# Distributions

## for Local Quantum Physics

Bharath Ron

# Table of Contents

# 1 | Differentiation of Functions

Vector spaces model spaces on which we can make sense of straight lines. Maps between vector spaces that take straight lines to straight lines are called linear maps. Linear maps are well understood. Differentiation is the approximation of a given function by a linear map. To talk about linear approximation we have to be working with vector spaces. To talk about approximation, we first need a notion of nearness and a notion of limit. Although Frechet spaces are the suitable spaces to with these properties, we will be working with Banach spaces in this chapter.

## 1.1 | Derivative of Functions on $\mathbb{R}^n$

A continuous function $f$ from an open set $U \subset X$ of a Banach space to a Banach space $Y$ is differentiable at $x \in U$ if there exists a linear map $df(x) : X \to Y$ such that,

$$\big\|f(x+h) - f(x) - df(x)(h)\big\|_Y \to 0, \text{ as } \|h\|_X \to 0.$$

Since norm of only the zero vector is zero, it follows that $df(x)$, if it exists is unique. The space of continuous linear transformations from the Banach space $X$ to $Y$ is denoted by $\mathcal{B}(X, Y)$, which is itself a Banach space with the norm,

$$\|T\|_{\mathcal{B}(X,Y)} = \sup_{\|x\| \leqq 1} \big\{\|Tx\|_Y\big\}.$$

The function is said to be differentiable on $U$ if it is differentiable for every $x \in U$. In such a case the derivative of $f$ is defined to be the map, $f' : x \mapsto df(x)$ which is a map from $U$ to $\mathcal{B}(X, Y)$. A differentiable function $f$ is said to be continuously differentiable on $U \subset X$ if the derivative map $f' : x \mapsto df(x)$ is continuous. The space of continuously differentiable functions from $U \subset X$ to $Y$ is denoted by $\mathcal{C}^1(U, Y)$.

### 1.1.1 | Elementary Properties

The basic properties of differentiation are immediate consequences of properties of the norm $\|\cdot\|_Y$. By triangle inequality the addition of continuously differentiable functions will be continuously differentiable in appropriate domain.

$$\big\|f(x+h) + g(x+h) - f(x) - g(x) - \big(df(x)(h) + dg(x)(h)\big)\big\|_Y$$

Since norms satisfy triangle inequality this is

$$\leqq \big\|f(x+h) - f(x) - df(x)(h)\big\|_Y + \big\|g(x+h) - g(x) - dg(x)(h)\big\|_Y.$$

Hence as $\|h\|_X \to 0$ the above term goes to zero and hence if $f$ and $g$ are differentiable at $x$ then so is $f + g$ and we have,

$$d(f+g)(x) = df(x) + dg(x)$$

Similarly, consider the product $f \cdot g(x) = f(x)g(x)$, consider

$$\left\|f(x+h)g(x+h) - f(x)dg(x)(h) - g(x)df(x)(h)\right\|_Y$$

The space of continuously differentiable functions $\mathcal{C}^1(U, Y)$ is an $\mathbb{R}$-algebra if we are working with complex field then $\mathbb{C}$-algebra.

The derivative of composition of functions is given by the chain rule.

### 1.1.2 | The Main Theorems

Suppose $f$ is a differentiable function from a connected open set $I$ of $\mathbb{R}$ to some Banach space $Y$. If we join $f(y)$ and $f(x)$ by a line, intuitively, the slope of the line should be expected to be less than the maximum derivative of the function. This is called the mean value theorem stated below. See [**?**] for a simple proof. We will give below a more complicated proof as in [1].

**Theorem 1.1.1.** *$f : \mathbb{R} \supset I \to Y$ be a differentiable. Then,*

$$\frac{\left\|f(y) - f(x)\right\|_Y}{\|y - x\|_{\mathbb{R}}} \leqq \underbrace{\left[\sup_{0 \leqq t \leqq 1} \left\{\left\|f'(x + t(y - x))\right\|_{\mathcal{B}(X,Y)}\right\}\right]}_{M}$$

**Proof**

The continuous map $f$ maps the line $[x, y] = x + t(y - x)$ to the curve $f(x + t(y - x))$. The distance of $f(x+t(y-x))$ from $f(x)$ is then given by, $\left\|f(x+t(y-x))-f(x)\right\|_Y$. The distance between the point $x$ and $x + t(y - x)$ is $\|x + t(y - x) - x\|_{\mathbb{R}} = t\|y - x\|_{\mathbb{R}}$. The slope is then given by,

$$\frac{\left\|f(x + t(y - x)) - f(x)\right\|_W}{t\|y - x\|_{\mathbb{R}}}.$$

To prove the theorem we must show that, $\left\|f(x + t(y - x)) - f(x)\right\|_W / t\|y - x\|_{\mathbb{R}} \leqq M$. holds for $t = 1$. We will show that the set

$$E = \left\{t \in I \;\middle|\; \tfrac{\|f(x+t(y-x))-f(x)\|_W}{t\|y-x\|_{\mathbb{R}}} \leqq M\right\}.$$

contains $1 \in \mathbb{R}$.

Since $f$ is continuous, $g(t) = \|f(x + t(y - x)) - f(x\|)/t\|y - x\|$ is a continuous function. This is a composition, addition, etc. of continuous functions. So, $\{g(t) \leqq M\}$ is a a closed set, and hence its pre-image which is $E$, must be closed. So, $E$ has a largest element $s$, note first that $E$ is non-empty, and contains 0. Then for $t > s$ with $t - s$ sufficiently small, we have,

$$\begin{aligned}\|f(x + t(y - x)) - f(x)\|_W &\leqq \|f(x + t(y - x)) - f(x + s(y - x))\|_W + \|f(x + s(y - x)) - f(x)\|_W \\ &\leqq M(t - s)\|y - x\|_{\mathbb{R}} + Ms\|y - x\|_{\mathbb{R}} = Mt\|y - x\|\end{aligned}$$

So, it should also contain $t$, which cannot happen since we assumed $s$ to be the largest element. Hence $s = 1$. $\square$

This can be extended to differentiable maps from lines in some Banach space to a target Banach space with a composition. Given a function $f : U \to Y$ that is differentiable at every point on the line segment joining $x$ and $y$ denoted by $[x, y]$, then we have the following generalised mean value theorem [1],

**Theorem 1.1.2. (Mean Value Theorem)** *If $f : U \to Y$ is differentiable on $[x, y]$ and $T \in \mathrm{Hom}(X, Y)$ then,*

$$\frac{\left\|f(y) - f(x) - T(y - x)\right\|_Y}{\left\|y - x\right\|_X} \leqq \left[\sup_{0<t<1} \left\{\left\|f'(x + t(y - x)) - T\right\|_{\mathcal{B}(X,Y)}\right\}\right]$$

Note that the supremum on the right hand side is a limit applied to smaller intervals. Let $\{f_i\}$ be a collection of differentiable functions which converge to a function $f$, suppose $\{f_i'\}$ converge to $g$ uniformly[1]. Then by applying the generalised mean value theorem to the function $f_i$ and setting $T = f_i'(x)$ and letting $i \to \infty$ we have,

$$\frac{\left\|f(y) - f(x) - g(x)(y - x)\right\|_Y}{\|y - x\|_X} \leqq \left[\sup_{0<t<1} \left\{\left\|g(x + t(y - x)) - g(x)\right\|_{\mathcal{B}(X,Y)}\right\}\right]$$

As $y \to x$, the right hand side goes to zero, and hence $f$ is differentiable at $x$ with $f'(x) = g(x)$.

Similar to mean value theorem, many of the properties of a differentiable function are closely related to the properties of its linear approximation. The inverse function theorem makes this precise and it is this theorem which allows usage of linear algebra for studying differentiable functions. See [**?**] for a proof.

**Theorem 1.1.3. (Inverse Function Theorem)** *Suppose $f \in \mathcal{C}^1(U, Y)$ and $f'(x) = df(x)$ is invertible, then there exists a neighborhood of $x$ on which $f$ has a smooth inverse.*

**Proof**

Let $f$ be a continuously differentiable function on $U$. Then we have a new map which sends each point in $U$ $x$ to the linear map $f'(x) \in \mathcal{B}(X, Y)$. So the second differential at $x$ is the linear map $f''(x)$ such that,

$$\left\|f'(x + h) - f'(x) - f''(x)(h)\right\|_{\mathcal{B}(X,Y)} \to 0, \text{ as } \|h\|_X \to 0.$$

where $f''(x) \in \mathcal{B}(X, \mathcal{B}(X, Y))$, since $\mathcal{B}(X, \mathcal{B}(X, Y)) \cong \mathcal{B}(X, X; Y)$, this is a multilinear map. The higher derivatives hence belong to this complicated vector space,

$$f^{(k)} \in \mathcal{C}(U, \mathcal{B}^k(X, Y)).$$

Since where $\mathcal{B}(X, X; Y)$ consist of all bilinear maps from $X \times X$ to $Y$.

[1] for any $\epsilon > 0$, there exist some $N$ such that for every $i > N$, $\|f_i'(x) - g(x)\|_\infty < \epsilon$ for all $x$.

## 1.2 | Differentiable Functions on $\mathbb{R}^n$

### 1.2.1 | Space of Smooth Functions $\mathcal{C}^\infty(\mathbb{R}^n)$

## 1.3 | Differentiable Functions on Manifolds

### 1.3.1 | Sheaves in Geometry

### 1.3.2 | Differentiable Manifolds

#### 1.3.2.1 | Differentiable Structure

#### 1.3.2.2 | The Sheaf of Smooth Functions

## A | Topological Facts

# 2 | Test Functions & Distributions

One can think of continuous functions as linear functionals on the space of compactly supported smooth function. This allows us to use the nice properties of compactly supported smooth functions to study other functions.

## 2.1 | Space of Test Functions $\mathcal{C}_c^\infty(\mathbb{R}^n)$

The space of complex valued smooth functions with compact support on $\mathbb{R}^n$ is called the space of test functions, denoted by $\mathcal{C}_c^\infty(\mathbb{R}^n)$. Elements of $\mathcal{C}_c^\infty(\mathbb{R}^n)$ are called test functions. $\mathcal{C}_c^\infty(\mathbb{R}^n)$ is non-empty. For each $x \in \mathbb{R}^n$ there exist compactly supported smooth functions $\varphi_x$ such that for $\varphi_x(x) > 0$. See §2.1.2.

Let $\lambda$ be the Lebesgue measure on $\mathbb{R}^n$, then each continuous function $f \in \mathcal{C}(\mathbb{R}^n)$ defines a linear functional on the space of test functions, given by,

$$\langle f, \cdot \rangle : \varphi \mapsto \langle f, \varphi \rangle \coloneqq \int_{\mathbb{R}^n} f(x)\varphi(x) d\lambda(x).$$

Since $\varphi$ is compactly supported, the above integral is well-defined, and the map is a bounded linear functional on $\mathcal{C}_c^\infty(\mathbb{R}^n)$. Note that the nicer the test function is, the more integrable the product will be and can can allow us to study functions that are not so nice.

**Lemma 2.1.1.** *The map $f \mapsto \langle f, \cdot \rangle$ is injective.*

**Proof**

If $f, g$ be two continuous functions that are not equal, then $h = f - g \neq 0$. So there exists some $y \in X$ such that $h(y) \neq 0$. By continuity, there exists a neighborhood $U_y$ around $y$ where $h$ is non-zero. Consider a test function that is non-negative with $\varphi(y) > 0$, such that $\operatorname{supp}(\varphi) \subseteq U_y$. In such a case we have,

$$\begin{aligned} \langle f, \varphi \rangle - \langle g, \varphi \rangle &= \int_{\mathbb{R}^n} \big(f(x) - g(x)\big)\varphi(x) dx \\ &= \int_{\mathbb{R}^n} h(x)\varphi(x) dx \neq 0 \end{aligned}$$

which means $\langle f, \varphi \rangle - \langle g, \varphi \rangle \neq 0$. So, $f \mapsto \langle f, \cdot \rangle$ is injective. □

This is a very useful property, it tells us if two functions are equal by checking if the functionals they give rise to are equal. Heuristically speaking, the test functions can see if the functions are equal or not. This property also holds when continuity of $f$ is replaced by local integrability, however in this case the test functions can see if two functions are equal everywhere except a measure zero set, see §1.2 in [1]. The idea of distribution theory is to study the space of linear functionals on the space of test functions instead of the space of continuous functions itself, since every continuous function gives rise to a linear functional, the space we are working in is larger and more stuff can be studied.

### 2.1.1 | Convolutions

If $f, g \in \mathcal{C}(\mathbb{R}^n)$ with either one with compact support, then the convolution $f * g$ is the continuous function defined by,

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

Convolution $f * g$ is the superposition of weighted translates of $f$ weighted by $g$. At each point $y$, the weight is $g(y)d\lambda(y)$. Hence $f * g$ can be thought of as $f$ smeared out by the values of the function $g$.

By the translational invariance of the Lebesgue measure, and taking $(x-y)$ as a new integration variable, we obtain that, $f * g = g * f$. However it is useful to see this using test functions. Considering the action of the convolution on test functions,

$$\begin{aligned}\big\langle f * g, \varphi \big\rangle &= \int_{\mathbb{R}^n} \Big( \int_{\mathbb{R}^n} f(x-z)g(z)d\lambda(z) \Big) \varphi(x)d\lambda(x) \\ &= \iint_{\mathbb{R}^n} f(x)g(y)\varphi(x+y)d\lambda(x)d\lambda(y) = \big\langle g * f, \varphi \big\rangle.\end{aligned}$$

Using injectivity of the map $f \mapsto \langle f, \cdot \rangle$ it follows that $f * g = g * f$. Note that convolution cannot be performed on general manifolds. It needs a group structure and an invariant measure. Convolution makes sense on locally compact groups. The commutativity of convolution only holds if the locally compact group is abelian.

Let $f, g$, and $h$ be three continuous functions with atleast one of them with compact support, we have,

$$\big\langle (f * g) * h, \varphi \big\rangle = \iiint_{\mathbb{R}^n} f(x)g(y)h(z)\varphi(x+y+z)d\lambda(x)d\lambda(y)d\lambda(z).$$

It follows by associativity of the group operation that the convolution operation is associative,

$$(f * g) * h = f * (g * h).$$

Convolutions synthesise the nice properties of the component functions, so the convolution is a function that has nice properties. This allows us to bring in the tools that can be used on nice functions to be used on not so nice functions. Suppose $f$ and $g$ are differentiable, and by linearity of integral, it follows that for any $\alpha \in \mathbb{N}^n$,

$$\begin{aligned}\big\langle \partial^\alpha (f * g), \varphi \big\rangle &= \int_{\mathbb{R}^n} \Big( \partial^\alpha \int_{\mathbb{R}^n} f(x-z)g(z)d\lambda(z) \Big) \varphi(x)d\lambda(x) \\ &= \int_{\mathbb{R}^n} \Big[ \int_{\mathbb{R}^n} \big(\partial^\alpha f(x-z)\big) g(z)d\lambda(z) \Big] \varphi(x)d\lambda(x) = \big\langle \big((\partial^\alpha f) * g\big), \varphi \big\rangle.\end{aligned}$$

Here, we are using the notational convention,

$$\partial^\alpha f \equiv \partial^{|\alpha|} f / \partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}.$$

Intuitively, $f * g$ is a weighted superposition of translates of $f$, weighted by the values of the function $g$. So, the rates of changes of the weighted translates is the same as the rate of change of the function and then taking the weighted translates. Since the convolution $f * g$ can also be thought of as weighted translates of of $g$ weighted by the values of the function $f$, the differentiation of the convolution can also be expressed as convolution of $f$ with differentiation of $g$.

$$\partial^\alpha(f * g) = (\partial^\alpha f) * g = f * (\partial^\alpha g).$$

So the differentiability of convolutions depend on the combined differentiability of either of the functions. If $f \in \mathcal{C}^j(\mathbb{R}^n)$ and $g \in \mathcal{C}^k(\mathbb{R}^n)$ then we have for $|\alpha| \leqq j$ and $|\beta| \leqq k$,

$$\partial^{\alpha+\beta}(f * g) = \big(\partial^\alpha f\big) * (\partial^\beta g).$$

Summing up, we have proved the following theorem,

**Theorem 2.1.2.** *If $f \in \mathcal{C}_c^j(\mathbb{R}^n)$, $g \in \mathcal{C}^k(\mathbb{R}^n)$ then $f * g \in C_c^{j+k}(\mathbb{R}^n)$.*

Note that the compact support condition on $f$ can be replaced by local integrability and the result will still hold, see [1]. This property of convolutions being more differentiable than individual functions is a very useful tool and can be used to approximate functions by more differentiable ones. This is called regularisation.

Given $f \in \mathcal{C}_c^j(\mathbb{R}^n)$ consider a test function $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ such that $\varphi(x) \geq 0$ and

$$\int_{\mathbb{R}^n} \varphi(x) d\lambda(x) = 1.$$

Define a new function $f_\varphi$ by the convolution,

$$f_\varphi = f * \varphi.$$

Since $\varphi$ is a smooth function, $f * \varphi$ is also a smooth function.

The product $f(x-y)\varphi(y)$ is non-zero when both $f(x-y)$ and $\varphi(y)$ are non-zero. If the support of $\varphi$ is denoted by $p_\varphi$, then we have $f(x-y)\varphi(y) = f(x-y)\varphi(y)\chi_{p_\varphi}(y)$. By definition of convolution we have

$$(f * \varphi)(x) = \int_{\mathbb{R}^n} f(x-y)\varphi(y) d\lambda(y) = \int_{p_\varphi} f(x-y)\varphi(y) d\lambda(y).$$

Note that $f * \varphi(x)$ is a weighted superposition of all the values $f(x-y)$ for $y \in p_\varphi$. So, to make the value $f * \varphi(x)$ closer to $f(x)$ we have to reduce the support of the test function $\varphi$. Heuristically, this makes $f * \varphi$ less smeared out, and the values more closer to the original function. How closely the convolution approximates to the original function depends on the support properties of the test function $\varphi$.

**Theorem 2.1.3. (Regularization)** *Let $0 \leqq \varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \varphi(x) d\lambda(x) = 1$. If $f \in \mathcal{C}_c^j(\mathbb{R}^n)$ then whenever $\partial^\alpha f$ is defined*

$$\partial^\alpha f \to \partial^\alpha f_\varphi$$

*uniformly as the support of $\varphi$ tends to a point.*

**Proof**

Using $\int_{\mathbb{R}^n}\varphi(x)d\lambda(x)=1$ and the definition of $f_\varphi$, we have

$$\begin{aligned}\big|f(x)-f_\varphi(x)\big| &= \left|f(x)\int_{\mathbb{R}^n}\varphi(x)d\lambda(x)-\int_{\mathbb{R}^n}f(x-y)\varphi(y)d\lambda(y)\right| \\ &= \left|\int_{p_\varphi}\big(f(x)-f(x-y)\big)\varphi(y)d\lambda(y)\right| \leqq \Big[\sup_{y\in p_\varphi}\big|f(x)-f(x-y)\big|\Big]\left|\int_{\mathbb{R}^n}\varphi(x)d\lambda(x)\right|.\end{aligned}$$

So we have,

$$\big|f(x)-f_\varphi(x)\big| \leqq \sup_{y\in p_\varphi}\big|f(x)-f(x-y)\big|.$$

By uniform continuity of $f$, for any $\epsilon>0$, there exists some $\delta$ such that, $|f(x)-f(x-y)|<\epsilon$ whenever $|y|<\delta$. Choose $\varphi$ with a small support $p_\varphi$ such that if $y\in p_\varphi$ then $|y|<\delta$. Then we have,

$$\big|f(x)-f_\varphi(x)\big| \leqq \sup_{y\in p_\varphi}\big|f(x)-f(x-y)\big|<\epsilon.$$

By assumption, $f\in\mathcal{C}_c^j(\mathbb{R}^n)$, and in particular Since $f\in\mathcal{C}_c^j(\mathbb{R}^n)$, it has a compact support, and hence, we can cover its support by finitely many $\delta$ balls. Hence we have,

$$\sup\big|f(x)-f_\varphi(x)\big|<\epsilon.$$

$f_\varphi$ converges uniformly to $f$ as support of $\varphi$ shrinks. For the case of $\partial^\alpha f$, since $f\in\mathcal{C}_c^j(\mathbb{R}^n)$, $\partial^\alpha f$ is uniformly continuous for $|\alpha|\leqq j$, by replacing $f$ with $\partial^\alpha f$ it follows that $\partial^\alpha f$ converges uniformly to $\partial^\alpha f_\varphi$ as support of $\varphi$ shrinks. □

Note that instead of $f$ being compactly supported, if we can instead assume that it is integrable, we can approximate integrable functions by differentiable functions. However this involves differentiations under integral signs for measurable functions and the proof requires some measure theory. In fact, if $f$ is an $L^p$ function then $f_\varphi$ is a continuously differentiable function, and as the support of $\varphi$ gets smaller

$$\big\|f-f_\varphi\big\|_{L^p}\to 0$$

where,

$$\big\|f\big\|_{L^p}:=\left[\int_{\mathbb{R}^n}\big|f(x)\big|^p d\lambda(x)\right]^{\frac{1}{p}}.$$

Hence the collection of compactly supported smooth functions are dense in space of $L^p$ functions. If $f\in L^p(\mathbb{R}^n)$, then we have

$$\begin{aligned}\big\|f(x)-f_\varphi(x)\big\|_{L^p} &= \left[\int_{\mathbb{R}^n}\left|\int_{p_\varphi}\big(f(x)-f(x-y)\big)\varphi(y)d\lambda(y)\right|^p d\lambda(z)\right]^{\frac{1}{p}} \\ &\leqq \Big[\sup_{y\in p_\varphi}\big|f(x)-f(x-y)\big|\Big]\left[\int_{p_\varphi}\big|\varphi(x)\big|^p d\lambda(x)\right]^{\frac{1}{p}}.\end{aligned}$$

The last term tends to zero as support of $\varphi_i$ shrinks. Hence every $L^p$ function can be approximated by compactly supported smooth functions.

### 2.1.2 | Partitions of Unity

## 2.2 | Topology on $\mathcal{C}_c^\infty(\mathbb{R}^n)$

To get a meaningful notion of approximation of functions, the topology should capture the important properties of the space of function under consideration, while also not being too restrictive. We have so far used the topology of uniform convergence. In the uniform topology, two functions are close to each other if their values at each point in the space are close to each other. This is however a very strong requirement, two continuous functions might have values close to each other in some regions and differ more outside the region, we would also like to think of such functions as being close to each other in that region.

So, the appropriate topology for our purpose is to start with a collection of semi-norms, indexed by compact sets of $\mathbb{R}^n$. For any compact set $K \Subset \mathbb{R}^n$, and any continuous function, define,

$$\left\|f\right\|_K = \sup_{x\in K} \left|f(x)\right|.$$

Since $|\lambda f(x)| = |\lambda|\, |f(x)|$ and $|\cdot|$ satisfies triangle inequality, we have,

$$\|\lambda f\|_K = |\lambda|\, \|f\|_K, \ \ \|f+g\|_K \leqq \|f\|_K + \|g\|_K.$$

So, $\|\cdot\|_K$ defines a semi-norm on $\mathcal{C}(\mathbb{R}^n)$. It is not a norm because a function might be zero in $K$ and not zero outside $K$, such non-zero functions have zero when 'measured' by $\|\cdot\|_K$. The topology generated by the collection of semi-norms $\mathcal{K}$ consisting of all semi-norms of the form $\|\cdot\|_K$ where $K \Subset \mathbb{R}^n$ is compact, is called the topology of compact convergence.

Two functions $f$ and $g$ are $\epsilon$-close to each other in this topology if for every compact set $K$ we have,

$$\|f-g\|_K < \epsilon$$

$f$ is $\epsilon$ close to $g$ on the compact set $K$ if $\|f-g\|_K < \epsilon$. Denote the set of all functions $g$ which are $\epsilon$-close to $f$ on a compact set $K$ by $B_K(f,\epsilon)$. The neighborhood basis of $f \in \mathcal{C}(\mathbb{R}^n)$ in this topology consists of subsets of $\mathcal{C}(\mathbb{R}^n)$, of the form,

$$B_{\mathcal{K}}(f,\epsilon) = \bigcap_{\|\cdot\|_K \in \mathcal{K}} B_K(f,\epsilon).$$

where $B_K(f,\epsilon) = \{g \in \mathcal{C}(\mathbb{R}^n) \mid \left\|f-g\right\|_K < \epsilon\}$. This topology also allows us to study unbounded functions by multiplying them with a cutoff function and is hence versatile and more useful than the uniform topology which is too restrictive, and can only be used to study very special class of well-behaved functions. We can use partitions of unity, see [1], to carefully restrict the functions to the regions of the space that are of interest to us.

The topology of compact convergence is not induced by a norm on $\mathcal{C}(\mathbb{R}^n)$. Because if it was induced by a norm $\|\cdot\|$, then for any $f \in \mathcal{C}(\mathbb{R}^n)$ we can choose a constant $c$ with $\|f\| < |c|$ in which case $f-c$ will be invertible (the series $\sum (f/c)^i$ converges and is the inverse) which is absurd if $f$ is unbounded.

## 2.3 | Distributions in $\mathbb{R}^n$

As we stated at the start, the idea of distribution theory is to study functions of interest to us as linear functionals on the space of test functions. Each locally integrable function $f$ gives rise to a map on the space of test functions $\mathcal{C}_c^\infty(\mathbb{R}^n)$ of the form,

$$\varphi \mapsto \langle f, \varphi \rangle \equiv \int_{\mathbb{R}^n} f(x)\varphi(x)d\lambda(x).$$

We will call these linear functionals of integral type.

Heuristically, we are trying to study the functions $f$ based on how test functions can percieve them. The theory of distributions allow us to construct functions that satisfy some expected properties. The approach is to study the expected properties abstractly as linear functionals on the space of test functions[1], and then approximate the linear functional by linear functionals of the above integral type. To be able to approximate we need an appropriate topology on the space of linear functionals on $\mathcal{C}_c^\infty(\mathbb{R}^n)$. The important requirement is that the **properties** we were interested in studying are not lost in the process of approximation. So after approximation, the resulting function will indeed have the **properties** we wanted. Linear functionals of interest to us are continuous linear functionals in this topology. The **properties** are those which can be detected by compactly supported smooth functions, via the integral above. We will call a linear functional on $\mathcal{C}_c^\infty(\mathbb{R}^n)$ a distribution if it assigns close by values to close by test functions. Hence the natural topology on $\mathcal{C}_c^\infty(\mathbb{R}^n)$ is our starting point.

A linear functional $\kappa$ on $\mathcal{C}_c^\infty(\mathbb{R}^n)$ is called a distribution if for every net of test functions $\{\varphi_i\}_{i\in F}$, converging to $\varphi$, the net of numbers $\{\kappa(\varphi_i)\}$ converges to $\kappa(\varphi)$.

$$\{\varphi_i\} \to \varphi \Rightarrow \{\kappa(\varphi_i)\} \to \kappa(\varphi).$$

where the topology on $\mathcal{C}_c^\infty(\mathbb{R}^n)$ is the natural topology, and the topology on $\mathbb{C}$ is the standard Euclidean topology. A convinient notation for distributions is the pairing notation,

$$\kappa(\varphi) \equiv \langle \kappa, \varphi \rangle.$$

This notation is convinient to express and use properties. The collection of all distributions on $\mathbb{R}^n$ is denoted by $\mathcal{D}(\mathbb{R}^n)$. $\mathcal{D}(\mathbb{R}^n)$ comes equipped with the weak$^*$ topology, defined by the semi-norms,

$$p_\varphi : \kappa \mapsto \langle \kappa, \varphi \rangle$$

A sequence of distributions $\{\kappa_i\}$ converges to $\kappa$ if $\kappa_i(\varphi)$ to $\kappa(\varphi)$ as complex numbers for all $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$. To discuss structure of the space of distributions we will start with an equivalent, more useful characterization of distributions.

---

[1] The general idea behind distribution theory is similar to the Grothendieck-Yoneda philosophy of category theory, where a mathematical objects is studied by its relation to other objects. Categorically speaking, one can view the objects of a category $\mathcal{C}$ as functors in $\mathbf{Sets}^{\mathcal{C}}$, via the Yoneda embedding. The objects of $\mathcal{C}$ represent the 'representable functors'. The case of distributions is similar. We are trying to study an object, which in our case are integrable functions $f$, we are trying to study its relation to other functions via the integral,

$$\langle f, \cdot \rangle : \varphi \mapsto \int_{\mathbb{R}^n} f(x)\varphi(x)d\lambda(x)$$

The functional $\langle f, \cdot \rangle$ can be thought of as a 'representable functional', represented by the function $f$. Note that in this case, we do not care about all the relations $f$ has with other objects, by considering the above integral we are trying to focus on the relations that matter for us.

**Theorem 2.3.1.** *$\kappa$ is a distribution iff $\forall \varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ and $\|\cdot\|_K \in \mathcal{T}$, $\exists\, c_\kappa, k$ such that,*

$$\left|\langle \kappa, \varphi \rangle\right| \leqq c_\kappa \textstyle\sum_{|\alpha| \leqq k} \left\|\partial^\alpha \varphi\right\|_K. \tag{distribution}$$

**Proof**

$\Leftarrow$ If $\{\varphi_i\}$ is a sequence of test functions that tend to the zero function in the natural topology of $\mathcal{C}_c^\infty(\mathbb{R}^n)$ then, $\|\partial^\alpha \varphi_i\|_K \to 0$. If $\kappa$ satisfies the inequality (distribution), then we have, $|\langle \kappa, \varphi_i \rangle| \leqq c_\kappa \sum_{|\alpha| \leqq k} \|\partial^\alpha \varphi_i\|_K$ which means $\langle \kappa, \varphi_i \rangle \to 0$.

$\Rightarrow$ Suppose the condition is not true then there must exist some compact set $K$ such that $|\langle \kappa, \varphi \rangle| \leqq c_\kappa \sum_{|\alpha| \leqq k} \|\partial^\alpha \varphi\|_K$ does not hold for any $c_\kappa$ and $k$. To arrive at a contradiction, consider the term $\varphi_j$, and take $c_\kappa = k = j$, we have,

$$\left|\langle \kappa, \varphi_j \rangle\right| > j \textstyle\sum_{|\alpha| \leqq j} \left\|\partial^\alpha \varphi_j\right\|_K.$$

For the sake of simplicity assume by rescaling that $\langle \kappa, \varphi_j \rangle = 1$. Then whenever $j \geq |\alpha|$, we have, $|\partial^\alpha \varphi_j| \leqq 1/j$ or the sequence $\{\varphi_j\}$ converges to zero even though $\{\langle \kappa, \varphi_j \rangle\}$ does not. So, the condition cannot be false. $\square$

Hörmander [1], uses this inequality as the definition and gives a few other characterizations. Although unintuitive, this equivalent characterisation is convinient as it uses inequalities instead of convergences of sequences, and is easier to use and manipulate.

### 2.3.1 | Space of Distributions $\mathcal{D}(\mathbb{R}^n)$

We can identify the equivalence classes of locally integrable functions with functionals, $[f] \mapsto \langle f, \cdot \rangle$, where the equivalence relation is that the functions are the same almost everywhere. So, we have an embedding of equivalence classes of functions inside the space of distributions. Since each distribution $\kappa$ has to satisfy the inequality,

$$\left|\langle \kappa, \varphi \rangle\right| \leqq c_\kappa \textstyle\sum_{|\alpha| \leqq k} \left\|\partial^\alpha \varphi\right\|_K.$$

The constant $c_\kappa$ is not so useful to us. We can however try to learn what distributions would be if we constrain the number $k$. A distribution $\kappa$ is said to be of order $k$ if the same $k$ can be used for every compact set $K$. The space of all distributions of order $k$ is denoted by $\mathcal{D}_k(\mathbb{R}^n)$. Since the sum and scaling of linear functionals is a linear functional, taking modulus of sum and scales of the functionals, it follows that they also satisfy the inequality above for the same $k$. Hence the space of all $k$th order distributions is a vector space. Their union $\mathcal{D}_F(\mathbb{R}^n) = \bigcup_k \mathcal{D}_k(\mathbb{R}^n)$ is called the space of finite order distributions.

Note that the space of continuous linear functionals on $\mathcal{C}_c^k(\mathbb{R}^n)$ is complete. Since every function $f \in \mathcal{C}_c^k(\mathbb{R}^n)$ can be approximated by smooth functions in $\mathcal{C}_c^\infty(\mathbb{R}^n)$ by (2.1.3), it follows that we can construct a linear functional on $\mathcal{C}_c^k(\mathbb{R}^n)$ using distributions by,

$$\left\langle \widehat{\kappa}, f \right\rangle \coloneqq \lim_{i \to \infty} \left\langle \kappa, f_{\varphi_i} \right\rangle. \tag{extension}$$

This is a unique extension of $\kappa$ to $\mathcal{C}_c^k(\mathbb{R}^n)$ which also satisfies

$$\left|\langle \widehat{\kappa}, f \rangle\right| \leqq c_\kappa \textstyle\sum_{|\alpha| \leqq k} \left\|\partial^\alpha f\right\|_K.$$

To show that the limit does not depend on the choice of sequence consider the new sequence $\{f_{\varphi_i} - f_{\psi_j}\}$, this tends to zero, and hence $\lim_{i\to\infty} \left\langle \kappa, f_{\varphi_i} - f_{\psi_j} \right\rangle \to 0$.

Riesz-Markov theorem says that every positive linear functional $\widehat{\mu}$ on the space $\mathcal{C}_c(\mathbb{R}^n)$ of compactly supported continous functions with the uniform topology corresponds to a positive Radon measure $\mu$ such that

$$\langle \widehat{\mu}, f \rangle = \int_{\mathbb{R}^n} f(x) d\mu(x)$$

for all $f \in \mathcal{C}_c(\mathbb{R}^n)$.

A distribution $\kappa$ is positive if $\langle \kappa, \varphi \rangle \geq 0$ for every $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ with $\varphi(x) \geq 0$ for all $x$. Let $\varphi$ be any test function (not necessarily positive), since $\varphi$ has a compact support, we can choose a compact set $K$ such that $\operatorname{supp} \varphi \subset K$. Choose a cutoff function $\chi_K$ with $\chi_K(x) = 1$ for $x \in K$ and smoothly vanishes outside. Using $\varphi$ we can construct a new function,

$$\varphi^+ = e^{i\theta} \|\varphi\|_{\sup} \chi_K + \varphi$$

The choice of $\theta$ is made so that $\varphi^+$ vanishes when supremum of $\varphi$ is attained. This function is useful.

**Theorem 2.3.2.** *If $\kappa$ is a positive distribution, then it is a positive measure.*

### Proof

We have to show that this gives rise to a positive functional on the space of compactly supported continuous functions. Firstly we show that it extends uniquely to a bounded linear functional on $\mathcal{C}_c(\mathbb{R}^n)$. By (extension), it follows that if $\langle \kappa, \varphi \rangle \leqq c_K \|\varphi\|_K$ then it can be uniquely extended. Then by Riesz-Markov theorem it follows that it should correspond to a unique measure.

Let $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$. Consider the action of $\kappa$ on the new smooth function $\varphi$,

$$\langle \kappa, \varphi^+ \rangle = e^{i\theta} \|\varphi\|_{\sup} \langle \kappa, \chi_K \rangle + \langle \kappa, \varphi \rangle$$

By taking modulus, this should be a non-negative number, hence we get,

$$\left| \langle \kappa, \varphi \rangle \right| \leqq c_K \|\varphi\|_K$$

Note that here $\| \cdot \|_{\sup}$ and $\| \cdot \|_K$ are the same because support of $\varphi$ is inside $K$. The constant $c_K$ only depends on the function $\chi_K$. Hence, there exists a unique extension of $\kappa$ to a positive linear functional on $\mathcal{C}_c(\mathbb{R}^n)$ and by Riesz-Markov theorem corresponds to a positive measure. □

**Theorem 2.3.3.** *Let $\{\kappa_i\} \subset \mathcal{D}(\mathbb{R}^n)$ be a sequence of distributions. If $\lim_{i \in \mathbb{N}} \langle \kappa_i, \varphi \rangle$ exists for each $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ then*

$$\langle \kappa, \varphi \rangle = \lim_{i \in \mathbb{N}} \langle \kappa_i, \varphi \rangle$$

*defines a distribution.*

### Proof

The idea is to use the uniform boundedness principle for Frechet spaces. To get a uniform bound we consider the case for compact subsets $K \Subset \mathbb{R}^n$, and then exhaust $\mathbb{R}^n$ by compact sets. In this case the topology on $\mathcal{C}_c^\infty(K)$ coincides with the topology defined by the semi-norms

$$\|\varphi\|_\alpha = \sup_{x\in K} \big|\partial^\alpha \varphi(x)\big|.$$

With this topology $\mathcal{C}_c^\infty(K)$ is a metric space. By the mean value theorem,

$$\big\|f(y) - f(x) - g(x)(y-x)\big\| \leqq \|y-x\| \Big[\sup_{t\in[0,1]} \big\|g(x+t(y-x)) - g(x)\big\|\Big].$$

By induction, if $\varphi_i \to \varphi$ then $\partial^\alpha \varphi_i \to \partial^\alpha \varphi$. $\mathcal{C}_c^\infty(K)$ is complete and hence a Frechet space.

Let $\{\kappa_i\} \subset \mathcal{D}(K)$ be a sequence of distributions. Then for each $\kappa_i$ we have,

$$\big|\langle \kappa_i, \varphi\rangle\big| \leqq c_{\kappa_i} \sum_{|\alpha|\leqq k_i} \|\varphi\|_\alpha,$$

for $c_{\kappa_i}$ and $k_i$ depending on $i$.

The restriction of $\kappa_i$ to $\mathcal{C}_c^\infty(K)$ gives a continuous linear functional on $\mathcal{C}_c^\infty(K)$. For all $\varphi \in \mathcal{C}_c^\infty(K)$ by definition of distributions, there exist $c_{\kappa_i}$ and $k_i$ such that or $\{\kappa_i\}$ is pointwise bounded on $\mathcal{C}_c^\infty(K)$. Hence by uniform boundedness principle

$$\varphi \mapsto \sup_{\kappa_i} \big\{\langle \kappa_i, \varphi\rangle\big\}$$

is a semi-norm.

We have to show that if $\{\kappa_i\} \subset \mathcal{D}(\mathbb{R}^n)$ is such that $\langle \kappa, \varphi\rangle := \lim_i \langle \kappa_i, \varphi\rangle$ exists for all $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$, then $\kappa$ defines a distribution. For

**Theorem 2.3.4.** *Given $\kappa \in \mathcal{D}(\mathbb{R}^n)$ there exists a sequence of test functions $\{\varphi_i\}_{i\in\mathbb{N}} \subset \mathcal{C}_c^\infty(\mathbb{R}^n)$ such that*

$$\langle \kappa, \varphi \rangle = \lim_{i\in\mathbb{N}} \int_{\mathbb{R}^n} \varphi_i(x)\varphi(x)d\lambda(x).$$

### 2.3.2 | Calculus of Distributions

We have $L^1_{\text{loc}}(\mathbb{R}^n) \subseteq \mathcal{D}(\mathbb{R}^n)$, hence $\mathcal{C}^\infty_c(\mathbb{R}^n) \subseteq \mathcal{D}(\mathbb{R}^n)$ since every test function is integrable. Since every distribution can be written as a limit of test functions, the operations on test functions extend to distributions. we can define operations on distributions Operations on test function can be extended to distributions. Each locally integrable function $f$ gives rise to a distribution of the form,

$$\langle f, \varphi \rangle \equiv \int_{\mathbb{R}^n} f(x)\varphi(x)d\lambda(x).$$

#### 2.3.2.1 | Adjoint Identities

The other method of definition operations on distributions is to define the operation on test functions, and

#### 2.3.2.2 | Differential Operators

Let $f$ be a continuous function and $\varphi$ be a compactly supported function. Then $x \mapsto \int_{\mathbb{R}^n} f(y)\varphi(y)d\lambda(y)$ is a constant function its derivative must be zero. By the translation invariance of the Lebesgue measure, we get

$$\int_{\mathbb{R}^n} \Big(\frac{\partial f(x)}{\partial x_j}\Big)\varphi(x)d\lambda(x) = -\int_{\mathbb{R}^n} f(x)\Big(\frac{\partial \varphi(x)}{\partial x_j}\Big)d\lambda(x).$$

Hence for locally integrable functions we have,

$$\big\langle \partial^\alpha f, \varphi \big\rangle = (-1)^{|\alpha|}\big\langle f, \partial^\alpha \varphi \big\rangle.$$

For any distribution $\kappa$ define $\partial^\alpha \kappa$ to be the unique continuous linear functional on the space of compactly supported smooth functions such that

$$\langle \partial^\alpha \kappa, \varphi \rangle = (-1)^{|\alpha|}\langle \kappa, \partial^\alpha \varphi \rangle,$$

for every $\varphi \in \mathcal{C}^\infty_c(\mathbb{R}^n)$.

Similarly, let $f$ be a locally integrable function and $g$ be a smooth function. Then for any compactly supported smooth function we have,

$$\int_{\mathbb{R}^n} \big(g(x)f(x)\big)\varphi(x)d\lambda(x) = \int_{\mathbb{R}^n} f(x)\big(g(x)\varphi(x)\big)d\lambda(x)$$

Hence for any distribution $\kappa$ and smooth function $g$, define $g \cdot \kappa$ to be the unique linear functional on the space of test functions given by the rule,

$$\langle g \cdot \kappa, \varphi \rangle = \langle \kappa, g \cdot \varphi \rangle.$$

### 2.3.3 | Peetre's Theorem

### 2.3.4 | The Sheaf of Distributions

Let $V \subset U \subset \mathbb{R}^n$. Note that any compact subset of $V$ is also a compact subset of $U$. Hence the space of compactly supported smooth functions $V$ is contained in that of $U$.

$$\mathcal{C}_c^\infty(V) \subset \mathcal{C}_c^\infty(U).$$

Let $\kappa$ be a distribution on $U$. It is a continuous linear functional on the space $\mathcal{C}_c^\infty(U)$. Since $\mathcal{C}_c^\infty(V)$ is a subspace of $\mathcal{C}_c^\infty(U)$, the linear map can be restricted to the subspace. The restriction of $\kappa$ to the open set $V$ denoted by $\kappa\big|_V$ is defined by its action,

$$\langle \kappa\big|_V, \varphi\rangle = \langle \kappa, \varphi\rangle,$$

for all $\varphi \in \mathcal{C}_c^\infty(V) \subset \mathcal{C}_c^\infty(U)$.

Let $\mathcal{O}(\mathbb{R}^n)$ be the category of open sets of $\mathbb{R}^n$ with inclusion maps as morphisms. The space of distributions then corresponds to a contravariant functor,

$$\mathcal{D} : \mathcal{O}(\mathbb{R}^n)^{\mathrm{op}} \to \mathbf{Sets}$$

which assigns to each open set $U$, the set $\mathcal{D}(U)$ of distributions over $U$. Where the inclusion map $V \subset U$ is sent to the restriction map, $\mathcal{D}(V \subset U) \equiv |_V : \mathcal{D}(U) \to \mathcal{D}(V)$. Hence $\mathcal{D}$ is a pre-sheaf.

As the next theorem will show distributions can also be patched up if they agree on intersections, that is to say if there is a collection of distributions $\{\kappa_i\}_{i\in F}$ defined on open sets $\{U_i\}_{i\in F}$ such that $\kappa_i|_{U_i\cap U_j} = \kappa_j|_{U_i\cap U_j}$ for all $i, j$, then there exists a unique distribution $\kappa$ such that $\kappa_i = \kappa|_{U_i}$.

This can diagramatically formulated as follows. Each cover $\{U_i\}_{i\in F}$ gives rise to a map,

$$e : \kappa \mapsto (\kappa|_i)_{i\in F} \in \prod_i \mathcal{D}(U_i).$$

and the restriction to the intersections gives us two more maps

$$\prod_{i\in F} \mathcal{D}(U_i) \overset{q}{\underset{p}{\rightrightarrows}} \prod_{i,j} \mathcal{D}(U_i \cap U_j).$$

given by,

$$p(\textstyle\prod_i \kappa_i) = \prod_{i,j} \kappa_i|_{U_i\cap U_j}, \qquad q(\prod_i \kappa_i) = \prod_{j,i} \kappa_i|_{U_i\cap U_j}.$$

If $\kappa_i$ can be patched up to a single distribution, it means that whenever $\kappa_i|_{U_i\cap U_j} = \kappa_j|_{U_i\cap U_j}$ we have a patched up distribution $\kappa$ such that $\kappa|_{U_i} = \kappa_i$. Hence the composite diagram,

$$\mathcal{D}(U) \overset{e}{\dashrightarrow} \prod_i \mathcal{D}(U_i) \overset{q}{\underset{p}{\rightrightarrows}} \prod_{i,j} \mathcal{D}(U_i \cap U_j).$$

is an equaliser, or $\mathcal{D}$ is a sheaf.

**Theorem 2.3.5.** *$\mathcal{D}$ is a sheaf. That is, for every cover $\{U_i\}_{i\in F}$ of $U$, the diagram*

$$\mathcal{D}(U) \overset{e}{\dashrightarrow} \prod_i \mathcal{D}(U_i) \overset{q}{\underset{p}{\rightrightarrows}} \prod_{i,j} \mathcal{D}(U_i \cap U_j).$$

*is an equaliser.*

**Proof**

Let $\kappa$ be a distribution on $\mathbb{R}^n$ and let $U \subset \mathbb{R}^n$ be such that for every $x \in U$ there exists a neighborhood $U_x$ such that $\kappa|_{U_x} \equiv 0$. For any $\varphi \in \mathcal{C}_c^\infty(U_x)$ we have, $\langle \kappa, \varphi \rangle = 0$. Let $\varphi \in \mathcal{C}_c^\infty(U)$, then the support of $\varphi$ belongs to $U$ and it can be covered by neighborhoods $U_x$ as above. Since the support is compact there exists a finite subcover $\{U_{x_i}\}_{i \in F}$. Choose a partition of unity $\psi_i$ for this finite subcover. If we let $\varphi_i = \psi_i \varphi$ we have

$$\varphi = \textstyle\sum_{i \in F} \varphi_i.$$

where support of each $\varphi_i$ is contained in $U_{x_i}$. By assumption we have $\langle \kappa|_U, \varphi_i \rangle = \langle \kappa|_{U_{x_i}}, \varphi_i \rangle = 0$ and hence we have

$$\langle \kappa|_U, \varphi \rangle = \textstyle\sum_{i \in F} \langle \kappa|_U, \varphi_i \rangle = 0.$$

This means that if the distribution is zero in all the subregions of the region then it must be zero. This indicates the local nature of distributions, and that the distribution can be constructed by patching up.

Let $U \subseteq \mathbb{R}^n$ and for all open cover $\{U_i\}$ of $U$ if $\kappa|_{U_i} = \tau|_{U_i}$ for all $i$ then and $\kappa \equiv \tau$ on $U$. On the intersections we have, $\kappa|_{U_i}|_{U_j} = \kappa|_{U_j}|_{U_i}$. If $\kappa_i \in \mathcal{D}(U_i)$ such that $\langle \kappa_i, \varphi \rangle = \langle \kappa_j, \varphi \rangle$ for every $\varphi \in \mathcal{C}_c^\infty(U_i \cap U_j)$ then it means that there exists a distrbution $\kappa \in \mathcal{D}(U)$ such that $\kappa_i = \kappa|_{U_i}$. The maps $\kappa_i \in \mathcal{D}(U_i)$ and $\kappa_j \in \mathcal{D}(U_j)$ represent the restriction of same distribution $\kappa$ if,

$$\kappa|_{U_i \cap U_j} = \kappa_i|_{U_j} = \kappa_j|_{U_i}.$$

If we have is an $F$-indexed family of distributions $(\kappa_i)_{i \in F} \in \prod_{i \in F} \mathcal{D}(U_i)$, and two maps

$$p(\textstyle\prod_i \kappa_i) = \prod_{i,j} \kappa_i|_{U_i \cap U_j}, \qquad q(\prod_i \kappa_i) = \prod_{j,i} \kappa_i|_{U_i \cap U_j}.$$

Note that the order of $i$ and $j$ is important here and thats what distinguishes the two maps. The above property of existence of the function $\kappa$ implies that $\kappa|_{U_j}|_{U_i \cap U_j} = \kappa|_{U_i}|_{U_i \cap U_j}$ which means that there is a map $e$ from $\mathcal{D}(U)$ to $\prod_i \mathcal{D}(U_i)$ such that $pe = pq$.

$$\mathcal{D}(U) \xrightarrow{e} \prod_i \mathcal{D}(U_i) \underset{q}{\overset{p}{\rightrightarrows}} \prod_{i,j} \mathcal{D}(U_i \cap U_j).$$

This is the collation property of sheaves. The map $e$ is called the equalizer of $p$ and $q$. Note that the existence of partition of unity was essential in this. Hence we have proved the theorem. □

# [[Finish the Proof]]

Page no. 43 Hormander.

**Theorem 2.3.6.** *The pre-sheaf $\mathcal{D}$ is also a sheaf on $\mathbb{R}^n$, that is to say, the diagram*

$$\mathcal{D}(U) \xrightarrow{e} \prod_i \mathcal{D}(U_i) \underset{q}{\overset{p}{\rightrightarrows}} \prod_{i,j} \mathcal{D}(U_i \cap U_j).$$

*where the maps $e, p, q$ are as above is an equaliser diagram.*

If $\kappa \in \mathcal{D}(\mathbb{R}^n)$ then the support of the distribution $\kappa$, denoted by $\operatorname{supp}(\kappa)$, is the set of all points in $\mathbb{R}^n$ having no open neighborhood to which the restriction of $\kappa$ is zero. Equivalently, $\mathbb{R}^n \setminus \operatorname{supp}(\kappa)$ is the open set of points such that $\kappa|_{\mathbb{R}^n \setminus \operatorname{supp}(\kappa)} \equiv 0$.

### 2.3.5 | Compactly Supported Distributions

**Theorem 2.3.7.** *The set of distributions in $\mathbb{R}^n$ with compact support is identical to the dual space of $\mathcal{C}^\infty$*

## 2.4 | Distributions on Manifolds

## 2.5 | Structure Theorems

# A | Review of Fréchet Spaces

A Fréchet space is a metrizable, complete locally convex vector space. If $F$ is a Fréchet space with a metric $d$ then for any two $x, y \in F$ let $d(x, y)$ be the distance between them, by choosing as radius a number smaller than half of this distance, we can construct two open sets around $x$ and $y$ that do not intersect. Hence every Fréchet space is Hausdorff. It is also locally compact because we can choose a neighborhood basis as a collection of balls, and around any point there exists a open set entirely contained in a compact set.

**Theorem 2.5.1. (Baire)** *If $F$ is a locally compact Hausdorff space, the intersection of countably many dense open sets of $F$ is dense in $F$.*

**Proof**

Let $\{U_i\}_{\mathbb{N}}$ be a countable collection of dense open sets. We have to show that every open set $U \subseteq F$ intersects $\bigcap_{\mathbb{N}} U_i$. It is sufficient to show for neighborhood basis. We need to show that for every neighborhood $B_r(x)$ of $x \in F$, $U_i \cap B_r(x)$ is non-empty for every $r \in (0, 1]$.

If $U$ is dense open set, for every open ball $B_r(x)$, $U \cap B_r(x)$ is non-empty and open. We need to show that $\bigcap_i U_i \cap B_r(x)$ is non-empty. Construct inductively open balls $B_{r_i}(x_i)$ such that

$$B_{r_{i+1}}(x_{i+1}) \subset U_i \cap B_{r_i}(x_i)$$

for all $i$, where $r_0 = 1$ and $x_0 = x$. Consider $\overline{B}_{r_i}(x_i)$, this gives us a nested collection of closed balls and since $F$ is locally compact, and Hausdorff their intersection is non-empty.

$$\overline{x} \in \bigcap_{i \in \mathbb{N}} B_{r_i}(x_i) \subset \Big(\bigcap_{i \in \mathbb{N}} U_i\Big) \cap B_r(x).$$

Hence $\bigcap_i U_i$ is dense in $F$. □

A set $C \subset F$ is said to be rare or nowhere dense if it has no interior, that is, the closure of $C$ does not contain any open set. Suppose $F$ is a complete metric space. Then by Baire's theorem $F$ cannot be written as a countable union of nowhere dense sets. Because otherwise, if

$$F = \bigcup_{i \in \mathbb{N}} C_i$$

with $C_i$ nowhere dense, then we can consider the countable collection of dense sets $\{F \backslash C_i\}$, the intersection of these sets would be dense, which cannot be true. This can be taken as an alternative statement of Baire's theorem.

**Theorem 2.5.2. (Baire)** *Complete metric space cannot be a countable union of rare sets.*

An important application of Baire's theorem is the Banach-Steinhaus theorem also known as the unform boundedness principle. Let $F$ be a Fréchet space and $Y$ be a locally convex vector space with a semi-norm $\| \cdot \|_Y$. Suppose we have a bounded linear map, $T : F \to Y$, since $T$ is bounded, the composition

$$F \xrightarrow{T} Y \xrightarrow{\|\cdot\|_Y} \mathbb{R}$$

defines a semi-norm on $F$ using $\|\cdot\|_Y$. Banach-Steinhaus theorem or uniform boundedness theorem allows us to construct such semi-norms[2] out of a collection of well behaved maps with a uniform bound.

**Theorem 2.5.3. (Banach-Steinhaus)** *Let $F$ be a Fréchet space, and $\|\cdot\|_Y$ be a semi-norm on a vector space $Y$. Let $\mathcal{S} \subseteq \operatorname{Hom}(F, Y)$ be such that for all $x \in F$,*

$$\sup_{T\in\mathcal{S}} \left\{\|Tx\|_Y\right\} < \infty,$$

*if and only if*

$$\|x\|_{Y,\mathcal{S}} \coloneqq \sup_{T\in\mathcal{S}} \left\{\|Tx\|_Y\right\}$$

*defines a semi-norm.*

### Proof

Note that $\|\cdot\|_{Y,\mathcal{S}}$ defined above will satisfy the semi-norm conditions because of the properties of sup. We have to make sure that it is defined whenever $\sup_{T\in\mathcal{S}}\left\{\|Tx\|_Y\right\} < \infty$ holds for all $x \in F$. We only have to show that $\|\cdot\|_{Y,\mathcal{S}}$ is well defined, that is for all $x \in F$, $\|x\|_{Y,\mathcal{S}} < \infty$ whenever $\sup_{T\in\mathcal{S}}\left\{\|Tx\|_Y\right\} < \infty$.

To prove this let

$$F_i = \{x \in F \mid \|Tx\|_Y \leq i \ \forall T \in \mathcal{S}\}.$$

Since $T$ are bounded linear maps, $F_i$ are closed sets. We have,

$$F_1 \subseteq F_2 \subseteq \cdots \subseteq F_i \subseteq F_{i+1} \subseteq \cdots, \ \bigcup_{i\in\mathbb{N}} F_i = F$$

Since $F$ is a complete metric space, not all $F_i$ can be rare or nowhere dense. One of these sets $F_k$ contains an open ball $B_r(y) \subset F_k$. Hence we have, $\|Tz\|_Y \leq k$ for all $z \in B_r(y)$. Every vector $x \in F$ can be reached by a line from $y$. For $\lambda = r/2\|x\|$,

$$x_\lambda \equiv y + \lambda x \in B_r(y)$$

Hence we have,

$$\big\|Tx\big\|_Y = \lambda^{-1}\big\|(Ty - tx_k)\big\|_Y \leq \lambda^{-1}\Big(\underbrace{\big\|Ty\big\|_Y - \big\|Tx_k\big\|_Y}_{2k}\Big)$$

Since $\lambda^{-1} = 2\|x\|/r$, we have,

$$\|Tx\| \leq \left(\tfrac{4k}{r}\right)\|x\|$$

Hence $T \in \mathcal{S}$ are uniformly bounded. The other side is clear. □

[2]Semi-norm are similar to norms in that they can be used to distinguish between vectors. However unlike norms they cannot distinguish between all vectors. The semi-norm for example could not detect a subspace of the vector space, and cannot distinguish points in some subspace. Semi-norms in this sense can be thought of as the subspace. An example would be the semi-norm, $(x_i)_i \mapsto |x_j|$. This can only detect the $j$th coordinate subspace.

**Theorem 2.5.4. (Banach-Alaoglu)** *$F$ is a normed space,*

$$B_1(\mathcal{A}^*) = \{\varphi \in \mathcal{A}^* \mid \|\varphi\| = 1\}$$

*is a compact Hausdorff space with respect to weak $^*$-topology.*

**Proof**

The idea is to embed $B_1(\mathcal{A}^*)$ in $Z = \overline{D}^{B_1(\mathcal{A})}$, where,

$$\overline{D}^{B_1(\mathcal{A})} = \prod_{i \in B_1(\mathcal{A})} \overline{D} = \{f : B_1(\mathcal{A}) \to \overline{D}\}.$$

which is compact because $\overline{D}$ is compact and by Tychonoff's theorem product of compact sets $\prod_{i \in B_1(\mathcal{A})} \overline{D}$ is compact.

If $\varphi \in B_1(\mathcal{A}^*)$, define a map, $F : B_1(\mathcal{A}^*) \to Z$, given by,

$$\varphi \mapsto \{\varphi(A)\}_{A \in B_1(\mathcal{A})} \in \overline{D}^{B_1(\mathcal{A})}$$

With $B_1(\mathcal{A}^*)$ equipped with the weak$^*$-topology, and $Z$ equipped with the product topology, the above map is continuous, i.e., if $\{\varphi_i\}_i \to \varphi$ if and only if $\{\varphi_i(A)\} \to \varphi(A)$ for all $A \in B_1(\mathcal{A})$. Since $F$ is 1-1, it maps homeomorphically onto its image in $Z$.

Since $Z$ is compact Hausdorff space, it's enough to show that range$(F)$ is closed in $Z$.

$$K_{A,B,\alpha,\beta} = \{f \in Z \mid f(\alpha A + \beta B) = \alpha f(A) + \beta f(B)\}$$

is a closed set. If $f$ is in the range of $F$ then there exists some $\varphi$ such that $\{\varphi(A)\}_{A \in B_1(\mathcal{A})} = f$. $f \in K_{A,B,\alpha,\beta}$ when defined.

$$\text{range}(F) = \bigcap_{A,B,\alpha,\beta} K_{A,B,\alpha,\beta}$$

and it's closed as it's the intersection of closed sets. $\square$

A topological space $X$ is said to be exhaustible by compact sets if there exists a sequence of compact subsets $\{K_i\}_{i\in F}$ such that,

$$K_1 \subset K_2^\circ \subset K_2 \subset \cdots \subset K_j^\circ \subset K_j \subset \cdots, \quad \bigcup_{i\in F} K_i^\circ = X.$$

If $X$ is locally compact, and $\sigma$-compact, then it is exhaustible by compact sets. Instead of proving this general result we will provide an example of an exhaustion by compact sets for non-empty open subset of $\mathbb{R}^n$. Define,

$$K_i = \{x \in X \mid |x| \leqq j, d(x, X^c) \geq \tfrac{1}{j}\}. \tag{exhaustion}$$

Clearly this set

This topology is however metrizable. Since $X$ locally compact, and $\sigma$-compact we can exhaust $X$ by compact sets. $K_i \subseteq K_{i+1}$ with $K_i \subset \mathrm{int}(K_{i+1})$ such that $\bigcup_i K_i = X$. Using the semi-norms $\|\cdot\|_{K_i}$ a metric can be constructed by

$$d(x,y) = \textstyle\sum_i \frac{1}{2^i} \frac{\|x-y\|_{K_i}}{1+\|x-y\|_{K_i}}.$$

The topology coincides with the topology of compact convergence because every compact set $K$ is a subset of one of these $K_i$s. Since it is a metric space it is sufficient to work with nets indexed by countable sets as every metric space is first countable. Together with this metric $\mathcal{C}(\mathbb{R}^n)$ is a Frechet space.[3] Once we have the Frechet space structure, certain theorems, such as uniform boundedness principle of analysis are applicable.

Topology on the space $\mathcal{C}_c^\infty(\mathbb{R}^n)$ of test functions can be built using the topology of compact convergence on the space of continuous functions. In this case, when we say two functions to be close to each other we expect their derivatives to be also close to each other.

We say a sequence of test functions $\{f_i\}$ converges to $f$ if there exists a semi-norm $\|\cdot\|_K$ such that,

$$\left\|\partial^\alpha f_i - \partial^\alpha f\right\|_K \to 0.$$

for all $\alpha$. This topology can also be thought of as being generated by a new collection of semi-norms defined by,

$$\|f\|_{(\alpha,K)} \coloneqq \|\partial^\alpha f\|_K.$$

We will denote this collection of semi-norms by $\mathcal{T}_\partial$ and call the topology it generates the natural topology on $\mathcal{C}_c^\infty(\mathbb{R}^n)$.

[3] we can define this topology for any open set $U \subset \mathbb{R}^n$, and $\mathcal{C}(U)$ is a Frechet space. So, the sheaf of algebras of continuous functions on $\mathbb{R}^n$ has extra structure, and can be thought of as a sheaf of Frechet spaces.

# 3 | Convolutions & Tensor Products

## 3.1 | Convolution of Distributions

We can define convolution of two continuous functions $f$ and $g$ when either one has a compact support. Using the pairing this can be extended to distributions. For $f$ and $g$ continuous, with atleast one of them compactly supported, for any $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ we have,

$$\langle f * g, \varphi \rangle = \iint_{\mathbb{R}^n} f(x)g(y)\varphi(x+y)d\lambda(x)d\lambda(y) = \langle f, g * \varphi \rangle.$$

is defined as

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

### 3.1.1 | with Smooth Functions

### 3.1.2 | with Distributions

#### 3.1.2.1 | The Theorem of Supports

## 3.2 | Distributions in Product Spaces

Let $\varphi \in \mathcal{C}_c^\infty(\mathcal{M} \times \mathcal{N})$ such that the support of $\varphi$ is contained in $K_{\mathcal{M}} \times K_{\mathcal{N}}$. Then for fixed $y \in \mathcal{N}$, $\varphi(x, y) \in \mathcal{C}_c^\infty(\mathcal{M})$. For any distributions $\kappa \in \mathcal{D}(\mathcal{M})$ we have,

$$y \mapsto \langle \kappa, \varphi(\cdot, y) \rangle \in \mathcal{C}_c^\infty(\mathcal{N}).$$

Hence for any $\nu \in \mathcal{D}(\mathcal{N})$ by definition of destribution we have,

$$\big| \langle \nu(y), \langle \kappa(x), \varphi(x, y) \rangle \rangle \big| \leq c_\nu \sum_y$$

### 3.2.1 | Tensor Products

### 3.2.2 | The Kernel Theorem

# 4 | The Fourier Transformation

Continuous functions on a space contain information about the space. Characters contain the same information at each point of the space. The idea of Fourier transform of a function is to adjust the weightage of characters at each point so that the information content of the weighted superposition of characters matches the information contained in the function. In this sense, the Fourier transform decomposes a given function into a continuous family of normalised characters.

## 4.1 | Normalised Characters on $\mathbb{R}^n$

The space of interest to us is $\mathbb{R}^n$ which acts on itself by translation. Characters are functions which are eigenfunctions for translations or transform under translation by multiplication by a factor. A character $c$ on $\mathbb{R}^n$ is such that for every $y \in \mathbb{R}^n$, $c(x+y) = k(y)c(x)$, for all $x \in \mathbb{R}^n$. $c$ is completely determined by $k$ once $c(0)$ is known.

A character $c$ is said to be normalized if $c(0) = 1$. In which case, we have $c(x) = c(x+0) = k(x)c(0) = k(x)$. We will denote normalised characters by $\xi$. For normalised characters, we have

$$\xi(x+y) = \xi(x)\xi(y).$$

The characters should be expected to have good behavior under convolutions as their value at translations is given by product of its value at these points. Suppose $\xi$ is a continuous normalized character, and $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ such that

$$\int_{\mathbb{R}^n} \xi(-y)\varphi(y)d\lambda(y) = 1$$

then,

$$\xi(x) = \int_{\mathbb{R}^n} \xi(x)\xi(-y)\varphi(y)d\lambda(y) = \int_{P_f} \xi(x-y)\varphi(y)d\lambda(y) = \xi * \varphi(x).$$

where $\lambda$ is the Lebesgue measure and $P_\varphi$ is the support of $\varphi$. So, $\xi$ can be expressed as a convolution with the test function $\varphi$. Since convolutions with test functions are smooth functions the continuous normalised character $\xi$ is smooth, $\xi \in \mathcal{C}^\infty(\mathbb{R}^n)$. If $e_j$ is an orthonormal basis on $\mathbb{R}^n$,

$$\xi(x) = \xi\big(\textstyle\sum_j x_j e_j\big) = \displaystyle\prod_j \xi_j(x_j),$$

where $\xi_j(x_j) \coloneqq \xi(x_j e_j)$. $\xi_j$ are continuous functions, and are themselves continuous characters on $\mathbb{R}$. Since any continuous character can be written as a convolution with a suitable test function, $\xi_j$s are smooth functions. Differentiating, we get,

$$\frac{d\xi_j(x_j+y_j)}{dy_j} = \frac{d\big(\xi_j(x_j)\xi_j(y_j)\big)}{dy_j} = \xi_j(x_j)\Big[\frac{d\xi_j(y_j)}{dy_j}\Big].$$

At $y_j = 0$, we have
$$\Big[\frac{d\xi_j(x_j)}{dy_j}\Big] = \xi_j(x_j)\Big[\frac{d\xi_j(0)}{dy_j}\Big] = -i\chi_j\xi_j(x_j).$$
So, the value of the derivative at each point is also determined by its value at origin. The choice $-i\chi_j$ is used for notational convenience for later.

The unique differentiable function that satisfies this ordinary differential equation is the exponential function, so, we have $\xi_j(x_j) = \xi_j(0)e^{-ix_j\chi_j}$. If $\xi$ is a normalised continuous character, then $\xi_i(0) = 1$. So we have

$$\xi_j(x_j) = e^{-ix_j\chi_j}.$$

Any continuous normalised character on $\mathbb{R}^n$ is of the form,

$$\xi(x) = \prod_i \xi_i(x_i) = e^{-i\sum_j x_j\chi_j} := e^{-i\langle x,\chi\rangle}.$$

The collection of all continuous normalised characters on $\mathbb{R}^n$ is an abelian group, and we denote it by $\widehat{\mathbb{R}}^n$. It is a locally compact abelian group. We have an isomorphism of locally compact abelian groups which assigns to each $\chi \in \mathbb{C}^n$ a continuous normalised character. We have proved the following,

**Lemma 4.1.1.** *Any continuous character on $\mathbb{R}^n$ is of the form $x \mapsto e^{-i\langle x,\chi\rangle}$ for $\chi \in \mathbb{C}^n$.*

Since we are interested in studying functions and distributions that are fairly well-behaved at $\infty$ we would only need bounded characters. If $i\chi_j$ is not purely complex, then $e^{-ix_j\chi_j}$ is unbounded. If the character is bounded then $\chi_j$ will have to be real.

**Lemma 4.1.2.** *Any bounded character on $\mathbb{R}^n$ is of the form $x \mapsto e^{-i\langle x,\chi\rangle}$ for $\chi \in \mathbb{R}^n$.*

Denote the collection of continuous bounded normalised characters by $(\mathbb{R}^n)'$. The above discussion gives us an isomorphism, $e^{-i\langle\cdot,\cdot\rangle} : \mathbb{R}^n \to (\mathbb{R}^n)'$ which sends $\chi \in \mathbb{R}^n$ to the character $e^{-i\langle\cdot,\chi\rangle}$. If the boundedness requirement is removed, then we have a character for each $\chi \in \mathbb{C}^n$. As locally compact abelian groups we have $\mathbb{R}^n \cong (\mathbb{R}^n)' \subset \mathbb{C}^n \cong \widehat{\mathbb{R}}^n$.

### 4.1.1 | Fourier Transform on $L^1(\mathbb{R}^n)$

The idea of Fourier analysis is to exploit the underlying symmetry of the space to study functions. This is done by decomposing the functions of interest as sums or integrals of functions that transform in simple ways under the action of the underlying symmetry group.

If $f$ is an integrable function on $\mathbb{R}^n$ then its decomposition in terms of characters is,

$$\mathcal{F}f(\chi) \equiv \widehat{f}(\chi) := \int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} f(x)d\lambda(x).$$

The pairing $\langle x, \chi\rangle$ allows for easier manipulations. One starts with a function, works with its Fourier transform, and inverts back the manipulated function via the Fourier inversion formula give by

$$\mathcal{G}g(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{i\langle x,\chi\rangle} g(\chi)d\lambda(\chi).$$

By linearity of integration we have,

$$\mathcal{F}(f+g)(\chi) = \mathcal{F}f(\chi) + \mathcal{F}g(\chi),\ \mathcal{F}(\lambda f)(\chi) = \lambda \mathcal{F}f(\chi).$$

Hence $\mathcal{F}$ is a linear map. Similarly, $\mathcal{G}$ is also a linear map.

#### 4.1.1.1 | Basic Properties $\mathcal{F}$ and $\mathcal{G}$

Fourier transforms are useful because they let us exchange certain properties of functions. This allows us to use different tools available for studying different properties.

**Theorem 4.1.3.** *If $f, g$ are integrable then*

$$\int (\mathcal{F}f)g = \int f(\mathcal{F}g).$$

**Proof**

For any integrable function $f \in L^1(\mathbb{R}^n)$, we have

$$\|\mathcal{F}f\|_\infty \leqq \sup_\chi \left| \int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} f(x) d\lambda(x) \right| \leqq \int_{\mathbb{R}^n} |e^{-i\langle x,\chi\rangle}||f(x)| d\lambda(x) = \|f\|_{L^1}.$$

This implies,

$$\int_{\mathbb{R}^n} \big|(\mathcal{F}f)(x)g(x)\big| d\lambda(x) \leqq \int_{\mathbb{R}^n} \|f\|_{L^1} \big|g(x)\big| d\lambda(x) = \|f\|_{L^1} \|g\|_{L^1}.$$

So, the product $\mathcal{F}f \cdot g$ is an integrable function. So, we can apply Fubini's theorem, gives us the following theorem,

$$\begin{aligned}
\int_{\mathbb{R}^n} (\mathcal{F}f)(x)g(x)d\lambda(x) &= \int_{\mathbb{R}^n} \left[ \int_{\mathbb{R}^n} e^{-i\langle \chi,x\rangle} f(\chi) d\lambda(\chi) \right] g(x) d\lambda(x) \\
&= \int_{\mathbb{R}^n} f(\chi) \left[ \int_{\mathbb{R}^n} e^{-i\langle \chi,x\rangle} g(x) d\lambda(x) \right] d\lambda(\chi) \\
&= \int_{\mathbb{R}^n} f(x)(\mathcal{F}g)(x) d\lambda(x).
\end{aligned}$$

□

Fourier transform exchanges the differentiability properties of a function with growth properties. If $f$ is differentiable then $\mathcal{F}f$ decays, and conversely, if $f$ decays then $\mathcal{F}f$ is differentiable. This is formalised by the following theorem,

**Theorem 4.1.4. (Exchange Formulas)** *If $f$ and its derivatives are integrable then,*

$$\begin{aligned}
\mathcal{F}(\partial f/\partial x_i)(\chi) &= i\chi_j\big(\mathcal{F}f(\chi)\big) \\
\big(\partial(\mathcal{F}f)/\partial \chi_i\big)(\chi) &= i\mathcal{F}(x_j f)(\chi)
\end{aligned}$$

*and for the Fourier inversion formula,*

$$\begin{aligned}
\mathcal{G}(\partial g/\partial \chi_i)(x) &= i x_j \big(\mathcal{G}g\big)(x) \\
i\mathcal{G}(\chi_j g)(x) &= \big(\partial(\mathcal{G}g)/\partial x_j\big)(x).
\end{aligned}$$

**Proof**

Using the fact that the integral $\int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} f(x) d\lambda(x)$ is a constant function in $x$ and by taking the differentiation inside the integral sign, and using the translation invariance of the Lebesgue measure $\lambda$, we have

$$\int_{\mathbb{R}^n} (\partial/\partial x_i)\big(e^{-i\langle x,\chi\rangle} f(x)\big) d\lambda(x) = (\partial/\partial x_i) \int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} f(x) d\lambda(x) = 0.$$

By the product rule we have $\int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} \big[\big(\partial f(x)/\partial x_j\big) - i\chi_j f(x)\big] d\lambda(x) = 0$. This gives us,

$$\begin{aligned}
\mathcal{F}(\partial f/\partial x_i)(\chi) &= \int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} \big(\partial f/\partial x_j\big)(x) d\lambda(x) \\
&= i\chi_j \int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} f(x) d\lambda(x) = i\chi_j \big(\mathcal{F} f(\chi)\big).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\big(\partial \mathcal{F} f/\partial \chi_j\big)(\chi) &= (\partial/\partial \chi_j)\Big[\int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} f(x) d\lambda(x)\Big] \\
&= \int_{\mathbb{R}^n} \big(\partial e^{-i\langle x,\chi\rangle}/\partial \chi_j\big) f(x) d\lambda(x) \\
&= i \int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} \big(x_i f(x)\big) d\lambda(x) = i\mathcal{F}(x_j f)(\chi).
\end{aligned}$$

The exchange formulas for Fourier inversion follow similarly. □

Characters convert the addition on the space they act to multiplication. If $f, g \in L^1(\mathbb{R}^n)$, recall that the convolution of $f$ and $g$ is defined by,

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x-y) g(y) d\lambda(y).$$

Since products of measurable functions are measurable it follows that $f(x-y)g(y)$ is measurable. By Tonelli's theorem and by the definition of convolution, we have

$$\begin{aligned}
\int_{\mathbb{R}^{2n}} \big|f * g(y)\big| d\lambda(y) &= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} f(x-y) g(y) d\lambda(x) \right| \lambda(y) \\
&= \int_{\mathbb{R}^n} \left[ \int_{\mathbb{R}^n} \big|f(x-y)\big|\big|g(y)\big| d\lambda(x) \right] \lambda(y) \\
&= \underbrace{\left[ \int_{\mathbb{R}^n} \big|f(x)\big| d\lambda(x) \right]}_{\|f\|_{L^1}} \underbrace{\left[ \int_{\mathbb{R}^n} \big|g(y)\big| d\lambda(y) \right]}_{\|g\|_{L^1}} < \infty.
\end{aligned}$$

Hence it follows that,

$$\big\|f * g\big\|_{L^1} = \int_{\mathbb{R}^n} \big|f * g(x)\big| d\lambda(x) \leqq \|f\|_{L^1} \|g\|_{L^1}.$$

$L^1$ space together with $\|\cdot\|_{L^1}$ is a Banach space, and the above inquality says that with convolution as the product it is also a Banach algebra. Fourier transforms allows us to exchange convolution and product properties, formalised by the following theorem,

**Theorem 4.1.5.** *If $f, g$ are $L^1$ functions, then,*

$$\widehat{f * g}(\chi) = \widehat{f}(\chi) \cdot \widehat{g}(\chi)$$

**Proof**

By Fubini's theorem $e^{-i\langle x,\chi\rangle} f(x-y)g(y)$ is integrable, and we have,

$$\begin{aligned}
\widehat{f * g}(\chi) &= \int_{\mathbb{R}^n} \left[ e^{-i\langle x,\chi\rangle} f(x-y)g(y)d\lambda(y) \right] d\lambda(x) \\
&= \int_{\mathbb{R}^n} \left[ \int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} f(x-y)d\lambda(x) \right] g(y)d\lambda(y) \\
&= \int_{\mathbb{R}^n} \left[ \int_{\mathbb{R}^n} e^{-i\langle z,\chi\rangle} f(x-y)d\lambda(z) \right] e^{-i\langle y,\chi\rangle} g(y)d\lambda(y) = \widehat{f}(\chi)\widehat{g}(\chi).
\end{aligned}$$

□

The other product-convolution exchange,

$$\widehat{(f \cdot g)}(\chi) = (2\pi)^{-n} \widehat{f} * \widehat{g}(\chi)$$

require more tools to prove, or the proof Fourier inversion formula, we will hence postpone it for later. These properties are useful and it is natural to try to extend Fourier transform to distributions, which allow us to study a wider range of functions.

## 4.2 | Schwartz Functions

The idea for defining Fourier transform on distributions is to define it for test functions, and define Fourier transform of distributions as the adjoint of its action on test functions via the pairing, $\langle \mathcal{F}\kappa, \varphi \rangle = \langle \kappa, \mathcal{F}\varphi \rangle$. For this to make sense, the term $\langle \kappa, \mathcal{F}\varphi \rangle$ should be well defined. In particular, $\mathcal{F}\varphi$ should be a test function. So, the space of test functions should be invariant under Fourier transform.

**Lemma 4.2.1.** *There exist some $\varphi \in \mathcal{C}_c^\infty(\mathbb{R})$ such that $\mathcal{F}(\varphi) \notin \mathcal{C}_c^\infty(\mathbb{R})$.*

**Proof**

Let $\varphi$ be a smooth function with compact support. Without loss of generality assume the support is contained in the closed interval $[-\epsilon, \epsilon]$. The Fourier transform of $\varphi$ is given by,

$$\begin{aligned}
\mathcal{F}(\varphi)(\chi) = \int_{-\epsilon}^{\epsilon} e^{-i\langle x,\chi\rangle} \varphi(x) d\lambda(x) &= \int_{-\epsilon}^{\epsilon} \left[ \textstyle\sum_k (-i\langle x,\chi\rangle)^k (k!)^{-1} \right] \varphi(x)d\lambda(x) \\
&= \textstyle\sum_k \underbrace{\left[ (-i)^k (k!)^{-1} \displaystyle\int_{-\epsilon}^{\epsilon} x^k \varphi(x) d\lambda(x) \right]}_{c_k} \chi^k.
\end{aligned}$$

Hence it follows that,

$$\mathcal{F}(\varphi)(\chi) = \textstyle\sum_k c_k \chi^k$$

We also have,

$$|c_k| = (k!)^{-1}\left|\int_{-\epsilon}^{\epsilon} x^k\varphi(x)d\lambda(x)\right| \leqq (k!)^{-1}\|\varphi\|_{\text{sup}}\left|\int_{-\epsilon}^{\epsilon} x^k d\lambda(x)\right| \leqq 2(k!)^{-1}\epsilon^{k+1}\|\varphi\|_{\text{sup}}.$$

By the ratio test, the radius of convergence of $\sum_k c_k\chi^k$ is given by

$$\lim_k |c_k|/|c_{k+1}| = 2\epsilon \lim_{k\to\infty}(k+1) = \infty$$

So, the Fourier transform of $\varphi$ can be written as a power series with infinite radius of convergence. So, the Fourier transform of $\varphi$ does not have a compact support, or $\mathcal{F}(\varphi) \notin \mathcal{C}_c^\infty(\mathbb{R})$. $\square$

So, the space of compactly supported smooth functions cannot be used to define Fourier transform on distributions. The reason for the failure is that the radius of convergence of the Fourier transform could not be bounded.

We need a space of functions larger than $\mathcal{C}_c^\infty(\mathbb{R}^n)$ such that the Fourier transform also belongs to it. We need to find the small enough space of functions which contains $\mathcal{C}_c^\infty(\mathbb{R}^n)$ and remains closed under Fourier transforms. We also need to be able to use the exchange formulas, so we should require the space of functions to be such that if $f$ belongs to it, then so must $x^\beta\partial^\alpha f$. The smallness here is needed to ensures that we can define Fourier transform on a dense subset of the space of distributions. Since the purpose of theory of distributions is to study differentiability properties, we should first require the new space of test functions to have the nice differentiability properties.

### 4.2.1 | The Schwartz Space $\mathcal{S}(\mathbb{R}^n)$

We want the space of test functions to be such that if $f$ is a test function so is $x^\beta\partial^\alpha f$, and it is closed under Fourier transforms. Such functions are called rapidly decreasing functions (decrease faster than any polynomial function) or Schwartz functions.

A smooth function $f$ is said to be a Schwartz function if

$$\sup_x \left|x^\beta\partial^\alpha\big(f(x)\big)\right| < \infty,$$

for all multi-indices $\alpha$ and $\beta$. The collection of all Schwartz functions is denoted by $\mathcal{S}(\mathbb{R}^n)$.

The linearity of derivatives gives us $x^\beta\partial^\alpha(f+g) = x^\beta\partial^\alpha f + x^\beta\partial^\alpha g$. Using the fact that $|\cdot|$ satisfies triangle inequality, we have $|x^\beta\partial^\alpha f + x^\beta\partial^\alpha g| \leqq |x^\beta\partial^\alpha f| + |x^\beta\partial^\alpha f|$, hence, $\sup_x \left|x^\beta\partial^\alpha(f+g)\right| < \infty$. Similarly, $\sup_x \left|x^\beta\partial^\alpha\lambda f(x)\right| < |\lambda|\sup_x \left|x^\beta\partial^\alpha\big(f(x)\big)\right| < \infty$. Hence we have,

$$f, g \in \mathcal{S}(\mathbb{R}^n) \Rightarrow \lambda f,\ f+g \in \mathcal{S}(\mathbb{R}^n).$$

Hence $\mathcal{S}(\mathbb{R}^n)$ is a vector space. To build a topology on $\mathcal{S}(\mathbb{R}^n)$, for every multi-indices $\alpha$ and $\beta$ let $\|\cdot\|_{\alpha,\beta}$ be defined by,

$$\|f\|_{\alpha,\beta} \coloneqq \sup_x \left|x^\beta\partial^\alpha\big(f(x)\big)\right|$$

for every multi-indices $\alpha$ and $\beta$. It follows that,

$$\|\lambda f\|_{\alpha,\beta} = |\lambda|\ \|f\|_{\alpha,\beta},$$

$$\|f+g\|_{\alpha,\beta} \leqq \|f\|_{\alpha,\beta} + \|g\|_{\alpha,\beta}.$$

Hence $\|\cdot\|_{\alpha,\beta}$ is a semi-norm for every multi-indices $\alpha$ and $\beta$. Together with the topology generated by these semi-norms the Schwartz space is a locally convex vector space. Since there are only countably many $\alpha$s and $\beta$s, the collection $\{\|\cdot\|_{\alpha,\beta}\}$ is a countable and hence the topology generates is metrizable. It is hence separable, that is, it has a countable dense set. The Schwartz space is a metrizable locally convex vector space.

By taking $\alpha = \beta = 0$ we also have that every Schwartz function is bounded. If $\varphi$ is a compactly supported smooth function then $x^\beta \partial^\alpha \varphi(x)$ also has compact support and hence,

$$\mathcal{C}_c^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n).$$

If $f$ is a Schwartz function, since $f$ decreases rapidly the value of $x^\beta \partial^\alpha f$ decreases fast enough so that it is integrable for every multi-indices $\alpha$ and $\beta$. To prove this let $K$ be the closed unit ball around the origin $0 \in \mathbb{R}^n$.

Since $f$ is a Schwartz function there exists $M_{\alpha,\beta}$ such that,

$$\left|x^\beta\right|\left|\partial^\alpha f(x)\right| \leqq \left\|x^\beta \partial^\alpha f\right\| \leqq M_{\alpha,\beta}.$$

Now $\mathbb{R}^n$ is the disjoint union of $K$ and $K^c$, and hence we have,

$$\begin{aligned}\int_{\mathbb{R}^n} \left|x^\beta \partial^\alpha f(x)\right| d\lambda(x) &= \left[\int_K + \int_{K^c}\right] \left|x^\beta \partial^\alpha f(x)\right| d\lambda(x) \\ &= \underbrace{\left[\int_K \left|x^\beta \partial^\alpha f(x)\right| d\lambda(x)\right]}_{\lambda_{\alpha,\beta}(K)} + \underbrace{\left[\int_{K^c} \left|x^\beta \partial^\alpha f(x)\right| d\lambda(x)\right]}_{\lambda_{\alpha,\beta}(K^c)}.\end{aligned}$$

Since $K$ is the closed unit cube its Lebesgue measure is $\lambda(K) = 1$.

**Lemma 4.2.2.** *For every $f \in \mathcal{S}(\mathbb{R}^n)$ the functions $x^\beta \partial^\alpha f$ is integrable for all $\alpha$ and $\beta$.*

**Proof**

Since $f$ is a Schwartz function $|x^\beta \partial^\alpha f|$ is bounded by $M_{\alpha,\beta}$, we have,

$$\lambda_{\alpha,\beta}(K) = \int_K \left|x^\beta \partial^\alpha f(x)\right| d\lambda(x) \leqq M_{\alpha,\beta} \cdot \underbrace{\left|\int_K d\lambda(x)\right|}_{\lambda(K)=1}.$$

To show the second term is bounded the idea is to use the fact that $r^{-2}$ is integrable and use polar coordinates to simplify the integral.

Let $\beta'$ be such that $|x^{\beta'}| > |x|^{n+1}\,|x^\beta|$ for $x \in K^c$. Then we have,

$$\begin{aligned}\lambda_{\alpha,\beta}(K^c) &= \int_{K^c} \left|x^\beta \partial^\alpha f(x)\right| d\lambda(x) \\ &\leqq \int_{K^c} \left[\left|x^{\beta'} \partial^\alpha f(x)\right| / |x|^{n+1}\right] d\lambda(x) \\ &= M_{\alpha,\beta'} \cdot \int_{K^c} d\lambda(x)/|x|^{n+1} < M_{\alpha,\beta'} c.\end{aligned}$$

here $c = \sigma(S^{n-1})\rho(r \geq 1)$ where $\sigma$ is the unique measure on $S^{n-1}$ invariant under rotations and $\rho(E) = \int_E r^{n-1} dr$. Hence we have

$$\int_{\mathbb{R}^n} |x^\beta \partial^\alpha f(x)| d\lambda(x) < \infty$$

$\square$

**Theorem 4.2.3.** *$\mathcal{C}_c^\infty(\mathbb{R}^n)$ is dense in $\mathcal{S}(\mathbb{R}^n)$, and hence $\overline{\mathcal{S}(\mathbb{R}^n)} = L^1(\mathbb{R}^n)$.*

**Proof**

Let $f \in \mathcal{S}(\mathbb{R}^n)$. Consider a compactly supported smooth function $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ such that $\varphi(x) = 1$ on $B_1(y)$, for some $y$ in the support of $f$. Using $\varphi$ we can construct a family of compactly supported functions given by

$$\varphi_t(x) = \varphi(tx).$$

In this case the support is scaled by $1/t$, and $\varphi_t(x) = 1$ on $B_{1/t}(y)$. Consider a new function,

$$f_t = f(x)\varphi(tx)$$

The support of this function is the intersection of the supports of $f$ and $\varphi_t$, and hence compact. Since each $f$ and $\varphi_t$ are smooth functions, $f_t$ is a compactly supported smooth function. $f_t(x) - f(x) = f(x)\big(\varphi(tx) - 1\big) = 0$ on $B_{1/t}(y)$. Hence we have $f_t \to f$ as $t \to 0$ for $f_t \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ or

$$\overline{\mathcal{C}_c^\infty(\mathbb{R}^n)} \subset \mathcal{S}(\mathbb{R}^n)$$

□

**Theorem 4.2.4.** *$\mathcal{S}(\mathbb{R}^n)$ is complete as a metric space.*

**Proof**

Let $\{f_i\}$ be a Cauchy sequence in $\mathcal{S}(\mathbb{R}^n)$. Since the seminorms are uniform norm, for each multi-index $\alpha$ and $\beta$ there must exist a continuous function $f_{\alpha,\beta}$ such that

$$x^\beta \partial^\alpha f_i \to f_{\alpha,\beta}.$$

uniformly. We now have to show that all functions $f_{\alpha,\beta}$ are of the form $x^\beta \partial^\alpha f$ for some Schwartz function $f$. This is a Schwartz function follows because $|x^\beta \partial^\alpha f| = \lim_i |x^\beta \partial^\alpha f_i| < \infty$.

We prove the existence of $f$ by induction. Assume there exists $f$ such that $x^\beta \partial^\alpha f = f_{\alpha,\beta}$ for $|\alpha| + |\beta| \leqq n$, we now construct a function such that $x^\beta \partial^{\alpha+j} f_j = f_{\alpha+i,\beta}$ and $x^{\beta+i}\partial^\alpha f_i = f_{\alpha,\beta+i}$.

For $x^\beta \partial^{\alpha+j} f_j = f_{\alpha+i,\beta}$, we want a function such that, $\partial^j f = f_{\alpha+j,\beta}$ Let

$$f_j(x) = \int_0^1 \partial^j f(x_1, \ldots, tx_j, \ldots, x_n) d\lambda(x)$$

Define

$$f_i(x) =$$

# [[Finish the Proof]]

### 4.2.2 | Fourier Transforms on $\mathcal{S}(\mathbb{R}^n)$

Since Fourier transform can be defined on integrable functions we can define Fourier transform on the Schwartz space. Since $\partial^\alpha f, x^\beta f$ are integrable, it follows that we can define Fourier transform of $\partial f/\partial x_j$ and $x^\beta f$, and by the (4.1.4) we have,

$$\mathcal{F}(\partial f/\partial x_j)(\chi) = i\chi_j(\mathcal{F}f)(\chi),$$
$$\big(\partial(\mathcal{F}f)/\partial\chi_i\big)(\chi) = i\mathcal{F}(x_j f)(\chi)$$

Using these exchange formulas we have

$$|\chi^\beta\partial^\alpha(\mathcal{F}f)(\chi)| = |\chi^\beta\big(\mathcal{F}(x^\alpha f)(\chi)\big)| = |\mathcal{F}(\partial^\beta(x^\alpha f))(\chi)| = |\mathcal{F}(\partial^\beta(x^\alpha f))(\chi)|.$$

If $f$ is a Schwartz function then $\partial^\beta\big(x^\alpha f(x)\big)$ is also a Schwartz function. We however need to verify that the Fourier transform keeps $\mathcal{S}(\mathbb{R}^n)$ invariant.

**Lemma 4.2.5.** *$\mathcal{F}$ is continuous, and $\mathcal{F}(\mathcal{S}(\mathbb{R}^n)) \subseteq \mathcal{S}(\mathbb{R}^n)$.*

**Proof**

Let $f$ be a Schwartz function. Then we have

$$\begin{aligned}|\mathcal{F}(\partial^\beta(x^\alpha f))(\chi)| &= \left|\int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle}\Big(\partial^\beta\big(x^\alpha f(x)\big)\Big)d\lambda(x)\right| \\ &\leqq \int_{\mathbb{R}^n}\big|e^{-i\langle x,\chi\rangle}\big|\big|\partial^\beta\big(x^\alpha f(x)\big)\big|d\lambda(x) = \int_{\mathbb{R}^n}\big|\partial^\beta\big(x^\alpha f(x)\big)\big|d\lambda(x).\end{aligned}$$

Since $\partial^\beta\big(x^\alpha f(x)\big)$ is a Schwartz function it is integrable, and hence $|\chi^\beta\partial^\alpha(\mathcal{F}f)(\chi)| < \infty$. Hence we have $\mathcal{F}(\mathcal{S}(\mathbb{R}^n)) \subseteq \mathcal{S}(\mathbb{R}^n)$.

If $\{f_i\}$ is a sequence of Schwartz functions such that $f_i \to 0$ we have

$$\begin{aligned}\big|\mathcal{F}f_i(\chi)\big| &= \left|\int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} f_i(x)d\lambda(x)\right| \\ &\leqq \int_{\mathbb{R}^n}\big|e^{-i\langle x,\chi\rangle}\big|\big|f_i(x)\big|d\lambda(x) \to 0.\end{aligned}$$

Hence $\mathcal{F}f_i \to 0$ or $\mathcal{F}$ is continuous. □

Let $\mathcal{G}$ be the Fourier inversion formula defined by,

$$\big(\mathcal{G}g\big)(x) = (2\pi)^{-n}\int_{\mathbb{R}^n} e^{i\langle x,\chi\rangle} g(\chi)d\lambda(\chi).$$

We need to show that $\mathcal{G}$ is the inverse of $\mathcal{F}$. $\mathcal{G}$ is linear because integration is linear. It is injective for reasons similar to why $\mathcal{F}$ is injective. So, $\mathcal{F}\circ\mathcal{G}$ is also an injective linear map. Consider the composition,

$$\begin{aligned}\big((\mathcal{G}\circ\mathcal{F})f\big)(x) &= (2\pi)^{-n}\int_{\mathbb{R}^n} e^{i\langle x,\chi\rangle}\mathcal{F}f(\chi)d\lambda(\chi) \\ &= (2\pi)^{-n}\int_{\mathbb{R}^n} e^{i\langle x,\chi\rangle}\left[\int_{\mathbb{R}^n} e^{-i\langle y,\chi\rangle} f(y)d\lambda(y)\right]d\lambda(\chi) \\ &= (2\pi)^{-n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\Big[e^{i\langle x-y,\chi\rangle} f(y)d\lambda(y)\Big]d\lambda(\chi)\end{aligned}$$

$e^{i\langle x-y,\chi\rangle}f(y)$ is not integrable, and hence we cannot use Fubini's theorem. The idea now is to use an integrating factor or damping factor and prove the theorem as a limit case. Since this is a frequently used trick we will discuss the trick.

**The Delta Distribution**

The idea now is to apply dominated convergence theorem. The dominated convergence theorem states that whenever we have a sequence of functions $\{h_i\}$ with $|h_i(x)| \leqq |g(x)|$ such that $h_i(x) \to h(x)$ then it follows that, $\lim_{i\to\infty} \int h_i = \int h$. In order to do this, we introduce an integrating factor or damping factor $\varphi_t(x)$ that depends on a parameter $t$, integrate the convolution with this integrating factor, show that the required property holds for every $t$ and then get the required result for the limit.

For any Schwartz function $\varphi$, consider a new function,

$$\varphi_t(x) = t^{-n}\varphi\left(\frac{x}{t}\right).$$

Since $|x^\beta \partial^\alpha \varphi_t| = |t^{-|\alpha|} x^\beta \partial^\alpha \varphi|$ it follows that $\varphi_t$ is also a Schwartz function for all $t > 0$. By change of variable we have, $d\lambda(y) = t^n d\lambda(y/t)$ and hence we have,

$$\begin{aligned}\int_{\mathbb{R}^n} \varphi_t(x) d\lambda(x) &= \int_{\mathbb{R}^n} t^{-n}\varphi(x/t) d\lambda(x) \\ &= \int_{\mathbb{R}^n} t^{-n}\varphi(x/t) t^n d\lambda(x/t) = \int_{\mathbb{R}^n} \varphi(x) d\lambda(x).\end{aligned}$$

**Lemma 4.2.6.** *Let $0 \leqq \varphi \in \mathcal{S}(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \varphi(x) d\lambda(x) = 1$, then for every $1 \geq \epsilon > 0$ there exists a compact set $K^\epsilon \subset \mathbb{R}^n$ such that,*

$$\int_{K^\epsilon} \varphi(x) d\lambda(x) = \epsilon.$$

**Proof**

Assume without loss of generality that $\varphi(0) > 0$ then there exists some closed ball $K$ around origin where its is non-zero. Hence $\int_K \varphi(x) d\lambda(x) > 0$. Now we can parametrise balls around origin by their radius, denote by $K_s$ a closed ball around origin of radius $s$. Then we have,

$$\begin{aligned}\int_{\mathbb{R}^n} \varphi(x) d\lambda(x) &= \left[\int_{K_s} + \int_{K_s^c}\right] \varphi(x) d\lambda(x) \\ &= \left[\int_{K_s} \varphi(x) d\lambda(x)\right] + \left[\int_{K_s^c} \varphi(x) d\lambda(x)\right].\end{aligned}$$

It is sufficient to show that the function

$$s \mapsto \lambda_\varphi(s) := \int_{K_s} \varphi(x) d\lambda(x)$$

is continuous, that is $\int_{K_s} \varphi(x) d\lambda(x) \to \int_{K_r} \varphi(x) d\lambda(x)$ as $s \to r$. Assume that $s \leqq r$ then,

$$\begin{aligned}\left|\lambda_\varphi(r) - \lambda_\varphi(s)\right| &= \left|\left[\int_{K_r} - \int_{K_s}\right] \varphi(x) d\lambda(x)\right| \\ &= \left|\int_{K_r \setminus K_s} \varphi(x) d\lambda(x)\right| \leqq \left[\lambda(K_r) - \lambda(K_s)\right] \left\|\varphi\right\|_\infty\end{aligned}$$

As $s \to r$, we have $\lambda(K_s) \to \lambda(K_r)$. Hence we have, $|\lambda_\varphi(r) - \lambda_\varphi(s)| \to 0$. □

Let $\varphi$ be such that $\varphi(x) \geq 0$ and

$$\varphi(y) = 1, \int_{\mathbb{R}^n} \varphi(x) d\lambda(x) = 1.$$

By the previous lemma, let $s$ be such that $\int_{K_s} \varphi(x) d\lambda(x) < 1 - \epsilon$, where $K_s$ is a closed ball centered at $y \in \mathbb{R}^n$ instead of origin. Notice that for $\varphi_t$, for $s/t$ we have,

$$\int_{K_{s/t}} \varphi(x) d\lambda(x) < 1 - \epsilon.$$

If $h$ is a bounded function we have

$$\begin{aligned}
\left| h(y) - \int_{\mathbb{R}^n} h(x)\varphi_t(x) d\lambda(x) \right| &= \left| h(y) \int_{\mathbb{R}^n} \varphi_t(x) d\lambda(x) - \int_{\mathbb{R}^n} h(x)\varphi_t(x) d\lambda(x) \right| \\
&= \left| \int_{\mathbb{R}^n} \big(h(y) - h(x)\big)\varphi_t(x) d\lambda(x) \right| \\
&\leqq \left[ \int_{K_{s/t}} + \int_{K^c_{s/t}} \right] \big|h(y) - h(x)\big|\big|\varphi(x)\big| d\lambda(x) \\
&\leqq (1 - \epsilon)\Big[ \sup_{K_{s/t}} \big|h(y) - h(x)\big| \Big] + 2\epsilon \|h\|_\infty.
\end{aligned}$$

Hence as $t$ increases $K_{s/t}$ becomes small. If $h$ is a continuous function then the above tells us that $\lim_{t\to\infty} \int_{\mathbb{R}^n} h(x)\varphi_t(x) d\lambda(x) = h(y)$. Using the functions $\varphi_t$ we can construct a one parameter collection of distributions,

$$f \mapsto \langle \varphi_t, f \rangle \equiv \int_{\mathbb{R}^n} f(x)\varphi_t(x) d\lambda(x).$$

Since the space of tempered distributions is complete with respect to weak topology the collection $\lim\langle \varphi_t, \cdot \rangle$ defines a unique tempered distribution called the delta distribution denoted by $\delta_y$.

$$\langle \delta_y, f \rangle \coloneqq \lim_{i\to\infty} \langle \varphi_{t_i}, f \rangle$$

This distributions is such that for every Schwartz function $f \in \mathcal{S}(\mathbb{R}^n)$ we have

$$\langle \delta_y, f \rangle = f(y).$$

The Schwartz function $\varphi$ above is called a damping factor.

#### 4.2.2.1 | The Fourier Inversion Theorem

We can now prove the Fourier inversion theorem. The idea of proof is to use the identity, $\int (\mathcal{F}f)g = \int f(\mathcal{F}g)$ to transfer the Fourier transform on a random Schwartz function to the more well behaved Schwartz function $\varphi$. By definition,

$$(\mathcal{G} \circ \mathcal{F})f = \int_{\mathbb{R}^n} e^{i\langle x,\chi \rangle} (\mathcal{F}f)(\chi) d\lambda(\chi).$$

Consider a damping factor $\varphi \in \mathcal{S}(\mathbb{R}^n)$, such that $\varphi(0) = 1$ and

$$\varphi(x) \geq 0, \int_{\mathbb{R}^n} \varphi(\chi) d\lambda(\chi) = 1.$$

If $h(\chi) = e^{i\langle x,\chi\rangle}\mathcal{F}f(\chi)$ we have,

$$\begin{aligned}\langle \varphi_t, h\rangle &= \int_{\mathbb{R}^n} e^{i\langle x,\chi\rangle}(\mathcal{F}f)(\chi)\varphi_t(\chi)d\lambda(\chi)\\ &= \int_{\mathbb{R}^n} e^{i\langle x,\chi\rangle}\left[\int_{\mathbb{R}^n} e^{-i\langle y,\chi\rangle} f(y)d\lambda(y)\right]\varphi_t(\chi)d\lambda(\chi)\\ &= \int_{\mathbb{R}^n}\left[\int_{\mathbb{R}^n} e^{-i\langle y-x,\chi\rangle} f(y)\varphi_t(\chi)d\lambda(y)\right]d\lambda(\chi).\end{aligned}$$

Clearly $|e^{-i\langle y-x,\chi\rangle}f(y)\varphi_t(\chi)| \leq |f(y)||\varphi_t(\chi)|$, and hence we can apply dominated convergence theorem, and we have,

$$\lim_{t\to\infty}\langle\varphi_t, h\rangle = h(y).$$

The bracketed term is integrable, and hence Fubini's theorem is applicable. If we have $\int_{\mathbb{R}^n} e^{i\langle x,\chi\rangle}\mathcal{F}f(\chi)\varphi_t(\chi)d\lambda(\chi) \to h(0) = \mathcal{F}f(\chi)$.

$$\int_{\mathbb{R}^n} e^{i\langle x,\chi\rangle}(\mathcal{F}f)(\chi)\varphi_t(\chi)d\lambda(\chi)$$

# [[Fix the Confusion]]

For every Schwartz function $h$ and Schwartz function $\varphi$ with

$$\varphi(x) \geq 0, \int_{\mathbb{R}^n}\varphi(x)d\lambda(x) = 1,$$

For any two Schwartz functions $h$ and $\varphi$ with $\varphi \geq 0$ and $\int_{\mathbb{R}^n} f(x)\varphi(x)d\lambda(x) = 1$, we have

$$\int_{\mathbb{R}^n}(\mathcal{F}h)(x)\varphi_t(x)d\lambda(x) = \int_{\mathbb{R}^n} h(x)(\mathcal{F}\varphi_t)(x)d\lambda(x).$$

we have,

$$\int_{\mathbb{R}^n}(\mathcal{F}h)(y-x)\varphi_t(x)d\lambda(x) = \int_{\mathbb{R}^n} h(y-x)(\mathcal{F}\varphi_t)(x)d\lambda(x).$$

$h(\chi) = e^{i\langle x,\chi\rangle}\mathcal{F}f(\chi)$ is a bounded function since $\|\mathcal{F}f\|_\infty \leqq \|f\|_{L^1}$. Let

$$h_i(\chi) = h_{\varphi_i}(\chi) = \int_{\mathbb{R}^n} h(\chi-\nu)\varphi_i(\nu)d\lambda(\nu).$$

**Theorem 4.2.7. (Fourier Inversion)** *$\mathcal{F}$ is invertible with $\mathcal{F}^{-1} \equiv \mathcal{G}$ and*

$$(\mathcal{F}\circ\mathcal{F})(f) = (2\pi)^n(\mathcal{G}f).$$

**Proof**

For any $f \in \mathcal{S}(\mathbb{R}^n)$ we must show that

$$\big((\mathcal{G}\circ\mathcal{F})f\big)(x) = (2\pi)^{-n}\int_{\mathbb{R}^n} e^{i\langle x,\chi\rangle}\Big[\int_{\mathbb{R}^n} e^{-i\langle y,\chi\rangle}f(y)d\lambda(y)\Big]d\lambda(\chi) = f(x).$$

Let

$$h(\chi) = e^{i\langle x,\chi\rangle}\int_{\mathbb{R}^n} e^{-i\langle y,\chi\rangle} f(y)d\lambda(y)$$

Hence $|h(\chi)| \leqq |\mathcal{F}f| \leqq \|f\|_{L^1}$. For a Schwartz function $\varphi$ with $\varphi(x) \geq 0$ and $\int_{\mathbb{R}^n}\varphi(x)d\lambda(x)$ let $h_i(\chi) = \int_{\mathbb{R}^n} h(\chi-\nu)\varphi_i(\nu)d\lambda(\nu)$. Since both $h$ and $\varphi$ are bounded we have $|h_i(\chi)| \leqq \|h\|_\infty\|\varphi\|_\infty$, and $h_i(\chi)\to h(\chi)$ for all $\chi$. Hence by dominated convergence,

$$\lim_{i\to\infty}\int_{\mathbb{R}^n} h_i(\chi)d\lambda(\chi) = \int_{\mathbb{R}^n} h(\chi)d\lambda(\chi).$$

We must find a function $f$ such that $f_i(x)\to f(x)$. We have,

$$\begin{aligned}\lim_{t\to 0}\int_{\mathbb{R}^n}\varphi(t\chi)\Big[\int_{\mathbb{R}^n} e^{-\langle y,\chi\rangle}f(y)d\lambda(y)\Big]d\lambda(\chi) &= \lim_{t\to 0}\int_{\mathbb{R}^n} e^{i\langle x,\chi\rangle}\varphi(t\chi)\mathcal{F}f(\chi)d\lambda(\chi)\\ &= \int_{\mathbb{R}^n} e^{i\langle x,\chi\rangle}\mathcal{F}f(\chi)d\lambda(\chi)\end{aligned}$$

# [[Finish the Proof]]

**A Formula for Delta Distribution**

### 4.2.2.2 | Parseval's Formula

**Theorem 4.2.8. (Parseval)** *If $f, g \in \mathcal{S}(\mathbb{R}^n)$ then,*

$$\int_{\mathbb{R}^n} f(x)\overline{g(x)}d\lambda(x) = (2\pi)^{-n}\int_{\mathbb{R}^n}\widehat{f}(\chi)\overline{\widehat{g}(\chi)}d\lambda(\chi).$$

**Proof**

This follows directly by expansion,

$$\begin{aligned}\int_{\mathbb{R}^n}(\mathcal{F}f)(\chi)\overline{(\mathcal{F}g)(\chi)}d\lambda(\chi) &= \int_{\mathbb{R}^n}\left[\int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle}f(x)d\lambda(x)\right]\overline{\left[\int_{\mathbb{R}^n} e^{-i\langle y,\chi\rangle}g(y)d\lambda(y)\right]}d\lambda(\chi)\\ &= \int_{\mathbb{R}^n}\left[\int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle}f(x)d\lambda(x)\right]\left[\int_{\mathbb{R}^n} e^{i\langle y,\chi\rangle}\overline{g(y)}d\lambda(y)\right]d\lambda(\chi)\\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\left[\int_{\mathbb{R}^n} e^{-i\langle x-y,\chi\rangle}d\lambda(\chi)\right]f(x)\overline{g(y)}d\lambda(x)d\lambda(y)\\ &= \int_{\mathbb{R}^n} f(x)\underbrace{\left[\int_{\mathbb{R}^n}\delta(y-x)\overline{g(y)}d\lambda(y)\right]}_{\overline{g(x)}}d\lambda(x) = \int_{\mathbb{R}^n} f(x)\overline{g(x)}d\lambda(x).\end{aligned}$$

□

**Theorem 4.2.9.** *For $f, g \in \mathcal{S}(\mathbb{R}^n)$,*

$$\widehat{(f\cdot g)}(\chi) = (2\pi)^{-n}\big(\widehat{f} * \widehat{g}\big)(\chi)$$

**Proof**

$$\begin{aligned}
(\widehat{f} * \widehat{g})(\chi) &= \int_{\mathbb{R}^n} \left[ \int_{\mathbb{R}^n} e^{-i\langle x, \chi - \nu \rangle} f(x) d\lambda(x) \right] \left[ \int_{\mathbb{R}^n} e^{-\langle y, \nu \rangle} g(y) d\lambda(y) \right] d\lambda(\nu) \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-i\langle x, \chi \rangle} e^{-i\langle x, \nu \rangle} f(x) e^{-\langle y, -\nu \rangle} g(y) d\lambda(y) d\lambda(\nu) d\lambda(x) \\
&= \int_{\mathbb{R}^n} e^{-i\langle x, \chi \rangle} f(x) \left[ \int_{\mathbb{R}^n} \left[ \int_{\mathbb{R}^n} e^{-\langle y - x, -\nu \rangle} d\lambda(\nu) \right] g(y) d\lambda(y) \right] d\lambda(x) \\
&= \int_{\mathbb{R}^n} e^{-i\langle x, \chi \rangle} f(x) g(x) d\lambda(x) = \widehat{(f \cdot g)}(\chi).
\end{aligned}$$

$\square$

## 4.3 | Tempered Distributions

A tempered distribution is any continuous linear functional $\tau : f \mapsto \langle \tau, f\rangle$, on the Schwartz space $\mathcal{S}(\mathbb{R}^n)$, where topology on $\mathcal{S}(\mathbb{R}^n)$ is the locally convex topology given by the countable collection of semi-norms, $\|\cdot\|_{\alpha,\beta}$. The space of all tempered distributions is denoted by $\mathcal{T}(\mathbb{R}^n)$.

**Theorem 4.3.1.** *$\tau$ is a tempered distribution if and only if for all $f \in \mathcal{S}(\mathbb{R}^n)$ there exist $c_\tau$ and $k$ such that,*

$$\big|\langle \tau, f\rangle\big| \leqq c_\tau \sum_{|\alpha|+|\beta|\leqq k} \big\|f\big\|_{\alpha,\beta}.$$

**Proof**

$\Leftarrow$ Let $\{f_i\}$ is a sequence of test functions that tend to the zero function in the topology generated by the semi-norms $\|\cdot\|_{\alpha,\beta}$. Then $\|f_i\|_{\alpha,\beta} \to 0$ for all multi-indices $\alpha$ and $\beta$. If $\tau$ satisfies the inequality the above inequality, then we have,

$$|\langle \tau, f_i\rangle| \leqq c_\tau \textstyle\sum_{|\alpha|+|\beta|\leqq k} \|f_i\|_{\alpha,\beta}$$

which means $\langle \tau, f_i\rangle \to 0$.
$\Rightarrow$ Suppose the condition is not true, that is, $\langle \tau, \cdot\rangle$ is a continuous functional and the inequality $\big|\langle \tau, f\rangle\big| \leqq c_\tau \sum_{|\alpha|+|\beta|\leqq k} \big\|f\big\|_{\alpha,\beta}$ is not true for any $c_\tau$ and $k$. To arrive at a contradiction, take $c_\tau = k = j$ and let $f_j$ be a function such that

$$\big|\langle \tau, f_j\rangle\big| > j \textstyle\sum_{|\alpha|+|\beta|\leqq j} \big\|f_j\big\|_{\alpha,\beta}.$$

For the sake of simplicity assume by rescaling that $\langle \tau, f_j\rangle = 1$. So, whenever $j \geq |\alpha| + |\beta|$, we have, $|x^\beta \partial^\alpha f_j| \leqq 1/j$. This means that the sequence $\{f_j\}$ converges to zero even though $\{\langle \tau, f_j\rangle\}$ does not. However we had assumed that $\langle \tau, \cdot\rangle$ is continuous, and hence we have arrived at a contradiction. $\square$

Note that each integrable function $\varphi$ consider the functional

$$f \mapsto \langle \varphi, f\rangle \equiv \int_{\mathbb{R}^n} f(x)\varphi(x)d\lambda(x).$$

In this case we have

$$\begin{aligned}\big|\langle \varphi, f\rangle\big| = \left|\int_{\mathbb{R}^n} f(x)\varphi(x)d\lambda(x)\right| &\leqq \int_{\mathbb{R}^n} \big|f(x)\big|\big|\varphi(x)\big|d\lambda(x)\\ &\leqq \sup_{x\in\mathbb{R}^n} \big|f(x)\big| \underbrace{\left[\int_{\mathbb{R}^n} |\varphi(x)|d\lambda(x)\right]}_{c_\varphi}.\end{aligned}$$

Note that $\sup_x |f(x)| = \|f\|_{0,0}$ and hence $\langle \varphi, \cdot\rangle$ defines a tempered distribution with $k = 0$ and $c_\tau$ and above. This is clearly continuous since if $f_i \to f$ in the Schwartz space then $\sup_x |f_i(x) - f(x)| \to 0$ and hence $\langle \varphi, f_i\rangle \to \langle \varphi, f\rangle$. In this sense tempered distributions generalise continuous functionals of this integral type.

### 4.3.1 | The Space of Tempered Distributions

Since every compactly supported smooth function is also a Schwartz function, the restriction of a tempered distribution $\tau$ restricted to $\mathcal{C}_c^\infty(\mathbb{R}^n)$ is a continuous linear functional on $\mathcal{C}_c^\infty(\mathbb{R}^n)$. Hence each tempered distribution can be restricted to get a distribution. Or, the space of tempered distributions is a subset of the space of distributions. We have,

$$\mathcal{C}_c^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \Rightarrow \mathcal{D}(\mathbb{R}^n) \supset \mathcal{T}(\mathbb{R}^n).$$

The restriction $|_{\mathcal{C}_c^\infty(\mathbb{R}^n)} : \mathcal{T}(\mathbb{R}^n) \to \mathcal{D}(\mathbb{R}^n)$ is an injective map.

The space of compactly supported smooth functions $\mathcal{C}_c^\infty(\mathbb{R}^n)$ is dense in the space of Schwartz function $\mathcal{S}(\mathbb{R}^n)$. Hence

**Theorem 4.3.2.** $\mathcal{T}(\mathbb{R}^n)$ *is complete in its natural topology.*

**Theorem 4.3.3.** $\mathcal{D}(\mathbb{R}^n)$

it is then extended to other distributions via limits using completeness of $\mathcal{D}(\mathbb{R}^n)$. For this idea to work, we need tempered distributions to be dense in $\mathcal{D}(\mathbb{R}^n)$.

### 4.3.2 | Fourier Transform on $\mathcal{T}(\mathbb{R}^n)$

The idea is to define the Fourier transform $\mathcal{F}\tau$ of a distribution $\tau$ as a continuous linear functional on the space of test functions that satisfies the expected properties of Fourier transforms. Each integrable function $\varphi$ gives rise to a continuous linear functional on the Schwartz space given by

$$f \mapsto \langle \varphi, f \rangle := \int_{\mathbb{R}^n} f(x)\varphi(x) d\lambda(x).$$

For integrable functions we have,

$$\int_{\mathbb{R}^n} (\mathcal{F}f)(x)g(x)d\lambda(x) = \int_{\mathbb{R}^n} f(x)(\mathcal{F}g)(x)d\lambda(x).$$

If $\tau$ is a tempered distribution on $\mathbb{R}^n$, then the Fourier transform $\mathcal{F}\tau$ of $\tau$ is defined as the unique continuous linear functional $\mathcal{F}\tau$ such that for all Schwartz functions $f$,

$$\langle \mathcal{F}\tau, f \rangle = \langle \tau, \mathcal{F}f \rangle.$$

Note that $\mathcal{F}\tau$ is uniquely determined because $\mathcal{F}$ is an isomorphism on $\mathcal{S}(\mathbb{R}^n)$. $\mathcal{G}$ is similarly defined as $\langle \mathcal{G}\tau, f \rangle = \langle \tau, \mathcal{G}f \rangle$. Both $\mathcal{F}$ and $\mathcal{G}$ are bijective linear maps on $\mathcal{T}(\mathbb{R}^n)$. The pairing gives us the immediate properties of the Fourier transform of tempered distributions.

The bilinearity of the pairing also implies that $\langle \mathcal{F}(\tau+\kappa), f \rangle = \langle \tau+\kappa, \mathcal{F}f \rangle = \langle \mathcal{F}\tau + \mathcal{F}\kappa, f \rangle$. Hence $\mathcal{F}$ is a linear map. Suppose $\tau_i \to \tau$ that is for all Schwartz functions $f$, $\langle \tau_i, f \rangle \to \langle \tau, f \rangle$. Then we have,

$$\langle \mathcal{F}\tau_i, f \rangle = \langle \tau_i, \mathcal{F}f \rangle \to \langle \tau, \mathcal{F}f \rangle = \langle \mathcal{F}\tau, f \rangle$$

Since $\mathcal{F}\tau$ is determined by its action on the Schwartz space we have,

$$\langle (\mathcal{G} \circ \mathcal{F})\tau, f \rangle = \langle \tau, (\mathcal{G} \circ \mathcal{F})f \rangle.$$

Since $\mathcal{G} \circ \mathcal{F}$ is identity on the Schwartz space it follows that $\mathcal{G} \circ \mathcal{F} = \mathbb{I}$. Using $\mathcal{F} \circ \mathcal{F}f = (2\pi)^n(\mathcal{G}f)$ and by the bilinearity of the pairing we have,

$$\langle (\mathcal{F} \circ \mathcal{F})\tau, f \rangle = \langle \tau, (\mathcal{F} \circ \mathcal{F})f \rangle = \langle \tau, (2\pi)^n(\mathcal{G}f) \rangle = \langle (2\pi)^n(\mathcal{G}\tau), f \rangle.$$

Hence we have,

$$(\mathcal{F} \circ \mathcal{F})\tau = (2\pi)^n \mathcal{G}\tau.$$

We have proved the following theorem,

**Theorem 4.3.4.** *$\mathcal{F}$ is a continuous linear isomorphism on $\mathcal{T}(\mathbb{R}^n)$.*

The exchange formulas of Fourier transforms on the Schwartz space give the exchange formulas for tempered distributions.

## 4.4 | Analyticity & Support

Decay properties of tempered distributions are closely related to smoothness of the Fourier transform. Paley-Weiner theorems say that compactly supported functions have analytic Fourier transform.

The Fourier transform of a compactly supported distribution $\kappa$ is defined by,

$$\langle \mathcal{F}\kappa, \varphi \rangle = \langle \rangle$$

### 4.4.1 | Fourier-Laplace Transform on $\mathcal{E}(\mathbb{R}^n)$

Recall that $\mathcal{E}(\mathbb{R}^n)$ is the space of all continuous linear functionals on the space of all smooth functions $\mathcal{C}^\infty(\mathbb{R}^n)$.[1] The Fourier transform of a distribution $\nu \in \mathcal{E}(\mathbb{R}^n)$ is defined

Note that we can 'vaguely' extend the Fourier transform to the complex plane as follows,

$$\begin{aligned}\mathcal{F}f(\chi + i\nu) &= \int_{\mathbb{R}^n} e^{-i\langle x, \chi + i\nu\rangle} f(x) d\lambda(x) \\ &= \int_{\mathbb{R}^n} e^{-i\langle x, \chi\rangle} \underbrace{e^{\langle x, \nu\rangle} f(x)}_{h(x)} d\lambda(x) = \mathcal{F}\big(e^{-\langle x, \nu\rangle} f(x)\big)(\chi).\end{aligned}$$

### 4.4.2 | Paley-Wiener-Schwartz Theorem

**Theorem 4.4.1. (Paley-Wiener-Schwartz)** *Let $K$ be a compact convex subset of $\mathbb{R}^n$ with supporting function $H$. If $\tau$ is a compactly supported distribution of order $n$ with support contained in $K$, then for all $\chi \in \mathbb{C}^n$,*

$$\big|\mathcal{F}\tau(\chi)\big| \leqq \lambda_\tau (1 + |\chi|)^n e^{H(Im(\chi))}.$$

*Conversely, every analytic function in $\mathbb{C}^n$ satisfying the above inequality is the Fourier-Laplace transform of a distribution with support in $K$.*

## 4.5 | Applications

[1] The topology on $\mathcal{C}^\infty$ is such that a sequence $f_i$ converges to $f$ if and only if

$$\sup_{x \in K} \big|\partial^\alpha f_i(x) - \partial^\alpha f(x)\big| \equiv \big\|f_i - f\big\|_{\alpha, K} \to 0.$$

where $K$ is a compact subset of $\mathbb{R}^n$.

## 4.6 | Checklist for Chapter

- ✓ Fourier on integrable functions
- ✓ Schwartz Space
- Fourier Inversion Formula (confused)
- ✓ Tempered Distribution
- Topological Properties, inclusions, density theorems, etc.
- ✓ Fourier Transform on Tempered
- Fourier-Laplace transform
- Paley-Wiener-Schwartz

# II

# Hyperfunctions

# 5 | Holomorphic Functions

## 5.1 | Cauchy-Riemann Equations

## 5.2 | Holomorphic Functions

### 5.2.1 | Main Theorems

### 5.2.2 | Runge's Theorem

## 5.3 | Holomorhic Functions in Several Variables

### 5.3.1 | Main Theorems

### 5.3.2 | Hartog's Extension

## 5.4 | Domains of Holomorphy

## 5.5 | Global Theorems

### 5.5.1 | The Tube Theorem

### 5.5.2 | The Edge of the Wedge Theorem

### 5.5.3 | The Double Cone Theorem

[1] L Hörmander, The Analysis of Linear Partial Differential Operators I, Second Edition, Springer-Verlag, 1989