Notes on Distribution Theory

References ① Functional Analysis, Sobolev spaces and Partial Differential Equations by Haim Brezis, Springer. (Chapter 8)

② Real Analysis - Modern Techniques and Their Applications by Gerald B. Folland, Second edition, John Wiley & Sons. (Chapter 9)

③ Partial Differential Equations by Jeffrey Rauch, Springer-Verlag. (Appendix)

④ Functional Analysis by Walter Rudin, Second edition, McGraw-Hill. (Chapters 6, 7)

In the nineteenth Century, physicists and engineers had to deal with practical problems involving non-differentiable functions. However, it was found that the usual rules of Calculus, applied formally, were quite successful in the solution of a majority of such problems. A rigorous mathematical theory to explain such operations was therefore needed at that time. Laurent Schwartz's theory of distributions, developed in 1940s and 1950s, has turned out to be a useful method in answering the question and is one of the mainstay of modern analysis.

Schwartz's attempt in solving the above problem was to consider a class of objects containing the "classical" differentiable functions and extend the usual rules of calculus on such objects. These objects turned out to be continuous linear functionals on certain "test" functions. We refer to these objects as "distributions" or "generalized functions". We should be able to differentiate these objects and for differentiable functions, the two notions of derivatives should match.

To motivate the need for such linear functionals, consider the following situation. In many problems of Mechanics and Electricity, we need to work with physical observables u which depends on the location. It therefore becomes impossible to measure $u(x)$ for all x, due to finite size/time requirement of the measurement. In such situations, we consider an "average" value $\int u(x)\,\phi(x)\,dx$, against some "standard" weight functions ϕ. Call such ϕ as "test" functions. For an appropriate class of test functions ϕ, we now observe u as a linear functional

$$\phi \longmapsto \int u(x)\,\phi(x)\,dx.$$

Of course, in the above consideration, we require $\int u(x)\,\phi(x)\,dx$ to be well-defined. The test functions ϕ should also be "smooth" or be differentiable of an appropriate order. To allow finiteness of $\int u(x)\phi(x)dx$ where u is "large" at infinity, we shall also impose that the support of the test functions should be compact.

Note ①: Above problems shall be considered on $\mathbb{R}^d$ or some open subset of $\mathbb{R}^d$. The domain Ω, of u and ϕ, shall therefore be $\mathbb{R}^d$ or some open subset of $\mathbb{R}^d$, if not stated otherwise. Ω should be non-empty.

Note ②: The functions that we consider are real or complex valued. The function spaces appearing in our discussion below shall be taken as $\mathbb{R}$ or $\mathbb{C}$ - vector spaces, as appropriate.

Note ③: The integrals $\int u(x)\,\phi(x)\,dx$ shall be interpreted as Lebesgue integrals, i.e. with respect to the Lebesgue measure on Ω.

Definition ① (Support of a function)

Let $f: \Omega \longrightarrow \mathbb{R}$ (or $\mathbb{C}$) be a function. The closure of $\{x \in \Omega \mid f(x) \neq 0\}$

is referred to as the support of f and is denoted by $\text{supp}(f)$.

Notation ①: We write $C_c^\infty(\Omega)$ to denote the class of infinitely differentiable functions (on Ω) whose support is compact and contained in Ω. Following Schwartz, we also write $\mathcal{D}(\Omega)$ to denote $C_c^\infty(\Omega)$.

Notation ②: We write $C^\infty(\Omega)$ to denote the class of infinitely differentiable functions on Ω. Following Schwartz, we also write $\mathcal{E}(\Omega)$ to denote $C^\infty(\Omega)$.

Note ④: To talk about the usual rules of calculus, involving differentiation and convergence theorems, we need the linear functionals u to be "continuous". Thus we are led to the topology of $\mathcal{D}(\Omega)$ etc., before we can discuss the "continuity" of linear functionals u on $\mathcal{D}(\Omega)$.

Note ⑤: In the interest of time, we

discuss the topology of $\mathcal{D}(\Omega)$ only through the convergence of sequences. The reader may refer to the book by G. Folland or W. Rudin for a detailed discussion.

Notation ③: $x = (x_1, x_2, \ldots, x_d)^t \in \mathbb{R}^d$ denotes an arbitary element of $\mathbb{R}^d$.

Notation ④: $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \in \{0, 1, 2, \ldots\}^d$ denotes a multi-index α. We write

$$|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_d.$$

Notation ⑤: For any multi-index α and $\phi \in \mathcal{D}(\Omega)$, we write $\partial^\alpha \phi := \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \cdots \partial_{x_d}^{\alpha_d} \phi$.

Note that $\partial^\alpha \phi \in \mathcal{D}(\Omega)$. If $\alpha_j = 0$ for some j, then the differentiation with respect to the j-th co-ordinate is ignored in $\partial^\alpha \phi$.

Note ⑥: If $\phi \in \mathcal{D}(\Omega)$ with $\mathrm{supp}(\phi) \subseteq K$ for some compact $K \subseteq \Omega$, then

$$\mathrm{supp}(\partial^\alpha \phi) \subseteq K$$

for all multi-indices α.

Notation ⑥: For any integer $n \geq 0$ and compact subsets K of Ω, we define

$$\|\phi\|_{n,K} := \sum_{|\alpha| \leq n} \sup_{x \in K} |\partial^\alpha \phi|$$

for all $\phi \in \mathcal{D}(\Omega)$.

Note ⑦: If $\{\phi_j\}_j$ is a sequence in $\mathcal{D}(\Omega)$ converging pointwise to $\phi \in \mathcal{D}(\Omega)$ and if there exists a compact subset K of Ω such that $supp(\phi_j) \subseteq K \ \forall j$, then we have $supp(\phi) \subseteq K$.

Definition ② (Convergence in $\mathcal{D}(\Omega)$)

We say that a sequence $\{\phi_j\}_j$ in $\mathcal{D}(\Omega)$ converges in $\mathcal{D}(\Omega)$ to ϕ if there exists a compact subset K of Ω such that $Supp(\phi_j) \subseteq K \ \forall j$ and $\{\partial^\alpha \phi_j\}_j$ converges uniformly to $\partial^\alpha \phi$ for all α.

Note ⑧: In Definition ② above, the statement for uniform convergence is equivalent to $\lim_{j \to \infty} \|\phi_j - \phi\|_{n,K} = 0$ for all $n \geq 0$, with K as in Definition ②.

Definition ③ (Distribution)

A distribution T is a continuous linear functional on $\mathcal{D}(\Omega)$. The space of all distributions is the dual space of

$\mathcal{D}(\Omega)$ and is denoted by $\mathcal{D}'(\Omega)$.

Note ⑨: Definition ③ may be restated as follows. A distribution T is a linear functional on $\mathcal{D}(\Omega)$ and is continuous in the following sense: if $\{\phi_j\}_j$ converges in $\mathcal{D}(\Omega)$ to ϕ, then $\{T(\phi_j)\}_j$ converges to $T(\phi)$.

Note ⑩: If we consider complex valued linear functionals, then $\mathcal{D}'(\Omega)$ is a complex vector space. If we consider only real valued linear functionals, then $\mathcal{D}'(\Omega)$ is a real vector space.

Notation ⑦: We simply write $\mathcal{D}'$ instead of $\mathcal{D}'(\mathbb{R}^d)$. The dimension d is usually understood from the context.

Notation ⑧: For $\phi \in \mathcal{D}(\Omega)$ and $T \in \mathcal{D}'(\Omega)$, we write $\langle T, \phi \rangle$ to denote $T(\phi)$. To clarify further, we have

$$\langle T, \phi \rangle = {}_{\mathcal{D}'(\Omega)}\langle T, \phi \rangle_{\mathcal{D}(\Omega)} = T(\phi) = T\phi .$$

An immediate consequence of Definition ③ is the following result.

Proposition ①: A linear functional T on $\mathcal{D}(\Omega)$ is a distribution if and only if for every compact subset K of Ω, there exists an integer $n = n(K,T) \geq 0$ and $c = c(K,T) > 0$ such that

$$|\langle T, \phi \rangle| \leq c \cdot \|\phi\|_{n,K}$$

for all $\phi \in \mathcal{D}(\Omega)$ with $\operatorname{supp}(\phi) \subseteq K$.

Example ① Fix $f \in \mathcal{D}(\Omega)$ and consider the linear functional T_f defined by

$$\langle T_f, \phi \rangle := \int f(x)\,\phi(x)\,dx \quad \forall\, \phi \in \mathcal{D}(\Omega).$$

Now $|\langle T_f, \phi \rangle| \leq \left(\int_K |f(x)|\,dx \right) \cdot \|\phi\|_{0,K}$

for all $\phi \in \mathcal{D}(\Omega)$ with $\operatorname{supp}(\phi) \subseteq K$, where K is some compact subset of Ω. Here, every element of $\mathcal{D}(\Omega)$ is identified in $\mathcal{D}'(\Omega)$.

Notation ⑨: $L^1_{loc}(\Omega)$ denotes the space of all measurable functions on Ω such that $\int_K |f(x)|\,dx < \infty$ for all compact subset K of Ω.

Example ② Fix $f \in L^1_{loc}(\Omega)$ and consider the linear functional T_f defined by

$$\langle T_f, \phi \rangle := \int f(x)\phi(x)\,dx, \ \forall \phi \in \mathcal{D}(\Omega).$$

As discussed in Example ①, we have $T_f \in \mathcal{D}'(\Omega)$. From now onwards, we shall write f instead of T_f – in which case, $\langle f, \phi \rangle = \langle T_f, \phi \rangle$.

<u>Note ⑪</u>: We say that a distribution T is "represented by a function f" or "is a function f" if $T = T_f$. Example ① and Example ② are examples of this type.

<u>Example ③</u>: let μ be a Radon measure on Ω, i.e. it is a Borel measure on Ω such that $\mu(K) < \infty$ for all compact subset K of Ω. All finite measures, including probability measures, are examples of Radon measures. Consider the linear functional T_μ defined by

$$\langle T_\mu, \phi \rangle := \int_\Omega \phi(x)\,d\mu(x), \ \forall \phi \in \mathcal{D}(\Omega).$$

We have $|\langle T_\mu, \phi \rangle| \le \mu(K) \cdot \|\phi\|_{0,K}$ for all $\phi \in \mathcal{D}(\Omega)$ with $\mathrm{supp}(\phi) \subseteq K$, where K is some compact subset of Ω. Thus T_μ

is a distribution. As in Example ②, we shall write μ instead of T_μ.

Example ④ (Dirac Distribution)

This is a special case of Example ③. Fix $x \in \Omega$. Consider the Dirac measure δ_x on Ω. As in Example ③, take the distribution given by

$$\langle \delta_x, \phi \rangle = \int \phi(y)\, d\delta_x(y)$$

$$= \phi(x), \quad \forall \phi \in \mathcal{D}(\Omega).$$

This distribution is not represented by a function.

Note ⑫: The topology on $\mathcal{D}'(\Omega)$ is taken as the topology of pointwise convergence on $\mathcal{D}(\Omega)$. This is also referred to as the weak* topology. We state this in terms of convergence of distributions in the next definition.

Definition ④ (Convergence of a Sequence of distributions)

We say that a sequence $\{T_n\}_n$ in

$\mathcal{D}'(\Omega)$ converges to T in $\mathcal{D}'(\Omega)$ if for all $\phi \in \mathcal{D}(\Omega)$, we have $\lim_{n\to\infty} \langle T_n, \phi \rangle = \langle T, \phi \rangle$.

Exercise ①: Consider a sequence of distributions in $\mathcal{D}'(\mathbb{R})$ given by the Gaussian density functions

$$f_n(x) := \frac{n}{\sqrt{2\pi}} \exp\left(-\frac{n^2 x^2}{2}\right), \quad \forall\, x \in \mathbb{R}, \quad n = 1, 2, \ldots$$

corresponding to mean 0 and variance $\frac{1}{n^2}$. Check that $\{f_n\}_n$ converges in $\mathcal{D}'$ to δ_0.

We now consider various important operations on distributions, which allow us to construct further examples.

I. Addition and Scalar multiplication

Since $\mathcal{D}'(\Omega)$ is a vector space, for $T, S \in \mathcal{D}'(\Omega)$ and scalars α, β, we have $\alpha T + \beta S \in \mathcal{D}'(\Omega)$ given by

$$\langle \alpha T + \beta S, \phi \rangle = \alpha \langle T, \phi \rangle + \beta \langle S, \phi \rangle \quad \forall\, \phi \in \mathcal{D}(\Omega).$$

II. (Multiplication by Smooth functions)

Fix $\psi \in \mathcal{E}(\Omega)$. Then for any $\phi \in \mathcal{D}(\Omega)$, we have $\psi\phi \in \mathcal{D}(\Omega)$ (Exercise).

Moreover, for any $T \in \mathcal{D}'(\Omega)$, the linear functional $\phi \longmapsto \langle T, \psi\phi \rangle$ is continuous (Exercise). We define ψT in $\mathcal{D}'(\Omega)$ as $\langle \psi T, \phi \rangle := \langle T, \psi\phi \rangle, \ \forall \phi \in \mathcal{D}(\Omega)$.

Note ⑬: A motivation for the above definition goes as follows. Suppose $T \in \mathcal{D}'(\Omega)$ is given by $f \in \mathcal{D}(\Omega)$. Then with ϕ, ψ as above observe the identity

$$\begin{aligned} \langle f, \psi\phi \rangle &= \int f(x)\big(\psi(x)\,\phi(x)\big)\,dx \\ &= \int \big(\psi(x) f(x)\big)\,\phi(x)\,dx. \\ &= \langle \psi f, \phi \rangle \end{aligned}$$

This identity helps us in defining the multiplication by smooth functions.

III. (Translation)

For simplicity, we first discuss this

operation on $\mathcal{D}' = \mathcal{D}'(\mathbb{R}^d)$.

Fix $x \in \mathbb{R}^d$ and for any function f defined on $\mathbb{R}^d$, consider the function $\tau_x f$ on $\mathbb{R}^d$ defined by $(\tau_x f)(y) := f(y-x) \; \forall y \in \mathbb{R}^d$. observe that τ_x maps $\mathcal{D}(\mathbb{R}^d)$ into $\mathcal{D}(\mathbb{R}^d)$.

For any $f, \phi \in \mathcal{D}(\mathbb{R}^d)$, observe that

$$\begin{aligned}
\langle \tau_x f, \phi \rangle &= \int_{\mathbb{R}^d} (\tau_x f)(y) \, \phi(y) \, dy \\
&= \int f(y-x) \, \phi(y) \, dy \\
&= \int f(y) \, \phi(y+x) \, dy \\
&= \int f(y) \, (\tau_{-x} \phi)(y) \, dy \\
&= \langle f, \tau_{-x} \phi \rangle .
\end{aligned}$$

The above identity motivates the translation of distribution as follows. For $T \in \mathcal{D}'(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$, define the translated distribution $\tau_x T \in \mathcal{D}'(\mathbb{R}^d)$ by

$$\langle \tau_x T, \phi \rangle := \langle T, \tau_{-x} \phi \rangle \; \forall \phi \in \mathcal{D}(\mathbb{R}^d).$$

<u>Note (14)</u>: Consider the Dirac distribution δ_0 in $\mathcal{D}'(\mathbb{R}^d)$. For any fixed $x \in \mathbb{R}^d$, we have

$$\langle \tau_x \delta_0, \phi \rangle = \langle \delta_0, \tau_{-x} \phi \rangle$$

$$= (\tau_{-x}\phi)(0)$$
$$= \phi(x)$$
$$= \langle \delta_x, \phi \rangle, \quad \forall\, \phi \in \mathcal{D}(\mathbb{R}^d).$$

Therefore $\tau_x \delta_0 = \delta_x$.

Note 15: It is possible to define translation on $\mathcal{D}'(\Omega)$. For $x \in \mathbb{R}^d$, consider the open set
$$\Omega + x = \{ y + x \mid y \in \Omega \}.$$
For any $\phi \in \mathcal{D}(\Omega + x)$, observe that $\tau_{-x}\phi$ is in $\mathcal{D}(\Omega)$. For $T \in \mathcal{D}'(\Omega)$, we can define $\tau_x T \in \mathcal{D}'(\Omega + x)$ by
$$_{\mathcal{D}'(\Omega+x)}\langle \tau_x T, \phi \rangle_{\mathcal{D}(\Omega+x)} := {}_{\mathcal{D}'(\Omega)}\langle T, \tau_{-x}\phi \rangle_{\mathcal{D}(\Omega)}$$
$$\forall\, \phi \in \mathcal{D}(\Omega + x).$$

IV. (Differentiation)

This is a very important operation in the goal of extending the rules of Calculus to distributions. In order to understand this operation, first consider the following simple case.

Let $f, \phi \in \mathcal{D}(\mathbb{R})$. Then, using integration by parts,

$$\langle f', \phi \rangle = \int f'(x)\,\phi(x)\,dx$$
$$= - \int f(x)\,\phi'(x)\,dx$$
$$= -\langle f, \phi' \rangle.$$

More generally, for $f, \phi \in \mathcal{D}(\Omega)$ and for any multi-index α, we have

$$\langle \partial^\alpha f, \phi \rangle = \int \partial^\alpha f(x)\,\phi(x)\,dx$$
$$= (-1)^{|\alpha|} \int f(x)\,\partial^\alpha \phi(x)\,dx$$
$$= (-1)^{|\alpha|} \langle f, \partial^\alpha \phi \rangle.$$

With this identity at hand, given T in $\mathcal{D}'(\Omega)$ and a multi-index α, we define the derivative $\partial^\alpha T \in \mathcal{D}'(\Omega)$ by

$$\langle \partial^\alpha T, \phi \rangle := (-1)^{|\alpha|} \langle T, \partial^\alpha \phi \rangle \quad \forall \phi \in \mathcal{D}(\Omega).$$

We refer to this notion of derivatives as "distributional derivatives".

<u>Note (16):</u> Fix $f \in \mathcal{D}(\Omega)$. As discussed above, for any multi-index α, the distributional derivative $\partial^\alpha f$ is given by the linear functional on $\mathcal{D}(\Omega)$

$$\phi \longmapsto \langle \partial^\alpha f, \phi \rangle = \int \partial^\alpha f(x)\, \phi(x)\, dx .$$

Hence, the distribution $\partial^\alpha f$ is represented by the function $\partial^\alpha f \in \mathcal{D}(\Omega)$. As such, the notion of distributional derivative agrees with the "classical" notion of derivative for functions in $\mathcal{D}(\Omega)$.

Note 17: (i) A distribution has derivatives of all orders.

(ii) Given any two multi-indices α & β, we have $\partial^\alpha \partial^\beta T = \partial^\beta \partial^\alpha T = \partial^{\alpha+\beta} T$, for all $T \in \mathcal{D}'(\Omega)$ where $\alpha+\beta = (\alpha_1+\beta_1, \dots, \alpha_2+\beta_2)$.

Example 5 Consider the Heaviside step function H on $\mathbb{R}$ as $H(x) := \mathbb{1}_{(0,\infty)}(x)$, $x \in \mathbb{R}$. Since $H \in \mathcal{L}^1_{loc}(\mathbb{R})$, we have $H \in \mathcal{D}'(\mathbb{R})$. Note that H is not differentiable at 0. However, we may compute the distributional derivatives as follows. For $\phi \in \mathcal{D}(\mathbb{R})$,

$$\begin{aligned}
\langle \partial H, \phi \rangle &= - \langle H, \partial \phi \rangle \\
&= - \int_{-\infty}^{\infty} H(x)\, \phi'(x)\, dx \\
&= - \int_0^{\infty} \phi'(x)\, dx
\end{aligned}$$

$$= \phi(0)$$

$$= \langle \delta_0, \phi \rangle.$$

Hence, $\partial H = \delta_0$.

<u>Notation (10)</u>: Given a function f on $\mathbb{R}^d$, we denote by $\tilde{f}$ its reflection in the origin defined by $\tilde{f}(x) := f(-x)$, $\forall x \in \mathbb{R}^d$.

<u>V. (Convolution)</u>

Fix $\psi \in \mathcal{D}(\mathbb{R}^d)$ and consider the set

$$V := \{x \in \mathbb{R}^d \mid x - y \in \Omega \text{ for } y \in \operatorname{Supp}(\psi)\}.$$

Given any $f \in \mathcal{L}^1_{loc}(\Omega)$, we have for $x \in V$

$$f * \psi(x) = \int f(x-y)\,\psi(y)\,dy$$

$$= \int f(y)\,\psi(x-y)\,dy$$

$$= \int f(y)\,(\tau_x \tilde{\psi})(y)\,dy$$

$$= \langle f, \tau_x \tilde{\psi} \rangle.$$

This motivates the following definition. For any $T \in \mathcal{D}'(\Omega)$, define the convolution $T * \psi$ as a function defined on V as

$$T * \psi(x) := \langle T, \tau_x \tilde{\psi} \rangle.$$

<u>Note (18)</u> (i) V may be an empty set.

(ii) A special case arises if we take $T = \delta_0$. Then

$$\delta_0 * \psi(x) = \langle \delta_0, \tau_x \tilde{\psi} \rangle$$
$$= (\tau_x \tilde{\psi})(0)$$
$$= \psi(x) \quad \forall x \in V.$$

(iii) From the above discussion, we also have the following connection between the convolution and the translation operation. For $\phi, \psi \in \mathcal{D}(\mathbb{R}^d)$, we have

$$\phi * \psi(x) = \int \phi(y)\, \psi(x-y)\, dy$$
$$= {}_{\mathcal{D}'(\mathbb{R}^d)}\langle \phi, \tau_x \tilde{\psi} \rangle_{\mathcal{D}(\mathbb{R}^d)}$$
$$= {}_{\mathcal{D}'(\mathbb{R}^d)}\langle \tau_x \tilde{\psi}, \phi \rangle_{\mathcal{D}(\mathbb{R}^d)}.$$

Proposition 2: Continue with ψ and V as above. Fix $T \in \mathcal{D}'(\Omega)$. Then

(i) $T * \psi \in \mathcal{E}(\Omega)$

(ii) $\partial^\alpha (T * \psi) = (\partial^\alpha T) * \psi = T * (\partial^\alpha \psi)$ for any multi-index α.

(iii) For any $\phi \in \mathcal{D}(V)$,

$$\int (T * \psi)(x)\, \phi(x)\, dx = \langle T, \phi * \tilde{\psi} \rangle.$$

Note 19: We shall discuss the notion of localization of distributions. To discuss this notion, we first introduce some terminology.

Definition ⑤ (Equality of Distributions on open subsets)

(i) Let $T, S \in \mathcal{D}'(\Omega)$ and let U be an open subset of Ω. We say that $T = S$ on U if $\langle T, \phi \rangle = \langle S, \phi \rangle \ \forall \phi \in \mathcal{D}(U)$. Here, every $\phi \in \mathcal{D}(U)$ is thought of as an element of $\mathcal{D}(\Omega)$, extending ϕ as 0 on $\Omega \cap U^c$.

(ii) Let T and U be as in (i). We say that T is equal to zero on U if $\langle T, \phi \rangle = 0 \ \forall \phi \in \mathcal{D}(U)$.

Theorem ①: (Locally finite partition of Unity in Ω)

Let Γ be a collection of open sets in $\mathbb{R}^d$ such that the union of them is Ω, i.e., Γ is an open cover of Ω.

Then there exists a sequence of non-negative functions $\{\psi_j\}_j$ in $\mathcal{D}(\Omega)$ such that

(i) For each j, $\operatorname{supp}(\psi_j) \subseteq U$ for some $U \in \Gamma$.

(ii) $\sum_{j=1}^{\infty} \psi_j(x) = 1 \ \forall x \in \Omega$

(iii) Given any compact subset K of Ω, there exist an integer m and an open set U with $K \subset U \subset \Omega$ with

$$\sum_{j=1}^{m} \psi_j(x) = 1 \quad \forall x \in U.$$

Such a collection of functions in $\mathcal{D}(\Omega)$ is called a locally finite partition of unity in Ω, subordinate to the open cover Γ.

VI. (localization)

Theorem ②: let Γ be an open cover of Ω. Suppose for each $U \in \Gamma$ we have $T_U \in \mathcal{D}'(U)$ such that $T_U = T_V$ on $U \cap V$, whenever $U \cap V \neq \phi$. Then there exists a unique $T \in \mathcal{D}'(\Omega)$ such that $T = T_U$ on U $\forall U \in \Gamma$.

Note ⑳: The product of distributions is usually not defined.

Definition ⑥ (support of a distribution)

let $T \in \mathcal{D}'(\Omega)$. We define the support of T, denoted by $\operatorname{supp}(T)$, as the complement of $\{x \in \Omega \mid T$ is equal to zero on a neighbourhood of $x\}$.

Note (21): (i) $\operatorname{supp}(\delta_x) = \{x\}$.

(ii) If a distribution T is represented by a continuous function f, then

$$\operatorname{supp}(T) = \operatorname{supp}(f).$$

(iii) For $T \in \mathcal{D}'(\Omega)$ and any multi-index α,

$$\operatorname{supp}(\partial^\alpha T) \subseteq \operatorname{supp}(T).$$

(iv) For $T \in \mathcal{D}'(\Omega)$ and $\psi \in \mathcal{D}(\Omega)$,

$$\operatorname{supp}(\psi T) \subseteq \operatorname{supp}(\psi) \cap \operatorname{supp}(T).$$

Proposition (3): If $\phi \in \mathcal{D}(\Omega)$ and $T \in \mathcal{D}'(\Omega)$ be such that $\operatorname{supp}(\phi) \cap \operatorname{supp}(T) = \phi$, then $\langle T, \phi \rangle = 0$.

Notation (11): $\mathcal{E}'(\Omega) := \{T \in \mathcal{D}'(\Omega) \mid \operatorname{supp}(T) \text{ is compact}\}$.

Definition (7) (Compactly supported distributions)

Elements of $\mathcal{E}'(\Omega)$ are called distributions with compact support or compactly supported distributions.

Note (22): (i) $\mathcal{E}'(\Omega)$, as defined above, is a specific subset of $\mathcal{D}'(\Omega)$. In particular, the linear functionals in $\mathcal{E}'(\Omega)$ are defined on $\mathcal{D}(\Omega)$. In the subsequent discussion, we show that $\mathcal{E}'(\Omega)$ is actually the dual space of $\mathcal{E}(\Omega)$.

(ii) We consider a topology on $\mathcal{E}(\Omega)$ through the notion of convergence. It is the topology of uniform convergence of functions, together with their derivatives, on compact subsets of Ω. It is possible to describe the same topology through a class of seminorms. The interested reader may consider the book by Folland for the details.

Proposition (4): $\mathcal{D}(\Omega)$ is dense in $\mathcal{E}(\Omega)$.

Theorem (3): $T \in \mathcal{D}'(\Omega)$ has compact support if and only if T extends uniquely to a continuous linear functional on $\mathcal{E}(\Omega)$.

Note (23): Theorem (3) justifies the usage of the notation $\mathcal{E}'(\Omega)$ (as defined in Notation (11)) Also refer to Note (22) above.

Example (6): We now discuss examples of compactly supported distributions. Refer to Note (21) above in this regard.

(i) The Dirac distribution δ_x.

(ii) Any $f \in \mathcal{D}(\Omega)$, as a distribution, is an example of this type.

(iii) If $T \in \mathcal{D}'(\Omega)$ with compact support, then $\partial^\alpha T \in \mathcal{D}'(\Omega)$ also has compact support for any multi-index α. This raises the question about defining a derivative on $\mathcal{E}'(\Omega)$ — which we shall discuss shortly.

(iv) If $T \in \mathcal{D}'(\Omega)$ and $\psi \in \mathcal{D}(\Omega)$, then $\psi T \in \mathcal{D}'(\Omega)$ has compact support. This raises the question about multiplication by functions on $\mathcal{E}'(\Omega)$ - which we shall discuss shortly.

Similar to the discussion for $\mathcal{D}'(\Omega)$, we now discuss various operations on $\mathcal{E}'(\Omega)$.

I. (Addition and Scalar Multiplication)

Since $\mathcal{E}'(\Omega)$ is a vector space, for $T, S \in \mathcal{E}'(\Omega)$ and scalars α, β, we have

$$ {}_{\mathcal{E}'(\Omega)}\langle \alpha T + \beta S, \phi \rangle_{\mathcal{E}(\Omega)} := \alpha \langle T, \phi \rangle + \beta \langle S, \phi \rangle, \quad \forall \phi \in \mathcal{E}(\Omega). $$

II. (Multiplication by smooth functions)

For $\psi \in \mathcal{E}(\Omega)$ and $T \in \mathcal{E}'(\Omega)$, we have $\psi T \in \mathcal{E}'(\Omega)$ defined by

$$ {}_{\mathcal{E}'(\Omega)}\langle \psi T, \phi \rangle_{\mathcal{E}(\Omega)} := {}_{\mathcal{E}'(\Omega)}\langle T, \psi\phi \rangle_{\mathcal{E}(\Omega)} \quad \forall \phi \in \mathcal{E}(\Omega). $$

III. (Translation)

For $x \in \mathbb{R}^d$ and $T \in \mathcal{E}'(\mathbb{R}^d)$, consider the translated distribution $\tau_x T \in \mathcal{E}'(\mathbb{R}^d)$ defined by

$$ {}_{\mathcal{E}'(\mathbb{R}^d)}\langle \tau_x T, \phi \rangle_{\mathcal{E}(\mathbb{R}^d)} := \langle T, \tau_{-x} \phi \rangle \quad \forall \phi \in \mathcal{E}(\mathbb{R}^d). $$

IV. (Differentiation)

For any multi-index α and $T \in \mathcal{E}'(\Omega)$, consider $\partial^\alpha T \in \mathcal{E}'(\Omega)$, defined by

$$ {}_{\mathcal{E}'(\Omega)}\langle \partial^\alpha T, \phi \rangle_{\mathcal{E}(\Omega)} := (-1)^{|\alpha|} \langle T, \partial^\alpha \phi \rangle, \quad \forall \phi \in \mathcal{E}(\Omega). $$

V. (Convolution)

(i) For any $\psi \in \mathcal{D}(\mathbb{R}^d)$ and $T \in \mathcal{E}'(\mathbb{R}^d)$, then the convolution $T * \psi$ is a function defined by

$$T * \psi(x) := \langle T, \tau_x \tilde{\psi} \rangle.$$

In fact, $T * \psi \in \mathcal{D}(\mathbb{R}^d)$.

(ii) For any $\psi \in \mathcal{E}(\mathbb{R}^d)$ and $T \in \mathcal{E}'(\mathbb{R}^d)$, the definition in (i), extends to a function $T * \psi$ defined by

$$T * \psi(x) := \langle T, \tau_x \tilde{\psi} \rangle.$$

In fact, $T * \psi \in \mathcal{E}(\mathbb{R}^d)$.

(iii) For $T \in \mathcal{E}'(\mathbb{R}^d)$ and $S \in \mathcal{D}'(\mathbb{R}^d)$, we define $T * S$ and $S * T$ in $\mathcal{D}'(\mathbb{R}^d)$ by

$${}_{\mathcal{D}'(\mathbb{R}^d)}\langle T * S, \phi \rangle_{\mathcal{D}(\mathbb{R}^d)} := \langle T, \tilde{S} * \phi \rangle$$

and $${}_{\mathcal{D}'(\mathbb{R}^d)}\langle S * T, \phi \rangle_{\mathcal{D}(\mathbb{R}^d)} := \langle S, \tilde{T} * \phi \rangle$$

$$\forall \phi \in \mathcal{D}(\mathbb{R}^d)$$

where $\tilde{S} \in \mathcal{D}'(\mathbb{R}^d)$ is defined by

$${}_{\mathcal{D}'(\mathbb{R}^d)}\langle \tilde{S}, \phi \rangle_{\mathcal{D}(\mathbb{R}^d)} := \langle S, \tilde{\phi} \rangle, \ \forall \phi \in \mathcal{D}(\mathbb{R}^d).$$

<u>Proposition 5</u>: If $T \in \mathcal{E}'(\mathbb{R}^d)$, then there exist a positive integer n, scalars c_α for multi-indices α with $|\alpha| \leq n$ and a continuous function f on $\mathbb{R}^d$ vanishing at infinity such that

$$T = \sum_{|\alpha| \leq n} c_\alpha \partial^\alpha f.$$

<u>Note (24)</u>: Fourier transformation is an important operation on functions. It is therefore a natural question to ask whether we can define such an operation for distributions. Given $\phi \in \mathcal{D}(\mathbb{R}^d)$, consider the Fourier transform $\hat{\phi}$ given by

$$\hat{\phi}(x) := (2\pi)^{-\frac{d}{2}} \int e^{-i x \cdot y} \phi(y)\, dy$$
$$\forall x \in \mathbb{R}^d.$$

Unfortunately, $\hat{\phi} \notin \mathcal{D}(\mathbb{R}^d)$. In fact, if $\hat{\phi}$ vanishes on some non-empty open set then $\phi \equiv 0$. Instead of working with $\mathcal{D}(\mathbb{R}^d)$, we consider a larger space of functions, viz. the Schwartz class which remains invariant under the Fourier transform.

Definition ⑧ (Schwartz space $\mathcal{S}(\mathbb{R}^d)$)

The Schwartz space $\mathcal{S}(\mathbb{R}^d)$ is defined as the set of infinitely differentiable functions ϕ on $\mathbb{R}^d$ such that

$$\max_{|\alpha| \leq n} \sup_x (1+|x|^2)^n |\partial^\alpha \phi(x)| < \infty$$
$$\forall n = 0, 1, 2, \ldots$$

Note 25 (i) By definition, $\mathcal{D}(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d) \subset \mathcal{E}(\mathbb{R}^d)$.

(ii) The function $x \mapsto \exp(-|x|^2)$ is in $\mathcal{S}(\mathbb{R}^d)$, but not in $\mathcal{D}(\mathbb{R}^d)$.

(iii) In literature, Schwartz class functions are also referred to as "rapidly decreasing smooth functions".

(iv) $\mathcal{S}(\mathbb{R}^d)$ forms a vector space.

(v) $\mathcal{S}(\mathbb{R}^d)$ is closed under differentiation.

(vi) $\mathcal{S}(\mathbb{R}^d)$ is closed under multiplication by polynomials.

(vii) $\mathcal{S}(\mathbb{R}^d) \subset \mathcal{L}^2(\mathbb{R}^d)$.

Theorem ④ : The Fourier transform is a continuous, linear, bijective mapping on $\mathcal{S}(\mathbb{R}^d)$, with continuous inverse.

Proposition ⑥: For $\phi, \psi \in \mathcal{S}(\mathbb{R}^d)$, we have

(i) $\phi\psi \in \mathcal{S}(\mathbb{R}^d)$

(ii) $\phi * \psi \in \mathcal{S}(\mathbb{R}^d)$

(iii) $(\phi\psi)^\wedge = \hat{\phi} * \hat{\psi}$.

Definition ⑨ (Topology on $\mathcal{S}(\mathbb{R}^d)$)

On $\mathcal{S}(\mathbb{R}^d)$, consider the seminorms

$$\phi \longmapsto \max_{|\alpha| \leq n} \sup_{x} (1+|x|^2)^n |\partial^\alpha \phi(x)|$$

for $n = 0, 1, 2, \ldots$. We consider the topology defined by these seminorms.

Proposition ⑦: $\mathcal{D}(\mathbb{R}^d)$ is dense in $\mathcal{S}(\mathbb{R}^d)$.

Definition ⑩ (Tempered distributions)

A tempered distribution is a continuous linear functional on $\mathcal{S}(\mathbb{R}^d)$. The space of all tempered distributions is the dual space of $\mathcal{S}(\mathbb{R}^d)$ and is denoted by $\mathcal{S}'(\mathbb{R}^d)$.

Example ⑦: $\mathcal{E}'(\mathbb{R}^d) \subset \mathcal{S}'(\mathbb{R}^d)$.

<u>Example ⑧</u>: A measurable function f on $\mathbb{R}^d$ satisfying

$$\int (1+|x|^2)^{-n} |g(x)|^p \, dx < \infty$$

for some $p \in [1,\infty)$ and $n>0$ generates a tempered distribution by the action

$$\phi \in \mathcal{S}(\mathbb{R}^d) \longmapsto \int f(x)\,\phi(x)\,dx.$$

<u>Example ⑨</u>: Special cases of Example ⑧ include all $\mathcal{L}^p(\mathbb{R}^d)$ functions and all polynomials.

<u>Note (26)</u>: We associate the weak* topology on $\mathcal{S}'(\mathbb{R}^d)$ – which is the topology of pointwise convergence on $\mathcal{S}(\mathbb{R}^d)$.

<u>Note (27)</u>: (i) If $T \in \mathcal{S}'(\mathbb{R}^d)$, then

$$T\big|_{\mathcal{D}(\mathbb{R}^d)} \in \mathcal{D}'(\mathbb{R}^d).$$

(ii) Using Proposition ②, any $T \in \mathcal{S}'(\mathbb{R}^d)$ is uniquely determined by $T\big|_{\mathcal{D}(\mathbb{R}^d)}$.

(iii) Combining (i) & (ii), we conclude that $\mathcal{S}'(\mathbb{R}^d)$ is the set of distributions that can be extended continuously from $\mathcal{D}(\mathbb{R}^d)$ to $\mathcal{S}(\mathbb{R}^d)$.

We now discuss operations on the space of tempered distributions.

I. (Addition and Scalar multiplication)

Note that $\mathcal{S}'(\mathbb{R}^d)$ is a vector space. Given $T, S \in \mathcal{S}'(\mathbb{R}^d)$ and scalars α, β, we have

$$\langle \alpha T + \beta S, \phi \rangle_{\mathcal{S}'(\mathbb{R}^d)\,\mathcal{S}(\mathbb{R}^d)} := \alpha \langle T, \phi \rangle + \beta \langle S, \phi \rangle \quad \forall \phi \in \mathcal{S}(\mathbb{R}^d).$$

II. (Multiplication by smooth functions)

It is, in general, not true that $\phi \in \mathcal{S}(\mathbb{R}^d)$, $\psi \in \mathcal{E}(\mathbb{R}^d) \Rightarrow \psi\phi \in \mathcal{S}(\mathbb{R}^d)$. To claim this implication, we in addition require that ψ is slowly increasing, i.e. ψ and all its derivatives have at most polynomial growth at infinity. Under this assumption, we can define multiplication by ψ as

$$\langle \psi T, \phi \rangle_{\mathcal{S}'(\mathbb{R}^d)\,\mathcal{S}(\mathbb{R}^d)} := \langle T, \psi\phi \rangle \quad \forall \phi \in \mathcal{S}(\mathbb{R}^d)$$

for any $T \in \mathcal{S}'(\mathbb{R}^d)$.

III. (Translation)

For any $x \in \mathbb{R}^d$ and $T \in \mathcal{S}'(\mathbb{R}^d)$, we have

$$ {}_{\mathcal{S}'(\mathbb{R}^d)}\langle \tau_x T, \phi \rangle_{\mathcal{S}(\mathbb{R}^d)} := \langle T, \tau_{-x} \phi \rangle \quad \forall \phi \in \mathcal{S}(\mathbb{R}^d). $$

IV. (Differentiation)

For any multi-index α and $T \in \mathcal{S}'(\mathbb{R}^d)$, we have

$$ {}_{\mathcal{S}'(\mathbb{R}^d)}\langle \partial^\alpha T, \phi \rangle_{\mathcal{S}(\mathbb{R}^d)} := (-1)^{|\alpha|} \langle T, \partial^\alpha \phi \rangle \quad \forall \phi \in \mathcal{S}(\mathbb{R}^d). $$

V. (convolution)

For any $\psi \in \mathcal{S}(\mathbb{R}^d)$ and $T \in \mathcal{S}'(\mathbb{R}^d)$, the function $T * \psi$ defined by

$$ T * \psi(x) := \langle T, \tau_x \tilde{\psi} \rangle, \quad x \in \mathbb{R}^d $$

is in $\mathcal{E}(\mathbb{R}^d)$ and slowly increasing. Hence it is a tempered distribution.

VI. (Fourier transform)

For $T \in \mathcal{S}'(\mathbb{R}^d)$, we define its Fourier transform $\hat{T} \in \mathcal{S}'(\mathbb{R}^d)$ by

$$ {}_{\mathcal{S}'(\mathbb{R}^d)}\langle \hat{T}, \phi \rangle_{\mathcal{S}(\mathbb{R}^d)} := \langle T, \hat{\phi} \rangle \quad \forall \phi \in \mathcal{S}(\mathbb{R}^d). $$

This definition can again be motivated by a similar duality action identity for Schwartz class functions. Here, Theorem 4 extends to the following result.

Theorem 5: The Fourier transform is a continuous, linear, bijective mapping on $\mathcal{S}'(\mathbb{R}^d)$, with continuous inverse.

Proposition 8: For $\psi \in \mathcal{S}(\mathbb{R}^d)$ and $T \in \mathcal{S}'(\mathbb{R}^d)$,

(i) $\partial^\alpha (T * \psi) = \partial^\alpha T * \psi = T * \partial^\alpha \psi$ for any multi-index α (see the convolution operation above)

(ii) $(T * \psi)^\wedge = \hat{T}\hat{\psi}$

(iii) $(\psi T)^\wedge = \hat{T} * \hat{\psi}$.

We now wish to look at classes of functions (and distributions) upto a certain order of smoothness - taken in terms of the distributional derivatives.

Note 28: Formally, consider the space of functions

$$H^n := \{ f \in \mathcal{L}^2(\mathbb{R}^d) \mid \partial^\alpha f \in \mathcal{L}^2(\mathbb{R}^d) \ \forall \alpha \text{ with } |\alpha| \leq n \}$$

for $n = 0, 1, 2, \ldots$

Equivalently, using the Fourier transform we may look at

$$H^n = \{ f \in \mathcal{L}^2(\mathbb{R}^d) \mid \xi \mapsto \xi^\alpha \hat{f}(\xi) \in \mathcal{L}^2(\mathbb{R}^d) \ \forall \alpha \text{ with } |\alpha| \leq n \}$$

for $n = 0, 1, \ldots$

where ξ^α denotes $\xi_1^{\alpha_1} \xi_2^{\alpha_2} \cdots \xi_d^{\alpha_d}$. The conditions in this description may be equivalently summarized as in the following definition.

Definition 4 (H^s for $s \in \mathbb{R}$)

$$H^s := \{ f \in \mathcal{S}'(\mathbb{R}^d) \mid \xi \mapsto (1+|\xi|^2)^{s/2} \hat{f}(\xi) \in \mathcal{L}^2(\mathbb{R}^d) \}.$$

Note 29: Here, we consider H^s as a subset of $\mathcal{S}'(\mathbb{R}^d)$ to take into account the "negative" regularities "s". Also observe that $\xi \mapsto (1+|\xi|^2)^{s/2}$ is slowly increasing and in $\mathcal{E}(\mathbb{R}^d)$ and, as such is a continuous linear operator on $\mathcal{S}'(\mathbb{R}^d)$. We also use the fact that $\mathcal{S}'(\mathbb{R}^d)$ remains invariant under the Fourier transform.

Notation 12: Consider the mapping

$$\langle \cdot, \cdot \rangle^{(s)} : H^s \times H^s \longrightarrow \mathbb{C}$$

defined by

$$\langle f, g \rangle^{(s)} = \int \left(1+|\xi|^2\right)^s \hat{f}(\xi)\, \overline{\hat{g}(\xi)}\, d\xi$$

$$\forall f, g \in H^s.$$

Proposition 9: (i) $\langle \cdot, \cdot \rangle^{(s)}$ is an inner product on H^s. Moreover, H^s becomes a Hilbert space with respect to this inner product.

(ii) $H^0 = \mathcal{L}^2(\mathbb{R}^d)$.

(iii) $\mathcal{S}(\mathbb{R}^d)$ is dense in H^s, $\forall s$.

(iv) For $s<t$, $\|\cdot\|^{(s)} \leq \|\cdot\|^{(t)}$ and H^t is a dense subspace of H^s.

(v) For any multi-index α,

$\partial^\alpha : H^s \longrightarrow H^{s-|\alpha|}$ is a bounded linear operator.

Theorem 6: Fix $f \in H^{-s}$ for some $s \in \mathbb{R}$ and consider the linear functional

$\phi \in \mathcal{S}(\mathbb{R}^d) \longmapsto \langle f, \phi \rangle_{\mathcal{S}'(\mathbb{R}^d)\ \mathcal{S}(\mathbb{R}^d)}$. This extends to a continuous linear functional on H^s and every element in $(H^s)^*$ is of this form. This identification between H^{-s} to $(H^s)^*$, which extends the duality between $\mathcal{S}'$ and $\mathcal{S}$, is a unitary isomorphism.

<u>Theorem ⑦</u>: (Sobolev Embedding)

If $s > k + \frac{d}{2}$, then $H^s \subset C_0^k$, where

$$C_0^k = \{ f \in C^k(\mathbb{R}^d) \mid \partial^\alpha f \in C_0 \ \forall \alpha \text{ with } |\alpha| \leq k \}.$$

Moreover, the inclusion map $H^s \hookrightarrow C_0^k$ is continuous.

<u>Theorem ⑧</u>: If $f \in \bigcap_{s \in \mathbb{R}} H^s$, then $f \in \mathcal{E}(\mathbb{R}^d)$.

<u>Note ㉚</u>: We have used the Schwartz space $\mathcal{S}(\mathbb{R}^d)$, the space of tempered distributions $\mathcal{S}'(\mathbb{R}^d)$ and the Fourier transforms — all of which requires the whole Euclidean space $\mathbb{R}^d$. To work with subsets Ω, we have to use $\mathcal{D}(\Omega)$ and $\mathcal{D}'(\Omega)$.

let Ω be an open set in $\mathbb{R}^d$.

Definition (11) ($W^{1,p}(\Omega)$ for $p \in [1,\infty]$)

$$W^{1,p}(\Omega) := \left\{ f \in \mathcal{L}^p(\Omega) \,\middle|\, \partial_j f \in \mathcal{L}^p(\Omega) \ \forall\, j=1,2,\dots,d \right\}$$

Also consider the norm

$$f \longmapsto \|f\|_{W^{1,p}(\Omega)} := \|f\|_{\mathcal{L}^p(\Omega)} + \sum_{j=1}^{d} \|\partial_j f\|_{\mathcal{L}^p(\Omega)}$$

Notation (13): Set $H^1(\Omega) := W^{1,2}(\Omega)$.

Exercise (2): The norm on $H^1(\Omega)$ arises from an inner product. Identify it.

Theorem (9): (i) $W^{1,p}(\Omega)$ is a Banach space. It is separable for $p \in [1,\infty)$

(ii) $H^1(\Omega)$ is a separable Hilbert space.

Note (31): $H^1(\mathbb{R}^d) = H^1$, as defined above.

Definition (12) ($W^{m,p}(\Omega)$ for $p \in [1,\infty]$ and $m = 2,3,\dots$)

$W^{m,p}(\Omega)$ is defined inductively as

$$W^{m,p}(\Omega) := \left\{ f \in W^{m-1,p}(\Omega) \,\middle|\, \partial_j f \in W^{m-1,p}(\Omega),\ j=1,2,\dots,d \right\}$$

Also consider the norm

$$f \longmapsto \|f\|_{W^{m,p}(\Omega)} := \sum_{|\alpha| \le m} \|\partial^\alpha f\|_{\mathcal{L}^p(\Omega)}$$

Notation (14): $H^m(\Omega) := W^{m,2}(\Omega)$.

Note (32): As mentioned in the case of $H^1(\Omega)$, $H^m(\Omega)$ also has a natural inner product.

Theorem (10): (i) $W^{m,p}(\Omega)$ is a Banach space.

(ii) $H^m(\Omega)$ is a Hilbert space.

Note (33): Under regularity conditions on Ω, in terms of the boundary (e.g. Ω is of class C^1), we can extend the functions in $W^{1,p}(\Omega)$ to $W^{1,p}(\mathbb{R}^d)$. This useful result helps us in simplifying many arguments involving $W^{1,p}(\Omega)$.

Theorem (11): Let Ω be of class C^1 and let $p \in [1, \infty)$. Then, the restrictions to Ω of functions in $\mathcal{D}(\mathbb{R}^d)$ is a dense subspace of $W^{1,p}(\Omega)$.

Note (34): In the interest of time, we skip the discussions on Sobolev Embedding Theorems for $W^{m,p}$.

Note (35): Due to the boundary behaviour of Ω, we can not directly consider the dual of $W^{1,p}(\Omega)$. In practice, we take the space $W^{-1,q}(\Omega)$, the dual space of $W_0^{1,p}(\Omega)$, for $1 \leq p < \infty$ where $\frac{1}{p} + \frac{1}{q} = 1$ (for $p \in (1,\infty)$) and $q = \infty$ (for $p = 1$) and, $W_0^{1,p}(\Omega)$ is the closure of $C_c^1(\Omega)$ in $W^{1,p}(\Omega)$.

Note (36): There are many notions of "$W^{s,p}(\Omega)$" for $s \in \mathbb{R}$. The area of fractional Sobolev spaces is an active area of research.