in \mathbb{R}^n,$$

where $x \cdot \xi = x_1\xi_1 + \ldots + x_n\xi_n$ is the usual dot product in $\mathbb{R}^n$.

Remark 1.2. Note that $\widehat{f}(\xi)$ is well-defined for each $\xi \in \mathbb{R}^n$ since

$$|f(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}| = |f(x)|$$

and $f \in \mathrm{L}^1(\mathbb{R}^n)$. Observe that

- $|\widehat{f}(\xi)| \leqslant \|f\|_1$ for all $\xi \in \mathbb{R}^n$ and $\widehat{f}(0) = \int_{\mathbb{R}^n} f(x)\,\mathrm{d}x$, so if $f \geqslant 0$, then $\widehat{f}(0) = \|f\|_1$,
- $\widehat{f} \in \mathrm{C}(\mathbb{R}^n)$ (this is easy to see),
- $\widehat{f}(\xi) \to 0$ as $|\xi| \to \infty$ (by the Riemann–Lebesgue lemma that we state and prove below).

The precise range $\mathcal{F}\big(\mathrm{L}^1(\mathbb{R}^n)\big) = \{\widehat{f} : f \in \mathrm{L}^1(\mathbb{R}^n)\}$ is not so easy to describe in terms not involving the Fourier transform. It can be shown that it is strictly smaller than

$$\mathrm{C}_0(\mathbb{R}^n) := \left\{ g \in \mathrm{C}(\mathbb{R}^n) : \lim_{|x|\to\infty} g(x) = 0 \right\}.$$

The range $\mathcal{F}\big(\mathrm{L}^1(\mathbb{R}^n)\big)$ is called the *Wiener algebra*.

One reason that we are interested in the Fourier transform here is its ability to transform partial derivatives to an algebraic operation.

Proposition 1.3. ***[Differentiation Rule]*** *Let $f \in \mathrm{L}^1(\mathbb{R}^n)$ and assume that for some $j \in \{1, \ldots, n\}$ the distributional partial derivative $\partial_j f \in \mathrm{L}^1(\mathbb{R}^n)$. Then*

$$\widehat{\partial_j f}(\xi) = \mathrm{i}\xi_j \widehat{f}(\xi).$$

Proof. Let $\phi \in \mathscr{D}(\mathbb{R}^n)$ be such that $\phi(x) = 1$ for $|x| \leqslant 1$. We then calculate

$$\begin{aligned}
\widehat{\partial_j f}(\xi) &= \int_{\mathbb{R}^n} \partial_j f(x) \mathrm{e}^{-\mathrm{i}x\cdot\xi} \,\mathrm{d}x \\
&\overset{\text{DCT}}{=} \lim_{r\to\infty} \int_{\mathbb{R}^n} \partial_j f(x) \mathrm{e}^{-\mathrm{i}x\cdot\xi} \phi\left(\frac{x}{r}\right) \,\mathrm{d}x \\
&= \lim_{r\to\infty} \left\langle \partial_j f,\, \mathrm{e}^{-\mathrm{i}(\cdot)\cdot\xi} \phi\left(\frac{\cdot}{r}\right) \right\rangle \\
&= \lim_{r\to\infty} \left\langle f,\, \mathrm{i}\xi_j \mathrm{e}^{-\mathrm{i}(\cdot)\cdot\xi} \phi\left(\frac{\cdot}{r}\right) - \mathrm{e}^{-\mathrm{i}(\cdot)\cdot\xi} (\partial_j \phi)\left(\frac{\cdot}{r}\right) \frac{1}{r} \right\rangle \\
&= \lim_{r\to\infty} \int_{\mathbb{R}^n} f(x) \left(\mathrm{i}\xi_j \mathrm{e}^{-\mathrm{i}x\cdot\xi} \phi\left(\frac{x}{r}\right) - \mathrm{e}^{-\mathrm{i}x\cdot\xi} (\partial_j \phi)\left(\frac{x}{r}\right) \frac{1}{r} \right) \mathrm{d}x \\
&\overset{\text{DCT}}{=} \int_{\mathbb{R}^n} f(x) \mathrm{i}\xi_j \mathrm{e}^{-\mathrm{i}x\cdot\xi} \,\mathrm{d}x = \mathrm{i}\xi_j \hat{f}(\xi).
\end{aligned}$$

□

Example 1.4. Let $f = \mathbf{1}_{(-1,1)}$. Clearly $f \in \mathrm{L}^1(\mathbb{R})$, and

$$\widehat{f}(\xi) = 2\mathrm{sinc}(\xi) := \begin{cases} \dfrac{2\sin\xi}{\xi} & \text{for } \xi \neq 0 \\ 2 & \text{for } \xi = 0. \end{cases}$$

Recall that *sinus cardinalis* function sinc is not absolutely integrable over $\mathbb{R}$, so $\hat{f} \notin \mathrm{L}^1(\mathbb{R})$.

We can generalize this to indicator functions of rectangles in $\mathbb{R}^n$ by a straightforward application of Fubini's theorem: if

$$R = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$$

then clearly $\mathbf{1}_R \in \mathrm{L}^1(\mathbb{R}^n)$ and

$$\widehat{\mathbf{1}_R}(\xi) = \prod_{j=1}^{n} \frac{\mathrm{e}^{-\mathrm{i}a_j\xi_j} - \mathrm{e}^{-\mathrm{i}b_j\xi_j}}{\mathrm{i}\xi_j}$$

for $\xi_j \neq 0$ for all j (but $\widehat{\mathbf{1}_R}$ is continuous so we can extend to the coordinate planes by continuity). By inspection, $\widehat{\mathbf{1}_R}(\xi) \to 0$ as $|\xi| \to \infty$. Thus $\widehat{\mathbf{1}_R} \in \mathrm{C}_0(\mathbb{R}^n)$, but as in the 1-dimensional case $\widehat{\mathbf{1}_R} \notin \mathrm{L}^1(\mathbb{R}^n)$.

Lemma 1.5. ***[The Riemann-Lebesgue Lemma]*** *Let $f \in \mathrm{L}^1(\mathbb{R}^n)$. Then the Fourier transform $\widehat{f} \in \mathrm{C}_0(\mathbb{R}^n)$.*

Proof. We have already remarked that the dominated convergence theorem easily gives that $\widehat{f}$ is continuous. (One can also prove it more directly from the definition of the Lebesgue integral: Exercise.) In order to prove that $\widehat{f}(\xi) \to 0$ as $|\xi| \to \infty$ we use that step functions are dense in

$\mathrm{L}^1(\mathbb{R}^n)$. Let $\varepsilon > 0$ and take a step function $s\colon \mathbb{R}^n \to \mathbb{R}$ so $\|f - s\|_1 < \varepsilon/2$. By Example 1.4 we infer that $\widehat{s} \in \mathrm{C}_0(\mathbb{R}^n)$ and so for some $r > 0$ we have $|\widehat{s}(\xi)| < \varepsilon/2$ for $|\xi| > r$. We then have for all $\xi \in \mathbb{R}^n$ with $|\xi| > r$:

$$\begin{aligned}
|\widehat{f}(\xi) &\leqslant |\widehat{f}(\xi) - \widehat{s}(\xi)| + |\widehat{s}(\xi)| \\
&< |\widehat{(f - s)}(\xi)| + \varepsilon/2 \\
&\leqslant \|f - s\|_1 + \varepsilon/2 \\
&< \varepsilon.
\end{aligned}$$

The proof is finished. □

You will be asked for another proof of the Riemann-Lebesgue lemma on Problem Sheet 1. While the result is quite elementary and easy to prove, it is useful: on several occasions we will prove that a function is continuous by showing that it is the Fourier transform of an integrable function.

Example 1.6. Let $\rho \in \mathscr{D}(\mathbb{R})$ be the standard mollifier kernel on $\mathbb{R}$ (so in particular, ρ is an even function satisfying $0 \leqslant \rho \leqslant 1$, $\rho(x) > 0$ for $|x| < 1$, $\mathrm{supp}(\rho) = [-1, 1]$, and $\int \rho = 1$). Then

$$\widehat{\rho}(\xi) = 2\int_0^1 \rho(x)\cos(x\xi)\,\mathrm{d}x.$$

It is not hard to check that $\widehat{\rho} \in \mathrm{C}^\infty(\mathbb{R})$, but $\mathrm{supp}(\hat{\rho})$ is not compact. Therefore $\widehat{\rho} \notin \mathscr{D}(\mathbb{R})$. We shall return to this point when discussing the uncertainty principle later in the course (see also Problem Sheet 1). However,

$$\widehat{\rho}(\xi) \to 0 \text{ as } |\xi| \to \infty,$$

and in fact, for any $k, m \in \mathbb{N}_0$ we have

$$|\xi|^k \frac{\mathrm{d}^m}{\mathrm{d}\xi^m}\widehat{\rho}(\xi) \to 0 \text{ as } |\xi| \to \infty.$$

Again this is not difficult to show and will be a consequence of a more general result proved later, so we leave it for now.

We would like to extend the Fourier transform to distributions, and to that end we seek an adjoint identity.

Proposition 1.7. ***[The Product Rule]*** *Let f, $g \in \mathrm{L}^1(\mathbb{R}^n)$. Then*

$$\int_{\mathbb{R}^n} f(x)\widehat{g}(x)\,\mathrm{d}x = \int_{\mathbb{R}^n} \widehat{f}(x)g(x)\,\mathrm{d}x.$$

Remark 1.8. Note that both sides are well-defined since the Fourier transform of an L^1 function is bounded and continuous. We thus have an identity that looks like an adjoint identity with $S = \mathcal{F} = T$, but there is an issue with the domain: if we start with f, $g \in \mathscr{D}(\mathbb{R}^n)$, then we will not have $\widehat{f}$, $\widehat{g} \in \mathscr{D}(\mathbb{R}^n)$, that is, unless $f \equiv g \equiv 0$. Indeed, in the example with the standard mollifier kernel we had $\widehat{\rho} \notin \mathscr{D}(\mathbb{R})$. In order to address this issue efficiently we must define a new class of test functions and distributions.

Proof. This is an easy application of Fubini:

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)\widehat{g}(x)\,\mathrm{d}x &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(y)\mathrm{e}^{-\mathrm{i}x\cdot y}\,\mathrm{d}y\,\mathrm{d}x \\ &\overset{\text{Fubini}}{=} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(y)\mathrm{e}^{-\mathrm{i}x\cdot y}\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{\mathbb{R}^n} \widehat{f}(y)g(y)\,\mathrm{d}y\end{aligned}$$

□

Before addressing the issues with the domain and the appropriate class of test functions, let us investigate the properties of the Fourier transform on L^1 functions a little more. It will turn out to be a useful source of insight.

To us the *orthogonal group* $\mathrm{O}(n)$ consists of $n \times n$ matrices with real entries such that its columns form an orthonormal basis for $\mathbb{R}^n$: thus $\theta \in \mathrm{O}(n)$ if and only if $\theta \in \mathrm{M}_{n\times n}(\mathbb{R})$ and $\theta^\dagger\theta = \mathrm{I}$. A matrix θ is a *special orthogonal matrix* of dimension n, $\theta \in \mathrm{SO}(n)$, if and only if $\theta \in \mathrm{O}(n)$ and $\det(\theta) = 1$. In B4.3 we defined the operation θ_* on functions and distributions. When $f \in \mathrm{L}^1(\mathbb{R}^n)$ and $\theta \in \mathrm{O}(n)$ we put $\theta_* f(x) := f(\theta x)$. We then have the following invariance property of the Fourier transform:

Proposition 1.9. ***[Invariance under orthogonal maps]*** *Let $f \in \mathrm{L}^1(\mathbb{R}^n)$ and $\theta \in \mathrm{O}(n)$. Then $\widehat{\theta_* f} = \theta_* \hat{f}$.*

Proof. This is a simple calculation where we substitute $y = \theta x$, use $\theta^{-1} = \theta^\dagger$ and exploit that $\det\theta = \pm 1$:

$$\begin{aligned}\widehat{\theta_* f}(\xi) &= \int_{\mathbb{R}^n} f(\theta x)\mathrm{e}^{-\mathrm{i}\xi\cdot x}\,\mathrm{d}x \\ &= \int_{\mathbb{R}^n} f(y)\mathrm{e}^{-\mathrm{i}\xi\cdot\theta^{-1}y}\,\mathrm{d}y \\ &= \int_{\mathbb{R}^n} f(y)\mathrm{e}^{-\mathrm{i}\theta\xi\cdot y}\,\mathrm{d}y \\ &= \theta_*\hat{f}(\xi).\end{aligned}$$

□

Remark 1.10. We emphasize the special case of *reflection through the origin*: $\tilde{f}(x) := f(-x)$ corresponding to $\theta = -\mathrm{I} \in \mathrm{O}(n)$. In this case we have

$$\mathcal{F}(\tilde{f}) = \widetilde{\mathcal{F}(f)}$$

for $f \in \mathrm{L}^1(\mathbb{R}^n)$.

Proposition 1.11. ***[Translation Rules]*** *Let $f \in \mathrm{L}^1(\mathbb{R}^n)$ and denote $\big(\tau_h f\big)(x) := f(x+h)$ for $h \in \mathbb{R}^n$. Then*

$$\mathcal{F}(\tau_h f)(\xi) = \mathrm{e}^{\mathrm{i}\xi\cdot h}\widehat{f}(\xi)$$

and

$$\mathcal{F}(\mathrm{e}^{-\mathrm{i}x\cdot h}f(x))(\xi) = \tau_h\widehat{f}(\xi)$$

for any $h \in \mathbb{R}^n$.

Proof. We simply calculate

$$\mathcal{F}(\tau_h f)(x) = \int_{\mathbb{R}^n} f(x+h)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\,\mathrm{d}x \overset{y=x+h}{=} \int_{\mathbb{R}^n} f(y)\mathrm{e}^{-i(y-h)\cdot\xi}\,\mathrm{d}y = \mathrm{e}^{\mathrm{i}h\cdot\xi}\widehat{f}(\xi),$$

and

$$\int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}x\cdot h}f(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\,\mathrm{d}x = \int_{\mathbb{R}^n} f(x)\mathrm{e}^{-\mathrm{i}x\cdot(\xi+h)}\,\mathrm{d}x = \widehat{f}(\xi+h) = \tau_h\widehat{f}(\xi).$$

□

Proposition 1.12. ***[Dilation Rules]*** *Let $f \in \mathrm{L}^1(\mathbb{R}^n)$ and denote*

$$(d_r f)(x) = f(rx)$$

for $r > 0$. Then

$$\mathcal{F}(d_r f)(\xi) = r^{-n}\widehat{f}(r^{-1}\xi) = r^{-n}(d_{\frac{1}{r}}\widehat{f})(\xi)$$

and

$$(d_r\widehat{f})(\xi) = \mathcal{F}(r^{-n}d_{\frac{1}{r}}f)(\xi).$$

Remark 1.13. Recall the notation ρ_ε for the standard mollifier on $\mathbb{R}^n$. We shall adapt it in general and write f_r for $r^{-n}d_{\frac{1}{r}}f$:

$$f_r := \tfrac{1}{r^n}d_{\frac{1}{r}}f.$$

We often refer to f_r as the L^1 *dilation* of f because $\|f_r\|_1 = \|f\|_1$ holds for all $r > 0$ when $f \in \mathrm{L}^1(\mathbb{R}^n)$.

Proof. The proof is a simple calculation as in the previous lemma (and we employ the notation from the previous remark):

$$\begin{aligned}
\mathcal{F}(d_r f)(\xi) &= \int_{\mathbb{R}^n} f(rx)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\,\mathrm{d}x \\
&\underset{\mathrm{d}y=r^n\mathrm{d}x}{\overset{y=rx}{=}} \int_{\mathbb{R}^n} f(y)\mathrm{e}^{-\mathrm{i}\frac{y}{r}\cdot\xi}r^{-n}\,\mathrm{d}y \\
&= r^{-n}\widehat{f}\left(\frac{\xi}{r}\right) \\
&= (\widehat{f})_r(\xi),
\end{aligned}$$

and

$$
\begin{aligned}
(d_r\widehat{f})(\xi) &= \widehat{f}(r\xi) \\
&= \int_{\mathbb{R}^n} f(x)\mathrm{e}^{-\mathrm{i}x\cdot r\xi}\,\mathrm{d}x \\
&\overset{y=rx}{\underset{\mathrm{d}y=r^n\mathrm{d}x}{=}} \int_{\mathbb{R}^n} f\left(\frac{y}{r}\right)\mathrm{e}^{-\mathrm{i}\frac{y}{r}\cdot r\xi}r^{-n}\,\mathrm{d}y \\
&= r^{-n}\int_{\mathbb{R}^n}(d_{\frac{1}{r}}f)(y)\mathrm{e}^{-\mathrm{i}y\cdot\xi}\,\mathrm{d}y \\
&= \widehat{(f_r)}(\xi).
\end{aligned}
$$

□

Proposition 1.14. ***[Convolution Rule]*** *Let $f,\ g \in \mathrm{L}^1(\mathbb{R}^n)$. Then $f * g \in \mathrm{L}^1(\mathbb{R}^n)$ and*

$$\mathcal{F}(f*g)(\xi) = \widehat{f}(\xi)\widehat{g}(\xi).$$

Proof. By Fubini,

$$
\begin{aligned}
\mathcal{F}(f*g)(\xi) &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x-y)g(y)\,\mathrm{d}y\,\mathrm{e}^{-\mathrm{i}x\cdot\xi}\,\mathrm{d}x \\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x-y)\mathrm{e}^{-i(x-y)\cdot\xi}\,\mathrm{d}x\,g(y)\mathrm{e}^{-iy\cdot\xi}\,\mathrm{d}y \\
&\overset{z=x-y}{\underset{\mathrm{d}z=\mathrm{d}x}{=}} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(z)\mathrm{e}^{-iz\cdot\xi}\,\mathrm{d}z\,g(y)\mathrm{e}^{-iy\cdot\xi}\,\mathrm{d}y \\
&= \widehat{f}(\xi)\widehat{g}(\xi).
\end{aligned}
$$

□

Proposition 1.15. ***[Reverse Differentiation Rule]*** *Let $f \in \mathrm{L}^1(\mathbb{R}^n)$ and assume $x_j f(x) \in \mathrm{L}^1(\mathbb{R}^n)$ for some $j \in \{1,\dots,n\}$. Then the distributional partial derivative $\partial_j\hat{f}$ is a continuous function and*

$$(\partial_j\widehat{f})(\xi) = \mathcal{F}(-\mathrm{i}x_j f(x))(\xi).$$

In fact, $\partial_j\hat{f}$ exists classically.

Proof. Let us start with the last statement. Fix $\xi \in \mathbb{R}^n$, $h \in \mathbb{R}\setminus\{0\}$ and consider the following difference quotient:

$$
\begin{aligned}
\triangle_{he_j}\hat{f}(\xi)/h &:= \frac{\widehat{f}(\xi+he_j)-\widehat{f}(\xi)}{h} \\
&= \int_{\mathbb{R}^n} f(x)\triangle_{he_j}\mathrm{e}^{-\mathrm{i}x\cdot(\cdot)}(\xi)/h\,\mathrm{d}x \\
&\underset{h\to 0}{\overset{\mathrm{DCT}}{\longrightarrow}} \int_{\mathbb{R}^n} -\mathrm{i}x_j f(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\,\mathrm{d}x \\
&= \mathcal{F}_{x\to\xi}(-\mathrm{i}x_j f(x)),
\end{aligned}
$$

so the partial derivative $\partial_j \hat{f}$ exists classically at ξ. By the Riemann-Lebesgue lemma, $\mathcal{F}_{x\to\xi}(-ix_j f(x))$ is continuous as the Fourier transform of an L^1 function, so $\partial_j \hat{f}$ is continuous. This is also the distributional partial derivative since, as we have seen in B4.3, $\triangle_{he_j}\widehat{f}/h \to \partial_j \widehat{f}$ in $\mathscr{D}'(\mathbb{R}^n)$ as $h \to 0$. Recall that this amounts to

$$\langle \triangle_{he_j}\widehat{f},\, \varphi\rangle \underset{h\to 0}{\longrightarrow} \langle \partial_j \widehat{f},\, \varphi\rangle$$

for every $\varphi \in \mathscr{D}(\mathbb{R}^n)$. It is not difficult to see that $\triangle_{he_j}\widehat{f}(\xi)/h \to \partial_j \widehat{f}(\xi)$ as $h \to 0$ locally uniformly in $\xi \in \mathbb{R}^n$ so the classical and distributional partial derivatives therefore coincide. □

We generalize the differentiation rules on L^1 to include linear partial differential operators as follows. Recall that these are conveniently written in terms of multi-index notation as discussed in B4.3: for a multi-index $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ and $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, $\partial = (\partial_1, \dots, \partial_n)$, we wrote

$$x^\alpha := x_1^{\alpha_1} \dots x_n^{\alpha_n}$$

and

$$\partial^\alpha := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}.$$

When $p(x)$ is a polynomial of degree at most k in n indeterminates, then

$$p(x) = \sum_{|\alpha|\leqslant k} c_\alpha x^\alpha,$$

where $c_\alpha \in \mathbb{C}$ and we sum over all multi-indices $\alpha \in \mathbb{N}_0^n$ of length $|\alpha| \leqslant k$. Corresponding to the polynomial $p(x)$ there is a *linear partial differential operator* defined by replacing x by ∂ throughout:

$$p(\partial) := \sum_{|\alpha|\leqslant k} c_\alpha \partial^\alpha.$$

If $c_\alpha \neq 0$ for some $\alpha \in \mathbb{N}_0^n$ with $|\alpha| = k$, then we say $p(\partial)$ has order k (so the order of $p(\partial)$ is simply the degree of $p(x)$). Sometimes we also write $p(\mathrm{i}\partial)$ or $p(-\mathrm{i}\partial)$, the notation being self explanatory:

$$p(\mathrm{i}\partial) = \sum_{|\alpha|\leqslant k} c_\alpha (\mathrm{i}\partial)^\alpha = \sum_{|\alpha|\leqslant k} c_\alpha i^{|\alpha|}\partial^\alpha,$$

and so on.

Corollary 1.16. ***[Generalized Differentiation Rules]*** *Let $p(x) \in \mathbb{C}[x]$ be a polynomial in n variables.*

(G1) If $f \in \mathrm{L}^1(\mathbb{R}^n)$ and $p(\partial)f \in \mathrm{L}^1(\mathbb{R}^n)$, then

$$\widehat{p(\partial)f}(\xi) = p(i\xi)\hat{f}(\xi).$$

(G2) If $f \in \mathrm{L}^1(\mathbb{R}^n)$ and $p(-\mathrm{i}x)f(x) \in \mathrm{L}^1(\mathbb{R}^n)$, then

$$\mathcal{F}_{x\to\xi}\big(p(-\mathrm{i}x)f(x)\big) = (p(\partial)\hat{f})(\xi).$$

The expressions $p(\partial)f$ and $p(\partial)\hat{f}$ are understood distributionally.

Sketch of proof. (G1): We can apply the differentiation rule in a straight forward manner to prove this *provided* we know that $\partial^\alpha f \in \mathrm{L}^1(\mathbb{R}^n)$ for all multi-indices $\alpha \in \mathbb{N}_0^n$ that appear in $p(\partial)$. However, knowing that f, $p(\partial)f \in \mathrm{L}^1(\mathbb{R}^n)$ does not necessarily mean that the individual partial derivatives $\partial^\alpha f$ that make up $p(\partial)f$ are in $\mathrm{L}^1(\mathbb{R}^n)$. The way around this problem is to mollify f: put $f_\varepsilon := \rho_\varepsilon * f$ for $\varepsilon > 0$, where $\big(\rho_\varepsilon\big)_{\varepsilon>0}$ is the standard mollifier on $\mathbb{R}^n$. Now $f_\varepsilon \in \mathrm{C}^\infty(\mathbb{R}^n)$ and for any multi-index $\alpha \in \mathbb{N}_0^n$ we have

$$\partial^\alpha f_\varepsilon = \big(\partial^\alpha \rho_\varepsilon\big) * f \in \mathrm{L}^1(\mathbb{R}^n)$$

as the convolution of two $\mathrm{L}^1(\mathbb{R}^n)$ functions. (Note that the L^1 norms might not stay bounded when $\varepsilon \searrow 0$, but that is not important because we shall apply the differentiation rules to f_ε for each fixed $\varepsilon > 0$.) The identity in (G1) therefore holds with f_ε in place of f and we conclude by taking $\varepsilon \searrow 0$, using $p(\partial)f_\varepsilon = \rho_\varepsilon * \big(p(\partial)f\big) \to p(\partial)f$ in $\mathrm{L}^1(\mathbb{R}^n)$ and that the Fourier transform $\mathcal{F}\colon \mathrm{L}^1(\mathbb{R}^n) \to \mathrm{L}^\infty(\mathbb{R}^n)$ is continuous. The details are left as an exercise.

(G2): We can apply the reverse differentiation rule in a straight forward manner to prove this provided we know that $x^\alpha f \in \mathrm{L}^1(\mathbb{R}^n)$ for all multi-indices α that appear in $p(x)$. We only know that f, $p(-\mathrm{i}x)f(x) \in \mathrm{L}^1(\mathbb{R}^n)$ and this will not be enough to guarantee that. In this case we *localize* f by multiplying it with the indicator function of the open ball centered at 0 and radius $j \in \mathbb{N}$:

$$f_j := f\mathbf{1}_{B_j(0)}.$$

Clearly $x^\alpha f_j(x) \in \mathrm{L}^1(\mathbb{R}^n)$ for all multi-indices $\alpha \in \mathbb{N}_0^n$ and so, by linearity and the reverse differentiation rule,

$$\mathcal{F}_{x\to\xi}\big(p(-\mathrm{i}x)f_j(x)\big) = p(\partial)\hat{f}_j(\xi) \tag{1}$$

holds for all $j \in \mathbb{N}$. Now $p(-\mathrm{i}x)f_j(x) \to p(-\mathrm{i}x)f(x)$ pointwise a.e. as $j \to \infty$ and $\big|p(-\mathrm{i}x)f_j(x)\big| \leqslant \big|p(-\mathrm{i}x)f(x)\big|$ a.e., so by Lebesgue's dominated convergence theorem we infer that $p(-\mathrm{i}x)f_j(x) \to p(-\mathrm{i}x)f(x)$ in $\mathrm{L}^1(\mathbb{R}^n)$ as $j \to \infty$. Clearly also $f_j \to f$ in $\mathrm{L}^1(\mathbb{R}^n)$ as $j \to \infty$, so $\hat{f}_j \to \hat{f}$ uniformly and hence in particular in the sense of distributions on $\mathbb{R}^n$. By $\mathscr{D}'$ continuity of differentiation, the right-hand side of (1) converges in $\mathscr{D}'(\mathbb{R}^n)$ to $p(\partial)\hat{f}(\xi)$. It is not difficult to conclude from here and we leave the details as an exercise. □

We shall give a more general version of the differentiation rules once we have developed the theory a bit further. The proof will then also be much more streamlined!

1.2 The Schwartz space

We are now ready to address the domain issue in the adjoint identity for the Fourier transform, the product rule (see Proposition 1.7), and line up with some definitions and results to this effect.

Definition 1.17. A function $f\colon \mathbb{R}^n \to \mathbb{C}$ is said to be *rapidly decreasing* if for every $m \in \mathbb{N}$ there exist r_m, $c_m > 0$ such that

$$|f(x)| \leqslant c_m |x|^{-m}$$

holds for all $|x| \geqslant r_m$.

Remark 1.18. It is not difficult to see that a continuous function f is rapidly decreasing if and only if for any polynomial $p(x) \in \mathbb{C}[x]$ the function $x \mapsto p(x)f(x)$ is bounded on $\mathbb{R}^n$:

$$\sup_{x\in\mathbb{R}^n} |p(x)f(x)| < \infty.$$

In fact, it suffices to take any monomial x^α here for the sufficiency. We emphasize that a rapidly decreasing function $f\colon \mathbb{R} \to \mathbb{R}$ *need not* be decreasing in the usual sense: rapidly decreasing means that $x^m f(x) \to 0$ as $|x| \to \infty$ for all $m \in \mathbb{N}$. This is a bit of an understatement as a rapidly decreasing function $f\colon \mathbb{R} \to \mathbb{R}$ is decreasing if and only if $f \equiv 0$ on $\mathbb{R}$. The point here is that a rapidly decreasing function need not be monotone from a certain point either, as the rapid approach to 0 can happen in an oscillatory manner as it does for instance with $\mathrm{e}^{-|x|}\cos x$.

Example 1.19. As functions defined for $x \in \mathbb{R}^n$ we have that $\frac{1}{1+|x|^m}$ is *not* rapidly decreasing for any $m \in \mathbb{N}$, whereas both $\mathrm{e}^{-|x|}$ and $\mathrm{e}^{-|x|^2}$ are rapidly decreasing.

The following definition is modelled on the properties of the Fourier transform of the standard mollifier kernel and is due to Laurent Schwartz (1940s):

Definition 1.20. [Schwartz test functions and the Schwartz Space]
A function $\varphi\colon \mathbb{R}^n \to \mathbb{C}$ is a *Schwartz test function on* $\mathbb{R}^n$ if

(i) $\varphi \in \mathrm{C}^\infty(\mathbb{R}^n)$, and

(ii) $\partial^\alpha \varphi$ is rapidly decreasing for all multi-indices $\alpha \in \mathbb{N}_0^n$.

The set of all Schwartz test functions on $\mathbb{R}^n$ is denoted by $\mathscr{S}(\mathbb{R}^n)$ and called the *Schwartz space.*

Example 1.21. As functions defined for $x \in \mathbb{R}^n$ we have

- $\mathrm{e}^{-|x|^2} \in \mathscr{S}(\mathbb{R}^n) \setminus \mathscr{D}(\mathbb{R}^n)$,
- $\mathrm{e}^{-|x|} \notin \mathscr{S}(\mathbb{R}^n)$ because it is not differentiable at zero,

- $\hat{\rho}(\xi) \in \mathscr{S}(\mathbb{R}^n) \setminus \mathscr{D}(\mathbb{R}^n)$, where ρ is the standard mollifier kernel on $\mathbb{R}^n$. Strictly speaking we have actually not really proved this yet, but the proof will follow soon.

The next result collects elementary properties of the class of Schwartz test functions.

Proposition 1.22.

(i) $\mathscr{S}(\mathbb{R}^n)$ *is a vector space and a commutative ring without unit (with the usual definitions of operations).*

(ii) If $p(x) \in \mathbb{C}[x]$ *and* $\varphi \in \mathscr{S}(\mathbb{R}^n)$*, then* $p\varphi \in \mathscr{S}(\mathbb{R}^n)$*.*

(iii) If $p(x) \in \mathbb{C}[x]$ *and* $\varphi \in \mathscr{S}(\mathbb{R}^n)$*, then* $p(\partial)\varphi \in \mathscr{S}(\mathbb{R}^n)$*.*

Proof. (i): It is clear that $\mathscr{S}(\mathbb{R}^n)$ is a vector subspace of $\mathrm{C}^\infty(\mathbb{R}^n)$. In order to check that it is a subring of $\mathrm{C}^\infty(\mathbb{R}^n)$ it suffices to show that $\phi\psi \in \mathscr{S}(\mathbb{R}^n)$ whenever ϕ, $\psi \in \mathscr{S}(\mathbb{R}^n)$. First we note that the product of two rapidly decreasing functions is rapidly decreasing and that linear combinations of rapidly decreasing functions are rapidly decreasing. Then we use the Leibniz rule to write for $\alpha \in \mathbb{N}_0^n$:

$$\partial^\alpha \bigl(\phi\psi\bigr) = \sum_{\beta \leqslant \alpha} \binom{\alpha}{\beta} \partial^\beta \phi \partial^{\alpha-\beta}\psi.$$

Since derivatives of Schwartz test functions are rapidly decreasing we see that $\partial^\alpha\bigl(\phi\psi\bigr)$ is rapidly decreasing, and hence, since α was arbitrary, that $\phi\psi \in \mathscr{S}(\mathbb{R}^n)$.

(ii): Clearly $p\varphi \in \mathrm{C}^\infty(\mathbb{R}^n)$. For each $1 \leqslant j \leqslant n$ and $\beta \in \mathbb{N}_0^n$ we have

$$\partial^\beta\bigl(x_j\varphi\bigr) = \beta_j \partial^{\beta-\beta_j e_j}\varphi + x_j \partial^\beta \varphi. \tag{2}$$

The derivative $\partial^\beta\bigl(x_j\varphi\bigr)$ is therefore rapidly decreasing and since this is true for all β we have shown that $x_j\varphi \in \mathscr{S}(\mathbb{R}^n)$. But then we get by iteration of this that $x^\alpha\varphi \in \mathscr{S}(\mathbb{R}^n)$ for all $\alpha \in \mathbb{N}_0^n$, and so by the vector space property, that $p\varphi \in \mathscr{S}(\mathbb{R}^n)$.

(iii): This is clear. □

For calculations with Schwartz test functions the following class of norms is very useful. Besides this they are used to define the notion of convergence in $\mathscr{S}(\mathbb{R}^n)$ below.

Definition 1.23. Let $\varphi \in \mathrm{C}^\infty(\mathbb{R}^n)$. Then we define for multi-indices α, $\beta \in \mathbb{N}_0^n$,

$$S_{\alpha,\beta}(\varphi) := \sup_{x \in \mathbb{R}^n} \bigl|x^\alpha \partial^\beta \varphi(x)\bigr|$$

and for nonnegative integers k, $l \in \mathbb{N}_0$ we put

$$\overline{S}_{k,l}(\varphi) := \max_{|\alpha| \leqslant k,\, |\beta| \leqslant l} S_{\alpha,\beta}(\varphi).$$

Remark 1.24. The quantities $S_{\alpha,\beta}(\varphi)$ and $\overline{S}_{k,l}(\varphi)$ are well-defined as extended real numbers when $\varphi \in \mathrm{C}^\infty(\mathbb{R}^n)$ and are finite when $\varphi \in \mathscr{S}(\mathbb{R}^n)$. In fact, their finiteness characterize the Schwartz space:

$$\mathscr{S}(\mathbb{R}^n) = \Big\{\varphi \in \mathrm{C}^\infty(\mathbb{R}^n) : \, S_{\alpha,\beta}(\varphi) < \infty \ \text{ for all } \ \alpha, \, \beta \in \mathbb{N}_0^n\Big\},$$

and likewise

$$\mathscr{S}(\mathbb{R}^n) = \Big\{\varphi \in \mathrm{C}^\infty(\mathbb{R}^n) : \, \overline{S}_{k.l}(\varphi) < \infty \ \text{ for all } \ k, \, l \in \mathbb{N}_0\Big\}.$$

It is easy to check that $S_{\alpha,\beta}$ and $\overline{S}_{k.l}$ are all norms on $\mathscr{S}(\mathbb{R}^n)$.

We record the following bounds that can be viewed as quantitative forms of (ii) and (iii) from Proposition 1.22.

Proposition 1.25. *Let $p(x) \in \mathbb{C}[x]$ be a polynomial of degree at most d:*

$$p(x) = \sum_{|\gamma| \leqslant d} c_\gamma x^\gamma.$$

Then for all $\varphi \in \mathscr{S}(\mathbb{R}^n)$ and all $k, \, l \in \mathbb{N}_0$ we have

$$\overline{S}_{k,l}(p\varphi) \leqslant \big(l+1\big)^d \Bigg(\sum_{|\gamma| \leqslant d} |c_\gamma|\Bigg) \overline{S}_{k+d,l}(\varphi) \tag{3}$$

and

$$\overline{S}_{k,l}(p(\partial)\varphi) \leqslant \Bigg(\sum_{|\gamma| \leqslant d} |c_\gamma|\Bigg) \overline{S}_{k,l+d}(\varphi). \tag{4}$$

Proof. In view of (2) we have for each $1 \leqslant j \leqslant n$ and multi-indices $|\alpha| \leqslant k$, $|\beta| \leqslant l$ that

$$\begin{aligned} S_{\alpha,\beta}(x_j\varphi) &\leqslant \beta_j S_{\alpha,\beta-\beta_j e_j}(\varphi) + S_{\alpha+e_j,\beta}(\varphi) \\ &\leqslant (\beta_j+1)\overline{S}_{|\alpha|+1,|\beta|}(\varphi) \\ &\leqslant (|\beta|+1)\overline{S}_{|\alpha|+1,|\beta|}(\varphi), \end{aligned}$$

hence $\overline{S}_{k,l}(x_j\varphi) \leqslant (l+1)\overline{S}_{k+1,l}(\varphi)$. For $\gamma \in \mathbb{N}_0^n$ with $|\gamma| \leqslant d$ we get (provided $\gamma_j \geqslant 1$ in the first line)

$$\begin{aligned} \overline{S}_{k,l}(x^\gamma\varphi) &\leqslant (l+1)\overline{S}_{k+1,l}(x^{\gamma-e_j}\varphi) \\ &\leqslant (l+1)^{\gamma_j}\overline{S}_{k+\gamma_j,l}(x^{\gamma-\gamma_j e_j}\varphi) \\ &\leqslant (l+1)^{|\gamma|}\overline{S}_{k+|\gamma|,l}(\varphi) \\ &\leqslant \big(l+1\big)^d\overline{S}_{k+d,l}(\varphi). \end{aligned}$$

The bound (3) follows easily from this. Finally the bound (4) follows easily from

$$S_{\alpha,\beta}(\partial^\gamma\varphi) = S_{\alpha,\beta+\gamma}(\varphi).$$

□

Definition 1.26. [Convergence in the sense of Schwartz test functions] Let φ_j, $\varphi \in \mathscr{S}(\mathbb{R}^n)$. Then we say φ_j converges to φ in the sense of Schwartz test functions, and write $\varphi_j \to \varphi$ in $\mathscr{S}(\mathbb{R}^n)$, if

$$S_{\alpha,\beta}(\varphi - \varphi_j) \longrightarrow 0$$

as $j \to \infty$ for all α, $\beta \in \mathbb{N}_0^n$. This can also be stated in terms of $\bar{S}_{k,l}$.

Remark 1.27. **[A metric on $\mathscr{S}(\mathbb{R}^n)$]** Define for $\varphi, \psi \in \mathscr{S}(\mathbb{R}^n)$

$$d(\varphi, \psi) := \sum_{k,l \in \mathbb{N}_0} 2^{-k-l} \frac{\bar{S}_{k,l}(\varphi - \psi)}{1 + \bar{S}_{k,l}(\varphi - \psi)}.$$

Then d is a metric on $\mathscr{S}(\mathbb{R}^n)$, and we have $\varphi_j \to \varphi$ in $\mathscr{S}(\mathbb{R}^n)$ if and only if $d(\varphi_j, \varphi) \to 0$. Note that d is translation invariant, meaning that,

$$d(\varphi + \eta, \psi + \eta) = d(\varphi, \psi)$$

holds for all φ, ψ, $\eta \in \mathscr{S}(\mathbb{R}^n)$. It can furthermore be shown that $(\mathscr{S}(\mathbb{R}^n), d)$ is complete and that the vector space operations are continuous (such a space is called a *Fréchet space*).

Remark 1.28. As in the case of compactly supported test functions the notion of convergence in $\mathscr{S}(\mathbb{R}^n)$ is severe: it requires a lot for a sequence to converge in the sense of Schwartz test functions. However, since clearly $\mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$ and the inclusion is strict it is not difficult to see that if φ_j, $\varphi \in \mathscr{D}(\mathbb{R}^n)$ and $\varphi_j \to \varphi$ in $\mathscr{D}(\mathbb{R}^n)$, then $\varphi_j, \varphi \in \mathscr{S}(\mathbb{R}^n)$ and $\varphi_j \to \varphi$ in $\mathscr{S}(\mathbb{R}^n)$. The converse is clearly false.

It is not difficult to show that $\mathscr{D}(\mathbb{R}^n)$ is an $\mathscr{S}$ dense subspace of $\mathscr{S}(\mathbb{R}^n)$: for each $\phi \in \mathscr{S}(\mathbb{R}^n)$ we can find $\phi_j \in \mathscr{D}(\mathbb{R}^n)$ so $\phi_j \to \phi$ in $\mathscr{S}(\mathbb{R}^n)$. We leave the details of this as an exercise (see Problem Sheet 2).

Example 1.29. Let $\varphi \in \mathscr{S}(\mathbb{R}^n)$ and $1 \leqslant j \leqslant n$. Then

$$\frac{\Delta_{he_j}\varphi}{h} \to \partial_j \varphi \text{ in } \mathscr{S}(\mathbb{R}^n) \text{ as } h \to 0.$$

Indeed using the fundamental theorem of calculus twice we find for α, $\beta \in \mathbb{N}_0^n$ that

$$S_{\alpha,\beta}\left(\frac{\Delta_{he_j}\varphi}{h} - \partial_j \varphi\right) \leqslant 2^{\alpha_j - 1}\Big(S_{\alpha,\beta+2e_j}(\varphi) + |h| S_{0,\beta+2e_j}(\varphi)\Big)|h|$$

and the assertion follows.

Example 1.30. Let $\left(\rho_\varepsilon\right)_{\varepsilon>0}$ be the standard mollifier on $\mathbb{R}^n$. Then for each $\varphi \in \mathscr{S}(\mathbb{R}^n)$ we have

$$\rho_\varepsilon * \varphi \to \varphi \text{ in } \mathscr{S}(\mathbb{R}^n) \text{ as } \varepsilon \searrow 0.$$

Using the fundamental theorem of calculus we find for α, $\beta \in \mathbb{N}_0^n$ that

$$S_{\alpha,\beta}\big(\rho_\varepsilon * \varphi - \varphi\big) \leqslant 2^{|\alpha|} \sum_{j=1}^{n} \Big(S_{\alpha,\beta+e_j}(\varphi) + \varepsilon^{|\alpha|} S_{0,\beta+e_j}(\varphi)\Big)\varepsilon$$

and the assertion follows. (Note also that $S_{\alpha,\beta}\big(\rho_\varepsilon * \varphi\big) < \infty$ follows by use of the triangle inequality.)

Corollary 1.31. *Let $p(x) \in \mathbb{C}[x]$ be a polynomial in n indeterminates. Then the maps $\varphi \mapsto p\varphi$ and $\varphi \mapsto p(\partial)\varphi$ are $\mathscr{S}$ continuous linear maps of $\mathscr{S}(\mathbb{R}^n)$ into $\mathscr{S}(\mathbb{R}^n)$.*

This is immediate from Proposition 1.25.

Proposition 1.32. *For all $p \in [1, \infty]$ we have $\mathscr{S}(\mathbb{R}^n) \subset \mathrm{L}^p(\mathbb{R}^n)$ and the inclusion map is continuous. More precisely, for each $p \in [1, \infty)$ there exists a constant $c = c(n, p)$ such that for all $\varphi \in \mathscr{S}(\mathbb{R}^n)$ we have*

$$\|\varphi\|_p \leqslant c(n,p)\overline{S}_{n+1,0}(\varphi).$$

For $p = \infty$ we have simply $\|\varphi\|_\infty = S_{0,0}(\varphi)$.

Proof. The point here is that the function $x \mapsto \left(1+|x|^2\right)^{-\frac{n+1}{2}}$ is integrable over $\mathbb{R}^n$: integrating in polar coordinates over $\mathbb{R}^n$ we find

$$\begin{aligned}\int_{\mathbb{R}^n} \frac{\mathrm{d}x}{\left(1+|x|^2\right)^{\frac{n+1}{2}}} &= \int_0^\infty \int_{\partial B_r(0)} \frac{\mathrm{d}S_x}{\left(1+|x|^2\right)^{\frac{n+1}{2}}}\,\mathrm{d}r \\ &= \omega_{n-1} \int_0^\infty \frac{r^{n-1}}{\left(1+r^2\right)^{\frac{n+1}{2}}}\,\mathrm{d}r < \infty.\end{aligned}$$

Next, writing for $p < \infty$

$$|\varphi(x)|^p = \left(1+|x|^2\right)^{-\frac{n+1}{2}} \left(\left(1+|x|^2\right)^{\frac{n+1}{2}} |\varphi(x)|^p \right)$$

we find

$$\|\varphi\|_p \leqslant \left(\int_{\mathbb{R}^n} \left(1+|x|^2\right)^{-\frac{n+1}{2}}\,\mathrm{d}x \right)^{\frac{1}{p}} \sup_{x \in \mathbb{R}^n} \left(\left(1+|x|^2\right)^{\frac{n+1}{2}} |\varphi(x)|^p \right)^{\frac{1}{p}}.$$

Elementary estimations show that we can take

$$c(n,p) = 2^{\frac{n}{p}} \left(1 + n^{\frac{n+1}{2}}\right)^{\frac{1}{p}} \left(\int_{\mathbb{R}^n} \left(1+|x|^2\right)^{-\frac{n+1}{2}}\,\mathrm{d}x \right)^{\frac{1}{p}},$$

but the exact value of the constant is not important here. □

Example 1.33. Recall that we defined the Sobolev space $\mathrm{W}^{k,p}(\mathbb{R}^n)$ for all $k \in \mathbb{N}$ and $p \in [1, \infty]$ to be all L^p functions whose distributional derivatives up to and including the order k are also L^p functions. Now since $\partial^\alpha \varphi \in \mathscr{S}(\mathbb{R}^n)$ for all $\alpha \in \mathbb{N}_0^n$ when $\varphi \in \mathscr{S}(\mathbb{R}^n)$ it follows from Lemma 1.32 that also $\mathscr{S}(\mathbb{R}^n) \subset \mathrm{W}^{k,p}(\mathbb{R}^n)$, and combining with Lemma 1.25 we see that the inclusion map is continuous (in fact we could write down bounds for the inclusion map).

Theorem 1.34. *The Fourier transform $\mathcal{F} \colon \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is a linear and $\mathscr{S}$ continuous map. The latter means that if $\varphi_j \to \varphi$ in $\mathscr{S}(\mathbb{R}^n)$, then also $\widehat{\varphi}_j \to \widehat{\varphi}$ in $\mathscr{S}(\mathbb{R}^n)$. The $\mathscr{S}$*

continuity is expressed more precisely through the Fourier bounds on $\mathscr{S}(\mathbb{R}^n)$*: for $k,\ l \in \mathbb{N}_0$ there exists a constant $c = c(n,k,l)$ so*

$$\overline{S}_{k,l}(\widehat{\varphi}) \leqslant c\overline{S}_{l+n+1,k}(\varphi) \tag{5}$$

for all $\varphi \in \mathscr{S}(\mathbb{R}^n)$.

Proof. Let $\varphi \in \mathscr{S}(\mathbb{R}^n)$. Then we have in particular that $x_j\varphi,\ \partial_j\varphi \in \mathscr{S}(\mathbb{R}^n)$ and since $\mathscr{S}(\mathbb{R}^n) \subset \mathrm{L}^1(\mathbb{R}^n)$ the differentiation rules give

$$\partial_j\widehat{\varphi}(\xi) = \mathcal{F}_{x\to\xi}\big(-\mathrm{i}x_j\varphi(x)\big) \text{ and } \xi_j\widehat{\varphi}(\xi) = -\mathrm{i}\widehat{\partial_j\varphi}(\xi).$$

It follows from the Riemann-Lebesgue lemma that $\partial_j\widehat{\varphi},\ \xi_j\widehat{\varphi} \in \mathrm{C}_0(\mathbb{R}^n)$. By induction on the length of multi-indices we find for $\alpha,\ \beta \in \mathbb{N}_0^n$ that

$$\partial^\beta\widehat{\varphi}(\xi) = \mathcal{F}_{x\to\xi}\big((-\mathrm{i}x)^\beta\varphi(x)\big) \text{ and } \xi^\alpha\widehat{\varphi}(\xi) = (-\mathrm{i})^{|\alpha|}\widehat{\partial^\alpha\varphi}(\xi).$$

both belong to $\mathrm{C}_0(\mathbb{R}^n)$, and hence that

$$\xi^\alpha\partial^\beta\widehat{\varphi}(\xi) = (-\mathrm{i})^{|\alpha|}\mathcal{F}_{x\to\xi}\Big(\partial^\alpha\big((-\mathrm{i}x)^\beta\varphi(\xi)\big)\Big) \in \mathrm{C}_0(\mathbb{R}^n). \tag{6}$$

Thus $S_{\alpha,\beta}(\widehat{\varphi}) < \infty$ and since in particular β was arbitrary also $\widehat{\varphi} \in \mathrm{C}^\infty(\mathbb{R}^n)$. But then $\widehat{\varphi} \in \mathscr{S}(\mathbb{R}^n)$. That the Fourier transform is $\mathscr{S}$ continuous follows of course if we can establish the Fourier bounds (5). First we recall that

$$\|\widehat{\psi}\|_\infty \leqslant \|\psi\|_1 \leqslant c\overline{S}_{n+1,0}(\psi)$$

for all $\psi \in \mathscr{S}(\mathbb{R}^n)$, where $c = c(n,1)$ is the constant from Proposition 1.32. Combining this bound with (6) we arrive at

$$\sup_{\xi\in\mathbb{R}^n}\big|\xi^\alpha\partial^\beta\widehat{\varphi}\big| \leqslant c\overline{S}_{n+1,0}\big(\partial^\alpha\big((-\mathrm{i}x)^\beta\varphi\big)\big).$$

The last term can be estimated by use of Proposition 1.25 whereby we find

$$\sup_{\xi\in\mathbb{R}^n}\big|\xi^\alpha\partial^\beta\widehat{\varphi}\big| \leqslant c\big(|\alpha|+1\big)^{|\beta|}\overline{S}_{n+1+|\beta|,|\alpha|}(\varphi).$$

Consequently we have shown that (5) holds with the constant $c = c(n,k,l) = c(n,1)\big(k+1\big)^l$, where $c(n,1)$ is the constant from Proposition 1.32. □

Remark 1.35. We record the following principle that is implicit in the above proof.

(a) Let $m \in \mathbb{N}_0$. If $f \in \mathrm{W}^{m,1}(\mathbb{R}^n)$ (with the convention $\mathrm{W}^{0,p}(\mathbb{R}^n) := \mathrm{L}^p(\mathbb{R}^n)$), then

$$\sup_{\xi\in\mathbb{R}^n}\big(1+|\xi|^2\big)^{\frac{m}{2}}|\hat{f}(\xi)| \leqslant c\|f\|_{\mathrm{W}^{m,1}},$$

where $c = c(n,m)$ is a constant. In fact, the Riemann-Lebesgue lemma tells us that the function $\xi \mapsto \big(1+|\xi|^2\big)^{\frac{m}{2}}\hat{f}(\xi)$ belongs to $\mathrm{C}_0(\mathbb{R}^n)$.

(b) Let $m \in \mathbb{N}$, $m \geqslant n+1$. If $\left(1+|x|^2\right)^{\frac{m}{2}} f(x) \in \mathrm{L}^\infty(\mathbb{R}^n)$, then $\hat{f} \in \mathrm{C}^{m-n-1}(\mathbb{R}^n)$ and $\partial^\alpha \hat{f}(\xi) \in \mathrm{C}_0(\mathbb{R}^n)$ for each multi-index α with $|\alpha| \leqslant m-n-1$.

There is clearly a gap of $n+1$ derivatives between (a) and (b), but in the proof of Theorem 1.34 it did not matter because the definition of a Schwartz function involves C^∞ smoothness and rapid decrease. The gap of $n+1$ derivatives between (a) and (b) is incurred when we go from L^∞ to L^1 using the bound from Proposition 1.32. Indeed, if instead of the L^∞ assumption in (b) we use an L^1 assumption as in (a) the result improves: If $\left(1+|x|^2\right)^{\frac{m}{2}} f(x) \in \mathrm{L}^1(\mathbb{R}^n)$, then $\hat{f} \in \mathrm{C}^m(\mathbb{R}^n)$ and $\partial^\alpha \hat{f} \in \mathrm{C}_0(\mathbb{R}^n)$ for each multi-index α with $|\alpha| \leqslant m$.

1.3 The Fourier inversion formula in $\mathscr{S}(\mathbb{R}^n)$ and in $\mathrm{L}^1(\mathbb{R}^n)$

Theorem 1.36. ***[Fourier Inversion Formula in $\mathscr{S}(\mathbb{R}^n)$]***

The Fourier transform $\mathcal{F}\colon \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is bijective with inverse given by

$$(\mathcal{F}^{-1}\psi)(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \psi(\xi) \mathrm{e}^{\mathrm{i}x\cdot\xi} \, \mathrm{d}\xi.$$

In symbols we have $\mathcal{F}^{-1} = (2\pi)^{-n}\widetilde{\mathcal{F}}$, where we recall that the reflection in the origin $\widetilde{\varphi}(x) := \varphi(-x)$ commutes with the Fourier transform.

Remark 1.37. So the Fourier transform $\mathcal{F}\colon \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is an $\mathscr{S}$ continuous linear bijective map with an $\mathscr{S}$ continuous and linear inverse. We also record that $\mathcal{F}^2\varphi = (2\pi)^n\widetilde{\varphi}$ so $\mathcal{F}^4 = (2\pi)^{2n}\mathrm{I}$ on $\mathscr{S}(\mathbb{R}^n)$.

The key to the proof of the inversion formula is the product rule: for all ϕ, $\psi \in \mathrm{L}^1(\mathbb{R}^n)$ we have

$$\int_{\mathbb{R}^n} \widehat{\phi}\psi \, \mathrm{d}x = \int_{\mathbb{R}^n} \phi\widehat{\psi} \, \mathrm{d}x.$$

The strategy is to make a good choice of ψ that will allow us to relate $\widehat{\phi}$ and ϕ. We start with the good choice:

Lemma 1.38. *If $G(x) = \mathrm{e}^{-\frac{|x|^2}{2}}$, $x \in \mathbb{R}^n$, then $\widehat{G} = (2\pi)^{\frac{n}{2}} G$.*

Proof. We reduce to the one-dimensional case by writing

$$G(x) = \mathrm{e}^{-\frac{|x|^2}{2}} = \prod_{j=1}^n \mathrm{e}^{-\frac{x_j^2}{2}},$$

and calculating by Tonelli's theorem

$$\hat{G}(\xi) = \prod_{j=1}^n \mathcal{F}_{x_j \to \xi_j}\left(\mathrm{e}^{-\frac{x_j^2}{2}}\right)(\xi_j).$$

Hence the conclusion follows if we can prove it for $n = 1$. In the remainder of this proof we therefore assume that $G(x) = \mathrm{e}^{-\frac{x^2}{2}}$, $x \in \mathbb{R}$. Note that $G(0) = 1$ and $G'(x) = -xG(x)$, so $G'(x) + xG(x) = 0$, that is, G is a solution to the initial value problem

$$\begin{cases} y' + xy = 0, & x \in \mathbb{R} \\ y(0) = 1. \end{cases} \tag{7}$$

Now by Fourier transforming the ODE and using the differentiation rules we find $0 = \mathrm{i}\xi\widehat{G} + \mathrm{i}\widehat{G}'$, or $\widehat{G}' + \xi\widehat{G} = 0$ for $\xi \in \mathbb{R}$. Also $\widehat{G}(0) = \int_{\mathbb{R}} G(x)\,\mathrm{d}x = \sqrt{2\pi}$ since

$$\begin{aligned} \widehat{G}(0)^2 = \int_{\mathbb{R}} G(x_1)\,\mathrm{d}x_1 \int_{\mathbb{R}} G(x_2)\,\mathrm{d}x_2 &= \int_{\mathbb{R}^2} \mathrm{e}^{-\frac{x_1^2+x_2^2}{2}}\,\mathrm{d}(x_1,x_2) \\ &= \int_0^\infty \int_{\partial B_r(0)} \mathrm{e}^{-\frac{r^2}{2}}\,\mathrm{d}S_x\,\mathrm{d}r \\ &= \int_0^\infty \mathrm{e}^{-\frac{r^2}{2}}\,2\pi r\,\mathrm{d}r \\ &= 2\pi. \end{aligned}$$

Consequently also $\widehat{G}/\sqrt{2\pi}$ satisfies (7), so by the uniqueness result for such solutions we conclude. (How do you reduce the needed uniqueness result to the constancy theorem via Leibniz' rule?) $\square$

The next result we need is an approximation result. We formulate it in a slightly more general form than is needed for present purposes.

Lemma 1.39. *Let $K \in \mathrm{L}^1(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} K\,\mathrm{d}x = 1$. Denoting by K_t the L^1 dilation of K by $t > 0$ we have for $\phi \in \mathscr{S}(\mathbb{R}^n)$*

$$K_t * \phi \to \phi \quad \text{in } \mathrm{L}^1(\mathbb{R}^n) \text{ and uniformly on } \mathbb{R}^n \text{ as } t \searrow 0.$$

For $f \in \mathrm{L}^1(\mathbb{R}^n)$ we have

$$K_t * f \to f \quad \text{in } \mathrm{L}^1(\mathbb{R}^n) \text{ as } t \searrow 0.$$

Remark 1.40. The family $(K_t)_{t>0}$ is called an approximate unit since the lemma in particular says that $K_t \to \delta_0$ in $\mathscr{D}'(\mathbb{R}^n)$ as $t \searrow 0$.

Proof. **[The proof is not examinable.]** Fix $\phi \in \mathscr{S}(\mathbb{R}^n)$ and let $\varepsilon > 0$. For $t > 0$ and $x \in \mathbb{R}^n$ we estimate

$$|K_t * \phi(x) - \phi(x)| \leqslant \int_{\mathbb{R}^n} |K(y)||\phi(x - ty) - \phi(x)|\,\mathrm{d}y.$$

The integral is split into two parts corresponding to $|y| \leqslant m$ and $|y| > m$, respectively, where $m > 0$ is chosen so

$$\int_{|y|>m} |K(y)|\,\mathrm{d}y < \frac{\varepsilon}{2(1 + 2\|\phi\|_\infty)}.$$

Hereby

$$\begin{aligned}\left|K_t * \phi(x) - \phi(x)\right| &\leqslant \left(\int_{|y|>m} + \int_{|y|\leqslant m}\right)|K(y)||\phi(x-ty) - \phi(x)|\,\mathrm{d}y\\ &\leqslant 2\|\phi\|_\infty \int_{|y|>m}|K(y)|\,\mathrm{d}y + \int_{|y|\leqslant m}|K(y)||\phi(x-ty)-\phi(x)|\,\mathrm{d}y\\ &< \tfrac{\varepsilon}{2} + \int_{|y|\leqslant m}|K(y)||\phi(x-ty)-\phi(x)|\,\mathrm{d}y.\end{aligned}$$

The second integral is estimated by use of the fundamental theorem of calculus, whereby we for each $x \in \mathbb{R}^n$, $t > 0$ and $|y| \leqslant m$ have

$$\left|\phi(x-ty) - \phi(x)\right| \leqslant \int_0^1 \left|\nabla\phi(x - sty)\right| tm\,\mathrm{d}s \leqslant \|\nabla\phi\|_\infty tm,$$

hence

$$\begin{aligned}\left|K_t * \phi(x) - \phi(x)\right| &\leqslant \tfrac{\varepsilon}{2} + \int_{|y|\leqslant m}|K(y)|\|\nabla\phi\|_\infty tm\,\mathrm{d}y\\ &\leqslant \tfrac{\varepsilon}{2} + \|K\|_1\|\nabla\phi\|_\infty mt < \varepsilon\end{aligned}$$

provided $t \in (0, \frac{\varepsilon}{2(1+\|K\|_1\|\nabla\phi\|_\infty m)})$. This establishes the uniform convergence, $K_t * \phi \to \phi$ on $\mathbb{R}^n$ as $t \searrow 0$. To establish the convergence in L^1 we proceed in a similar way, this time taking $m > 0$ so

$$\int_{|y|>m}|K(y)|\,\mathrm{d}y < \frac{\varepsilon}{2(1 + 2\|\phi\|_1)}.$$

Then we estimate, using Tonelli's theorem to swap integration orders in the last line:

$$\begin{aligned}\|K_t * \phi - \phi\|_1 &= \int_{\mathbb{R}^n}\left|\int_{\mathbb{R}^n} K(y)\big(\phi(x-ty) - \phi(x)\big)\,\mathrm{d}y\right|\,\mathrm{d}x\\ &\leqslant \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}|K(y)|\left|\phi(x-ty)-\phi(x)\right|\,\mathrm{d}y\,\mathrm{d}x\\ &= \int_{\mathbb{R}^n}|K(y)|\int_{\mathbb{R}^n}\left|\phi(x-ty)-\phi(x)\right|\,\mathrm{d}x\,\mathrm{d}y.\end{aligned}$$

We split the y-integral as in the previous part of the proof and with the above choice of m we estimate as before resulting in

$$\|K_t * \phi - \phi\|_1 \leqslant \tfrac{\varepsilon}{2} + \|K\|_1\|\nabla\phi\|_1 mt < \varepsilon$$

provided $t \in (0, \frac{\varepsilon}{2(1+\|K\|_1\|\nabla\phi\|_1 m)})$.

Finally for $f \in \mathrm{L}^1(\mathbb{R}^n)$ we take $\phi \in \mathscr{S}(\mathbb{R}^n)$ so $\|f - \phi\|_1 < \varepsilon/2$. Then

$$\begin{aligned}\|K_t * f - f\|_1 &\leqslant \|K_t * (f-\phi)\| + \|K_t * \phi - \phi\|_1 + \|\phi - f\|_1\\ &\leqslant 2\|f - \phi\|_1 + \|K_t * \phi - \phi\|_1\\ &< \varepsilon + \|K_t * \phi - \phi\|_1\end{aligned}$$

and the conclusion follows from the result for $\phi \in \mathscr{S}(\mathbb{R}^n)$. □

Proof of Theorem 1.36. First we note that since by Lemma 1.38

$$\int_{\mathbb{R}^n}\widehat{G}\,\mathrm{d}\xi = \int_{\mathbb{R}^n}(2\pi)^{n/2}G\,\mathrm{d}\xi = (2\pi)^{n/2}\widehat{G}(0) = (2\pi)^n$$

we can use Lemma 1.39 with $K = (2\pi)^{-n}\widehat{G}$. Accordingly we have for $\phi \in \mathscr{S}(\mathbb{R}^n)$ that

$$\Big(\big((2\pi)^{-n}\widehat{G}\big)_t * \phi\Big)(x) \to \phi(x) \text{ uniformly in } x \in \mathbb{R}^n \text{ as } t \searrow 0.$$

Here we can also write the left-hand side, using the dilation, product and translation rules (in that order):

$$\begin{aligned}\Big(\big((2\pi)^{-n}\widehat{G}\big)_t * \phi\Big)(x) =&(2\pi)^{-n}\int_{\mathbb{R}^n}\phi(x-y)\big(\widehat{G}\big)_t(y)\,\mathrm{d}y\\ =&(2\pi)^{-n}\int_{\mathbb{R}^n}\mathcal{F}_{\xi\to y}\big(\phi(x-\xi)\big)G(ty)\,\mathrm{d}y\\ =&(2\pi)^{-n}\int_{\mathbb{R}^n}\widehat{\phi}(-y)\mathrm{e}^{-\mathrm{i}y\cdot x}G(ty)\,\mathrm{d}y.\end{aligned}$$

By use of Lebesgue's dominated convergence theorem we have, as $t\searrow 0$,

$$\begin{aligned}(2\pi)^{-n}\int_{\mathbb{R}^n}\widehat{\phi}(-y)\mathrm{e}^{-\mathrm{i}y\cdot x}G(ty)\,\mathrm{d}y \to&(2\pi)^{-n}\int_{\mathbb{R}^n}\widehat{\phi}(-y)\mathrm{e}^{-\mathrm{i}y\cdot x}\,\mathrm{d}y\\ =&(2\pi)^{-n}\int_{\mathbb{R}^n}\widehat{\phi}(y)\mathrm{e}^{\mathrm{i}y\cdot x}\,\mathrm{d}y\end{aligned}$$

concluding the proof. □

The above method is easily adapted to also give the Fourier Inversion Formula in L^1. While this result will also be a consequence of the much more general Fourier Inversion Formula that we establish in the next section we have chosen to present it here for two reasons. First, we shall use it to prove a convolution rule in this section. Second, and more important, it serves as an illustration that even though the general definitions we make in the next section look quite soft they are indeed well chosen and have interesting ramifications!

Theorem 1.41. ***[Fourier Inversion Formula in*** $\mathrm{L}^1(\mathbb{R}^n)$***]***

Let $f\in\mathrm{L}^1(\mathbb{R}^n)$*. Then*

$$f(x)=\lim_{t\searrow 0}(2\pi)^{-n}\int_{\mathbb{R}^n}\hat{f}(\xi)\mathrm{e}^{\mathrm{i}x\cdot\xi-\frac{t^2|\xi|^2}{2}}\,\mathrm{d}\xi \quad \textit{in} \quad \mathrm{L}^1(\mathbb{R}^n). \tag{8}$$

Consequently, when also $\hat{f}\in\mathrm{L}^1(\mathbb{R}^n)$*, then*

$$f(x)=(2\pi)^{-n}\int_{\mathbb{R}^n}\hat{f}(\xi)\mathrm{e}^{\mathrm{i}x\cdot\xi}\,\mathrm{d}\xi \tag{9}$$

holds almost everywhere.

Proof. As above we have by the product, translation and dilation rules for each $t>0$ and all $x\in\mathbb{R}^n$:

$$\big(f*\big((2\pi)^{-n}\hat{G}\big)_t\big)(x)=(2\pi)^{-n}\int_{\mathbb{R}^n}\hat{f}(\xi)\mathrm{e}^{\mathrm{i}x\cdot\xi-\frac{t^2|\xi|^2}{2}}\,\mathrm{d}\xi.$$

By the Lemma 1.39 the left-hand side converges to f in $\mathrm{L}^1(\mathbb{R}^n)$ as $t\searrow 0$, hence (8) holds. If also $\hat{f}\in\mathrm{L}^1(\mathbb{R}^n)$, then Lebesgue's dominated convergence theorem allows us to conclude (9). Note that in this case the Riemann-Lebesgue lemma tells us that the right-hand side of (9) belongs to $\mathrm{C}_0(\mathbb{R}^n)$. □

Proposition 1.42. ***[The convolution rule]*** *If φ, $\psi \in \mathscr{S}(\mathbb{R}^n)$, then*

$$\widehat{(\varphi\psi)} = (2\pi)^{-n}\widehat{\varphi} * \widehat{\psi}.$$

Proof. Since $\widehat{\varphi}$, $\widehat{\psi}$ both belong to $\mathscr{S}(\mathbb{R}^n)$ they are in particular integrable, and so we can use the convolution rule we derived for L^1 functions. Hereby we find, using the Fourier Inversion Formula in $\mathscr{S}$ in the second step,

$$\begin{aligned}\widehat{\widehat{\varphi} * \widehat{\psi}} =& \widehat{\widehat{\varphi}}\widehat{\widehat{\psi}} \\ =& (2\pi)^{2n}\widetilde{\varphi}\widetilde{\psi},\end{aligned}$$

and the conclusion follows by use of the Fourier Inversion Formula in L^1. □

Corollary 1.43. *If ϕ, $\psi \in \mathscr{S}(\mathbb{R}^n)$, then $\phi * \psi \in \mathscr{S}(\mathbb{R}^n)$ too.*

Using the Fourier bounds (5) one can, for instance, show bounds of the form

$$\overline{S}_{k,l}(\phi * \psi) \leqslant c\overline{S}_{k+n+1,l+n+1}(\phi)\overline{S}_{k+n+1,0}(\psi),$$

where ϕ, $\psi \in \mathscr{S}(\mathbb{R}^n)$, k, $l \in \mathbb{N}_0$ and $c = c(n,k,l)$ is a constant.

Exercise: Prove that for multi-indices α, $\beta \in \mathbb{N}_0^n$ and ϕ, $\psi \in \mathscr{S}(\mathbb{R}^n)$ we have

$$S_{\alpha,\beta}(\phi * \psi) \leqslant 2^{|\alpha|}c(n,1)\overline{S}_{|\alpha|,|\beta|}(\phi)\overline{S}_{n+1+|\alpha|,0}(\psi),$$

where $c(n,1)$ is the constant from Proposition 1.32. Conclude that $\phi * \psi \in \mathscr{S}(\mathbb{R}^n)$.
(*Hint:* Write for x, $y \in \mathbb{R}^n$,

$$x^\alpha = (x-y+y)^\alpha = \sum_{\gamma \leqslant \alpha} \binom{\alpha}{\gamma}(x-y)^{\alpha-\gamma}y^\gamma,$$

and deduce that

$$x^\alpha(\phi * \psi)(x) = \sum_{\gamma \leqslant \alpha} \binom{\alpha}{\gamma}\Big(\big((\cdot)^{\alpha-\gamma}\phi\big) * \big((\cdot)^\gamma\psi\big)\Big)(x).$$

The result follows from this.)

1.4 Tempered distributions

The product rule that was established for L^1 functions becomes an adjoint identity for the Fourier transform if we restrict it to Schwartz test functions. We can then consistently extend the Fourier transform to a class of distributions:

Definition 1.44. [Tempered Distributions] A functional $u\colon \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ is a tempered distribution if

(i) u is linear, and

(ii) u is $\mathscr{S}$ continuous: if $\phi_j \to \phi$ in $\mathscr{S}(\mathbb{R}^n)$, then $u(\phi_j) \to u(\phi)$.

The set of all tempered distributions on $\mathbb{R}^n$ is denoted by $\mathscr{S}'(\mathbb{R}^n)$.

Bracket notation: As for the other classes of distributions we often write $\langle u, \phi \rangle = u(\phi)$.

Note that when $u \colon \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ is a linear functional, then the $\mathscr{S}$ continuity (ii) will follow provided we can prove it holds for $\phi = 0$.

Remark 1.45. Because $\mathscr{D}(\mathbb{R}^n) < \mathscr{S}(\mathbb{R}^n) < \mathrm{C}^\infty(\mathbb{R}^n)$ (where we recall that '$<$' signifies 'proper subspace of') it follows that

$$\mathscr{E}'(\mathbb{R}^n) < \mathscr{S}'(\mathbb{R}^n) < \mathscr{D}'(\mathbb{R}^n).$$

Indeed, if $u \in \mathscr{S}'(\mathbb{R}^n)$, then the restriction $u|_{\mathscr{D}(\mathbb{R}^n)}$ of course remains linear and if $\phi_j \to 0$ in $\mathscr{D}(\mathbb{R}^n)$, then the convergence also holds in the $\mathscr{S}(\mathbb{R}^n)$ sense, whereby $\langle \big(u|_{\mathscr{D}(\mathbb{R}^n)}\big), \phi_j \rangle = \langle u, \phi_j \rangle \to 0$ proving that $u|_{\mathscr{D}(\mathbb{R}^n)} \in \mathscr{D}'(\mathbb{R}^n)$. The subspace test easily gives that $\mathscr{S}'(\mathbb{R}^n)$ is a subspace of $\mathscr{D}'(\mathbb{R}^n)$. To see that it is a proper subspace we show that $\mathrm{e}^{|x|^2} \in \mathscr{D}'(\mathbb{R}^n) \setminus \mathscr{S}'(\mathbb{R}^n)$: we only need to argue that it is not a tempered distribution. Assume it were and denote it by T. Then we would have $\langle T, \phi \rangle = \int_{\mathbb{R}^n} \phi(x) \mathrm{e}^{|x|^2} \, \mathrm{d}x$ for $\phi \in \mathscr{D}(\mathbb{R}^n)$. Now $\mathrm{e}^{-|x|^2} \in \mathscr{S}(\mathbb{R}^n)$ and taking $\phi_j = \mathrm{e}^{-|x|^2} \chi_j \in \mathscr{D}(\mathbb{R}^n)$, where $\chi_j = \rho * \mathbf{1}_{B_j(0)}$ (ρ is the standard mollifier kernel on $\mathbb{R}^n$), we see that $\phi_j \to \mathrm{e}^{-|x|^2}$ in $\mathscr{S}(\mathbb{R}^n)$. The specific construction also means that $\phi_j(x) \nearrow \mathrm{e}^{-|x|^2}$ pointwise in $x \in \mathbb{R}^n$, so using Lebesgue's monotone convergence theorem and that T is $\mathscr{S}$ continuous we get a contradiction, namely

$$\begin{aligned} \langle T, \mathrm{e}^{-|x|^2} \rangle &= \lim_{j \to \infty} \langle T, \phi_j \rangle = \lim_{j \to \infty} \int_{\mathbb{R}^n} \mathrm{e}^{|x|^2} \phi_j(x) \, \mathrm{d}x \\ &= \int_{\mathbb{R}^n} \mathbf{1}_{\mathbb{R}^n}(x) \, \mathrm{d}x = \infty. \end{aligned}$$

The proof that the space of compactly supported distributions forms a proper subspace of the tempered distributions is left as an exercise.

Remark 1.46. We emphasize that a tempered distribution is uniquely determined by its values on $\mathscr{D}(\mathbb{R}^n)$, namely if $u \in \mathscr{S}'(\mathbb{R}^n)$ and $\langle u, \phi \rangle = 0$ holds for all $\phi \in \mathscr{D}(\mathbb{R}^n)$, then $u = 0$. This follows because $\mathscr{D}(\mathbb{R}^n)$ is $\mathscr{S}$ dense in $\mathscr{S}(\mathbb{R}^n)$ and u is $\mathscr{S}$ continuous.

Example 1.47. [L^p **functions as tempered distributions**] We have $\mathrm{L}^p(\mathbb{R}^n) < \mathscr{S}'(\mathbb{R}^n)$ for all $p \in [1, \infty]$, where the tempered distribution corresponding to $f \in \mathrm{L}^p(\mathbb{R}^n)$ will be denoted by T_f and given by the rule

$$\langle T_f, \varphi \rangle = \int_{\mathbb{R}^n} f \varphi \, \mathrm{d}x, \quad \varphi \in \mathscr{S}(\mathbb{R}^n).$$

We must show that it is well-defined, linear and $\mathscr{S}$ continuous. Fix $\varphi \in \mathscr{S}(\mathbb{R}^n)$. Clearly $x \mapsto f(x)\varphi(x)$ is measurable. If q is the Hölder conjugate exponent to p, so $\frac{1}{p} + \frac{1}{q} = 1$, then

Hölder's inequality yields

$$\int_{\mathbb{R}^n} |f\varphi| \, \mathrm{d}x \leqslant \|f\|_p \|\varphi\|_q \leqslant c(n,q)\|f\|_p \overline{S}_{n+1,0}(\varphi) < \infty.$$

The bound is slightly better when $q = \infty$, see Proposition 1.32. In any case, T_f is well-defined on $\mathscr{S}(\mathbb{R}^n)$, and is then clearly also linear there. Furthermore, the above bound implies that

$$\left|\langle T_f, \varphi \rangle\right| \leqslant c\overline{S}_{n+1,0}(\varphi)$$

holds for all $\varphi \in \mathscr{S}(\mathbb{R}^n)$, where c is a constant. Consequently T_f is $\mathscr{S}$ continuous and therefore a tempered distribution (indeed, if $\phi_j \to 0$ in $\mathscr{S}(\mathbb{R}^n)$, then the above bound gives $\langle T_f, \phi_j \rangle \to 0$).

Because the tempered distribution T_f is uniquely determined by its values on $\mathscr{D}(\mathbb{R}^n)$ it follows from the fundamental lemma of the calculus of variations (see B4.3) that T_f uniquely determines f as an L^p function. We therefore also in the case of tempered distributions identify T_f directly with f, and write $T_f = f$.

Example 1.48. **[Finite Borel measures as tempered distributions]** If μ is a finite Borel measure on $\mathbb{R}^n$, then we define

$$\langle T_\mu, \varphi \rangle = \int_{\mathbb{R}^n} \varphi \, \mathrm{d}\mu, \quad \varphi \in \mathscr{S}(\mathbb{R}^n).$$

It is easy to see that T_μ is well-defined and linear on $\mathscr{S}(\mathbb{R}^n)$, and since also $\left|\langle T_\mu, \varphi \rangle\right| \leqslant \mu(\mathbb{R}^n) S_{0,0}(\varphi)$ holds for all $\varphi \in \mathscr{S}(\mathbb{R}^n)$ it is also $\mathscr{S}$ continuous, hence is a tempered distribution. As in the $\mathscr{D}'(\mathbb{R}^n)$ context the distribution T_μ uniquely determines the measure μ, so in the sequel we identify T_μ with μ and write simply $T_\mu = \mu$.

In particular we record that the Diract measures δ_a concentrated at points $a \in \mathbb{R}^n$ are tempered distributions.

As we have seen that the function $f = \mathrm{e}^{|x|^2} \in \mathrm{L}^\infty_{\mathrm{loc}}(\mathbb{R}^n)$ does not correspond to a tempered distribution we must conclude that in order to be a tempered distribution, a function cannot grow too fast at infinity. This is admittedly quite vague, but it has to be and we will return to this point later. Meanwhile we introduce the natural replacement for local L^p functions in the tempered context:

Definition 1.49. [Tempered L^p functions and measures] Let $p \in [1,\infty]$. Then $f \in \mathrm{L}^p_{\mathrm{loc}}(\mathbb{R}^n)$ is a *tempered* L^p *function* if there exists $m \in \mathbb{N}_0$ so

$$\frac{f(x)}{\left(1+|x|^2\right)^{\frac{m}{2}}} \in \mathrm{L}^p(\mathbb{R}^n).$$

A Borel measure μ on $\mathbb{R}^n$ is a *tempered measure* if there exists $m \in \mathbb{N}_0$ so $\left(1+|x|^2\right)^{-\frac{m}{2}}\mu$ is a finite measure on $\mathbb{R}^n$:

$$\int_{\mathbb{R}^n} \frac{\mathrm{d}\mu(x)}{\left(1+|x|^2\right)^{\frac{m}{2}}} < \infty.$$

Example 1.50. **[Tempered L^p functions and measures as tempered distributions]**

Fix $p \in [1,\infty]$ and assume that f is a tempered L^p function, say $f(x)/\big(1+|x|^2\big)^{m/2} \in \mathrm{L}^p(\mathbb{R}^n)$. If $\phi \in \mathscr{S}(\mathbb{R}^n)$, then from

$$f(x)\phi(x) = \frac{f(x)}{\big(1+|x|^2\big)^{\frac{m}{2}}}\big(1+|x|^2\big)^{\frac{m}{2}}\phi(x)$$

we get by use of Hölder's inequality

$$\begin{aligned}\int_{\mathbb{R}^n} |f(x)\phi(x)|\,\mathrm{d}x \leqslant & \left\| \frac{f(\cdot)}{\big(1+|\cdot|^2\big)^{\frac{m}{2}}} \right\|_p \left\| \big(1+|\cdot|^2\big)^{\frac{m}{2}}\phi(\cdot) \right\|_q \\ \leqslant & c\overline{S}_{m+n+1,0}(\phi) < \infty,\end{aligned}$$

where c is a constant that depends on f, n, q and obtained from Proposition 1.32. It follows that

$$\langle f,\phi\rangle = \int_{\mathbb{R}^n} f\phi\,\mathrm{d}x, \quad \phi \in \mathscr{S}(\mathbb{R}^n),$$

is well-defined, linear and $\mathscr{S}$ continuous, hence that it is a tempered distribution on $\mathbb{R}^n$.

Likewise if μ is a tempered measure on $\mathbb{R}^n$, then

$$\langle \mu,\phi\rangle = \int_{\mathbb{R}^n} \phi\,\mathrm{d}\mu, \quad \phi \in \mathscr{S}(\mathbb{R}^n)$$

is well-defined, linear and $\mathscr{S}$ continuous, so also a tempered distribution on $\mathbb{R}^n$.

As in the case of the $\mathscr{D}'$ distributions the crucial continuity property of tempered distributions can be recast as a boundedness property.

Proposition 1.51. ***[The boundedness property of tempered distributions]***

Let $u\colon \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ be linear. Then $u \in \mathscr{S}'(\mathbb{R}^n)$ if and only if there exist constants $c \geqslant 0$, k, $l \in \mathbb{N}_0$ such that

$$\big|\langle u,\, \varphi\rangle\big| \leqslant c\bar{S}_{k,l}(\varphi) \tag{10}$$

holds for all $\varphi \in \mathscr{S}(\mathbb{R}^n)$.

Remark 1.52. It follows that tempered distributions have finite order: when (10) holds the order of u is at most l.

Proof. The *if* part is clear. To prove the *only if* statement, assume u is $\mathscr{S}$ continuous but that (10) fails for all $c = k = l = j \in \mathbb{N}$: there exist $\varphi_j \in \mathscr{S}(\mathbb{R}^n)$ such that

$$|\langle u,\, \varphi_j\rangle| > j\bar{S}_{j,j}(\varphi_j).$$

Clearly $\varphi_j \neq 0$, so $\bar{S}_{j,j}(\varphi_j) > 0$ and we may define

$$\psi_j = \frac{\varphi_j}{j\bar{S}_{j,j}(\varphi_j)} \in \mathscr{S}(\mathbb{R}^n).$$

Fix $\alpha, \beta \in \mathbb{N}_0^n$. Then for $j \geqslant |\alpha|$, $|\beta|$ we have $S_{\alpha,\beta}(\psi_j) \leqslant j^{-1}$, so $S_{\alpha,\beta}(\psi_j) \to 0$ as $j \to \infty$. Since α, β were arbitrary we conclude that $\psi_j \to 0$ in $\mathscr{S}(\mathbb{R}^n)$ and so, by $\mathscr{S}$ continuity of u, $\langle u, \psi_j \rangle \to 0$. But this contradicts $|\langle u, \psi_j\rangle| > 1$ for all $j \in \mathbb{N}$. □

Definition 1.53. **[Convergence of Tempered Distributions]** For a sequence (u_j) in $\mathscr{S}'(\mathbb{R}^n)$ and $u \in \mathscr{S}'(\mathbb{R}^n)$ we write

$$u_j \longrightarrow u \;\text{ in } \mathscr{S}'(\mathbb{R}^n)$$

if $\langle u_j, \varphi\rangle \to \langle u, \varphi\rangle$ for each fixed $\varphi \in \mathscr{S}(\mathbb{R}^n)$.

Remark 1.54. Since $\mathscr{D}(\mathbb{R}^n) < \mathscr{S}(\mathbb{R}^n)$ this is stronger than convergence in $\mathscr{D}'(\mathbb{R}^n)$, but as in that case is otherwise a very weak notion of convergence. Note that if $K \in \mathrm{L}^1(\mathbb{R}^n)$ and $\int_{\mathbb{R}^n} K \,\mathrm{d}x = 1$, then $K_t \to \delta_0$ in $\mathscr{S}'(\mathbb{R}^n)$ as $t \searrow 0$. This follows from Lemma 1.39.

1.4.1 The adjoint identity scheme in the tempered context.

Using the adjoint identity scheme in $\mathscr{S}(\mathbb{R}^n)$ we may consistently extend operations on Schwartz test functions to tempered distributions. The procedure is exactly the same as in the $\mathscr{D}$ context: we have an *operation* on Schwartz test functions that we would like to extend to tempered distributions. This is a linear map $T\colon \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ and we assume that there exists a linear and $\mathscr{S}$ continuous map $S\colon \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ so that the *adjoint identity*

$$\int_{\mathbb{R}^n} T(\phi)\psi \,\mathrm{d}x = \int_{\mathbb{R}^n} \phi S(\psi)\,\mathrm{d}x$$

holds for all ϕ, $\psi \in \mathscr{S}(\mathbb{R}^n)$. We may then define $\bar{T}\colon \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ by the rule: for $u \in \mathscr{S}'(\mathbb{R}^n)$ put

$$\big\langle \bar{T}(u), \phi\big\rangle := \big\langle u, S(\phi)\big\rangle, \quad \phi \in \mathscr{S}(\mathbb{R}^n).$$

We note that hereby $\bar{T}(u)\colon \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is linear and $\mathscr{S}$ continuous, so $\bar{T}\colon \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ is well-defined. It is then clearly also linear. It is $\mathscr{S}'$ continuous, since the definitions easily give that if $u_j \to u$ in $\mathscr{S}'(\mathbb{R}^n)$, then also $\bar{T}(u_j) \to \bar{T}(u)$ in $\mathscr{S}'(\mathbb{R}^n)$. The adjoint identity guarantees consistency, meaning that $\bar{T}|_{\mathscr{S}(\mathbb{R}^n)} = T$, and as we did in the $\mathscr{D}$ context we shall usually skip the bar and simply denote the extension $\bar{T}$ by T again.

Definition 1.55. For $\alpha \in \mathbb{N}_0^n$, $p(x) \in \mathbb{C}[x]$, $\theta \in \mathscr{S}(\mathbb{R}^n)$ and $u \in \mathscr{S}'(\mathbb{R}^n)$ we define the tempered distributions $\partial^\alpha u$, pu, $\hat{u}$ and $u * \theta$ by the rules

$$\langle \partial^\alpha u, \varphi\rangle \coloneqq (-1)^{|\alpha|}\langle u, \partial^\alpha \varphi\rangle,$$

$$\langle pu,\,\varphi\rangle := \langle u,\,p\varphi\rangle,$$
$$\langle \hat{u},\,\varphi\rangle := \langle u,\,\hat{\varphi}\rangle,$$
$$\langle u*\theta,\,\varphi\rangle := \langle u,\,\tilde{\theta}*\varphi\rangle.$$

We also define composition with orthogonal maps $\theta_* u$, $\tilde{u}$, the translation $\tau_h u$, and dilations $d_r u$, u_r as on $\mathscr{D}'(\mathbb{R}^n)$. All of the above operations are linear and continuous in the sense of $\mathscr{S}'(\mathbb{R}^n)$, and, as we will see, the rules for the Fourier Transform on $\mathscr{S}(\mathbb{R}^n)$ also hold on $\mathscr{S}'(\mathbb{R}^n)$.

Remark 1.56. **[Consistency with definition of $\mathcal{F}$ on L^1]** We have consistency on the space of Schwartz test functions by the adjoint identity scheme, but we should also check that we have consistency with our definition of the Fourier transform on L^1. To make the discussion clearer we shall in this remark revert to the notation T_g for the tempered distribution corresponding to the tempered L^p function g. Fix $f \in \mathrm{L}^1(\mathbb{R}^n)$ and consider the corresponding tempered distribution T_f. Now $\widehat{T_f} \in \mathscr{S}'(\mathbb{R}^n)$ is given by the rule $\langle \widehat{T_f}, \phi\rangle := \int_{\mathbb{R}^n} f\hat{\phi}\,\mathrm{d}x$ for $\phi \in \mathscr{S}(\mathbb{R}^n)$. By the product rule on L^1 we have for $\phi \in \mathscr{S}(\mathbb{R}^n)$:

$$\int_{\mathbb{R}^n} \widehat{f}\phi\,\mathrm{d}x = \int_{\mathbb{R}^n} f\widehat{\phi}\,\mathrm{d}x = \langle \widehat{T}_f, \phi\rangle,$$

and consequently $\widehat{T}_f = T_{\widehat{f}}$.

Theorem 1.57. ***[Fourier Inversion Formula on $\mathscr{S}'(\mathbb{R}^n)$]***
The Fourier transform $\mathcal{F}\colon \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ is a linear bijection with inverse $\mathcal{F}^{-1} = (2\pi)^{-n}\tilde{\mathcal{F}}$.

Proof. We check that $\big((2\pi)^{-n}\tilde{\mathcal{F}}\big) \circ \mathcal{F} = \mathcal{F} \circ \big((2\pi)^{-n}\tilde{\mathcal{F}}\big) = \mathrm{I}$, the identity on $\mathscr{S}'(\mathbb{R}^n)$. Our definitions allow us to deduce it from the Fourier Inversion Formula on $\mathscr{S}(\mathbb{R}^n)$ as follows. Fix $u \in \mathscr{S}'(\mathbb{R}^n)$. Then for $\varphi \in \mathscr{S}(\mathbb{R}^n)$ we have

$$\begin{aligned}\big\langle (2\pi)^{-n}\tilde{\mathcal{F}}\mathcal{F}u,\,\varphi\big\rangle =& \big\langle u,\,(2\pi)^{-n}\mathcal{F}\tilde{\mathcal{F}}\varphi\big\rangle\\ =& \big\langle u,\,\varphi\big\rangle\\ =& \big\langle u,\,(2\pi)^{-n}\tilde{\mathcal{F}}\mathcal{F}\varphi\big\rangle\\ =& \big\langle (2\pi)^{-n}\mathcal{F}\tilde{\mathcal{F}}u,\,\varphi\big\rangle.\end{aligned}$$

□

Example 1.58. **[Fourier transforms of finite Borel measures]** Let μ be a finite Borel measure on $\mathbb{R}^n$, so that in particular $\mu \in \mathscr{S}'(\mathbb{R}^n)$. For $\phi \in \mathscr{S}(\mathbb{R}^n)$ we have

$$\begin{aligned}\langle \hat{\mu}, \phi\rangle =& \langle \mu, \hat{\phi}\rangle = \int_{\mathbb{R}^n} \hat{\phi}(\xi)\,\mathrm{d}\mu(\xi)\\ =& \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \phi(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\,\mathrm{d}x\,\mathrm{d}\mu(\xi).\end{aligned}$$

Note that

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\left|\phi(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\right|\,\mathrm{d}x\,\mathrm{d}\mu(\xi) = \|\phi\|_1\mu(\mathbb{R}^n) \leqslant c(n,1)\mu(\mathbb{R}^n)\overline{S}_{n+1,0}(\phi) < \infty$$

where $c(n,1)$ is the constant from Proposition 1.32. Hence by Fubini's theorem we may swap the integration orders and conclude that

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}x\cdot\xi}\,\mathrm{d}\mu(x), \quad \xi \in \mathbb{R}^n.$$

It is not difficult to prove that $\hat{\mu}$ is a uniformly continuous function on $\mathbb{R}^n$. In general it will however not vanish at infinity and so not belong to $\mathrm{C}_0(\mathbb{R}^n)$. For instance we have that $\widehat{\delta_a} = \mathrm{e}^{-\mathrm{i}a\cdot\xi}$ and in particular that

$$\widehat{\delta_0} = 1$$

or more precisely it is the function $\mathbf{1}_{\mathbb{R}^n}$. If we use the Fourier Inversion Formula in $\mathscr{S}'(\mathbb{R}^n)$ we get $\widehat{\mathbf{1}_{\mathbb{R}^n}} = (2\pi)^n\delta_0$. Note that we can write the latter as

$$\delta_0 = (2\pi)^{-n}\lim_{r\to\infty}\int_{B_r(0)} \mathrm{e}^{-\mathrm{i}x\cdot\xi}\,\mathrm{d}x, \tag{11}$$

where the convergence is understood in $\mathscr{S}'(\mathbb{R}^n)$. The formula (11) is the *Fourier-Gel'fand formula* for Dirac's delta-function.

Proposition 1.59. *The rules stated in Propositions 1.3 (differentiation rule), 1.9 (invariance under orthogonal maps), 1.11 (translation rules), 1.12 (dilation rules), 1.15 (reverse differentiation rule) and Corollary 1.16 (generalized differentiation rules) remain true when* $u \in \mathscr{S}'(\mathbb{R}^n)$.

The proof is straight forward and left as an exercise. In particular note how the proof of Corollary 1.16 simplifies in the present more general context!

Example 1.60. We defined the principal value distribution $\mathrm{pv}\big(\frac{1}{x}\big)$ by

$$\big\langle \mathrm{pv}\big(\tfrac{1}{x}\big), \phi\big\rangle = \lim_{\varepsilon\searrow 0}\left(\int_{-\infty}^{-\varepsilon} + \int_{\varepsilon}^{\infty}\right)\frac{\phi(x)}{x}\,\mathrm{d}x, \quad \phi \in \mathscr{D}(\mathbb{R}).$$

It is not difficult to see that we may take $\phi \in \mathscr{S}(\mathbb{R})$ above and that $\mathrm{pv}\big(\frac{1}{x}\big)$ hereby is a tempered distribution. We can also note that $\log|x|$ is a tempered $\mathrm{L}^1(\mathbb{R})$ function, hence a tempered distribution and that its distributional derivative equals $\mathrm{pv}\big(\frac{1}{x}\big)$. The latter must therefore also be a tempered distribution. Using the differentiation rules for the Fourier transform on the identity $x\mathrm{pv}\big(\frac{1}{x}\big) = 1$ that holds in $\mathscr{S}'(\mathbb{R})$ and results from Example 1.58 we find

$$2\pi\delta_0 = \widehat{\mathbf{1}_{\mathbb{R}}} = \widehat{x\mathrm{pv}(\tfrac{1}{x})} = \mathrm{i}\tfrac{\mathrm{d}}{\mathrm{d}\xi}\widehat{\mathrm{pv}(\tfrac{1}{x})},$$

hence

$$\widehat{\mathrm{pv}(\tfrac{1}{x})} = -2\pi \mathrm{i} H + c,$$

where $H = H(\xi)$ is Heaviside's function and $c \in \mathbb{C}$. In order to find the constant c we note that the distribution $\mathrm{pv}(\frac{1}{x})$ is odd (u is odd/even if $\widetilde{u} = -u/\widetilde{u} = u$) so

$$\begin{aligned} 0 = \widehat{\mathrm{pv}(\tfrac{1}{x})} + \widetilde{\widehat{\mathrm{pv}(\tfrac{1}{x})}} &= -2\pi \mathrm{i} H + c - 2\pi \mathrm{i} \widetilde{H} + c \\ &= -2\pi \mathrm{i} + 2c, \end{aligned}$$

so $c = \pi \mathrm{i}$ and $\widehat{\mathrm{pv}\big(\frac{1}{x}\big)} = \pi \mathrm{i} - 2\pi \mathrm{i} H$, or

$$\widehat{\mathrm{pv}\big(\tfrac{1}{x}\big)}(\xi) = -\mathrm{i}\pi \mathrm{sgn}(\xi).$$

Example 1.61. What is the Fourier transform of Heaviside's function H? We Fourier transform $H' = \delta_0$ by use of the differentiation rule to get $1 = \widehat{H'} = \mathrm{i}\xi \widehat{H}$. This is an inhomogeneous linear equation and we see that it has $\mathrm{pv}(\frac{1}{\mathrm{i}\xi})$ as a particular solution. The general solution to the homogeneous equation is found by use of a result from B4.3 that says any distribution with support in $\{0\}$ must be a linear combination of δ_0 and its derivatives. By inspection we check it can only be $c\delta_0$, $c \in \mathbb{C}$, so we infer that $\widehat{H} = -\mathrm{i}\mathrm{pv}(\frac{1}{\xi}) + c\delta_0$ for some $c \in \mathbb{C}$. Now $1 = H + \widetilde{H}$ so by Fourier transform we get

$$2\pi \delta_0 = \widehat{H} + \widetilde{\widehat{H}} = 2c\delta_0,$$

whereby $c = \pi$. Therefore

$$\widehat{H} = -\mathrm{i}\mathrm{pv}(\tfrac{1}{\xi}) + \pi \delta_0.$$

1.4.2 Multiplication by moderate C^∞ functions

You will have noticed that we only defined the product of a tempered distribution with a polynomial so far. The issue is of course that we cannot multiply with a general C^∞ function and we need to restrict to a subclass of these.

Definition 1.62. A function $f \colon \mathbb{R}^n \to \mathbb{C}$ is of *polynomial growth* if there exist constants $c \geqslant 0$, $m \in \mathbb{N}_0$ such that

$$|f(x)| \leqslant c\big(1 + |x|^2\big)^{\frac{m}{2}}$$

holds for all $x \in \mathbb{R}^n$.

Remark 1.63. Obviously f is of polynomial growth if and only if there exists a polynomial p on $\mathbb{R}^n$ such that $|f(x)| \leqslant |p(x)|$ for all x.

Example 1.64. The example with $u = \mathrm{e}^{|x|^2}$ showed us that we cannot expect to make sense of general $\mathrm{L}^p_{\mathrm{loc}}(\mathbb{R}^n)$ functions as tempered distributions. We concluded, heuristically, that to be a tempered distribution, a function cannot grow too fast at infinity. As remarked already, this is vague and it has to be. Indeed, consider the L^∞ function $u = \cos(\mathrm{e}^x)$ on $\mathbb{R}$. Clearly we may consider $u \in \mathscr{S}'(\mathbb{R})$ and hence also its distributional derivative $u' \in \mathscr{S}'(\mathbb{R})$. Now it is easy to see that $u' = -\sin(\mathrm{e}^x)\mathrm{e}^x$, where we must understand its action on $\phi \in \mathscr{S}(\mathbb{R}^n)$ as an improper integral:

$$\langle u', \phi\rangle = -\lim_{s\to\infty} \int_{-\infty}^{s} \sin(\mathrm{e}^x)\mathrm{e}^x \phi(x)\,\mathrm{d}x$$

since $u' \notin \mathrm{L}^1(\mathbb{R})$. The function u' is obviously not polynomially bounded.

Definition 1.65. A function $a\colon \mathbb{R}^n \to \mathbb{C}$ is said to be a *moderate* C^∞ *function* if $a \in \mathrm{C}^\infty(\mathbb{R}^n)$ and all its partial derivatives $\partial^\alpha a$, $\alpha \in \mathbb{N}_0^n$, have polynomial growth: for each multi-index $\alpha \in \mathbb{N}_0^n$ we can find constants $c_\alpha \geqslant 0$, $m_\alpha \in \mathbb{N}_0$ so

$$\big|(\partial^\alpha a)(x)\big| \leqslant c_\alpha\big(1 + |x|^2\big)^{\frac{m_\alpha}{2}} \tag{12}$$

holds for all $x \in \mathbb{R}^n$.

Example 1.66. Schwartz test functions $\phi \in \mathscr{S}(\mathbb{R}^n)$ and polynomials $p(x) \in \mathrm{C}[x]$ are clearly moderate C^∞ functions, as are $\cos(p(x))$, $\sin(p(x))$ and so on. The function $\mathrm{e}^{|x|^2}$ is C^∞ but is not moderate.

Proposition 1.67. *If a, $b\colon \mathbb{R}^n \to \mathbb{C}$ are moderate C^∞ functions and $\lambda \in \mathbb{C}$, then also $a + \lambda b$ and ab are moderate C^∞ functions. Furthermore, $\partial^\alpha a$ is a moderate C^∞ function for any multi-index α.*

Sketch of proof. That $a + \lambda b$ and $\partial^\alpha a$ are moderate C^∞ functions follow from the definition immediately. For the product, ab, one must use the generalized Leibniz rule. □

The key result for moderate C^∞ functions follows:

Proposition 1.68. *Let $a\colon \mathbb{R}^n \to \mathbb{C}$ be a moderate C^∞ function. Then for all k, $l \in \mathbb{N}_0$ we have that*

$$\overline{S}_{k,l}(a\phi) \leqslant 2^l c_l (n+1)^{m_l} \overline{S}_{k+m_l,l}(\phi)$$

holds for all $\phi \in \mathscr{S}(\mathbb{R}^n)$, where the constants $c_l := \max_{|\gamma|\leqslant l} c_\gamma$ and $m_l := \max_{|\gamma|\leqslant l} m_\gamma$ and c_γ, m_γ are the constants in the bound (12) for a.

It follows in particular that $\phi \mapsto a\phi$ is a linear and $\mathscr{S}$ continuous map from $\mathscr{S}(\mathbb{R}^n)$ to itself.

Proof. Fix α, $\beta \in \mathbb{N}_0^n$ with $|\alpha| \leqslant k$, $|\beta| \leqslant l$. Then for $\phi \in \mathscr{S}(\mathbb{R}^n)$ we compute using the Leibniz rule and estimating as usual:

$$\begin{aligned}
\left|x^{\alpha}\partial^{\beta}(a\varphi)\right| &= \left|\sum_{\gamma\leqslant\beta}\binom{\beta}{\gamma}(\partial^{\gamma}a)x^{\alpha}(\partial^{\beta-\gamma}\phi)\right| \\
&\leqslant \sum_{\gamma\leqslant\beta}\binom{\beta}{\gamma}\left|\partial^{\gamma}a\right|\left|x^{\alpha}(\partial^{\beta-\gamma}\phi)\right| \\
&\leqslant \sum_{\gamma\leqslant\beta}\binom{\beta}{\gamma}c_{\gamma}\big(1+|x|\big)^{m_{\gamma}}\left|x^{\alpha}(\partial^{\beta-\gamma}\phi)\right| \\
&\leqslant \sum_{\gamma\leqslant\beta}\binom{\beta}{\gamma}c_{\gamma}(n+1)^{m_{\gamma}-1}\big(1+\sum_{j=1}^{n}|x_j|^{m_{\gamma}}\big)\left|x^{\alpha}(\partial^{\beta-\gamma}\phi)\right| \\
&\leqslant c_l(n+1)^{m_l-1}\sum_{\gamma\leqslant\beta}\binom{\beta}{\gamma}\big(\overline{S}_{k,l}(\phi)+n\overline{S}_{k+m_l,l}(\phi)\big) \\
&\leqslant c_l(n+1)^{m_l}2^{|\beta|}\overline{S}_{k+m_l,l}(\phi) \\
&\leqslant c_l(n+1)^{m_l}2^{l}\overline{S}_{k+m_l,l}(\phi).
\end{aligned}$$

The remaining assertions all follow easily from this bound. □

In view of this proposition we may define the product of a tempered distribution u with a moderate C^{∞} function a: if $u \in \mathscr{S}'(\mathbb{R}^n)$ and $a\colon \mathbb{R}^n \to \mathbb{C}$ is a moderate C^{∞} function, then

$$\langle au,\, \varphi\rangle := \langle u,\, a\varphi\rangle, \quad \varphi \in \mathscr{S}(\mathbb{R}^n).$$

Clearly, $au \in \mathscr{S}'(\mathbb{R}^n)$ and with the obvious definition of ua we clearly have $ua = au$. Furthermore, the map

$$\mathscr{S}'(\mathbb{R}^n) \ni u \longmapsto au \in \mathscr{S}'(\mathbb{R}^n)$$

is linear and $\mathscr{S}'$ continuous. It is also easy to check that the Leibniz rule remains valid in the present context:

$$\partial_j\big(au\big) = (\partial_j a)u + a\partial_j u \quad (1 \leqslant j \leqslant n).$$

1.4.3 Mollification and approximation

We defined the tempered distribution $u * \theta = \theta * u$ by the adjoint identity scheme when $u \in \mathscr{S}'(\mathbb{R}^n)$ and $\theta \in \mathscr{S}(\mathbb{R}^n)$.

Proposition 1.69. ***[Convolution with Schwartz test function.]*** *If $u \in \mathscr{S}'(\mathbb{R}^n)$ and $\theta \in \mathscr{S}(\mathbb{R}^n)$, then $u * \theta$ is a moderate C^{∞} function and*

$$(u*\theta)(x) = \big\langle u, \theta(x-\cdot)\big\rangle, \quad x \in \mathbb{R}^n.$$

Furthermore we have for each $\alpha \in \mathbb{N}_0^n$ that

$$\partial^{\alpha}\big(u*\theta\big) = \big(\partial^{\alpha}u\big)*\theta = u*\big(\partial^{\alpha}\theta\big). \tag{13}$$

Proof. The proof that $u * \theta \in \mathrm{C}^\infty(\mathbb{R}^n)$, $(u * \theta)(x) = \langle u, \theta(x - \cdot)\rangle$ and (13) is identical to the proof given for a similar result in B4.3 and so we skip the details here. It remains to show that $u * \theta$ is moderate. In view of (13) it suffices to show that it has polynomial growth. In order to do that we invoke the boundedness property of u: according to Proposition 1.51 there exist constants $c \geqslant 0$, k, $l \in \mathbb{N}_0$ so

$$\left|\langle u, \phi\rangle\right| \leqslant c\overline{S}_{k,l}(\phi)$$

holds for all $\phi \in \mathscr{S}(\mathbb{R}^n)$. We apply this with $\phi = \theta(x - \cdot)$, whereby $\left|\langle u, \theta(x - \cdot)\rangle\right| \leqslant c\overline{S}_{k,l}(\theta(x - \cdot))$ results for each $x \in \mathbb{R}^n$. Now for multi-indices $|\alpha| \leqslant k$, $|\beta| \leqslant l$ we estimate for x, $y \in \mathbb{R}^n$:

$$\begin{aligned}
\left|y^\alpha \partial_y^\beta \theta(x - y)\right| &= \left|y^\alpha (\partial^\beta \theta)(x - y)\right| \\
&= \left|(y - x + x)^\alpha (\partial^\beta \theta)(x - y)\right| \\
&\leqslant \sum_{\gamma \leqslant \alpha} \binom{\alpha}{\gamma} \left|(x - y)^\gamma (\partial^\beta \theta)(x - y)\right| |x^{\alpha - \gamma}| \\
&\leqslant \sum_{\gamma \leqslant \alpha} \binom{\alpha}{\gamma} S_{\gamma,\beta}(\theta) |x^{\alpha - \gamma}| \\
&\leqslant \overline{S}_{k,l}(\theta) \sum_{\gamma \leqslant \alpha} \binom{\alpha}{\gamma} |x^{\alpha - \gamma}| \\
&= \overline{S}_{k,l}(\theta) \prod_{j=1}^n \left(1 + |x_j|\right)^{\alpha_j} \\
&\leqslant \overline{S}_{k,l}(\theta) \left(1 + |x|\right)^k.
\end{aligned}$$

Consequently we have for all $x \in \mathbb{R}^n$,

$$\left|\langle u, \theta(x - \cdot)\rangle\right| \leqslant c\overline{S}_{k,l}(\theta)\left(1 + |x|\right)^k,$$

concluding the proof. (Note that $\left(1 + |x|^2\right)^{\frac{1}{2}} \leqslant 1 + |x| \leqslant \sqrt{2}\left(1 + |x|^2\right)^{\frac{1}{2}}$ holds for all x, so it is not important that we stated the polynomial growth in terms of $\left(1 + |x|^2\right)^{\frac{1}{2}}$ rather than $1 + |x|$.) ☐

We turn to the issue of approximation.

Proposition 1.70. ***[Mollification of tempered distributions.]*** *Let $u \in \mathscr{S}'(\mathbb{R}^n)$ and let $\left(\rho_\varepsilon\right)_{\varepsilon > 0}$ be the standard mollifier on $\mathbb{R}^n$. Then*

$$u * \rho_\varepsilon \to u \ \text{ in } \ \mathscr{S}'(\mathbb{R}^n) \ \text{ as } \ \varepsilon \searrow 0.$$

Furthermore, we can find $u_\varepsilon \in \mathscr{D}(\mathbb{R}^n)$ (note: compactly supported test functions), so

$$u_\varepsilon \to u \ \text{ in } \ \mathscr{S}'(\mathbb{R}^n) \ \text{ as } \ \varepsilon \searrow 0.$$

The proof is similar to a proof in the $\mathscr{D}$ setting and one should use Example 1.30. We leave the details as an exercise.

1.5 The Fourier transform on L^2

1.5.1 Plancherel's theorem

Theorem 1.71. ***[Plancherel's Theorem]*** *The Fourier transform $\mathcal{F}\colon \mathrm{L}^2(\mathbb{R}^n) \to \mathrm{L}^2(\mathbb{R}^n)$ is bijective, and $(2\pi)^{-n/2}\mathcal{F}$ is* unitary *(isometric and onto). That is, $\mathcal{F}(\mathrm{L}^2) = \mathrm{L}^2$ and*

$$\|\hat{f}\|_2 = (2\pi)^{n/2}\|f\|_2 \tag{14}$$

for $f \in \mathrm{L}^2(\mathbb{R}^n)$, and more generally

$$\int_{\mathbb{R}^n} f(x)\overline{g(x)}\,\mathrm{d}x = (2\pi)^{-n}\int_{\mathbb{R}^n} \hat{f}(\xi)\overline{\hat{g}(\xi)}\,\mathrm{d}\xi. \tag{15}$$

for $f,\, g \in \mathrm{L}^2(\mathbb{R}^n)$. Furthermore, the Fourier transform $\widehat{f}$ is given by

$$\widehat{f}(\xi) = \lim_{j\to\infty}\int_{B_j(0)} f(x)\mathrm{e}^{-\mathrm{i}\xi\cdot x}\,\mathrm{d}x \tag{16}$$

with convergence in $\mathrm{L}^2(\mathbb{R}^n)$.

The two identities (14), (15) are often called *Parseval's formulae.*

Remark 1.72. It is important to realize that when $f \in \mathrm{L}^2(\mathbb{R}^n)$, then $x \mapsto f(x)\mathrm{e}^{-\mathrm{i}\xi\cdot x}$ need not be integrable on $\mathbb{R}^n$ so that the Fourier transform $\widehat{f}$ must be defined as for a tempered distribution (that is by the adjoint identity scheme as in Definition 1.55). This also means that the convergence in (16) takes place in $\mathrm{L}^2(\mathbb{R}^n)$ and we emphasize that this does *not* imply convergence pointwise almost everywhere in ξ. However, by a standard consequence of L^2 convergence (how?), there exists a subsequence (j_k) of (j) (that will depend on f in general) such that

$$\widehat{f}(\xi) = \lim_{k\to\infty}\int_{B_{j_k}(0)} f(x)\mathrm{e}^{-\mathrm{i}\xi\cdot x}\,\mathrm{d}x$$

holds pointwise almost everywhere in $\xi \in \mathbb{R}^n$.

Proof. We start by observing that for $\varphi,\, \psi \in \mathscr{S}(\mathbb{R}^n)$

$$\int_{\mathbb{R}^n} \varphi\bar{\psi}\,\mathrm{d}x = (2\pi)^{-n}\int_{\mathbb{R}^n} \hat{\varphi}\bar{\hat{\psi}}\,\mathrm{d}\xi,$$

and in particular for $\varphi = \psi$

$$\int_{\mathbb{R}^n} |\varphi|^2\,\mathrm{d}x = (2\pi)^{-n}\int_{\mathbb{R}^n} |\hat{\varphi}|^2\,\mathrm{d}\xi. \tag{17}$$

This follows from the product rule and the Fourier Inversion Formula on $\mathscr{S}(\mathbb{R}^n)$: clearly $\bar{\psi} \in \mathscr{S}(\mathbb{R}^n)$, so $\mathcal{F}^{-1}(\bar{\psi}) \in \mathscr{S}(\mathbb{R}^n)$ and so

$$\int_{\mathbb{R}^n} \varphi\bar{\psi}\,\mathrm{d}x = \int_{\mathbb{R}^n} \varphi\mathcal{F}(\mathcal{F}^{-1}\bar{\psi})\,\mathrm{d}x = \int_{\mathbb{R}^n} \hat{\varphi}\mathcal{F}^{-1}(\bar{\psi})\,\mathrm{d}x.$$

Now

$$\mathcal{F}^{-1}(\bar{\psi})(x) = (2\pi)^{-n}\int_{\mathbb{R}^n}\bar{\psi}(y)\mathrm{e}^{ix\cdot y}\,\mathrm{d}y = (2\pi)^{-n}\overline{\int_{\mathbb{R}^n}\psi(y)\mathrm{e}^{-ix\cdot y}\,\mathrm{d}y} = (2\pi)^{-n}\overline{\hat{\psi}(x)}.$$

If now $f \in \mathrm{L}^2(\mathbb{R}^n)$ we know that there exist $f_j \in \mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$ so $\|f - f_j\|_2 \to 0$. Clearly this means in particular that $f_j \to f$ in $\mathscr{S}'(\mathbb{R}^n)$, and thus by $\mathscr{S}'$-continuity of the Fourier transform, $\hat{f}_j \to \hat{f}$ in $\mathscr{S}'(\mathbb{R}^n)$. By (17) we see that

$$\int_{\mathbb{R}^n}|\hat{f}_j - \hat{f}_k|^2\,\mathrm{d}\xi = (2\pi)^n\int_{\mathbb{R}^n}|f_j - f_k|^2\,\mathrm{d}x,$$

so $(\hat{f}_j)$ is Cauchy in L^2. It is thus convergent in L^2 by the Riesz–Fischer theorem (see below) $\hat{f}_j \to g$ in $\mathrm{L}^2(\mathbb{R}^n)$ for some $g \in \mathrm{L}^2(\mathbb{R}^n)$. Clearly then $\hat{f}_j \to g$ in $\mathscr{S}'(\mathbb{R}^n)$ too, and so $g = \hat{f}$. Finally we note that $f\mathbf{1}_{B_j(0)} \to f$ in $\mathrm{L}^2(\mathbb{R}^n)$ so the Fourier transform can be found as an L^2 limit as asserted. □

Theorem 1.73. ***[The Riesz–Fischer theorem]*** *Let $p \in [1,\infty]$. Then $\mathrm{L}^p(\mathbb{R}^n)$ with the norm $\|\cdot\|_p$ is complete: if (f_j) is a Cauchy sequence in $\mathrm{L}^p(\mathbb{R}^n)$, then there exists $f \in \mathrm{L}^p(\mathbb{R}^n)$ so $\|f - f_j\|_p \to 0$.*

So far we have shown that $\mathcal{F}\colon \mathrm{L}^1(\mathbb{R}^n) \to \mathrm{C}_0(\mathbb{R}^n) \subset \mathrm{L}^\infty(\mathbb{R}^n)$ and $\mathcal{F}\colon \mathrm{L}^2(\mathbb{R}^n) \to \mathrm{L}^2(\mathbb{R}^n)$ as continuous linear maps. What happens on the other L^p spaces? Note that if $f \in \mathrm{L}^p(\mathbb{R}^n)$ for some $p \in (1,2)$, then we may write $f = f_1 + f_2$, where

$$f_1 = \begin{cases} f & \text{if } |f| \geqslant 1, \\ 0 & \text{if } |f| < 1, \end{cases} \qquad \text{and} \qquad f_2 = \begin{cases} 0 & \text{if } |f| \geqslant 1, \\ f & \text{if } |f| < 1, \end{cases}$$

and since $\|f_1\|_1 \leqslant \|f\|_p$, $\|f_2\|_2 \leqslant \|f\|_p$ it follows that $\widehat{f} = \widehat{f_1} + \widehat{f_2} \in \mathrm{C}_0(\mathbb{R}^n) + \mathrm{L}^2(\mathbb{R}^n)$. But in fact a much more precise result holds true:

Theorem 1.74. ***[Hausdorff–Young]*** *For $p \in (1,2)$ and $\frac{1}{p} + \frac{1}{q} = 1$ we have for $f \in \mathrm{L}^p(\mathbb{R}^n)$ that $\hat{f} \in \mathrm{L}^q(\mathbb{R}^n)$ with*

$$\|\hat{f}\|_q \leqslant (2\pi)^{n/q}\|f\|_p.$$

We will not prove the result here and merely remark that the Fourier transform on Lebesgue spaces is onto a Lebesgue space only when $p = 2$. For $p > 2$ it can be shown that the image $\mathcal{F}(\mathrm{L}^p(\mathbb{R}^n))$ contains tempered distributions of positive orders.

1.5.2 L^2 based Sobolev spaces

A benefit of distribution theory is that it allows us to find solutions to problems that have no classical solution. But we often want our solutions to be regular distributions and preferably as nice as possible. The theory of Sobolev spaces can often be used quite efficiently for that.

Proposition 1.75. ***[An*** L^2 ***identity for the Laplacian]***
Let $u \in \mathrm{L}^2(\mathbb{R}^n)$ *and assume* $\Delta u \in \mathrm{L}^2(\mathbb{R}^n)$*. Then* $\partial_j \partial_k u \in \mathrm{L}^2(\mathbb{R}^n)$ *for all* $1 \leqslant j, k \leqslant n$*, and*

$$\sum_{j,k=1}^{n} \|\partial_j \partial_k u\|_2^2 = \|\Delta u\|_2^2. \tag{18}$$

Remark 1.76. If we only assume that $v \in \mathscr{S}'(\mathbb{R}^n)$ and $\Delta v \in \mathrm{L}^2(\mathbb{R}^n)$, then it is an exercise to show that for some harmonic polynomial $p \in \mathbb{C}[x]$ the conclusion of Proposition 1.75 applies to $u = v + p$.

Proof. First we note that according to the differentiation rule $\widehat{\Delta u} = -|\xi|^2 \widehat{u}$ and $\widehat{\partial_j \partial_k u} = -\xi_j \xi_k \widehat{u}$ hold in $\mathscr{S}'(\mathbb{R}^n)$. By Plancherel's theorem $\widehat{u}$, $\widehat{\Delta u} \in \mathrm{L}^2(\mathbb{R}^n)$, so

$$\widehat{\partial_j \partial_k u} = \frac{\xi_j \xi_k}{|\xi|^2}\big(-|\xi|^2 \widehat{u}\big) = \frac{\xi_j \xi_k}{|\xi|^2} \widehat{\Delta u} \in \mathrm{L}^2(\mathbb{R}^n),$$

and therefore by Plancherel's theorem again, $\partial_j \partial_k u \in \mathrm{L}^2(\mathbb{R}^n)$. Next, we use Parseval's formula:

$$\begin{aligned}
\sum_{j,k=1}^{n} \|\partial_j \partial_k u\|_2^2 =& (2\pi)^{-\frac{n}{2}} \sum_{j,k=1}^{n} \|\widehat{\partial_j \partial_k u}\|_2^2 \\
=& (2\pi)^{-\frac{n}{2}} \sum_{j,k=1}^{n} \int_{\mathbb{R}^n} \left(\frac{\xi_j \xi_k}{|\xi|^2}\right)^2 \big|\widehat{\Delta u}\big|^2 \,\mathrm{d}\xi \\
=& (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \big|\widehat{\Delta u}\big|^2 \,\mathrm{d}\xi \\
=& \|\Delta u\|_2^2,
\end{aligned}$$

where we used that $\sum_{j,k=1}^{n} \xi_j^2 \xi_k^2 = \big(\sum_{j=1}^{n} \xi_j^2\big)^2 = |\xi|^4$. □

Proposition 1.77. ***[An interpolation inequality]***
Assume u, $\Delta u \in \mathrm{L}^2(\mathbb{R}^n)$*. Then also* $\partial_j u \in \mathrm{L}^2(\mathbb{R}^n)$ *for each* $j \in \{1, \dots, n\}$ *and*

$$\sum_{j=1}^{n} \|\partial_j u\|_2^2 \leqslant \tfrac{n}{2}\|u\|_2^2 + \tfrac{1}{2}\|\Delta u\|_2^2.$$

Proof. Let $u_\varepsilon = \rho_\varepsilon * u$, where $\big(\rho_\varepsilon\big)_{\varepsilon>0}$ is the standard mollifier on $\mathbb{R}^n$. For ε', $\varepsilon'' > 0$ we put $\phi = u_{\varepsilon'} - u_{\varepsilon''}$ and record that $\phi \in \mathrm{C}^\infty(\mathbb{R}^n) \cap \mathscr{S}'(\mathbb{R}^n)$ and ϕ, $\Delta\phi \in \mathrm{L}^2(\mathbb{R}^n)$. The latter follows from $\Delta\phi = \big(\rho_{\varepsilon'} - \rho_{\varepsilon''}\big) * \Delta u$ and properties of mollification. Then by the differentiation rule $\widehat{\partial_j \phi} = \mathrm{i}\xi_j \widehat{\phi}$, and as $\widehat{\phi} \in \mathrm{L}^2(\mathbb{R}^n)$ by Plancherel's theorem, the distribution $\widehat{\partial_j \phi}$ is a tempered L^2

function and we can estimate

$$\begin{aligned}|\widehat{\partial_j\phi}| = |\xi_j||\widehat{\phi}| &\leqslant \tfrac{1+\xi_j^2}{2}|\widehat{\phi}| \\ &\leqslant \tfrac{1+|\xi|^2}{2}|\widehat{\phi}| \\ &= \tfrac{1}{2}\big(|\widehat{\phi}| + |\widehat{\Delta\phi}|\big).\end{aligned}$$

By Plancherel's theorem again we infer that the n functions $\partial_j\phi$ are L^2, hence by Parseval's formula

$$\begin{aligned}\sum_{j=1}^n \|\partial_j\phi\|_2^2 =&(2\pi)^{-\frac{n}{2}}\sum_{j=1}^n \|\widehat{\partial_j\phi}\|_2^2 \\ \leqslant&(2\pi)^{-\frac{n}{2}}\left(\frac{n}{2}\|\widehat{\phi}\|_2^2 + \tfrac{1}{2}\|\widehat{\Delta\phi}\|_2^2\right) \\ \leqslant&\tfrac{n}{2}\|\phi\|_2^2 + \tfrac{1}{2}\|\Delta\phi\|_2^2.\end{aligned}$$

Because u, $\Delta u \in \mathrm{L}^2(\mathbb{R}^n)$ the families (u_ε) and (Δu_ε) are both Cauchy in L^2 as $\varepsilon \searrow 0$, and the above estimate then implies that also the n families $(\partial_j u_\varepsilon)$ are Cauchy in L^2 as $\varepsilon \searrow 0$. But then by the Riesz–Fischer theorem there exist n functions $g_j \in \mathrm{L}^2(\mathbb{R}^n)$ so $\partial_j u_\varepsilon \to g_j$ in $\mathrm{L}^2(\mathbb{R}^n)$ as $\varepsilon \searrow 0$. Because $u_\varepsilon \to u$ in $\mathrm{L}^2(\mathbb{R}^n)$ also $u_\varepsilon \to u$ in $\mathscr{S}'(\mathbb{R}^n)$, and so by $\mathscr{S}'$ continuity of differentiation, $\partial_j u_\varepsilon \to \partial_j u$ in $\mathscr{S}'(\mathbb{R}^n)$. It follows that $g_j = \partial_j u$ and the interpolation inequality is then an easy consequence of the differentiation rule and Parseval's formula. □

Propositions 1.75 and 1.77 show that when u, $\Delta u \in \mathrm{L}^2(\mathbb{R}^n)$, then u is a $\mathrm{W}^{2,2}$ Sobolev function: $\partial^\alpha u \in \mathrm{L}^2(\mathbb{R}^n)$ for each multi-index $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leqslant 2$. The Plancherel theorem makes it possible to define the $\mathrm{W}^{k,2}$ spaces in terms of the Fourier transform:

$$\mathrm{W}^{k,2}(\mathbb{R}^n) = \left\{u \in \mathscr{S}'(\mathbb{R}^n) :\ \big(1+|\xi|^2\big)^{\frac{k}{2}}\widehat{u} \in \mathrm{L}^2(\mathbb{R}^n)\right\}.$$

The norm

$$\|u\|_{\mathrm{H}^k} := \Big\|\big(1+|\xi|^2\big)^{\frac{k}{2}}\widehat{u}\Big\|_2$$

is equivalent to the norm $\|\cdot\|_{\mathrm{W}^{k,2}}$ that we defined in B4.3 as

$$\|u\|_{\mathrm{W}^{k,2}} = \left(\sum_{|\alpha|\leqslant k}\|\partial^\alpha u\|_2^2\right)^{\frac{1}{2}}.$$

Both norms come from inner products, where we record that $\|u\|_{\mathrm{H}^k} = \sqrt{(u,u)_{\mathrm{H}^k}}$ and

$$(u,v)_{\mathrm{H}^k} := \int_{\mathbb{R}^n}\widehat{u}(\xi)\overline{\widehat{v}(\xi)}\big(1+|\xi|^2\big)^k\,\mathrm{d}\xi.$$

The fact that we have this equivalent definition in terms of the Fourier transform allows us to define a scale of L^2 based Sobolev spaces that includes any differentiation order $s \in \mathbb{R}$:

Definition 1.78. **[Sobolev spaces of real order** s**]** Let $s \in \mathbb{R}$. Then

$$\mathrm{H}^s(\mathbb{R}^n) := \left\{ u \in \mathscr{S}'(\mathbb{R}^n) : \left(1 + |\xi|^2\right)^{\frac{s}{2}} \widehat{u} \in \mathrm{L}^2(\mathbb{R}^n) \right\}$$

with inner product

$$(u, v)_{\mathrm{H}^s} := \int_{\mathbb{R}^n} \widehat{u}(\xi)\overline{\widehat{v}(\xi)}\left(1 + |\xi|^2\right)^s \,\mathrm{d}\xi$$

and corresponding norm $\|u\|_{\mathrm{H}^s} := \sqrt{(u, u)_{\mathrm{H}^s}}$.

$\mathrm{H}^s(\mathbb{R}^n)$ is an example of a Hilbert space (that is, a normed vector space where the norm is defined by an inner product and where the corresponding metric space is complete). The theory of Hilbert spaces is discussed in *Functional Analysis 1 & 2*.

We record that $\mathrm{H}^0(\mathbb{R}^n) = \mathrm{L}^2(\mathbb{R}^n)$ and that the scale is nested: when $s < t$ we have

$$\mathrm{H}^t(\mathbb{R}^n) < \mathrm{H}^s(\mathbb{R}^n).$$

The regularity of distributions in $\mathrm{H}^s(\mathbb{R}^n)$ increases with $s \in \mathbb{R}$ in the following precise sense:

Proposition 1.79. *Let* $u \in \mathrm{H}^s(\mathbb{R}^n)$ *for some* $s > \frac{n}{2}$. *Then* $u \in \mathrm{C}^k(\mathbb{R}^n)$ *for each* $k \in \mathbb{N}_0$ *with* $k < s - \frac{n}{2}$, *and in fact* $\partial^\alpha u \in \mathrm{C}_0(\mathbb{R}^n)$ *for each multi-index with* $|\alpha| < s - \frac{n}{2}$.

Proof. Let $\alpha \in \mathbb{N}_0^n$ satisfy $|\alpha| < s - \frac{n}{2}$. By the differentiation rule $\widehat{\partial^\alpha u} = (\mathrm{i}\xi)^\alpha \widehat{u}$, and since in particular $\widehat{u} \in \mathrm{L}^2(\mathbb{R}^n)$ and

$$\begin{aligned} |\xi^\alpha| = \prod_{j=1}^n |\xi_j|^{\alpha_j} &\leqslant \prod_{j=1}^n |\xi|^{\alpha_j} = |\xi|^{|\alpha|} \\ &\leqslant \left(1 + |\xi|^2\right)^{\frac{|\alpha|}{2}} \end{aligned}$$

we may estimate

$$\begin{aligned} |\xi^\alpha \widehat{u}| &\leqslant \left(1 + |\xi|^2\right)^{\frac{|\alpha|}{2}} |\widehat{u}| \\ &= \left(1 + |\xi|^2\right)^{\frac{|\alpha|-s}{2}} \left(1 + |\xi|^2\right)^{\frac{s}{2}} |\widehat{u}|. \end{aligned}$$

By the Cauchy–Schwarz inequality

$$\begin{aligned} \int_{\mathbb{R}^n} |\xi^\alpha \widehat{u}| \,\mathrm{d}\xi &\leqslant \left\| \left(1 + |\xi|^2\right)^{\frac{|\alpha|-s}{2}} \right\|_2 \left\| \left(1 + |\xi|^2\right)^{\frac{s}{2}} |\widehat{u}| \right\|_2 \\ &= \left\| \left(1 + |\xi|^2\right)^{\frac{|\alpha|-s}{2}} \right\|_2 \|u\|_{\mathrm{H}^s}. \end{aligned}$$

We check that the first factor is finite by calculating in polar coordinates:

$$\begin{aligned}\left\|\left(1+|\xi|^2\right)^{\frac{|\alpha|-s}{2}}\right\|_2^2 &= \int_{\mathbb{R}^n} \left(1+|\xi|^2\right)^{|\alpha|-s}\,\mathrm{d}\xi \\ &= \int_0^\infty \int_{\partial B_r(0)} \left(1+r^2\right)^{|\alpha|-s}\,\mathrm{d}S_\xi\,\mathrm{d}r \\ &= \omega_{n-1} \int_0^\infty \frac{r^{n-1}}{(1+r^2)^{s-|\alpha|}}\,\mathrm{d}r.\end{aligned}$$

Here the exponent $s-|\alpha| > \frac{n}{2}$ so the integral is finite, proving that $\widehat{\partial^\alpha u} \in \mathrm{L}^1(\mathbb{R}^n)$. By the Fourier Inversion Formula in $\mathscr{S}'(\mathbb{R}^n)$ we have $(2\pi)^n \widetilde{\partial^\alpha u} = \widehat{\widehat{\partial^\alpha u}}$ and the latter belongs to $\mathrm{C}_0(\mathbb{R}^n)$ by virtue of the Riemann–Lebesgue lemma. But then clearly also $\partial^\alpha u \in \mathrm{C}_0(\mathbb{R}^n)$ concluding the proof. □

1.6 The Fourier transform and convolutions

Proposition 1.80. *If $u \in \mathscr{S}'(\mathbb{R}^n)$ and $\theta \in \mathscr{S}(\mathbb{R}^n)$, then $\widehat{u*\theta} = \widehat{u}\widehat{\theta}$ and $\widehat{u\theta} = (2\pi)^{-n}\widehat{u}*\widehat{\theta}$.*

Proof. For $\phi \in \mathscr{S}(\mathbb{R}^n)$ the definitions give

$$\langle \widehat{u*\theta}, \phi\rangle = \langle u, \widetilde{\theta} * \widehat{\phi}\rangle.$$

We continue by use of the Fourier Inversion Formula on $\mathscr{S}(\mathbb{R}^n)$, then the convolution rule (Proposition 1.42) and then definitions:

$$\begin{aligned}\langle u, \widetilde{\theta} * \widehat{\phi}\rangle &= (2\pi)^{-n}\langle u, \widehat{\widehat{\theta}} * \widehat{\phi}\rangle \\ &= \langle u, \widehat{(\widehat{\theta}\phi)}\rangle \\ &= \langle \widehat{u}, \widehat{\theta}\phi\rangle \\ &= \langle \widehat{u}\widehat{\theta}, \phi\rangle.\end{aligned}$$

For the other identity we note that $\widehat{u} \in \mathscr{S}'(\mathbb{R}^n)$, $\widehat{\theta} \in \mathscr{S}(\mathbb{R}^n)$ so by the above result and the Fourier Inversion formulas:

$$\begin{aligned}\widehat{\widehat{u}*\widehat{\theta}} = \widehat{\widehat{u}}\widehat{\widehat{\theta}} &= (2\pi)^{2n}\widetilde{u}\widetilde{\theta} \\ &= (2\pi)^{2n}\widetilde{u\theta} \\ &= (2\pi)^n \widehat{\widehat{u\theta}},\end{aligned}$$

and the result follows from the Fourier Inversion formula in $\mathscr{S}'(\mathbb{R}^n)$. □

Before the general convolution rule we record the following result.

Proposition 1.81. *Let $u \in \mathscr{E}'(\mathbb{R}^n)$. Then $\widehat{u}$ is a moderate C^∞ function and*

$$\widehat{u}(\xi) = \langle u, \mathrm{e}^{-\mathrm{i}(\cdot)\cdot\xi} \rangle, \quad \xi \in \mathbb{R}^n.$$

Proof. Take $\theta \in \mathscr{D}(\mathbb{R}^n)$ so $\theta = 1$ near $\mathrm{supp}(u)$. Then $\theta u = u$ and by Proposition 1.80 we get

$$\widehat{u} = \widehat{\theta u} = (2\pi)^{-n} \widehat{u} * \widehat{\theta}.$$

It follows from Proposition 1.69 that $\widehat{u}$ is a moderate C^∞ function. We conclude the proof using the rules for the Fourier transform as follows:

$$\begin{aligned}
\widehat{u}(\xi) = (2\pi)^{-n}(\widehat{u} * \widehat{\theta})(\xi) &= (2\pi)^{-n}\langle \widehat{u}, \widehat{\theta}(\xi - \cdot)\rangle \\
&= (2\pi)^{-n}\langle \widehat{u}, \widetilde{\widehat{\theta}}(\cdot - \xi)\rangle \\
&= (2\pi)^{-2n}\langle \widehat{u}, \mathcal{F}^3\theta(\cdot - \xi)\rangle \\
&= (2\pi)^{-2n}\langle \widehat{u}, \tau_{-\xi}\mathcal{F}^3\theta(\cdot)\rangle \\
&= (2\pi)^{-2n}\langle u, \mathrm{e}^{-\mathrm{i}(\cdot)\cdot\xi}\mathcal{F}^4\theta\rangle \\
&= \langle u, \mathrm{e}^{-\mathrm{i}(\cdot)\cdot\xi}\theta\rangle \\
&= \langle u, \mathrm{e}^{-\mathrm{i}(\cdot)\cdot\xi}\rangle.
\end{aligned}$$

□

Theorem 1.82. ***[The general convolution rule]*** *Let $u \in \mathscr{S}'(\mathbb{R}^n)$ and $v \in \mathscr{E}'(\mathbb{R}^n)$. Then $u * v \in \mathscr{S}'(\mathbb{R}^n)$ and*

$$\widehat{u * v} = \widehat{u}\widehat{v}.$$

Proof. We recall from B4.3 that for $\phi \in \mathscr{D}(\mathbb{R}^n)$ we defined $\langle u * v, \phi\rangle = \langle u, \widetilde{v} * \phi\rangle$ and that it was shown that $\widetilde{v} * \phi \in \mathscr{D}(\mathbb{R}^n)$. Let (ϕ_j) be a sequence in $\mathscr{D}(\mathbb{R}^n)$ so $\phi_j \to 0$ in $\mathscr{S}(\mathbb{R}^n)$. We must show that $\widetilde{v} * \phi_j \to 0$ in $\mathscr{S}(\mathbb{R}^n)$. Fix multi-indices α, $\beta \in \mathbb{N}_0^n$.

Let K be a compact neighbourhood of $\mathrm{supp}(v)$. Then we can find constants $c \geqslant 0$, $m \in \mathbb{N}_0$ so

$$\big|\langle v, \psi\rangle\big| \leqslant c \sum_{|\gamma| \leqslant m} \sup_K |\partial^\gamma \psi|$$

for all $\psi \in \mathrm{C}^\infty(\mathbb{R}^n)$. With $\psi = \partial^\beta \phi_j(\cdot - x)$ we get:

$$\begin{aligned}
\big|\partial^\beta(\widetilde{v} * \phi_j)(x)\big| &= \big|\langle v, \partial^\beta \phi_j(\cdot - x)\rangle\big| \\
&\leqslant c \sum_{|\gamma| \leqslant m} \sup_{y \in K} |\partial^{\gamma+\beta}\phi_j(y - x)|.
\end{aligned}$$

Consequently we have for all $x \in \mathbb{R}^n$,

$$
\begin{aligned}
\left|x^\alpha \partial^\beta \big(\widetilde{v} * \phi_j\big)(x)\right| &\leqslant c \sum_{|\gamma|\leqslant m} \sup_{y\in K} |x^\alpha (\partial^{\gamma+\beta}\phi_j)(y-x)| \\
&\leqslant c \sum_{|\gamma|\leqslant m} \sup_{y\in K} \sum_{\sigma\leqslant\alpha} \binom{\alpha}{\sigma} \big|(y-x)^\sigma (\partial^{\gamma+\beta}\phi_j)(y-x)\big| |y^{\alpha-\sigma}| \\
&\leqslant c \sum_{|\gamma|\leqslant m} \sum_{\sigma\leqslant\alpha} \binom{\alpha}{\sigma} S_{\sigma,\gamma+\beta}(\phi_j) \sup_{y\in K} |y^{\alpha-\sigma}|,
\end{aligned}
$$

and thus $S_{\alpha,\beta}(\widetilde{v} * \phi_j) \to 0$. But then $u * v$ is $\mathscr{S}$ continuous on $\mathscr{D}(\mathbb{R}^n)$ and since the latter is $\mathscr{S}$ dense in $\mathscr{S}(\mathbb{R}^n)$ it follows that $u * v \in \mathscr{S}'(\mathbb{R}^n)$. Finally we compute its Fourier transform by use of the previously established rules: for $\phi \in \mathscr{S}(\mathbb{R}^n)$,

$$
\begin{aligned}
\langle \widehat{u * v}, \phi\rangle = \langle u * v, \widehat{\phi}\rangle &= \langle u, \widetilde{v} * \widehat{\phi}\rangle \\
&= (2\pi)^{-n} \langle u, \widehat{\widehat{v}} * \widehat{\phi}\rangle \\
&= \langle u, \widehat{\widehat{v}\phi}\rangle \\
&= \langle \widehat{u}, \widehat{v}\phi\rangle \\
&= \langle \widehat{u}\widehat{v}, \phi\rangle.
\end{aligned}
$$

□

Remark 1.83. **[An extension of the convolution product using Fourier transform]**

Inspired by the last result and its proof we record the following: if $u, v \in \mathscr{S}'(\mathbb{R}^n)$ and $\widehat{v}$ is a moderate C^∞ function, then we *define* the convolution $u * v$ by

$$
u * v := (2\pi)^{-n} \widetilde{\widehat{\widehat{u}\widehat{v}}}.
$$

It is clear that hereby $u * v \in \mathscr{S}'(\mathbb{R}^n)$ and by the generalized convolution rule it is an extension of the case when $v \in \mathscr{E}'(\mathbb{R}^n)$. Furthermore, with the obvious definition of $v * u$ we clearly have $u * v = v * u$, and using the rules for the Fourier transform we also see that $\partial^\alpha(u * v) = (\partial^\alpha u) * v = u * (\partial^\alpha v)$ holds for all $\alpha \in \mathbb{N}_0^n$.

Even if this notion of convolution looks general it can still be generalized, but we shall refrain from doing that here.

1.7 The Paley-Wiener theorem

When $u \in \mathscr{E}'(\mathbb{R}^n)$ we have seen that $\widehat{u}$ is a moderate C^∞ function and that

$$
\widehat{u}(\xi) = \langle u, \mathrm{e}^{-\mathrm{i}\xi\cdot(\cdot)}\rangle.
$$

Here we will show that $\widehat{u}$ admits an extension to $\mathbb{C}^n$ as a holomorphic function in the variables $\xi_1, \ldots, \xi_n$. This extension is often called the *Fourier–Laplace transform* of u. As a preparation for this, and in fact much more, we start with the case when u is a compactly supported test function. The result in this case is also of independent interest.

Theorem 1.84. ***[Paley–Wiener for test functions.]***
(1) If $\varphi \in \mathscr{D}(\mathbb{R}^n)$ and $\operatorname{supp}(\varphi) \subseteq \overline{B_R(0)}$, then

$$\widehat{\varphi}(\zeta) = \int_{\mathbb{R}^n} \varphi(x) \mathrm{e}^{-\mathrm{i}\zeta\cdot x} \, \mathrm{d}x$$

is defined for all $\zeta \in \mathbb{C}^n$ and is an entire function (in each of the variables $\zeta_1, \ldots, \zeta_n \in \mathbb{C}$ separately) with the property that for each $m \in \mathbb{N}$ there exists a constant $c_m \geqslant 0$ so

$$\bigl|\widehat{\varphi}(\zeta)\bigr| \leqslant c_m \bigl(1+|\zeta|\bigr)^{-m} \mathrm{e}^{R|\mathrm{Im}(\zeta)|} \tag{19}$$

holds for all $\zeta \in \mathbb{C}^n$. Here the imaginary part of ζ, $\mathrm{Im}(\zeta)$, is understood componentwise when $n > 1$.

(2) Conversely, if $\Phi \colon \mathbb{C}^n \to \mathbb{C}$ is an entire function with the boundedness property (19), then there exists $\varphi \in \mathscr{D}(\mathbb{R}^n)$ with support in $\overline{B_R(0)}$ so $\Phi = \widehat{\varphi}$.

Proof of (1). [**The proof is only examinable in the case $n = 1$.**] It is clear that $\widehat{\varphi}(\zeta)$ is well-defined for all $\zeta \in \mathbb{C}^n$, and by a standard use of Lebesgue's dominated convergence theorem it follows that this extension is C^1 and that it satisfies the Cauchy-Riemann equations in each of the variables $\zeta_j \in \mathbb{C}$. Therefore $\widehat{\varphi}$ is entire (in each variable ζ_j separately). If we write $\zeta = \xi + \mathrm{i}\eta$, then

$$\begin{aligned} \bigl|\widehat{\varphi}(\zeta)\bigr| &\leqslant \int_{B_R(0)} |\varphi(x)| \mathrm{e}^{x\cdot\eta} \, \mathrm{d}x \\ &\leqslant \|\varphi\|_1 \mathrm{e}^{R|\eta|} \end{aligned}$$

holds. For a multi-index $\alpha \in \mathbb{N}_0^n$ we have clearly $\partial^\alpha \varphi \in \mathscr{D}(\mathbb{R}^n)$ with support in $\overline{B_R(0)}$ so the above bound holds with $\partial^\alpha \varphi$ in place of φ too. Now integration by parts gives

$$\widehat{\bigl(\partial^\alpha \varphi\bigr)}(\zeta) = (i\zeta)^\alpha \widehat{\varphi}(\zeta)$$

for all $\zeta \in \mathbb{C}^n$, so

$$|\zeta^\alpha| \bigl|\widehat{\varphi}(\zeta)\bigr| \leqslant \|\partial^\alpha \varphi\|_1 \mathrm{e}^{R|\eta|}$$

holds. We combine this inequality with the elementary inequalities

$$\bigl(1+|\zeta|\bigr)^m \leqslant \bigl(1+|\zeta_1| + \ldots + |\zeta_n|\bigr)^m \leqslant (n+1)^{m-1}\bigl(1+|\zeta_1|^m + \ldots + |\zeta_n|^m\bigr)$$

that hold for $\zeta \in \mathbb{C}^n$. Hereby

$$\begin{aligned} \bigl(1+|\zeta|\bigr)^m |\widehat{\varphi}(\zeta)| &\leqslant (n+1)^{m-1}\left(1 + \sum_{j=1}^n |\zeta_j|^m\right) |\widehat{\varphi}(\zeta)| \\ &\leqslant (n+1)^{m-1}\left(\|\varphi\|_1 + \sum_{j=1}^n \|\partial_j^m \varphi\|_1\right) \mathrm{e}^{R|\eta|}. \end{aligned}$$

We conclude that (19) holds with

$$c_m = (n+1)^{m-1}\Big(\|\varphi\|_1 + \sum_{j=1}^n \|\partial_j^m \varphi\|_1\Big).$$

Proof of (2).[**The proof is only examinable in the case $n = 1$.**] We start by recalling that when $f\colon \mathbb{C} \to \mathbb{C}$ is entire, then we get from Cauchy's integral formula and the estimation lemma the *Cauchy inequalities* for the derivatives: $|f^{(k)}(z)| \leqslant k! \max_{|w-z|=1} |f(w)|$ for all $z \in \mathbb{C}$ and $k \in \mathbb{N}_0$. Put $\phi = \Phi|_{\mathbb{R}^n}$. Then $\phi \in \mathrm{C}^\infty(\mathbb{R}^n)$ and for a multi-index $\beta \in \mathbb{N}_0^n$ we have

$$\big(\partial^\beta \phi\big)(\xi) = \frac{\partial^{|\beta|}\Phi}{\partial \zeta^\beta}(\xi)$$

for $\xi \in \mathbb{R}^n$ (note that on the left-hand side the differentiation is with respect to the ξ-variables whereas it is with respect to the ζ-variables on the right-hand side). The right-hand side can be estimated using the Cauchy inequalities:

$$\left|\frac{\partial^{|\beta|}\Phi}{\partial \zeta^\beta}(\xi)\right| \leqslant \beta! \max_{|\zeta-\xi|=1} \big|\Phi(\zeta)\big|.$$

Let $\alpha \in \mathbb{N}_0^n$ be a multi-index and put $m = |\alpha|$. Then according to the boundedness property of Φ we can find a constant $c_m \geqslant 0$ so (19) holds for all $\zeta \in \mathbb{C}^n$. Since $\xi \in \mathbb{R}^n$ we have for $\zeta \in \mathbb{C}^n$ with $|\zeta - \xi| = 1$ that $|\mathrm{Im}(\zeta)| \leqslant 1$ and $|\zeta| \geqslant \big||\xi| - 1\big|$, so we get

$$\begin{aligned}\Big|\xi^\alpha \big(\partial^\beta \phi\big)(\xi)\Big| &\leqslant \beta! \mathrm{e}^R c_m |\xi^\alpha| \big(1 + \big||\xi|-1\big|\big)^{-m} \\ &\leqslant \beta! \mathrm{e}^R c_m \left(\frac{|\xi|}{\big(1+\big||\xi|-1\big|\big)}\right)^m \\ &\leqslant \beta! \mathrm{e}^R c_m,\end{aligned}$$

hence $S_{\alpha,\beta}(\phi) < \infty$. It follows that $\phi \in \mathscr{S}(\mathbb{R}^n)$, and that we therefore by virtue of the Fourier Inversion formula in $\mathscr{S}(\mathbb{R}^n)$ can find $\varphi \in \mathscr{S}(\mathbb{R}^n)$ with $\widehat{\varphi} = \phi$:

$$\begin{aligned}\varphi(x) =& (2\pi)^{-n} \int_{\mathbb{R}^n} \phi(\xi) \mathrm{e}^{\mathrm{i}x\cdot\xi}\,\mathrm{d}\xi \\ =& (2\pi)^{-n} \int_{\mathbb{R}^n} \Phi(\xi) \mathrm{e}^{\mathrm{i}x\cdot\xi}\,\mathrm{d}\xi.\end{aligned}$$

Since the function $\zeta \mapsto \Phi(\zeta)\mathrm{e}^{\mathrm{i}x\cdot\zeta}$ is entire we can use the Cauchy theorem to deform the integration contour, and because of the boundedness property (19) we can change it from $\mathbb{R}^n$ to $\mathbb{R}^n + \mathrm{i}\eta$ with any $\eta \in \mathbb{R}^n$. In particular, corresponding to $m = n+1$ we find a constant $c_{n+1} \geqslant 0$ so

$$\big|\Phi(\xi + \mathrm{i}\eta)\big| \leqslant c_{n+1}\big(1 + |\xi + \mathrm{i}\eta|\big)^{-n-1} \mathrm{e}^{R|\eta|} \leqslant c_{n+1}\big(1+|\xi|\big)^{-n-1} \mathrm{e}^{R|\eta|}.$$

Hereby we estimate

$$
\begin{aligned}
\big|\varphi(x)\big| &= \left|(2\pi)^{-n}\int_{\mathbb{R}^n}\Phi(\xi+\mathrm{i}\eta)\mathrm{e}^{\mathrm{i}x\cdot(\xi+\mathrm{i}\eta)}\,\mathrm{d}\xi\right| \\
&\leqslant (2\pi)^{-n}\int_{\mathbb{R}^n}\big|\Phi(\xi+\mathrm{i}\eta)\big|\,\mathrm{d}\xi\mathrm{e}^{-x\cdot\eta} \\
&\leqslant c_{n+1}(2\pi)^{-n}\int_{\mathbb{R}^n}\frac{\mathrm{d}\xi}{\big(1+|\xi|\big)^{n+1}}\mathrm{e}^{-x\cdot\eta+R|\eta|} \\
&=: C\mathrm{e}^{-x\cdot\eta+R|\eta|},
\end{aligned}
$$

defining the constant C in the last line. If we take $\eta = tx$ for a $t>0$ we get

$$
\big|\varphi(x)\big| \leqslant C\mathrm{e}^{-t|x|(|x|-R)}.
$$

Hence for $|x|>R$ taking $t\to\infty$ shows that $\varphi(x)=0$, so that necessarily φ is supported in $\overline{B_R(0)}$ as asserted. □

Example 1.85. Let $u\in\mathscr{E}'(\mathbb{R}^n)$ with $\operatorname{supp}(u)\subseteq\{0\}$. Then a result from B4.3 states that $u\in\operatorname{span}\big\{\partial^\alpha\delta_0 : \alpha\in\mathbb{N}_0^n\big\}$, so that u can be expressed as

$$
u = \sum_{|\alpha|\leqslant d} c_\alpha\partial^\alpha\delta_0,
$$

where $c_\alpha\in\mathbb{C}$ and $d\in\mathbb{N}_0$. Assuming that $c_\alpha\neq 0$ for some α with $|\alpha|=d$ the distribution u has order d. Now since $\widehat{\delta_0}=1$ we get by use of the differentiation rule:

$$
\widehat{u} = \sum_{|\alpha|\leqslant d} c_\alpha(\mathrm{i}\xi)^\alpha = p(\xi)
$$

with

$$
p(\xi) = \sum_{|\alpha|\leqslant d}\mathrm{i}^{|\alpha|}c_\alpha\xi^\alpha\in\mathbb{C}[\xi].
$$

Clearly we may extend $p(\xi)$ to $\mathbb{C}^n$ simply by replacing $\xi\in\mathbb{R}^n$ by $\zeta\in\mathbb{C}^n$: $\widehat{u}(\zeta)=p(\zeta)$ is an entire function. If we take $c:=\max_{|\alpha|\leqslant d}|c_\alpha|$, then we also have the bound

$$
\big|\widehat{u}(\zeta)\big| \leqslant c\big(1+|\zeta|\big)^d \tag{20}
$$

for all $\zeta\in\mathbb{C}^n$. Note that any entire function on $\mathbb{C}^n$ satisfying the bound (20) must be a polynomial of degree at most d by virtue of the Liouville theorem.

Theorem 1.86. ***[The Paley–Wiener theorem]***

(1) If u is a distribution of order m with support in $\overline{B_R(0)}$, then the Fourier–Laplace transform $\widehat{u}$ is an entire function satisfying for some constant $c\geqslant 0$,

$$
\big|\widehat{u}(\zeta)\big| \leqslant c\big(1+|\zeta|\big)^m\mathrm{e}^{R|\mathrm{Im}(\zeta)|} \tag{21}
$$

for all $\zeta \in \mathbb{C}^n$.

(2) Conversely, if $\Phi\colon \mathbb{C}^n \to \mathbb{C}$ is an entire function satisfying (21) for some $m \in \mathbb{N}_0$ and constant $c \geqslant 0$, then there exists $u \in \mathscr{E}'(\mathbb{R}^n)$ supported in $\overline{B_R(0)}$ with Fourier–Laplace transform $\widehat{u} = \Phi$.

Remark 1.87. While the results stated in Theorems 1.84 and 1.86 here are attributed to Paley and Wiener, which is customary, they are actually in the present form due to Laurent Schwartz. Paley and Wiener considered L^2 functions on $\mathbb{R}$ that vanish for negative reals and proved that their Fourier transforms extend to the upper half-plane as holomorphic functions satisfying a certain bound (constituting the analytic Hardy space $\mathcal{H}^2$).

Proof of (1). [**The proof is only examinable in the case $n = 1$.**] We know from the previous section that $\widehat{u}$ is a moderate C^∞ function and $\widehat{u}(\xi) = \langle u, \mathrm{e}^{-\mathrm{i}(\cdot)\cdot\xi}\rangle$ for $\xi \in \mathbb{R}^n$. The right-hand side is clearly well-defined also for complex arguments and we can *define*

$$\widehat{u}(\zeta) := \langle u, \mathrm{e}^{-\mathrm{i}(\cdot)\cdot\zeta}\rangle, \quad \zeta \in \mathbb{C}^n,$$

where we emphasize that u acts on the function

$$\mathbb{R}^n \ni x \mapsto \mathrm{e}^{-\mathrm{i}x\cdot\zeta} = \mathrm{e}^{-\mathrm{i}\left(x_1\zeta_1 + \ldots + x_n\zeta_n\right)}.$$

We refer to $\widehat{u}\colon \mathbb{C}^n \to \mathbb{C}$ as the *Fourier–Laplace transform* of u. If we write $\zeta = \xi + \mathrm{i}\eta$ with the understanding that ξ, $\eta \in \mathbb{R}^n$, then we can think of $\widehat{u}$ as a function of $(\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^n$ and we assert that it is C^1 and that its partial derivatives may be calculated by differentiation behind the distribution sign. Indeed, fix a direction $1 \leqslant k \leqslant n$ and let $\left(\mathbf{e}_j\right)_{j=1}^n$ be the standard basis for $\mathbb{R}^n$. Then for any multi-index $\alpha \in \mathbb{N}_0^n$ we have that

$$\partial_\eta^\alpha \frac{\mathrm{e}^{-\mathrm{i}x\cdot(\zeta + \mathrm{i}\mathbf{e}_k h)} - \mathrm{e}^{-\mathrm{i}x\cdot\zeta}}{h} \to \partial_\eta^{\alpha + \mathbf{e}_k}\left(\mathrm{e}^{-\mathrm{i}x\cdot\zeta}\right)$$

locally uniformly in $x \in \mathbb{R}^n$ as $h \to 0$. Using the boundedness property of the compactly supported distribution u we conclude that $\widehat{u}$ is differentiable with respect to η_k with $\partial_{\eta_k}\widehat{u}(\zeta) = \langle u, \partial_{\eta_k}\mathrm{e}^{-\mathrm{i}(\cdot)\cdot\zeta}\rangle$. The argument for the other variables is identical and it is not difficult to see that the partial derivatives are continuous, so that the function $\widehat{u}$ as asserted is C^1 on $\mathbb{R}^n \times \mathbb{R}^n$. We may now check the Cauchy–Riemann equation by differentiation behind the distribution sign:

$$\frac{\partial}{\partial \bar{\zeta}_k}\widehat{u}(\zeta) = \left\langle u, \frac{\partial}{\partial \bar{\zeta}_k}\mathrm{e}^{-\mathrm{i}(\cdot)\cdot\zeta}\right\rangle = 0.$$

Consequently the function $\zeta_k \mapsto \widehat{u}(\zeta)$ is holomorphic, so that $\widehat{u}$ is entire as a function in each of the variables $\zeta_1, \ldots, \zeta_n$ separately. It is customary to refer to such functions simply as entire functions on $\mathbb{C}^n$. In order to prove the bound (21) we invoke the boundedness property that follows from knowing that u is supported in $\overline{B_R(0)}$ and has order at most m. Indeed $\overline{B_{R+1}(0)}$ is a compact neighbourhood of $\overline{B_R(0)}$ so we may find a constant $c \geqslant 0$ such that

$$\left|\langle u, \phi\rangle\right| \leqslant c \sum_{|\alpha| \leqslant m} \sup_{|y| \leqslant R+1} \left|(\partial^\alpha \phi)(y)\right|$$

holds for all $\phi \in \mathrm{C}^\infty(\mathbb{R}^n)$. For $\varepsilon \in (0,1]$ to be specified below we put $\theta := \rho_\varepsilon * \mathbf{1}_{B_{R+\varepsilon}(0)}$, where ρ_ε is the standard mollifier on $\mathbb{R}^n$. Then $\theta = 1$ on $\overline{B_R(0)}$, $\theta = 0$ off $B_{R+2\varepsilon}(0)$ and

$$\left|\partial^\alpha \theta\right| = \left|\varepsilon^{-|\alpha|}\left(\partial^\alpha \rho\right)_\varepsilon * \mathbf{1}_{B_{R+\varepsilon}}\right| \leqslant \varepsilon^{-|\alpha|} \int_{\mathbb{R}^n} |\partial^\alpha \rho| \, \mathrm{d}x$$

for each $\alpha \in \mathbb{N}_0^n$. Using $u = \theta u$ we calculate

$$\begin{aligned} \left|\widehat{u}(\zeta)\right| = \left|\langle u, \mathrm{e}^{-\mathrm{i}(\cdot)\cdot\zeta}\rangle\right| &= \left|\langle u, \theta \mathrm{e}^{-\mathrm{i}(\cdot)\cdot\zeta}\rangle\right| \\ &\leqslant c \sum_{|\alpha| \leqslant m} \sup_{|y| \leqslant R+1} \left|\partial^\alpha\left(\theta(y)\mathrm{e}^{-\mathrm{i}y\cdot\zeta}\right)\right| \\ &\leqslant c \sum_{|\alpha| \leqslant m} \sup_{|y| \leqslant R+2\varepsilon} \sum_{\gamma \leqslant \alpha} \binom{\alpha}{\gamma} \varepsilon^{-|\gamma|} \|\partial^\gamma \rho\|_1 |\mathrm{e}^{-\mathrm{i}y\cdot\zeta}| |(-\mathrm{i}\zeta)^{\alpha-\gamma}| \\ &\leqslant c \sum_{|\alpha| \leqslant m} \sum_{\gamma \leqslant \alpha} \binom{\alpha}{\gamma} \varepsilon^{-|\gamma|} \|\partial^\gamma \rho\|_1 \mathrm{e}^{(R+2\varepsilon)|\eta|} |\zeta^{\alpha-\gamma}|. \end{aligned}$$

Put $c_m := \max_{|\alpha| \leqslant m} \|\partial^\alpha \rho\|_1$ and note that $|\zeta^{\alpha-\gamma}| \leqslant (1+|\zeta|)^{|\alpha|-|\gamma|}$ for multi-indices satisfying $\gamma \leqslant \alpha$, hence we have

$$\left|\widehat{u}(\zeta)\right| \leqslant c c_m \sum_{|\alpha| \leqslant m} \sum_{\gamma \leqslant \alpha} \binom{\alpha}{\gamma} \varepsilon^{-|\gamma|} \left(1+|\zeta|\right)^{|\alpha|-|\gamma|} \mathrm{e}^{(R+2\varepsilon)|\eta|}$$

and choosing $\varepsilon = 1/(1+|\zeta|)$ we find

$$\begin{aligned} \left|\widehat{u}(\zeta)\right| &\leqslant c c_m \sum_{|\alpha| \leqslant m} \sum_{\gamma \leqslant \alpha} \binom{\alpha}{\gamma} \left(1+|\zeta|\right)^{|\alpha|} \mathrm{e}^{R|\eta|+2} \\ &\leqslant C\left(1+|\zeta|\right)^m \mathrm{e}^{R|\eta|}, \end{aligned}$$

where $C := c c_m 2^m \mathrm{e}^2 \sum_{|\alpha| \leqslant m} 1$. □

Proof of (2). [**The proof is only examinable in the case** $n = 1$.] It follows in particular from (21) with $\zeta = \xi \in \mathbb{R}^n$ that $\phi = \Phi|_{\mathbb{R}^n}$ is a tempered L^∞ function, so defining $u := \mathcal{F}^{-1}\phi$ we have $u \in \mathscr{S}'(\mathbb{R}^n)$. Put $u_\varepsilon = \rho_\varepsilon * u$, where $\left(\rho_\varepsilon\right)_{\varepsilon>0}$ is the standard mollifier on $\mathbb{R}^n$. Then u_ε is a moderate C^∞ function and, by the convolution rule,

$$\widehat{u_\varepsilon} = \widehat{u} d_\varepsilon \widehat{\rho} = \phi d_\varepsilon \widehat{\rho}.$$

The last function can, by Theorem 1.84 and our hypothesis, be extended to $\mathbb{C}^n$ as a holomorphic function $\zeta \mapsto \Phi(\zeta)\widehat{\rho}(\varepsilon\zeta)$. For a given $N \in \mathbb{N}$ we can according to (19) find a constant $c_{m+N} \geqslant 0$ so

$$\begin{aligned} \left|\widehat{\rho}(\varepsilon\zeta)\right| &\leqslant c_{m+N}\left(1+|\varepsilon\zeta|\right)^{-m-N} \mathrm{e}^{|\mathrm{Im}(\varepsilon\zeta)|} \\ &\leqslant c_{m+N} \varepsilon^{-m-N}\left(1+|\zeta|\right)^{-m-N} \mathrm{e}^{\varepsilon|\mathrm{Im}(\zeta)|} \end{aligned}$$

for all $\zeta \in \mathbb{C}^n$ provided $\varepsilon \in (0,1)$. In combination with (21) we find that

$$\left|\Phi(\zeta)\widehat{\rho}(\varepsilon\zeta)\right| \leqslant c c_{m+N} \varepsilon^{-m-N}\big(1+|\zeta|\big)^{-N} \mathrm{e}^{(R+\varepsilon)|\mathrm{Im}(\zeta)|}$$

for all $\zeta \in \mathbb{C}^n$. Consequently, for each fixed $\varepsilon \in (0,1)$, the Paley-Wiener theorem for test functions yields $\varphi \in \mathscr{D}(\mathbb{R}^n)$ supported in $\overline{B_{R+\varepsilon}(0)}$ so its Fourier-Laplace transform $\widehat{\varphi} = \Phi d_\varepsilon \widehat{\rho}$. In particular, its Fourier transform must equal $\widehat{u_\varepsilon}$, so that by the Fourier Inversion formula we have $\varphi = u_\varepsilon$. Now since $u_\varepsilon \to u$ in $\mathscr{S}'(\mathbb{R}^n)$ as $\varepsilon \searrow 0$ the conclusion follows easily. □

Corollary 1.88. *Let u, $v \in \mathscr{E}'(\mathbb{R}^n)$ and assume that $u * v = 0$. Then $u = 0$ or $v = 0$.*

The proof follows by Fourier-Laplace transforming $u*v = 0$ and then using the identity theorem for holomorphic functions. It is not difficult to see that the result fails if one of the distributions is not of compact support.

Example 1.89. Recall that a distribution $E \in \mathscr{D}'(\mathbb{R}^n)$ is called a fundamental solution to the differential operator $p(\partial)$ provided $p(\partial)E = \delta_0$, where $p(x) \in \mathbb{C}[x]$ is a polynomial in n variables. Assume the polynomial $p(x)$ has nonzero degree. Could there exist a *compactly supported* fundamental solution to $p(\partial)$? Assume $E \in \mathscr{E}'(\mathbb{R}^n)$ and that $p(\partial)E = \delta_0$. Since $\mathscr{E}'(\mathbb{R}^n) < \mathscr{S}'(\mathbb{R}^n)$ we can then Fourier transform to get by the differentiation rule:

$$p(\mathrm{i}\xi)\widehat{E}(\xi) = 1.$$

By the Paley-Wiener theorem $\widehat{E}$ has an entire extension to $\mathbb{C}^n$ and clearly so does any polynomial, so we must have the above identity for all $\xi = \zeta \in \mathbb{C}^n$. But when the polynomial $p(\mathrm{i}\zeta)$ has nonzero degree it has complex roots and then $\widehat{E}$ cannot be entire. The contradiction shows that a partial differential operator with constant coefficients and of order at least one can never have a compactly supported fundamental solution.

Not examinable:

Theorem 1.90. *Let $p(x) \in \mathbb{C}[x]$ be a polynomial of degree $d \in \mathbb{N}$. If $f \in \mathscr{E}'(\mathbb{R}^n)$, then the PDE*

$$p(\partial)u = f$$

admits a solution $u \in \mathscr{E}'(\mathbb{R}^n)$ if and only if $\zeta \mapsto \widehat{f}(\zeta)/p(\mathrm{i}\zeta)$ can be extended to an entire function on $\mathbb{C}^n$. The solution is then unique.

Proof. The *only if* direction is easy: when $u \in \mathscr{E}'(\mathbb{R}^n)$ is a solution, then we get by Fourier transform of the PDE that $p(\mathrm{i}\xi)\widehat{u} = \widehat{f}$. Since both $\widehat{u}$, $\widehat{f}$ admit entire extensions we can also extend the Fourier transformed PDE to all of $\mathbb{C}^n$:

$$p(\mathrm{i}\zeta)\widehat{u}(\zeta) = \widehat{f}(\zeta)$$

holds for all $\zeta \in \mathbb{C}^n$. It follows that $\zeta \mapsto \widehat{f}(\zeta)/p(\mathrm{i}\zeta)$ admits an entire extension.

The uniqueness part is also easy as it follows from the identity theorem for holomorphic functions and the Fourier Inversion formula. We turn to the *if* direction where we assume that $g\colon \mathbb{C}^n \to \mathbb{C}$ is the entire extension of $\widehat{f}(\zeta)/p(\mathrm{i}\zeta)$. The proof relies on the following

Auxiliary Lemma: Let $h\colon \mathbb{C} \to \mathbb{C}$ be entire and $p(z) \in \mathbb{C}[z]$ a polynomial of degree $m \in \mathbb{N}$ with leading coefficient $a \in \mathbb{C} \setminus \{0\}$. Then

$$\left|ah(0)\right| \leqslant \max_{|z|=1}\left|p(z)h(z)\right|.$$

Proof of Auxiliary Lemma. Let $\bar{p}(z)$ denote the polynomial obtained from $p(z)$ by conjugating its coefficients. Put $q(z) = z^m \bar{p}(\frac{1}{z})$ so that $q(z)$ is a polynomial of degree m with $q(0) = \bar{a}$. Now applying the maximum modulus principle to the holomorphic function $z \mapsto q(z)h(z)$ we get

$$\left|ah(0)\right| = \left|q(0)h(0)\right| \leqslant \max_{|z|=1}\left|q(z)h(z)\right| = \max_{|z|=1}\left|p(z)h(z)\right|,$$

which is the assertion. □

We *assume* that the coefficient a of ζ_1^d in $p(\mathrm{i}\zeta)$ is not zero, where we recall that d was the overall degree of $p(\zeta)$. (That this can be achieved by a linear change of variables will not be discussed here.) The point of this assumption is that then the coefficient a must be a constant and independent of the remaining variables in ζ. If we apply the Auxiliary Lemma with $h(z) := g(\zeta_1 + z, \zeta_2, \ldots, \zeta_n)$ and the polynomial $z \mapsto p(\mathrm{i}(\zeta_1 + z), \mathrm{i}\zeta_2, \ldots, \mathrm{i}\zeta_n)$ of degree d and with leading coefficient $a \neq 0$, then we get

$$\begin{aligned} |a||g(\zeta)| &\leqslant \sup_{|z|=1} \left|p(\mathrm{i}(\zeta_1 + z), \mathrm{i}\zeta_2, \ldots, \mathrm{i}\zeta_n) g(\zeta_1 + z, \zeta_2, \ldots, \zeta_n)\right| \\ &= \sup_{|z|=1} \left|\widehat{f}(\zeta_1 + z, \zeta_2, \ldots, \zeta_n)\right|. \end{aligned}$$

Because $\widehat{f}$ satisfies an estimate of form (21) also g will satisfy such an estimate. But then the Paley-Wiener theorem yields $u \in \mathscr{E}'(\mathbb{R}^n)$ whose Fourier-Laplace transform $\widehat{u} = g$. Clearly u is the sought solution. □

Another application of the Paley–Wiener theorem is to the proof of

Theorem 1.91. ***[The Ehrenpreis–Malgrange theorem]*** *Let $p(x) \in \mathbb{C}[x]$ be a polynomial of n variables that is not identically 0. Then the corresponding partial differential operator $p(\partial)$ admits a fundamental solution $E \in \mathscr{D}'(\mathbb{R}^n)$: $p(\partial)E = \delta_0$.*

We omit the proof here.

Example 1.92. Let $p(x) \in \mathbb{C}[x]$ be a polynomial of n variables that is not identically 0 and assume Ω is a bounded open subset of $\mathbb{R}^n$. If $f \in \mathrm{C}^\infty(\mathbb{R}^n)$, then we assert that the PDE

$$p(\partial)u = f \quad \text{ in } \Omega,$$

admits a solution $u \in \mathrm{C}^\infty(\mathbb{R}^n)$. Indeed, by virtue of the Ehrenpreis-Malgrange theorem we can find $E \in \mathscr{D}'(\mathbb{R}^n)$ so $p(\partial)E = \delta_0$. Put $\chi := \rho * \mathbf{1}_{B(\Omega,1)}$ so that $\chi \in \mathscr{D}(\mathbb{R}^n)$ satisfies $\chi = 1$ on Ω (here ρ is the standard mollifier kernel on $\mathbb{R}^n$). Note $\chi f \in \mathscr{D}(\mathbb{R}^n)$ so that $u := E * (\chi f) \in \mathrm{C}^\infty(\mathbb{R}^n)$ satisfies the PDE:

$$p(\partial)u = \big(p(\partial)E\big) * \big(\chi f\big) = \delta_0 * \big(\chi f\big) = \chi f,$$

and since $\chi f = f$ on Ω the assertion follows.

Theorem 1.93. ***[Qualitative version of the uncertainty principle]***
If $u \in \mathscr{E}'(\mathbb{R}^n)$ and $\widehat{u} \in \mathscr{E}'(\mathbb{R}^n)$, then $u = 0$.

Proof. The Fourier-Laplace transform $\widehat{u}$ is an entire function by the Paley-Wiener theorem. Since the support of the Fourier transform $\widehat{u}$ is compact too we can find $r > 0$ so $\widehat{u}(\xi) = 0$ for all $\xi \in \mathbb{R}^n$ with $\max\big\{|\xi_1|, \ldots, |\xi_n|\big\} \geqslant r$. Write $\zeta = (\zeta_1, \zeta') \in \mathbb{C} \times \mathbb{C}^{n-1}$ and fix $\xi_0' \in \mathbb{R}^{n-1}$. Now the function $\zeta_1 \mapsto \widehat{u}(\zeta_1, \xi_0')$ is entire and

$$(-\infty, -r] \cup [r, \infty) \subset \big\{\zeta_1 \in \mathbb{C} : \widehat{u}(\zeta_1, \xi_0') = 0\big\},$$

so by the identity theorem we must have $\widehat{u}(\zeta_1, \xi_0') = 0$ for all $\zeta_1 \in \mathbb{C}$. Since $\xi_0' \in \mathbb{R}^{n-1}$ was arbitrary we have shown that $\widehat{u} = 0$ on $\mathbb{R}^n$, and so, by the Fourier inversion formula on $\mathscr{S}'$, $u = 0$ as asserted. □

1.8 The Heisenberg uncertainty principle

There are many manifestations of the uncertainty principle that all express that it is impossible for both u and $\widehat{u}$ to be concentrated on small sets. The most famous quantitative uncertainty principle is due to Heisenberg, who formulated it in the context of quantum mechanics. In our context, it says, heuristically speaking, that if a function is concentrated on a ball of radius r, then its Fourier transform cannot be concentrated on a scale much smaller than $1/r$. More precisely:

Theorem 1.94. ***[Heisenberg's inequality]*** *If $f \in \mathscr{S}(\mathbb{R}^n)$ and $x_0 \in \mathbb{R}^n$, $\xi_0 \in \mathbb{R}^n$, then*

$$(2\pi)^{\frac{n}{2}} \tfrac{n}{2} \|f\|_2^2 \leqslant \|(x - x_0)f\|_2 \|(\xi - \xi_0)\widehat{f}\|_2.$$

The equality holds if and only if f is a modulated Gaussian:

$$f(x) = c \mathrm{e}^{\mathrm{i}\xi_0 \cdot x} \mathrm{e}^{-\varepsilon(x - x_0)^2},$$

where $c \in \mathbb{C}$ and $\varepsilon > 0$ are arbitrary.

Proof. By virtue of the translation rules we can also assume that $x_0 = \xi_0 = 0$ (apply the basic case of the inequality to the function $\mathrm{e}^{-\mathrm{i}(\cdot)\cdot\xi_0}\tau_{x_0} f$ to obtain the inequality in the general case). By the differentiation rule we have $\mathrm{i}\xi\widehat{f}(\xi) = \widehat{\nabla f}(\xi)$ so using first Parseval's identity, then the Cauchy-Schwarz inequality, then the bound $|a + \mathrm{i}b| \geqslant |a|$ and then integration by parts we get

$$\begin{aligned}
\|xf\|_2 \|\xi\widehat{f}\|_2 = \|xf\|_2 \|\widehat{\nabla f}\|_2 &= (2\pi)^{\frac{n}{2}} \|xf\|_2 \|\nabla f\|_2 \\
&\geqslant (2\pi)^{\frac{n}{2}} \left| \int_{\mathbb{R}^n} xf(x) \cdot \overline{\nabla f(x)} \, \mathrm{d}x \right| \\
&\geqslant (2\pi)^{\frac{n}{2}} \left| \int_{\mathbb{R}^n} x \cdot \mathrm{Re}\big(f(x)\nabla\overline{f}(x)\big) \, \mathrm{d}x \right| \\
&= (2\pi)^{\frac{n}{2}} \left| \int_{\mathbb{R}^n} x \cdot \nabla\big(\tfrac{1}{2}|f(x)|^2\big) \, \mathrm{d}x \right| \\
&= (2\pi)^{\frac{n}{2}} \tfrac{n}{2} \|f\|_2^2.
\end{aligned}$$

From the known cases of equality in the Cauchy-Schwarz inequality we deduce that equality holds for f if and only if $xf(x)$ and $\nabla f(x)$ are proportional and (using that $|a + \mathrm{i}b| = |a|$ holds iff $b = 0$)

$$\int_{\mathbb{R}^n} x \cdot \mathrm{Im}\big(f(x)\overline{\nabla f(x)}\big) \, \mathrm{d}x = 0.$$

It is not difficult to see that the former implies that $f(x) = c\mathrm{e}^{\frac{c_0}{2}|x|^2}$ for constants c, $c_0 \in \mathbb{C}$. Since $f \in \mathscr{S}(\mathbb{R}^n)$ we infer that, provided $c \neq 0$ that we henceforth assume, $\mathrm{Re}(c_0) < 0$. Next we calculate

$$f(x)x \cdot \overline{\nabla f(x)} = |c|^2 \bar{c}_0 |x|^2 \mathrm{e}^{\mathrm{Re}(c_0)|x|^2},$$

hence from

$$0 = \int_{\mathbb{R}^n} x \cdot \mathrm{Im}\big(f(x)\overline{\nabla f(x)}\big)\,\mathrm{d}x = -\mathrm{Im}(c_0)|c|^2 \int_{\mathbb{R}^n} |x|^2 \mathrm{e}^{\mathrm{Re}(c_0)|x|^2}\,\mathrm{d}x$$

we get $\mathrm{Im}(c_0) = 0$. Conversely, it is clear that equality holds in Heisenberg's inequality for this type of functions. □

The Heisenberg inequality is closely related to a *Sobolev type inequality*, meaning loosely speaking an integral inequality that bounds a function in terms of its derivatives. Indeed using the differentiation rules and Plancherel's theorem we see that Heisenberg's inequality for $x_0 = \xi_0 = 0$ is equivalent to

$$(2\pi)^{\frac{n}{2}} \tfrac{n}{2} \|f\|_2^2 \leqslant \|\nabla \widehat{f}\|_2 \|\nabla f\|_2 \tag{22}$$

for all $f \in \mathscr{S}(\mathbb{R}^n)$. For comparison we state:

Theorem 1.95. ***[A basic Sobolev inequality]*** *Assume $n \geqslant 3$. Then*

$$S_n \|f\|_{\frac{2n}{n-2}} \leqslant \|\nabla f\|_2$$

holds for all $f \in \mathscr{S}(\mathbb{R}^n)$, where

$$S_n := \tfrac{n(n-2)}{4} \omega_{n-1}^{\frac{2}{n}}$$

and we recall that ω_{n-1} is the area of the unit sphere $\partial B_1(0)$ in $\mathbb{R}^n$. The constant S_n is sharp.

We omit the proof and merely remark that the exponent in the basic Sobolev inequality *must be* $2n/(n-2)$. Indeed this follows if we assume that $\|\nabla f\|_2 \geqslant c\|f\|_p$ holds for some constant c for all $f \in \mathscr{S}(\mathbb{R}^n)$ and then apply the inequality to the dilated functions $d_r f$ and consider what happens when $r > 0$ is large and when it is small.

Remark 1.96. Heisenberg's inequality holds for *all* $f \in \mathrm{L}^2(\mathbb{R}^n)$ if we define $\|(x - x_0)f\|_2 = \infty$ when $(x - x_0)f \notin \mathrm{L}^2(\mathbb{R}^n)$ and $\|(\xi - \xi_0)\widehat{f}\|_2 = \infty$ when $(\xi - \xi_0)\widehat{f} \notin \mathrm{L}^2(\mathbb{R}^n)$. In order to prove this we first note that the inequality is trivial when its right-hand side is ∞. We may therefore assume, in addition to $f \in \mathrm{L}^2(\mathbb{R}^n)$, that $(x - x_0)f$, $(\xi - \xi_0)\widehat{f} \in \mathrm{L}^2(\mathbb{R}^n)$. The reduction to the case $x_0 = \xi_0 = 0$ is then the same as before. By the differentiation rules and Plancherel's theorem we infer that f, $\widehat{f} \in \mathrm{H}^1(\mathbb{R}^n)$. For the standard mollifier $\big(\rho_\varepsilon\big)_{\varepsilon>0}$ on $\mathbb{R}^n$ we define $f_\varepsilon = \rho_\varepsilon * f$. Then the results on mollification from B4.3 yield that $f_\varepsilon \in \mathrm{C}^\infty(\mathbb{R}^n)$ and

$$\|f_\varepsilon\|_2 \leqslant \|f\|_2 \text{ and } \|\nabla f_\varepsilon\|_2 \leqslant \|\nabla f\|_2.$$

Observe that the proof of Theorem 1.94 still applies to such f_ε, whereby

$$(2\pi)^{\frac{n}{2}} \tfrac{n}{2} \|f_\varepsilon\|_2^2 \leqslant \|x f_\varepsilon\|_2 \|\xi \widehat{f_\varepsilon}\|_2.$$

We want to conclude by passing $\varepsilon \searrow 0$. Clearly this gives the wanted limit on the left-hand side. On the right-hand side we employ the differentiation rules and rewrite

$$\xi \widehat{f_\varepsilon} = -\mathrm{i}\widehat{\nabla f_\varepsilon} \text{ and } x f_\varepsilon = \mathrm{i}\nabla \widehat{f_\varepsilon}.$$

By Plancherel's theorem we see that $\widehat{\nabla f_\varepsilon} \to \widehat{\nabla f}$ in $\mathrm{L}^2(\mathbb{R}^n)$ as $\varepsilon \searrow 0$, and so by the differentiation rule again, $\|\xi \widehat{f_\varepsilon}\|_2 \to \|\xi \widehat{f}\|_2$ as $\varepsilon \searrow 0$. Finally, by the convolution and dilation rules $\widehat{f_\varepsilon} = d_\varepsilon \widehat{\rho} \widehat{f}$, hence $\nabla \widehat{f_\varepsilon} = \nabla\big(d_\varepsilon \widehat{\rho}\big)\widehat{f} + d_\varepsilon \widehat{\rho} \nabla \widehat{f}$. It follows from the triangle inequality that

$$\begin{aligned}\|\nabla \widehat{f} - \nabla \widehat{f_\varepsilon}\|_2 \leqslant & \varepsilon \| d_\varepsilon\big(\nabla \widehat{\rho}\big)\widehat{f}\|_2 \\ & + \|\big(1 - d_\varepsilon \widehat{\rho}\big)\nabla \widehat{f}\|_2 \to 0\end{aligned}$$

as $\varepsilon \searrow 0$. The conclusion follows from this.

2 Applications of the Fourier transform

2.1 Periodic distributions and Fourier series

Definition 2.1. Let $u \in \mathscr{D}'(\mathbb{R})$ and $T > 0$. Then we say that u is *periodic with period* T, or briefly T-periodic, if $\tau_T u = u$.

Example 2.2. Assume that u is a T-periodic distribution: $0 = \langle \tau_T u - u, \phi \rangle = \langle u, \tau_{-T}\phi - \phi \rangle$ for all $\phi \in \mathscr{D}(\mathbb{R})$. If u is a regular distribution, then this is, by the fundamental lemma of the calculus of variations, equivalent to $u(x+T) = u(x)$ for almost all $x \in \mathbb{R}$, so the definition coincides with the usual definition in this case. We also note, that in the general case of a distribution $u \in \mathscr{D}'(\mathbb{R})$ we have that u is T-periodic if and only if the dilated distribution $d_{\frac{T}{2\pi}} u$ is 2π-periodic since

$$\tau_{2\pi} d_{\frac{T}{2\pi}} u - d_{\frac{T}{2\pi}} u = d_{\frac{T}{2\pi}} \big(\tau_T u - u\big).$$

Intuitively, a T-periodic distribution is determined by its behaviour on $(0, T]$ or any interval of length T. In the following we shall confine attention to 2π-periodic distributions.

Example 2.3. Let

$$u = \sum_{j \in \mathbb{Z}} \mathrm{e}^{\mathrm{i}jx}.$$

We claim that this is a 2π-periodic distribution. In fact, we will show that it is a tempered distribution and that

$$\sum_{j=-m}^{j=n} \int_{-\infty}^{\infty} \phi(x) \mathrm{e}^{\mathrm{i}jx}\,\mathrm{d}x \to \langle u, \phi \rangle \text{ as } m,\, n \to \infty$$

holds for all $\phi \in \mathscr{S}(\mathbb{R})$. In order to prove this we recall the Fourier bounds. They imply in particular that we can find a constant $c \geqslant 0$ so $\overline{S}_{2,0}(\widehat{\phi}) \leqslant c \overline{S}_{2,2}(\phi)$ for all $\phi \in \mathscr{S}(\mathbb{R})$. Therefore

$$\begin{aligned}\big|\widehat{\phi}(-j)\big| = & \big(1 + j^2\big)\big|\widehat{\phi}(-j)\big| \tfrac{1}{1+j^2} \\ \leqslant & 2\overline{S}_{2,0}(\widehat{\phi}) \tfrac{1}{1+j^2}\end{aligned}$$

for all $j \in \mathbb{Z}$, and so the series $\sum_{j\in\mathbb{Z}}\langle e^{ijx}, \phi\rangle$ is absolutely convergent and we can estimate

$$\begin{aligned}\big|\langle u, \phi\rangle\big| &\leqslant 2c\overline{S}_{2,2}(\phi)\sum_{j\in\mathbb{Z}} \tfrac{1}{1+j^2} \\ &= 2c\left(1 + 2\sum_{j=1}^{\infty} \tfrac{1}{1+j^2}\right)\overline{S}_{2,2}(\phi).\end{aligned}$$

Consequently, $u \in \mathscr{S}'(\mathbb{R})$ has order at most 2 and it is clear that $\tau_{2\pi}u = u$.

It was no coincidence that the 2π-periodic distribution from the last example was tempered:

Lemma 2.4. *If $u \in \mathscr{D}'(\mathbb{R})$ is 2π-periodic, then $u \in \mathscr{S}'(\mathbb{R})$.*

Proof. Put $\chi = \rho * \mathbf{1}_{(-1,2\pi+1]}$, where ρ is the standard mollifier kernel on $\mathbb{R}$. We now define the *periodisation* of χ to be the function

$$P\chi(x) := \sum_{k\in\mathbb{Z}} \chi(x + 2\pi k), \quad x \in \mathbb{R}.$$

Clearly the series is a finite sum as it for each $x \in \mathbb{R}$ consists of at most 3 non-zero terms, hence $P\chi \in C^\infty(\mathbb{R})$. By the definition we have $P\chi(x + 2\pi) = P\chi(x)$ for all x and since $\chi = 1$ on $[0, 2\pi]$ we also have $P\chi(x) \geqslant 1$ for all x. Now define $\Psi := \chi/P\chi$. Then $\Psi \in \mathscr{D}(\mathbb{R})$ and its periodisation is

$$P\Psi(x) = \sum_{k\in\mathbb{Z}} \Psi(x + 2\pi k) \equiv 1$$

for all $x \in \mathbb{R}$. Now if $\phi \in \mathscr{D}(\mathbb{R})$, then we get (using that the sum is finite and that u is periodic):

$$\begin{aligned}\langle u, \phi\rangle = \left\langle u, \sum_{k\in\mathbb{Z}} \tau_{2\pi k}\Psi\phi\right\rangle &= \sum_{k\in\mathbb{Z}}\big\langle u, (\tau_{2\pi k}\Psi)\phi\big\rangle \\ &= \sum_{k\in\mathbb{Z}}\big\langle \tau_{-2\pi k}u, \Psi\tau_{-2\pi k}\phi\big\rangle \\ &= \sum_{k\in\mathbb{Z}}\big\langle u, \Psi\tau_{-2\pi k}\phi\big\rangle \\ &= \left\langle u, \Psi\sum_{k\in\mathbb{Z}} \tau_{-2\pi k}\phi\right\rangle \\ &= \langle u, \Psi P\phi\rangle. \end{aligned} \tag{23}$$

Put $K := [-2, 2\pi + 2]$ and use the boundedness property of u to find constants corresponding to K, $c = c_K \geqslant 0$, $m = m_K \in \mathbb{N}_0$ so

$$\big|\langle u, \varphi\rangle\big| \leqslant c\sum_{j=0}^{m} \sup|\varphi^{(j)}|$$

holds for all $\varphi \in \mathscr{D}(K)$. Since $\Psi P\phi \in \mathscr{D}(K)$ for all $\phi \in \mathscr{D}(\mathbb{R})$ we may take $\varphi = \Psi P\phi$ and get using Leibniz' rule

$$\begin{aligned}\big|\langle u,\phi\rangle\big| = \big|\langle u,\Psi P\phi\rangle\big| &\leqslant c\sum_{j=0}^{m}\sup\big|(\Psi P\phi)^{(j)}\big| \\ &\leqslant c\sum_{j=0}^{m}\sum_{s=0}^{j}\binom{j}{s}\sup\big|\Psi^{(j-s)}(P\phi)^{(s)}\big| \\ &\leqslant c2^{m+1}\max_{0\leqslant s\leqslant m}\max_{K}\big|\Psi^{(s)}\big|\max_{0\leqslant s\leqslant m}\max_{K}\big|(P\phi)^{(s)}\big| \\ &\leqslant c2^{m+1}\overline{S}_{0,m}(\Psi)\max_{0\leqslant s\leqslant m}\max_{|x|\leqslant 4\pi}\big|(P\phi)^{(s)}(x)\big|.\end{aligned}$$

Here we estimate the last factor as follows:

$$\begin{aligned}\max_{|x|\leqslant 4\pi}\big|(P\phi)^{(s)}(x)\big| &\leqslant \sum_{k\in\mathbb{Z}}\max_{|x|\leqslant 4\pi}\big|\phi^{(s)}(x+2\pi k)\big| \\ &\leqslant \sum_{k\in\mathbb{Z}}\max_{|x|\leqslant 4\pi}\left(\tfrac{1+(x+2\pi k)^2}{1+(x+2\pi k)^2}\big|\phi^{(s)}(x+2\pi k)\big|\right) \\ &\leqslant \sum_{k\in\mathbb{Z}}\max_{|x|\leqslant 4\pi}\tfrac{1}{1+(x+2\pi k)^2}2\overline{S}_{2,s}(\phi) \\ &\leqslant \left(\sum_{k\in\mathbb{Z}_-}\tfrac{1}{1+(2\pi k+4\pi)^2}+1+\sum_{k\in\mathbb{N}}\tfrac{1}{1+(2\pi k-4\pi)^2}\right)2\overline{S}_{2,s}(\phi) \\ &= 2\left(1+2\sum_{k\in\mathbb{N}}\tfrac{1}{1+4\pi^2(k-2)^2}\right)\overline{S}_{2,s}(\phi),\end{aligned}$$

and consequently we have shown that

$$\big|\langle u,\phi\rangle\big| \leqslant C\overline{S}_{2,m}(\phi)$$

holds for all $\phi \in \mathscr{D}(\mathbb{R})$ for some constant C. It follows that u extends by $\mathscr{S}$ continuity to $\mathscr{S}(\mathbb{R})$, and so is a tempered distribution. In fact, the extension is given by the formula in (23). □

Theorem 2.5. ***[The Poisson summation formula.]*** *If $\varphi \in \mathscr{S}(\mathbb{R})$, then*

$$\sum_{k\in\mathbb{Z}}\widehat{\varphi}(2\pi k) = \sum_{k\in\mathbb{Z}}\varphi(k)$$

holds.

The Poisson summation formula can also be stated as

$$\sum_{k\in\mathbb{Z}}\mathrm{e}^{-2\pi\mathrm{i}kx} = \sum_{k\in\mathbb{Z}}\delta_k \quad \text{in } \mathscr{S}'(\mathbb{R}), \tag{24}$$

where the two series are understood to converge in $\mathscr{S}'(\mathbb{R})$. For instance,

$$\sum_{k=-m}^{k=n} \langle \mathrm{e}^{-2\pi \mathrm{i}kx}, \phi \rangle \to \sum_{k\in\mathbb{Z}} \widehat{\phi}(2\pi k) \;\text{ as }\; m,\, n \to \infty$$

for each $\phi \in \mathscr{S}(\mathbb{R})$.

Proof. Put $u = \sum_{k\in\mathbb{Z}} \delta_{2\pi k}$. It is not difficult to see that the series converges in $\mathscr{S}'(\mathbb{R})$ and so by Fourier transformation $\widehat{u} = \sum_{k\in\mathbb{Z}} \mathrm{e}^{-2\pi \mathrm{i}k\xi}$. We record that $\tau_1 \widehat{u} = \widehat{u}$, so $\widehat{u}$ is 1-periodic, and also

$$\big(\mathrm{e}^{2\pi \mathrm{i}\xi} - 1\big)\widehat{u} = 0.$$

Because $\mathrm{e}^{-2\pi \mathrm{i}\xi} - 1 \neq 0$ when $\xi \in \mathbb{R} \setminus \mathbb{Z}$ it follows that $\widehat{u}$ is supported in $\mathbb{Z}$. In particular, the restriction $\widehat{u}|_{(-1,1)}$ has support in $\{0\}$, so by a result from B4.3 it follows that $\widehat{u}|_{(-1,1)}$ has the form $\sum_{j=0}^{m} c_j \delta_0^{(j)}$ for some constants $m \in \mathbb{N}_0$ and $c_j \in \mathbb{C}$. We assert that in fact $m = 0$ so that $\widehat{u}|_{(-1,1)} = c_0 \delta_0$. We see this as follows: if we have the general form of $\widehat{u}$, then we calculate for $\phi \in \mathscr{D}(-1,1)$,

$$\begin{aligned} 0 = \big\langle \big(\mathrm{e}^{2\pi \mathrm{i}\xi} - 1\big)\widehat{u}, \phi \big\rangle &= \sum_{j=0}^{m} (-1)^j c_j \tfrac{\mathrm{d}^j}{\mathrm{d}\xi^j}|_{\xi=0} \Big(\big(\mathrm{e}^{2\pi \mathrm{i}\xi} - 1\big)\phi \Big) \\ &= \sum_{j=1}^{m} (-1)^j c_j \sum_{s=1}^{j} \binom{j}{s} (2\pi \mathrm{i})^s \phi^{(j-s)}(0). \end{aligned}$$

Here we take $\phi(x) := \frac{x^j}{j!}\big(\rho_\varepsilon * \mathbf{1}_{(-\varepsilon,\varepsilon)}\big)(x)$, $x \in (-1,1)$, for a fixed $\varepsilon \in (0,1/2)$. Since $\phi \in \mathscr{D}(-1,1)$ and $\phi^{(s)}(0) = \delta_{s,j}$ (Kronecker delta) we find $c_j = 0$ for all $1 \leqslant j \leqslant m$, and thus $\widehat{u}|_{(-1,1)} = c_0 \delta_0$ for some $c_0 \in \mathbb{C}$ as asserted. But $\widehat{u}$ is 1-periodic, so

$$\widehat{u} = c_0 \sum_{k\in\mathbb{Z}} \delta_k. \tag{25}$$

In order to find the constant c_0 we evaluate the identity (25) at $\tau_x \phi$ for $\phi \in \mathscr{S}(\mathbb{R})$ and $x \in (0,1]$: the left-hand side is

$$\big\langle \widehat{u}, \tau_x \phi \big\rangle = \big\langle u, \widehat{\tau_x \phi} \big\rangle = \big\langle u, \mathrm{e}^{\mathrm{i}x(\cdot)} \widehat{\phi} \big\rangle,$$

and so equating with the result on the right-hand side we arrive at

$$\sum_{k\in\mathbb{Z}} \mathrm{e}^{2\pi \mathrm{i}kx} \widehat{\phi}(2\pi k) = c_0 \sum_{k\in\mathbb{Z}} \phi(k + x).$$

Note that the two series converge uniformly in $x \in (0,1]$ and so integrating over $x \in (0,1]$ we get by integration term-by-term that

$$\widehat{\phi}(0) = \int_0^1 \sum_{k\in\mathbb{Z}} \mathrm{e}^{2\pi\mathrm{i}kx}\widehat{\phi}(2\pi k)\,\mathrm{d}x = \int_0^1 c_0 \sum_{k\in\mathbb{Z}} \phi(k+x), \mathrm{d}x$$
$$= c_0 \sum_{k\in\mathbb{Z}} \int_k^{k+1} \phi(y)\,\mathrm{d}y$$
$$= c_0 \int_{-\infty}^{\infty} \phi(y)\,\mathrm{d}y.$$

Thus $c_0 = 1$ and we are done. □

Example 2.6. The Poisson summation formula (24) can also be written $\sum_{k\in\mathbb{Z}} \mathrm{e}^{2\pi\mathrm{i}kx} = \sum_{k\in\mathbb{Z}} \delta_k$. Apply for a $T > 0$ the dilation $d_{1/T}$ to both sides to obtain

$$\sum_{k\in\mathbb{Z}} \mathrm{e}^{\frac{2\pi}{T}\mathrm{i}kx} = T\sum_{k\in\mathbb{Z}} \delta_{kT}, \tag{26}$$

hence in particular we record this for $T = 2\pi$: $\sum_{k\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}kx} = 2\pi \sum_{k\in\mathbb{Z}} \delta_{2\pi k}$. Many other identities can be derived by such simple manipulations, for instance we can differentiate in the distributional sense to obtain $\sum_{k\in\mathbb{Z}} \mathrm{i}k\mathrm{e}^{\mathrm{i}kx} = 2\pi \sum_{k\in\mathbb{Z}} \delta'_{2\pi k}$ and so on.

Note that the distribution on the left-hand side of the identity (26) is continuous with respect to the norm $\overline{S}_{2,2}$ (by virture of the Fourier bounds) and that the distribution on the right-hand side is continuous with respect to the norm $\overline{S}_{2,0}$ and so in particular with respect to $\overline{S}_{2,2}$ too. It therefore follows that we can evaluate (26) on any function $f\colon \mathbb{R} \to \mathbb{C}$ of class C^2 satisfying $\overline{S}_{2,2}(f) < \infty$, that is, for any such f we have

$$\sum_{k\in\mathbb{Z}} \widehat{f}\big(\tfrac{2\pi}{T}k\big) = T\sum_{k\in\mathbb{Z}} f(kT).$$

In fact, it is not difficult to check that the identity holds for all continuous functions $f\colon \mathbb{R} \to \mathbb{C}$ for which $\overline{S}_{2,0}(f) + \overline{S}_{2,0}(\widehat{f}) < \infty$, a condition that does not explicitly involve derivatives of f.

Example 2.7. Let $u \in \mathscr{D}'(\mathbb{R})$ be 2π-periodic. Then $u \in \mathscr{S}'(\mathbb{R})$ and $\langle u, \phi\rangle := \langle u, \Psi P\phi\rangle$ for $\phi \in \mathscr{S}(\mathbb{R})$ by Lemma 2.4 and (23). Let us calculate its Fourier transform:

$$\langle \widehat{u}, \phi\rangle = \langle u, \widehat{\phi}\rangle = \langle u, \Psi P\widehat{\phi}\rangle = \left\langle u, \Psi \sum_{k\in\mathbb{Z}} \widehat{\phi}(\cdot + 2\pi k)\right\rangle$$
$$\text{translation rule } = \left\langle u, \Psi \sum_{k\in\mathbb{Z}} \mathcal{F}_{x\to 2\pi k}\big(\phi(x)\mathrm{e}^{-\mathrm{i}x(\cdot)}\big)\right\rangle$$
$$\text{Poisson summation } = \left\langle u, \Psi \sum_{k\in\mathbb{Z}} \phi(k)\mathrm{e}^{-\mathrm{i}k(\cdot)}\right\rangle$$
$$= \sum_{k\in\mathbb{Z}} \phi(k)\big\langle u, \Psi\mathrm{e}^{-\mathrm{i}k(\cdot)}\big\rangle,$$

where the last equality follows by $\mathscr{S}$ continuity of u and $\mathscr{S}$ convergence of the series. Consequently,

$$\widehat{u} = \sum_{k\in\mathbb{Z}} 2\pi c_k \delta_k \text{ with } c_k = \tfrac{1}{2\pi}\big\langle u, \Psi \mathrm{e}^{-\mathrm{i}k(\cdot)}\big\rangle, \tag{27}$$

where the series converges in $\mathscr{S}'(\mathbb{R})$, meaning that

$$\left\langle 2\pi \sum_{k=-m}^{k=n} c_k \delta_k, \phi \right\rangle \to \langle \widehat{u}, \phi\rangle \text{ as } m,\, n \to \infty$$

for each $\phi \in \mathscr{S}(\mathbb{R})$. Using the Fourier inversion formula in $\mathscr{S}'(\mathbb{R})$ we find

$$u = \tfrac{1}{2\pi}\widetilde{\mathcal{F}^2 u} = \sum_{k\in\mathbb{Z}} c_k \mathrm{e}^{\mathrm{i}kx} \tag{28}$$

again with convergence of the series in $\mathscr{S}'(\mathbb{R})$.

Let us now assume that u is a 2π-periodic regular distribution. We claim that the coefficients c_k are the usual Fourier coefficients of the function u. In order to show this we simply calculate for each $k \in \mathbb{Z}$:

$$\begin{aligned}
2\pi c_k = \big\langle u, \Psi \mathrm{e}^{-\mathrm{i}k(\cdot)}\big\rangle &= \int_{-\infty}^{\infty} u(x)\Psi(x)\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x = \sum_{j\in\mathbb{Z}} \int_{2\pi j}^{2\pi(j+1)} u(x)\Psi(x)\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x \\
&= \sum_{j\in\mathbb{Z}} \int_0^{2\pi} u(x-2\pi j)\Psi(x-2\pi j)\mathrm{e}^{-\mathrm{i}k(x-2\pi j)}\,\mathrm{d}x \\
&= \sum_{j\in\mathbb{Z}} \int_0^{2\pi} u(x)\Psi(x-2\pi j)\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x \\
&= \int_0^{2\pi} u(x)\mathrm{e}^{-\mathrm{i}kx} \sum_{j\in\mathbb{Z}} \Psi(x-2\pi j)\,\mathrm{d}x = \int_0^{2\pi} u(x)\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x,
\end{aligned}$$

where we used that the periodisation of Ψ converges uniformly.

Because of this it is natural to make the following

Definition 2.8. The numbers c_k defined in (27) are called the *Fourier coefficients* of u and (28) is the *Fourier series expansion* of u.

Proposition 2.9. *Let $(c_k)_{k\in\mathbb{Z}}$ be a doubly infinite sequence of complex numbers.*

(1) Then $(c_k)_{k\in\mathbb{Z}}$ are the Fourier coefficients for a 2π-periodic C^∞ *function $u\colon \mathbb{R} \to \mathbb{C}$ if and only if for all $m \in \mathbb{N}_0$ we have*

$$k^m c_k \to 0 \ \ \textit{as} \ \ |k| \to \infty. \tag{29}$$

(2) Then $(c_k)_{k\in\mathbb{Z}}$ are the Fourier coefficients for a 2π-periodic distribution if and only if there exist constants $c \geqslant 0$ and $N \in \mathbb{N}_0$ so $|c_k| \leqslant c\big(1+|k|\big)^N$ holds for all $k \in \mathbb{Z}$. In this case we say the doubly infinite sequence has moderate growth*.*

Proof of (1). If $u\colon \mathbb{R} \to \mathbb{C}$ is 2π-periodic and C^∞, then we get by partial integration m times for each $k \in \mathbb{Z} \setminus \{0\}$:

$$c_k = \big(\mathrm{i}k\big)^{-m} \tfrac{1}{2\pi} \int_0^{2\pi} u^{(m)}(x) \mathrm{e}^{-\mathrm{i}kx} \, \mathrm{d}x,$$

and so the fast decay (29) follows easily. Conversely, if we have the fast decay (29), then the series

$$\sum_{k\in\mathbb{Z}} \big(\mathrm{i}k\big)^m c_k \mathrm{e}^{\mathrm{i}kx}$$

is uniformly convergent in $x \in \mathbb{R}$ by the Weierstrass M-test. But then the function

$$u(x) := \sum_{k\in\mathbb{Z}} c_k \mathrm{e}^{\mathrm{i}kx},\ x \in \mathbb{R},$$

is the 2π-periodic C^∞ function with Fourier coefficients c_k. □

The proof of (2) is an exercise.

Example 2.10. Let $u = \sum_{k\in\mathbb{Z}} c_k \delta_k$, where $\big(c_k\big)_{k\in\mathbb{Z}}$ is a doubly infinite sequence of complex numbers. It is clear that we always have $u \in \mathscr{D}_0'(\mathbb{R})$. However, $u \in \mathscr{S}'(\mathbb{R})$ if and only if $\big(c_k\big)_{k\in\mathbb{Z}}$ has moderate growth. In view of Proposition 2.9 this is not surprising.

Let us briefly summarize our discussion of Fourier series so far. We emphasize three points:

- When $\varphi \in \mathscr{S}(\mathbb{R})$, then its periodisation $P\varphi$ is a 2π-periodic C^∞ function whose Fourier series was calculated using Poisson's summation formula

$$P\varphi(x) = \sum_{k\in\mathbb{Z}} \tfrac{\widehat{\varphi}(k)}{2\pi} \mathrm{e}^{\mathrm{i}kx},$$

Here the convergence is in the C^∞ sense, meaning local uniform convergence of the partial sums $\sum_{k=-m}^{n} \tfrac{\widehat{\varphi}(k)}{2\pi} \mathrm{e}^{\mathrm{i}kx}$ and all their derivatives locally uniformly in $x \in \mathbb{R}$ as $m, n \to \infty$. In fact, the above Fourier series for $P\varphi$ is equivalent to the Poisson summation formula and is perhaps easier to remember.

- Given any 2π-periodic C^∞ function ϕ its Fourier series converges to ϕ in the C^∞ sense. In view of Proposition 2.9(1) and the Fourier inversion formula in $\mathscr{S}$ there exists $\varphi \in \mathscr{S}(\mathbb{R})$ so $\phi = P\varphi$. We note that Proposition 2.9(1) also gives a characterization of the Fourier coefficients for 2π-periodic C^∞ functions.

- A 2π-periodic distribution $u \in \mathscr{D}'(\mathbb{R})$ is tempered and given for $\phi \in \mathscr{S}(\mathbb{R})$ by $\langle u, \phi\rangle := \langle u, \Psi P\phi\rangle$. We have

$$u = \sum_{k\in\mathbb{Z}} c_k \mathrm{e}^{\mathrm{i}kx}, \ \text{ where } \ c_k = \tfrac{1}{2\pi}\langle u, \Psi \mathrm{e}^{-\mathrm{i}k(\cdot)}\rangle,$$

and the convergence of partial sums is in $\mathscr{S}'(\mathbb{R})$. Also here Proposition 2.9(2) gives a characterization of the possible Fourier coefficients.

Theorem 2.11. ***[The Plancherel Theorem for Fourier series.]***
Assume that $u \in \mathrm{L}^2_{\mathrm{loc}}(\mathbb{R})$ is 2π-periodic and has Fourier coefficients c_k. Then

$$u(x) = \sum_{k\in\mathbb{Z}} c_k \mathrm{e}^{\mathrm{i}kx} \quad \text{in } \mathrm{L}^2\big(0, 2\pi\big)$$

and

$$\tfrac{1}{2\pi} \int_0^{2\pi} |u(x)|^2 \,\mathrm{d}x = \sum_{k\in\mathbb{Z}} |c_k|^2. \tag{30}$$

Conversely, if $(C_k)_{k\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$, then the series

$$\sum_{k\in\mathbb{Z}} C_k \mathrm{e}^{\mathrm{i}kx}$$

converges in $\mathrm{L}^2\big(0, 2\pi\big)$ to a 2π-periodic $\mathrm{L}^2_{\mathrm{loc}}$ function with Fourier coefficients C_k.

The identity (30) is often also called Parseval's formula (as were (14) and (15)).

Proof. Assume first that $u\colon \mathbb{R} \to \mathbb{C}$ is a 2π-periodic C^∞ function. Then we have from Proposition 2.9 (1) that its Fourier series converges uniformly:

$$u(x) = \sum_{k\in\mathbb{Z}} c_k \mathrm{e}^{\mathrm{i}kx} \qquad \text{uniformly in } x \in \mathbb{R}.$$

Consequently we calculate

$$\begin{aligned} \int_0^{2\pi} |u(x)|^2 \,\mathrm{d}x &= \int_0^{2\pi} \sum_{k,l\in\mathbb{Z}} c_k \overline{c_l} \mathrm{e}^{\mathrm{i}(k-l)x} \,\mathrm{d}x \\ &= \sum_{k,l\in\mathbb{Z}} c_k \overline{c_l} \int_0^{2\pi} \mathrm{e}^{\mathrm{i}(k-l)x} \,\mathrm{d}x \\ &= 2\pi \sum_{k\in\mathbb{Z}} |c_k|^2 \end{aligned}$$

as required. We next turn to the general case where $u\colon \mathbb{R} \to \mathbb{C}$ is a 2π-periodic $\mathrm{L}^2_{\mathrm{loc}}$ function (or rather one of its representatives). Put $u_t := \rho_t * u$, where $\big(\rho_t\big)_{t>0}$ is the standard mollifier on $\mathbb{R}$. Then $u_t\colon \mathbb{R} \to \mathbb{C}$ is a 2π-periodic C^∞ function and if $c_k(t)$ are its Fourier coefficients we know that $u_t(x) = \sum_{k\in\mathbb{Z}} c_k(t)\mathrm{e}^{\mathrm{i}kx}$ holds uniformly in $x \in \mathbb{R}$ and that $\int_0^{2\pi} |u(x) - u_t(x)|^2 \,\mathrm{d}x \to 0$ as $t \searrow 0$. If we apply what we just proved to the difference $u_s - u_t$, a 2π-periodic C^∞ function, we find

$$\tfrac{1}{2\pi} \int_0^{2\pi} |u_s(x) - u_t(x)|^2 \,\mathrm{d}x = \sum_{k\in\mathbb{Z}} |c_k(s) - c_k(t)|^2.$$

Consequently, as $(u_t)_{t>0}$ is Cauchy in L^2 as $t \searrow 0$ also $\big((c_k(t))_{k\in\mathbb{Z}}\big)_{t>0}$ is Cauchy in $\ell^2(\mathbb{Z})$ as $t \searrow 0$. But the latter is complete by the Riesz-Fischer theorem so for some $(a_k)_{k\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$ we have $\sum_{k\in\mathbb{Z}} |c_k(t) - a_k|^2 \to 0$ as $t \searrow 0$. Now we have in particular that $u_t \to u$ in $\mathscr{S}'(\mathbb{R})$ as $t \searrow 0$ so, by $\mathscr{S}'$ continuity of the Fourier transform, also $\widehat{u_t} \to \widehat{u}$ in $\mathscr{S}'(\mathbb{R})$ as $t \searrow 0$. Since

$$\widehat{u_t} = 2\pi \sum_{k\in\mathbb{Z}} c_k(t)\delta_k \ \text{ and } \ \widehat{u} = 2\pi \sum_{k\in\mathbb{Z}} c_k\delta_k,$$

it follows that $c_k(t) \to c_k$ as $t \searrow 0$ pointwise in $k \in \mathbb{Z}$. Therefore $c_k = a_k$ for all $k \in \mathbb{Z}$, and so taking $t \searrow 0$ in

$$\tfrac{1}{2\pi} \int_0^{2\pi} |u_t(x)|^2 \,\mathrm{d}x = \sum_{k\in\mathbb{Z}} |c_k(t)|^2$$

and Parseval's formula (30) follows. Finally we check that for $m,\, n \in \mathbb{N}$,

$$\tfrac{1}{2\pi} \int_0^{2\pi} \left| u(x) - \sum_{k=-m}^{n} c_k \mathrm{e}^{\mathrm{i}kx} \right|^2 \mathrm{d}x = \tfrac{1}{2\pi} \int_0^{2\pi} |u(x)|^2 \,\mathrm{d}x - \sum_{k=-m}^{n} |c_k|^2 \to 0 \ \text{ as } \ m,\, n \to \infty.$$

In the opposite direction we are given $(C_k)_{k\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$. Since then clearly $\sup_{k\in\mathbb{Z}} |C_k| < \infty$ we may by virtue of Proposition 2.9 (2) define

$$u := \sum_{k\in\mathbb{Z}} C_k \mathrm{e}^{\mathrm{i}kx}$$

with convergence in $\mathscr{S}'(\mathbb{R})$. It follows by the previous part of the proof that the series converges in $\mathrm{L}^2(0, 2\pi)$ so the proof is concluded. □

Example 2.12. When $f\colon \mathbb{R} \to \mathbb{C}$ is a 2π-periodic $\mathrm{H}^1_{\mathrm{loc}}$ function with $c_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,\mathrm{d}x$, then

$$\int_0^{2\pi} |f(x) - c_0|^2 \,\mathrm{d}x \leqslant \int_0^{2\pi} |f'(x)|^2 \,\mathrm{d}x$$

holds. Equality holds precisely when $f = c_0 + c_{-1}\mathrm{e}^{-\mathrm{i}x} + c_1\mathrm{e}^{\mathrm{i}x}$.

By Plancherel's theorem for Fourier series we have $f(x) = \sum_{k\in\mathbb{Z}} c_k \mathrm{e}^{\mathrm{i}x}$ in $\mathrm{L}^2(0, 2\pi)$ and $f'(x) = \sum_{k\in\mathbb{Z}} \mathrm{i}kc_k \mathrm{e}^{\mathrm{i}x}$ in $\mathrm{L}^2(0, 2\pi)$. Consequently Parseval's identity gives

$$\begin{aligned}
\int_0^{2\pi} |f(x) - c_0|^2 \,\mathrm{d}x &= 2\pi \sum_{k\in\mathbb{Z}\setminus\{0\}} |c_k|^2 \\
&= 2\pi \sum_{k\in\mathbb{Z}\setminus\{0\}} \tfrac{k^2}{k^2} |c_k|^2 \\
&\leqslant 2\pi \sum_{k\in\mathbb{Z}\setminus\{0\}} |\mathrm{i}kc_k|^2 \\
&= \int_0^{2\pi} |f'(x)|^2 \,\mathrm{d}x
\end{aligned}$$

as required. We note that equality occurs precisely when $c_k = 0$ for all $k \in \mathbb{Z} \setminus \{0, \pm 1\}$.

Remark 2.13. We have briefly discussed convergence properties of Fourier series and have emphasized that our results refer to convergence of $\sum_{k=-m}^{k=n} c_k \mathrm{e}^{\mathrm{i}kx}$ in various different senses (C^∞, $\mathrm{L}^2_{\mathrm{loc}}$ and $\mathscr{S}'$) as m, $n \to \infty$. Classically one has however approached the question of convergence differently and been interested in the convergence of the *symmetric* partial sums

$$S_n(x) := \sum_{k=-n}^{n} c_k \mathrm{e}^{\mathrm{i}kx}$$

as $n \to \infty$. In this connection it is also customary to discuss *summability methods*, including in particular *Cesàro summability* and *Abel summability.* The former refers to convergence of arithmetic means of the partial sums

$$F_n(x) := \tfrac{1}{n} \sum_{j=0}^{n-1} S_j(x)$$

and the latter to convergence of the Abel means

$$A_n(x, r) := \sum_{k=-n}^{n} c_k r^{|k|} \mathrm{e}^{\mathrm{i}kx},$$

where $r \in (0, 1)$.

2.2 Fundamental solutions

Theorem 2.14. *Let $n \geqslant 3$. Then*

$$E(x) = -\frac{1}{(n-2)\omega_{n-1}} |x|^{2-n},$$

$x \in \mathbb{R}^n \setminus \{0\}$, is a fundamental solution for Δ.

Remark 2.15. Note that E is C^∞ away from zero, $E \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n) \cap \mathscr{S}'(\mathbb{R}^n)$, and that

$$\partial_j E = \frac{1}{\omega_{n-1}} x_j |x|^{-n} \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n).$$

The constant ω_{n-1} is the surface area of $\mathbb{S}^{n-1}$ in $\mathbb{R}^n$. One can show that $\omega_{n-1} = n\mathcal{L}^n(B_1(0))$, and

$$\mathcal{L}^n(B_1(0)) = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2} + 1\right)}$$

for all $n \in \mathbb{N}$. In particular, we record the values for $n = 2$ and 3:

$$\omega_1 = 2\mathcal{L}^2(B_1(0)) = 2\pi, \qquad \omega_2 = 3\mathcal{L}^3(B_1(0)) = 4\pi.$$

The calculation uses $\Gamma(x+1) = x\Gamma(x)$ and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Note also that

- in $n = 1$ the Laplacian $\frac{\mathrm{d}^2}{\mathrm{d}x^2}$ has fundamental solution x^+,
- in $n = 2$ the Laplacian $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y}$ has fundamental solution $\frac{1}{2\pi} \log \sqrt{x^2 + y^2}$. This is called the *logarithmic potential*;
- in $n = 3$ the Laplacian $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ has fundamental solution

$$-\frac{1}{4\pi} \frac{1}{\sqrt{x^2 + y^2 + z^2}}.$$

This is called the *Newtonian potential*.

Proof. Fourier transforming $\Delta E = \delta_0$ we get

$$1 = \hat{\delta}_0 = \widehat{\Delta E} = -|\xi|^2 \hat{E}(\xi).$$

This is not enough to deduce that $\hat{E}(\xi) = -\frac{1}{|\xi|^2}$, only that

$$\hat{E}(\xi) = -\frac{1}{|\xi|^2} + \hat{T}$$

for some $T \in \mathscr{S}'(\mathbb{R}^n)$ satisfies $\Delta T = 0$. This means that $-|\xi|^2 \hat{T} = 0$ in $\mathscr{S}'(\mathbb{R}^n)$. Hence if $\varphi \in \mathscr{S}(\mathbb{R}^n)$ and $0 \notin \operatorname{supp} \varphi$, then

$$\psi(\xi) = -\frac{\varphi(\xi)}{|\xi|^2}$$

for $\xi \neq 0$ and $\psi(0) = 0$ belongs to $\mathscr{S}(\mathbb{R}^n)$, and so

$$\langle \hat{T}, \varphi \rangle = \langle -|\xi|^2 \hat{T}, \psi \rangle = 0.$$

We express this by $\operatorname{supp} \hat{T} = \{0\}$, that is $\hat{T}$ has support $\{0\}$. From B4.3 we know that this implies that $\hat{T} \in \operatorname{span}\{\partial^\alpha \delta_0 \, : \, \alpha \in \mathbb{N}_0^n\}$, and hence that $T \in \operatorname{span}\{(2\pi)^{-n}(ix)^\alpha \, : \, \alpha \in \mathbb{N}_0^n\} = \mathbb{C}[x]$. Since also $\Delta T = 0$, we see that T must be a harmonic polynomial. Note that implicit in this is the Liouville-type result saying that if $T \in \mathscr{S}'(\mathbb{R}^n)$ is harmonic, then T is a polynomial. Now we return to the quest for fundamental solutions:

$$\hat{E}(\xi) = -\frac{1}{|\xi|^2} + \hat{T}(\xi).$$

We only need one, so consider $\hat{E} = -\frac{1}{|\xi|^2}$. The result then follows from the following.

Lemma 2.16 (Auxiliary Lemma)**.** *Let $\alpha \in (-n, 0)$ and put $f(x) = |x|^\alpha$. Then $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n) \cap \mathscr{S}'(\mathbb{R}^n)$ and $\hat{f}(\xi) = c(n, \alpha)|\xi|^{-n-\alpha}$, where*

$$c(n, \alpha) = 2^{\alpha+n} \pi^{\frac{n}{2}} \frac{\Gamma\left(\frac{n+\alpha}{2}\right)}{\Gamma\left(-\frac{\alpha}{2}\right)}$$

and $-n < -n - \alpha < 0$.

Proof. We start with the observation that for $x \neq 0$

$$|x|^\alpha \Gamma\left(-\frac{\alpha}{2}\right) = |x|^\alpha \int_0^\infty t^{-\frac{\alpha}{2}-1} \mathrm{e}^{-t}\, \mathrm{d}t = \int_0^\infty s^{-\frac{\alpha}{2}-1} \mathrm{e}^{-s|x|^2}\, \mathrm{d}s,$$

where we made the substitution $t = s|x|^2$, and hence

$$|x|^\alpha = \frac{1}{\Gamma\left(-\frac{\alpha}{2}\right)} \int_0^\infty s^{-\frac{\alpha}{2}-1} \mathrm{e}^{-s|x|^2}\, \mathrm{d}s.$$

Note that

$$\frac{1}{\Gamma\left(-\frac{\alpha}{2}\right)} \int_0^j s^{-\frac{\alpha}{2}-1} \mathrm{e}^{-s|x|^2}\, \mathrm{d}s \underset{j\to\infty}{\longrightarrow} |x|^\alpha$$

in $\mathscr{S}'(\mathbb{R}^n)$ and Riemann sums for the integrals

$$\frac{1}{\Gamma\left(-\frac{\alpha}{2}\right)} \int_0^j s^{-\frac{\alpha}{2}-1} \mathrm{e}^{-s|x|^2}\, \mathrm{d}s.$$

for j fixed converge as mesh size tends to zero in the $\mathscr{S}'(\mathbb{R}^n)$ sense. Consequently we get by $\mathscr{S}'$-continuity and linearity of $\mathcal{F}$ that

$$\begin{aligned}
\mathcal{F}_{x\to\xi}(|x|^\alpha) &= \frac{1}{\Gamma\left(-\frac{\alpha}{2}\right)} \int_0^\infty s^{-\frac{\alpha}{2}-1} \mathcal{F}_{x\to\xi}(\mathrm{e}^{-s|x|^2})\, \mathrm{d}s \\
&= \frac{1}{\Gamma\left(-\frac{\alpha}{2}\right)} \int_0^\infty s^{-\frac{\alpha}{2}-1} \left(\frac{\pi}{s}\right)^{\frac{n}{2}} \mathrm{e}^{-\frac{|\xi|^2}{4s}}\, \mathrm{d}s \\
&= \frac{\pi^{\frac{n}{2}}}{\Gamma\left(-\frac{\alpha}{2}\right)} \left(\frac{|\xi|}{2}\right)^{-n-\alpha} \int_0^\infty t^{\frac{n+\alpha}{2}-1} \mathrm{e}^{-t}\, \mathrm{d}t \\
&= 2^{n+\alpha} \pi^{\frac{n}{2}} \frac{\Gamma\left(\frac{n+\alpha}{2}\right)}{\Gamma\left(-\frac{\alpha}{2}\right)} |\xi|^{-n-\alpha}.
\end{aligned}$$

□

□