L4: Distributions.

Contents:

• Distributions or generalized functions.
• Approximations to the identity.
• Calculus with distributions.
• The Schwartz class and tempered distributions.

1 Distributions or generalized functions.

We start with the following synthesis of the introduction of the classical book by Gelfand-Shilov [GeSh64, page 1].

Physicist (and engineers?) have long been using 'singular functions'. The simplest example is the delta function \( \delta(x - x_0) \).

These 'singular functions' occur as a rule only in intermediate stages of a solution of a problem and if some of them occurs in the final solution it is only in an integrand where it is multiplied by some other sufficiently good function.

Thus given a 'singular function' we know the result of its integration against a good function \( \psi \):

\[
\int_{\mathbb{R}} \delta(x - x_0) \varphi(x) \, dx = \varphi(x_0)
\]

This allows us to introduce this 'singular functions' as linear functionals over the space of good functions.

Let \( V \) be a vector space over \( \mathbb{R} \). Recall that the dual \( V^* \) is by definition the set of functions

\[
\lambda : V \rightarrow \mathbb{R}
\]

such that \( \lambda(av + bw) = a\lambda(v) + b\lambda(w) \) for all \( a, b \in \mathbb{R} \) and \( v, w \in V \).

It is also usual to write \( \lambda(v) = \langle \lambda, v \rangle \).

Let \( D \) be the vector space \( C_0^\infty \) of all \( C_0^\infty \) of functions \( \varphi : \mathbb{R} \rightarrow \mathbb{R} \) with compact support. That is to say, \( \varphi \in C_0^\infty \) if the set \( \text{supp}(\varphi) := \{x : \varphi(x) \neq 0\} \) is compact and \( \varphi \)
1.1 Functions as distributions

is infinitely derivable.

We say that the sequence \((\varphi_n), \varphi_n \in \mathcal{D}\) *converges to zero* if all this functions vanish outside a fixed interval \(I = [a, b]\), the same for all \(\varphi_n\) and converge to uniformly to zero together with their derivatives of any order.

We say that a linear functional \(L \in \mathcal{D}^*\) is continuous if \(L(\varphi_n) \to 0\) if \(\varphi_n\) converges to zero.

The set of continuous linear functionals is denoted by \(\mathcal{D}'\) and its elements are called *distributions* or *generalized functions*.

**Example 1.1.** Let \(\delta : \mathcal{D} \to \mathbb{R}\) be defined by

\[
\langle \delta, \varphi \rangle := \varphi(0).
\]

Then \(\delta\) is continuous and so \(\delta \in \mathcal{D}'\).

In the literature we will find also the integral notation. Namely, if \(L \in \mathcal{D}'\) then:

\[
L(\varphi) = \int_{\mathbb{R}} L(x) \varphi(x) dx = \int_{-\infty}^{+\infty} L(x) \varphi(x) dx = \langle L, \varphi \rangle
\]

1.1 Functions as distributions

Let \(f : \mathbb{R} \to \mathbb{R}\) be a ordinary function. The integral

\[
\langle L_f, \varphi \rangle = \int_{\mathbb{R}} f(x) \varphi(x) dx
\]

gives rise a distribution \(^1\) denoted by \(L_f\). This is the reason why distributions are called *generalized functions*.

**Warning** In many books the associated distribution \(L_f\) is also denoted by \(f\).

Notice that the \(\delta\) is not given by any function.

Here is another distribution \(PV\) which is not given by any function and which is also different from the \(\delta\) (i.e. \(PV^2\) do not comes from a mesure):

\(^1\) Under suitable conditions on \(f\), e.g. if \(\int_I |f(x)| dx < K\) for any interval \(I\).

\(^2\) PV comes from Principal Value. Actually, PV is the principal value of \(\frac{1}{x}\).
\begin{equation}
\langle PV, \varphi \rangle = \lim_{a \to 0} \int_{|\xi| > a} \frac{\varphi(\xi)}{\xi} d\xi
\end{equation}

This distribution is related to the so called Hilbert transform\(^3\).

### 1.2 Linear combinations of distributions

As in an abstract vector space we do not define a multiplication between distributions. But we do define the multiplication between \(C^\infty\) functions and distributions.

If \(f\) is a \(C^\infty\) function and \(L\) is a distribution we define its multiplication \(f.L \in \mathcal{D}'\) as follows

\[ \langle f.L, \varphi \rangle = \langle L, f\varphi \rangle \]

**Example 1.2.**

\[ f.\delta = f(0)\delta \]

Since we can add distributions we can construct linear combinations

\[ f_1L_1 + f_2L_2 + \cdots + f_kL_k \]

of distributions with coefficients in the ring of \(C^\infty\) functions. This is an example of what mathematician call a module over \(C^\infty(\mathbb{R})\).

**Warning:** Notice that for \(f(x) = x\) we have \(f.\delta = 0\) but neither \(f\) nor \(\delta\) are zero. This is different from what happens in a vector space, i.e. in a vector space \(r \overrightarrow{v} = \overrightarrow{0}\) imply either \(r = 0\) or \(\overrightarrow{v} = \overrightarrow{0}\). This is one of the main difference between modules and vector spaces.

**Proposition 1.3.** If \(L \in \mathcal{D}'\) satisfies \(x.L = 0\) then there exists a constant \(c\) such that \(L = c\delta\).

**Proof.** Fix \(\psi \in \mathcal{D}\) a function such that \(\psi(0) = 1\). We claim that \(L - c\delta = 0\) where \(c = \langle L, \psi \rangle\).

Indeed, any test \(\varphi\) can be written as

\[ \varphi = \varphi(0)\psi + x\overline{\varphi(x)} \]

where \(\overline{\varphi(x)} \in \mathcal{D}\)

\(^3\)http://en.wikipedia.org/wiki/Hilbert_transform
Then
\[
\langle L - c\delta, \varphi \rangle = \langle L - c\delta, \varphi(0)\psi + x\hat{\varphi}(x) \rangle \\
= \langle L - c\delta, \varphi(0)\psi \rangle + \langle L - c\delta, x\hat{\varphi}(x) \rangle \\
= \langle L - c\delta, \varphi(0)\psi \rangle \\
= \langle -c\delta, \varphi(0)\psi \rangle + \langle L, \varphi(0)\psi \rangle \\
= -c\varphi(0) + \varphi(0)\langle L, \psi \rangle \\
= 0
\]

This shows \( L - c\delta = 0 \) that is to say that \( L \) is a multiple of \( \delta \). \( \square \)

It is possible to prove a generalization of the above result for \( C^\infty \)-functions with simple roots. If \( \alpha \) is a root of \( C^\infty \)-function \( f(x) \) we say that it is a simple root if \( x - \alpha \frac{\hat{f}(x)}{f(x)} \) is \( C^\infty \) near \( \alpha \).

**Proposition 1.4.** Let \( f \in C^\infty \) be a function with simple roots and let \( Z \) be the set of zeros of \( f \). If \( L \in \mathcal{D}' \) satisfies \( f(x).L = 0 \) then
\[
L = \sum_{\alpha \in Z} c_\alpha \delta_\alpha
\]

The above is false without the hypothesis on the simple zeroes. Indeed, we will see that the solutions of \( x^2.L = 0 \) are more distributions than the multiples of the a single delta, i.e. \( L = c\delta + d\delta' \).
1.3 The Heaviside step function $H$.

The function $H$ is used in the mathematics of control theory and signal processing to represent a signal that switches on at a specified time and stays switched on indefinitely. It was named after the English polymath Oliver Heaviside\(^4\) (1850-1925).

To be precise $H \in \mathcal{D}'$ and acts on tests as follows:

$$\langle H, \varphi \rangle = \int_{0}^{\infty} \varphi(x)dx$$

The word 'step' comes from the fact that $H$ is the distribution associated to the step function $H(x)$ defined as follows:

$$H(x) = \begin{cases} 
0 & \text{if } x < 0, \\
1 & \text{if } x \geq 0. 
\end{cases}$$

That is to say

$$\langle H, \varphi \rangle = \int_{\mathbb{R}} H(x) \varphi(x)dx$$

We can translate the Heaviside distribution at $x_0$ and usually people write $H(x - x_0)$ to indicate such a translation.

A rectangular pulse of magnitude $a$ of duration $t$ is given by a function whose value is $a$ in an interval of length $t$. Such rectangular pulse $R_{a,t}(x)$ can be obtained by using the Heaviside function as follows:

$$R_{a,t}(x) = aH(x) - aH(x - t)$$

It is also called rectangle function, the gate function, pulse function, or window function.

It is useful to denote $\text{rec}(x) = H(x + \frac{1}{2}) - H(x + \frac{1}{2} - 1)$ the special case of rectangular pulse of magnitude and duration 1 starting at $x = -\frac{1}{2}$.

1.4 Translations and scaling

If $L$ is a distribution and $a \in \mathbb{R}$ we denote by $L_a$ its translation by $a$ defined as

$$\langle L_a, \varphi \rangle = \langle L, \tau_a \varphi \rangle$$

\(^4\)http://www-history.mcs.st-andrews.ac.uk/Biographies/Heaviside.html
where \( \tau_a \varphi(x) = \varphi(x + a) \).

**Warning:** Recall that \( \tau_a \) shift the picture to the \( a \) units to the left. Moreover, if \( L_f \) is the distribution associated to \( f \) then

\[
(L_f)_a = L_{\tau_a f}
\]

The scaling of \( s_r L \) of a distribution is defined as

\[
\langle s_r L, \varphi \rangle = \langle L, \frac{s_r \varphi}{r} \rangle
\]

where \( s_r \varphi(x) = \varphi\left(\frac{x}{r}\right) \).

The motivation was of course given by the following:

\[
L_{s_r f} = s_r L_f
\]


\section{Approximation to the identity}

\subsection{Weak convergence}

Let \( L_n \in \mathcal{D}' \) be a sequence of distributions. We say that \( L_n \) converge weakly to \( L \in \mathcal{D}' \) if for all test \( \varphi \in \mathcal{D} \)

\[
\lim_{n \to \infty} \langle L_n, \varphi \rangle = \langle L, \varphi \rangle
\]

\textbf{Example 2.1.} Let \( \varphi_n(x) := \varphi(x-n) \) be the sequence of functions obtained by translating a test function \( \varphi \). Then the distributions \( \varphi_n \) converge weakly to the zero distribution.

\textbf{Example 2.2.} Let \( f_n(x) := H(x) - H(x-n) \) be a sequence of functions. Then the associated distributions \( L_{f_n} \) converge weakly to the Heaviside function \( H \).

The properties of weak limits are similar to those of ordinary limits.

\subsection{Approximation to the identity}

A sequence \((f_n)\) of functions is called a \textit{approximation of the identity} or \(\delta\)-convergent if the following two conditions holds:

\begin{enumerate}
  \item \( \int_{\mathbb{R}} f_n(x) dx = 1 \) and \( \int_{\mathbb{R}} |f_n(x)| dx \leq K \),
  \item for any \( r > 0 \) \( \lim_{n \to \infty} \int_{\mathbb{R} \setminus [-r,r]} |f_n(x)| dx = 0 \)
\end{enumerate}

Here is the important theorem

\textbf{Theorem 2.3.} Under the above hypothesis

\[
\lim_{n \to \infty} L_{f_n} = \delta
\]

The proof is given in the Appendix A.

\textbf{Example 2.4.} The following are \(\delta\)-convergent sequences:

\begin{enumerate}
  \item \( f_n(x) = R_{\frac{1}{n}} \left(x + \frac{1}{2n}\right) \),
  \item \( f_n(x) = \frac{n}{\pi(1+n^2x^2)} \),
  \item \( f_n(x) = \sqrt{\frac{n}{\pi}} e^{-nx^2} \),
  \item \( f_n(x) = \frac{\sin(nx)}{\pi x} \).
\end{enumerate}
2.2 Approximation to the identity

Actually, all the above examples are constructed by the following simple idea. A function $K(x)$ is called a *kernel* if the following condition hold:

$$\int_{\mathbb{R}} K(x)dx = 1, \quad \int_{\mathbb{R}} |K(x)|dx \leq M.$$  

Then the sequence $f_n(x) := K(nx)n$ is a delta convergent sequence. 

Example $-1$ corresponds to the kernel $K(x) = R_{1,1}(x + \frac{1}{2}).$

Example $-2$ corresponds to the so called *Cauchy-Poisson* kernel $K(x) = \frac{1}{\pi(1+x^2)}.$

Example $-3$ corresponds to the so called *Gauss-Weierstrass* kernel $K(x) = \frac{1}{\sqrt{\pi}}e^{-x^2}.$

Notice that in this case $n$ was replaced by $\sqrt{n}.$

Finally, example $-4$ corresponds to the Dirichlet kernel $K(x) = \frac{\sin(x)}{\pi x}.$
3 The calculus of distributions

Here is the definition of the derivative $L'$ of a distribution $L$.

$$\langle L', \varphi \rangle := -\langle L, \frac{d \varphi}{dx} \rangle$$

Notice that $L' \in \mathcal{D}'$ and that the derivative is linear. Motivation for the above definition is provided by the 'integration by parts'.

Example 3.1. $\langle \delta', \varphi \rangle = -\varphi'(0)$ and in general

$$\langle \delta^{(n)}, \varphi \rangle = (-1)^n \frac{d^n \varphi}{dx^n}(0)$$

Theorem 3.2.

$$H' = \delta$$

Proof. Let $\varphi \in \mathcal{D}$ be a test function. We have to show that

$$\langle H', \varphi \rangle = \langle \delta, \varphi \rangle .$$

The RHS of the above equation is $\varphi(0)$ whilst the LHS is

$$\langle H', \varphi \rangle = -\langle H, \frac{d \varphi}{dx} \rangle = - \int_{-\infty}^{\infty} \frac{d \varphi}{dx} dx = \varphi(0)$$

due to the fundamental theorem of calculus. $\square$

Example 3.3. Here is the derivative of the rectangular impulse $R_{a,t}$:

$$R'_{a,t}(x) = a\delta(x) - a\delta(x-t)$$

If $f$ is right continuous and left continuous at a point $s$ we write $J(f, s) := f(s^+) - f(s^-)$ for the 'jump' at $s$.

Here is an important property of the derivative of piecewise derivable function with a discrete set of jumps.

**Theorem 3.4.** Let $f(x)$ be a function derivable outside of a discrete set $S = \{s_1, s_2, \cdots \}$ where $f$ jumps. Then $f'$ is the distribution

$$\frac{df}{dx} + \sum_{s \in S} J(f, s) \delta(x-s)$$

Proof. See Appendix B.
3.1 Derivative of a product.

Then we have the following result.

Theorem 3.5.

\[
(f.L)' = \frac{df}{dx}.L + f.L'
\]

Here is an interesting example:

Example 3.6.

\[
(x.H)' = H + x.\delta = H
\]

3.2 Graphical derivation
4 The Schwartz class $S$ and tempered distributions

We say that a function $f(x)$ is rapidly decreasing if for all $n$:

$$\lim_{x \to \infty} |x^n f(x)| = 0$$

The Schwartz $^5$ class $S$ consists of the $C^\infty$ rapidly decreasing functions $f$ all of whose derivatives are also rapidly decreasing.

Notice $\mathcal{D} \subset S$.

**Example 4.1.** $f(x) = e^{-x^2}$ is in $S$ but not in $\mathcal{D}$.

Notice that if $L \in S^*$ then $L \in \mathcal{D}^*$. So if $L \in S^*$ and $L \in \mathcal{D}'$ the we will said that $L$ is a tempered distribution. It is common to write $S'$ for the set of tempered distributions.

So a distribution is tempered if it can be also evaluated at functions of the Schwartz class. For example the distribution $L_f$ associated to the function $f(x) = e^{x^2}$ is not tempered since it value at $e^{-x^2}$ is infinite.

Let $SI(\mathbb{R})$ be the space of slowly increasing functions. Namely, the $C^\infty$ functions $g$ such that $gS \in S$. Then $S'$ is a $SI(\mathbb{R})$-module.

As we will see later tempered distributions are the natural model for the signals and their spectral analysis via Fourier transforms.

---

Appendix A.

We have to show that for a given $\varphi \in D$:

$$\lim_{n \to \infty} \langle f_n, \varphi \rangle = \varphi(0)$$

which is equivalent, by property $(i)$, to showing:

$$\lim_{n \to \infty} \int_{\mathbb{R}} f_n(x)(\varphi(x) - \varphi(0))dx = 0$$

Let $\epsilon > 0$ be arbitrarily small. Since $\varphi$ has compact support and is continuous there exists $r$ such that $|\varphi(x) - \varphi(0)| < \frac{\epsilon}{2K}$ if $|x| < r$ and $M$ such that $|\varphi(x)| \leq \frac{M}{2}$. Then

$$|\int_{\mathbb{R}} f_n(x)(\varphi(x) - \varphi(0))dx| = |\int_{\mathbb{R}\setminus[-r,r]} f_n(x)(\varphi(x) - \varphi(0))dx + \int_{[-r,r]} f_n(x)(\varphi(x) - \varphi(0))dx|$$

$$\leq \int_{\mathbb{R}\setminus[-r,r]} |f_n(x)||(|\varphi(x) - \varphi(0)|)dx + \int_{[-r,r]} |f_n(x)||(|\varphi(x) - \varphi(0)|)dx$$

$$\leq M \int_{\mathbb{R}\setminus[-r,r]} |f_n(x)|dx + \frac{\epsilon}{2K} \int_{[-r,r]} |f_n(x)|dx$$

$$\leq M \int_{\mathbb{R}\setminus[-r,r]} |f_n(x)|dx + \frac{\epsilon}{2}$$

where we have used property $(i)$. Then by using property $(ii)$ there exists $n_0$ such that

$$\int_{\mathbb{R}\setminus[-r,r]} |f_n(x)|dx \leq \frac{\epsilon}{2M}$$

if $n \geq n_0$. Then for $n \geq n_0$ it follows that

$$|\int_{\mathbb{R}} f_n(x)(\varphi(x) - \varphi(0))dx| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

This proves the theorem. □
Appendix B.

It is enough to show the theorem when \( S \) has just one point \( s \). Actually, we can assume w.l.o.g that \( s = 0 \). Then

\[
\langle f', \varphi \rangle = - \int_{\mathbb{R}} f(x) \frac{d\varphi}{dx} dx = \int_{\mathbb{R} \setminus [-r, r]} f(x) \frac{d\varphi}{dx} dx - \int_{[-r, r]} f(x) \frac{d\varphi}{dx} dx
\]

since \( f \) is bounded near zero we have that

\[
\langle f', \varphi \rangle = - \lim_{r \to 0} \int_{\mathbb{R} \setminus [-r, r]} f(x) \frac{d\varphi}{dx} dx = - \lim_{r \to 0} \left( \int_{-\infty}^{-r} f(x) \frac{d\varphi}{dx} dx + \int_{r}^{\infty} f(x) \frac{d\varphi}{dx} dx \right)
\]

\[
= - \lim_{r \to 0} \left( - \int_{-\infty}^{-r} \frac{df}{dx} \varphi(x) dx + (f \cdot \varphi)|_{-\infty}^{-r} + \int_{r}^{\infty} f(x) \frac{d\varphi}{dx} dx \right)
\]

\[
= \int_{-\infty}^{0} \frac{df}{dx} \varphi(x) dx - f(0^-) \varphi(0) - \lim_{r \to 0} \int_{r}^{\infty} f(x) \frac{d\varphi}{dx} dx,
\]

\[
= \int_{-\infty}^{0} \frac{df}{dx} \varphi(x) dx - f(0^-) \varphi(0) + \int_{0}^{\infty} \frac{df}{dx} \varphi(x) dx + f(0^+) \varphi(0),
\]

\[
= \int_{\mathbb{R}} \frac{df}{dx} \varphi(x) dx + J(f, 0) \varphi(0),
\]

\[
= \langle \frac{df}{dx} + J(f, 0) \delta, \varphi \rangle,
\]

which shows that \( f' = \frac{df}{dx} + J(f, 0) \delta \). □

References