Tempered Distributions

Lionel Fiske and Cairn Overturf

May 19, 2016

In the classical study of partial differential equations one requires a solution to be differentiable. While intuitively this requirement seems necessary, in practice one often finds this to be too restrictive. In reality, initial data is often discontinuous, or may have cusps, therefore, it is natural to develop the notion of a non-smooth solution to a PDE. The theory of distributions was born out of these considerations. In this paper we will focus on a particularly useful type of distribution the *tempered distribution.*

Tempered distributions will allow us to give a definition for the derivative of non-smooth functions such as the Heaviside function, as well as help to make rigorous mathematical objects such as the dirac delta. Additionally, in this paper we will briefly discuss the Fourier transform, how it is related to tempered distributions, and its applications to solving PDEs.

1 Basic Definitions

To understand tempered distributions we must first understand a few basic definitions and spaces. To begin we need the notion of a function space. A *function space* is a topological space whose elements are maps from a common domain to a common codomain. These spaces often form structures such as vector spaces and metric spaces and prove to be invaluable in fields such as harmonic analysis and partial differential equations.

Example 1. Let $C(\mathbb{R})$ denote the set of all continuous functions from $\mathbb{R} \to \mathbb{R}$. $C(\mathbb{R})$ is a function space.

Throughout the course of this paper we will be considering one function space in particular, the space of *Schwartz functions.* Informally, these are infinitely differentiable functions whose derivatives decay faster than any polynomial at infinity. A property which allows many functions to be integrated over $\mathbb{R}$ when multiplied by a Schwartz function.

Definition 1. *(Schwartz functions)*
We denote the space

$$\mathcal{S}(\mathbb{R}) = \{f \in C^\infty(\mathbb{R}) : \|f\|_{\alpha,\beta} < \infty, \forall \alpha, \beta \in \mathbb{N}\} \text{ where } \|f\|_{\alpha,\beta} = \sup_{x\in\mathbb{R}} |x^\alpha f^{(\beta)}(x)|\} \ .$$

as the space of *Schwartz functions.*

The map $\|\cdot\|_{\alpha,\beta}$ is called a *semi-norm* as it has most of the properties of a norm, however this map is not a true norm as it is possible for non-zero vectors to map to zero. Another useful property of the Schwartz functions is the fact that they form a vector space structure under the standard operations of point-wise addition and scalar multiplication. This additional structure allows us to take advantage of many theorems of linear algebra, in particular, we can make sense of linear functions defined on $\mathcal{S}(\mathbb{R})$. Also note that $\mathcal{S}(\mathbb{R})$ is closed under differentiation and multiplication by polynomials and bounded $C^\infty(\mathbb{R})$ functions.

Example 2. Let $R > 0$, then any $C^\infty(\mathbb{R})$ function $\varphi(x)$ satisfying $\varphi(x) = 0$ for $|x| \geqslant R$ is Schwartz.

As with most spaces, these function spaces become more interesting when we define functions on them. We call the map f a *functional* if

$$f : \Omega \to \mathbb{R}, \text{ where } \Omega \text{ is a function space.}$$

Function spaces and functionals are the basis of the branch of mathematics known as *functional analysis*.

Example 3. Let $I : C(\mathbb{R}) \to \mathbb{R}$ be defined as $I[u] = \int_0^1 u(x)dx$. I is a functional.

We define the functional

$$d(f, g) = \sum_{\alpha,\beta\in\mathbb{N}} 2^{-\alpha-\beta} \frac{\|f - g\|_{\alpha,\beta}}{1 + \|f - g\|_{\alpha,\beta}}.$$

Clearly, d is well defined, because for any $a > 0$ we have $0 < \frac{a}{a+1} < 1$ and we see

$$0 < \sum_{n,m\in\mathbb{N}} 2^{-n-m}$$

is the product of two convergent geometric series. The space of Schwartz functions forms a metric space when equipped with the metric d.

Lemma 1. The space of functions $\mathcal{S}(\mathbb{R})$ equipped with the metric $d : \mathcal{S}(\mathbb{R}) \times \mathcal{S}(\mathbb{R}) \to \mathbb{R}$ is a metric space.

Proof. To prove this is a metric space we must show that d satisfies the metric axioms. Let $f, g, h \in \mathcal{S}(\mathbb{R})$ then:

1. *(Non-negativity)* We have already shown the metric axiom $d(f, g) \geqslant 0$ holds.

2. *(Identity of Indiscernibles)* Since all of the terms of the sum are non-negative, $d(f, g) = 0$ implies that every term in the sum must be zero. In particular, $\|f, g\|_{0,0} = \sup_{x\in\mathbb{R}} |f(x) - g(x)| = 0$ and thus $f(x) = g(x)$ for all $x \in \mathbb{R}$. Conversely, if $f = g$ then $\|f - g\|_{\alpha,\beta} = 0$ so all the terms of d are 0 and thus $d(f, g) = 0$.

3. *(Symmetry)* Since $\|f - g\|_{\alpha,\beta} = \sup_{x\in\mathbb{R}} |x^\alpha D^\beta(f - g)| = \sup_{x\in\mathbb{R}} |x^\alpha D^\beta(g - f)| = \|g - f\|_{\alpha,\beta}$, we have d is symmetric.

4. *(Triangle Inequality)* Since $\frac{x}{1+x}$ is an increasing function, for the α, β term of $d(f, h)$ we have:

$$\begin{aligned} \frac{\|f - h\|_{\alpha,\beta}}{1 + \|f - h\|_{\alpha,\beta}} &\leqslant \frac{\|f - g\|_{\alpha,\beta} + \|g - h\|_{\alpha,\beta}}{1 + \|f - g\|_{\alpha,\beta} + \|g - h\|_{\alpha,\beta}} \\ &\leqslant \frac{\|f - g\|_{\alpha,\beta}}{1 + \|f - g\|_{\alpha,\beta}} + \frac{\|g - h\|_{\alpha,\beta}}{1 + \|g - h\|_{\alpha,\beta}}, \end{aligned}$$

the last inequality follows since the denominators are less than or equal to $1+\|f-g\|_{\alpha,\beta}+\|g-h\|$. Notice the last two terms together are the α, β term of $d(f, g) + d(g, h)$. Since the α, β term of $d(f, h)$ is less than or equal to the corresponding term of $d(f, g)+d(g, h)$ for all α, β, we have $d(f, h) \leqslant d(f, g)+d(g, h)$. Therefore the triangle inequality will holds for d.

$\square$

This metric is constructed such that for any sequence of Schwartz functions $(f_n)_{n=0}^{\infty}$ we have that f_n converges to $f \in \mathcal{S}(\mathbb{R})$ with respect to the metric d if and only if $\lim_{n\to\infty} \|f_n - f\|_{\alpha,\beta} = 0$ for any $\alpha, \beta \in \mathbb{N}$. Thus, showing convergence in $\mathcal{S}(\mathbb{R})$ it suffices to do computations with $\|\cdot\|_{\alpha,\beta}$. Furthermore, when proving properties about tempered distributions we take advantage of the rapid decay of Schwartz functions to move limits inside integrals over $\mathbb{R}$. We can do this because on any finite compact interval we have convergence in the *sup norm* metric which implies uniform convergence on that interval. Furthermore, since Schwartz functions decay so rapidly we know for any two Schwartz functions beyond a certain interval they will be $\frac{\varepsilon}{2}$ close to zero for any $\varepsilon > 0$ meaning they are at most ε close to each other. Thus, given a convergent sequence of functions it is possible to take a large enough interval (on which the sequence converges uniformly) so that outside the functions are so close together that they also converge uniformly. We omit a rigorous proof for brevity.

Now we come to the main definition, that of a tempered distribution.

Definition 2. *(Tempered Distributions)*
Let $T : \mathcal{S}(\mathbb{R}) \to \mathbb{R}$ be a functional. We say T is a *tempered distribution* if it is both linear and continuous.

We know that $\mathcal{S}(\mathbb{R})$ is a vector space, consequently we know from topology that the space of all continuous linear functionals on $S(\mathbb{R})$ is closed under the standard operations of addition and scalar multiplication. We adopt the notation $S'(\mathbb{R})$ for the space of tempered distributions. We remark that in general the set of linear and continuous functionals on a vector space is called the continuous dual of the space.

Example 4. Consider the object δ, defined for $\varphi \in S(\mathbb{R})$ by $\delta[\varphi] = \varphi(0)$. We claim that this object is a tempered distribution . Let $\varphi, \psi \in \mathcal{S}(\mathbb{R})$ and $c \in \mathbb{R}$. To check for linearity notice

$$\delta[\varphi + \psi] = (\varphi + \psi)(0) = \varphi(0) + \psi(0) = \delta[\varphi] + \delta[\psi].$$

For homogeneity we have

$$\delta[c\varphi] = (c\varphi)(0) = c \cdot \varphi(0) = c\delta[\varphi].$$

Continuity also follows easily. Let $(\varphi_n)_{n=0}^{\infty}$ be an arbitrary sequence of Schwartz functions that converges to φ. Since the sequence of functions converges with respect to d we know that φ_n converges uniformly to $\varphi \in \mathcal{S}(\mathbb{R})$ because it must converge with respect to the $\|\cdot\|_{0,0}$ norm. Thus, $\delta[\varphi_n] = \varphi_n(0)$ also converges to $\varphi(0)$ as $n \to \infty$. As δ is a linear continuous functional defined on $\mathcal{S}(\mathbb{R})$ it is a tempered distribution.

Example 5. We claim the function $I : S(\mathbb{R}) \to \mathbb{R}$ defined by the Riemann integral

$$I[\varphi] = \int_0^1 \varphi(x)dx$$

is a tempered distribution. Since Schwartz functions are $C^\infty(\mathbb{R})$ they are continuous and so I is well defined. Clearly I is also linear by the properties of the Riemann integral. To show I is continuous, let $(\varphi_n)_{n=m}^{\infty}$ be an arbitrary sequence of Schwartz functions such that $\varphi_n \to \varphi \in S(\mathbb{R})$ with respect to $\|\cdot\|_{\alpha,\beta}$. Then $\varphi_n \to \varphi$ uniformly and consequently

$$\lim_{n\to\infty} I[\varphi_n] = \lim_{n\to\infty} \int_0^1 \varphi_n(x)dx = \int_0^1 \lim_{n\to\infty} \varphi_n(x)dx = \int_0^1 \varphi(x)dx = I[\varphi].$$

Therefore, I is a tempered distribution.

Example 6. Let f be a function such that $f\varphi$ is integrable on $\mathbb{R}$ for all $\varphi \in \mathcal{S}(\mathbb{R})$. Then we denote the tempered distribution induced by f as $T_f : \mathcal{S}(\mathbb{R}) \to \mathbb{R}$ and define it as

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

The proof that T_f is indeed a tempered distribution is similar to Example 5.

2 Operations on Tempered Distributions

Differentiation

Next we turn our focus to defining a notion of differentiation for tempered distributions. To motivate this definition we consider a tempered distribution of the form (6)

$$T_f'[\varphi] = \int_{\mathbb{R}} f'(x)\varphi(x)dx$$

where f' denotes the derivative of f. Integration by parts yields

$$T_{f'}[\varphi] = \int_{\mathbb{R}} f'(x)\varphi(x)dx = f(x)\varphi(x)\Big|_{-\infty}^{\infty} - \int_{\mathbb{R}} f(x)\varphi'(x)dx = -\int_{\mathbb{R}} f(x)\varphi'(x)dx = -T_f[\varphi'].$$

The boundary terms vanish due to the integrability of $f\varphi$ over $\mathbb{R}$. Keeping this form in mind we define the distributional derivative.

Definition 3. *(Distributional Derivative)*
Let $T \in \mathcal{S}'(\mathbb{R})$. We define the distributional derivative of T.

$$T'[\phi] = -T[\varphi'].$$

Lemma 2. Let $T \in S'(\mathbb{R})$ then T' is a tempered distribution.

Proof. Let $\varphi, \psi \in S(\mathbb{R})$, then we have linearity since

$$T'[\varphi + \psi] = -T[(\varphi + \psi)'] = -T[\varphi' + \psi'] = -T[\varphi'] - T[\psi'] = T'[\varphi] + T'[\psi].$$

Now we will show continuity. Let φ_n be a sequence of Schwartz functions that converges to φ with respect to the metric d. Then by definition of d, $\|\varphi_n - \varphi\|_{\alpha,\beta} = \|\varphi_n' - \varphi'\|_{\alpha,\beta-1}$ converges to 0 for all $\alpha, \beta \geqslant 1$. Consequently φ_n' converges to φ' with respect to d. Thus

$$\lim_{n\to\infty} T'[\varphi_n] = \lim_{n\to\infty} -T[\varphi_n'] = -T[\varphi'] = T'[\varphi]$$

since T is continuous. Therefore T' is continuous and a tempered distribution. ☐

This definition allows us to make rigorous many common objects used in the study of differential equations and mathematical physics. Two of the most common examples of objects which are related through the distributional derivative are the Heaviside function and the delta distribution.

Example 7. Let the Heaviside function, $\theta(x) : \mathbb{R} \to \mathbb{R}$ be defined as

$$\theta(x) = \begin{cases} 0 \text{ if } x \in (-\infty, 0) \\ 1 \text{ if } x \in [0, \infty) \end{cases}.$$

This induces the distribution

$$H[\varphi] = \int_{\mathbb{R}} \theta(x)\varphi(x)dx.$$

Now we compute the derivative as:

$$H'[\varphi] = -\int_{\mathbb{R}} \theta(x)\varphi'(x)dx = -\int_0^\infty \varphi'(x)dx = -\varphi(x)\big|_0^\infty = \varphi(0) = \delta[\varphi].$$

Thus, in the sense of distributions the derivative of the Heaviside is the delta distribution.

Example 8. Let $\psi \in S(\mathbb{R})$ consider the distribution differential equation

$$\Lambda' = T_\psi.$$

Let $\Psi(x) = \int_{x_0}^x \psi(\xi)d\xi$. The distribution

$$\Lambda[\varphi] = T_\psi[\varphi]$$

is a solution to the differential equation because

$$\Lambda'[\varphi] = -\int_{\mathbb{R}} \Psi(x)\varphi'(x)dx = -\Psi(x)\varphi(x)\big|_{\mathbb{R}} + \int_{\mathbb{R}} \psi(x)\varphi(x)dx = \int_{\mathbb{R}} \psi(x)\varphi(x)dx = T_\psi[\varphi].$$

Fourier Transform

One useful operation defined on the Schwartz functions is the Fourier transform. This function can be thought of as the continuous analogue to the Fourier series.

Definition 4. *(Fourier transform)*
Let $\varphi \in \mathcal{S}(\mathbb{R})$. We define the function $\mathcal{F} : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ as

$$\mathcal{F}(\varphi)(y) = \widehat{\varphi}(y) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \varphi(x)e^{-ixy}dx$$

and denote $\widehat{\varphi}$ as the *Fourier transform* of φ.

We remark that the Fourier transform of a Schwartz function is always a Schwartz function. The proof for this fact is nontrivial but we omit it here for brevity. Stronger still, the Fourier transform is also a bijection from the Schwartz functions to themselves. Therefore, the Fourier transform has an inverse which we denote by $\check{f}$.

Theorem 1. (Fourier Inversion Theorem) Let φ be a Schwartz function then the inverse of the Fourier transform is

$$\mathcal{F}^{-1}(\varphi)(x) = \check{\varphi}(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \varphi(y)e^{ixy}dy \tag{1}$$

Proof. Let φ and ψ be Schwartz functions such that

$$\psi(0) = \int_R \widehat{\psi}(x)dx = 1,$$

Then we have, by interchanging the limit and integral

$$\lim_{\varepsilon\to 0} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \psi(\varepsilon y)\widehat{\varphi}(y)e^{ixy}dy = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{\varphi}(y)e^{ixy}dy.$$

Moving the limit inside the integral is valid since ψ is Schwartz. Furthermore, by Fubini's theorem we can interchange the order of integration to obtain

$$\begin{aligned}\lim_{\varepsilon\to 0}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\psi(\varepsilon y)\widehat{\varphi}(y)e^{ixy}dy &= \lim_{\varepsilon\to 0}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\psi(\varepsilon y)e^{ixy}\left(\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\varphi(t)e^{-iyt}dt\right)dy.\\ &= \lim_{\varepsilon\to 0}\int_{\mathbb{R}}\varphi(t)\left(\int_{\mathbb{R}}\psi(\varepsilon y)e^{-iy(t-x)}dy\right)dt\end{aligned}$$

After the change of variables $u = \varepsilon y$ the last integral becomes

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}}\varphi(t)\left(\int_{\mathbb{R}}\psi(u)e^{-i\frac{u}{\varepsilon}(t-x)}\frac{du}{\varepsilon}\right)dt = \lim_{\varepsilon\to 0}\int_{\mathbb{R}}\varphi(t)\widehat{\psi}\left(\frac{t-x}{\varepsilon}\right)\frac{dt}{\varepsilon}.$$

Performing another change of variables $v = (t-x)\varepsilon$ and interchanging the limit and the integral we have $\varphi(x)$ is equal to

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}}\varphi(x+\varepsilon v)\widehat{\psi}(v)dv = \varphi(x)\int_{\mathbb{R}}\widehat{\psi}(v)dv = \varphi(x),$$

the limit passes to the argument of φ since φ is continuous. Thus the Fourier transform is invertible with inverse given by (1).

□

Example 9. Let $G : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ be defined as $G(x) = e^{-x^2/2}$. The Fourier transform of G is given by

$$\begin{aligned}\widehat{G}(y) =& \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}e^{-x^2/2}e^{-ixy}dx \qquad (2)\\ =& \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-(x^2/2+ixy)/2}dx\\ =& \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-(x+iy)^2/2-y^2/2}dx\\ =& \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-(x+iy)^2/2}dx\end{aligned}$$

If we let $z = x + iy$, then we obtain

$$\begin{aligned}\widehat{G}(y) &= \frac{e^{-y^2/2}}{\sqrt{2\pi}}\int_{-\infty+iy}^{\infty+iy}e^{-z^2/2}dz\\ &= \frac{e^{-y^2/2}}{\sqrt{2\pi}}\cdot\sqrt{2\pi} = e^{-y^2/2}\end{aligned}$$

Therefore, $\widehat{G}(y) = G(y)$.

We can generalize this notion of the Fourier transform to tempered distributions by defining

$$\mathfrak{F}(T)[\varphi] = \widehat{T}[\varphi] = T[\widehat{\varphi}].$$

As with the Fourier transform of Schwartz functions, it can be shown that if $T \in S'(\mathbb{R})$ then $\mathfrak{F}(T) \in S'(\mathbb{R})$ and furthermore, $\mathfrak{F}$ is an isomorphism with inverse given by

$$\mathfrak{F}^{-1}(T)[\varphi] = \check{T}[\varphi] = T[\check{\varphi}],$$

for $\varphi \in S(\mathbb{R})$.

Example 10. Let δ denote the delta distribution and $\varphi \in \mathcal{S}(\mathbb{R})$. Consider

$$\begin{aligned}\widehat{\delta}[\varphi] &= \delta[\widehat{\varphi}] \\ &= \delta\left[\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \varphi(x)e^{-ixy}dx\right] \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \varphi(x)e^{0}dx \\ &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\varphi(x)dx = T_{\frac{1}{\sqrt{2\pi}}}[\varphi].\end{aligned}$$

Thus, the Fourier transform of the delta distribution is the tempered distribution induced by $\dfrac{1}{\sqrt{2\pi}}$. Unfortunately we do not obtain the nice identity $\widehat{\delta} = T_1$ with our normalization of the Fourier transform, however the convention we use is more convenient for solving PDEs.

One of the primary reason people work with Tempered distributions is the fact that the Fourier transform is an *isomorphism* on the vector space of Tempered distributions. This is to say that the Fourier transform is a function which maps $\mathcal{S}'(\mathbb{R})$ to $\mathcal{S}'(\mathbb{R})$, is linear, and is invertible.

Lemma 3. The Fourier transform is an isomorphism on $\mathcal{S}'(\mathbb{R})$.

Proof. We have that the Fourier transform is a bijection from $\mathcal{S}'(\mathbb{R}) \to \mathcal{S}'(\mathbb{R})$. Now we need to show linearity, let T_1 and T_2 be tempered distributions then

$$(\widehat{T_1 + T_2})[\varphi] = (T_1 + T_2)[\widehat{\varphi}] = T_1[\widehat{\varphi}] + T_2[\widehat{\varphi}] = \widehat{T}_1[\varphi] + \widehat{T}_2[\varphi].$$

Therefore $(\widehat{T_1 + T_2}) = \widehat{T}_1 + \widehat{T}_2$.

□

Convolutions

The last operation on tempered distributions we would like to discuss is convolutions. A convolution is a binary operation which returns another function which is often viewed as a modification to one of the original functions. These operations are used in Fourier analysis, partial differential equations, in the study of signals and systems, as well as many branches of engineering.

Definition 5. *Convolution of Schwartz functions*
Let f, g be functions. We define their convolution as

$$f * g(x) = \int_{-\infty}^{\infty} f(x-y)g(y)dy = g * f(x).$$

In order for the convolution to be well defined we need $f(x-y)g(y)$ to be integrable over $\mathbb{R}$ for all x, clearly this holds if f and g are Schwartz. The commutativity of the convolution follows from a simple change of variables. As with the previous operations, to define the convolution of a Schwartz function with a tempered distribution we pass the convolution to the argument. Precisely, if $\psi, \phi \in S(\mathbb{R})$ and $T \in S'(\mathbb{R})$ then the convolution of ψ and T is a distribution and acts on ϕ as

$$\psi * T[\phi] = T[\widetilde{\psi} * \varphi]$$

where $\widetilde{\psi}(x) = \psi(-x)$ is the reflection about 0. For this definition we use the fact that convolution of Schwartz functions is Schwartz but do not prove it.

Example 11. This example will show the delta distribution acts as an identity for the convolution operation. Let $\psi, \phi \in S(\mathbb{R})$ then

$$\begin{aligned}\psi * \delta[\varphi] &= \delta[\widetilde{\psi} * \varphi] \\ &= \delta\left[\int_{-\infty}^{\infty} \widetilde{\psi}(x-y)\varphi(y)dy\right] \\ &= \delta\left[\int_{-\infty}^{\infty} \psi(y-x)\varphi(y)dy\right] \\ &= \int_{-\infty}^{\infty} \psi(y)\varphi(y)dy = T_\psi[\varphi].\end{aligned}$$

Therefore we write $\psi * \delta = \psi$.

Theorem 2. (Derivative of a Convolution) Let $\varphi, \psi \in \mathcal{S}(\mathbb{R})$. We have for any $\alpha \in \mathbb{N}$

$$D^\alpha\big(\varphi * \psi\big)(x) = \big((D^\alpha\varphi) * \psi\big)(x) = \big(\varphi * (D^\alpha\psi)\big)(x)$$

Proof. Consider,

$$\begin{aligned}D^\alpha(\varphi * \psi)(x) = D^\alpha \int_{\mathbb{R}} \varphi(x-y)\psi(y)dy &= \int_{\mathbb{R}} D^\alpha\varphi(x-y)\psi(y)dy \\ &= \int_{\mathbb{R}} (D^\alpha\varphi(x-y))\psi(y)dy = \big((D^\alpha\varphi) * \psi\big)(x). \qquad \text{(Leibniz' rule)}\end{aligned}$$

The interchange of derivatives and integral above is justified because φ and ψ are in the Schwartz class. We also have the convolution operator is commutative so without loss of generality

$$D^\alpha(\varphi * \psi)(x) = D^\alpha(\psi * \varphi)(x) = (D^\alpha\psi * \varphi)(x)$$

From this the theorem follows. □

3 Applications to PDE

To make distributions useful to the theory of partial differential equations we must extend the Schwartz space to $\mathbb{R}^n$. Before we proceed we will need the notion of a multi-index. An *n dimensional mult-index* α is an n-tuple $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N} \times \dots \mathbb{N} = \mathbb{N}^n$, and the *order* of α is $|\alpha_1 + \dots + \alpha_n|$. Let α be an n dimensional multi-index and $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, then we define $x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}$, and for a sufficiently differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, we define

$$f^{(\alpha)}(x) = \frac{\partial^{|\alpha|} f(x)}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$$

Now we are ready to define $\mathcal{S}(\mathbb{R}^n)$

$$\mathcal{S}(\mathbb{R}^n) = \{f \in C^\infty(\mathbb{R}^n) : \|f\|_{\alpha,\beta} < \infty, \forall \alpha, \beta \in \mathbb{N}^n\} \text{ where } \|f\|_{\alpha,\beta} = \sup_{x \in \mathbb{R}^n} |x^\alpha f^{(\beta)}(x)|.$$

Notice that this definition coincides with the definition given at the beginning of Section 1 when $n = 1$.

There is a corresponding version of Lemma 1 for $\mathcal{S}(\mathbb{R}^n)$ but we will omit it. Therefore $S(\mathbb{R}^n)$ is a topological vector space with continuous dual space $S'(\mathbb{R}^n)$ which we will call the tempered distributions. This name is justified since this definition also coincides with our previous definition of tempered distributions when $n = 1$. Before we look at any PDEs we need to extend differentiation to $S'(\mathbb{R}^n)$. Let $\alpha \in \mathbb{N}^n$ be a multi-index and $T \in S'(\mathbb{R}^n)$ then we define $T^{(\alpha)}[\phi] = (-1)^{|\alpha|}T[\varphi^\alpha]$.

Example 12. For this example we seek a tempered distribution $U \in S'(\mathbb{R}^2)$ which is a solution the partial differential equation

$$U_{tt} - U_{xx} = 0. \qquad \text{(wave equation)}$$

We say U is a solution of the wave equation if

$$\begin{aligned} U_{tt}[\varphi] &= U_{xx}[\varphi] \\ \implies U[\varphi_{tt} - \varphi_{xx}] &= 0, \text{ for all } \varphi \in S(\mathbb{R}^2). \end{aligned}$$

Let us assume U is a tempered distribution induced by a function $u : \mathbb{R}^2 \to \mathbb{R}$, then the previous equation implies

$$\int_{\mathbb{R}^2} u(x,t)(\varphi_{tt}(x,t) - \varphi_{xx}(x,t))dxdt = 0, \text{ for all } \varphi \in \mathcal{S}'(\mathbb{R}). \qquad (3)$$

We claim $u = F(x-t) + F(x+t)$ for any bounded function $F : \mathbb{R}^2 \to \mathbb{R}$ solves (3). Substituting, we obtain

$$\int_{\mathbb{R}^2} (F(x-t) + F(x+t))(\varphi_{tt}(x,t) - \varphi_{xx}(x,t))dxdt$$

To show that this is identically zero we will use a trick which exploits the fact that the delta distribution is the identity under convolutions. First we need to introduce a mollifier χ which satisfies

$$\begin{aligned} &(i) \ \ \chi \in C_0^\infty(\mathbb{R}) \subset \mathcal{S}(\mathbb{R}^n) \\ &(ii) \int_{\mathbb{R}^n} \chi(x)dx = 1 \\ &(iii) \lim_{\epsilon\to\infty} \int_{\mathbb{R}^n} \frac{1}{\varepsilon^2}\chi\left(\frac{x}{\epsilon^2}\right)\varphi(x)dx = \delta[\varphi]. \end{aligned}$$

For brevity we define $\chi_\epsilon(x) = \chi\left(\frac{x}{\epsilon}\right)$. We note that for even a non-differentiable function f the convolution, $f^\epsilon = \chi_\epsilon * f$ is smooth because

$$D^\alpha(\chi_\epsilon * f) = (D^\alpha\chi_\epsilon) * f.$$

Thus $u^\epsilon(x,t) = F^\epsilon(x-t) + F^\epsilon(x+t)$ is a classical solution of the wave equation and therefore

$$0 = \lim_{\epsilon\to0} \int_{\mathbb{R}^2} \varphi(x,t)(u^\epsilon_{tt}(x,t) - u^\epsilon_{xx}(x,t))dxdt.$$

Moving the limit inside the integral we get

$$\begin{aligned} 0 &= \lim_{\epsilon\to0} \int_{\mathbb{R}^2} \varphi(x,t)(u^\epsilon_{tt}(x,t) - u^\epsilon_{xx}(x,t))dxdt \\ &= \lim_{\epsilon\to0} \int_{\mathbb{R}^2} u^\epsilon(x,t)(\varphi_{tt}(x,t) - \varphi_{xx}(x,t))dxdt \\ &= \int_{\mathbb{R}^2} \lim_{\epsilon\to0} u^\epsilon(x,t)(\varphi_{tt}(x,t) - \varphi_{xx}(x,t))dxdt \\ &= \int_{\mathbb{R}^2} u(x,t)(\varphi_{tt}(x,t) - \varphi_{xx}(x,t))dxdt, \end{aligned}$$

the second equality follows from performing integration by parts twice and the third equality holds because the delta distributions is the identity under convolution. Thus, we have that U satisfies the differential equation.

Using this example we can define non-smooth solutions to the wave equation. Consider the Weierstrass function

$$W(x) = \sum_{k=0}^{\infty} b^{-k\alpha} \cos(b^k \pi x)$$

where $\alpha = \frac{-ln(a)}{ln(b)}$, $0 < a < 1$ and b is defined so that $ab < 1 + \frac{3\pi}{2}$. One can verify that this function is *differentiable nowhere* on its domain. However, in the distributional sense, by example (12) the function

$$u(x,t) = \frac{W(x-t) + W(x+t)}{2}$$

is a solution to the wave equation with initial data

$$\begin{cases} u(x,0) = W(x) & \forall x \in \mathbb{R} \\ u_t(x,0) = 0 & \forall x \in \mathbb{R}. \end{cases}$$

In this way we have defined a solution to the wave equation which is completely non-smooth at $t = 0$ (and for all but countably many values of t after that).

4 Conclusion

In this paper we defined a space of functions, the *Schwartz functions*, and studied both that space as well as it continuous dual space. We then defined some operations on Schwartz functions and tempered distributions and showed how these objects can be used to allow non-smooth solutions to partial differential equations.

We intended this paper to give a brief overview of both the Schwartz functions and tempered distributions. Distributions are foundational to many modern branches of mathematics and for answering age old question in many other branches. These functions have many desirable integration and differentiation properties which we can be exploited to solve problems and answer questions in a wide variety of pure and applied mathematics.

References

Pereyra, Maria Cristina., and Lesley A. Ward. Harmonic Analysis: From Fourier to Wavelets. Providence, RI: American Mathematical Society, 2012. Print.

Lorenz, Jens. Lecture Notes. Dec.-Jan. 2015. Math 583. University of New Mexico, Albuquerque.

Evans, Lawrence C. Partial Differential Equations. Providence, RI: American Mathematical Society, 1998. Print.

John, Fritz. Partial Differential Equations. New York: Springer-Verlag, 1982. Print.