Analysis III - Distribution Theory
IN BA3 - Martin Werner Licht

Notes by Ali EL AZDI

November 6, 2024

Introduction

This document is designed to provide a concise, step-by-step introduction to *Distribution Theory*, often referred to as the theory of generalized functions. The notes follow a clean and clear exposition style, illustrating concepts with simple examples. Should you discover any inaccuracies or have suggestions for improvements, please reach out at `ali.elazdi@epfl.ch`. Alternatively, you can contribute via a pull request if you wish to propose significant changes, additions, or corrections to the course content.

Contents

Chapter 1

Introduction to Distribution Theory

This course aims to establish a rigorous framework for generalized functions, show why objects like the Dirac Delta require a broader viewpoint than classical functions, and equip readers with the essential tools to work with these distributions.

1.1 Course I - Introduction to Distribution Theory

Motivation. In classical analysis, certain "functions" (like the Dirac Delta) don't exist in the usual sense but arise naturally in differential equations and signal processing. Instead of evaluating pointwise, we probe functions by integrating them against well-chosen, smooth "test" functions, capturing broader features rather than individual values.

1.1.1 Probing Functions via Integration

One of the core ideas in Distribution Theory is the concept of *testing* or *probing* a function by integrating it against another. Rather than focusing on pointwise values, we explore how a given function behaves when combined with various smooth functions under the integral sign.

Key Idea: If $\varphi : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are sufficiently nice functions (e.g., continuous, smooth, etc.), then we can explore:

$$\int_{\mathbb{R}} \varphi(x) \, g(x) \, dx.$$

This measures how φ and g *interact* or *overlap*.

By viewing functions through the lens of integrals against "test functions," one can capture broad, "averaged" information about a function's shape or behavior, rather than relying on individual point values.
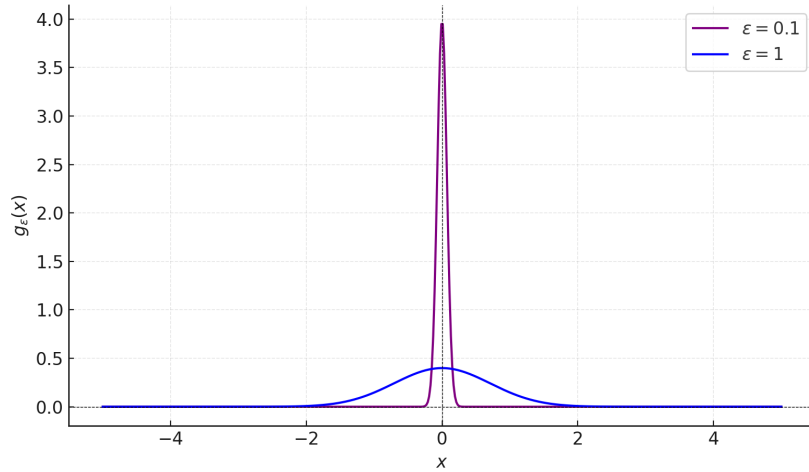
Example 1.1.1 (Gaussian Approximation.)**.**
For $\varepsilon > 0$, let's define

$$g_\varepsilon(x) = \frac{1}{\sqrt{2\pi}\,\varepsilon}\,e^{-\frac{x^2}{\varepsilon^2}}.$$

Each g_ε integrates to 1, and as $\varepsilon \to 0$, it becomes sharply peaked at $x = 0$. Hence, for a continuous function φ,

$$\lim_{\varepsilon \to 0} \int_\mathbb{R} \varphi(x)\,g_\varepsilon(x)\,dx = \varphi(0),$$

mimicking evaluation at a point but through integrals. Although no ordinary function does this exactly, this limiting behavior paves the way to defining the Dirac Delta as a functional.

Normalization of g_ε

Let's verify that g_ε integrates to one for each ε. Consider:

$$\int_\mathbb{R} g_\varepsilon(x)\,dx = \frac{1}{\sqrt{2\pi}\,\varepsilon} \int_\mathbb{R} \exp\!\left(-\frac{x^2}{\varepsilon^2}\right) dx.$$

By the substitution $u = x/sqrt\pi\varepsilon$, hence $dx = \varepsilon\,du$, we obtain

$$\int_\mathbb{R} \exp\!\left(-\frac{x^2}{\varepsilon^2}\right) dx = \varepsilon \int_\mathbb{R} e^{-u^2}\,du = \varepsilon\sqrt{\pi}.$$

Thus,

$$\int_\mathbb{R} g_\varepsilon(x)\,dx = \frac{1}{\sqrt{2\pi}\,\varepsilon}\left(\varepsilon\sqrt{\pi}\right) = 1.$$

Interpretation
Because for each ε, g_ε has unit integral, the quantity

$$\int_\mathbb{R} \varphi(x)\,g_\varepsilon(x)\,dx$$

can be viewed as a weighted average of φ, where the weights are dictated by the Gaussian shape g_ε. As $\varepsilon \to 0$, this integral concentrates around $x = 0$ and tends to evaluate φ at 0, at least for continuous φ:

$$\lim_{\varepsilon \to 0} \int_\mathbb{R} \varphi(x)\,g_\varepsilon(x)\,dx = \varphi(0).$$

1.1.2 The Dirac Delta and Generalized Functions

A major motivation for Distribution Theory is to rigorously handle the *Dirac Delta*, denoted δ_0, which captures the idea of "point evaluation at zero" in an integral.

Heuristic Motivation

We might ask if there is a classical function $g_0(x)$ such that

$$\int_{\mathbb{R}} g_0(x)\, \varphi(x)\, dx = \varphi(0) \quad \text{for all test functions } \varphi(x).$$

No ordinary function satisfies this. However, from the limiting behavior of the Gaussian family g_ε, we see a suggestion that in some generalized sense, one can define:

$$\delta_0[f] := f(0),$$

meaning δ_0 is a *functional* that takes a function f and returns its value at 0.

A Functional, Not a Function

The Dirac Delta is more properly viewed as a *linear functional*:

$$\delta_0 : C^0(\mathbb{R}) \to \mathbb{R}, \quad f \mapsto f(0).$$

It is not a genuine function in the classical sense but acts on functions by returning the value at a specific point.

Heuristic Notation: Physicists and engineers often write

$$\int_{\mathbb{R}} \delta_0(x)\, f(x)\, dx = f(0),$$

but strictly speaking, δ_0 is not a function. The equality holds in the distributional (generalized function) sense.
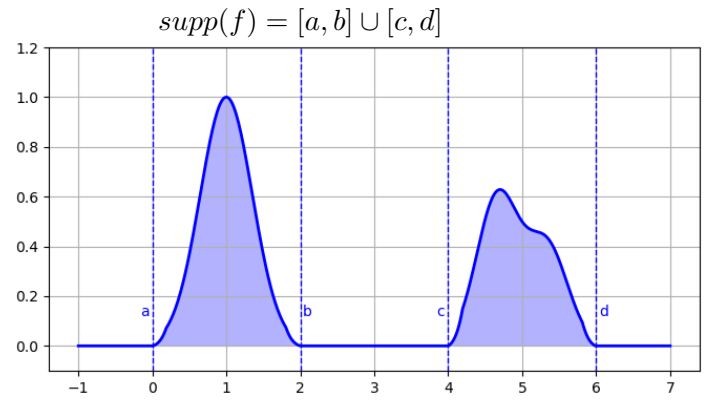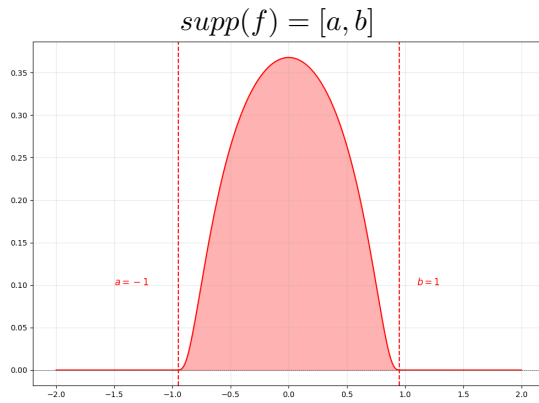
1.2 Foundational Definitions

To fully formalize these ideas, we need precise definitions of the underlying spaces of test functions and the generalized functions (distributions) themselves. Before diving into those spaces, we first clarify the notion of a function's *support*.

1.2.1 Support of a Function

Definition 1.2.1 (Support). *For a function $f : \mathbb{R} \to \mathbb{R}$, its* support *is defined as the closure of the set where f is nonzero:*

$$\text{supp}(f) = \overline{\{\, x \in \mathbb{R} \mid f(x) \neq 0 \,\}}.$$
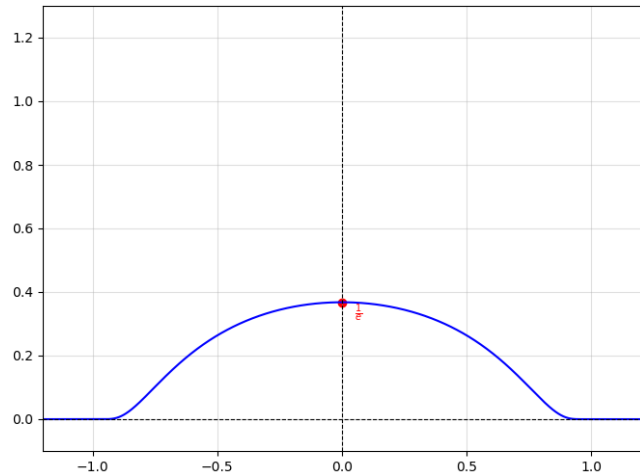
Example 1.2.1. *For example*

If a function f is nonzero only in a finite union of intervals, its support is then precisely that union of intervals (closed). If $f \equiv 0$ everywhere outside a certain bounded set, we say f has *compact support*.

A Closer Look at the Bump Function and Its Derivatives

The function

$$\varphi(x) = \begin{cases} e^{-\frac{1}{1-x^2}}, & |x| < 1, \\ 0, & |x| \geq 1, \end{cases}$$

is smooth and compactly supported in $[-1, 1]$. Despite looking piecewise, it belongs to $C^\infty(\mathbb{R})$ because both φ and all its derivatives vanish seamlessly at ± 1. Let us confirm this by examining its derivative.

Derivative Computation. For $|x| < 1$, using the chain rule, we compute:

$$\varphi'(x) = \frac{d}{dx}\left(e^{-\frac{1}{1-x^2}}\right) = e^{-\frac{1}{1-x^2}} \times (-1) \times \frac{d}{dx}\left[\frac{1}{1-x^2}\right].$$

Since

$$\frac{d}{dx}\left[\frac{1}{1-x^2}\right] = \frac{d}{dx}(1-x^2)^{-1} = -(1-x^2)^{-2} \times \left(-2x\right) = \frac{2x}{(1-x^2)^2},$$

we get

$$\varphi'(x) = e^{-\frac{1}{1-x^2}} \times \frac{2x}{(1-x^2)^2},$$

for $|x| < 1$, and $\varphi'(x) = 0$ for $|x| \geq 1$.

Behavior Near $x = \pm 1$.
Although the factor $\dfrac{2x}{(1-x^2)^2}$ blows up as $x \to \pm 1$, the exponential term $e^{-\frac{1}{1-x^2}}$ goes to zero *faster* than any power of $(1-x^2)$ goes to infinity. Specifically, as $x \to \pm 1$, we have $-\frac{1}{1-x^2} \to -\infty$, making $e^{-\frac{1}{1-x^2}} \to 0$. Hence the product remains finite and actually tends to 0 at $x = \pm 1$, guaranteeing continuity of φ' at those boundary points. Repeated differentiation yields higher-order derivatives that similarly vanish outside $(-1, 1)$ and remain smooth inside $(-1, 1)$, so indeed $\varphi \in C_c^\infty(\mathbb{R})$.

1.3 Preliminaries on Test Functions and Distributions

Distributions generalize the idea of functions by allowing us to make sense of certain *function-like* objects (such as the Dirac Delta) that are not strictly definable as ordinary functions.

1.3.1 The Space of Test Functions $\mathcal{D}$

In Distribution Theory, we work with a special space of "test" functions:

$$\mathcal{D} := C_c^\infty(\mathbb{R}) = \left\{\, \psi \colon \mathbb{R} \to \mathbb{R} \mid \psi \text{ is smooth and } \mathrm{supp}(\psi) \text{ is a bounded set} \,\right\}.$$

Any function $\psi \in \mathcal{D}$ is called a *test function*. We have just seen a prototype $\varphi(x)$ (the bump function) as one example: it is $\mathcal{C}^\infty$ and supported in $[-1, 1]$.

Why Compact Support?

Having compact support is crucial for many of the powerful theorems in Distribution Theory. It ensures that integrals

$$\int_{\mathbb{R}} \psi(x) \,(\text{some generalized object}) \, dx$$

will only "see" a finite region, thus avoiding certain convergence problems. Moreover, controlling both the smoothness and the support gives us a very flexible and well-behaved *testing* space.

1.3.2 Distributions as Continuous Linear Functionals on $\mathcal{D}$

A *distribution* is a linear functional on $\mathcal{D}$ that is continuous with respect to seminorms measuring the size of a test function and its derivatives on bounded supports.

Definition 1.3.1 (Distributions). *Let $\mathcal{D}'$ be the set of distributions over $\mathbb{R}$, the set of linear continous functionals over $\mathcal{D}$*

$$\mathcal{D}' = \{T \colon \mathcal{D} \to \mathbb{R} \mid T \text{ is linear and continuous}\}.$$

We denote the action of $T \in \mathcal{D}'$ on $\psi \in \mathcal{D}$ by $\langle T, \psi \rangle := T(\psi)$, analogous to an inner product.

A distribution $T \in \mathcal{D}'$ satisfies:

- **Boundedness**: For every $\psi \in \mathcal{D}$, $|T(\psi)|$ is finite.

- **Linearity**: For all scalars $\alpha, \beta \in \mathbb{R}$ and test functions $\psi, \varphi \in \mathcal{D}$,

$$T(\alpha\psi + \beta\varphi) = \alpha T(\psi) + \beta T(\varphi).$$

- **Continuity**:
  Meaning that "f changes little if the input changes little." Explicitly, for each interval $[a, b] \subseteq \mathbb{R}$, there exist constants $C > 0$ and $k \in \mathbb{N}_0$ such that:

$$\forall \varphi \in \mathcal{D}, \quad \operatorname{supp}(\varphi) \subseteq [a, b] \implies |f(\varphi)| \leq C \sum_{0 \leq i \leq k} \max_{x \in \mathbb{R}} \left| \partial^i \varphi(x) \right|.$$

1.3.3 Regular Distributions from (Locally Integrable) Functions

Definition 1.3.2 (Regular Distributions). *Let f be any (locally) integrable function over $\mathbb{R}$. We define a linear functional on the space of test functions $\varphi \in \mathcal{D}$ (infinitely differentiable functions with compact support) by*

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

*This functional is called a **regular** distribution.*

Properties:

- *Finiteness:* Since f is integrable and φ has compact support,

$$|\langle f, \varphi \rangle| = \left| \int_{\mathbb{R}} f(x)\, \varphi(x)\, dx \right| \leq \int_{\mathbb{R}} |f(x)|\, |\varphi(x)|\, dx \leq C \cdot \max_{x \in \mathbb{R}} |\varphi(x)|,$$

  where

$$C := \int_{\mathbb{R}} |f(x)|\, dx.$$

- *Linearity:* For scalars α, β and test functions $\varphi, \psi \in \mathcal{D}$,

$$\langle f, \alpha\varphi + \beta\psi \rangle = \alpha \langle f, \varphi \rangle + \beta \langle f, \psi \rangle.$$
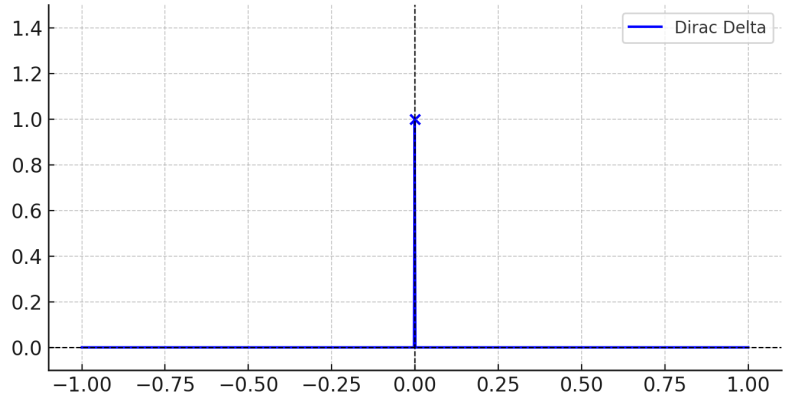
- *Continuity:* In the sense of distributions, continuity is verified by the inequality:

$$|\langle f, \varphi \rangle| \leq C \max_{x \in \mathbb{R}} |\varphi(x)|.$$

  Since any seminorm in the space of test functions can be bounded by a similar expression, this ensures continuity.

Definition 1.3.3 (Dirac Delta). *The* Dirac delta *at 0, denoted* δ_0, *is defined by*

$$\delta_0(\psi) := \psi(0), \quad \textit{for } \psi \in \mathcal{D}.$$

Interpretation: While δ_0 is not a genuine function in the usual sense, it is a linear functional. Every test function ψ is continuous and has compact support, so evaluating $\psi(0)$ is well-defined.

Continuity of δ_0: We have

$$|\delta_0(\psi)| = |\psi(0)| \leq \max_{x \in \mathbb{R}} |\psi(x)| \leq C \sum_{i=0}^{0} \max_{x \in \mathbb{R}} |\psi^{(i)}(x)|,$$

where we can choose $C = 1$ and $k = 0$. This satisfies the continuity condition in distribution theory.

Symbolic Notation: Often, one writes

$$\int_{\mathbb{R}} \delta_0(x)\,\psi(x)\,dx = \psi(0),$$

emphasizing the 'sampling' property of the delta at $x = 0$.

Definition 1.3.4 (Dirac Comb). *The* Dirac comb, *denoted* Δ_1, *is defined by*

$$\Delta_1(\varphi) := \sum_{n \in \mathbb{Z}} \varphi(n),$$

for each test function $\varphi \in \mathcal{D}$.

Properties:

- *Finiteness:* For $\text{supp}(\varphi) \subset [a,b]$, there are finitely many integers in $[a,b]$. Hence

$$\Delta_1(\varphi) = \sum_{n \in [a,b] \cap \mathbb{Z}} \varphi(n)$$

is a finite sum.

- *Linearity:*
$$\Delta_1(\alpha\varphi + \beta\psi) = \sum_{n \in \mathbb{Z}} \left[\alpha\,\varphi(n) + \beta\,\psi(n) \right] = \alpha\Delta_1(\varphi) + \beta\Delta_1(\psi).$$

- *Continuity:* For φ with support in $[a,b]$, let M be the number of integers in $[a,b]$. Then

$$|\Delta_1(\varphi)| \leq \sum_{n \in [a,b] \cap \mathbb{Z}} |\varphi(n)| \leq M \max_{x \in \mathbb{R}} |\varphi(x)|.$$

This verifies the required continuity condition in the sense of distributions.

10

Interpretation: Δ_1 can be viewed as a train of impulses (point masses) at every integer point, frequently used in signal processing to represent a periodic sampling process.

Example 1.3.1 (An Integral over $[0, \infty)$). *Consider the distribution T given by*

$$\langle T, \varphi \rangle := \int_0^\infty \varphi(x) \, dx.$$

Properties:

- Finiteness: *If $\operatorname{supp}(\varphi) \subset [a, b]$, then*

$$|\langle T, \varphi \rangle| = \left| \int_0^\infty \varphi(x) \, dx \right| \leq \int_a^b |\varphi(x)| \, dx \leq (b - a) \max_{x \in \mathbb{R}} |\varphi(x)|.$$

- Continuity: *For any compact interval $[a, b]$, we can choose $C = (b - a)$ and $k = 0$ to show*

$$|\langle T, \varphi \rangle| \leq C \max_{x \in \mathbb{R}} |\varphi(x)|.$$

- Linearity: *For $\alpha, \beta \in \mathbb{R}$,*

$$\langle T, \alpha\varphi + \beta\psi \rangle = \int_0^\infty [\alpha \, \varphi(x) + \beta \, \psi(x)] \, dx = \alpha \, \langle T, \varphi \rangle + \beta \, \langle T, \psi \rangle.$$

Example 1.3.2 (Evaluation of a Second Derivative at π). *Consider*

$$\langle T, \varphi \rangle := \varphi''(\pi).$$

Properties:

- Finiteness and Continuity:

$$|\langle T, \varphi \rangle| = |\varphi''(\pi)| \leq \max_{x \in \mathbb{R}} |\varphi''(x)|,$$

which fits the continuity requirement for distributions.

- Linearity: *Follows immediately from the linearity of the second derivative operator:*

$$\langle T, \alpha\varphi + \beta\psi \rangle = (\alpha\varphi + \beta\psi)''(\pi) = \alpha \, \varphi''(\pi) + \beta \, \psi''(\pi).$$

Example 1.3.3 (Shifted Dirac Delta δ_a). **Definition:** *For any $a \in \mathbb{R}$,*

$$\delta_a(\varphi) := \varphi(a).$$

This is a shift of the delta distribution at 0. In symbolic form, one often writes

$$\delta_a(x) := \delta(x - a).$$

Hence δ_0 is a special case with $a = 0$.

Example 1.3.4 (Scaled Dirac Comb Δ_L). **Definition:** *For a positive period $L > 0$, the scaled Dirac comb is given by*

$$\Delta_L(\varphi) := \sum_{n \in \mathbb{Z}} \varphi(nL),$$

which can be interpreted as a periodic sequence of impulses at intervals of length L.

1.4 Derivatives of Distributions

1.4.1 Definition and Motivation

Let $f \in C^\infty(\mathbb{R})$ be a smooth function and $\varphi \in \mathcal{D}$ be a test function whose support is contained in $[a, b]$. Then

$$\int_{-\infty}^{\infty} f'(x)\,\varphi(x)\,dx = \int_a^b f'(x)\,\varphi(x)\,dx.$$

Using integration by parts,

$$\int_a^b f'(x)\,\varphi(x)\,dx = [f(x)\,\varphi(x)]_a^b - \int_a^b f(x)\,\varphi'(x)\,dx.$$

Since φ vanishes at a and b (due to compact support),

$$\int_{-\infty}^{\infty} f'(x)\,\varphi(x)\,dx = -\int_{-\infty}^{\infty} f(x)\,\varphi'(x)\,dx.$$

Definition 1.4.1 (Distributional Derivative). *For a general distribution $f \in \mathcal{D}'$, its **distributional derivative** f' is defined by*

$$\langle f', \varphi \rangle := -\langle f, \varphi' \rangle,$$

for all $\varphi \in \mathcal{D}$.

Interpretation: This extends the classical idea of differentiation by effectively *transferring* the derivative onto the test function via an integration-by-parts formula.

1.4.2 Examples of Distributional Derivatives

Example 1.4.1 (A C^1 Function). *If $f \in C^1(\mathbb{R})$, then in the distribution sense,*

$$f'_{distributional} = f'_{classical}.$$

Hence the distributional derivative agrees with the usual one for sufficiently smooth functions.

Example 1.4.2 (Derivative of an Integral Functional). *Define*

$$\langle T, \varphi \rangle := \int_0^{\infty} \varphi(x)\,dx.$$

Then, using the definition of the distributional derivative,

$$\langle T', \varphi \rangle = -\langle T, \varphi' \rangle = -\int_0^{\infty} \varphi'(x)\,dx = \varphi(0),$$

by the fundamental theorem of calculus. This equals $\delta_0(\varphi)$, so

$$T' = \delta_0.$$

Alternatively, we can view T as coming from the Heaviside function

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

whose derivative (in the distribution sense) is δ_0.

Example 1.4.3 (Jump Distributions). *Let*

$$\langle T, \varphi \rangle := \int_0^1 \varphi(x)\, dx.$$

Then

$$\langle T', \varphi \rangle = -\int_0^1 \varphi'(x)\, dx = -\varphi(1) + \varphi(0),$$

which can be written as

$$\langle T', \varphi \rangle = -\delta_1(\varphi) + \delta_0(\varphi).$$

Equivalently, this distribution arises from the characteristic function of $[0,1]$:

$$g(x) = \begin{cases} 1, & 0 \le x \le 1, \\ 0, & otherwise. \end{cases}$$

Its distributional derivative captures the jumps at 0 and 1.

1.4.3   Distributional Derivatives of Piecewise Differentiable Functions

Suppose $f : \mathbb{R} \to \mathbb{R}$ is piecewise continuous and piecewise differentiable, with possible jump discontinuities at points $a_0 < a_1 < \cdots < a_n$. For each test function φ,

$$\langle f', \varphi \rangle = -\langle f, \varphi' \rangle = -\int_{-\infty}^{\infty} f(x)\, \varphi'(x)\, dx.$$

Breaking the integral into intervals and performing integration by parts on each sub-interval, we see that:

$$\langle f', \varphi \rangle = \int_{-\infty}^{a_0} f'(x)\, \varphi(x)\, dx + \sum_{i=1}^{n} \int_{a_{i-1}}^{a_i} f'(x)\, \varphi(x)\, dx + \int_{a_n}^{\infty} f'(x)\, \varphi(x)\, dx + \sum_{i=0}^{n} \varphi(a_i)\left[f(a_i^-) - f(a_i^+) \right].$$

Hence, the distributional derivative consists of:

- The classical derivative $f'(x)$ in each smooth region.

- Additional *Dirac delta* terms at each jump discontinuity a_i, weighted by the jump $f(a_i^-) - f(a_i^+)$.

1.4.4   Specific Examples

Example 1.4.4 (Hat Function). *Consider the **hat function**:*

$$f(x) = \begin{cases} 0, & x < -1, \\ x + 1, & -1 \le x < 0, \\ 1 - x, & 0 \le x < 1, \\ 0, & x \ge 1. \end{cases}$$

Since f is continuous and piecewise linear,

$$f'(x) = \begin{cases} 0, & x < -1, \\ 1, & -1 < x < 0, \\ -1, & 0 < x < 1, \\ 0, & x > 1. \end{cases}$$

This is also the weak derivative. *The* second *distributional derivative involves the jumps in $f'(x)$:*

$$f''(x) = \delta_{-1} - 2\,\delta_0 + \delta_1.$$

Example 1.4.5 (Heaviside Function). *The **Heaviside function** $H(x)$,*

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

has the distributional derivative

$$H'(x) = \delta_0.$$

Example 1.4.6 (Staircase Function). *Consider a **staircase function** with integer steps,*

$$f(x) = n, \quad if\ x \in [n-1, n),\ n \in \mathbb{Z}.$$

Its distributional derivative is

$$f'(x) = \sum_{n \in \mathbb{Z}} \delta_n,$$

often called a Dirac comb.

1.4.5 Derivatives of the Dirac Delta

For the Dirac delta δ_0, higher-order derivatives are defined by

$$\left\langle \delta_0^{(n)},\ \varphi \right\rangle = (-1)^n\ \varphi^{(n)}(0).$$

This highlights the principle that differentiating a delta distribution results in a new distribution that 'samples' higher-order derivatives of the test function at the origin.