Fourier analysis and distribution theory
Lecture notes, Fall 2022
(Version: June 1, 2025)

Pu-Zhao Kow

Department of Mathematics and Statistics, P.O. Box 35 (MaD), FI-40014 University of Jyväskylä, Finland

Email address: pu-zhao.pz.kow@jyu.fi

Preface

This lecture note is prepared for the course *Fourier Analysis* (MATS315) in the 1st period, which gives an introduction to Fourier analysis and distribution theory, mainly based on Mikko Salo's lecture note some parts of the monographs [**FJ98, Mit18**]. We also refer to some other materials such that the monographs [**Bre11, Str08**], as well as Richard Melrose's lecture note.

Title. Fourier analysis and distribution theory

Lectures. Thursday 10.15–12.00 in MaD380, Friday 10.15–12.00 in MaD381 (Exercise sessions are on Thursday at 08.30–10.00 in MaD355). Begins at September 1, 2022 and ends at October 20, 2022.

Language. Instruction and completion in English

Instructor. Pu-Zhao Kow, (Office: MaD369, Email: pu-zhao.pz.kow@jyu.fi)

Completion. The course can be taken for credit by attending the lectures, returning written solutions (at least 50%, please indicate your email) and writing a short report (not more than 10 pages) with an oral presentation on it at the end of the course. One possible topic for the final report is some theorems in the lecture note that the proofs are omitted. To get grade 5, you need to complete at least 80% of the returning written solutions.

There will be a following up course *Fourier Analysis II* (MATS5170) in the 2nd period, which will lectured by Tuomas Orponen. The course will be lectured in English.

https://sites.google.com/view/tuomaths/teaching/fourier-analysis-ii

Please see the website above for more specific information.

Contents

CHAPTER 1

Fourier series

1.1. Weak derivatives

Given any integer $n \geq 1$. For each multi-index $\alpha = (\alpha_1, \cdots, \alpha_n)$ with $\alpha_j \in \mathbb{Z}_{\geq 0}$, we define $\operatorname{supp}(\alpha) := \left\{ \ j \in \{1, \cdots, n\} \ \middle| \ \alpha_j \neq 0 \ \right\}$,

$$\alpha! := \alpha_1! \alpha_2! \cdots \alpha_n! \equiv \prod_{j=1}^{n} \alpha_j! \quad \text{and} \quad |\alpha| := \sum_{j=1}^{n} \alpha_j$$

$$\partial^\alpha := \prod_{j \in \operatorname{supp}(\alpha)} \partial_j^{\alpha_j} \text{ with the partial derivatives } \partial_j := \frac{\partial}{\partial x_j}$$

$$x^\alpha := \prod_{j \in \operatorname{supp}(\alpha)} x_j^{\alpha_j} \quad \text{for all } x = (x_1, \cdots, x_n) \in \mathbb{C}^n$$

with the convention

$$\partial^{(0,\cdots,0)} := \mathrm{Id} \quad \text{and} \quad x^{(0,\cdots,0)} := 1.$$

Let $\beta = (\beta_1, \cdots, \beta_n)$ be another multi-index such that

$$\beta \leq \alpha \quad (\text{that is, } \beta_j \leq \alpha_j \text{ for all } j),$$

we define the multi-index

$$\alpha - \beta := (\alpha_1 - \beta_1, \cdots, \alpha_n - \beta_n).$$

We have the following binomial theorem [**Mit18**]:

$$(x+y)^\gamma = \sum_{\alpha + \beta = \gamma} \frac{\gamma!}{\alpha! \beta!} x^\alpha y^\beta \quad \text{for all } x, y \in \mathbb{C}^n. \tag{1.1.1}$$

Let Ω be an open set in $\mathbb{R}^n$. For each $k \in \mathbb{Z}_{\geq 0} \cup \{\infty\}$, we define the following functional spaces:

$$\begin{aligned}
C^k(\Omega) &:= \left\{ \ \varphi : \Omega \to \mathbb{C} \ \middle| \ \partial^\alpha \varphi \text{ is continuous for all } \alpha \text{ with } |\alpha| \leq k \ \right\}, \\
C^k(\overline{\Omega}) &:= \left\{ \ \varphi|_{\overline{\Omega}} \ \middle| \ \varphi \in C^k(U) \text{ for some open set } U \supset \overline{\Omega} \ \right\}, \\
C_c^k(\Omega) &:= \left\{ \ \varphi \in C^k(\Omega) \ \middle| \ \operatorname{supp}(\varphi) \subset \Omega \text{ is compact } \right\} \\
&\equiv \left\{ \ \varphi \in C^k(\mathbb{R}^n) \ \middle| \ \operatorname{supp}(\varphi) \subset \Omega \text{ is compact } \right\}.
\end{aligned}$$

Given any $f \in C^1(\Omega)$, using (1-dimensional) integration by parts, we can easily compute that

$$\int_\Omega (\partial_j f) \varphi \, dx = - \int_\Omega f \partial_j \varphi \, dx \quad \text{for all } \varphi \in C_c^\infty(\Omega).$$

Using induction, for each $f \in C^{|\alpha|}(\Omega)$, we can easily verify that

$$\int_\Omega (\partial^\alpha f)\varphi \, dx = (-1)^{|\alpha|} \int_\Omega f \partial^\alpha \varphi \, dx \quad \text{for all } \varphi \in C_c^\infty(\Omega). \tag{1.1.2}$$

We see that the right-hand-side of (1.1.2) is actually well-defined for all locally L^1 functions f, i.e.

$$f \in L^1_{\rm loc}(\Omega) := \left\{ \, f : \Omega \to \mathbb{C} \mid \|f\|_{L^1(K)} \equiv \textstyle\int_K |f| \, dx < \infty \text{ for all compact set } K \subset \Omega \, \right\}.$$

This suggests the following definition:

DEFINITION 1.1.1. Let $f \in L^1_{\rm loc}(\Omega)$. A function $g \in L^1_{\rm loc}(\Omega)$ is called a *weak derivative* (of order α) of f if

$$\int_\Omega g\varphi \, dx = (-1)^{|\alpha|} \int_\Omega f \partial^\alpha \varphi \, dx \quad \text{for all } \varphi \in C_c^\infty(\Omega). \tag{1.1.3}$$

THEOREM 1.1.2 (Theorem 1.3 in [Mit18]). *If $g \in L^1_{\rm loc}(\Omega)$ satisfying $g = 0$ in Ω-distribution sense, i.e.*

$$\int_\Omega g\varphi \, dx = 0 \quad \textit{for all } \varphi \in C_c^\infty(\Omega), \tag{1.1.4}$$

then $g = 0$ a.e. in Ω.

REMARK 1.1.3. The converse of Theorem 1.1.2 is trivial. Here and after, we shall omit the notation "a.e." if there is no any ambiguity.

REMARK 1.1.4. In particular, for any multi-index α, each function $g \in L^1_{\rm loc}(\Omega)$ produced from $f \in L^1_{\rm loc}(\Omega)$ must be unique. Therefore, we can just simply write $\partial^\alpha f := g$. However, Theorem 1.1.2 does not guarantee the existence of such $\partial^\alpha f$.

PROOF OF THEOREM 1.1.2. Consider a function ϕ satisfying

$$\phi \in C_c^\infty(\mathbb{R}^n), \quad \phi \geq 0, \quad \operatorname{supp}(\phi) \subset \overline{B_1} \quad \text{and} \quad \int_{\mathbb{R}^n} \phi(x) \, dx = 1.$$

One concrete example is the function $\phi : \mathbb{R}^n \to \mathbb{R}$ defined by

$$\phi(x) = \begin{cases} C \exp\left(\dfrac{1}{|x|^2 - 1} \right) & \text{for } x \in B_1, \\ 0 & \text{otherwise,} \end{cases}$$

with $C := (\omega_{n-1} \int_0^1 \exp(\frac{1}{\rho^2-1}) \rho^{n-1} \, d\rho)^{-1}$, where $\omega_{n-1} = n|B_1| = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}$ is the surface area of the unit sphere ∂B_1. For each $\epsilon > 0$, we define (sometimes we refer it the *standard mollifier*)

$$\phi_\epsilon(x) = \frac{1}{\epsilon^n} \phi\left(\frac{x}{\epsilon} \right) \quad \text{for each } x \in \mathbb{R}^n.$$

Then for each $\epsilon > 0$ we have

$$\phi_\epsilon \in C_c^\infty(\mathbb{R}^n), \quad \phi_\epsilon \geq 0, \quad \operatorname{supp}(\phi_\epsilon) \subset \overline{B_\epsilon} \quad \text{and} \quad \int_{\mathbb{R}^n} \phi_\epsilon(x) \, dx = 1.$$

Fix $x \in \Omega$ and $0 < \epsilon < \operatorname{dist}(x, \partial\Omega)$, then $B_\epsilon(x) \subset \Omega$ and $\phi_\epsilon(x - \cdot) \in C_c^\infty(\Omega)$. Therefore from (1.1.4) we have

$$g * \phi_\epsilon(x) = \int_\Omega g(y) \phi_\epsilon(x-y) \, dy = 0 \quad \text{for all } x \in \Omega. \tag{1.1.5}$$

Therefore from Lebesgue's differentiation theorem we have

$$\begin{aligned}|g(x)| &= \left| \int_\Omega g(x)\phi_\epsilon(x-y)\,dy - \overbrace{\int_\Omega g(y)\phi_\epsilon(x-y)\,dy}^{=0 \text{ by } (1.1.5)} \right| \\ &\le \frac{1}{\epsilon^n}\int_{B_\epsilon(x)} |g(x)-g(y)|\phi\left(\frac{y}{\epsilon}\right)dy \\ &\le |B_1|\|\phi\|_{L^\infty(\mathbb{R}^n)} \frac{1}{|B_\epsilon(x)|}\int_{B_\epsilon(x)} |g(x)-g(y)|\,dy \to 0 \quad \text{as } \epsilon \to 0\end{aligned}$$

for a.e. $x \in \Omega$ (in particular, for all Lebesgue point x in Ω), which conclude our lemma. $\square$

When $f \in C^{|\alpha|}(\Omega)$, then the weak derivative $\partial^\alpha f$ is coincide with the usual derivative $\partial^\alpha f$ (which is continuous). The (classical) Laplacian is defined by

$$\Delta f := \sum_{j=1}^n \partial_j^2 f \quad \text{for all } f \in C^2(\Omega).$$

Using the weak derivatives as in Definition 1.1.1, we can define the weak Laplacian $\mathcal{L}f$ of f by

$$\int_\Omega (\mathcal{L}f)\varphi\,dx := \int_\Omega f\Delta\varphi\,dx \quad \text{for all } \varphi \in C_c^\infty(\Omega)$$

provided the weak derivatives of f exist. Similar ideas also can apply to other differential operators.

EXAMPLE 1.1.5. We now consider the Heaviside function

$$H(x) := \begin{cases} 1 & \text{for all } x > 0, \\ 0 & \text{for all } x \le 0. \end{cases} \tag{1.1.6}$$

It is easy to see that $H \in L^1_{\mathrm{loc}}(\mathbb{R})$. We define

$$f(x) := \begin{cases} x & \text{for all } x > 0, \\ 0 & \text{for all } x \le 0. \end{cases}$$

It is easy to see that

$$-\int_{\mathbb{R}} f(x)\varphi'(x)\,dx = -\int_0^\infty x\varphi'(x)\,dx = -x\varphi(x)\Big|_{x=0}^{x=\infty} + \int_0^\infty \varphi(x)\,dx = \int_{\mathbb{R}} H(x)\varphi(x)\,dx,$$

which shows that the Heaviside function H given in (1.1.6) is the weak derivative of order one of f.

However, not all $L^1_{\mathrm{loc}}(\Omega)$ function admits weak derivative:

EXAMPLE 1.1.6. We now show that the weak derivative of order 1 of the Heaviside function H given in (1.1.6) does not exist. Suppose the contrary, that H has a weak derivative of order 1, says $g \in L^1_{\mathrm{loc}}(\mathbb{R})$. We see that

$$\int_{-\infty}^\infty g(x)\varphi(x)\,dx = -\int_{-\infty}^\infty H(x)\varphi'(x)\,dx = -\int_0^\infty \varphi'(x)\,dx = \varphi(0) \tag{1.1.7}$$

for all $\varphi \in C_c^\infty(\mathbb{R})$. Hence we know that

$$\int_0^\infty g(x)\varphi(x)\,dx = 0 \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}\setminus\{0\}).$$

Using Theorem 1.1.2 with $\Omega = \mathbb{R}\setminus\{0\}$, we conclude that $g = 0$ a.e. in $\mathbb{R}$. Therefore from (1.1.7) we know that $\varphi(0) = 0$ for all $\varphi \in C_c^\infty(\mathbb{R})$, which leading to a contradiction.

Remark 1.1.7 (Theorem 1 in Section 4.3 of [**EG15**]). The following general integration by parts formula is well-known: Let Ω be a bounded Lipschitz domain in $\mathbb{R}^n$ and given any $1 \le p < \infty$. There exists a bounded linear operator

$$\mathrm{Tr} : W^{1,p}(\Omega) \to L^p(\partial\Omega; \mathcal{H}^{n-1}) \quad \text{such that} \quad \mathrm{Tr}\,(f) = f \text{ on } \partial\Omega \quad \text{for all } f \in W^{1,p}(\Omega) \cap C(\overline{\Omega}),$$

which is uniquely defined up to sets of $\mathcal{H}^{n-1}\lfloor\partial\Omega$ measure zero (It is called the *trace* of f on $\partial\Omega$). Furthermore, for all $\varphi \in (C^1(\mathbb{R}^n))^n$ and $f \in W^{1,p}(\Omega)$, we have

$$\int_\Omega f\,\mathrm{div}\,(\varphi)\,dx = -\int_\Omega \nabla f\cdot\varphi\,dx + \int_{\partial\Omega}(\nu\cdot\varphi)\mathrm{Tr}\,(f)\,d\mathcal{H}^{n-1},$$

where ν is the unit outer normal to $\partial\Omega$.

1.2. 1-dimensional Fourier series in L^2

Let $\Omega \subset \mathbb{R}^n$ be an open set and let $1 \le p \le \infty$. For each $m \in \mathbb{N}$, the Sobolev space $W^{m,p}(\Omega)$ is defined by

$$W^{m,p}(\Omega) := \left\{\ f \in L^p(\Omega) \ \middle|\ \partial^\alpha f \in L^p(\Omega) \text{ for all } |\alpha| \le m\ \right\},$$

equipped with the norm

$$\|f\|_{W^{m,p}(\Omega)} = \left(\sum_{|\alpha|\le m}\|\partial^\alpha f\|_{L^p(\Omega)}^p\right)^{\frac{1}{p}}.$$

where $\partial^\alpha f$ are the weak derivatives (Definition 1.1.1) of f. We set

$$H^m(\Omega) := W^{m,2}(\Omega).$$

Note that $H^m(\Omega)$ is indeed a Hilbert space equipped with the scalar product

$$(f,g)_{H^m(\Omega)} := \sum_{|\alpha|\le m}(\partial^\alpha f, \partial^\alpha g)_{L^2(\Omega)},$$

where

$$(f,g)_{L^2(\Omega)} = \int_\Omega f\overline{g}\,dx \quad \text{for all } f,g \in L^2(\Omega).$$

For $1 \le p < \infty$, we denote $W_0^{m,p}(\Omega)$ the closure of $C_c^\infty(\Omega)$ in $W^{m,p}(\Omega)$, and we set $H_0^1(\Omega) = W_0^{1,2}(\Omega)$. Here we refer to the monograph [**Bre11**] for properties of these Sobolev spaces, here we will not going to exhaust all of these details. The following theorem is an important fact in the Hilbert space theory, which can be found in [**Bre11**, Theorem 9.31] (see also [**Bre11**, Theorem 8.22] for 1-dimensional case).

Theorem 1.2.1 (Spectral decomposition of Dirichlet Laplacian). *Let Ω be a bounded Lipschitz domain. There exist a Hilbert basis $\{\phi_k\}_{k\in\mathbb{N}}$ of $L^2(\Omega)$ and a sequence of real numbers $\{\lambda_k\}_{k\in\mathbb{N}}$ with $0 < \lambda_1 \le \lambda_2 \le \cdots \to \infty$ such that*

$$\phi_k \in H_0^1(\Omega)\cap C^\infty(\Omega), \quad -\Delta\phi_k = \lambda_k\phi_k \text{ in } \Omega.$$

For each (real-valued) $f \in L^2(\Omega)$ we have

$$f = \sum_{k=1}^{\infty} (f, \phi_k)_{L^2(\Omega)} \phi_k \quad \text{converges in } L^2(\Omega). \tag{1.2.1}$$

The precise meaning of (1.2.1) is

$$\lim_{m\to\infty} \left\| f - \sum_{k=1}^{m} (f, \phi_k)_{L^2(\Omega)} \phi_k \right\|_{L^2(\Omega)} = 0.$$

If $(f, \phi_k)_{L^2(\Omega)} = 0$ for all $k \in \mathbb{N}$, then (1.2.1) implies that $f = 0$ a.e. in Ω. Moreover, the following Parseval-Bessel identity holds:

$$\|f\|_{L^2(\Omega)}^2 = \sum_{k=1}^{\infty} |(f, \phi_k)_{L^2(\Omega)}|^2.$$

In particular, the following general result on separable Hilbert space can be proved using the Hahn-Banach theorem [**Bre11**, Corollary 1.8]:

PROPOSITION 1.2.2. *Let $\big(H, (\cdot,\cdot)\big)$ be a separable Hilbert space, and let $\{\phi_k\}_{k\in\mathbb{N}}$ be an orthonormal subset of H. Then the following are equivalent:*

(1) *$\{\phi_k\}_{k\in\mathbb{N}}$ is an orthonormal (Hilbert) basis.*
(2) *The following Parseval identity holds:*
$$\|f\|^2 = \sum_{k\in\mathbb{N}} |(f, \phi_k)|^2.$$
(3) *If $f \in H$ and $(f, \phi_k) = 0$ for all $k \in \mathbb{N}$, then $f \equiv 0$.*

Choosing $n = 1$ and $\Omega = (0, \pi)$ in Theorem 1.2.1, we know that the sequence $\{\phi_k\}_{k\in\mathbb{N}}$ defined by

$$\phi_k(x) = \sqrt{\frac{2}{\pi}} \sin(kx) \quad \text{for } k = 1, 2, \cdots$$

is an orthonormal basis of $L^2(0, \pi)$. In particular, we compute that

$$\begin{aligned} \frac{2}{\pi} \int_0^{\pi} \sin(kx)\sin(kx)\, dx &= \frac{1}{\pi} \int_0^{\pi} \big(1 - \cos(2kx)\big)\, dx \\ &= 1 - \frac{1}{2k} \sin(2kx)\bigg|_{x=0}^{\pi} = 1 \end{aligned}$$

and for each $k_1 \neq k_2$ we have

$$\begin{aligned} &\frac{2}{\pi} \int_0^{\pi} \sin(k_1 x)\sin(k_2 x)\, dx \\ &= \frac{1}{\pi} \int_0^{\pi} \big(\cos((k_1 - k_2)x) - \cos((k_1 + k_2)x)\big)\, dx \\ &= \frac{1}{\pi} \left(\frac{1}{k_1 - k_2} \sin((k_1 - k_2)x) - \frac{1}{k_1 + k_2} \sin((k_1 + k_2)x) \right)\bigg|_{x=0}^{x=\pi} = 0. \end{aligned}$$

Hence given any real-valued $f \in L^2(0, \pi)$, we can write

$$f(x) = \sum_{k=1}^{\infty} a_k \sin(kx) \quad \text{converges in } L^2(0, \pi), \tag{1.2.2}$$

with

$$a_k = \frac{2}{\pi} \int_0^\pi \sin(kx) f(x)\, dx.$$

The expansion (1.2.2) is called the *Fourier sine expansion*. In fact, we can do the similar things for Laplacian with Neumann eigenvalues: The sequence $\{\tilde{\phi}_k\}_{k \in \mathbb{N} \cup \{0\}}$ defined by

$$\tilde{\phi}_k(x) = \sqrt{\frac{2}{\pi}} \cos(kx) \quad \text{for } k = 0, 1, 2, \cdots$$

is also an orthonormal basis of $L^2(0,\pi)$, and similar idea induces *Fourier cosine expansion*

$$f(x) = \frac{1}{2} b_0 + \sum_{k=1}^\infty b_k \cos(kx) \quad \text{converges in } L^2(0,\pi), \tag{1.2.3}$$

with

$$b_k = \frac{2}{\pi} \int_0^\pi \cos(kx) f(x)\, dx$$

see [**Bre11**, Comments on Chapter 5]. Indeed, we also compute that

$$\begin{aligned} &\frac{2}{\pi} \int_0^\pi \cos(kx)\cos(kx)\, dx = \frac{1}{\pi} \int_0^\pi \big(1 + \cos(2kx)\big)\, dx \\ &= 1 - \frac{1}{2k} \sin(2kx) \Big|_{x=0}^{x=\pi} = 1 \end{aligned}$$

and for each $k_1 \neq k_2$ we have

$$\begin{aligned} &\frac{2}{\pi} \int_0^\pi \sin(k_1 x) \sin(k_2 x)\, dx \\ &= \frac{1}{\pi} \int_0^\pi \big(\cos((k_1 - k_2)x) + \cos((k_1 + k_2)x) \big)\, dx \\ &= \frac{1}{\pi} \left(\frac{1}{k_1 - k_2} \sin((k_1 - k_2)x) + \frac{1}{k_1 + k_2} \sin((k_1 + k_2)x) \right) \Big|_{x=0}^{x=\pi} = 0. \end{aligned}$$

EXAMPLE 1.2.3. Let $f(x) = 1$ in the interval $(0,\pi)$. The function has a Fourier sine series with coefficients

$$\begin{aligned} a_k &= \frac{2}{\pi} \int_0^\pi \sin(kx)\, dx = -\frac{2}{\pi k} \cos(kx) \Big|_{x=0}^{x=\pi} \\ &= \frac{2}{\pi k}(1 - \cos k\pi) = \frac{2}{\pi k}(1 - (-1)^m), \end{aligned}$$

which in particular gives

$$a_k = \begin{cases} \frac{4}{k\pi} & \text{if } k \text{ is odd,} \\ 0 & \text{if } k \text{ is even.} \end{cases}$$

Thus

$$1 = \frac{4}{\pi} \sum_{m \in \mathbb{N}} \frac{1}{2m-1} \sin\big((2m-1)x\big) \quad \text{converges in } L^2(0,\pi).$$

The same function has a Fourier cosine series with coefficients

$$b_k = \frac{2}{\pi}\int_0^\pi \cos(kx)\,dx = \frac{2}{\pi k}\sin(kx)\Big|_{x=0}^{x=\pi} = 0 \quad \text{for all } k \neq 0$$

$$b_0 = \frac{2}{\pi}\int_0^\pi 1\,dx = 2.$$

This shows that the Fourier cosine series of this function is trivial.

EXAMPLE 1.2.4. Let $f(x) = x$ in $(0,\pi)$. Its Fourier sine series has the coefficients

$$\begin{aligned} a_k &= \frac{2}{\pi}\int_0^\pi x\sin kx\,dx = \left(-\frac{2x}{k\pi}\cos kx + \frac{2}{k^2\pi}\sin kx\right)\Big|_{x=0}^{x=\pi} \\ &= -\frac{2}{k}\cos k\pi = (-1)^{k+1}\frac{2}{k}. \end{aligned}$$

Thus in $(0,\pi)$ we have

$$x = 2\sum_{k=1}^{\infty}(-1)^{k+1}\frac{\sin(kx)}{k} \quad \text{converges in } L^2(0,\pi).$$

Its Fourier cosine series has the coefficients $b_0 = \frac{2}{\pi}\int_0^\pi x\,dx = \pi$ and

$$\begin{aligned} b_k &= \frac{2}{\pi}\int_0^\pi x\cos kx\,dx = \left(\frac{2x}{k\pi}\sin kx + \frac{2}{k^2\pi}\cos kx\right)\Big|_{x=0}^{x=\pi} \\ &= \frac{2}{k}\sin k\pi + \frac{2}{k^2\pi}(\cos k\pi - 1) = \frac{2}{k\pi}((-1)^k - 1), \end{aligned}$$

which gives

$$b_k = \begin{cases} -\frac{4}{k^2\pi} & \text{for } k \text{ odd}, \\ 0 & \text{for } k \neq 0 \text{ even}. \end{cases}$$

Thus in $(0,\pi)$ we have

$$x = \frac{\pi}{2} - \frac{4}{\pi}\sum_{m\in\mathbb{N}}\frac{1}{(2m-1)^2}\cos((2m-1)x) \quad \text{converges in } L^2(0,\pi).$$

Given any function $f : \mathbb{R} \to \mathbb{R}$, we can decompose it into the sum of even function and odd function by the following simple observation:

$$f(x) = \overbrace{\frac{f(x) - f(-x)}{2}}^{\text{odd function}} + \overbrace{\frac{f(x) + f(-x)}{2}}^{\text{even function}}. \tag{1.2.4}$$

Moreover, the decomposition (1.2.4) is unique: If $f(x) = f_{\text{odd}}(x) + f_{\text{even}}(x)$ for some odd function f_{odd} and even function f_{even}, then from (1.2.4) we can write

$$\overbrace{f_{\text{odd}}(x) - \frac{f(x) - f(-x)}{2}}^{\text{odd function}} = \overbrace{\frac{f(x) + f(-x)}{2} - f_{\text{even}}(x)}^{\text{even function}} \quad \text{for all } x \in \mathbb{R}.$$

This implies

$$f_{\text{odd}}(x) = \frac{f(x) - f(-x)}{2} \quad \text{and} \quad f_{\text{even}}(x) = \frac{f(x) + f(-x)}{2}.$$

We see that $\sin(kx)$ are odd functions, while $\cos(kx)$ are even functions. Therefore, it is natural to represent $f:(-\pi,\pi)\to\mathbb{C}$ by the Fourier series

$$f(x)=\frac{1}{2}B_0+\sum_{k=1}^{\infty}(\overbrace{A_k\sin(kx)}^{\text{odd part}}+\overbrace{B_k\cos(kx)}^{\text{even part}})\quad\text{for } x\in(-\pi,\pi) \tag{1.2.5}$$

in a suitable sense. In next subsection we will show that the series (1.2.5) as well as (1.2.6) are converges in $L^2(-\pi,\pi)$. In addition, if f is sufficiently smooth, in particular the convergence is point-wise. Using similar computations, the coefficients are given by

$$\begin{aligned}A_k&=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin(kx)\,dx\quad\text{for } k=1,2,\cdots\\ B_k&=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos(kx)\,dx\quad\text{for } k=0,1,2,\cdots\end{aligned}$$

Since $e^{i\theta}=\cos\theta+i\sin\theta$ for all $\theta\in\mathbb{R}$, we may alternatively consider the series

$$f(x)=\sum_{k=-\infty}^{\infty}c_ke^{ikx}\quad\text{for } x\in(-\pi,\pi) \tag{1.2.6}$$

with $c_k\in\mathbb{C}$. We can write (1.2.6) as

$$\begin{aligned}f(x)&=\sum_{k=-\infty}^{\infty}c_ke^{ikx}=c_0+\sum_{k=1}^{\infty}(c_ke^{ikx}+c_{-k}e^{-ikx})\\
&=c_0+\sum_{k=1}^{\infty}(\Re c_k+i\Im c_k)\big(\cos(kx)+i\sin(kx)\big)\\
&\quad+\sum_{k=1}^{\infty}(\Re c_{-k}+i\Im c_{-k})\big(\cos(kx)-i\sin(kx))\big)\\
&=c_0+\sum_{k=1}^{\infty}\big(\Re c_k\cos(kx)-\Im c_k\sin(kx)\big)+i\big(\Re c_k\sin(kx)+\Im c_k\cos(kx)\big)\\
&\quad+\sum_{k=1}^{\infty}\big(\Re c_{-k}\cos(kx)+\Im c_{-k}\sin(kx)\big)+i\big(-\Re c_{-k}\sin(kx)+\Im c_{-k}\cos(kx)\big)\\
&=c_0+\sum_{k=1}^{\infty}\big(-\Im c_k+\Im c_{-k}+i(\Re c_k-\Re c_{-k})\big)\sin(kx)\\
&+\sum_{k=1}^{\infty}\big(\Re c_k+\Re c_{-k}+i(\Im c_k+\Im c_{-k})\big)\cos(kx)\\
&=c_0+\sum_{k=1}^{\infty}i(c_k-c_{-k})\sin(kx)+\sum_{k=1}^{\infty}(c_k+c_{-k})\cos(kx).\end{aligned}$$

Therefore from (1.2.5) (the coefficient are unique) we have

$$c_0=\frac{1}{2}B_0,\quad c_k+c_{-k}=B_k\text{ and } i(c_k-c_{-k})=A_k\quad\text{for all } k\in\mathbb{N}.$$

Equivalently,

$$c_0 = \frac{1}{2}B_0, \quad c_k = \frac{1}{2}(B_k - iA_k) \text{ and } c_{-k} = \frac{1}{2}(B_k + iA_k) \quad \text{for all } k \in \mathbb{N}. \tag{1.2.7}$$

In particular, (1.2.7) is equivalent to

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx \quad \text{for all } k \in \mathbb{Z}.$$

REMARK 1.2.5. If f is real-valued, then

$$\overline{c_{-k}} = \overline{\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{ikx}\,dx} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx = c_k \quad \text{for all } k \in \mathbb{Z}.$$

Conversely, if $\overline{c_{-k}} = c_k$ for all $k \in \mathbb{Z}$, then from (1.2.6) we have

$$\overline{f(x)} = \sum_{k=-\infty}^{\infty} \overline{c_k e^{ikx}} = \sum_{k=-\infty}^{\infty} \overline{c_k} e^{-ikx} = \sum_{k=-\infty}^{\infty} \overline{c_{-k}} e^{ikx} = \sum_{k=-\infty}^{\infty} c_k e^{ikx} = f(x),$$

that is, f is real-valued.

EXAMPLE 1.2.6. Let $f(x) = x$ in the interval $(-\pi,\pi)$. Its full Fourier series has the coefficients

$$B_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} x\,dx = 0,$$

$$\begin{aligned} B_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} x\cos(kx)\,dx = \left(\frac{x}{k\pi}\sin(kx) + \frac{1}{k^2\pi}\cos(kx)\right)\bigg|_{x=-\pi}^{x=\pi} \\ &= \frac{1}{k^2\pi}(\cos(k\pi) - \cos(-k\pi)) = 0 \end{aligned}$$

and

$$\begin{aligned} A_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} x\sin(kx)\,dx = \left(-\frac{x}{k\pi}\cos(kx) + \frac{1}{k^2\pi}\sin(kx)\right)\bigg|_{x=-\pi}^{x=\pi} \\ &= -\frac{1}{k}\cos k\pi - \frac{1}{k}\cos(-k\pi) = (-1)^{k+1}\frac{2}{k}. \end{aligned}$$

This gives us exactly the same series as in Example 1.2.4, except that it is supposed to be valid in $(-\pi,\pi)$. Since f is a odd function, therefore the even part of (1.2.5) should vanishes.

1.3. n-dimensional Fourier series in L^2

The ideas for multi-variable case is also similar: If $f : \mathbb{R}^n \to \mathbb{C}$ is of 2π-periodic in each variable, we want to represent it by the Fourier series

$$f(x) = \sum_{k\in\mathbb{Z}^n} c_k e^{ik\cdot x} = \sum_{k\in\mathbb{Z}^n} c_k e^{ik_1x_1}\cdots e^{ik_2x_2} \quad \text{for all } x = (x_1,\cdots,x_n) \in \mathbb{R}^n \tag{1.3.1}$$

in some suitable sense. We now consider the cube $Q = [-\pi,\pi]^n$, and normalize the inner product on $L^2(Q)$ by

$$(f,g) \equiv (f,g)_{L^2(Q)} := \frac{1}{|Q|}\int_Q f\overline{g}\,dx \equiv \fint_Q f\overline{g}\,dx \quad \text{for all } f,g \in L^2(Q).$$

with $|Q| = (2\pi)^n$.

LEMMA 1.3.1. *The countable set $\{e^{ik\cdot x}\}_{k\in\mathbb{Z}^n}$ is an orthonormal subset of $L^2(Q)$.*

PROOF. For each $k,\ell\in\mathbb{Z}^n$, using Fubini's theorem we see that

$$
\begin{aligned}
(e^{ik\cdot x},e^{i\ell\cdot x}) &= (2\pi)^{-n}\int_Q e^{i(k-\ell)\cdot x}\,dx \\
&= (2\pi)^{-n}\prod_{j=1}^n\int_{-\pi}^{\pi} e^{i(k_j-\ell_j)x_j}\,dx_j = \begin{cases}1 & ,k=\ell,\\ 0 & ,k\neq\ell,\end{cases}
\end{aligned}
$$

which conclude the lemma. □

We will put much effort to prove the following proposition:

PROPOSITION 1.3.2. $\{e^{ik\cdot x}\}_{k\in\mathbb{Z}^n}$ *a* complete orthonormal basis *of* $L^2(Q)$.

REMARK 1.3.3. Given any $f\in L^2(Q)$ be such that $(f,e^{ik\cdot x})=0$ for all $k\in\mathbb{Z}^n$. If we can prove $f=0$ a.e. in Q, using Proposition 1.2.2 we conclude Proposition 1.3.2. Since $e^{ik\cdot x}=e^{ik_1x_1}\cdots e^{ik_nx_n}$, We see that

$$
0=\int_{[-\pi,\pi]^n} fe^{-ik\cdot x}\,dx=\int_{-\pi}^{\pi}\left(\int_{[-\pi,\pi]^{n-1}} f(x)e^{-ik'\cdot x'}\,dx'\right)e^{-ik_1x_1}\,dx_1 \quad \text{for all } k\in\mathbb{Z}^n, \tag{1.3.2}
$$

with $k'=(k_2,\cdots,k_n)$ and $x'=(x_2,\cdots,x_n)$. If we can show Proposition 1.3.2 for the case when $n=1$, using Proposition 1.2.2 we know that

$$
0=\int_{[-\pi,\pi]^{n-1}} f(x)e^{-ik'\cdot x'}\,dx' \quad \text{for all } k'\in\mathbb{Z}^{n-1}. \tag{1.3.3}
$$

Repeating the arguments that proving from (1.3.2) to (1.3.3), we conclude that $f=0$ a.e. in Q.

PROOF OF PROPOSITION 1.3.2 USING SPECTRAL THEORY. We decompose

$$
f(x)=f_{\rm odd}(x)+f_{\rm even}(x)
$$

for some odd function $f_{\rm odd}$ and even function $f_{\rm even}$ as in (1.2.4). Then Proposition 1.3.2 immediately follows by approximate $f_{\rm odd}$ using the Fourier sine series, while approximate $f_{\rm even}$ using the Fourier cosine series, as stated in Section 1.2. □

We will exhibit the classical proof of Proposition 1.3.2, which involving Dirichlet kernel, later. We are now ready to prove the main result of this section.

THEOREM 1.3.4 (Fourier series of L^2 functions). *If $f\in L^2(Q)$, then one has the Fourier series*

$$
f(x)=\sum_{k\in\mathbb{Z}^n}\hat{f}(k)e^{ik\cdot x} \text{ converges in } L^2(Q), \tag{1.3.4}
$$

with the Fourier coefficients

$$
\hat{f}(k)=(f,e^{ik\cdot x})=\fint_Q f(x)e^{-ik\cdot x}\,dx. \tag{1.3.5}
$$

One has the Parseval identity

$$
\|f\|_{L^2(Q)}^2=\sum_{k\in\mathbb{Z}^n}|\hat{f}(k)|^2.
$$

Conversely, if $c = (c_k) \in \ell^2(\mathbb{Z}^n)$, then the series $\sum_{k\in\mathbb{Z}^n} c_k e^{ik\cdot x}$ converges in $L^2(Q)$ to some $f \in L^2(Q)$ and it is necessarily $c_k = \hat{f}(k)$.

REMARK 1.3.5. Here we denote by $\ell^2(\mathbb{Z}^n)$ the space of the complex sequences $c = (c_k)_{k\in\mathbb{Z}^n}$ with norm

$$\|c\|_{\ell^2(\mathbb{Z}^n)} = \left(\sum_{k\in\mathbb{Z}^n} |c_k|^2\right)^{\frac{1}{2}}.$$

Theorem 1.3.4 says that there is a 1-1 corresponding between the elements in $L^2(Q)$ with the elements in $\ell^2(\mathbb{Z}^n)$. In other words, (1.3.5) can be viewed as the *discrete Fourier transform*, and the *inverse discrete Fourier transform* is given by the formula (1.3.4).

PROOF OF THEOREM 1.3.4. The first part of Theorem 1.3.4 is an immediate consequence of Proposition 1.3.2. For the converse, if $(c_k) \in \ell^2(\mathbb{Z}^n)$, then we see that the partial sum

$$S_N := \sum_{|k|\leq N} c_k e^{ik\cdot x}$$

is a Cauchy sequence in $L^2(Q)$. Since $L^2(Q)$ is complete, then we know that the series $\sum_{k\in\mathbb{Z}^n} c_k e^{ik\cdot x}$ converges in $L^2(Q)$ to some $f \in L^2(Q)$. For each $N \geq k$, we also see that

$$|c_k - \hat{f}(k)| = |(S_N - f, e^{ikx})| \leq C_n \|e^{ikx}\|_{L^2(Q)} \|S_N - f\|_{L^2(Q)} \to 0 \quad \text{as } N \to \infty,$$

which conclude $c_k = \hat{f}(k)$. □

We now give a direct proof to Proposition 1.3.2 (for the case when $n = 1$). We wish to construction a special kernel as follows:

DEFINITION 1.3.6. A sequence $\{Q_N(x)\}_{N\in\mathbb{N}}$ of 2π-period continuous functions on the real line is called an *approximate identity* if

(1) $Q_N \geq 0$ for all $N \in \mathbb{N}$,
(2) $⨍_{-\pi}^{\pi} Q_N(x)\,dx \equiv (2\pi)^{-1}\int_{-\pi}^{\pi} Q_N(x)\,dx = 1$ for all $N \in \mathbb{N}$, and
(3) for each $0 < \epsilon < \pi$ one has $\lim\limits_{N\to\infty} \sup\limits_{\epsilon\leq|x|\leq\pi} Q_N(x) = 0$.

We now prove the existence of such function satisfies Definition 1.3.6.

LEMMA 1.3.7. *The sequence*

$$Q_N(x) := c_N \left(\frac{1+\cos x}{2}\right)^N \quad \text{with} \quad c_N = 2\pi\left(\int_{-\pi}^{\pi} \left(\frac{1+\cos x}{2}\right)^N dx\right)^{-1}$$

is an approximate identity.

PROOF. It is easy to see that $Q_N \geq 0$ and $(2\pi)^{-1}\int_{-\pi}^{\pi} Q_N(x)\,dx = 1$ for all $N \in \mathbb{N}$. We estimate the constant c_N as followings:

$$\begin{aligned} 1 &= \frac{c_N}{2\pi}\int_{-\pi}^{\pi}\left(\frac{1+\cos x}{2}\right)^N dx = \frac{c_N}{\pi}\int_0^{\pi}\left(\frac{1+\cos x}{2}\right)^N dx \\ &\geq \frac{c_N}{\pi}\int_0^{\pi}\left(\frac{1+\cos x}{2}\right)^N \sin x\,dx \\ &= \frac{c_N}{\pi}\int_{-1}^{1}\left(\frac{1+t}{2}\right)^N dt = \frac{2c_N}{\pi}\int_0^1 s^N\,ds = \frac{2c_N}{\pi(N+1)}. \end{aligned}$$

Thus for each $0 < \epsilon < \pi$ we have

$$0 \le \sup_{\epsilon \le |x| \le \pi} Q_N(x) \le Q_N(\epsilon) = c_N \left(\frac{1+\cos\epsilon}{2}\right)^N \le \frac{\pi(N+1)}{2}\left(\frac{1+\cos\epsilon}{2}\right)^N \to 0 \quad \text{as } N \to \infty,$$

because $0 < \frac{1+\cos\epsilon}{2} < 1$. □

Let f and g be two 2π-period functions. Then we formally define the convolution $f * g$ by

$$(f * g)(x) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)g(x-y)\,dy \equiv \frac{1}{2\pi}\int_{-\pi}^{\pi} g(y)f(x-y)\,dy, \tag{1.3.6}$$

which is also a 2π-periodic function (the second identity holds only for 2π-period functions). The following lemma explains the naming of Definition 1.3.6.

Lemma 1.3.8. *Let Q_N be an approximate identity as in Definition 1.3.6 and let f be a 2π-periodic function. If f is continuous, then*

$$\lim_{N\to\infty} Q_N * f = f \quad \text{converges in } L^\infty(-\pi,\pi).$$

If $f \in L^p(-\pi,\pi)$ for some $1 \le p < \infty$, then

$$\lim_{N\to\infty} Q_N * f = f \quad \text{converges in } L^p(-\pi,\pi).$$

Proof of Lemma 1.3.8. We first observe that

$$(Q_N * f - f)(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} Q_N(y)\big(f(x-y) - f(x)\big)\,dy.$$

Case 1: f is a continuous 2π-periodic function. Given any $\epsilon > 0$, there exists $\delta(\epsilon) > 0$ such that

$$\sup_{|y|\le\delta(\epsilon)} \big|f(x-y) - f(x)\big| \le \epsilon \quad \text{for all } x \in \mathbb{R}$$

and

$$\sup_{\delta(\epsilon)\le|x|\le\pi} Q_N(x) \le \epsilon \quad \text{for all sufficiently large } N. \tag{1.3.7}$$

Then for all sufficiently large N we estimate

$$\begin{aligned}
|(Q_N * f - f)(x)| &\le \frac{1}{2\pi}\left(\int_{|y|\le\delta(\epsilon)} + \int_{\delta(\epsilon)\le|y|\le\pi}\right) Q_N(y)\big|f(x-y) - f(x)\big|\,dy \\
&\le \frac{\epsilon}{2\pi}\Bigg(\overbrace{\int_{|y|\le\delta(\epsilon)} Q_N(y)\,dy}^{\le 2\pi} + \overbrace{\int_{\delta(\epsilon)\le|y|\le\pi} \big|f(x-y) - f(x)\big|\,dy}^{\le 4\pi\|f\|_{L^\infty(\mathbb{R})}}\Bigg) \\
&\le \epsilon(1 + 2\|f\|_{L^\infty(\mathbb{R})}),
\end{aligned}$$

which gives

$$\limsup_{N\to\infty} |(Q_N * f - f)(x)| \le \epsilon(1 + 2\|f\|_{L^\infty(\mathbb{R})}).$$

By arbitrariness of $\epsilon > 0$, we conclude the first part of the lemma.

Case 2: $f \in L^p(-\pi,\pi)$ for some $1 \le p < \infty$. Using the Mikowski's inequality, we estimate

$$\begin{aligned}
\|Q_N * f - f\|_{L^p(-\pi,\pi)} &\le \frac{1}{2\pi}\left(\int_{-\pi}^{\pi}\left|\int_{-\pi}^{\pi} Q_N(y)\big(f(x-y)-f(x)\big)\,dy\right|^p dx\right)^{\frac{1}{p}} \\
&\le \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\int_{-\pi}^{\pi}\left|Q_N(y)\big(f(x-y)-f(x)\big)\right|^p dx\right)^{\frac{1}{p}} dy \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} Q_N(y)\|f(\cdot - y) - f\|_{L^p(-\pi,\pi)}\,dy.
\end{aligned}$$

Since $f \in L^p(-\pi,\pi)$, by approximate it by $C_c^\infty(-\pi,\pi)$ functions, given any $\epsilon > 0$, there exists $\delta(\epsilon) > 0$ such that

$$\sup_{|y|\le\delta(\epsilon)} \|f(\cdot - y) - f\|_{L^p(-\pi,\pi)} \le \epsilon.$$

Again using (1.3.7), then for all sufficiently large N we estimate

$$\begin{aligned}
\|Q_N * f - f\|_{L^p(-\pi,\pi)} &\le \frac{1}{2\pi}\left(\int_{|y|\le\delta(\epsilon)} + \int_{\delta(\epsilon)\le|y|\le\pi}\right) Q_N(y)\|f(\cdot - y) - f\|_{L^p(-\pi,\pi)}\,dy \\
&\le \frac{\epsilon}{2\pi}\left(\overbrace{\int_{|y|\le\delta(\epsilon)} Q_N(y)\,dy}^{\le 2\pi} + \overbrace{\int_{\delta(\epsilon)\le|y|\le\pi} \|f(\cdot - y) - f\|_{L^p(-\pi,\pi)}\,dy}^{\le 4\pi\|f\|_{L^p(-\pi,\pi)}}\right) \\
&\le \epsilon(1 + 2\|f\|_{L^p(\mathbb{R})}),
\end{aligned}$$

and we prove the second part of the lemma similar as in first part. $\square$

We are now ready to give a direct proof to Proposition 1.3.2.

Proof of Proposition 1.3.2 using Definition 1.3.6. As mentioned in Remark 1.3.3, it is suffice to prove the case when $n = 1$. Let $f \in L^2(-\pi,\pi)$ be such that

$$(f, e^{ikx}) = 0 \quad \text{for all } k \in \mathbb{Z}. \tag{1.3.8}$$

Using Proposition 1.2.2, it is suffice to show $f \equiv 0$ a.e. Using Lemma 1.3.7, in particular from (1.3.8) we have $Q_N * f = 0$ for all $N \in \mathbb{N}$. Since $Q_N * f \to f$ as $N \to \infty$ in $L^2(-\pi,\pi)$, we conclude our result. $\square$

1.4. Pointwise and uniform convergence

Although the convergence of Fourier series in other sense is not the main topic of this course, it may be of interest to mention a few classcial results. In order to simplify the analysis, here we only consider the case when $n = 1$. The convergence results in this section may not optimal, see e.g. [**Zyg02**, Chapter II] for more optimized results.

We first study the pointwise convergence of the Fourier series:

Theorem 1.4.1. *Let f be a 2π-periodic function in $\mathbb{R}$ which is of bounded variation. The Fourier coefficients are given by*

$$\hat{f}(k) = (f, e^{ikx}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx = ⨍_{-\pi}^{\pi} f(x)e^{-ikx}\,dx.$$

Then the (1-dimensional) Fourier series $\sum_{k\in\mathbb{Z}} \hat{f}(k)e^{ikx}$ *of* f *converges to*

$$\frac{1}{2}\big(f(x+) + f(x-)\big) \text{ pointwsely,}$$

where $f(x\pm) := \lim_{\theta\to 0_+} f(x\pm\theta)$.

Theorem 1.4.1 is a consequence of the Dirichlet-Jordan test [**Zyg02**, Theorem (8.1)(i), Section 8, Chapter II][1]. Rather explaining the precise definition of the term "bounded variation", we note that a piecewise continuous function also has bounded variation. For simplicity, here we only proof Theorem 1.4.1 for the special case when f is piecewise C^1 as in [**Str08**, Theorem 5.4.4∞].

Proof of Theorem 1.4.1 for piecewise C^1 functions. The partial sum of the Fourier series of a function $f \in L^1(-\pi,\pi)$ extended as a 2π-periodic function into $\mathbb{R}$, are given by

$$S_m f(x) := \sum_{k=-m}^{m} \hat{f}(k)e^{ikx} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y) \overbrace{\left(\sum_{k=-m}^{m} e^{ik(x-y)}\right)}^{=:D_m(x-y)} dy = (D_m * f)(x), \tag{1.4.1}$$

where the convolution is given in (1.3.6). The *Dirichlet kernel* $D_m(z)$ can be represented by

$$\begin{aligned} D_m(z) &= \sum_{k=-m}^{m} e^{ikz} = e^{-imz}\sum_{k=0}^{2m} e^{ikz} = e^{-imz}\frac{e^{i(2m+1)z}-1}{e^{iz}-1} \\ &= \frac{e^{i(m+\frac{1}{2})z} - e^{-i(m+\frac{1}{2})z}}{e^{\frac{i}{2}z} - e^{-\frac{i}{2}z}} = \frac{\sin((m+\frac{1}{2})z)}{\sin(\frac{1}{2}z)}. \end{aligned}$$

Since $\int_{-\pi}^{\pi} e^{ikz}\,dz = 0$ for all $k \neq 0$, then it is easy to verify that

$$⨍_{-\pi}^{\pi} D_m(z)\,dz = 1 \quad \text{for all } m \in \mathbb{N}.$$

Then we have

$$\begin{aligned} & S_m f(x) - \frac{1}{2}\big(f(x+) + f(x-)\big) \\ &= \frac{1}{2\pi}\int_0^{\pi} D_m(\theta)\big(f(x+\theta) - f(x+)\big)\,d\theta + \frac{1}{2\pi}\int_{-\pi}^{0} D_m(\theta)\big(f(x+\theta) - f(x-)\big)\,d\theta \\ &= \frac{1}{2\pi}\int_0^{\pi} g_+(\theta)h_m(\theta)\,d\theta + \frac{1}{2\pi}\int_{-\pi}^{0} g_-(\theta)h_m(\theta)\,d\theta \end{aligned} \tag{1.4.2}$$

with

$$g_\pm(\theta) = \frac{f(x+\theta) - f(x\pm)}{\sin(\frac{1}{2}\theta)} \quad \text{and} \quad h_m(\theta) = \sin\left(\left(m + \frac{1}{2}\right)\theta\right).$$

[1] I would like to thank Aleksi Harmoinen for pointing this out to me.

Using the mean value theorem, the fact $\lim_{\theta\to 0}\frac{\frac{1}{2}\theta}{\sin(\frac{1}{2}\theta)} = 1$ and the piecewise C^1 assumption on f, we see that

$$\begin{aligned}
\limsup_{\theta\to 0_+} |g_+(\theta)| &= 2\limsup_{\theta\to 0_+}\left|\frac{f(x+\theta)-f(x+)}{\theta}\right| < \infty,\\
\limsup_{\theta\to 0_-} |g_-(\theta)| &= 2\limsup_{\theta\to 0_-}\left|\frac{f(x+\theta)-f(x-)}{\theta}\right| < \infty,
\end{aligned}$$

then we see that $g_\pm$ is bounded in $(0,\pi)$ and $(-\pi,0)$ respectively. Since $\{h_m\}_{m\in\mathbb{N}}$ form an orthogonal set in the interval $(0,\pi)$ and $(-\pi,0)$, then we have the Bessel's inequality

$$\begin{aligned}
\frac{1}{\pi}\sum_{m=1}^{\infty}\left|\int_0^\pi g_+(\theta)h_m(\theta)\,d\theta\right|^2 &= \sum_{m=1}^{\infty}\frac{|\int_0^\pi g_+(\theta)h_m(\theta)\,d\theta|^2}{\int_0^\pi |h_m(\theta)|^2\,d\theta} \le \int_0^\pi |g_+(\theta)|^2\,d\theta < \infty,\\
\frac{1}{\pi}\sum_{m=1}^{\infty}\left|\int_{-\pi}^0 g_-(\theta)h_m(\theta)\,d\theta\right|^2 &= \sum_{m=1}^{\infty}\frac{|\int_{-\pi}^0 g_-(\theta)h_m(\theta)\,d\theta|^2}{\int_{-\pi}^0 |h_m(\theta)|^2\,d\theta} \le \int_{-\pi}^0 |g_-(\theta)|^2\,d\theta < \infty.
\end{aligned}$$

This implies

$$\begin{aligned}
&\left|S_m f(x) - \frac{1}{2}\big(f(x+)+f(x-)\big)\right|\\
&\le \frac{1}{2\pi}\left|\int_0^\pi g_+(\theta)h_m(\theta)\,d\theta\right| + \frac{1}{2\pi}\left|\int_{-\pi}^0 g_-(\theta)h_m(\theta)\,d\theta\right|\\
&\to 0 \quad \text{as } m\to\infty,
\end{aligned}$$

which complete the proof of Theorem 1.4.1. $\square$

From Theorem 1.4.1, we know that if f has a jump, then the Fourier series of f never converges to f uniformly. The following theorem, which is a consequence of the Dirichlet-Jordan test [**Zyg02**, Theorem (8.1)(ii), Section 8, Chapter II], shows that if the Fourier series of a continuous function converges uniformly:

THEOREM 1.4.2. *Let f be a continuous 2π-periodic function. The Fourier coefficients are given by*

$$\hat{f}(k) = (f, e^{ikx}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx = ⨍_{-\pi}^{\pi} f(x)e^{-ikx}\,dx.$$

Then the (1-dimensional) Fourier series $\sum_{k\in\mathbb{Z}}\hat{f}(k)e^{ikx}$ of f converges to f uniformly in $\mathbb{R}$.

EXERCISE 1.4.3. Prove Theorem 1.4.2 for the special case when $f\in C^1(\mathbb{R})$. (Hint: Compute the Fourier coefficients of f'.)

1.5. Pointwise convergence: Dini's criterion

We now want to prove a pointwise convergence result as in Mikko Salo's lecture note. We begin our discussions from the following fundamental result due to Riemann and Lebesgue:

LEMMA 1.5.1 (Riemann-Lebesgue). *If $f\in L^1(-\pi,\pi)$, then $\hat{f}(k)\to 0$ as $k\to\pm\infty$.*

Proof. Since both f and e^{-ikx} are periodic, we have

$$2\pi\hat{f}(k) = \overbrace{\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx}^{(1)} = \int_{-\pi}^{\pi} f\Big(x-\frac{\pi}{k}\Big)e^{-ik(x-\frac{\pi}{k})}\,dx = -\overbrace{\int_{-\pi}^{\pi} f\Big(x-\frac{\pi}{k}\Big)e^{-ikx}\,dx}^{(2)}.$$

Then

$$\begin{aligned}2\pi\hat{f}(k) &= \frac{1}{2}\Big(\overbrace{\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx}^{\text{using (1)}} - \overbrace{\int_{-\pi}^{\pi} f\Big(x-\frac{\pi}{k}\Big)e^{-ikx}\,dx}^{\text{using (2)}}\Big)\\ &= \frac{1}{2}\int_{-\pi}^{\pi}\Big[f(x)-f\Big(x-\frac{\pi}{k}\Big)\Big]e^{-ikx}\,dx.\end{aligned}$$

Since $f \in L^1(-\pi,\pi)$, given any $\epsilon > 0$, we choose a continuous periodic function g with $\|f-g\|_{L^1(-\pi,\pi)} \le \epsilon$. Then we see that

$$|\hat{f}(k)| \le |(f-g)^\wedge(k)| + |\hat{g}(k)| \le \epsilon + |\hat{g}(k)|.$$

By (uniform) continuity of g, we see that

$$\lim_{k\to\infty} 2\pi\hat{g}(k) = \frac{1}{2}\lim_{k\to\infty}\int_{-\pi}^{\pi}\Big[g(x)-g\Big(x-\frac{\pi}{k}\Big)\Big]e^{-ikx}\,dx = 0,$$

therefore by arbitrariness of $\epsilon > 0$, we conclude the lemma. □

Lemma 1.5.2 (Riemann localization principle). *For each $\delta > 0$ we have*

$$\lim_{m\to\infty}\int_{\delta<|y|<\pi} D_m(\theta)\big(f(x+\theta)-f(x)\big)\,d\theta = 0.$$

In particular, if $f \in L^1(-\pi,\pi)$ satisfies $f = 0$ near x (i.e. $f = 0$ a.e. in $(x-\delta, x+\delta)$ for some $\delta > 0$), then $\lim_{m\to\infty} S_m f(x) = \lim_{m\to\infty}\frac{1}{2\pi}\int_{\delta<|y|<\pi} D_m(\theta)\big(f(x+\theta)-f(x)\big)\,d\theta = 0$.

Proof. Using (1.4.2) we have

$$\int_{\delta<|y|<\pi} D_m(\theta)\big(f(x+\theta)-f(x)\big)\,d\theta = \int_{-\pi}^{\pi} g(y)h_m(y)\,dy$$

with

$$g(\theta) = \frac{f(x-y)}{\sin(\frac{1}{2}y)}\chi_{\{\delta<|y|<\pi\}} \quad \text{and} \quad h_m(\theta) = \sin\Big(\Big(m+\frac{1}{2}\Big)\theta\Big).$$

Since $g \in L^1(-\pi,\pi)$ and $\sin t = \frac{e^{it}-e^{-it}}{2i}$, we have

$$\frac{1}{2\pi}\int_{\delta<|y|<\pi} D_m(\theta)\big(f(x+\theta)-f(x)\big)\,d\theta = -\Big(\frac{e^{-i\frac{y}{2}}g}{2i}\Big)^\wedge(m) + \Big(\frac{e^{i\frac{y}{2}}g}{2i}\Big)^\wedge(m).$$

Hence the Riemann-Lebesgue lemma (Lemma 1.5.1) concludes Lemma 1.5.2. □

By assuming something slightly stronger than continuity, pointwise convergence holds:

Theorem 1.5.3 (Dini's criterion). *If $f \in L^1(-\pi,\pi)$ and let x be a point such that*

$$\int_{|y|<\delta}\left|\frac{f(x+y)-f(x)}{y}\right|dy < \infty \quad \text{for some } \delta > 0, \tag{1.5.1}$$

then $\lim_{m\to\infty} S_m f(x) = f(x)$.

REMARK 1.5.4. If f is α-Hölder continuous near x for some $\alpha > 0$, i.e.

$$|f(x) - f(y)| \leq C|x-y|^{\alpha} \quad \text{for all } y \text{ near } x,$$

then f satisfies (1.5.1). It is interesting to compare Theorem 1.5.3 (lower regularity assumptions, but need continuity) with Theorem 1.4.1 (allowing finitely many jumps).

PROOF OF THEOREM 1.5.3. Similar to (1.4.2), we have

$$S_m f(x) - f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} D_m(\theta)\big(f(x+\theta) - f(x)\big)\, d\theta$$
$$= \frac{1}{2\pi}\left(\int_{|y|<\delta} + \int_{\delta<|y|<\pi}\right) D_m(\theta)\big(f(x+\theta) - f(x)\big)\, d\theta.$$

We see that

$$\left|\int_{|y|<\delta} D_m(\theta)\big(f(x+\theta) - f(x)\big)\, d\theta\right| \leq C \int_{|y|<\delta} \left|\frac{f(x-y) - f(x)}{y}\right| dy.$$

From (1.5.1), given any $\epsilon > 0$, we can choose $\delta = \delta(\epsilon) > 0$ such that

$$\left|\int_{|y|<\delta} D_m(\theta)\big(f(x+\theta) - f(x)\big)\, d\theta\right| \leq C \int_{|y|<\delta} \left|\frac{f(x-y) - f(x)}{y}\right| dy < \epsilon.$$

Using Lemma 1.5.2, we have

$$\limsup_{m\to\infty} \big(S_m f(x) - f(x)\big) \leq \epsilon.$$

Then our result follows from the arbitrariness of $\epsilon > 0$. □

1.6. Gibbs-Wilbraham phenomenon

Despite the partial sum $S_m f$ converges pointwisely to the piecewise continuous function f, the partial sum $S_m f$ produces large peaks around the jump of f, which overshoot and undershoot the function's actual values. This approximation error approaches a limit of about 9%. This phenomenon is called the *Gibbs-Wilbraham phenomenon*, and we refer the details to the survey paper [**HH79, Onn82**]. The Gibbs phenomenon is not limited to Fourier expansion. In fact many researchers have reported the phenomenon in many different eigenfunction expansions, including those based on the Chebyshev and Legendre polynomials, which are special cases of the Gegenbauer polynomials [**GS97**]. In other words, given the Gegenbauer expansion coefficients of a piecewise analytic function, one faces the same convergence problem.

We now state the following theorem (without proof, which can be found in [**Onn82**][2]):

THEOREM 1.6.1 (Gibbs-Wilbraham, Theorem F of [**HH79**]). *Let f be the function given in Theorem 1.4.1. Let $\mathscr{D}$ be the set of discontinuities of f. For each $x \in \mathscr{D}$, let ℓ_x be the vertical line segment with*

$$\text{length } \frac{2}{\pi}\mathrm{Si}\,(\pi)|f(x+) - f(x-)| \text{ centered at } \frac{1}{2}\big(f(x+) - f(x-)\big).$$

Let $\mathcal{G}(g)$ be the graph of g. Then we have

$$\lim_{m\to\infty} \mathcal{G}(S_m f) = \mathcal{G}(f) \cup \bigcup_{x\in\mathscr{D}} \ell_x \quad \text{(limit as a set).}$$

[2] I would like to thank Aleksi Harmoinen for pointing this out to me.

REMARK 1.6.2. Note that

$$\frac{2}{\pi}\int_0^\pi \frac{\sin\theta}{\theta}\,d\theta = \frac{2}{\pi}\mathrm{Si}\,(\pi) = 1 + 2\big(\overbrace{0.0894\cdots}^{\substack{\text{about 9\%}\\ \text{overshoot}}}\big).$$

Sometimes $\mathrm{Si}\,(\pi)$ is known as the *Gibbs-Wilbraham constant*. In general, there are many jumps in signal, therefore this 9% overshoot actually causing significant noise in computation.

EXAMPLE 1.6.3 ([**Str08**]). The Fourier series

$$\sum_{n\geq 1 \text{ odd}} \frac{2}{n\pi}\sin n\pi$$

converges pointwisely to

$$f(x) = \begin{cases} \frac{1}{2} & \text{for } 0 < x < \pi, \\ 0 & \text{when } x = 0, \\ -\frac{1}{2} & \text{for } -\pi < x < 0. \end{cases}$$

Using (1.4.1), we know that the partial sum is given by

$$(S_m f)(x) = \frac{1}{4\pi}\left(\int_0^\pi - \int_{-\pi}^0\right)\frac{\sin(M(x-y))}{\sin(\frac{1}{2}(x-y))}\,dy,$$

where $M = m + \frac{1}{2}$. We consider the change of variable $\theta = M(x-y)$ in the first integral, while $\theta = M(y-x)$ in the second integral, we have

$$\begin{aligned}(S_m f)\left(\frac{\pi}{M}\right) &= \frac{1}{2\pi}\left(\int_{\pi-M\pi}^{\pi} - \int_{-\pi-M\pi}^{-\pi}\right)\frac{\sin\theta}{2M\sin(\frac{\theta}{2M})}\,d\theta \\ &= \frac{1}{2\pi}\left(\int_{-\pi}^{\pi} - \int_{-M\pi-\pi}^{-M\pi+\pi}\right)\frac{\sin\theta}{2M\sin(\frac{\theta}{2M})}\,d\theta \\ &= \frac{1}{2\pi}\left(\int_{-\pi}^{\pi} - \int_{M\pi-\pi}^{M\pi+\pi}\right)\frac{\sin\theta}{2M\sin(\frac{\theta}{2M})}\,d\theta, \end{aligned} \tag{1.6.1}$$

where we have changed the variable $\theta \mapsto -\theta$ in the last equality. When $M > 2$, we see that

$$\frac{\pi}{4} < \left(1 - \frac{1}{M}\right)\frac{\pi}{2} \leq \frac{\theta}{2M} \leq \left(1 + \frac{1}{M}\right)\frac{\pi}{2} < \frac{3\pi}{4} \quad \text{for all } \theta \in (M\pi - \pi, M\pi + \pi),$$

which implies

$$\sin\left(\frac{\theta}{2M}\right) > \frac{1}{\sqrt{2}} \quad \text{for all } \theta \in (M\pi - \pi, M\pi + \pi).$$

Therefore, we see that

$$\begin{aligned} &\limsup_{M\to\infty} \frac{1}{2\pi}\int_{M\pi-\pi}^{M\pi+\pi}\left|\frac{\sin\theta}{2M\sin(\frac{\theta}{2M})}\right|d\theta \\ &\leq \limsup_{M\to\infty}\frac{1}{2\sqrt{2}\pi M}\int_{M\pi-\pi}^{M\pi+\pi} d\theta = \limsup_{M\to\infty}\frac{1}{\sqrt{2}M} = 0. \end{aligned} \tag{1.6.2}$$

On the other hand, we see that

$$\lim_{M\to\infty} 2M\sin\frac{\theta}{2M} = \theta \quad \text{uniformly in } -\pi \leq \theta \leq \pi. \tag{1.6.3}$$

Combining (1.6.1), (1.6.2) and (1.6.3), we know that

$$\lim_{m\to\infty}(S_m f)\left(\frac{\pi}{m+\frac{1}{2}}\right) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\sin\theta}{\theta}\,d\theta = \frac{1}{\pi}\int_{0}^{\pi}\frac{\sin\theta}{\theta}\,d\theta = \frac{1}{\pi}\mathrm{Si}\,(\pi),$$

which verifies Theorem 1.6.1.

1.7. Cesàro summability of Fourier series in L^p

According to [**GS97**], the first attempt to resolve the Gibbs phenomenon was made by the Hungarian mathematician Lipót Fejér (or Leopold Fejér): In 1990 he discovered that the Cesàro means of the partial sums converge uniformly. Before explaining how to resolve the Gibbs phenomenon, let us begin our discussions from the following simple observation:

Lemma 1.7.1. *Let $\{c_m\}_{m=0,1,2,\cdots}$ be a sequence of complex numbers. Suppose that it converges to a limit $c \in \mathbb{C}$, then so its Cesàro sum:*

$$\lim_{N\to\infty}\frac{1}{N+1}\sum_{m=0}^{N} c_n = \ell.$$

Proof. Given any $\epsilon > 0$, there exists a sufficiently large $N_0 \in \mathbb{N}$ such that

$$|a_m - \ell| \le \epsilon \quad \text{for all } m > N_0.$$

For each $N > N_0$, we write

$$\begin{aligned}
\left|\frac{1}{N+1}\sum_{m=0}^{N} c_m - \ell\right| &\le \frac{1}{N}\sum_{m=0}^{N}|a_m-\ell| = \frac{1}{N}\sum_{m=0}^{N_0}|a_m-\ell| + \frac{1}{N}\sum_{m=N_0+1}^{N}|a_m-\ell| \\
&\le \frac{1}{N}\left((N_0+1)\sup_{0\le m\le N_0}|a_m-\ell|\right) + \overbrace{\frac{N-N_0}{N}}^{\le 1}\overbrace{\sup_{N_0<m\le N}|a_m-\ell|}^{\le\epsilon} \\
&\le \frac{1}{N}\left((N_0+1)\sup_{0\le m\le N_0}|a_m-\ell|\right) + \epsilon,
\end{aligned}$$

which implies

$$\limsup_{N\to\infty}\left|\frac{1}{N+1}\sum_{m=0}^{N} c_m - \ell\right| \le \epsilon.$$

Our lemma follows from the arbitrariness of $\epsilon > 0$. $\square$

Example 1.7.2 (Grandi's series). Let $a_m = (-1)^m$ for $m \ge 0$. Hence $\{a_m\}_{m=0}^{\infty}$ is the sequence $1, -1, 1, -1, \cdots$. Clearly the partial sum $S_m = \sum_{k=0}^{m} a_k$ does not converges, and in particular

$$\{S_m\}_{m=0}^{\infty} = \{1, 0, 1, 0, \cdots\}.$$

We see that the Cesàro sums are given by

$$\sigma_N = \frac{1}{N+1}\sum_{m=0}^{N} S_m = \begin{cases} \frac{M}{2M-1} & \text{if } N = 2M-1 \text{ is odd,} \\ \frac{1}{2} & \text{if } N = 2M \text{ is even,} \end{cases}$$

which implies

$$\lim_{N\to\infty}\sigma_N = \frac{1}{2}.$$

The above observation suggests that, instead of the partial sums $S_m f$, we consider the Cesàro sums

$$\sigma_N f(x) = \frac{1}{N+1}\sum_{m=0}^{N} S_m f(x).$$

This can be written in convolution form as

$$\sigma_N f(x) = \frac{1}{N+1}\sum_{m=0}^{N}(D_m * f)(x) = (F_N * f)(x),$$

where F_N is the *Fejér kernel*:

$$\begin{aligned} F_N(x) &= \frac{1}{N+1}\sum_{m=0}^{N}\frac{e^{i(m+\frac{1}{2})x} - e^{-i(m+\frac{1}{2})x}}{e^{i\frac{1}{2}x} - e^{-i\frac{1}{2}x}} \\ &= \frac{1}{N+1}\frac{e^{i\frac{1}{2}x}\left(\frac{e^{i(N+1)x}-1}{e^{ix}-1}\right) - e^{-i\frac{1}{2}x}\left(\frac{e^{-i(N+1)x}-1}{e^{-ix}-1}\right)}{e^{i\frac{1}{2}x} - e^{-i\frac{1}{2}x}} \\ &= \frac{1}{N+1}\frac{e^{i(N+1)x} - 1 + e^{-i(N+1)x} - 1}{(e^{i\frac{1}{2}x} - e^{-i\frac{1}{2}x})^2} \\ &= \frac{1}{N+1}\frac{\sin^2(\frac{N+1}{2}x)}{\sin^2(\frac{1}{2}x)}. \end{aligned}$$

EXERCISE 1.7.3. Verify that the Fejér kernel is an approximate identity as in Definition 1.3.6.

Our previous sections concerning how the partial sum $S_m f \equiv D_m * f$ converges to f in different senses. Since the Dirichlet kernel D_m takes negative values, it is not an approximate identity (Definition 1.3.6). However, using the summation method, we obtain an approximate kernel. In other words, the Cesàro sums "regularize" the kernel. Therefore using Lemma 1.3.8 we conclude the following theorem.

THEOREM 1.7.4 (Cesàro summability of Fourier series). *Let f be a 2π-periodic function in $\mathbb{R}$. If $f \in L^p(-\pi,\pi)$ for some $1 \le p < \infty$, then*

$$\lim_{N\to\infty} \|\sigma_N f - f\|_{L^p(-\pi,\pi)} = 0.$$

If f is continuous, then

$$\lim_{N\to\infty} \|\sigma_N f - f\|_{L^\infty(-\pi,\pi)} = 0. \tag{1.7.1}$$

Finally, we see that the uniform convergence (1.7.1) indicates that the lack of Gibbs-Wilbraham phenomena by considering the Cesàro sums $\sigma_N f$ of a continuous function f.

CHAPTER 2

Fourier transform

In previous chapter, we consider Fourier series for periodic functions. The main goal of this chapter is to study an analogue for non-periodic functions.

2.1. Motivations

We first perform some formal computations to bring out some motivations.

EXERCISE 2.1.1. Let $T > 0$ and let $f : \mathbb{R}^n \to \mathbb{C}$ be a function with period $2T$ on each variable. Show that the Fourier series of f is given by

$$f(x) = \sum_{k\in\mathbb{Z}^n} \hat{f}(k)e^{i\frac{\pi}{T}k\cdot x} \quad \text{with} \quad \hat{f}(k) = \fint_{[-T,T]^n} f(y)e^{-i\frac{\pi}{T}k\cdot y}\,dy, \tag{2.1.1}$$

where $\fint_{[-T,T]^n}$ is the average integral given by $\fint_{[-T,T]^n} \equiv \frac{1}{|[-T,T]^n|}\int_{[-T,T]^n} \equiv \frac{1}{(2T)^n}\int_{[-T,T]^n}$.

If we denote $\xi = k\frac{\pi}{T} \in \frac{\pi}{T}\mathbb{Z}$, then (2.1.1) is just simply

$$f(x) = \frac{1}{(2\pi)^n}\sum_{k\in\mathbb{Z}^n}\left(\int_{[-T,T]^n} f(y)e^{-i\xi\cdot y}\right)e^{i\xi\cdot x}\left(\frac{\pi}{T}\right)^n.$$

We observe that $(\frac{\pi}{T})^n$ is the volume of each square in the mesh $\frac{\pi}{T}\mathbb{Z}$. In view of Riemann integral, *formally* taking the limit $T \to \infty$ we see that

$$f(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} f(y)e^{-i\xi\cdot y}\right)e^{i\xi\cdot x}\,d\xi. \tag{2.1.2}$$

DEFINITION 2.1.2. The Fourier transform of $f \in L^1(\mathbb{R}^n)$ is defined by $(\mathscr{F}f)(\xi) \equiv \hat{f}(\xi) := \displaystyle\int_{\mathbb{R}^n} f(y)e^{-i\xi\cdot y}\,dy$.

REMARK 2.1.3. It is easy to see that $\|\hat{f}\|_{L^\infty(\mathbb{R}^n)} \leq \|f\|_{L^1(\mathbb{R}^n)}$ and $\hat{f} \in C^0(\mathbb{R}^n)$.

From (2.1.2) we *formally* have the Fourier inversion formula

$$f(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\hat{f}(\xi)e^{i\xi\cdot x}\,d\xi, \tag{2.1.3}$$

which is indeed true in some sense. In this course, we will focus on Fourier series in Euclidean space $\mathbb{R}^n$. Here we remark that the Fourier series we consider in previous chapter is indeed equivalent to the Fourier transform on torus $\mathbb{T}^n$.

Indeed, we often approximate Fourier transform by Fourier series in practical engineering application (e.g. signal processing).

EXERCISE 2.1.4 (Riemann-Lebesgue). Prove that if $f \in L^1(\mathbb{R}^n)$, then $\lim_{|\xi|\to\infty}\hat{f}(\xi) = 0$. [Hint: $C_c^\infty(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n)$, and consider the Laplacian.][1]

[1]See Lemma 1.5.1 for corresponding lemma for Fourier series

2.2. Schwartz space $\mathscr{S}(\mathbb{R}^n)$

Using Fubini's theorem, it is easy to see that

$$\int_{\mathbb{R}^n} \hat{f}(\xi)g(\xi)\,d\xi = \int_{\mathbb{R}^n} f(x)\hat{g}(x)\,dx \quad \text{for all } f,g \in L^1(\mathbb{R}^n). \tag{2.2.1}$$

Identity (2.2.1) might suggest defining the Fourier transform $\hat{f}$ of a distribution $f \in \mathscr{D}'(\mathbb{R}^n) \equiv (C_c^\infty(\mathbb{R}^n))'$. However, there is a series problem to implement this idea, since

$$\mathscr{F}(C_c^\infty(\mathbb{R}^n)) \not\subset C_c^\infty(\mathbb{R}^n),$$

see Corollary 2.3.7. To overcome this difficulty, we consider the functional space $\mathscr{S}(\mathbb{R}^n)$, which is the set of those smooth functions which, together with their derivatives, decrease more rapidly than the inverse of any polynomial. Precisely see the following definition:

Definition 2.2.1. The *Schwartz class* of rapidly decreasing functions is defined as

$$\mathscr{S}(\mathbb{R}^n) := \left\{ \varphi \in C^\infty(\mathbb{R}^n) \;\middle|\; \begin{array}{l} [\varphi]_{\alpha,\beta} := \sup\limits_{x\in\mathbb{R}^n} |x^\beta \partial^\alpha \varphi(x)| < \infty \\ \text{for all multi-indices } \alpha,\beta \end{array} \right\}. \tag{2.2.2}$$

The elements of $\mathscr{S}(\mathbb{R}^n)$ are called the *Schwartz function*.

Remark 2.2.2. For each $m \in \mathbb{N}$, we see that there exists $C \geq 1$ such that

$$C^{-1}|x|^m \leq \sum_{|\gamma|=m} |x^\gamma| \leq C|x|^m \quad \text{for all } x \in \mathbb{R}^n,$$

since the restriction of the function $g(x) := \sum_{|\gamma|=m} |x^\gamma|$ on $\mathcal{S}^{n-1}$ attains a nonzero minimum. Therefore each smooth function φ belongs to $\mathscr{S}(\mathbb{R}^n)$ if and only if

$$\sup_{x\in\mathbb{R}^n} \left|\langle x\rangle^m \partial^\alpha \varphi(x)\right| < \infty \quad \text{with} \quad \langle x\rangle = (1+|x|^2)^{\frac{1}{2}} \tag{2.2.3}$$

for all $m \in \mathbb{Z}_{\geq 0}$ and for all multi-index α with $|\alpha| \leq m$. In other words,

$$\mathscr{S}(\mathbb{R}^n) = \left\{ \varphi \in C^\infty(\mathbb{R}^n) \;\middle|\; \begin{array}{l} \|\varphi\|_m := \sum\limits_{|\alpha|\leq m} \sup\limits_{x\in\mathbb{R}^n} \left|\langle x\rangle^m \partial^\alpha \varphi(x)\right| < \infty \\ \text{for all } m \in \mathbb{N} \end{array} \right\}. \tag{2.2.4}$$

Here we remark that $[\cdot]_{\alpha,\beta}$ is a *semi-norm* and $\|\cdot\|_m$ is a *norm*.

Exercise 2.2.3. Prove that for each fixed number $a \in (0,\infty)$ the function $f(x) = e^{-a|x|^2}$ ($x \in \mathbb{R}^n$) belongs to $\mathscr{S}(\mathbb{R}^n)$. Therefore $C_c^\infty(\mathbb{R}^n) \subsetneq \mathscr{S}(\mathbb{R}^n) \subsetneq C^\infty(\mathbb{R}^n)$. However note that $e^{-|x|}$ is not in Schwartz space since it is not C^∞ near the origin.

We already define Schwartz class $\mathscr{S}(\mathbb{R}^n)$ as a set. We now define a topology for it (i.e. define open sets in $\mathscr{S}(\mathbb{R}^n)$), in order to make the notion of "continuous" make sense. Using the norms $\|\cdot\|_m$, we now define

$$\mathsf{d}_{\mathscr{S}(\mathbb{R}^n)}(\varphi,\psi) := \sum_{m=0}^\infty 2^{-m} \frac{\|\varphi-\psi\|_m}{1+\|\varphi-\psi\|_m} \quad \text{for all } \varphi,\psi \in \mathscr{S}(\mathbb{R}^n). \tag{2.2.5}$$

Exercise 2.2.4. Verify that (2.2.5) is a metric. [Hint: If $\|\cdot\|$ is a norm on a vector space, show that $\frac{\|u+v\|}{1+\|u+v\|} \leq \frac{\|u\|}{1+\|u\|} + \frac{\|v\|}{1+\|v\|}$.]

In particular, $\{\varphi_j\} \subset \mathscr{S}(\mathbb{R}^n)$ converges to zero in $\mathscr{S}(\mathbb{R}^n)$ if and only if

$$\|\varphi_j\|_m \to 0 \quad \text{for all } m \in \mathbb{Z}_{\geq 0},$$

or equivalently $[\varphi_j]_{\alpha,\beta} \to 0$ for all multi-indices α, β. Consequently, we have the following observation:

LEMMA 2.2.5. *The linear operator $T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is continuous if and only if*

$$\|T\varphi_j\|_m \to 0 \text{ provided } \|\varphi_j\|_m \to 0 \text{ for all } m \in \mathbb{Z}_{\geq 0}, \tag{2.2.6}$$

or equivalently

$$[T\varphi_j]_{\alpha,\beta} \to 0 \text{ provided } [\varphi_j]_{\alpha,\beta} \to 0 \text{ for all multi-indices } \alpha, \beta.$$

PROOF. If $T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is continuous, then it is clearly that (2.2.6) holds. Conversely, we now assume that (2.2.6). Since $\sum_{m=0}^{\infty} 2^{-m} < \infty$, given any $\epsilon > 0$, there exists $N(\epsilon) > 0$ such that $\sum_{m=N(\epsilon)+1}^{\infty} 2^{-m} < \epsilon$, which implies

$$\mathsf{d}_{\mathscr{S}(\mathbb{R}^n)}(\varphi, \psi) - \sum_{m=0}^{N(\epsilon)} 2^{-m} \frac{\|\varphi - \psi\|_m}{1 + \|\varphi - \psi\|_m} = \sum_{m=N(\epsilon)+1}^{\infty} 2^{-m} \frac{\|\varphi - \psi\|_m}{1 + \|\varphi - \psi\|_m} \leq \sum_{m=N(\epsilon)+1}^{\infty} 2^{-m} < \epsilon.$$

Hence from (2.2.6) we have

$$\limsup_{d(\varphi,\psi)\to 0} \mathsf{d}_{\mathscr{S}(\mathbb{R}^n)}(T\varphi, T\psi) \leq \epsilon + \sum_{m=0}^{N(\epsilon)} 2^{-m} \limsup_{\|\varphi - \psi\|_m \to 0} \frac{\|T(\varphi - \psi)\|_m}{1 + \|T(\varphi - \psi)\|_m} = \epsilon.$$

By arbitrariness of $\epsilon > 0$, we conclude our lemma. □

THEOREM 2.2.6. *Let $\mathsf{d}_{\mathscr{S}(\mathbb{R}^n)}$ be the metric given by (2.2.5). Then $(\mathscr{S}(\mathbb{R}^n), \mathsf{d}_{\mathscr{S}(\mathbb{R}^n)})$ is a Fréchet space space, that is, it is a complete metric space.*

REMARK 2.2.7. Note that $\mathscr{S}(\mathbb{R}^n)$ is a (Grothedieck) nuclear space. Since each infinitely dimensional Banach space spaces are not nuclear, then we cannot define a norm on $\mathscr{S}(\mathbb{R}^n)$.

PROOF OF THEOREM 2.2.6. Let $\{\varphi_j\}$ be a Cauchy sequence in $\mathscr{S}(\mathbb{R}^n)$. Given any $\epsilon > 0$ and multi-indices α, β, using the above observation, there exists $M > 0$ such that

$$[\varphi_j - \varphi_k]_{\alpha,\beta} \equiv \|x^\alpha \partial^\beta \varphi_j - x^\alpha \partial^\beta \varphi_k\|_{L^\infty(\mathbb{R}^n)} < \epsilon \quad \text{for all } j, k \geq M.$$

Hence the sequence $\{x^\alpha \partial^\beta \varphi_j\}$ is a Cauchy in the complete space $\big(C^0(\mathbb{R}^n), \|\cdot\|_{L^\infty(\mathbb{R}^n)}\big)$, then there exists a unique $g_{\alpha,\beta} \in C^0(\mathbb{R}^n)$ such that

$$\lim_{j\to\infty} \|x^\alpha \partial^\beta \varphi_j - g_{\alpha,\beta}\|_{L^\infty(\mathbb{R}^n)} = 0.$$

Let $g := g_{0,0}$, using the fact that $C^m(\mathbb{R}^n)$ is complete, then we see that $\partial^\beta g = g_{0,\beta}$ inductively. By the uniqueness of the limit, we see that $x^\alpha \partial^\beta g = g_{\alpha,\beta}$ and $g \in \mathscr{S}(\mathbb{R}^n)$, and consequently $\varphi_j \to g$ in $\mathscr{S}(\mathbb{R}^n)$. □

DEFINITION 2.2.8. The space of *slowly increasing functions* in $\mathbb{R}^n$ is defined as

$$\mathscr{O}_{\mathrm{M}}(\mathbb{R}^n) := \left\{ f \in C^\infty(\mathbb{R}^n) \;\middle|\; \begin{array}{l} \text{for each multi-index } \alpha, \text{ there exists} \\ M \in \mathbb{Z}_{\geq 0} \text{ such that } \sup_{x\in\mathbb{R}^n} \big|\langle x\rangle^{-M} \partial^\alpha f(x)\big| < \infty \end{array} \right\}$$

EXERCISE 2.2.9. Prove that for each $s \in \mathbb{R}$ the function $f(x) := \langle x\rangle^s$ ($x \in \mathbb{R}^n$) belongs to $\mathscr{O}_{\mathrm{M}}(\mathbb{R}^n)$.

EXERCISE 2.2.10. Prove that the function $f(x) = e^{i|x|^2}$ $(x \in \mathbb{R}^n)$ belongs to $\mathscr{O}_{\mathrm{M}}(\mathbb{R}^n)$.

Some other basic properties of the Schwartz class are collected in the next proposition.

PROPOSITION 2.2.11. *If $f \in \mathscr{O}_{\mathrm{M}}(\mathbb{R}^n)$ and $v \in \mathscr{S}(\mathbb{R}^n)$, then the following operations are continuous maps from $\mathscr{S}(\mathbb{R}^n)$ into $\mathscr{S}(\mathbb{R}^n)$:*

1. ***Reflection.*** *$\varphi \mapsto \tilde{\varphi}$ with $\tilde{\varphi}(x) = \varphi(-x)$,*
2. ***Conjugation.*** *$\varphi \mapsto \overline{\varphi}$,*
3. ***Translation.*** *$\varphi \mapsto \tau_{x_0}\varphi$ with $\tau_{x_0}\varphi(x) = \varphi(x - x_0)$,*
4. ***Derivative.*** *$\varphi \mapsto \partial^\alpha\varphi$, in other words, $\mathscr{S}(\mathbb{R}^n)$ is stable under differentiation (using similar arguments, we know that $\mathscr{O}_{\mathrm{M}}(\mathbb{R}^n)$ is also stable under differentiation),*
5. ***Multiplication.*** *$\varphi \mapsto f\varphi$.*

PROOF. Part (1) and (2) are clear. For (3), by observing that

$$x^\alpha = (x - x_0 + x_0)^\alpha = \sum_{\gamma \le \alpha} c_\gamma (x - x_0)^\gamma,$$

we see that

$$\begin{aligned}[]
[\tau_{x_0}\varphi]_{\alpha,\beta} &= \sup_{x \in \mathbb{R}^n} |x^\alpha \partial^\beta \varphi(x - x_0)| \\
&\le C \sum_{\gamma \le \alpha} \sup_{x \in \mathbb{R}^n} |(x - x_0)^\gamma \partial^\beta \varphi(x - x_0)| = C \sum_{\gamma \le \alpha} [\varphi]_{\gamma,\beta}.
\end{aligned}$$

Therefore (3) follows from Lemma 2.2.5. Part (4) is an immediate consequence of the following identity:

$$[\partial^\beta \varphi]_{\alpha',\beta'} = [\varphi]_{\alpha',\beta'+\beta}.$$

Since $f \in \mathscr{O}_{\mathrm{M}}$, given any β we may choose C and N such that $\left|\langle x\rangle^{-N}\partial^\gamma f(x)\right| \le C$ whenever $\gamma \le \beta$. Now we have

$$\begin{aligned}[]
[f\varphi]_{\alpha,\beta} = \|x^\alpha \partial^\beta (f\varphi)\|_{L^\infty(\mathbb{R}^n)} &= \left\| x^\alpha \sum_{\gamma \le \beta} c_\gamma (\partial^{\beta-\gamma} f)(\partial^\gamma \varphi) \right\|_{L^\infty(\mathbb{R}^n)} \\
&\le C \sum_{\gamma \le \beta} \left\| x^\alpha \langle x\rangle^N \overbrace{\left(\langle x\rangle^{-N} \partial^{\beta-\gamma} f\right)}^{\le C} (\partial^\gamma \varphi) \right\|_{L^\infty(\mathbb{R}^n)} \\
&\le C \sum_{\gamma \le \beta} \left\| x^\alpha \langle x\rangle^N \partial^\gamma \varphi \right\|_{L^\infty(\mathbb{R}^n)}.
\end{aligned}$$

Therefore (5) follows from Lemma 2.2.5. □

LEMMA 2.2.12. *For each $1 \le p < \infty$, the space $\mathscr{S}(\mathbb{R}^n)$ is continuous embedded in $L^p(\mathbb{R}^n)$, i.e. $\mathscr{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ as a set and the inclusion mapping $\iota : \mathscr{S}(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ is continuous.*

PROOF. Let $\varphi \in \mathscr{S}(\mathbb{R}^n)$. For $p = 1$, the claim follows from

$$\begin{aligned}
\|\varphi\|_{L^1(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} \langle x\rangle^{-n-1} \left(\langle x\rangle^{n+1} |\varphi(x)|\right) dx \\
&\le \|\langle x\rangle^{n+1}\varphi\|_{L^\infty(\mathbb{R}^n)} \overbrace{\int_{\mathbb{R}^n} \langle x\rangle^{-n-1}\, dx}^{<\infty} \\
&\equiv C \|\langle x\rangle^{n+1}\varphi\|_{L^\infty(\mathbb{R}^n)}.
\end{aligned}$$

For $p > 1$, the inequality

$$\|\varphi\|_{L^p(\mathbb{R}^n)} = \left(\int_{\mathbb{R}^n} |\varphi(x)||\varphi(x)|^{p-1}\, dx \right)^{\frac{1}{p}} \leq \|\varphi\|_{L^1(\mathbb{R}^n)}^{\frac{1}{p}} \|\varphi\|_{L^\infty(\mathbb{R}^n)}^{1-\frac{1}{p}}$$

implies the result. □

2.3. Fourier transform on Schwartz space

In view of Lemma 2.2.12, the Fourier transform is well-defined on $\mathscr{S}(\mathbb{R}^n)$ by restriction.

Exercise 2.3.1. Let $\phi_n(x) = e^{-\frac{1}{2}|x|^2}$, which is in $\mathscr{S}(\mathbb{R}^n)$ by Exercise 2.2.3. Prove that $\hat{\phi}_n = (2\pi)^{\frac{n}{2}} \phi_n$ and $\phi_n(0) = (2\pi)^{-n} \int_{\mathbb{R}^n} \hat{\phi}_n(x)\, dx$.

We first show that $\hat{f}$ is smooth whenever $f \in \mathscr{S}(\mathbb{R}^n)$:

Lemma 2.3.2. *For any $f \in \mathscr{S}(\mathbb{R}^n)$, the Fourier transform $\hat{f}$ is in $C^\infty(\mathbb{R}^n)$ and $\partial^\alpha \hat{f} \in L^\infty(\mathbb{R}^n)$ for all multi-index α.*

Proof. From Lemma 2.2.12 and the definition of the Fourier transform, we know that $\hat{f} \in L^\infty(\mathbb{R}^n)$.

Observe that

$$\begin{aligned}
\frac{\hat{f}(\xi + he_k) - \hat{f}(\xi)}{h} &= \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x) \frac{e^{-ihx_k} - 1}{h}\, dx \\
&= \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x) \left(\frac{1}{h} \int_0^{x_k} \frac{d}{dt}(e^{-iht})\, dt \right) dx \\
&= -i \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x) \left(\int_0^{x_k} e^{-iht}\, dt \right) dx.
\end{aligned}$$

Since $|\int_0^{x_k} e^{-iht}\, dt| \leq |x_k|$, the Lebesgue dominated convergence theorem implies

$$\begin{aligned}
\partial_{\xi_k} \hat{f}(\xi) &= \lim_{h\to 0} \frac{\hat{f}(\xi + he_k) - \hat{f}(\xi)}{h} \\
&= -i \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x) \left(\lim_{h\to 0} \int_0^{x_k} e^{-iht}\, dt \right) dx \\
&= \int_{\mathbb{R}^n} e^{-ix\cdot\xi} \big(-ix_k f(x) \big)\, dx = \mathscr{F}(-ix_k f(x)).
\end{aligned}$$

Since $-ix_k f(x)$ is also in $\mathscr{S}(\mathbb{R}^n)$, the using the observation in the first line of the proof we see that $\partial_{\xi_k} \hat{f}(\xi)$ is in $L^\infty(\mathbb{R}^n)$. Inductively, by observing that $x^\alpha f(x)$ is in $\mathscr{S}(\mathbb{R}^n)$, we conclude that $\partial_\xi^\alpha \hat{f}(\xi)$ is in $L^\infty(\mathbb{R}^n)$. □

Proposition 2.3.3 (Basic properties of Fourier transform). *Let $f \in \mathscr{S}(\mathbb{R}^n)$, $x_0, \xi_0 \in \mathbb{R}^n$, $c > 0$ and multi-indices α, β. Then the following identities hold:*

1. ***Translation.*** $(\tau_{x_0} f)^\wedge(\xi) = e^{-ix_0\cdot\xi} \hat{f}(\xi)$ *with* $\tau_{x_0} f(x) = f(x - x_0)$,
2. ***Modulation.*** $(e^{ix\cdot\xi_0} f)^\wedge(\xi) = \tau_{\xi_0} \hat{f}(\xi)$,
3. ***Scaling.*** $(f_c)^\wedge(\xi) = c^{-n} \hat{f}(\xi/c)$ *with* $f_c(x) = f(cx)$,
4. ***Derivative.*** $(\partial_x^\alpha f)^\wedge(\xi) = (i\xi)^\alpha \hat{f}(\xi)$,
5. ***Polynomial.*** $((-ix)^\beta f)^\wedge(\xi) = \partial_\xi^\beta \hat{f}(\xi)$.

REMARK 2.3.4. In some context of pseudo-differential operator, some authors denote $D_{x_k} = \frac{1}{i}\partial_{x_k}$. In this case, we write $(D_x^\alpha f)^\wedge(\xi) = \xi^\alpha \hat{f}(\xi)$ and $((-x)^\beta f)^\wedge(\xi) = D_\xi^\beta \hat{f}(\xi)$.

EXERCISE 2.3.5. Proof Proposition 2.3.3.

We are now able to prove (2.1.2) in a rigorous sense as follows:

THEOREM 2.3.6. *The mapping $\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is an algebraic and topological isomorphism (i.e. it is bijective, continuous and its inverse is also continuous). In addition, its inverse is the operator $\mathscr{F}^{-1} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is given by the formula*

$$(\mathscr{F}^{-1}g)(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{ix\cdot\xi} g(\xi)\, d\xi \quad \text{for all } g \in \mathscr{S}(\mathbb{R}^n) \text{ and } x \in \mathbb{R}^n, \tag{2.3.1}$$

that is, $(\mathscr{F}^{-1}g)(\zeta) = (2\pi)^{-n}(\mathscr{F}g)(-\zeta)$.

PROOF. Let $f \in \mathscr{S}(\mathbb{R}^n)$. From Lemma 2.3.2, we already know that $\hat{f} \in C^\infty(\mathbb{R}^n)$. Given any multi-indices α and β, we see that

$$\begin{aligned}
[\hat{f}]_{\alpha,\beta} &= \sup_{\xi\in\mathbb{R}^n} |\xi^\alpha \partial_\xi^\beta \hat{f}(\xi)| = \sup_{\xi\in\mathbb{R}^n} |(i\xi)^\alpha \partial_\xi^\beta \hat{f}(\xi)| \\
&= \sup_{\xi\in\mathbb{R}^n} \left| \left[\partial_x^\alpha \big((-ix)^\beta f\big)\right]^\wedge(\xi)\right| \\
&\le \|\partial_x^\alpha \big((-ix)^\beta f\big)\|_{L^1(\mathbb{R}^n)}.
\end{aligned}$$

Using Leibniz rule (a.k.a. product rule), we see that $\partial_x^\alpha\big((-ix)^\beta f\big) = \sum_{k=1}^m c_k x^{\alpha_k} \partial_x^{\beta_k} f(x)$ for some constants c_k and multi-indices α_k, β_k, so

$$[\hat{f}]_{\alpha,\beta} \le C \sum_{k=1}^m \|x^{\alpha_k} \partial_x^{\beta_k} f\|_{L^1(\mathbb{R}^n)} \le C \sum_{k=1}^m \|x^{\alpha_k+n+1} \partial_x^{\beta_k} f\|_{L^\infty(\mathbb{R}^n)}.$$

By arbitrariness of α and β, and using Lemma 2.2.5, we conclude that $\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is continuous.

To prove Theorem 2.3.6, it is remain to show (2.3.1). Fixing any $\varphi \in \mathscr{S}(\mathbb{R}^n)$ and $c > 0$. Choosing $g(x) = \varphi(x/c)$ in (2.2.1) gives

$$\int_{\mathbb{R}^n} \hat{f}(x)\varphi(x/c)\, dx = \int_{\mathbb{R}^n} f(y) c^n \hat{\varphi}(cy)\, dy = \int_{\mathbb{R}^n} f(y/c)\hat{\varphi}(y)\, dy.$$

Taking the limit $c \to \infty$ (Lebesgue dominated convergence theorem) in the equality above, we have

$$\varphi(0) \int_{\mathbb{R}^n} \hat{f}(x)\, dx = f(0) \int_{\mathbb{R}^n} \hat{\varphi}(y)\, dy.$$

We now choose φ to be the Gaussian ϕ_n in Exercise 2.3.1, then we obtain that

$$f(0) = (2\pi)^{-n} \int_{\mathbb{R}^n} \hat{f}(x)\, dx,$$

which proves (2.3.1) for $x = 0$. Therefore, from Proposition 2.3.3 we know that

$$f(x) = (\tau_x f)(0) = (2\pi)^{-n} \int_{\mathbb{R}^n} (\tau_x f)^\wedge(x)\, dx = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{ix\cdot\xi} \hat{f}(\xi)\, d\xi, \tag{2.3.2}$$

which implies the theorem[2]. □

[2] The surjectivity can be seen by rephrase (2.3.2) as $(\mathscr{F}^2 f)(-x) = (2\pi)^n f(x)$ for all $x \in \mathbb{R}^n$.

COROLLARY 2.3.7. *If $\varphi \in C_c^\infty(\mathbb{R}^n)$ and $\hat{\varphi} \in C_c^\infty(\mathbb{R}^n)$, then $\varphi \equiv 0$.*

PROOF. Suppose that $\varphi \in C_c^\infty(\mathbb{R}^n)$ with $\hat{\varphi} \in C_c^\infty(\mathbb{R}^n)$, and let $x^* = (x_1^*, \cdots, x_n^*) \in \mathbb{R}^n$. Define $\Phi : \mathbb{C} \to \mathbb{C}$ by setting

$$\Phi(z) := \int_{\mathbb{R}^n} e^{-i(zx_1 + \sum_{j=2}^n x_j^* x_j)} \varphi(x_1, \cdots, x_n)\, dx_1 \cdots dx_n \quad \text{for } z \in \mathbb{C}.$$

Then Φ is analytic in $\mathbb{C}$ and $\Phi(t) = \hat{\varphi}(t, x_2^*, \cdots, x_n^*)$ for every $t \in \mathbb{R}$. Since $\hat{\varphi}$ has compact support, then $\Phi = 0$ in $\mathbb{R} \setminus [-R, R]$ if $R > 0$ sufficiently large. Using the unique continuation property of analytic function, we conclude that $\Phi \equiv 0$ in $\mathbb{C}$, which gives

$$\hat{\varphi}(x^*) = \Phi(x_1^*) = 0.$$

By arbitrariness of $x^* \in \mathbb{R}^n$, and using the injectivity of the Fourier transform on $C_c^\infty(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$, we conclude our lemma. □

We finally end this subsection by the following proposition.

PROPOSITION 2.3.8. *For each $f, g \in \mathscr{S}(\mathbb{R}^n)$, one has*

(1) ***Symmetry.*** *$\mathscr{F}^2 f = (2\pi)^n \tilde{f}$ with $\tilde{f}(x) = f(-x)$. Consequently, $\mathscr{F}^4 f = (2\pi)^{2n} f$.*
(2) ***Parseval's identity.*** $\displaystyle\int_{\mathbb{R}^n} \hat{f}(x) g(x)\, dx = \int_{\mathbb{R}^n} f(x) \hat{g}(x)\, dx.$
(3) ***Parseval's identity.*** $\displaystyle\int_{\mathbb{R}^n} f(x) \overline{g(x)}\, dx = (2\pi)^{-n} \int_{\mathbb{R}^n} \hat{f}(\xi) \overline{\hat{g}(\xi)}\, d\xi.$ *Consequently,* $\displaystyle\int_{\mathbb{R}^n} |f(x)|^2\, dx = (2\pi)^{-n} \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2\, d\xi.$

EXERCISE 2.3.9. Prove Proposition 2.3.8.

2.4. The space of tempered distributions $\mathscr{S}'(\mathbb{R}^n)$

We now want to define the corresponding class of distributions, namely the tempered distributional Fourier transform.

DEFINITION 2.4.1. Let $\mathscr{S}'(\mathbb{R}^n)$ be the set of continuous (w.r.t. the metric (2.2.5)) linear functional on $\mathscr{S}(\mathbb{R}^n)$, i.e. dual space of $\mathscr{S}(\mathbb{R}^n)$. Precisely,

$$\mathscr{S}'(\mathbb{R}^n) := \left\{ T : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C} \;\middle|\; \begin{array}{l} T \text{ linear and } T(\varphi_j) \to 0 \\ \text{whenever } \varphi_j \to 0 \text{ in } \mathscr{S}(\mathbb{R}^n) \end{array} \right\}.$$

The elements of $\mathscr{S}'(\mathbb{R}^n)$ are called *tempered distributions*.

We first show that any tempered distribution has finite order in the following lemma:

LEMMA 2.4.2. *For any $T \in \mathscr{S}'(\mathbb{R}^n)$, there exist $C > 0$ and $N \in \mathbb{N}$ such that*

$$|T(\varphi)| \le C \sum_{|\beta| \le N} \|\langle x \rangle^N \partial^\beta \varphi\|_{L^\infty(\mathbb{R}^n)} \quad \textit{for all } \varphi \in \mathscr{S}(\mathbb{R}^n).$$

PROOF. Suppose the contrary, that for any $N > 0$ there is a $\varphi_N \in \mathscr{S}(\mathbb{R}^n)$ such that

$$|T(\varphi_N)| \ge N \sum_{|\beta| \le N} \|\langle x \rangle^N \partial^\beta \varphi_N\|_{L^\infty(\mathbb{R}^n)}. \tag{2.4.1}$$

If we define

$$\psi_N(x) := \frac{1}{N}\left(\sum_{|\beta|\le N} \|\langle x\rangle^N \partial^\beta \varphi_N\|_{L^\infty(\mathbb{R}^n)}\right)^{-1} \varphi_N(x) \quad \text{for all } x \in \mathbb{R}^n,$$

from (2.4.1) it is easy to see that

$$|T(\psi_N)| \ge 1 \quad \text{for all } N \in \mathbb{N}. \tag{2.4.2}$$

On the other hand, for each fixed multi-index β_0 we have

$$\|\langle x\rangle^N \partial^{\beta_0}\psi_N\|_{L^\infty(\mathbb{R}^n)} = \frac{1}{N}\left(\sum_{|\beta|\le N} \|\langle x\rangle^N \partial^\beta \varphi_N\|_{L^\infty(\mathbb{R}^n)}\right)^{-1} \|\langle x\rangle^N \partial^{\beta_0}\varphi_N\|_{L^\infty(\mathbb{R}^n)} \le \frac{1}{N}$$

for all sufficiently large N. By arbitrariness of β_0, from Lemma 2.2.5 we know that $\psi_N \to 0$ in $\mathscr{S}(\mathbb{R}^n)$. Since $T \in \mathscr{S}'(\mathbb{R}^n)$, then

$$\lim_{N\to\infty} T(\psi_N) = 0,$$

which contradicts with (2.4.2). □

Example 2.4.3. If $f : \mathbb{R}^n \to \mathbb{C}$ any measurable polynomially bounded function f, in the sense that $|f(x)| \le C\langle x\rangle^N$ for a.e. $x \in \mathbb{R}^n$, define

$$T_f : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}, \quad T_f(\varphi) = \int_{\mathbb{R}^n} f\varphi\,dx.$$

Since for any $\varphi \in \mathscr{S}(\mathbb{R}^n)$ we have

$$|T_f(\varphi)| = \left|\int_{\mathbb{R}^n} f\varphi\,dx\right| \le C\int_{\mathbb{R}^n} \langle x\rangle^N |\varphi(x)|\,dx \le C\|\langle x\rangle^{N+n+1}\varphi\|_{L^\infty(\mathbb{R}^n)}.$$

Using Lemma 2.2.5, we know that $T(\varphi_j) \to 0$ whenever $\varphi_j \to 0$ in $\mathscr{S}(\mathbb{R}^n)$. Moreover, it is possible to identify the distribution T_f with the function f, since the condition $T_{f_1} = T_{f_2}$ implies that

$$\int_{\mathbb{R}^n} (f_1 - f_2)\varphi\,dx = 0 \quad \text{for all } \varphi \in \mathscr{S}(\mathbb{R}^n),$$

which implies that $f_1 = f_2$ a.e., hence it is legitimate to denote $f \in \mathscr{S}'(\mathbb{R}^n)$. Therefore, we can identify $\mathscr{O}_M(\mathbb{R}^n)$ as a subspace of $\mathscr{S}'(\mathbb{R}^n)$.

Exercise 2.4.4. Prove that for each $a \in (-n, \infty)$ the function $|x|^a$ is a tempered distribution in $\mathbb{R}^n$. Therefore, we know that $\mathscr{O}_M(\mathbb{R}^n) \subsetneq \mathscr{S}'(\mathbb{R}^n)$.

Example 2.4.5 (Measures as distributions). Let μ be either a complex Borel measure or a positive Borel measure[3] on $\mathbb{R}^n$. We say that the measure μ is polynomially bounded if for some N the total variation $|\mu|$ satisfies

$$\int_{\mathbb{R}^n} \langle x\rangle^{-N} d|\mu|(x) < \infty.$$

[3] For complex measures, the measure can take on complex values, infinite values are not allowed. In contrast, infinite values are allowed for positive measures. In particular, a finite positive measure is a special case of a complex measures.

Any polynomial bounded measure μ and for any $\varphi \in \mathscr{S}(\mathbb{R}^n)$, we see that

$$\left| \int_{\mathbb{R}^n} \varphi(x)\, d\mu(x) \right| \leq \int_{\mathbb{R}^n} |\varphi(x)|\, d|\mu|(x) \leq \|\langle x \rangle^N \varphi(x)\|_{L^\infty(\mathbb{R}^n)} \int_{\mathbb{R}^n} \langle x \rangle^{-N}\, d|\mu|(x),$$

which shows μ can be identify as an element T_μ in $\mathscr{S}'(\mathbb{R}^n)$ given by

$$T_\mu(\varphi) := \int_{\mathbb{R}^n} \varphi(x)\, d\mu(x),$$

therefore it is legitimate to denote $\mu \in \mathscr{S}'(\mathbb{R}^n)$.

EXAMPLE 2.4.6 (L^p functions as distributions). For each $1 \leq p \leq \infty$, we have $L^p(\mathbb{R}^n)$ can be identify as a subspace of $\mathscr{S}'(\mathbb{R}^n)$ by identifying f with the element

$$T_f(\varphi) := \int_{\mathbb{R}^n} f(x)\varphi(x)\, dx.$$

In particular, we see that

$$|T_f(\varphi)| \leq \|f\|_{L^p(\mathbb{R}^n)} \|\varphi\|_{L^{p'}(\mathbb{R}^n)} \leq C \|f\|_{L^p(\mathbb{R}^n)} \|\varphi\|_{\mathscr{S}(\mathbb{R}^n)},$$

where we used Hölder's inequality and Lemma 2.2.12.

DEFINITION 2.4.7. Let $\{T_j\}_{j=1}^\infty \subset \mathscr{S}'(\mathbb{R}^n)$ and $T \in \mathscr{S}'(\mathbb{R}^n)$. We say that $T_j \to T$ in $\mathscr{S}'(\mathbb{R}^n)$ if

$$T_j(\varphi) \to T(\varphi) \quad \text{for any } \varphi \in \mathscr{S}(\mathbb{R}^n).$$

LEMMA 2.4.8 (Convergence in $\mathscr{S}'(\mathbb{R}^n)$). *The followings are true:*

(1) *If $T_j \to T$ in $\mathscr{S}'(\mathbb{R}^n)$ and $T_j \to S$ in $\mathscr{S}'(\mathbb{R}^n)$, then $T \equiv S$.*
(2) *If $\{\varphi_j\}$ is a sequence in $\mathscr{S}(\mathbb{R}^n)$ (resp. $L^p(\mathbb{R}^n)$ for some $1 \leq p \leq \infty$) with $\varphi_j \to \varphi$ in $\mathscr{S}(\mathbb{R}^n)$ (resp. in $L^p(\mathbb{R}^n)$), then $\varphi_j \to \varphi$ in $\mathscr{S}'(\mathbb{R}^n)$.*

EXERCISE 2.4.9. Prove Lemma 2.4.8.

The operations on tempered distribution can be induced from Proposition 2.2.11:

PROPOSITION 2.4.10. *Let $f \in \mathscr{O}_{\mathrm{M}}(\mathbb{R}^n)$. The following operations map $\mathscr{S}'(\mathbb{R}^n)$ into $\mathscr{S}'(\mathbb{R}^n)$, and they extend the corresponding operations on $\mathscr{S}(\mathbb{R}^n)$:*

(1) ***Reflection.*** *$\tilde{T}(\varphi) = T(\tilde{\varphi})$ with $\tilde{\varphi}(x) = \varphi(-x)$,*
(2) ***Conjugation.*** *$\overline{T}(\varphi) = \overline{T(\overline{\varphi})}$,*
(3) ***Translation.*** *$(\tau_{x_0} T)(\varphi) = T(\tau_{-x_0}\varphi)$ with $\tau_{-x_0}\varphi(x) = \varphi(x + x_0)$,*
(4) ***Distributional derivative.*** *$(\partial^\alpha T)(\varphi) = (-1)^{|\alpha|} T(\partial^\alpha \varphi)$,*
(5) ***Multiplication.*** *$(fT)(\varphi) = T(f\varphi)$.*

REMARK 2.4.11 (Distributional derivatie v.s. weak derivative). Perhaps the most striking point is that any tempered distribution has distributional derivatives of any order, and these derivatives are still tempered distributions. We consider the Heaviside function $H \in L^\infty(\mathbb{R}) \subset \mathscr{S}'(\mathbb{R})$ (see Example 2.4.6) given in (1.1.6). According to Proposition 2.4.10, the distributional derivative of H is given by

$$H'(\varphi) = -H(\varphi') = -\int_{-\infty}^{\infty} H(x)\varphi'(x)\, dx = -\int_0^\infty \varphi'(x)\, dx = \varphi(0) \quad \text{for all } \varphi \in \mathscr{S}(\mathbb{R}),$$

therefore, we have $H' = \delta_0$. It is worth-mentioning that the weak derivative (Definition 1.1.1) of H does not exist (see Example 1.1.6). Each weak derivative also a distributional derivative, but not the converse. Without any ambiguity, here and after, we denote ∂^α the distributional derivatives.

2.5. Fourier transform on $\mathscr{S}'(\mathbb{R}^n)$

Parseval's identity in Proposition 2.3.8 shows that the following definition extends the Fourier transform:

DEFINITION 2.5.1. The Fourier transform of any tempered distribution $T \in \mathscr{S}'(\mathbb{R}^n)$ is the tempered distribution $\hat{T} = \mathscr{F}T$ defined by

$$\hat{T}(\varphi) = T(\hat{\varphi}).$$

Similarly, the inverse Fourier transform of $T \in \mathscr{S}'(\mathbb{R}^n)$ is the distribution $\check{T} = \mathscr{F}^{-1}T$ for which $\check{T}(\varphi) = T(\check{\varphi})$.

EXAMPLE 2.5.2. The Fourier transform of the Dirac measure δ_{x_0} is the tempered distribution given by

$$(\delta_{x_0})^\wedge(\varphi) = \delta_{x_0}(\hat{\varphi}) = \hat{\varphi}(x_0) = \int_{\mathbb{R}^n} e^{-ix_0\cdot\xi}\varphi(\xi)\,d\xi.$$

Thus, $(\delta_{x_0})^\wedge$ can be identify with the function $\xi \mapsto e^{-ix_0\cdot\xi}$. In particular, $\mathscr{F}\delta_0 = 1$.

EXAMPLE 2.5.3. Using Proposition 2.3.3, the derivative of Dirac measure can be computed as followings:

$$\begin{aligned}(\partial^\alpha\delta_0)^\wedge(\varphi) &= (\partial^\alpha\delta_0)(\hat{\varphi}) = (-1)^{|\alpha|}\delta_0(\partial^\alpha\hat{\varphi})\\ &= (-1)^{|\alpha|}\delta_0\big(((-ix)^\alpha\varphi)^\wedge\big)\\ &= \delta_0\big(((ix)^\alpha\varphi)^\wedge\big) = \int_{\mathbb{R}^n}(i\xi)^\alpha\varphi(\xi)\,d\xi,\end{aligned}$$

which conclude $(\partial^\alpha\delta_0)^\wedge = (i\xi)^\alpha$.

Similar to the Fourier transform on Schwartz space, it is easy (and natural) to see that the Fourier transform is also isomorphism on the space of the tempered distributions. Here we record this observation as a theorem:

THEOREM 2.5.4 (Fourier inversion theorem). *The Fourier transform is a bijective map from $\mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$. It is continuous in the sense that*

$$T_j \to T \text{ in } \mathscr{S}'(\mathbb{R}^n) \implies \hat{T}_j \to \hat{T} \text{ in } \mathscr{S}'(\mathbb{R}^n).$$

One has the inversion formula

$$T(\tilde{\varphi}) = (2\pi)^{-n}\hat{T}(\hat{\varphi}) \quad \text{for all } \varphi \in \mathscr{S}(\mathbb{R}^n) \tag{2.5.1}$$

with $\tilde{\varphi}(x) = \varphi(-x)$.

PROOF. The inversion formula (2.5.1) immediately follows from Proposition 2.3.8. From (2.5.1) and since $\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is continuous bijective, we know that $\mathscr{F} : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ is also continuous bijective. □

From Proposition 2.3.3, it is easy to see the following properties:

PROPOSITION 2.5.5 (Basic properties of Fourier transform). *Let $T \in \mathscr{S}'(\mathbb{R}^n)$, $\varphi \in \mathscr{S}(\mathbb{R}^n)$, $x_0, \xi_0 \in \mathbb{R}^n$, $c > 0$ and multi-indices α, β. Then the Fourier transform on $\mathscr{S}'(\mathbb{R}^n)$ has the following properties:*

(1) ***Translation.*** $(\tau_{x_0} T)^\wedge = e^{-ix_0\cdot\xi}\hat{T}$
(2) ***Modulation.*** $(e^{ix\cdot\xi_0}T)^\wedge = \tau_{\xi_0}\hat{T}$,
(3) ***Derivative.*** $(\partial_x^\alpha T)^\wedge = (i\xi)^\alpha \hat{T}$,
(4) ***Polynomial.*** $((-ix)^\beta T)^\wedge = \partial_\xi^\beta \hat{T}$.

Here the operators are given in Proposition 2.4.10.

2.6. Fourier transform on L^2

To extend the Fourier transform on $L^2(\mathbb{R}^n)$, we prove the following extension lemma for densely defined bounded linear transform:

LEMMA 2.6.1. *Let X and Y be Banach spaces and let X_0 be a dense subspace of X. If $T_0 : X_0 \to Y$ be a linear mapping satisfies*

$$\|T_0 x\|_Y \le C\|x\|_X \quad \textit{for all } x \in X_0, \tag{2.6.1}$$

then there exists a unique bounded linear mapping $T : X \to Y$ with $T|_{X_0} \equiv T_0$ such that

$$\|Tx\|_Y \le C\|x\|_X \quad \textit{for all } x \in X \tag{2.6.2}$$

and

$$Tx = \lim_{j\to\infty} T_0 x_j \quad \textit{for all } \{x_j\} \subset X_0 \textit{ with } x_j \to x \textit{ in } X. \tag{2.6.3}$$

PROOF. Let $\{x_j\} \subset X_0$ be such that $x_j \to x$ in X. Using (2.6.1), it is easy to see that

$$\limsup_{j,k\to\infty} \|T_0 x_j - T_0 x_k\|_Y \le \limsup_{j,k\to\infty} \|x_j - x_k\|_X = 0,$$

that is $\{T_0 x_j\}$ is a Cauchy sequence in Y. Since Y is a Banach space, then there exists a unique $y \in Y$ such that

$$\lim_{j\to\infty} T_0(x_j) = y \quad \text{in } Y.$$

Suppose that $\{x_j'\} \subset X_0$ is another sequence such that $x_j' \to x$. Using (2.6.1), we see that

$$\limsup_{j\to\infty} \|T_0 x_j - T_0 x_j'\|_Y \le \limsup_{j\to\infty} \|x_j - x_j'\|_X = 0.$$

Therefore, the unique extension T given in (2.6.3) is well-defined. On the other hand, we see that

$$\begin{aligned}
\|Tx\|_Y &\le \limsup_{j\to\infty} \Big(\|Tx - T_0 x_j\|_Y + \|T_0 x_j\|_Y \Big) \\
&= \limsup_{j\to\infty} \|T_0 x_j\|_Y \le \limsup_{j\to\infty} C\|x_j\|_X \\
&\le C\Big(\limsup_{j\to\infty} \|x_j - x\|_X + \|x\|_X \Big) = C\|x\|_X,
\end{aligned}$$

which conclude (2.6.2). □

Using the fact that $\mathscr{S}(\mathbb{R}^n)$ is a dense subspace of $L^p(\mathbb{R}^n)$ for $1 \le p < \infty$, we now show that the restriction of $\mathscr{F} : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ on $L^1(\mathbb{R}^n)$ is consistent with the Fourier transform on $L^1(\mathbb{R}^n)$.

THEOREM 2.6.2. *The Fourier transform is a continuous map from $L^1(\mathbb{R}^n)$ into $C^0(\mathbb{R}^n)$. For any $f \in L^1(\mathbb{R}^n)$ the Fourier transform is given by the usual formula*

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\, dx \quad \textit{for all } \xi \in \mathbb{R}^n. \tag{2.6.4}$$

PROOF. If $f \in \mathscr{S}(\mathbb{R}^n)$ then we already know that $\hat{f} \in C^0(\mathbb{R}^n)$ such that $\|\hat{f}\|_{L^\infty(\mathbb{R}^n)} \leq \|f\|_{L^1(\mathbb{R}^n)}$. This means that $\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to C^0(\mathbb{R}^n)$ is a bounded linear map from a dense subspace of $L^1(\mathbb{R}^n)$ to $C^0(\mathbb{R}^n)$, hence using Lemma 2.6.1 there exists a unique bounded extension

$$\Phi : L^1(\mathbb{R}^n) \to C^0(\mathbb{R}^n) \quad \text{with } \|\Phi(f)\|_{L^\infty(\mathbb{R}^n)} \leq \|f\|_{L^1(\mathbb{R}^n)}.$$

We wish to show that $\Phi = \mathscr{F}|_{L^1(\mathbb{R}^n)}$, where $\mathscr{F}$ is the Fourier transform on $\mathscr{S}'(\mathbb{R}^n)$. For this we take any $f \in L^1(\mathbb{R}^n)$ and choose a sequence $\{f_j\} \subset \mathscr{S}(\mathbb{R}^n)$ such that $f_j \to f$ in $L^1(\mathbb{R}^n)$. Then

$$\mathscr{F} f_j \to \Phi(f) \text{ in } L^\infty(\mathbb{R}^n) \implies \mathscr{F} f_j \to \Phi(f) \text{ in } \mathscr{S}'(\mathbb{R}^n).$$

Using Theorem 2.5.4, we know that $\mathscr{F} f_j \to \mathscr{F} f$ in $\mathscr{S}'(\mathbb{R}^n)$, then the uniqueness of limit gives

$$\mathscr{F} f = \Phi(f) \quad \text{in } \mathscr{S}'(\mathbb{R}^n).$$

The formula (2.6.4) is given by

$$\Phi(f)(\xi) = \lim_{j\to\infty} \hat{f}_j(\xi) = \lim_{j\to\infty} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f_j(x)\, dx = \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\, dx,$$

where the last equality follows since $\|f_j - f\|_{L^1(\mathbb{R}^n)} \to 0$. $\square$

THEOREM 2.6.3 (Plancherel). *The Fourier transform is an isomorphism from $L^2(\mathbb{R}^n)$ onto $L^2(\mathbb{R}^n)$. It is isometric in the sense that*

$$\|\hat{f}\|_{L^2(\mathbb{R}^n)} = (2\pi)^{\frac{n}{2}} \|f\|_{L^2(\mathbb{R}^n)}.$$

The transform is given by

$$\hat{f}(\xi) = \lim_{R\to\infty} \int_{|x|\leq R} f(x) e^{-ix\cdot\xi}\, dx \quad \textit{in } L^2(\mathbb{R}^n). \tag{2.6.5}$$

PROOF. Using the Parseval's identity, we know that $\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is an isometry from a dense subset of $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n)$. Therefore from Lemma 2.6.1, it extends uniquely to an isometry $\Phi : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$. Using a similar argument, we can show that Φ and $\mathscr{F}|_{L^2(\mathbb{R}^n)}$ coincide. For any $f \in L^2(\mathbb{R}^n)$, we have

$$\chi_{B_R} f \to f \quad \text{in } L^2(\mathbb{R}^n) \text{ as } R \to \infty.$$

Hence Parseval's identity gives

$$\int_{|x|\leq R} e^{-ix\cdot\xi} f(x)\, dx = (\chi_{B_R} f)^\wedge \to \hat{f} \quad \text{in } L^2(\mathbb{R}^n) \text{ as } R \to \infty,$$

where we used the fact that $\chi_{B_R} f \in L^1(\mathbb{R}^n)$. $\square$

2.7. The space of compactly supported distributions $\mathscr{E}'(\Omega)$

To study the local behavior of tempered distributions, we introduce the following concepts.

DEFINITION 2.7.1. For any open set $\Omega \subset \mathbb{R}^n$, the distribution $T \in \mathscr{S}'(\mathbb{R}^n)$ is said to *vanish* on Ω, written $T = 0$ in Ω, if

$$T(\varphi) = 0 \quad \text{for all } \varphi \in C_c^\infty(\Omega).$$

Two distributions T_1 and T_2 are said to be *equal* in Ω if $T_1 - T_2$ vanish in Ω.

We recall the following proposition regarding to the partition of unity [**Mit18**, Theorem 14.42]:

PROPOSITION 2.7.2 (Partition of unity for arbitrary open covers). *Let $\{\Omega_j\}_{j\in J}$ be an arbitrary family of open sets in $\mathbb{R}^n$ and set $\Omega := \bigcup_{j\in J} \Omega_j$. Then there exists an at most countable collection $\{\varphi_i\}_{i\in I} \subset C^\infty(\Omega)$ of non-zero functions satisfying the following properties:*

(1) *For every $i \in I$, there exists $j \in J$ such that $\operatorname{supp}(\varphi_i) \subset \Omega_j$;*
(2) *For every $i \in I$, one has $0 \le \varphi_i \le 1$ in Ω;*
(3) *The family of sets $\left\{ x \in \Omega \,\middle|\, \varphi_i(x) \neq 0 \right\}$, indexed by $i \in I$, is locally finite in Ω;* [4]
(4) *$\sum_{i\in I} \varphi_i(x) = 1$ for every $x \in \Omega$.*

The family $\{\varphi_i\}_{i\in I}$ is called a partition of unity subordinate to the family $\{\Omega_j\}_{j\in J}$.

REMARK 2.7.3 (Reindexing). First of all, for those $j \in J$ such that there does not exist i such that $\operatorname{supp}(\varphi_i) \subset \Omega_j$, we define $\psi_j \equiv 0$. Let $\tilde{J} \subset J$ be the index set such that for each $j \in \tilde{J}$ there exists $i \in I$ such that $\operatorname{supp}(\varphi_i) \subset \Omega_j$. Since I is countable, then $\tilde{J}$ is also countable, therefore we can identify $\tilde{J} \cong \mathbb{N}$. We define [5]

$$\begin{aligned} I_1 &:= \left\{ i \in I \,\middle|\, \operatorname{supp}(\varphi_i) \subset \Omega_1 \right\}, \\ I_j &:= \left\{ i \in I \,\middle|\, \operatorname{supp}(\varphi_i) \subset \Omega_j \right\} \setminus I_{j-1} \quad \text{for all } j = 2, 3, \cdots . \end{aligned}$$

We see that I_j are disjoint and $\bigcup_{j\in\mathbb{N}} I_j = I$. We now consider the family of functions $\{\psi_j\}_{j\in J}$ defined by

$$\psi_j = \sum_{i\in I_j} \varphi_i.$$

We see that $\psi_j \in C^\infty(\Omega_j)$, $0 \le \psi_j \le 1$, such that any compact set $K \subset \Omega$ has a neighborhood U where only finite many ψ_j are not identically zero, and

$$\sum_{j\in J} \psi_j(x) = 1 \quad \text{for all } x \in U.$$

We now able to prove the following lemma:

LEMMA 2.7.4. *If $\{\Omega_j\}_{j\in J}$ is a family of open sets in $\mathbb{R}^n$, and if T vanishes in each Ω_j, then T vanishes on $\Omega := \bigcup_{j\in J} \Omega_j$.*

[4] A family $\{A_i\}_{i\in I}$ of subsets of $\mathbb{R}^n$ is said to be locally finite in $E \subset \mathbb{R}^n$ provided every $x \in E$ has a neighborhood $U \subset \mathbb{R}^n$ with the property that the set $\left\{ i \in I \,\middle|\, A_i \cap U \neq \emptyset \right\}$ is finite.

[5] Here we remark that it is possible to have $\operatorname{supp}(\varphi_i) \subset \Omega_{j_1} \cap \Omega_{j_2}$ with $j_1 \neq j_2$.

PROOF. Let $\{\psi_j\}_{j\in J}$ be the family of functions described in Remark 2.7.3. Let $\varphi \in C_c^\infty(\Omega)$ and write $K = \mathrm{supp}\,(\varphi)$. We can now write

$$\varphi = \sum_{j\in J} \psi_j \varphi$$

where only finitely many terms of the sum are nonzero. Thus

$$T(\varphi) = \sum_{j\in J} T(\psi_j\varphi) = 0,$$

using the fact that T vanishes on each Ω_j. □

This lemma ensures the following make sense:

DEFINITION 2.7.5. The *support* of a distribution $T \in \mathscr{S}'(\mathbb{R}^n)$, denoted by $\mathrm{supp}\,(T)$, is the complement of the largest open subset of $\mathbb{R}^n$ where T vanishes.

We now want to give a characterization of the tempered distributions with compact support. Let $\mathscr{E}(\mathbb{R}^n) = C^\infty(\mathbb{R}^n)$. We define

$$[f]_N = \sum_{|\alpha|\le N} \|\partial^\alpha f\|_{L^\infty(B_N(0))} \quad \text{for each } N \in \mathbb{Z}_{\ge 0},$$

which are semi-norms on $\mathscr{E}(\mathbb{R}^n)$. Similarly to Exercise 2.2.4, we induce a metric on $\mathscr{E}(\mathbb{R}^n)$ defined by

$$\mathsf{d}_{\mathscr{E}(\mathbb{R}^n)}(\varphi,\psi) := \sum_{N=0}^\infty 2^{-N} \frac{[\varphi-\psi]_N}{1+[\varphi-\psi]_N} \quad \text{for all } \varphi,\psi \in \mathscr{E}(\mathbb{R}^n), \tag{2.7.1}$$

and that $f_j \to f$ in $\mathscr{E}(\mathbb{R}^n)$ if and only if $\partial^\alpha f_j \to \partial^\alpha f$ uniformly on compact subsets of $\mathbb{R}^n$ for any multi-index α. Therefore, using a similar argument as in Theorem 2.2.6, we have the following:

THEOREM 2.7.6. *Let $\mathsf{d}_{\mathscr{E}(\mathbb{R}^n)}$ be the metric given by (2.7.1). Then $(\mathscr{E}(\mathbb{R}^n), \mathsf{d}_{\mathscr{E}(\mathbb{R}^n)})$ is a Fréchet space space (i.e. complete metric space), and the identity map $\iota : \mathscr{S}(\mathbb{R}^n) \to \mathscr{E}(\mathbb{R}^n)$ is continuous.*

REMARK 2.7.7. Similar to Remark 2.2.7, since $\mathscr{E}(\mathbb{R}^n)$ is a (Grothedieck) nuclear space, then we cannot define a norm on $\mathscr{E}(\mathbb{R}^n)$.

Similar to Definition 2.4.1, we also consider the following definition:

DEFINITION 2.7.8. Let $\mathscr{E}'(\mathbb{R}^n)$ be the set of continuous (w.r.t. the metric (2.7.1)) linear functional on $\mathscr{E}(\mathbb{R}^n)$, i.e. dual space of $\mathscr{E}(\mathbb{R}^n)$. Precisely,

$$\mathscr{E}'(\mathbb{R}^n) := \left\{ T : \mathscr{E}(\mathbb{R}^n) \to \mathbb{C} \;\middle|\; \begin{array}{l} T \text{ linear and } T(\varphi_j) \to 0 \\ \text{whenever } \varphi_j \to 0 \text{ in } \mathscr{E}(\mathbb{R}^n) \end{array} \right\}.$$

Similar to Lemma 2.4.2, we have the following lemma:

LEMMA 2.7.9. *For any $T \in \mathscr{E}'(\mathbb{R}^n)$, there exist $C > 0$ and $N \in \mathbb{N}$ such that*

$$|T(f)| \le C \sum_{|\alpha|\le N} \|\partial^\alpha \varphi\|_{L^\infty(B_N(0))} \quad \textit{for all } f \in \mathscr{E}(\mathbb{R}^n).$$

EXERCISE 2.7.10. Prove Lemma 2.7.9.

Each element $S \in \mathscr{E}'(\mathbb{R}^n)$ induces $T := S \circ \iota \in \mathscr{S}'(\mathbb{R}^n)$, where $\iota : \mathscr{S}(\mathbb{R}^n) \to \mathscr{E}(\mathbb{R}^n)$ is given in Theorem 2.7.6. Since $\mathscr{S}(\mathbb{R}^n)$ is dense in $\mathscr{E}(\mathbb{R}^n)$ (for any $f \in \mathscr{E}(\mathbb{R}^n)$ just take a sequence $\{f_j\} \subset \mathscr{S}(\mathbb{R}^n)$ such that $f_j = f$ in $B_j(0)$), then S induces a unique $T = S \circ \iota$. Moreover, the mapping $\iota : \mathscr{S}(\mathbb{R}^n) \to \mathscr{E}(\mathbb{R}^n)$ is continuous, therefore $\mathscr{E}'(\mathbb{R}^n)$ is continuously embedded in $\mathscr{S}'(\mathbb{R}^n)$, i.e. the topology are compatible. The following theorem is the main result of this section, it states that $\mathscr{E}'(\mathbb{R}^n)$ is exactly the compactly supported (tempered) distributions:

THEOREM 2.7.11. *Let $T \in \mathscr{S}'(\mathbb{R}^n)$. The following are equivalent:*

(1) *T has compact support,*
(2) *T can be extended to an element in $\mathscr{E}'(\mathbb{R}^n)$.*

REMARK 2.7.12. Accordingly, we can define Fourier transform on $\mathscr{E}'(\mathbb{R}^n)$ as in Definition 2.5.1.

PROOF OF THEOREM 2.7.11. **(1) $\Longrightarrow$ (2).** Suppose $T \in \mathscr{S}'(\mathbb{R}^n)$ has compact support, and choose $\psi \in C_c^\infty(\mathbb{R}^n)$ so that $\psi = 1$ on some open set containing $\operatorname{supp}(T)$[6], and we denote $K := \operatorname{supp}(\psi)$. Then we see that

$$T(\varphi) = T(\psi\varphi) \quad \text{for all } \varphi \in \mathscr{S}(\mathbb{R}^n),$$

and we can extend T on $\mathscr{E}(\mathbb{R}^n)$ by defining

$$T(f) = T(\psi f) \quad \text{for all } f \in \mathscr{E}(\mathbb{R}^n).$$

We now want to show $T \in \mathscr{E}'(\mathbb{R}^n)$. Since $T \in \mathscr{S}'(\mathbb{R}^n)$, using Lemma 2.4.2 there exist C and N such that

$$|T(\varphi)| \leq C \sum_{|\alpha| \leq N} \|\langle x \rangle^N \partial^\alpha \varphi\|_{L^\infty(\mathbb{R}^n)} \quad \text{for all } \varphi \in \mathscr{S}(\mathbb{R}^n).$$

Since for any $f \in \mathscr{E}(\mathbb{R}^n)$, the function $\psi f \in C_c^\infty(\mathbb{R}^n)$ satisfies $\operatorname{supp}(\psi f) \subset K$, this implies that

$$|T(f)| = |T(\psi f)| \leq C' \sum_{|\alpha| \leq N} \|\partial^\alpha(\psi f)\|_{L^\infty(\mathbb{R}^n)} \leq C'' \sum_{|\alpha| \leq N} \|\partial^\alpha f\|_{L^\infty(K)},$$

which implies that $T \in \mathscr{E}'(\mathbb{R}^n)$.

(2) $\Longrightarrow$ (1). For the converse, we suppose that $T \in \mathscr{E}'(\mathbb{R}^n)$. Using Lemma 2.7.9, there exist $C > 0$ and $N \in \mathbb{N}$ such that

$$|T(f)| \leq C \sum_{|\alpha| \leq N} \|\partial^\alpha \varphi\|_{L^\infty(B_N(0))} \quad \text{for all } f \in \mathscr{E}(\mathbb{R}^n).$$

If T does not have compact support, then for any M there is a function $\varphi \in C_c^\infty(\mathbb{R}^n \setminus \overline{B_M(0)})$ for which $T(\varphi) \neq 0$, this clearly contradicts the above inequality. $\square$

The extreme case of a distribution with compact support is one whose support is a point. The following theorem characterizes all distributions with support consisting of one point, which can be found in [**FJ98**, Theorem 3.2.1]:

[6]Sometimes, we simply say $\psi = 1$ near $\operatorname{supp}(T)$.

THEOREM 2.7.13. *Suppose that $T \in \mathscr{S}'(\mathbb{R}^n)$ such that $\operatorname{supp}(T) = \{0\}$. Then there is a non-negative integer N such that*

$$T = \sum_{|\alpha| \leq N} c_\alpha \partial^\alpha \delta,$$

where c_α are complex numbers.

For each $u \in \mathscr{S}'(\mathbb{R}^n)$, we can define the distributional Laplacian

$$\Delta u := \sum_{j=1}^n \partial_j^2 u \in \mathscr{S}'(\mathbb{R}^n)$$

by using Proposition 2.4.10. Then we say that $\Delta u = 0$ in distribution sense if $(\Delta u)(\varphi) = 0$ for all $\varphi \in \mathscr{S}(\mathbb{R}^n)$. As a consequence, we obtain a generalization of the standard Liouville theorem which states that any bounded harmonic function is constant.

COROLLARY 2.7.14 (Liouville-type theorem for distributions). *If $u \in \mathscr{S}'(\mathbb{R}^n)$ satisfies $\Delta u = 0$ in distribution sense, then u is a polynomial.*

REMARK 2.7.15. The only bounded polynomial is constant function. Therefore, if $u \in L^\infty(\mathbb{R}^n)$ satisfying $\Delta u = 0$ (in distribution sense), then $u \equiv$ constant.

PROOF. Using Proposition 2.5.5, taking Fourier transform in the equation $\Delta u = 0$ implies that $|\xi|^2 \hat{u} = 0$ in distribution sense. Hence we know that

$$\hat{u}(\varphi) = |\xi|^2 \hat{u}(|\xi|^{-2}\varphi) = 0 \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}^n \setminus \{0\}),$$

which shows that $\operatorname{supp}(\hat{u}) = \{0\}$. Using Theorem 2.7.13, we know that

$$\hat{u} = \sum_{|\alpha| \leq N} c_\alpha \partial^\alpha \delta.$$

Taking the inverse Fourier transform and using Proposition 2.5.5, we see that u is a polynomial. $\square$

2.8. The space of test functions $\mathscr{D}(\Omega)$ and distributions $\mathscr{D}'(\Omega)$

Fixing any compact set K in $\mathbb{R}^n$, we denote

$$\mathscr{D}_K := \left\{ \varphi \in C^\infty(\mathbb{R}^n) \,\middle|\, \operatorname{supp}(\varphi) \subset K \right\}.$$

For each fixed $N \in \mathbb{Z}_{\geq 0}$, it is easy to see that

$$\|\varphi\|_{N,K} := \sum_{|\alpha| \leq N} \|\partial^\alpha \varphi\|_{L^\infty(K)}$$

is a norm defined on $\mathscr{D}_K$. Similar to (2.2.5), Exercise 2.2.4 and Theorem 2.2.6, we have the following lemma:

LEMMA 2.8.1. *$\mathscr{D}_K$ is a Fréchet space (i.e. complete metric space) equipped with the metric*

$$\mathsf{d}_{\mathscr{D}_K}(\varphi, \psi) := \sum_{N=0}^\infty 2^{-N} \frac{\|\varphi - \psi\|_{N,K}}{1 + \|\varphi - \psi\|_{N,K}}. \tag{2.8.1}$$

REMARK 2.8.2. Similar to Remark 2.2.7, since $\mathscr{D}_K$ is a (Grothedieck) nuclear space, then we cannot define a norm on $\mathscr{D}_K$.

Let $\Omega \subset \mathbb{R}^n$ be an open set. We now define the set of test functions by

$$\mathscr{D}(\Omega) := C_c^\infty(\Omega) \equiv \bigcup_{K \subset \Omega \text{ compact}} \mathscr{D}_K.$$

Similarly, let us introduce the norms

$$\|\varphi\|_N \equiv \|\varphi\|_{N,\Omega} := \sum_{|\alpha| \le N} \|\partial^\alpha \varphi\|_{L^\infty(\Omega)}.$$

Similarly, we can equip $\mathscr{D}(\Omega)$ with the metric

$$\mathsf{d}_{\mathscr{D}(\Omega)}(\varphi, \psi) := \sum_{N=0}^\infty 2^{-N} \frac{\|\varphi - \psi\|_N}{1 + \|\varphi - \psi\|_N}.$$

However, this topology has the disadvantage of *not being complete*.

EXAMPLE 2.8.3. Take $n = 1$ and $\Omega = \mathbb{R}$. Let $\phi \in \mathscr{D}(\mathbb{R})$ with $\mathrm{supp}\,(\phi) \subset [0,1]$ and $\phi > 0$ in $(0,1)$. Define

$$\psi_m(x) := \phi(x-1) + \frac{1}{2}\phi(x-2) + \cdots + \frac{1}{m}\phi(x-m).$$

Note that $\{\psi_m\}$ is a Cauchy sequence in $(\mathscr{D}(\mathbb{R}), \mathsf{d}_{\mathscr{D}(\mathbb{R})})$, but the limit $\lim_{m\to\infty} \psi_m(x)$ does not have compact support.

We usually equip $\mathscr{D}(\Omega)$ by another (locally convex) topology τ in which Cauchy sequences do converge. The fact the topology τ is not metrizable is only a minor inconvenience. The following fact can be found in [**Rud91**, Chapter 6] (see also [**Mit18**, Appendix 14.1] as well as Theorem 3.6.2 in Mikko Salo's lecture note):

THEOREM 2.8.4. *There exists a topology τ on $\mathscr{D}(\Omega)$ which is a vector space topology (i.e. addition and scalar multiplication are continuous operations) and has the following properties:*

1. *A sequence $\{\varphi_j\}$ in $\mathscr{D}(\Omega)$ converges if and only if $\{\varphi_j\} \subset \mathscr{D}_K$ for some fixed compact set $K \subset \Omega$ and $\{\varphi_j\}$ converges in $\mathscr{D}_K$,*
2. *$\mathscr{D}(\Omega)$ is a complete topological space (i.e. any Cauchy sequence, or net, in $\mathscr{D}(\Omega)$ converges).*

We now introduce the usual operations on the space $\mathscr{D}(\Omega)$, which can be proved similar to the case of Schwartz functions:

PROPOSITION 2.8.5. *Let $\Omega \subset \mathbb{R}^n$ be an open set, and we consider the topological space $(\mathscr{D}(\Omega), \tau)$, where τ is the topology given in Theorem 2.8.4. If $f \in C^\infty(\Omega)$, then the following operations are continuous maps from $(\mathscr{D}(\Omega), \tau)$ into $(\mathscr{D}(\Omega), \tau)$:*

1. ***Conjugation.*** *$\varphi \mapsto \overline{\varphi}$,*
2. ***Derivative.*** *$\varphi \mapsto \partial^\alpha \varphi$,*
3. ***Multiplication.*** *$\varphi \mapsto f\varphi$.*

If $\Omega = \mathbb{R}^n$, then additionally the following operations are continuous from $(\mathscr{D}(\mathbb{R}^n), \tau)$ into $(\mathscr{D}(\mathbb{R}^n), \tau)$:

1. ***Reflection.*** *$\varphi \mapsto \tilde{\varphi}$ with $\tilde{\varphi}(x) = \varphi(-x)$,*
2. ***Translation.*** *$\varphi \mapsto \tau_{x_0}\varphi$ with $\tau_{x_0}\varphi(x) = \varphi(x - x_0)$,*

The following theorem is a special case of [**Rud91**, Chapter 6] (see also [**Mit18**, Appendix 14.6] as well as Theorem 3.6.3 in Mikko Salo's lecture note):

THEOREM 2.8.6. *Let T be a linear map from $(\mathscr{D}(\Omega),\tau)$ into $(\mathbb{C},|\cdot|)$. The following statements are equivalent:*

(1) *T is continuous (with respect to the topology τ given in Theorem 2.8.4),*
(2) *$\lim_{j\to\infty} T(\varphi_j)=0$ whenever $\varphi_j\to 0$ in $(\mathscr{D}(\Omega),\tau)$,*
(3) *$T|_{\mathscr{D}_K}$ is continuous for each compact subset K in Ω wit respect to the metric d_K given in (2.8.1).*

Here and after, we do not explicit state the topology τ of the space of test function $\mathscr{D}(\Omega)$. The following definition suggested by Theorem 2.8.6:

DEFINITION 2.8.7 (Distributions). The set of continuous linear functionals on $\mathscr{D}(\Omega)$ is denoted by $\mathscr{D}'(\Omega)$ and its elements are called *distributions* on Ω.

The following fact can be proved using similar ideas as in Lemma 2.4.2:

LEMMA 2.8.8. *If $T\in\mathscr{D}'(\Omega)$, then for any compact set $K\subset\Omega$ there exist $C>0$ and $N>0$ (depending on K) such that*

$$|T(\varphi)|\le C\sum_{|\alpha|\le N}\|\partial^\alpha\varphi\|_{L^\infty(K)}\quad\textit{for all } \varphi\in\mathscr{D}_K. \tag{2.8.2}$$

DEFINITION 2.8.9. If there is a fixed N such that (2.8.2) is satisfies for any K (i.e. N is independent of K), then such distribution T is said to be of *order* $\le N$, and if N is the least such integer then T is said to be of *order* N.

EXAMPLE 2.8.10. Each element $f\in L^1_{\rm loc}(\Omega)$ can be identify with $T_f\in\mathscr{D}'(\Omega)$ defined by

$$T_f(\varphi):=\int_{\mathbb{R}^n} f(x)\varphi(x)\,dx,$$

by using Theorem 2.8.6 with the estimate

$$|T_f(\varphi)|\le\int_K|f(x)\varphi(x)|\,dx\le\|\varphi\|_{L^\infty(K)}\int_K|f(x)|\,dx\quad\text{for all }\varphi\in\mathscr{D}_K.$$

In particular, any continuous function gives rise to a distribution.

EXAMPLE 2.8.11 (Measures). Let μ be either a complex Borel measure on Ω, or a Borel positive measure on Ω that is locally finite (i.e. satisfies $|\mu|(K)<\infty$ for every compact $K\subset\Omega$)[7]. Consider the linear mapping $T_\mu:\mathscr{D}(\Omega)\to\mathbb{C}$ defined by

$$T_\mu(\varphi)=\int_\Omega\varphi(x)\,d\mu(x)\quad\text{for all }\varphi\in\mathscr{D}(\Omega)\equiv C_c^\infty(\Omega).$$

It is easy to see that

$$|T_\mu(\varphi)|\le|\mu|(K)\|\varphi\|_{L^\infty(K)}\quad\text{for all }\varphi\in\mathscr{D}_K,$$

where the positive Borel measure $|\mu|$ is the total variation of μ. Therefore from Theorem 2.8.6 we know that $T_\mu\in\mathscr{D}'(\Omega)$ and it is of order 0.

Conversely, if $T\in\mathscr{D}'(\mathbb{R}^n)$ has order 0, using [**Mit18**, Proposition 2.16], in particular T determines a (necessarily unique) measure μ. Precisely, the statement reads:

[7]For complex measures, the measure can take on complex values, infinite values are not allowed. In contrast, infinite values are allowed for positive measures. In particular, a finite positive measure is a special case of a complex measures.

THEOREM 2.8.12. *Let $T \in \mathscr{D}'(\Omega)$ has order 0. Then the distribution T extends uniquely to linear map $T_\mu : C_c^0(\Omega) \to \mathbb{C}$ that is locally bounded in the following sense: for each compact set $K \subset \Omega$ there exists $C_K > 0$ such that*

$$|T_\mu(\varphi)| \leq C_K \|\varphi\|_{L^\infty(K)} \quad \textit{for all } \varphi \in C_c^0(\Omega).$$

In addition, the functional T_μ satisfies the following properties:

(1) *Let $\{K_j\}_{j\in\mathbb{N}}$ be a* compact exhaustion[8] *of Ω, that is, a sequence of compact subsets of Ω satisfying $K_j \subset \operatorname{int}(K_{j+1})$ and $\Omega = \bigcup_{j=1}^\infty K_j$. Then there exists a sequence of complex regular Borel measures μ_j on K_j such that*
   (a) *$\mu_j(E) = \mu_\ell(E)$ for every $\ell \in \mathbb{N}$, every Borel set $E \subset \operatorname{int}(K_\ell)$ and every $j \geq \ell$,*
   (b) *for each $j \in \mathbb{N}$ one has*

$$T_\mu(\varphi) = \int_{K_j} \varphi \, d\mu_j \quad \textit{for all } \varphi \in C^0(\Omega) \textit{ with } \operatorname{supp}(\varphi) \subset K_j.$$

(2) *There exist two Radon measures μ_1, μ_2, taking Borel sets from Ω into $[0, \infty]$, such that*

$$\Re(T_\mu(\varphi)) = \int_\Omega \varphi \, d\mu_1 - \int_\Omega \varphi \, d\mu_2 \quad \textit{for all real-valued } \varphi \in C_c^0(\Omega).$$

   Furthermore, a similar conclusion is valid for $\Im(T_\mu(\varphi))$.

Hence we can identify Radon measures with distributions of order 0.

EXAMPLE 2.8.13 (Continuous embedding). Clearly, $\mathscr{S}'(\mathbb{R}^n)$ is a subset of $\mathscr{D}'(\mathbb{R}^n)$. We now want to show the topology are compatible, that is, we want to show that $\mathscr{S}'(\mathbb{R}^n)$ is continuously embedded in $\mathscr{D}'(\mathbb{R}^n)$: We want to show if $T \in \mathscr{S}'(\mathbb{R}^n)$ and $\varphi_j \to 0$ in $\mathscr{D}(\mathbb{R}^n)$, then $T(\varphi_j) \to 0$. Using Theorem 2.8.4, there exists a compact set $K \subset \mathbb{R}^n$ such that $\operatorname{supp}(\varphi_j) \subset K$ for all j and $\partial^\alpha \varphi_j \to 0$ uniformly on K for any multi-index α. It is easy to see that

$$\lim_{j\to\infty} \|\langle x \rangle^N \partial^\alpha \varphi_j\|_{L^\infty(K)} = 0 \quad \text{for any } N \text{ and } \alpha,$$

showing that $\varphi_j \to 0$ in $\mathscr{S}(\mathbb{R}^n)$. Since $T \in \mathscr{S}'(\mathbb{R}^n)$, then $T(\varphi_j) \to 0$. Together with Theorem 2.7.11, we have the following continuous embedding:

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

In particular, the distribution derivative in Proposition 2.4.10 is a special case of the following definition:

DEFINITION 2.8.14. For any $T \in \mathscr{D}'(\Omega)$, the *distribution derivative* $\partial^\alpha T \in \mathscr{D}'(\Omega)$ of T is defined by

$$(\partial^\alpha T)(\varphi) := (-1)^{|\alpha|} T(\partial^\alpha \varphi) \quad \text{for all } \varphi \in \mathscr{D}(\Omega).$$

REMARK 2.8.15. Despite some authors also called it the *weak derivative*, remember not to be confused with Definition 1.1.1. We again refer to Remark 2.4.11.

PROPOSITION 2.8.16. *$\mathscr{D}(\mathbb{R}^n)$ is dense in $\mathscr{S}(\mathbb{R}^n)$. Consequently, from Example* 2.8.13*, we know that*

$$u, v \in \mathscr{S}'(\mathbb{R}^n) \textit{ with } u = v \textit{ in } \mathscr{D}'(\mathbb{R}^n) \implies u = v \textit{ in } \mathscr{S}'(\mathbb{R}^n).$$

---

[8]For example, if we define $K_j := \Omega \setminus \left(\{ \, x \mid |x| > m \, \} \cup \bigcup_{z\in\mathbb{R}^n\setminus\Omega} B_{1/m}(z) \right)$, then $\{K_j\}_{j\in\mathbb{N}}$ is a compact exhaustion of Ω.

PROOF. Fixing any $\varphi \in \mathscr{S}(\mathbb{R}^n)$. We choose $\rho \in C_c^\infty(\mathbb{R}^n)$ such that $\rho = 1$ for $|x| \leq 1$ and set

$$\varphi_j(x) = \rho(x/j)\varphi(x) \quad \text{where } j \in \mathbb{N}.$$

Then there are constants C_N such that

$$\sum_{|\alpha|\leq N, |\beta|\leq N} [\varphi - \varphi_j]_{\alpha,\beta} \leq C_N \sum_{|\alpha|\leq N, |\beta|\leq N} \sup_{|x|\geq j} |x^\alpha \partial^\beta \varphi(x)|.$$

Since $\varphi \in \mathscr{S}(\mathbb{R}^n)$, then the right-hand-side converges to 0 as $j \to \infty$. □

Since $\partial^\alpha : \mathscr{D}(\Omega) \to \mathscr{D}(\Omega)$ is continuous (Proposition 2.8.5), then using Theorem 2.8.6 we know that $\partial^\alpha T$ is continuous (with respect to the topology τ given in Theorem 2.8.4), that is,

$$\partial^\alpha T \in \mathscr{D}'(\Omega) \quad \text{for any multi-index } \alpha,$$

showing that Definition 2.8.14 is well-defined for all multi-index α. In other words, any distribution has well defined derivatives of any order even if it arises from a function which is not differentiable in the classical sense (as well as in the sense of Definition 1.1.1).

EXERCISE 2.8.17. Prove that for every $c \in \mathbb{R}$ one has

$$(e^{-c|x|})' = -ce^{-cx}H(x) + ce^{cx}H(-x) \quad \text{in } \mathscr{D}'(\mathbb{R}).$$

EXERCISE 2.8.18. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} x \ln|x| - x & \text{for } x \neq 0, \\ 0 & \text{for } x = 0. \end{cases}$$

Prove that f is a continuous function and compute its distributional derivative (of order 1) f'.

EXERCISE 2.8.19. Let $n = 1$ and $T = \sum_{j=1}^\infty \partial^j \delta_j \in \mathscr{D}'(\mathbb{R})$, that is,

$$T(\varphi) = \sum_{j=1}^\infty (-1)^j \varphi^{(j)}\Big|_{x=j} \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R}).$$

Show that T does not have finite order.

An immediate consequence of the definition is that this map is also sequentially continuous:

THEOREM 2.8.20 (Theorem 2.1.1 in [**FJ98**]). *If $\{T_j\}$ be a sequence of distributions in $\mathscr{D}'(\Omega)$ converges to T in $\mathscr{D}'(\Omega)$. For each multi-index α, one has*

$$\partial^\alpha T_j \to \partial^\alpha T \quad \textit{in } \mathscr{D}'(\Omega).$$

PROOF. It follows from the Theorem 2.8.6 and Definition 2.8.14 that as $j \to \infty$,

$$\lim_{j\to\infty} \partial^\alpha T_j(\phi) = (-1)^{|\alpha|} \lim_{j\to\infty} T_j(\partial^\alpha \phi) = (-1)^{|\alpha|} T(\partial^\alpha \phi) = \partial^\alpha T(\phi)$$

for all $\phi \in \mathscr{D}(\Omega) \equiv C_c^\infty(\Omega)$. □

The final operation on distributions that we wish to introduce here is multiplication by functions. This is easy to define since if $f \in C^\infty(\Omega)$ then fT is a well-defined distribution if

$$(fT)(\varphi) := T(f\varphi) \quad \text{for all } \varphi \in \mathscr{D}(\Omega) \equiv C_c^\infty(\Omega).$$

We summarize that we have done in the following proposition:

PROPOSITION 2.8.21. *If $f \in C^\infty(\Omega)$, then the following operations are well-defined maps from $\mathscr{D}'(\Omega)$ into $\mathscr{D}'(\Omega)$:*

(1) ***Reflection.*** $\tilde{T}(\varphi) = T(\tilde{\varphi})$ *with* $\tilde{\varphi}(x) = \varphi(-x)$,
(2) ***Conjugation.*** $\overline{T}(\varphi) = \overline{T(\overline{\varphi})}$,
(3) ***Translation.*** $(\tau_{x_0}T)(\varphi) = T(\tau_{-x_0}\varphi)$ *with* $\tau_{-x_0}\varphi(x) = \varphi(x + x_0)$,
(4) ***Distributional derivative.*** $(\partial^\alpha T)(\varphi) = (-1)^{|\alpha|}T(\partial^\alpha \varphi)$,
(5) ***Multiplication.*** $(fT)(\varphi) = T(f\varphi)$.

To study the local behavior of distributions we introduce the following concepts similar to Definition 2.7.1:

DEFINITION 2.8.22. For any open set $V \subset \Omega$, the distribution $T \in \mathscr{D}'(\Omega)$ is said to *vanish* on V, written $T = 0$ in V, if

$$T(\varphi) = 0 \quad \text{for all } \varphi \in \mathscr{D}(V) \equiv C_c^\infty(V).$$

Two distributions T_1 and T_2 are said to be *equal* in V if $T_1 - T_2$ vanish in V.

It is an important fact that if the local behavior of a distribution is known at each point, then the distribution is uniquely determined globally. The proof uses a partition of unity.

THEOREM 2.8.23. *Let $\{\Omega_i\}$ be an open cover of Ω and let $\{T_i\}$ be a family of distributions such that $T_i \in \mathscr{D}'(\Omega_i)$, and suppose that for any Ω_i, Ω_j with $\Omega_i \cap \Omega_j \neq \emptyset$, one has*

$$T_i = T_j \quad \textit{on } \Omega_i \cap \Omega_j.$$

Then there is a unique $T \in \mathscr{D}'(\Omega)$ for which $T = T_i$ on each Ω_i.

PROOF. Let $\{\psi_i\}$ be the partition of unity subordinate to $\{\Omega_i\}$ as in Remark 2.7.3 (Theorem 2.7.2). We see that $\psi_i \in C^\infty(\Omega_i)$, $0 \leq \psi_i \leq 1$, such that any compact set $K \subset \Omega$ has a neighborhood U where only finite many ψ_i are not identically zero, and

$$\sum_i \psi_i(x) = 1 \quad \text{for all } x \in U.$$

Accordingly, we define the distribution T by

$$T(\varphi) = \sum_i T_i(\psi_i \varphi) \quad \text{for all } \varphi \in \mathscr{D}(\Omega) \equiv C_c^\infty(\Omega).$$

Then we can easily verify that $T = T_i$ in Ω_i by testing $\varphi \in \mathscr{D}_K$ with compact subset $K \subset \Omega_i$. Using similar ideas, the uniqueness also follows. □

2.9. Convolution of functions

DEFINITION 2.9.1. The *convolution* of two measurable functions $f, g : \mathbb{R}^n \to \mathbb{C}$ is the function $f * g : \mathbb{R}^n \to \mathbb{C}$ given by

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x - y)\, dy$$

provided that the integral exists almost everywhere.

REMARK 2.9.2. A change of variable gives that $f * g = g * f$. Moreover, we also see that

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x - y)\, dy = \int_{\mathbb{R}^n} f(y)\tilde{g}(y - x)\, dy = \int_{\mathbb{R}^n} f(y)(\tau_x \tilde{g})(y)\, dy. \tag{2.9.1}$$

We denote[9]

$$\begin{aligned}
L_{\mathrm{poly}}(\mathbb{R}^n) &:= \left\{ f:\mathbb{R}^n\to\mathbb{C} \;\middle|\; \begin{array}{l}\text{there exists } C>0 \text{ such that}\\ |f(x)|\le C\langle x\rangle^N \text{ for a.e. } x\in\mathbb{R}^n\end{array} \right\},\\
C_{\mathrm{poly}}(\mathbb{R}^n) &:= \left\{ f\in L_{\mathrm{poly}}(\mathbb{R}^n) \;\middle|\; f \text{ is continuous} \right\},\\
C_{\infty}(\mathbb{R}^n) &:= \left\{ f\in C^0(\mathbb{R}^n) \;\middle|\; \langle x\rangle^N f\in L^\infty(\mathbb{R}^n) \text{ for all } N\in\mathbb{N} \right\},\\
C_{\infty}^k(\mathbb{R}^n) &:= \left\{ f:\mathbb{R}^n\to\mathbb{C} \;\middle|\; \partial^\alpha f\in C_\infty(\mathbb{R}^n) \text{ for all } |\alpha|\le k \right\},\\
C_{\mathrm{poly}}^k(\mathbb{R}^n) &:= \left\{ f:\mathbb{R}^n\to\mathbb{C} \;\middle|\; \partial^\alpha f\in C_{\mathrm{poly}}(\mathbb{R}^n) \text{ for all } |\alpha|\le k \right\}.
\end{aligned}$$

We will first prove the following theorem to ensure the well-definedness of convolution:

THEOREM 2.9.3. *The convolution is a mapping*

(1) $L_{\mathrm{loc}}^1(\mathbb{R}^n)\times C_c^k(\mathbb{R}^n)\to C^k(\mathbb{R}^n)$,
(2) $C^j(\mathbb{R}^n)\times C_c^k(\mathbb{R}^n)\to C^{j+k}(\mathbb{R}^n)$,
(3) $C_c^j(\mathbb{R}^n)\times C_c^k(\mathbb{R}^n)\to C_c^{j+k}(\mathbb{R}^n)$,
(4) $L_{\mathrm{poly}}(\mathbb{R}^n)\times C_\infty^k(\mathbb{R}^n)\to C_{\mathrm{poly}}^k(\mathbb{R}^n)$,
(5) $C_{\mathrm{poly}}^j(\mathbb{R}^n)\times C_\infty^k(\mathbb{R}^n)\to C_{\mathrm{poly}}^{j+k}(\mathbb{R}^n)$,
(6) $C_\infty^j(\mathbb{R}^n)\times C_\infty^k(\mathbb{R}^n)\to C_\infty^{j+k}(\mathbb{R}^n)$.

In addition, we have

$$\partial^{\alpha+\beta}(f*g)=(\partial^\alpha f)*(\partial^\beta g) \quad \textit{whenever } |\alpha|\le j \textit{ and } |\beta|\le k. \tag{2.9.2}$$

We choose $j=0$ in (1) and (4).

We need the following auxiliary lemma to prove Theorem 2.9.3:

LEMMA 2.9.4. *Given any positive integer N and a compact set $K\subset\mathbb{R}^n$, we have*

$$\langle x-y\rangle^N\le C\langle x\rangle^N \quad \textit{for all } y\in K \tag{2.9.3}$$

for some constant $C=C_{K,N}>0$. Consequently, if $\langle\cdot\rangle^N f\in L^\infty(\mathbb{R}^n)$, then there exists a constant $C=C_{K,N}>0$ such that

$$\sup_{y\in K}\sup_{x\in\mathbb{R}^n}\left|\langle x\rangle^N f(x+y)\right|\equiv\sup_{y\in K}\sup_{x\in\mathbb{R}^n}\left|\langle x-y\rangle^N f(x)\right|\le C\|\langle x\rangle^N f\|_{L^\infty(\mathbb{R}^n)}.$$

PROOF. Note that there exists $R=R(K)>0$ such that $K\subset\overline{B_R(0)}$. We first consider the case when $N=2m$ to be an even integer. The expression

$$\langle x-y\rangle^{2m}=(1+|x-y|^2)^m=\big((1+|x|^2)+(-2x\cdot y+|y|^2)\big)^m$$

may be expanded using binomial theorem into

$$\langle x-y\rangle^{2m}=\sum_{j=0}^m\binom{m}{j}(1+|x|^2)^{m-j}(-2x\cdot y+|y|^2)^j.$$

The condition $|y|\le R$ implies that $|-2x\cdot y+|y|^2|\le C_R\langle x\rangle$. We thus have the estimate

$$\langle x-y\rangle^{2m}\le C\langle x\rangle^{2m} \quad \text{for all } y\in K. \tag{2.9.4}$$

If $N=2m+1$ is an odd integer, we write

$$\langle x-y\rangle^{2m+1}=\left(\langle x-y\rangle^{2m}\right)^{\frac{2m+1}{2m}}\le\left(C\langle x\rangle^{2m}\right)^{\frac{2m+1}{2m}}\le C'\langle x\rangle^{2m+1} \quad \text{for all } y\in K. \tag{2.9.5}$$

[9] Some authors denote $L^0(\Omega)\equiv L(\Omega)$ be the set of measurable functions $f:\Omega\to\mathbb{C}$ with $|f(x)|<\infty$ for a.e. $x\in\Omega$.

Combining (2.9.4) and (2.9.5), we conclude (2.9.3). □

With Lemma 2.9.4 at hand, we are now ready to prove Theorem 2.9.3.

Proof of Theorem 2.9.3 **(1)**. Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $g \in C_c^k(\mathbb{R}^n)$. For each fixed $x \in \mathbb{R}^n$, using (2.9.1) we know that

$$(f * g)(x) = \int_{\mathrm{supp}\,(\tau_x \tilde{g})} f(y)(\tau_x \tilde{g})(y)\,dy \text{ is well-defined.}$$

On the other hand, for each $h \in \mathbb{R} \setminus \{0\}$, we see that

$$\frac{(f*g)(x+he_j) - (f*g)(x)}{h} = \int_{\mathbb{R}^n} f(y) \frac{g(x-y+he_j) - g(x-y)}{h}\,dy. \tag{2.9.6}$$

We see that there exists a compact set K such that

$$\bigcup_{|h|\leq 1} \mathrm{supp}\left(\frac{g(x-\cdot+he_j) - g(x-\cdot)}{h}\right) \subset K.$$

In addition, for each $|h| \leq 1$, using Taylor's theorem (in this particular case, simply the mean value theorem), we have

$$\frac{g(x-y+he_j) - g(x-y)}{h} = \frac{\partial g}{\partial x_j}(x-y+\theta e_j) \quad \text{for some } |\theta| \leq 1. \tag{2.9.7}$$

Since $\frac{\partial g}{\partial x_j}$ is uniformly bounded, then we can apply Lebesgue dominated convergence theorem on (2.9.6) to obtain

$$\frac{\partial (f*g)}{\partial x_j}(x) = \left(f * \frac{\partial g}{\partial x_j}\right)(x). \tag{2.9.8}$$

Iterating this argument gives that

$$\partial^\beta (f*g) = f * (\partial^\beta g) \quad \text{for all } |\beta| \leq k$$

and $f * g \in C^k(\mathbb{R}^n)$. □

Proof of Theorem 2.9.3 **(2)**. The same argument as in (1) shows that

$$\partial^\alpha (f*g) = (\partial^\alpha f) * g \quad \text{for all } |\alpha| \leq j,$$

and we also conclude (2.9.2). □

Proof of Theorem 2.9.3 **(3)**. Differentiability follows from (2), and the support condition follows from the inclusion

$$\mathrm{supp}\,(f*g) \subset \overbrace{\underbrace{\mathrm{supp}\,(f)}_{\text{compact}} + \underbrace{\mathrm{supp}\,(g)}_{\text{compact}}}^{\text{closed}}.$$

This fact can be shown by noting that: if $x \notin \mathrm{supp}\,(f) + \mathrm{supp}\,(g)$, then

$$y \in \mathrm{supp}\,(f) \implies x - y \notin \mathrm{supp}\,(g),$$

and then $(f*g)(x) = 0$ by the definition of convolution, i.e. $x \notin \mathrm{supp}\,(f*g)$. □

Exercise 2.9.5. Show that if $A \subset \mathbb{R}^n$ is compact and $B \subset \mathbb{R}^n$ is closed, then $A + B$ is a closed set in $\mathbb{R}^n$.

Proof of Theorem 2.9.3 **(4)**. Let $f \in L_{\rm poly}(\mathbb{R}^n)$ and $g \in C_\infty^k(\mathbb{R}^n)$. Then $\langle \cdot \rangle^{-N} f \in L^1(\mathbb{R}^n)$ for some large enough N, and for any fixed $x \in \mathbb{R}^n$, by using (2.9.1) we have

$$|(f*g)(x)| \le \int_{\mathbb{R}^n} |f(y)g(x-y)|\,dy \le \|\langle \cdot \rangle^{-N} f\|_{L^1(\mathbb{R}^n)} \|\langle \cdot \rangle^N \tau_x \tilde{g}\|_{L^\infty(\mathbb{R}^n)},$$

which shows that $f*g$ is well-defined. On the other hand, if $|h| \le 1$, using (2.9.7) we have

$$\left| f(y) \frac{g(x-y+he_j) - g(x-y)}{h} \right| \le \overbrace{|\langle y \rangle^{-N} f(y)|}^{\in L^1(\mathbb{R}^n)} \overbrace{\left| \langle y \rangle^N \frac{\partial g}{\partial x_j}(x-y+\theta e_j) \right|}^{\text{uniformly bounded by Lemma 2.9.4}} \quad \text{for some } |\theta| \le 1.$$

Then we can apply Lebesgue dominated convergence theorem on (2.9.6) to obtain (2.9.8). Similar to (1), it follows that $f*g \in C^k(\mathbb{R}^n)$.

To show $f*g \in C_{\rm poly}^k(\mathbb{R}^n)$, using an iterative argument, it is suffice to show $f*g \in C_{\rm poly}(\mathbb{R}^n)$. We see that

$$\begin{aligned} |\langle x \rangle^{-N} (f*g)(x)| &= \left| \langle x \rangle^{-N} \int_{\mathbb{R}^n} f(x-y) g(y)\,dy \right| \\ &\le \int_{\mathbb{R}^n} \langle x-y \rangle^{-N} |f(x-y)| \frac{\langle x-y \rangle^N}{\langle x \rangle^N} |g(y)|\,dy \\ &\le C \|\langle \cdot \rangle^{-N} f\|_{L^1(\mathbb{R}^n)} \sup_{y \in \mathbb{R}^n} \frac{\langle x-y \rangle^N}{\langle x \rangle^N \langle y \rangle^N}. \end{aligned}$$

Since

$$\langle x-y \rangle^2 = 1 + |x-y|^2 \le 1 + 2(|x|^2+|y|^2) \le 2(1+|x|^2)(1+|y|^2) = 2\langle x \rangle^2 \langle y \rangle^2,$$

then we see that

$$\sup_{x,y \in \mathbb{R}^n} \frac{\langle x-y \rangle^N}{\langle x \rangle^N \langle y \rangle^N} = \left(\sup_{x,y \in \mathbb{R}^n} \frac{\langle x-y \rangle^2}{\langle x \rangle^2 \langle y \rangle^2} \right)^{\frac{N}{2}} \le 2^{\frac{N}{2}},$$

thus $|\langle x \rangle^{-N} (f*g)(x)| \le C_N \|\langle \cdot \rangle^{-N} f\|_{L^1(\mathbb{R}^n)}$, which conclude our result. □

Proof of Theorem 2.9.3 **(5)**. This follows similarly as in (4). □

Proof of Theorem 2.9.3 **(6)**. By (5), it is enough to show that $f*g \in C_\infty(\mathbb{R}^n)$ whenever $f,g \in C_\infty(\mathbb{R}^n)$. The binomial expansion (1.1.1) gives

$$(x-y+y)^\alpha = \sum_{i=1}^k c_i (x-y)^{\alpha_i} y^{\beta_i}$$

for some constants c_i and some multi-indices α_i and β_i, so we have

$$\begin{aligned} |x^\alpha (f*g)(x)| &\le \int_{\mathbb{R}^n} \left| (x-y+y)^\alpha f(y) g(x-y) \right| dy \\ &\le \sum_{i=1}^k |c_i| \int_{\mathbb{R}^n} \left| y^{\beta_i} f(y) (x-y)^{\alpha_i} g(x-y) \right| dy \\ &\le \sum_{i=1}^k |c_i| \left(\sup_{z \in \mathbb{R}^n} |z^{\alpha_i} g(z)| \right) \int_{\mathbb{R}^n} |z^{\beta_i} f(z)|\,dz. \end{aligned} \tag{2.9.9}$$

This implies that $\langle x \rangle^N (f*g)(x)$ is a bounded function for any $N \in \mathbb{N}$, so the claim follows. □

THEOREM 2.9.6. *The convolution is a* separately continuous *map*[10]

(1) $\mathscr{D}(\mathbb{R}^n) \times \mathscr{D}(\mathbb{R}^n) \to \mathscr{D}(\mathbb{R}^n)$,
(2) $\mathscr{E}(\mathbb{R}^n) \times \mathscr{D}(\mathbb{R}^n) \to \mathscr{E}(\mathbb{R}^n)$,
(3) $\mathscr{S}(\mathbb{R}^n) \times \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$.

PROOF. Theorem 2.9.3 immediately gives that the ranges in (1)–(3) are correct. It remains to show continuity.

If $f \in \mathscr{E}(\mathbb{R}^n)$ and $\varphi \in \mathscr{D}_K$, then for any compact subset $K_0 \subset \mathbb{R}^n$ we have

$$\begin{aligned}\sup_{x\in K_0} |\partial^\alpha (f*\varphi)(x)| = \sup_{x\in K_0} |(f*\partial^\alpha\varphi)(x)| &\le \sup_{x\in K_0} \int_K |f(x-y)(\partial^\alpha\varphi)(y)|\,dy \\ &\le |K| \|f\|_{L^\infty(K_1)} \|\partial^\alpha\varphi\|_{L^\infty(K)},\end{aligned}$$

where $K_1 := K_0 - K$ is compact. Taking $\varphi = \varphi_k$ with $\varphi_k \to 0$ in $\mathscr{D}_K$, we conclude the mapping

$$\varphi \in \mathscr{D}(\mathbb{R}^n) \mapsto f*\varphi \in \mathscr{E}(\mathbb{R}^n) \tag{2.9.10}$$

is continuous by using Theorem 2.8.6.

Similarly, for each fixed $\varphi \in \mathscr{D}(\mathbb{R}^n)$, we can obtain the estimate

$$\sup_{x\in K_0} |\partial^\alpha (f*\varphi)(x)| \le |K| \|\partial^\alpha f\|_{L^\infty(K_1)} \|\varphi\|_{L^\infty(K)}$$

Taking $f = f_k$ with $f_k \to 0$ in $\mathscr{E}(\mathbb{R}^n)$, we conclude the mapping

$$f \in \mathscr{E}(\mathbb{R}^n) \mapsto f*\varphi \in \mathscr{E}(\mathbb{R}^n) \tag{2.9.11}$$

is continuous. Then we conclude (1) and (2) by (2.9.10) and (2.9.11).

Using (2.9.9), we know that

$$\sup_{x\in\mathbb{R}^n} |x^\alpha (f*g)(x)| \le \rho(g) := \sum_{i=1}^k |c_i| \left(\sup_{z\in\mathbb{R}^n} |z^{\alpha_i} g(z)| \right) \int_{\mathbb{R}^n} |z^{\beta_i} f(z)|\,dz.$$

Note that ρ is a continuous semi-norm on $\mathscr{S}(\mathbb{R}^n)$. This implies that

$$[f*g]_{\alpha,\beta} = \sup_{x\in\mathbb{R}^n} |x^\alpha (f*\partial^\beta g)(x)| \le \rho(\partial^\beta g)$$

and we see that $g \mapsto \rho(\partial^\beta g)$ is also a continuous semi-norm on $\mathscr{S}(\mathbb{R}^n)$. Then we conclude (3) from Lemma 2.2.5. □

2.10. Tensor products

Before we define convolution between distributions, we first introduce some concept of tensor products. Given any functions f, g, the tensor product $f \otimes g$ is defined by

$$(f\otimes g)(x,y) := f(x)g(y).$$

If $T_f \in \mathscr{D}'(\Omega_1)$ (resp. $T_g \in \mathscr{D}'(\Omega_2)$) is the corresponding distribution of f (resp. g), then we define

$$(T_f \otimes T_g)(\varphi) := \int_{\Omega_1\times\Omega_2} f(x)g(y)\varphi(x,y)\,dx\,dy \quad \text{for all } \varphi \in \mathscr{D}(\Omega_1\times\Omega_2).$$

[10]This means that convolution satisfies the followings:

(1) for each fixed ψ, the mapping $\varphi \mapsto \varphi * \psi$ is continuous,
(2) for each φ, the mapping $\psi \mapsto \varphi * \psi$ is continuous.

If one take $\varphi(x,y) = (\varphi_1 \otimes \varphi_2)(x,y) \equiv \varphi_1(x)\varphi_2(y)$, here one obtains the identity

$$(T_f \otimes T_g)(\varphi_1 \otimes \varphi_2) = T_f(\varphi_1)T_g(\varphi_2).$$

The above ideas are in fact rigorous by the following facts (here we list them without proof):

Proposition 2.10.1 (Corollaries 4.1.1 and 4.1.2 in [**FJ98**]). *Let $U, V \subset \mathbb{R}^n$ be open sets. Let $T \in \mathscr{D}'(U)$ and let $\phi \in \mathscr{E}(U \times V)$ satisfy the following hypothesis: each point $y' \in V$ has a neighborhood $\Omega(y') \subset V$ such that* $\operatorname{supp}(\phi(\cdot, y))$ *is contained in a compact set $K = K(y')$ if $y \in \Omega(y')$. Then*

$$v(y) := T(\phi(\cdot, y)) \text{ is in } C^\infty(V)$$

and for each multi-index α we have

$$\partial_y^\alpha v(y) = T(\partial_y^\alpha \phi(\cdot, y)).$$

In particular, we have the following special cases:

(1) *If $T \in \mathscr{D}'(U)$ and $\phi \in \mathscr{D}(U \times V)$, then $v \in \mathscr{D}(V)$.*
(2) *If $T \in \mathscr{E}'(U)$ and $\phi \in \mathscr{E}(U \times V)$, then $v \in \mathscr{E}(V)$.*

The following theorem ensures the well-definedness of the tensor product of distributions:

Theorem 2.10.2 (Theorem 4.3.2 in [**FJ98**]). *Let $U, V \subset \mathbb{R}^n$ be open sets. Given any $T \in \mathscr{D}'(U)$ and $S \in \mathscr{D}'(V)$. There exists a unique element $T \otimes S \in \mathscr{D}'(U \times V)$, called the tensor product of T and S, written as $T \times S$, such that*

$$(T \otimes S)(\varphi_1 \otimes \varphi_2) = T(\varphi_1)S(\varphi_2) \quad \text{for all } \varphi_1 \in \mathscr{D}(U) \text{ and } \varphi_2 \in \mathscr{D}(V).$$

Proposition 2.10.3 (Theorem 4.3.3 in [**FJ98**]). *Let $U, V \subset \mathbb{R}^n$ be open sets. Given any $T \in \mathscr{D}'(U)$, $S \in \mathscr{D}'(V)$ and $\phi \in \mathscr{D}(U \times V)$.*

(1) *The tensor product $T \otimes S \in \mathscr{D}'(U \times V)$ given in Theorem 2.10.2 can be computed as*

$$\begin{aligned} (T \otimes S)(\varphi) &= S(v), \quad v(y) = T(\varphi(\cdot, y)) \text{ for each } y \in V, \\ (T \otimes S)(\varphi) &= T(u), \quad u(x) = S(\varphi(x, \cdot)) \text{ for each } x \in U, \end{aligned}$$

for all $\varphi \in \mathscr{D}(U \times V)$.
(2) $\operatorname{supp}(T \otimes S) = \operatorname{supp}(T) \times \operatorname{supp}(S)$.
(3) *Given any multi-indices α and β, we have*

$$\partial_x^\alpha \partial_y^\beta (T \otimes S) = \partial_x^\alpha T \otimes \partial_y^\beta S.$$

(4) *The tensor product is a* separately continuous *bilinear form on $\mathscr{D}'(U) \times \mathscr{D}'(V)$.*

2.11. Convolution of distributions

We now want to define convolution between distributions as a special case of tensor products. The usual requirement that the operation should extend the convolution of functions leads to the following: Given any functions f, g, and let $T_f, T_g \in \mathscr{D}'(\mathbb{R}^n)$ be the corresponding distributions, we see that

$$\begin{aligned}(T_f * T_g)(\varphi) &= \int_{\mathbb{R}^n} (f*g)(z)\varphi(z)\,dz = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} g(y)f(z-y)\varphi(z)\,dy\,dz \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(y)\varphi(x+y)\,dy\,dx \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(y)\phi(x,y)\,dy\,dx \quad \text{where } \phi(x,y)=\varphi(x+y) \\ &= (T_f\otimes T_g)(\phi).\end{aligned}$$

This suggests us to define the convolution of $T,S \in \mathscr{D}'(\mathbb{R}^n)$ by

$$(T*S)(\varphi) := (T\otimes S)(\phi) \quad \text{for all } \varphi\in\mathscr{D}(\mathbb{R}^n). \tag{2.11.1}$$

However, the convolution (2.11.1) may not exist, since in general $\phi \in \mathscr{E}(\mathbb{R}^n\times\mathbb{R}^n)$ does not have compact support even when $\varphi \in C_c^\infty(\mathbb{R}^n)$. A simply way to overcome this difficulty is to assume that one of the distributions T and S has compact support.

Let us assume $T\in\mathscr{E}'(\mathbb{R}^n)$ and $S\in\mathscr{D}'(\mathbb{R}^n)$. We choose $\rho\in C_c^\infty(\mathbb{R}^n)$ such that $\rho=1$ on a neighborhood of $\operatorname{supp}(T)$, see Theorem 2.7.11. For each $\varphi\in\mathscr{D}(\mathbb{R}^n)$, we define the function

$$\phi_\rho(x,y) := \rho(x)\varphi(x+y). \tag{2.11.2}$$

Note that

$$\begin{aligned}&(x,y)\in\operatorname{supp}(\phi_\rho) \\ \Longleftrightarrow\ & x\in\operatorname{supp}(\rho) \text{ and } x+y\in\operatorname{supp}(\varphi) \\ \Longleftrightarrow\ & x\in\operatorname{supp}(\rho) \text{ and } y\in\operatorname{supp}(\varphi)-x,\end{aligned}$$

which implies that

$$\operatorname{supp}(\phi_\rho) = \operatorname{supp}(\rho)\times\big(\operatorname{supp}(\varphi)-\operatorname{supp}(\rho)\big),$$

which shows that $\phi_\rho\in\mathscr{D}(\mathbb{R}^n\times\mathbb{R}^n)$. Plugging (2.11.2) in Proposition 2.10.3, we conclude the followings:

THEOREM 2.11.1. *The convolution is a* separately continuous *map*

(1) $\mathscr{E}'(\mathbb{R}^n)\times\mathscr{D}'(\mathbb{R}^n)\to\mathscr{D}'(\mathbb{R}^n)$,
(2) $\mathscr{E}'(\mathbb{R}^n)\times\mathscr{E}'(\mathbb{R}^n)\to\mathscr{E}'(\mathbb{R}^n)$,

Here, the convolution is given by

$$(T*S)(\varphi) := (T\otimes S)(\phi_\rho) \quad \textit{for all } \varphi\in\mathscr{D}(\mathbb{R}^n) \tag{2.11.3}$$

with (2.11.2) and the cut-off function $\rho\in C_c^\infty(\mathbb{R}^n)$ satisfies $\rho=1$ on a neighborhood of $\operatorname{supp}(T)$. *In particular, (2.11.3) is independent*[11] *of choices of the cut-off function ρ. Moreover, it can be computed as follows:*

$$\begin{aligned}(T*S)(\varphi) &= S(v), \quad v(y)=T(\phi_\rho(\cdot,y)) \textit{ for each } y\in\mathbb{R}^n, \\ (T*S)(\varphi) &= T(u), \quad u(x)=S(\phi_\rho(x,\cdot)) \textit{ for each } x\in\mathbb{R}^n,\end{aligned}$$

for all $\varphi\in\mathscr{D}(\mathbb{R}^n)$, and consequently $T*S=S*T$. In addition, we have

$$\operatorname{supp}(T*S)\subset\operatorname{supp}(T)+\operatorname{supp}(S) \quad \textit{for all } T\in\mathscr{E}'(\mathbb{R}^n) \textit{ and } S\in\mathscr{D}'(\mathbb{R}^n). \tag{2.11.4}$$

[11] If $\sigma\in C_c^\infty(\mathbb{R}^n)$ be another cut-off function satisfies $\sigma=1$ on a neighborhood of $\operatorname{supp}(T)$, for each x near $\operatorname{supp}(T)$ we see that $\big(\rho(x)-\sigma(x)\big)\phi(x+y)=0$, and we have $(T\otimes S)(\phi_\rho)=(T\otimes S)(\phi_\sigma)$ for all $\varphi\in\mathscr{D}(\mathbb{R}^n)$.

REMARK 2.11.2. Since $\operatorname{supp}(T)$ is compact and $\operatorname{supp}(S)$ is closed, then $\operatorname{supp}(T) + \operatorname{supp}(S)$ is closed, see Exercise 2.9.5.

PROOF. It is remain to show that 2.11.4. To show this, it is suffice to show that if $x \notin \operatorname{supp}(T) + \operatorname{supp}(S)$, then $x \notin \operatorname{supp}(T * S)$, equivalently, we want to show

$$T * S = 0 \quad \text{in } \mathbb{R}^n \setminus (\operatorname{supp}(T) + \operatorname{supp}(S)).$$

From Proposition 2.10.3(2), we know that

$$(x, y) \in \operatorname{supp}(T \otimes S) \implies x + y \in \operatorname{supp}(T) + \operatorname{supp}(S).$$

Hence we know that if the support of $\phi \in \mathscr{D}(\mathbb{R}^n)$ is disjoint from $\operatorname{supp}(T) + \operatorname{supp}(S)$, then the support of $(x, y) \mapsto \phi(x + y)$ is disjoint from

$$\operatorname{supp}(T \otimes S) = \operatorname{supp}(T) \times \operatorname{supp}(S),$$

which conclude our theorem. $\square$

Having defined the convolution on fairly general spaces, we now summarize some of the properties of the operation.

PROPOSITION 2.11.3. *Let $T \in \mathscr{E}'(\mathbb{R}^n)$ and $S \in \mathscr{D}'(\mathbb{R}^n)$.*

(1) ***Associativity.*** *If $R \in \mathscr{E}'(\mathbb{R}^n)$, then $R * (S * T) = (R * S) * T = (R * T) * S$.*
(2) ***Translation.*** *If $x \in \mathbb{R}^n$, then $\tau_x(T * S) = (\tau_x T) * S = T * (\tau_x S)$.*
(3) ***Differentiation.*** *If α is a multi-index, then $\partial^\alpha(T * S) = (\partial^\alpha T) * S = T * (\partial^\alpha S)$.*
(4) ***Translation.*** *If $x \in \mathbb{R}^n$, then $S * \delta_x = \delta_x * S = \tau_x S$. In particular when $x = 0$, the Dirac measure δ_0 is an identity element for the convolution operation.*

These are all easily derived from Theorem 2.11.1, the proofs are left to the reader.

EXERCISE 2.11.4. Let $T_1 = 1$, $T_2 = \delta_0'$ and $T_3 = H$ (the Heaviside unit step function) and show that

$$(T_1 * T_2) * T_3 \text{ and } (T_2 * T_3) * T_1 \text{ both exist but they are not identical.}$$

This exercise emphasizes that in general the associativity property in Proposition 2.11.3(1) only valid with the condition on supports.

The next theorem asserts that it is separately sequentially continuous with an appropriate definition of convergence in $\mathscr{E}'(\mathbb{R}^n)$.

PROPOSITION 2.11.5. *The following are true:*

(1) *Suppose that $T \in \mathscr{E}'(\mathbb{R}^n)$ and the sequence $\{S_j\}_{j\in\mathbb{N}}$ converges to S in $\mathscr{D}'(\mathbb{R}^n)$. Then*

$$T * S_j \to T * S \quad \text{in } \mathscr{D}'(\mathbb{R}^n). \tag{2.11.5}$$

(2) *Suppose that $T \in \mathscr{D}'(\mathbb{R}^n)$ and the sequence $\{S_j\}_{j\in\mathbb{N}}$ converges to S in $\mathscr{D}'(\mathbb{R}^n)$ such that $\operatorname{supp}(S_j) \subset K$ for some compact set K independent of j. Then (2.11.5) holds.*

PROOF. **(1).** If $T \in \mathscr{E}'(\mathbb{R}^n)$ and $\varphi \in C_c^\infty(\mathbb{R}^n)$, then $v(y) = T(\varphi(\cdot + y))$ is in $\mathscr{D}(\mathbb{R}^n)$ by Proposition 2.10.1. Hence it follows from Theorem 2.11.1 that

$$\lim_{j\to\infty} (T * S_j)(\varphi) = \lim_{j\to\infty} S_j(v) = S(v) = (T * S)(\varphi) \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R}^n).$$

(ii). Choosing a cut-off function $\rho \in C_c^\infty(\mathbb{R}^n)$ with $\rho = 1$ in a neighborhood of K. Then one has, again from Theorem 2.11.1 that

$$\lim_{j\to\infty}(T * S_j)(\varphi) = \lim_{j\to\infty} S_j(\rho v) = S(\rho v) = (T * S)(\varphi) \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R}^n),$$

which conclude our proposition. □

2.12. Convolution between distributions and functions

So far, we discussed the convolution between functions, and between distributions. We now consider the convolution between a function and a distribution. Fixing any $\varphi \in C_c^\infty(\mathbb{R}^n)$. Using Theorem 2.11.1 (with $T = T_\rho$), we have

$$(T_\rho * S)(\varphi) = S(v), \quad v(y) = T_\rho(\phi_\sigma(\cdot, y)) \text{ for each } y \in \mathbb{R}^n,$$

with $\phi_\sigma(x,y) = \sigma(x)\varphi(x+y)$, where $\sigma \in C_c^\infty(\mathbb{R}^n)$ is a cut-off function with $\sigma = 1$ on a neighborhood of $\operatorname{supp}(\rho)$. We compute that

$$\begin{aligned} v(y) = T_\rho(\phi_\rho(\cdot, y)) &= \int_{\mathbb{R}^n} \rho(x)\varphi(x+y)\,dx \\ &= \int_{\mathbb{R}^n} \rho(x-y)\varphi(x)\,dx. \end{aligned}$$

Using Proposition 2.10.1 we know that the function $w(x) := S(\rho(x - \cdot))$ is actually is in C^∞, therefore we see that

$$(T_\rho * S)(\varphi) = \int_{\mathbb{R}^n} w(x)\varphi(x)\,dx \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}^n).$$

Therefore, in this case, we can just simply identify the distribution $T_\rho * S$ with the function w. In this case, we just simply denote $w \equiv \rho * S$, i.e. we define

$$(\rho * S)(x) := S(\rho(x - \cdot)) \quad \text{for all } x \in \mathbb{R}^n, \tag{2.12.1}$$

and we reach the following theorem:

THEOREM 2.12.1. *If $S \in \mathscr{D}'(\mathbb{R}^n)$ and $\rho \in \mathscr{D}(\mathbb{R}^n)$, then $\rho * S \in C^\infty(\mathbb{R}^n)$.*

REMARK 2.12.2. In particular when $S \in \mathscr{E}'(\mathbb{R}^n)$, then $\rho * S \in C_c^\infty(\mathbb{R}^n)$. See also Theorem 2.14.6.

From (2.12.1), we also see that $(\rho * S)^\sim(x) = (\rho * S)(-x) = S(\rho(-x - \cdot)) = S(\tilde\rho(x + \cdot))$. Let $T \in \mathscr{E}'(\mathbb{R}^n)$. Using Proposition 2.11.1, we conclude the following corollary:

COROLLARY 2.12.3. *Let $T \in \mathscr{E}'(\mathbb{R}^n)$ and $S \in \mathscr{D}'(\mathbb{R}^n)$ (or $T \in \mathscr{D}'(\mathbb{R}^n)$ and $S \in \mathscr{E}'(\mathbb{R}^n)$), then $T\big((\rho * S)^\sim\big) = (T * S)(\tilde\rho)$ for all $\rho \in \mathscr{D}(\mathbb{R}^n)$.*

We can now prove one of the principal results in the theory of distributions.

THEOREM 2.12.4 (Density). *$C_c^\infty(\mathbb{R}^n)$ is dense in $\mathscr{D}'(\mathbb{R}^n)$.*

PROOF. Given any $T \in \mathscr{D}'(\mathbb{R}^n)$. Fixing any $\psi \in C_c^\infty(\mathbb{R}^n)$ with $\int \psi(x)\,dx = 1$, and set $\psi_j(x) = j^n\psi(jx)$ for $j = 1, 2, \cdots$. Then $\psi_j \to \delta_0$ in $\mathscr{D}'(\mathbb{R}^n)$ and $\operatorname{supp}(\psi_j) \subset \operatorname{supp}(\psi)$ for all j. Hence by Proposition 2.11.5 we have

$$\psi_j * T \to \delta_0 * T \equiv T \quad \text{in } \mathscr{D}'(\mathbb{R}^n).$$

By Theorem 2.12.1, we know that $\psi_j * T \in C^\infty(\mathbb{R}^n)$.

Now take $\chi \in C_c^\infty(\mathbb{R}^n)$ such that $\chi = 1$ in $B_1(0)$, and put

$$T_j := \chi(x/j)\psi_j * T \quad \text{for all } j \in \mathbb{N}.$$

If $\varphi \in C_c^\infty(\mathbb{R}^n)$, then

$$T_j(\varphi) := \int_{\mathbb{R}^n} (\psi_j * T)(x)\varphi(x)\, dx \quad \text{for all sufficiently large } j.$$

Hence we know that $T_j \in C_c^\infty(\mathbb{R}^n)$ and converges to T in $\mathscr{D}'(\mathbb{R}^n)$, which conclude the proof of the theorem. $\square$

The above theorem can be extended to distributions defined on an open set $\Omega \subset \mathbb{R}^n$:

THEOREM 2.12.5 (Density). *If $\Omega \subset \mathbb{R}^n$ is an open set, then $C_c^\infty(\Omega)$ is dense in $\mathscr{D}'(\Omega)$.*

PROOF. Let $T \in \mathscr{D}'(\Omega)$. Let $\{K_j\}_{j\in\mathbb{N}}$ be a compact exhaustion of Ω (described in Theorem 2.8.12). For each j, choosing $\rho_j \in C_c^\infty(\Omega)$ such that $\rho_j = 1$ in a neighborhood of K_j. We define

$$T_j := \rho_j T \quad \text{for all } j \in \mathbb{N}.$$

Then $T_j \in \mathscr{E}'(\Omega)$ extends trivially to an element of $\mathscr{E}'(\mathbb{R}^n)$. Now choose ψ such that

$$\psi \in C_c^\infty(\mathbb{R}^n), \quad \operatorname{supp}(\psi) \subset \overline{B_1(0)}, \quad \int_{\mathbb{R}^n} \psi(x)\, dx = 1.$$

Using (2.11.4) (in Theorem 2.11.1), one can find a decreasing sequence of positive real numbers $\{\epsilon_j\}$, which tending to zero, such that, if one sets

$$\psi_j(x) = \epsilon_j^{-n}\psi(x/\epsilon_j) \quad \text{for all } j \in \mathbb{N},$$

then $\psi_j * T_j$ are supported in Ω and hence are elements of $C_c^\infty(\Omega)$.

It is remains to prove $\psi_j * T_j \to T$ in $\mathscr{D}'(\Omega)$. For each $\varphi \in C_c^\infty(\Omega)$, there is a k such that $T(\varphi) = T_k(\varphi)$ and

$$(\psi_j * T_j)(\varphi) = T_j(u_j) = T_k(u_j) = (\psi_j * T_k)(\varphi) \quad \text{for all } j \geq k,$$

with $u_j(x) = \int_{\mathbb{R}^n} \psi_j(y)\varphi(x+y)\, dy$. By observing that $\psi_j \to \delta$ in $\mathscr{D}'(\mathbb{R}^n)$ and that $\operatorname{supp}(\psi_j) \subset \operatorname{supp}(\psi_1)$ for all $j \in \mathbb{N}$, using Proposition 2.11.5 we have

$$\lim_{j\to\infty} (\psi_j * T_j)(\varphi) = \lim_{j\to\infty} (\psi_j * T_k)(\varphi) = (\delta * T_k)(\varphi) = T_k(\varphi) = T(\varphi),$$

and our theorem follows by arbitrariness of $\varphi \in C_c^\infty(\Omega)$. $\square$

2.13. Convolution of distributions with non-compact supports

So far, it has been assumed that at least one of the distributions T and S has compact support, in order to ensure the existence of the convolution $T * S$. This can be replaced by a condition that is more symmetric, and extends to any finite set of distributions. Let $m \in \mathbb{Z}_{\geq 2}$ and let

$$\mu(x^{(1)}, \cdots, x^{(m)}) := x^{(1)} + \cdots + x^{(m)} \quad \text{for all } x^{(j)} \in \mathbb{R}^n. \tag{2.13.1}$$

DEFINITION 2.13.1. Let $A_1, \cdots, A_m$ be closed sets in $\mathbb{R}^n$. We shall say that the restriction of the mapping μ (as in (2.13.1)) to $A_1 \times \cdots \times A_m$ is *proper* if, for any $\delta > 0$, there is a $\delta' > 0$ such that

$$\|\mu\|_{L^\infty(A_1\times\cdots\times A_m)} \leq \delta \implies \sup_{j=1,\cdots,m} |x^{(j)}| \leq \delta'.$$

EXAMPLE 2.13.2. Suppose that $A_j \subset \{\ x \in \mathbb{R}^n \mid x_i \geq 0 \text{ for all } i = 1, \cdots, n\ \}$, then the restriction of the mapping μ (as in (2.13.1)) to $A_1 \times \cdots \times A_m$ is proper by choosing $\delta' = \delta$.

LEMMA 2.13.3. *Let $A_1, \cdots, A_m$ be closed subsets of $\mathbb{R}^n$. Let $\epsilon > 0$ and let $A_j^\epsilon = A_j + \overline{B_\epsilon(0)}$ be the closed neighborhood of A_j. Assume that the restriction of the mapping μ (as in (2.13.1)) to $A_1 \times \cdots \times A_m$ is proper, then its restriction to $A_1^\epsilon \times \cdots \times A_j^\epsilon$ is also proper.*

PROOF. Fixing any $\delta > 0$. Suppose that $x^{(j)} \in A_j^\epsilon$ and $|x^{(1)} + \cdots + x^{(m)}| \leq \delta$. There exist $x_0^{(j)} \in A_j$ such that $|x^{(j)} - x_0^{(j)}| \leq \epsilon$, hence

$$|x_0^{(1)} + \cdots + x_0^{(m)}| \leq \delta + m\epsilon.$$

Since the restriction of the mapping μ (as in (2.13.1)) to $A_1 \times \cdots \times A_m$ is proper, then there exists $\delta'' > 0$ such that

$$\sup_{j=1,\cdots,m} |x_0^{(j)}| \leq \delta'' \quad \text{which implies} \quad \sup_{j=1,\cdots,m} |x^{(j)}| \leq \delta' \equiv \delta'' + \epsilon,$$

which conclude the lemma. □

Let $T_1, \cdots, T_m \in \mathscr{D}'(\mathbb{R}^n)$, and suppose that the restriction of the mapping μ (as in (2.13.1)) to $\operatorname{supp}(T_1) \times \cdots \times \operatorname{supp}(T_m)$ is proper. Let $\epsilon > 0$, $\varphi \in \mathscr{D}(\mathbb{R}^n)$ and we set $\phi := \varphi \circ \mu$. By Lemma 2.13.3, the set

$$K_\epsilon(\varphi) := \left(\left(\operatorname{supp}(T_1)\right)^\epsilon \times \cdots \times \left(\operatorname{supp}(T_m)\right)^\epsilon \right) \cap \operatorname{supp}(\phi)$$

is compact[12]. One can choose cut-off functions $\rho_1, \cdots, \rho_m \in C_c^\infty(\mathbb{R}^n)$ such that

$$\rho(x^{(1)}, \cdots, x^{(m)}) := (\rho_1 \otimes \cdots \otimes \rho_m)(x^{(1)}, \cdots, x^{(m)}) \equiv \rho_1(x^{(1)}) \cdots \rho_m(x^{(m)})$$

is supported in $K_\epsilon(\varphi)$ and $\rho = 1$ in a neighborhood of $K_0(\varphi)$. Then we we conclude the following theorem:

THEOREM 2.13.4 (Convolution of distributions with non-compact supports). *Let $T_1, \cdots, T_m \in \mathscr{D}'(\mathbb{R}^n)$. If restriction of the mapping μ (as in (2.13.1)) to $\operatorname{supp}(T_1) \times \cdots \times \operatorname{supp}(T_m)$ is proper, then we can define the convolution $T_1 * \cdots * T_m \in \mathscr{D}'(\mathbb{R}^n)$ by*

$$(T_1 * \cdots * T_m)(\varphi) := (T_1 \otimes \cdots \otimes T_m)(\rho\phi) \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R}^n). \tag{2.13.2}$$

The definition (2.13.2) is independent of the choice of cut-off functions ρ.

REMARK 2.13.5. It reduces to the convolution as defined in Theorem 2.11.1 when $m = 2$ and $T_1 \in \mathscr{E}'(\mathbb{R}^n)$.

We also list some basic properties of this extended version of convolution as a proposition.

PROPOSITION 2.13.6. *Let $T_1, \cdots, T_m \in \mathscr{D}'(\mathbb{R}^n)$ and assume that the restriction of the mapping μ (as in (2.13.1)) to $\operatorname{supp}(T_1) \times \cdots \times \operatorname{supp}(T_m)$ is proper.*

(1) ***Associativity.*** *If I and J are disjoint subsets of $\{1, \cdots, m\}$ whose union is $\{1, \cdots, m\}$, then*

$$T_1 * \cdots * T_m = (*T_i)_{i \in I} * (*T_j)_{j \in J}.$$

(2) ***Support.*** $\operatorname{supp}(T_1 * \cdots * T_m) \subset \operatorname{supp}(T_1) + \cdots + \operatorname{supp}(T_m)$.

[12] Clearly $K_\epsilon(\varphi)$ is closed. We only need to show boundedness, which is easily follows from Lemma 2.13.3.

2.14. Structure of $\mathscr{D}'(\mathbb{R}^n)$, $\mathscr{E}'(\mathbb{R}^n)$ and $\mathscr{S}'(\mathbb{R}^n)$

We first prove structure theorem for $\mathscr{D}'(\mathbb{R}^n)$ using some fundamental solutions of some differential operators. Recall that the distributional derivative of the Heaviside function H (given in (1.1.6)) is simply the Dirac measure, i.e.

$$\partial H = \delta_0.$$

In other words, H is the fundamental solution of the differential operator ∂. This can easily be extended as follows: By observing that

$$x_+ = xH(x) \quad \text{for all } x \in \mathbb{R},$$

we see that $\partial x_+ = H$, and so

$$\partial^k \left(\frac{x_+^{k-1}}{(k-1)!} \right) = \delta_0 \quad \text{for each } k \in \mathbb{N}.$$

Using Proposition 2.10.3(iii), if one sets

$$E_k(x) := \frac{(x_1)_+^{k-1} \cdots (x_n)_+^{k-1}}{((k-1)!)^n} \quad \text{for all } x \in \mathbb{R}^n,$$

then one has

$$(\partial_1 \cdots \partial_n)^k E_k = \delta_0 \quad \text{in } \mathscr{D}'(\mathbb{R}^n), \tag{2.14.1}$$

that is, E_k is the fundamental solution of the differential operator $(\partial_1 \cdots \partial_n)^k$. This can be used to prove the following so-called *structure theorem*:

THEOREM 2.14.1 (Structure of $\mathscr{D}'(\mathbb{R}^n)$ restricted on a bounded open set). *Let Ω be a bounded open set in $\mathbb{R}^n$. For each $T \in \mathscr{D}'(\mathbb{R}^n)$, then*

$$T|_\Omega = \partial^\alpha f$$

for some multi-index α and for some function $f \in C^0(\mathbb{R}^n)$. Precisely,

$$T(\varphi) = (-1)^{|\alpha|} \int_{\mathbb{R}^n} f(x) \partial^\alpha \varphi(x)\, dx \quad \textit{for all } \varphi \in C_c^\infty(\Omega).$$

REMARK 2.14.2. From the proof, we see that we can choose $\operatorname{supp}(f)$ in any pre-assigned neighborhood of $\overline{\Omega}$. In particular, the structure theorem also can be formulated in terms of Laplacian, see Theorem 2.15.6.

PROOF. Since Ω is bounded, one can find a cut-off function $\psi \in C_c^\infty(\mathbb{R}^n)$ such that $\psi = 1$ in Ω. It is easy to see that

$$T = \psi T \quad \text{in } \Omega.$$

Since ψT has compact support, then it is of finite order N (as in Definition 2.8.9). From (2.14.1) (as well as Proposition 2.10.3(iii)), one has

$$\psi T = (\partial_1 \cdots \partial_n)^{N+2} E_{N+2} * (\psi T).$$

So the theorem will follow, with multi-index $\alpha = (N+2, \cdots, N+2)$, once it is shown that the distribution $E_{N+2} * (\psi T)$ can be identify with a continuous function.

Let $\rho \in C_c^\infty(\mathbb{R}^n)$ be such that $\rho \geq 0$, $\operatorname{supp}(\rho) \subset \overline{B_1(0)}$ and $\int_{\mathbb{R}^n} \rho(x)\, dx = 1$. We set $\rho_\epsilon(x) := \epsilon^{-n} \rho(x/\epsilon)$ where $\epsilon > 0$. We define

$$f_\epsilon := \big(E_{N+2} * (\psi T)\big) * \rho_\epsilon \in C^\infty(\mathbb{R}^n) \quad \text{(by Theorem 2.12.1)}.$$

Since ψT and ρ_ϵ have compact support, the associative law applies (Proposition 2.11.3), and gives

$$f_\epsilon = \psi T * (E_{N+2} * \rho_\epsilon).$$

Using Theorem 2.12.1, we see that $E_{N+2} * \rho_\epsilon \in C^\infty(\mathbb{R}^n)$, and therefore we compute that

$$f_\epsilon(x) = \psi T((E_{N+2} * \rho)(x - \cdot)).$$

In fact, we can show that

$$E_{N+2} * \rho_\epsilon \to E_{N+2} \quad \text{in } C^N(\mathbb{R}^n) \tag{2.14.2}$$

as $\epsilon \to 0$ (left as exercise). Since ψT is of order N using Definition 2.8.9 we see that for each compact set K

$$\lim_{\epsilon \to 0} \|f_\epsilon - E_{N+2} * (\psi T)\|_{L^\infty(K)} = 0,$$

thus the mapping

$$f(x) := E_{N+2} * (\psi T)(x) = \psi T(E_{N+2}(x - \cdot)) \quad \text{is in } C^0(\mathbb{R}^n).$$

On the other hand, using Proposition 2.11.5, f_ϵ also converges to $E_{N+2} * (\psi T)$ in $\mathscr{D}'(\mathbb{R}^n)$ as $\epsilon \to 0$. It is clear that the limits are the same in this case, therefore we conclude our proof. $\square$

EXERCISE 2.14.3. Prove (2.14.2). [Hint: Note that ρ_ϵ is simply the standard mollifier. See the proof of properties of mollifier.]

COROLLARY 2.14.4 (Structure of $\mathscr{E}'(\mathbb{R}^n)$). *Let $T \in \mathscr{E}'(\mathbb{R}^n)$. Then there is an integer $m \geq 0$ and a set of continuous functions $\{f_\alpha\}_{|\alpha| \leq m}$ such that*

$$T = \sum_{|\alpha| \leq m} \partial^\alpha f_\alpha.$$

Precisely,

$$T(\varphi) = \sum_{|\alpha| \leq m} (-1)^{|\alpha|} \int_{\mathbb{R}^n} f_\alpha(x) \partial^\alpha \varphi(x)\, dx \quad \text{for all } \varphi \in \mathscr{E}(\mathbb{R}^n).$$

PROOF. Let $\sigma \in C_c^\infty(\mathbb{R}^n)$ be such that $\sigma = 1$ in a neighborhood of $\operatorname{supp}(T)$. Let Ω be an open set such that $\operatorname{supp}(T) \subset \Omega$. Using Theorem 2.14.1, we see that

$$T(\psi) = (-1)^{|\alpha|} \int_{\mathbb{R}^n} f(x) \partial^\alpha \psi(x)\, dx \quad \text{for all } \psi \in \mathscr{D}(\Omega).$$

Therefore, we see that

$$T(\varphi) = T(\sigma\varphi) = (-1)^{|\alpha|} \int_{\mathbb{R}^n} f(x) \partial^\alpha (\sigma\varphi)(x)\, dx \quad \text{for all } \varphi \in \mathscr{E}(\mathbb{R}^n),$$

and thus we obtain our corollary. $\square$

We now prove structure theorem for $\mathscr{S}'(\mathbb{R}^n)$ using Fourier transform, which stated that every tempered distribution is a distributional derivative of finite order of some continuous function of polynomial growth, can be found in [**FJ98**, Theorem 8.3.1]:

THEOREM 2.14.5 (Structure of $\mathscr{S}'(\mathbb{R}^n)$). *Any $T \in \mathscr{S}'(\mathbb{R}^n)$ can be written as $T = \partial^\alpha f$ for some multi-index α and some polynomially bounded continuous function f.*

PROOF. It is suffice to prove this for $T \in \mathscr{S}'(\mathbb{R}^n)$ supported in $\Omega = \{ x \in \mathbb{R}^n \mid x_1 > 0, \cdots, x_n > 0 \}$, i.e.

$$T(\varphi) = 0 \quad \text{for all } \varphi \in \mathscr{S}(\mathbb{R}^n) \text{ with } \mathrm{supp}\,(\varphi) \subset \mathbb{R}^n \setminus \Omega.$$

For this implies the result when $T \in \mathscr{S}'(\mathbb{R}^n)$ is supported in $\{ x \in \mathbb{R}^n \mid \sigma_1 x_1 > -\delta, \cdots, \sigma_n x_n > -\delta \}$ where $\delta > 0$ and each $\sigma_j = \pm 1$ (using Lemma 2.13.3), and, via a partition of unity, any tempered distribution can be written as the sum of 2^n distributions of this type.

Let $T \in \mathscr{S}'(\mathbb{R}^n)$ with $\mathrm{supp}\,(T) \subset \Omega$. Using Lemma 2.4.2 (as well as Remark 2.2.2), we have

$$|T(\varphi)| \leq C \sum_{|\alpha|,|\beta| \leq N} \sup_{x \in \mathbb{R}^n} x^\alpha |D^\beta \varphi(x)| \quad \text{for all } \varphi \in \mathscr{S}(\mathbb{R}^n). \tag{2.14.3}$$

Choosing $\varphi = \rho\phi$ with $\phi \in C_c^\infty(\mathbb{R}^n)$ and $\rho \in C^\infty(\mathbb{R}^n)$ with $\rho = 1$ on a neighborhood of $\mathrm{supp}\,(T)$, we have (with different constant C)

$$|T(\phi)| \leq C \sum_{|\alpha|,|\beta| \leq N} \sup_{x \in \Omega} \big(x^\alpha |D^\beta \phi(x)| \big) \quad \text{for all } \phi \in C_c^\infty(\mathbb{R}^n),$$

since $|x^\alpha| = x^\alpha$ when $x \in \Omega$.

We define $E_{N+2} \in \mathscr{D}'(\mathbb{R}^n)$ by

$$E_{N+2}(x) := \frac{(x_1)_+^{N+1} \cdots (x_n)_+^{N+1}}{((N+1)!)^n}.$$

Since restriction of the mapping μ (as in (2.13.1)) to $\mathrm{supp}\,(E_{N+2}) \times \mathrm{supp}\,(T)$ is proper (see Example 2.13.2), using Theorem 2.13.4 the convolution $E_{N+2} * T$ is well-defined. From Theorem 2.11.1 and (2.14.1), we see that

$$\mathrm{supp}\,(E_{N+2} * T) \subset \Omega, \quad T = (\partial_1 \cdots \partial_n)^{N+2} (E_{N+2} * T).$$

As in the proof of Theorem 2.14.1, we can show that $f := E_{N+2} * T \in C^0(\mathbb{R}^n)$ and that from (2.14.3) we have

$$|f(x)| \leq C' \sum_{|\alpha|,|\beta| \leq N} \sup_{t \in \Omega} t^\alpha \partial^\beta \big((x_1 - t_1)_+^{N+1} \cdots (x_n - t_n)_+^{N+1} \big)$$

with $C' = C/((N+1)!)^n$. An elementary computation, which is left to the reader, shows that $|f(x)| \leq C'' \langle x \rangle^M$ with $M = 2N+1$ for some constant $C'' > 0$, so the theorem is proved. □

With the structure theorem at hand, we can prove the following result (see e.g. Theorem 3.9.1 of Mikko Salo's lecture note).

THEOREM 2.14.6. *If $T \in \mathscr{S}'(\mathbb{R}^n)$ and $f \in \mathscr{S}(\mathbb{R}^n)$, then $f * T \in \mathscr{O}_{\mathrm{M}}(\mathbb{R}^n)$.*

PROOF. By Theorem 2.14.5, there is a multi-index α and a polynomially bounded continuous function h such that $T = \partial^\alpha h$, i.e.

$$T(\varphi) = (-1)^{|\alpha|} \int_{\mathbb{R}^n} h(y) \partial_y^\alpha \varphi(y)\, dy.$$

From (2.12.1), we have

$$(f * T)(x) = T(f(x - \cdot)) = (-1)^{|\alpha|} \int_{\mathbb{R}^n} h(y) \partial_y^\alpha f(x-y)\, dy = (-1)^{|\alpha|} \big(h * \partial^\alpha f \big)(x),$$

which implies our desired result. $\square$

2.15. Fourier transform on $\mathscr{E}'(\mathbb{R}^n)$ and the convolution theorem

With the structure theorem at hand, we are now able to show that the Fourier transform of compactly supported distribution (see Definition 2.5.1 and Theorem 2.7.11) is actually a smooth function in $\mathbb{R}^n$. This illustrates the fact that the Fourier transform exchanges decay properties with smoothness.

THEOREM 2.15.1. *Let $T \in \mathscr{E}'(\mathbb{R}^n)$, then its Fourier transform $\hat{T} = T_F$ with $F(\xi) = T(\phi(\cdot,\xi))$ with $\phi(x,\xi) = e^{-ix\cdot\xi}$ and $F \in \mathscr{O}_{\mathrm{M}}(\mathbb{R}^n)$.*

REMARK 2.15.2. Accordingly, we can identify $\hat{T}$ with the function $\hat{T}(\xi) := T(\phi(\cdot,\xi))$ for all $\xi \in \mathbb{R}^n$.

PROOF OF THEOREM 2.15.1. Let $T \in \mathscr{E}'(\mathbb{R}^n)$. Using the structure theorem in Corollary 2.14.4, we can write $T = \sum_{|\alpha|\le N} \partial^\alpha f_\alpha$ with $f_\alpha \in C_c(\mathbb{R}^n)$. Then by properties of the Fourier transform on Schwartz functions (Proposition 2.3.3), for each $\varphi \in \mathscr{S}(\mathbb{R}^n)$, we have

$$\begin{aligned}\hat{T}(\varphi) = T(\hat{\varphi}) &= \sum_{|\alpha|\le N} (-1)^\alpha \int_{\mathbb{R}^n} f_\alpha(x)\partial^\alpha\hat{\varphi}(x)\,dx \\ &= \sum_{|\alpha|\le N} \int_{\mathbb{R}^n} f_\alpha(x)((i\xi)^\alpha\varphi)^\wedge(x)\,dx \\ &= \int_{\mathbb{R}^n} \bigg(\sum_{|\alpha|\le N} (i\xi)^\alpha \hat{f}_\alpha(\xi)\bigg)\varphi(\xi)\,d\xi.\end{aligned}$$

On the other hand, we compute that

$$\begin{aligned}\sum_{|\alpha|\le N} (i\xi)^\alpha \hat{f}_\alpha(\xi) &= \sum_{|\alpha|\le N} (i\xi)^\alpha \int_{\mathbb{R}^n} f_\alpha(\xi)e^{-ix\cdot\xi}\,d\xi \\ &= \sum_{|\alpha|\le N} (-1)^{|\alpha|} \int_{\mathbb{R}^n} f_\alpha(\xi)\partial^\alpha e^{-ix\cdot\xi}\,d\xi \\ &= T(\phi(\cdot,\xi)) \quad \text{with } \phi(x,\xi) = e^{-ix\cdot\xi}, \end{aligned} \tag{2.15.1}$$

which gives

$$\hat{T}(\varphi) = \int_{\mathbb{R}^n} T(\phi(\cdot,\xi))\varphi(\xi)\,d\xi.$$

Since $\hat{f} \in L^\infty(\mathbb{R}^n)$, from (2.15.1) we immediately see that the mapping $\xi \mapsto T(\phi(\cdot,\xi))$ is in $\mathscr{O}_{\mathrm{M}}(\mathbb{R}^n)$. $\square$

COROLLARY 2.15.3. *If $T, S \in \mathscr{E}'(\mathbb{R}^n)$, then $(T * S)^\wedge(\xi) = \hat{T}(\xi)\hat{S}(\xi)$.*

PROOF. Using Theorem 2.11.1, we know that $T * S \in \mathscr{E}'(\mathbb{R}^n)$. Using Theorem 2.15.1, one can computes follows:

$$\begin{aligned}(T * S)^\wedge(\xi) &= (T * S)(\phi(\cdot,\xi)) \quad \text{with } \phi(x,\xi) = e^{-ix\cdot\xi} \\ &= (T \otimes S)(\psi(\cdot,\xi)) \quad \text{with } \psi(x,y,\xi) = e^{-i(x+y)\cdot\xi} = \phi(x,\xi)\phi(y,\xi) \\ &= T(\phi(\cdot,\xi))S(\phi(\cdot,\xi)) = \hat{T}(\xi)\hat{S}(\xi),\end{aligned}$$

which conclude our corollary. □

We now establish a version of the convolution theorem for distributions.

THEOREM 2.15.4. *Let $T \in \mathscr{S}'(\mathbb{R}^n)$ and $S \in \mathscr{E}'(\mathbb{R}^n)$. Then $T*S \in \mathscr{S}'(\mathbb{R}^n)$ and $(T*S)^\wedge = \hat{T}\hat{S}$.*

REMARK 2.15.5. Recall that $\mathscr{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is an isomorphism. Using Schwartz's inequality, it is easy to show that $u * v \in L^\infty(\mathbb{R}^n) \subset \mathscr{S}'(\mathbb{R}^n)$ is well-defined. Using Parseval's identity, we also can easily see that $\hat{u}\hat{v} \in L^1(\mathbb{R}^n) \subset \mathscr{S}'(\mathbb{R}^n)$. In particular, the convolution theorem also holds for $L^2(\mathbb{R}^n)$ functions: $(u * v)^\wedge = \hat{u}\hat{v}$ for all $u, v \in L^2(\mathbb{R}^n)$, see e.g. [**FJ98**, Theorem 9.2.3].

PROOF OF THEOREM 2.15.4. By Proposition 2.4.10 and Theorem 2.15.1, one has $\hat{T}\hat{S} \in \mathscr{S}'(\mathbb{R}^n)$. Since $\mathscr{F} : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ is bijective (Theorem 2.5.4), there exists a unique $R \in \mathscr{S}'(\mathbb{R}^n)$ such that $\hat{R} = \hat{T}\hat{S}$. To compute R, we use the Fourier inversion formula in Theorem 2.5.4 and Corollary 2.15.3: For each $\varphi \in \mathscr{D}(\mathbb{R}^n)$, we have

$$\begin{aligned} R(\tilde{\varphi}) &\overset{(2.5.1)}{=} (2\pi)^{-n}\hat{R}(\hat{\varphi}) = (2\pi)^{-n}(\hat{T}\hat{S})(\hat{\varphi}) = (2\pi)^{-n}\int_{\mathbb{R}^n} \hat{T}(\xi)\hat{S}(\xi)\hat{\varphi}(\xi)\,d\xi \\ &= (2\pi)^{-n}\int_{\mathbb{R}^n} \hat{T}(\xi)(S*\varphi)^\wedge(\xi)\,d\xi = (2\pi)^{-n}\hat{T}((S*\varphi)^\wedge) \overset{(2.5.1)}{=} T\big((S*\varphi)^\sim\big) = (T*S)(\tilde{\varphi}), \end{aligned}$$

where the last equality follows from Corollary 2.12.3. Hence we conclude our result from density result in Proposition 2.8.16. □

We now also able to obtain a structure theorem for $\mathscr{D}'(\mathbb{R}^n)$ in terms of Laplacian (see also Theorem 2.14.1).

THEOREM 2.15.6 (Structure of $\mathscr{D}'(\mathbb{R}^n)$). *If $T \in \mathscr{D}'(\mathbb{R}^n)$ and Ω is a bounded open set in $\mathbb{R}^n$. Then there is a $f \in C^0(\mathbb{R}^n)$ and an integer $N \geq 0$ such that*

$$T = (1-\Delta)^N f \quad \textit{in } \Omega.$$

PROOF. One can choose a cut-off function $\rho \in C_c^\infty(\mathbb{R}^n)$ such that $\rho = 1$ in Ω. As $\rho T \in \mathscr{E}'(\mathbb{R}^n)$, Theorem 2.15.1 implies that $(\rho T)^\wedge \in \mathscr{O}_{\mathrm{M}}(\mathbb{R}^n)$. Hence one can find an integer N such that

$$g(\xi) := (1+|\xi|^2)^{-N}(\rho T)^\wedge(\xi) \quad \text{is in } L^1(\mathbb{R}^n).$$

Since $L^1(\mathbb{R}^n) \subset \mathscr{S}'(\mathbb{R}^n)$ and $\mathscr{F} : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ is bijective continuous, then there exists $f \in C^0(\mathbb{R}^n)$ such that such that $\hat{f} = g$, which gives

$$(\rho T)^\wedge(\xi) = (1+|\xi|^2)^N \hat{f}(\xi) \quad \text{for all } \xi \in \mathbb{R}^n.$$

Hence the Fourier inversion formula implies that

$$\rho T = (1-\Delta)^N f \quad \text{in } \mathbb{R}^n,$$

which conclude our theorem. □

2.16. Convolution of pv tempered distributions and Fourier transform

As a preamble, observe that $f(x) := \frac{1}{x}$ for $x \neq 0$ is not $L^1_{\rm loc}(\mathbb{R})$, therefore we cannot associate it by a distribution using the way described in Example 2.8.10. Nonetheless, it is possible to associate to this function a certain using some tricks.

EXAMPLE 2.16.1. We consider the mapping $T : \mathscr{D}(\mathbb{R}) \to \mathbb{C}$ by

$$T(\varphi) := \lim_{\epsilon \to 0_+} \int_{|x| \geq \epsilon} \frac{\varphi(x)}{x}\,dx \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R}). \tag{2.16.1}$$

We first prove the mapping (2.16.1) is well-defined. Given any $\varphi \in \mathscr{D}(\mathbb{R})$ and suppose $R > 0$ is such that $\operatorname{supp}(\varphi) \subset (-R, R)$. Fixing any $0 < \epsilon < R$, and since $\frac{1}{x}$ is odd, we have

$$\int_{|x| \geq \epsilon} \frac{\varphi(x)}{x}\,dx = \int_{\epsilon \leq |x| \leq R} \frac{\varphi(x)}{x}\,dx = \int_{\epsilon \leq |x| \leq R} \frac{\varphi(x) - \varphi(0)}{x}\,dx.$$

Since $|\frac{\varphi(x)-\varphi(0)}{x}| \leq \sup_{|y| \leq R} |\varphi'(y)|$ for each $x \in \mathbb{R}\setminus\{0\}$, using Lebesgue dominated convergence theorem, we see that

$$T(\varphi) = \lim_{\epsilon \to 0_+} \int_{\epsilon \leq |x| \leq R} \frac{\varphi(x) - \varphi(0)}{x}\,dx = \int_{|x| \leq R} \frac{\varphi(x) - \varphi(0)}{x}\,dx, \tag{2.16.2}$$

this shows that (2.16.1) is well-defined, and it can expressed in terms of (2.16.2). From the estimate

$$|T(\varphi)| \leq 2R \sup_{|x| \leq R} |\varphi'(x)| \quad \text{for all } \varphi \in \mathscr{D}((-R,R)),$$

we know that T is a distribution in $\mathbb{R}$ of order *at most* one (see Definition 2.8.9).

We are left with showing that T does not have order 0. Consider the compact set $K = [0,1]$ and for each $j \in \mathbb{N}$, let $\varphi_j \in \mathscr{D}((0,1))$ be such that $0 \leq \varphi_j \leq 1$ and $\varphi_j = 1$ on $[\frac{1}{j+2}, 1 - \frac{1}{j+2}]$. Then we see that

$$T(\varphi_j) = \int_0^1 \frac{\varphi_j(x)}{x}\,dx \geq \int_{\frac{1}{j+2}}^{1-\frac{1}{j+2}} \frac{1}{x}\,dx = \ln(j+1) \quad \text{for each } j \in \mathbb{N}.$$

Since $\sup_{x \in K} |\varphi_j(x)| \leq 1$ and $\lim_{j \to \infty} \ln(j+1) = \infty$, this shows that there is no constant $C > 0$ with the property that

$$|T(\varphi)| \leq C \sup_{x \in K} |\varphi(x)| \quad \text{for all } \varphi \in \mathscr{D}_K,$$

this proves that T does not have order 0. Therefore we conclude that T of order 1. We usually denote such distribution T by $\operatorname{pv}\frac{1}{x}$, called the *principal value* $\frac{1}{x}$. Many authors (including myself) just simply ignore the notation "pv", so we need to understand it by ourselves when reading these literature.

EXERCISE 2.16.2. Show that $\operatorname{pv}\frac{1}{x} \in \mathscr{S}'(\mathbb{R})$.

EXERCISE 2.16.3. Prove that $(\ln|x|)' = \operatorname{pv}\frac{1}{x}$ in $\mathscr{S}'(\mathbb{R})$.

To generalize the above idea to general case, we need to introduce some definition.

DEFINITION 2.16.4. A nonempty open set $\mathcal{O}$ in $\mathbb{R}^n$ is called a *cone-like region* if $tx \in \mathcal{O}$ whenever $x \in \mathcal{O}$ and $t > 0$. Given a cone-like region $\mathcal{O} \subset \mathbb{R}^n$, call a function $f : \mathcal{O} \to \mathbb{C}$ *positive homogeneous of degree* $k \in \mathbb{R}$ if $f(tx) = t^k f(x)$ for every $t > 0$ and every $x \in \mathcal{O}$.

From Example 2.16.1, we see that the key features of the function $\Theta(x) := \frac{1}{x}$ $(x \in \mathbb{R}\setminus\{0\})$ that allowed us to define $\operatorname{pv}\frac{1}{x}$ as a tempered distribution on the real line as follows:

$$\Theta \in C^0(\mathbb{R}\setminus\{0\}), \quad \text{positive homogeneous of degree } -1, \quad \Theta(1)+\Theta(-1)=0.$$

EXERCISE 2.16.5. Prove that if $f \in C^0(\mathbb{R}^n\setminus\{0\})$ is positive homogeneous of degree $k \in \mathbb{R}$, then

$$|f(x)| \le \|f\|_{L^\infty(\mathcal{S}^{n-1})}|x|^k \quad \text{for all } x \in \mathbb{R}^n\setminus\{0\}.$$

We now want to generalize the above to $\mathbb{R}^n$:

PROPOSITION 2.16.6. *Let $n \ge 2$ be an integer, and let Θ be a function satisfying*

$$\Theta \in C^0(\mathbb{R}^n\setminus\{0\}), \quad \textit{positive homogeneous of degree } -n, \quad \int_{\mathcal{S}^{n-1}} \Theta\,ds = 0. \tag{2.16.3}$$

Then the linear map $\operatorname{pv}\Theta : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ given by

$$(\operatorname{pv}\Theta)(\varphi) := \lim_{\epsilon\to 0_+}\int_{|x|\ge\epsilon}\Theta(x)\varphi(x)\,dx \quad \textit{for all } \varphi \in \mathscr{S}(\mathbb{R}^n) \tag{2.16.4}$$

is well-defined and is in $\mathscr{S}'(\mathbb{R}^n)$. In addition,

$$\operatorname{pv}\Theta\Big|_{\mathbb{R}^n\setminus\{0\}} = \Theta\Big|_{\mathbb{R}^n\setminus\{0\}} \quad \textit{in } \mathscr{D}'(\mathbb{R}^n\setminus\{0\}). \tag{2.16.5}$$

PROOF. Writing $\hat{x} = x/|x|$ for all $x \ne 0$. Fixing any radially symmetric function $\psi \in C^1(\mathbb{R}^n)$ with $\psi(0)=1$ and

$$|\psi(x)| \le C|x|^{-\epsilon_0} \quad \text{for all } x \text{ with } |x| \ge R$$

for some $\epsilon_0 > 0$, $R > 0$ and $C > 0$. For each $\epsilon > 0$, we compute

$$\begin{aligned}
\int_{|x|\ge\epsilon}\Theta(x)\psi(x)\,dx &= \int_{|x|\ge\epsilon}\frac{\Theta(\hat{x})}{|x|^n}\psi(x)\,dx \quad (\text{positive homogeneous of degree } -n) \\
&= \int_\epsilon^\infty \frac{\psi(r)}{r}\Bigg(\overbrace{\int_{\mathcal{S}^{n-1}}\Theta(\hat{x})\,ds(\hat{x})}^{=0}\Bigg)\,dr = 0.
\end{aligned}$$

Then we have

$$\int_{|x|\ge\epsilon}\Theta(x)\varphi(x)\,dx = \int_{|x|\ge\epsilon}\Theta(x)\big(\varphi(x)-\varphi(0)\psi(x)\big)\,dx = \int_{|x|\ge\epsilon}\Theta(\hat{x})\frac{\varphi(x)-\varphi(0)\psi(x)}{|x|^n}\,dx$$

Since $\Theta \in C^0(\mathbb{R}^n\setminus\{0\})$ and using Exercise 2.16.5, then we see that for each $\varphi \in \mathscr{S}(\mathbb{R}^n)$ we know that

$$\text{the mapping } x \mapsto \Theta(x)\big(\varphi(x)-\varphi(0)\psi(x)\big) = \Theta(\hat{x})\frac{\varphi(x)-\varphi(0)\psi(x)}{|x|^n} \text{ is in } L^1(\mathbb{R}^n).$$

Therefore using Lebesgue dominated convergence theorem, we have

$$(\operatorname{pv}\Theta)(\varphi) = \int_{\mathbb{R}^n}\Theta(x)\big(\varphi(x)-\varphi(0)\psi(x)\big)\,dx \quad \text{for all } \varphi \in \mathscr{S}(\mathbb{R}^n). \tag{2.16.6}$$

Using mean value theorem on the mapping $t \mapsto \varphi(tx)\psi((1-t)x)$, it is easy to see that

$$|(\operatorname{pv}\Theta)(\varphi)| \le C\sup_{|\alpha|\le 1,|\beta|\le 1}[\varphi]_{\alpha,\beta},$$

where $[\cdot]_{\alpha,\beta}$ is given by (2.2.2), which shows that $\mathrm{pv}\,\Theta \in \mathscr{S}'(\mathbb{R}^n)$. The fact (2.16.5) is immediate from definitions. $\square$

REMARK 2.16.7 (Representation). By choosing $\psi = 1$ in $\overline{B_1(0)}$ in (2.16.6), we can compute that

$$(\mathrm{pv}\,\Theta)(\varphi) = \int_{|x|\leq 1} \Theta(x)\big(\varphi(x) - \varphi(0)\big)\,dx + \int_{|x|>1} \Theta(x)\varphi(x)\,dx \quad \text{for all } \varphi \in \mathscr{S}(\mathbb{R}^n). \tag{2.16.7}$$

EXAMPLE 2.16.8. If $j \in \{1, \cdots, n\}$, the function Θ defined by $\Theta(x) := \frac{x_j}{|x|^{n+1}}$ for each $x \in \mathbb{R}^n \setminus \{0\}$ satisfies (2.16.3). By Proposition 2.16.6, we have

$$\mathrm{pv}\,\frac{x_j}{|x|^{n+1}} \text{ belongs to } \mathscr{S}'(\mathbb{R}^n).$$

From (2.16.7), we have the representation

$$\begin{aligned}\left(\mathrm{pv}\,\frac{x_j}{|x|^{n+1}}\right)(\varphi) &= \lim_{\epsilon\to 0_+} \int_{|x|\geq\epsilon} \frac{x_j\varphi(x)}{|x|^{n+1}}\,dx \\ &= \int_{|x|\leq 1} \frac{x_j(\varphi(x) - \varphi(0))}{|x|^{n+1}}\,dx + \int_{|x|>1} \frac{x_j\varphi(x)}{|x|^{n+1}}\,dx.\end{aligned}$$

From Proposition 2.16.6, we know that whenever Θ is a function satisfying the conditions in (2.16.3), the principal value distribution $\mathrm{pv}\,\Theta \in \mathscr{S}'(\mathbb{R}^n)$. As such, its Fourier transform makes sense in $\mathscr{S}'(\mathbb{R}^n)$, see Theorem 2.5.4. Here we do not walk through all the computations (they are quite technical). The results following are adopted from [**Mit18**, Section 4.5].

Before discussing the Fourier transform of principal value distributions $\mathrm{pv}\,\Theta$, we first introduce an auxiliary function as in [**Mit18**, Therem 4.74]:

PROPOSITION 2.16.9 (Theorem 4.74 in [**Mit18**]). *Let Θ be a function satisfying the conditions in (2.16.3). Then the function given by the formula*

$$m_\Theta(\xi) := -\int_{\mathcal{S}^{n-1}} \Theta(\omega)\log(i\xi\cdot\omega)\,ds(\omega) \quad \text{for all } \xi \in \mathbb{R}^n \setminus \{0\},$$

where we choose the branch

$$\log(i\xi\cdot\omega) = \ln|\xi\cdot\omega| + i\frac{\pi}{2}\mathrm{sgn}\,(\xi\cdot\omega),$$

is well-defined, positive homogeneous of degree zero, satisfies

$$\int_{\mathcal{S}^{n-1}} m_\Theta(\xi)\,ds(\xi) = 0, \quad \|m_\Theta\|_{L^\infty(\mathbb{R}^n)} \leq C_n\|\Theta\|_{L^\infty(\mathcal{S}^{n-1})}$$

with positive constant $C_n := \frac{\pi|\mathcal{S}^{n-1}|}{2} + \int_{\mathcal{S}^{n-1}} |\ln|\frac{\xi}{|\xi|}\cdot\omega||\,ds(\omega)$ and $\mathscr{F}^{-1}(m_\Theta) = m_{\mathscr{F}^{-1}\mathrm{pv}\,\Theta}$. In addition, if $\Theta \in C^k(\mathbb{R}^n \setminus \{0\})$ for some $k \in \mathbb{Z}_{\geq 0} \cup \{\infty\}$, then

$$m_\Theta\Big|_{\mathbb{R}^n\setminus\{0\}} \in C^k(\mathbb{R}^n \setminus \{0\}).$$

If Θ is an even function in the sense of $\Theta(\omega) = \Theta(-\omega)$ for all $\omega \in \mathcal{S}^{n-1}$, then $\overline{m_\Theta} = m_{\overline{\Theta}}$.

In fact, the function introduce above is the Fourier transform of the tempered distribution $\mathrm{pv}\,\Theta$:

THEOREM 2.16.10 (Theorem 4.74 in [**Mit18**]). *Let Θ be a function satisfying the conditions in (2.16.3). Then the Fourier transform of the tempered distribution* $\mathrm{pv}\,\Theta$ *is given by*

$$\mathscr{F}(\mathrm{pv}\,\Theta) = m_\Theta \quad \text{in } \mathscr{S}'(\mathbb{R}^n).$$

The next proposition elaborates on the manner in which principal value tempered distributions convoluted with Schwartz functions.

PROPOSITION 2.16.11. *Let Θ be a function satisfies (2.16.3). Then for each $\varphi \in \mathscr{S}(\mathbb{R}^n)$ one has that $(\mathrm{pv}\,\Theta) * \varphi \in \mathscr{O}_{\mathrm{M}}(\mathbb{R}^n)$ and*

$$\big((\mathrm{pv}\,\Theta) * \varphi\big)(x) = \lim_{\epsilon\to 0_+} \int_{|x-y|\geq\epsilon} \Theta(x-y)\varphi(y)\,dy \quad \text{for all } x \in \mathbb{R}^n.$$

PROOF. Using Proposition 2.16.6, we know that $\mathrm{pv}\,\Theta \in \mathscr{S}'(\mathbb{R}^n)$. For each $\varphi \in \mathscr{S}(\mathbb{R}^n)$, using Theorem 2.14.6, we know that $(\mathrm{pv}\,\Theta)*\varphi \in \mathscr{O}_{\mathrm{M}}(\mathbb{R}^n)$. On the other hand, using (2.12.1), we compute that

$$((\mathrm{pv}\,\Theta) * \varphi)(x) = (\mathrm{pv}\,\Theta)(\varphi(x-\cdot)) = \lim_{\epsilon\to 0_+} \int_{|z|\geq\epsilon} \Theta(z)\varphi(x-z)\,dz,$$

which conclude our proposition. □

Let Θ_1, Θ_2 be functions satisfy (2.16.3). From Proposition 2.16.6, we know that $\mathrm{pv}\,\Theta_j \in \mathscr{S}'(\mathbb{R}^n)$ for $j = 1,2$. We now want to define the convolution $(\mathrm{pv}\,\Theta_1)*(\mathrm{pv}\,\Theta_2)$. Let $\psi \in \mathscr{D}(\mathbb{R}^n)$ such that $\psi \equiv 1$ near the origin, and we write

$$\begin{aligned}
T_{00} &:= \big(\psi \mathrm{pv}\,\Theta_1\big) * \big(\psi \mathrm{pv}\,\Theta_2\big),\\
T_{01} &:= \big(\psi \mathrm{pv}\,\Theta_1\big) * \big((1-\psi)\mathrm{pv}\,\Theta_2\big),\\
T_{10} &:= \big((1-\psi)\mathrm{pv}\,\Theta_1\big) * \big(\psi \mathrm{pv}\,\Theta_2\big),\\
T_{11} &:= \big((1-\psi)\mathrm{pv}\,\Theta_1\big) * \big((1-\psi)\mathrm{pv}\,\Theta_2\big).
\end{aligned}$$

Note that T_{00}, T_{01} and T_{10} are well-defined by Theorem 2.11.1. In particular, T_{11} is well-defined as follows:

EXERCISE 2.16.12. Show that $T_{11} = f_1 * f_2$ where $f_j = (1-\psi)\Theta_j$ for $j = 1,2$, are functions belonging in $L^2(\mathbb{R}^n)$.

From this, we know that u_{11} is well-defined in $\mathscr{S}'(\mathbb{R}^n)$. Then it is make sense to define

$$(\mathrm{pv}\,\Theta_1) * (\mathrm{pv}\,\Theta_2) := T_{00} + T_{01} + T_{10} + T_{11} \quad \text{in } \mathscr{S}'(\mathbb{R}^n).$$

Since $\mathscr{F} : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ is a continuous bijective, then $\mathscr{F}\big((\mathrm{pv}\,\Theta_1) * (\mathrm{pv}\,\Theta_2)\big)$ is make sense. In fact, we have

$$\mathscr{F}\big((\mathrm{pv}\,\Theta_1) * (\mathrm{pv}\,\Theta_2)\big) = m_{\Theta_1} m_{\Theta_2} \quad \text{in } \mathscr{S}'(\mathbb{R}^n),$$

see [**Mit18**, Section 4.6] for more details.

EXAMPLE 2.16.13. For each $j \in \{1,\cdots,n\}$, the operators R_j defined by

$$R_j\varphi := \left(\mathrm{pv}\,\frac{x_j}{|x|^{n+1}}\right) * \varphi \quad \text{for all } \varphi \in \mathscr{S}(\mathbb{R}^n),$$

are called the *Riesz transform* in $\mathbb{R}^n$. In the particular case when $n = 1$ the corresponding operator

$$H\varphi := \left(\operatorname{pv}\frac{1}{x}\right) * \varphi \quad \text{for all } \varphi \in \mathscr{S}(\mathbb{R}),$$

is called the *Hilbert transform*. These operators play a fundamental role in harmonic analysis. In particular, the Riesz transform can be extended as a bounded linear operator

$$R_j : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$$

and its Fourier transform is given by

$$(R_j\varphi)^\wedge(\xi) = C_n \frac{i\xi_j}{|\xi|}\hat{\varphi}(\xi) \quad \text{in } \mathscr{S}'(\mathbb{R}^n) \quad \text{with } C_n = -\frac{\pi^{\frac{n+1}{2}}}{\Gamma(\frac{n+1}{2})},$$

see e.g. [**Mit18**, Theorem 4.97] for more details. By writing $\nabla = (\partial_1, \cdots, \partial_n)$ and $R = (R_1, \cdots, R_n)$, since $(\partial_j\varphi)^\wedge(\xi) = i\xi_j\hat{\varphi}(\xi)$ and $(-\Delta\varphi)^\wedge(\xi) = |\xi|^2\hat{\varphi}(\xi)$, therefore sometimes we simply denote

$$R_j = C_n\partial_j(-\Delta)^{-\frac{1}{2}}$$

or in vector notation $R = C_n\nabla(-\Delta)^{-\frac{1}{2}}$.

EXERCISE 2.16.14. Show that if $f \in C^0(\mathbb{R}^n \setminus \{0\})$ is positive homogeneous of degree $k \in \mathbb{R}$ with $k > -n$, then $f \in \mathscr{S}'(\mathbb{R}^n)$.

EXERCISE 2.16.15. Prove that if $f \in C^0(\mathbb{R}^n \setminus \{0\})$ is positive homogeneous of degree $1-n$ on $\mathbb{R}^n \setminus \{0\}$ and $g \in C^0(\mathcal{S}^{n-1})$, then

$$\int_{\partial B_R(0)} g(x/R)f(x)\,ds(x) = \int_{\mathcal{S}^{n-1}} g(x)f(x)\,ds(x) \quad \text{for all } R > 0.$$

Based on the above observation, we discusses a basic class of principal value tempered distributions.

EXAMPLE 2.16.16. Let $\Phi \in C^1(\mathbb{R}^n \setminus \{0\})$ be positive homogeneous of degree $1-n$ (and hence $\Phi \in \mathscr{S}'(\mathbb{R}^n)$ by Exercise 2.16.14). Then for each $j \in \{1, \cdots, n\}$ it follows that $\partial_j\Phi$ satisfies the conditions in (2.16.3). Consequently, $\operatorname{pv}(\partial_j\Phi)$ is a well-defined tempered distribution by Proposition 2.16.6. The condition $\int_{\mathcal{S}^{n-1}}(\partial_j\Phi)(\hat{x})\,ds(\hat{x}) = 0$ can be verified using Exercise 2.16.15 as follows:

$$\begin{aligned}
0 &= \int_{|x|=2} \Phi(x)\frac{x_j}{2}\,ds(x) - \int_{|x|=1}\Phi(x)x_j\,ds(x) = \int_{1<|x|<2}\partial_j\Phi(x)\,dx \\
&= \int_1^2\int_{\mathcal{S}^{n-1}}(\partial_j\Phi)(r\hat{x})r^{n-1}ds(\hat{x})\,dr = \left(\int_1^2\frac{1}{r}\,dr\right)\int_{\mathcal{S}^{n-1}}(\partial_j\Phi)(\hat{x})\,ds(\hat{x}) \\
&= \ln 2\int_{\mathcal{S}^{n-1}}(\partial_j\Phi)(\hat{x})\,ds(\hat{x}).
\end{aligned}$$

Principal value tempered distributions often arise when differentiating certain types of functions exhibiting a point singularity. Specifically, we have the following theorem:

THEOREM 2.16.17. *Let $\Phi \in C^1(\mathbb{R}^n \setminus \{0\})$ be a function that is positive homogeneous of degree $1-n$. Then for each $j \in \{1, \cdots, n\}$, the distributional derivative $\partial_j\Phi$ satisfies*

$$\partial_j\Phi = \left(\int_{\mathcal{S}^{n-1}}\Phi(\hat{x})\hat{x}_j\,ds(\hat{x})\right)\delta_0 + \operatorname{pv}(\partial_j\Phi) \quad \textit{in } \mathscr{S}'(\mathbb{R}^n). \tag{2.16.8}$$

Proof. Using Proposition 2.8.16, it is suffice to prove (2.16.8) in $\mathscr{D}'(\mathbb{R}^n)$. Fixing any $\varphi \in \mathscr{D}(\mathbb{R}^n)$, we see that

$$\begin{aligned}(\partial_j \Phi)(\varphi) &= -\int_{\mathbb{R}^n} \Phi(x)\partial_j\varphi(x)\,dx = -\lim_{\epsilon\to 0_+}\int_{|x|\geq\epsilon}\Phi(x)\partial_j\varphi(x)\,dx\\ &= \lim_{\epsilon\to 0_+}\int_{|x|\geq\epsilon}\partial_j\Phi(x)\varphi(x)\,dx + \lim_{\epsilon\to 0_+}\int_{|x|=\epsilon}\hat{x}_j\Phi(x)\varphi(x)\,ds(x)\\ &= \big(\mathrm{pv}\,(\partial_j\Phi)\big)(\varphi) + \lim_{\epsilon\to 0_+}\int_{|x|=\epsilon}\frac{x_j}{\epsilon}\Phi(x)\varphi(x)\,ds(x).\end{aligned}$$

For each $\epsilon > 0$, using Exercise 2.16.15 we see that

$$\begin{aligned}&\int_{|x|=\epsilon}\frac{x_j}{\epsilon}\Phi(x)\varphi(x)\,ds(x)\\ &= \int_{|x|=\epsilon}\frac{x_j}{\epsilon}\Phi(x)\big(\varphi(x)-\varphi(0)\big)\,ds(x) + \varphi(0)\int_{|x|=\epsilon}\frac{x_j}{\epsilon}\Phi(x)\,ds(x)\\ &= \int_{|x|=\epsilon}\frac{x_j}{\epsilon}\Phi(x)\big(\varphi(x)-\varphi(0)\big)\,ds(x) + \varphi(0)\int_{\mathcal{S}^{n-1}}\hat{x}\Phi(\hat{x})\,ds(\hat{x}).\end{aligned}$$

In addition, using Exercise 2.16.5 we may estimate

$$\begin{aligned}&\left|\int_{|x|=\epsilon}\frac{x_j}{\epsilon}\Phi(x)\big(\varphi(x)-\varphi(0)\big)\,ds(x)\right|\\ &\leq \epsilon\|\nabla\varphi\|_{L^\infty(\mathbb{R}^n)}\|\Phi\|_{L^\infty(\mathcal{S}^{n-1})}\int_{|x|=\epsilon}\frac{1}{|x|^{n-1}}\,ds(x)\\ &= \epsilon\|\nabla\varphi\|_{L^\infty(\mathbb{R}^n)}\|\Phi\|_{L^\infty(\mathcal{S}^{n-1})}|\mathcal{S}^{n-1}| \to 0 \quad \text{as } \epsilon\to 0_+.\end{aligned}$$

Combining the equations above, we conclude our theorem. □

Bibliography

[Bre11] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011. MR2759829, doi:10.1007/978-0-387-70914-7.

[EG15] L. C. Evans and R. Gariepy. *Measure theory and fine properties of functions*. Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition, 2015. MR3409135, doi:10.1201/b18333.

[FJ98] F. G. Friedlander and M. Joshi. *Introduction to the theory of distributions*. Cambridge University Press, Cambridge, second edition, 1998. MR1721032, Zbl:0971.46024.

[GS97] D. Gottlieb and C.-W. Shu. On the Gibbs phenomenon and its resolution. *SIAM Rev.*, 39(4):644–668, 1997. MR1491051, Zbl:0885.42003, doi:10.1137/S0036144596301390.

[HH79] E. Hewitt and R.E. Hewitt. The Gibbs-Wilbraham phenomenon: an episode in Fourier analysis. *Arch. Hist. Exact Sci.*, 21(2):129–160, 1979. MR0555102, doi:10.1007/BF00330404.

[Mit18] D. Mitrea. *Distributions, partial differential equations, and harmonic analysis*. Universitext. Springer, Cham, second edition, 2018. MR3887685, doi:10.1007/978-3-030-03296-8.

[Onn82] C. W. Onneweer. The Wilbraham-Gibbs phenomenon in Fourier analysis. *Math. Medley*, 10(2):47–54, 1982. MR0697668, Zbl:0517.01023, `https://sms.math.nus.edu.sg/volume-10-1982/`.

[Rud91] W. Rudin. *Functional Analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, second edition, 1991. MR1157815.

[Str08] W. A. Strauss. *Partial differential equations: An introduction*. John Wiley & Sons, Ltd., Chichester, second edition, 2008. MR2398759.

[Zyg02] A. Zygmund. *Trigonometric series. Vol. I, II.* Cambridge Math. Lib. Cambridge University Press, Cambridge, third edition, 2002. MR1963498, Zbl:1084.42003, doi:10.1017/CBO9781316036587.