# DIPLOMARBEIT

Titel der Diplomarbeit

"The Kernel Theorem and Microlocal Analysis for Distributions on Manifolds"

Verfasser

Vera Ganglberger

angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag.rer.nat)

Wien, 2010

| | |
|---|---|
| Studienkennzahl lt. Studienblatt: | A 406 |
| Studienrichtung lt. Studienblatt: | Mathematik |
| Betreuer: | Ao.Univ.-Prof. Dr. Günter Hörmann |

# Abstract

This thesis develops several aspects of the theory of kernel distributions on manifolds. Beginning with a section on prerequisites, it shortly outlines the most important concepts needed from the fields of functional analysis, differential geometry and distribution theory.

The second chapter then deals with the kernel theorem by Laurent Schwartz, stating that all linear and continuous operators $\mathcal{D}(X) \to \mathcal{D}'(Y)$ (for $X$ and $Y$ two manifolds) are in fact kernel maps given by a distribution $K \in \mathcal{D}'(X \times Y)$ via $(K \cdot \varphi)(\psi) = \langle K, \varphi \otimes \psi \rangle$ for $\varphi \in \mathcal{D}(X)$, $\psi \in \mathcal{D}(Y)$. After proving this theorem, regular and regularising kernel operators are examined as special cases of kernel operators that allow an extension of the kernel map to distribution spaces.

The last chapter then moves on to microlocal analysis on manifolds, where first of all an intrinsic, coordinate-invariant definition of the wave front set for distributions on a manifold is given. Ultimately, the thesis returns to kernel operators and investigates their microlocal properties, showing how the wave front set of a distribution of the form $K \cdot u \in \mathcal{D}'(Y)$ for regular kernels $K$ and compactly supported distributions $u \in \mathcal{E}'(X)$ is controlled by the wave front sets of $K$ and $u$.

# Zusammenfassung

Diese Diplomarbeit behandelt mehrere Aspekte der Theorie der Kerndistributionen auf Mannigfaltigkeiten. Dazu werden einleitend die wichtigesten Begriffe aus den Gebieten der Funktionalanalysis, Differentialgeometrie und Distributionentheorie wiederholt.

Danach behandelt der Text den Satz vom distributionellen Kern auf Mannigfaltigkeiten von Laurent Schwartz, welcher besagt, dass alle stetigen und linearen Abbildungen des Raumes $\mathcal{D}(X)$ in den Raum $\mathcal{D}'(Y)$, für $X$ und $Y$ zwei Mannigfaltigkeiten, durch eine Kerndistribution $K \in \mathcal{D}'(X \times Y)$ und die Abbildung $(K \cdot \varphi)(\psi) = \langle K, \varphi \otimes \psi \rangle$ für $\varphi \in \mathcal{D}(X)$ und $\psi \in \mathcal{D}(Y)$ gegeben werden kann. Anschließend werden die Spezialfälle der regulären und regularisierenden Kerndistributionen besprochen, welche eine Ausdehnung des Kernoperators auf Räume von Distributionen erlauben.

Schließlich wird das Thema mikrolokale Analysis im Hinblick auf Distributionen auf Mannigfaltigkeiten und insbesondere für Kernoperatoren besprochen. Dazu wird zuerst eine Koordinaten-invariante Definition der Wellenfrontmenge hergeleitet, welche es ermöglicht von der Wellenfrontmenge einer Distribution auf einer Mannigfaltigkeit zu sprechen. Danach wird die Wellenfrontmenge von Distributionen untersucht, die von einem Kernoperator stammen, d.h. die der Form $K \cdot u \in \mathcal{D}'(Y)$ für einen regulären Kern $K$ und eine Distribution $u \in \mathcal{E}'(X)$ entsprechen.

# Acknowledgements

Above all I want to thank Günther Hörmann for being the most patient and helpful supervisor I could have hoped for. Thank you for always taking the time to answer my questions, read all my various drafts and redrafts and for always offering constructive criticism. But most of all, thank you for never giving me the impression that I was stealing your doubtlessly precious time.

Secondly, I want to thank the members of the DIANA group at the University of Vienna for the many interesting and enlightening hours I spent in your lectures, and for putting thought not only into the content but also into the didactics.

And last but by no means least I want to thank my parents for enabling all this through not only funding, but also by never giving me a reason to believe there was anything I could not do.

# Contents

# 1 Introduction

Kernel distributions are generalisations of kernels in integral transforms of the type

$$f \mapsto kf(y) := \int_X k(x,y)f(x)dx,$$

which map functions defined on $X$ to functions defined on $Y$ using a kernel function $k$ whose domain space is the product $X \times Y$. In a standard $L^2$ setting for instance we have that transforms of this type are linear and continuous under suitable conditions on $k$, but by far not every bounded operator of the form $K : L^2(X) \to L^2(Y)$ can be given by a kernel function $k$ on $X \times Y$. In a distributional setting however, we have the kernel theorem by Laurent Schwartz stating that we have a kernel distribution. To be more precise, let $A : \mathcal{D}(X) \to \mathcal{D}'(Y)$ be a linear and continuous operator (with respect to the usual topologies), then the kernel theorem states that there exists a distribution $K \in \mathcal{D}'(X \times Y)$ such that we have

$$(A(\varphi))(\psi) = \langle K, \varphi \otimes \psi \rangle \quad \forall \varphi \in \mathcal{D}(X), \forall \psi \in \mathcal{D}(Y).$$

The first major aim of this thesis is to give a proof of this statement for $X$ and $Y$ two manifolds, which, to begin with, requires an introduction into the field of distributions on manifolds. After having proved the kernel theorem, we investigate several special cases of these kernel operators, namely regular and regularising kernel operators. These kernels allow an extension of the operator to larger, and even distributional, domain spaces. Regularising kernel operators, which display the "nicest" behaviour, are the $\mathcal{C}^\infty$-analogue to integral transforms as introduced above.

After having explored the mapping properties of kernel operators to a satisfactory degree, we turn to microlocal analysis and the wave front set of distributions, which is a refinement of the concept of singular support. Subsequently a suitable theory for microlocal analysis on manifolds is introduced, and the ultimate goal is to investigate the wave front sets of distributions that are images of distributions under a regular kernel operator, i.e. gaining information on the set $\mathrm{WF}(K \cdot u)$ for $K \in \mathcal{D}'(X \times Y)$ regular and $u \in \mathcal{D}'(X)$ compactly supported. On manifolds the major issue is finding a coordinate-invariant, intrinsic definition of the wave front set for a distribution on the manifold.

The thesis mainly follows J. Dieudonné's "Treatise on Analysis Vol. 7" [5], additional references for specific sources and results are given in the text.

# 2 Prerequisites

## 2.1 Functional Analysis

This section is a short collection of results in functional analysis that will be needed at various points later.

**Theorem 2.1.** (Banach-Steinhaus) *Let $E$ be a Fréchet space, $F$ a normed space. If $H$ is a set of continuous and linear operators $E \to F$ such that*

$$\sup_{u \in H} \|u(x)\| < \infty \quad \forall x \in E,$$

*then $H$ is equicontinuous.*

*Proof.* see [3, 12.16.4] □

**Remark 2.2.** In a distribution theoretic setting we will later read this theorem in the following way: Let $B$ be a bounded set of distributions. Then we have that

$$\forall K \subset\subset X \; \exists m \in \mathbb{N} \; \exists C > 0 \; \forall u \in B \; \forall \varphi \in \mathcal{D}(K) : |\langle u, \varphi \rangle| \leq C \sum_{|\alpha \leq m} \|\partial^\alpha \varphi\|_{L^\infty(K)},$$

by simply applying the Banach Steinhaus theorem to the Fréchet spaces $\mathcal{D}(K)$. This in particular implies that all the distributions in $B$ are of order less or equal to $m$.

The following two lemmas will be needed for the proof of the kernel theorem. They both deal with continuity issues for maps defined on Fréchet spaces or products of Fréchet spaces.

**Lemma 2.3.** *Let $E$ and $F$ be Fréchet spaces, $G$ a Banach-space and let $B$ be a bilinear map $B : E \times F \to G$ with the property that both of the partial linear maps $B(x, .)$ and $B(., y)$ are continuous for $x \in E$ resp. for $y \in F$. Then $B$ is continuous on $E \times F$.*

*Proof.* For a proof of this lemma see for instance [15, 5.1, p. 51] □

**Lemma 2.4.** *Let $(E, \tau_E)$, $(F, \tau_F)$ be Fréchet spaces and $P : E \to F$ be linear. If we know that $P$ is continuous with respect to some Hausdorff topologies $\tau$ (on $E$) and $\tau'$ (on $F$) both weaker than the original Fréchet space topologies, then it follows that $P : (E, \tau_E) \to (F, \tau_F)$ is also continuous.*

*Proof.* The closed graph theorem (cf. [3, 12.16.13., p. 103]) tells us that it suffices to show that the graph of $P$ is closed in $E \times F$ (endowed with the product topology $\tau_E \times \tau_F$). This we will in turn show by proving that if $(u_n) \to u$ (w.r.t. $\tau_E$) and $Pu_n \to v$ (w.r.t. $\tau_F$, then $Pu = v$. We know from the assumption that $(u_n, Pu_n) \to (u, v)$ in the coarser, yet Hausdorff, product topology, and so it follows that $Pu$ is in fact equal to $v$. □

**Proposition 2.5.** *Let $E$ be a compact metric space, $F$ a normed space and $(f_n)$ an equicontinuous sequence of functions $\mathcal{C}(E) \to F$. If $(f_n)$ converges to some $g$ on $E$, then the convergence is even uniform on $E$.*

*Proof.* (cf. [3, Theorem 7.5.6]) □

**Proposition 2.6.** *Any bounded subset of $\mathcal{E}(U)$ is relatively compact in $\mathcal{E}(U)$.*

*Proof.* (cf. [4, Theorem 17.1.2]) □

**Theorem 2.7.** *Let $E$ be a metrizable and separable locally convex vector space and $H$ an equicontinuous subset of the dual space $E'$. Then the weakly closed hull of $H$ in $E'$ is a metrizable compact space with respect to the weak topology.*

*Proof.* (cf. [3, Theorem 12.15.7]) □

## 2.2 Manifolds

As the main part of this thesis will be dealing with distributions on manifolds, a short introduction to the theory of abstract manifolds will be provided here. For more details I recommend Michael Kunzinger's lecture notes for his course on "Differential geometry" [13, Chapter 2].

An *(abstract) differentiable manifold* $X$ is defined to be a set $X$ together with an equivalence class of *atlases*. An atlas is a family of *charts* $\mathcal{A} = \{(\psi_i, V_i) : i \in I\}$, i.e. $\psi_i$ is a bijective map $V_i \to \psi_i(V_i)$ for $V_i \subseteq X$ and $\psi_i(V_i) \subseteq \mathbb{R}^n$ open, that are *pair-wise compatible*, which means that $\psi_1(V_1 \cap V_2)$ and $\psi_2(V_1 \cap V_2)$ are open subsets of $\mathbb{R}^n$ and the change of charts

$$\psi_2 \circ \psi_1^{-1} : \psi_1(V_1 \cap V_2) \to \psi_2(V_1 \cap V_2)$$

is a $\mathcal{C}^\infty$-diffeomorphism. In addition we have to have that $M = \bigcup_{i \in I} V_i$. Two atlases $\mathcal{A}_1$, $\mathcal{A}_2$ are called *equivalent* if $\mathcal{A}_1 \cup \mathcal{A}_2$ itself is an atlas of M, i.e., if all charts of $\mathcal{A}_1 \cup \mathcal{A}_2$ are compatible. An equivalence class of atlases is called a *differentiable (or $\mathcal{C}^\infty$-) structure on $X$*.

We can define a topology on the manifold $X$ with the help of the chart domains of a so-called *maximal atlas*, i.e. an atlas to which all possible compatible charts have been added. The chart domains then give a basis of a uniquely defined *manifold topology of* $X$, with respect to which, the charts $(\psi, V)$ are all homeomorphisms of $V \to \psi(V)$. (In the special case that the manifold $X$ is embedded into some $\mathbb{R}^m$ this manifold topology coincides with the trace topology of $\mathbb{R}^m$ on $X$.)

The topology on $X$ allows us to say what continuous functions are. Smooth functions $f : X \to Y$, for $X$ and $Y$ two $\mathcal{C}^\infty$-manifolds, are defined in the following way: $f$ is called *smooth (or $\mathcal{C}^\infty$)* if it is continuous and if for all $p \in X$ there exists a chart $\phi$ of $X$ around $p$ and a chart $\psi$ of $Y$ around $f(p)$ such that

$$\psi \circ f \circ \phi^{-1}$$

is smooth. $f$ is called a *diffeomorphism* if it is bijective and if $f$ and $f^{-1}$ are both smooth.

Note that this definition is fairly easy because we simply transfer the problem of differentiability to open sets in $\mathbb{R}^n$ and $\mathbb{R}^m$, where we know what smoothness of a function is. In most situations this notion of differentiability will suffice, however in Chapter 4 we will need a more refined concept, which I will shortly outline here: Before one can say what the derivative of a function on a manifold is, one needs to locally approximate the manifold by a vector space in order to have access to some sort of linear structure. We stick to the idea that the derivative of a function is the linear best approximation, and therefore we define the following concepts: Let $M$ be a manifold, $p \in M$ with $(\psi, V)$ a chart around $p$ and let $I \subseteq \mathbb{R}$ be an interval:

(i) Two $\mathcal{C}^\infty$-curves $c_1, c_2 : I \to M$ with $c_1(0) = p = c_2(0)$ are called *tangential* at $p$ with respect to $\psi$ if $(\psi \circ c_1)'(0) = (\psi \circ c_2)'(0)$.

This notion is however independent of the respective chart and so we may define an equivalence relation on the set of all curves $c$ near $p$ by:

$$c_1 \sim c_2 :\Leftrightarrow c_1 \text{ and } c_2 \text{ are tangential at } p.$$

We denote the equivalence classes by $[c]_p$.

(ii) The *tangent space* of $M$ at $p$ is defined as $T_pM := \{[c]_p \mid c : I \to M\ \mathcal{C}^\infty, c(0) = p\}$.

(iii) Let $N$ be another manifold, $f : M \to N$ $\mathcal{C}^\infty$. Then $T_p f : T_pM \to T_{f(p)}N, [c]_p \mapsto [f \circ c]_{f(p)}$ is the *derivative* of $f$ at $p$.

Using the derivative of a chart $(\psi, V)$ one obtains a bijection

$$T_p\psi : T_pM \to T_{\psi(p)}\psi(V) \cong \mathbb{R}^n$$

that induces a well-defined vector space structure on $T_pM$, with respect to which the derivative as defined above is indeed linear.

Ultimately we define:

- The *tangent bundle of a manifold* $M$ is defined as $TM := \bigcup\{p\} \times T_pM$.
- The *cotangent bundle of a manifold* $M$ is defined as $T^*M := \bigcup\{p\} \times (T_pM)^*$, i.e. the fibres of the cotangent bundle are the dual spaces of the fibres of the tangent bundle.
- Let $f \in \mathcal{C}^\infty(M, \mathbb{R})$, then the *exterior derivative* $df$ *of* $f$ is given by the map $df : M \to T^*M,\ p \mapsto T_pf$.

Hence the exterior derivative of a function is a smooth section of the cotangent bundle and as such, it is a one-form.

We stop our discussion of differentiation on manifolds and return to their topological properties: First of all manifolds always fulfil the first separation axiom, i.e. they are $T_1$. Unfortunately $T_2$ is not necessarily given, but is often assumed by convention because one does not want to work with topologies that are not Hausdorff. In particular, if the topology of a manifold is known to be Hausdorff, then the manifold is locally compact, a very useful property in topological constructions.

What is more, abstract manifolds fulfil the first axiom of countability and they have a countable basis of the topology if and only if there exists a countable atlas. (Therefore all compact manifolds are in particular second countable.)

We will assume all manifolds to be Hausdorff and second countable, not only because it is generally more convenient but also because these two properties furnish us with some very crucial tools for distribution theory. Firstly, we will need the partition of unity:

We have that if the manifold $X$ is second countable and Hausdorff, then for any open cover $\mathcal{U}$ of $X$ there exists a partition of unity $\{\chi_j \mid j \in \mathbb{N}\}$ subordinate to $\mathcal{U}$ such that, for all $j$, $\mathrm{supp}\chi_j$ is compact and contained in a chart domain.

Consequently, this will guarantee us the existence of cut-off functions around compact sets $K$ whose support lies in some open neighbourhood around $K$.

And finally, we will need that any second countable Hausdorff manifold possesses an exhaustion by compact sets, i.e. there exists a family of compact subsets $(K_j)_{j\in\mathbb{N}}$ of $M$ with $K_j \subseteq K_{j+1}^{\circ}$ for all $j$, such that $M = \bigcup_{j\in\mathbb{N}} K_j$.

## 2.3 Distributions on $\mathbb{R}^n$

Since $n$-dimensional manifolds locally look like open subsets of $\mathbb{R}^n$ and charts allow us to transfer the topological structure in a nice way, distributions on manifolds share most of their properties with distributions on $\mathbb{R}^n$. Indeed, for many of the statements later we will initially argue why it suffices to show the statement for distributions on open sets. Therefore, a survey of the most important results in distribution theory on $\mathbb{R}^n$ will be given here. For details and proofs see for instance the book by Friedlander and Joshi [8].

For $\Omega$ an open subset of $\mathbb{R}^n$ the space of *test functions* is defined as

$$\mathcal{D}(\Omega) := \{\varphi \in \mathcal{C}^\infty(\Omega) |\ \mathrm{supp}(\varphi) \text{ is compact}\},$$

where

$$\mathrm{supp}(\varphi) := \overline{\{x \in \Omega \mid \varphi(x) \neq 0\}}^{\Omega}$$

is called the *support of* $\varphi$. We endow the space of test functions with the countable strict inductive limit topology induced by the Fréchet spaces

$$\mathcal{D}(K_i) := \{\varphi \in \mathcal{C}^\infty(\Omega) \mid\ \mathrm{supp}(\varphi) \subseteq K_i\},$$

for $(K_i)_{i\in I}$ an exhaustion by compact sets of $\Omega$. Such a topology is also referred to as an *(LF)-structure.*

A note on this topology: Roughly speaking a strict inductive limit topology is a topology on a vector space $E$ that is induced by an increasing sequence $(E_n)$ of locally convex vector spaces with the properties that $E = \bigcup E_n$ and the topology induced by $E_{n+1}$ on $E_n$ is identical to the topology initially given on $E_n$. We define on $E$ the locally convex topology given by the collection of zero-neighbourhoods $\mathcal{A}$ of all convex sets $A$ in $E$ such that $A \cap E_n$ is a neighbourhood of 0 in $E_n$ for each $n \in \mathbb{N}$.

If all $E_n$ are Fréchet spaces we obtain an inductive limit topology with the following properties (cf. Trèves [18, chapter 13]):

- $E_n$ is a topological subspace of $E$. (This remains true even if the 'building blocks' are only locally convex vector spaces.)

- $E$ is complete.
- A set $B$ is bounded in $E$ if and only if it lies in some $E_n$ and is bounded there.
- Let $u : E \to F$ be a linear map into some locally convex vector space $F$. Then $u$ is continuous if and only if, for each $n$, the restriction of $u$ to $E_n$ is a continuous linear map into $F$.

These facts are important for the theory of distribution as they give that sequential continuity of a distribution is equivalent to continuity: Let $u$ be a sequentially continuous functional on $\mathcal{D}(\Omega)$. Then the restriction of $u$ to any $\mathcal{D}(K_i)$ is sequentially continuous, but on the Fréchet spaces $\mathcal{D}(K_i)$ this is equivalent to continuity. Therefore $u$ is continuous on $\mathcal{D}(\Omega)$.

Thus what needs to be known about the (LF)-structure on $\mathcal{D}(\Omega)$ is that it gives rise to the following concept of convergence: Let $(\varphi_n)$ be a sequence in $\mathcal{D}(\Omega)$, $\varphi \in \mathcal{D}(\Omega)$, then

$$\varphi_n \to \varphi \iff \begin{cases} (1) \;\; \exists K \subset\subset \Omega : \varphi_n,\ \varphi \in \mathcal{D}(K)\ \forall n, \text{ and} \\ (2) \;\; \forall \alpha : \partial^\alpha \varphi_n \to \partial^\alpha \varphi \text{ uniformly on } K. \end{cases}$$

*Distributions* are linear and (sequentially) continuous functionals on $\mathcal{D}(\Omega)$. Continuity can be characterized by the semi-norm estimate

$$\forall K \subset\subset \Omega\ \exists C > 0\ \exists m \in \mathbb{N} : |\langle u, \varphi \rangle| \leq C \sum_{|\alpha| \leq m} \|\partial^\alpha \varphi\|_{L^\infty(K)} \qquad \forall \varphi \in \mathcal{D}(K).$$

A distribution $u$ is called *regular*, if there is a function $f \in \mathcal{C}^\infty(\Omega)$ such that the action of the distribution on a test function is given by

$$\langle u, \varphi \rangle = \int_\Omega f(x)\varphi(x)dx \qquad \forall \varphi \in \mathcal{D}(\Omega).$$

A distribution $u \in \mathcal{D}'(\Omega)$ is said to be *of order* $m$ if, in the semi-norm estimate of $u$, $m$ can be chosen independently of $K$ and $m$ is minimal with that property. That means that the action of the distribution on any test function $\varphi$ only depends on derivatives of $\varphi$ of order less or equal to $m$.

A sequence of distributions $u_n$ is said to *converge to* $u \in \mathcal{D}'(\Omega)$ if it converges pointwise, i.e. $\lim \langle u_n, \varphi \rangle = \langle u, \varphi \rangle \quad \forall \varphi \in \mathcal{D}'(\Omega)$.

Even though distributions in $\mathcal{D}'(\Omega)$ cannot be evaluated at points $x \in \Omega$, they can be localised to open subsets $\Omega'$ of $\Omega$: $u|_{\mathcal{D}(\Omega')} \equiv u|_{\Omega'}$ is defined as the map

$$\begin{aligned} u : \mathcal{D}(\Omega') &\to \mathbb{C} \\ \varphi &\mapsto \langle u, \varphi \rangle. \end{aligned}$$

Two distributions $u, v \in \mathcal{D}(\Omega)$ are said to *coincide on* $\Omega'$, if $u|_{\Omega'} = v|_{\Omega'}$. Consequently we may define the *support of a distribution* $u \in \mathcal{D}(\Omega)$ as

$$\operatorname{supp}(u) := \Omega \backslash \{x \in \Omega | u = 0 \text{ on a neighbourhood of } x\}.$$

Such localisations to open subsets of $\Omega$ retain enough information about the distribution, that if we know a distribution on an open cover of $\Omega$, the distribution is uniquely determined. More precisely:
Let $(\Omega_i)_{i \in I}$ be an open cover of $\Omega$ and $u_i \in \mathcal{D}'(\Omega_i)$ such that

$$u_i|_{\Omega_i \cap \Omega_j} = u_j|_{\Omega_i \cap \Omega_j} \quad \forall i, j \text{ with } \Omega_i \cap \Omega_j \neq \emptyset.$$

Then there exists a unique distribution $u \in \mathcal{D}(\Omega)$ with $u|_{\Omega_i} = u_i \quad \forall i \in I$.

Having defined the support of a distribution, we can also look at distributions whose support is compact. These can be uniquely extended to functionals on $\mathcal{C}^\infty(\Omega)$ and satisfy the semi-norm estimate

$$\exists K \subset\subset \Omega \; \exists m \in \mathbb{N} \; \exists C > 0 : |\langle u, \varphi \rangle| \leq C \sum_{|\alpha| \leq m} \|\partial^\alpha \varphi\|_{L^\infty(K)} \quad \forall \varphi \in \mathcal{C}^\infty(\Omega).$$

This in particular implies that distributions with compact support are of finite order.

We move on to operations that can be defined on distribution spaces. The definitions of such operations are chosen in a way so that they turn out to be extensions of the corresponding operators on test functions.

Firstly, any distribution $u \in \mathcal{D}(\Omega)$ can be multiplied with a $\mathcal{C}^\infty$-function $f$ by the definition

$$\langle f \cdot u, \varphi \rangle := \langle u, f \cdot \varphi \rangle.$$

Nearly analogously we define derivatives of distributions

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

Defining the tensor product and convolution of distributions is a bit more difficult, and requires us to think about test functions that depend on additional parameters. Briefly put, one wants to define the tensor product $u \otimes v \in \mathcal{D}'(\Omega_1 \times \Omega_2)$ of two distributions

$u \in \mathcal{D}'(\Omega_1)$ and $v \in \mathcal{D}'(\Omega_2)$ acting on some $\chi \in \mathcal{D}(\Omega_1 \times \Omega_2)$ by

$$\langle u \otimes v, \chi \rangle := \langle v(y), \langle u(x), \chi(x,y) \rangle \rangle$$

because this gives for splitting tensors that

$$\langle u \otimes v, \varphi \otimes \psi \rangle := \langle u, \varphi \rangle \cdot \langle v, \psi \rangle \qquad \forall \varphi \in \mathcal{D}(\Omega_1),\ \forall \psi \in \mathcal{D}(\Omega_2).$$

Note that for this definition it has to be verified that $y \mapsto \langle u(x), \chi(x,y) \rangle \equiv \langle u, \chi(.,y) \rangle$ is a test function on $\Omega_2$, which it indeed turns out to be. The tensor product is bilinear and separately continuous.

For defining the convolution of two distributions, some sacrifices have to be made in order to maintain a compact support on the right-hand side of the bracket. A straightforward way of achieving this is by assuming that one of the distributions has compact support and multiplying the right-hand side by a cut-off around that support. So let $u \in \mathcal{D}'(\mathbb{R}^n)$ and $v \in \mathcal{D}'(\mathbb{R}^n)$ with $\operatorname{supp}(v)$ compact and choose a cut-off $\rho \in \mathcal{D}(\mathbb{R}^n)$ with $\rho \equiv 1$ on a neighbourhood of $\operatorname{supp}(v)$. Then the convolution $u * v \in \mathcal{D}'(\mathbb{R}^n)$ is defined by

$$\langle u * v, \varphi \rangle := \langle u(x) \otimes v(y), \rho(y)\varphi(x+y) \rangle \qquad \varphi \in \mathcal{D}(\mathbb{R}^n).$$

This definition does not depend on the choice of the cut-off and the convolution is separately continuous.

Convolution can be used to regularise distributions: More precisely, let $\rho \in \mathcal{C}^\infty(\mathbb{R}^n)$, $u \in \mathcal{D}'(\mathbb{R}^n)$ and one of them with compact support. Then

$$(u * \rho)(x) = \langle u(y), \rho(x-y) \rangle \in \mathcal{C}^\infty(\mathbb{R}^n).$$

This property, separate continuity and the fact that the $\delta$-distribution is the neutral element of convolution can be used to show that the space of test functions $\mathcal{D}(\mathbb{R}^n)$ is dense in $\mathcal{D}'(\mathbb{R}^n)$, resp. that $\mathcal{D}(\Omega)$ is dense in $\mathcal{D}'(\Omega)$. (One convolutes any distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ with mollifiers $\rho_n \to \delta$ (for $n \to \infty$). Then $\mathcal{D}(\mathbb{R}^n) \ni u * \rho_n \to u * \delta = u$.)

Finally we turn to the Fourier transform for distributions and again we start by looking at how the Fourier transform acts on test functions: The Fourier transform of a function $\varphi \in \mathcal{D}(\Omega) \hookrightarrow \mathcal{D}(\mathbb{R}^n)$ is given by

$$\hat{\varphi}(\xi) = \int_{\mathbb{R}^n} e^{-i\langle x, \xi \rangle} \varphi(x) dx.$$

So one would be tempted to define the Fourier transform of a distribution $u \in \mathcal{D}'(\Omega)$ by simply setting

$$\langle \hat{u}, \varphi \rangle := \langle u, \hat{\varphi} \rangle,$$

yet there is a major problem with this. The Fourier transform of a test function is never compactly supported, unless $\varphi \equiv 0$. Therefore one either has to limit the definition to $u \in \mathcal{E}'(\Omega)$, the space of compactly supported distributions which can act on any $\mathcal{C}^\infty$-function, or introduce *tempered distributions*. These are defined as the dual space of the Schwartz space

$$\mathcal{S}(\mathbb{R}^n) := \{\psi \in \mathcal{C}^\infty(\mathbb{R}^n) \mid \sup_{x \in \mathbb{R}^n} |x^\alpha D^\beta \psi(x)| < \infty \ \forall \alpha, \beta \in \mathbb{N}^n\}.$$

Fourier transform is an isomorphism on $\mathcal{S}(\mathbb{R}^n)$, and so there the definition works too. For $u \in \mathcal{E}'(\Omega)$ we even have that $\hat{u} \in \mathcal{O}_M(\mathbb{R}^n) \subseteq \mathcal{C}^\infty(\mathbb{R}^n)$ where

$$\mathcal{O}_M := \{f \in \mathcal{C}^\infty(\mathbb{R}^n) \mid \forall \alpha \in \mathbb{N}_0^n \ \exists N \in \mathbb{N} \ \exists C > 0 \ \forall x \in \mathbb{R}^n : |\partial^\alpha f(x)| \leq C(1 + |x|)^N\}$$

and

$$\hat{u}(\xi) = \langle u, e^{-i\langle \cdot, \xi \rangle} \rangle.$$

## 2.4 Distributions on Manifolds

On $\mathbb{R}^n$, every locally integrable function defines a distribution by the natural assignment

$$\varphi \mapsto \int f(x) \cdot \varphi(x) dx. \tag{2.1}$$

However on manifolds we are faced with the problem that there is no a-priori notion of integration. In order to keep the natural identification of functions and distributions, the integral of the product has to be rendered meaningful in some way or other.

Therefore one-densities come into play as objects that can naturally be integrated on manifolds, as they carry information on the volume element. (Even though a certain measure or $n$-form (on oriented manifolds) for integration would basically work, such an approach would be restrictive and impractical.) Before going into detail let me point out that the main objective is to make the product in (2.1) a one-density so that it becomes integrable. As the product of a one-density with a function is again a one-density, this leaves us with a choice of putting the burden of being integrable either on the test function or on the distribution. What is more, the product of a $q$-density and a $q'$-density is a

$(q+q')$-density, so the "volume element" could also be shared by both objects in question. All these three approaches are valid and which one to choose is more a matter of taste. We will work with the case where the test objects remain functions and the distributions carry the density-character.

In addition when defining distributions on manifolds one can on a more general level consider test objects that map from $M$ into some arbitrary vector bundle $E$, instead of simply into the complex plane. As this causes only little inconvenience in notation the general case will be introduced here, although later we will be working only with test objects that map into $\mathbb{C}$. We now begin with a more detailed discussion (for more information cf. [9, 3.1]):

**Definition 2.8.** Let $W$ be a real vector space of dimension $n$ and $\Lambda^n W$ be the *$n$-fold antisymmetrised tensor product of $W$*. A *one-density $\mu$ on $W$* is a map

$$\mu : \Lambda^n W \setminus \{0\} \to \mathbb{C}$$

such that for all $0 \neq s \in \mathbb{R}$ and for all $0 \neq \omega \in \Lambda^n W$ we have

$$\mu(s\omega) = |s| \cdot \mu(\omega).$$

Since $\Lambda^n W$ is one-dimensional, a one-density $\mu$ on $W$ is uniquely determined by its value on a particular $\omega \neq 0$. The vector space of all one-densities on $W$ is denoted by $\mathrm{Vol}(W)$.

Because we want to use these densities for integrating, we want to know about their transformation behaviour under a change of variables, which is just as we would require it to be. Namely let $(v_i)_{i=1}^n$ and $(w_i)_{i=1}^n$ be two bases of $W$ and $v_i = \sum a_{ij} w_j$, then we have

$$\mu(v_1 \wedge \cdots \wedge v_n) = |\det A| \cdot \mu(w_1 \wedge \cdots \wedge w_n). \tag{2.2}$$

(Compare this to the transformation behaviour of $n$-forms on a manifold, which do not give the modulus of the determinant. Hence $n$-forms can only be used to define the integral of *orientable* manifolds, where the positive sign of the determinant can be forced by choosing an oriented atlas. On oriented manifolds the concepts of densities and $n$-forms coincide.)

We can use the given transformation behaviour to define one-densities on manifolds as sections of a (complex) line bundle whose transition functions are the analogue of (2.2). Bearing in mind that a given family of transition functions uniquely determines a vector bundle over $X$ we only need to recall some basic facts of vector bundles and sections before we can give the definition of the volume bundle over $X$.

A vector bundle with base space $X$ will be denoted by $(E, X, \pi)$ and for a chart $(V, \psi)$ in $X$, a *vector bundle chart* $(V, \Psi)$ over $\psi$ will be written in the form

$$\begin{aligned}\Psi : \pi^{-1}(V) &\to \psi(V) \times \mathbb{K}^{n'} \\ z &\mapsto (\psi(p), \psi^1(z), \ldots, \psi^{n'}(z) \equiv (\psi(p), \boldsymbol{\psi}(z)),\end{aligned}$$

where $p = \pi(z)$ and the typical fibre is $\mathbb{K}^{n'}$. Let $(V_\alpha, \Psi_\alpha)$ denote a vector bundle atlas, then we write

$$\Psi_\alpha \circ \Psi_\beta^{-1}(y, w) = (\psi_{\alpha\beta}(y), \boldsymbol{\psi}_{\alpha\beta}(y)w),$$

where $\psi_{\alpha\beta} := \psi_\alpha \circ \psi_\beta^{-1}$ is the change of chart on the base and $\boldsymbol{\psi}_{\alpha\beta} : \psi_\beta(V_\alpha \cap V_\beta) \to \mathrm{GL}(n', \mathbb{K})$ denotes the *transition function*.

For a given vector bundle $E$ over $X$ we will denote the *dual bundle*, that is the bundle whose fibres are the dual spaces of the fibres of $E$, by $E^*$.

We will denote spaces of $\mathcal{C}^k$*-sections* $(0 \leq k \leq \infty)$ in the bundle $(E, X, \pi)$ (with compact support) by $\Gamma^k(X, E)$ (resp. by $\Gamma^k_c(X, E)$), and local sections over the open set $V \subseteq X$ by $\Gamma^k(V, E)$ (resp. by $\Gamma^k_c(V, E)$). The *local expression* of a section $u \in \Gamma(X, E)$ in a chart $(V_\alpha, \Psi_\alpha)$ is defined as

$$\begin{aligned}u_\alpha \equiv (\Psi_\alpha)_* u|_{V_\alpha} \equiv \Psi_\alpha \circ u|_{V_\alpha} \circ \psi_\alpha^{-1} &\in \Gamma(\psi_\alpha(V_\alpha), \mathbb{R}^{n'}) \\ x &\mapsto (x, u^1_\alpha(x), \ldots, u^{n'}_\alpha(x)),\end{aligned}$$

where $u^i_\alpha = \psi^i_\alpha \circ u \circ \psi_\alpha^{-1} \in \mathcal{C}^\infty(\psi_\alpha(V_\alpha))$.

After this prelude we can now define the volume bundle over a manifold.

**Definition 2.9.** Let $(V_\alpha, \psi_\alpha)_\alpha$ be an atlas for $X$. The *volume bundle over* $X$ denoted by $\mathrm{Vol}(X)$ is the one-dimensional complex vector bundle given by the following cocycle of transition functions

$$\begin{aligned}\boldsymbol{\psi}_{\alpha\beta} : \psi_\beta(V_\alpha \cap V_\beta) &\to \mathbb{R} \setminus \{0\} = \mathrm{GL}(1, \mathbb{R}) \\ \boldsymbol{\psi}_{\alpha\beta}(y) :&= |\det D(\psi_\alpha \circ \psi_\beta^{-1})(y)|^{-1} \\ &= |\det D(\psi_\beta \circ \psi_\alpha^{-1})(x)|\end{aligned}$$

where $x = \psi_{\alpha\beta}(y)$.

It can be verified that the definition via the transition functions actually gives the vector bundle $\bigcup_{p \in X} \{p\} \times \mathrm{Vol}(T_p X)$.

**Definition 2.10.** For $\mu \in \Gamma_c^0(X, \mathrm{Vol}(X))$, $(V_\alpha, \psi_\alpha)_\alpha$ an atlas ($V_\alpha$ relatively compact) for $X$ and $(f_\alpha)_\alpha$ a partition of unity subordinate to $(V_\alpha)_\alpha$ we can finally define the integral of $\mu$ by

$$\int_X \mu = \sum_\alpha \int_{V_\alpha} f_\alpha \mu := \sum \int_{\psi_\alpha(V_\alpha)} f_\alpha(\psi_\alpha^{-1}(x))\mu_\alpha(x)dx,$$

where $\mu_\alpha$ is the local expression of $\mu$ in the chart $(V_\alpha, \psi_\alpha)$.

That this definition works and is independent of the choice of the atlas and the partition of unity is due to the fact that the densities so nicely emulate the transformation behaviour of multiple integrals. In two steps this definition can be extended to so-called *non-negative* densities that take values only in $\mathbb{R}_0^+$ and then to any non-compactly supported densities. Since every one-density on a manifold gives rise to a measure by which we can integrate, we will sometimes write $\int_X f(x)d\mu(x)$ instead of $\int_X f\mu$.

Having settled the question of how to integrate on manifolds we can now turn to defining test functions and distributions. We want our test objects to be compactly supported sections in a bundle over the manifold $X$, therefore we need to define a suitable topology on section spaces. To do this we repeat the construction of the $\mathbb{R}^n$ case with slight modulations.

One first looks at the spaces $\Gamma^\infty(V, F) \equiv \Gamma(V, F)$ and endows them with a topology that turns them into Fréchet spaces. More precisely, fix a vector bundle atlas $(V_\alpha, \Psi_\alpha)_{\alpha \in A}$ and a fundamental sequence of compact subsets $(K_m^\alpha)_m$ of $\psi_\alpha(V_\alpha)$ and then define the family of semi-norms

$$p_{s,m,\alpha}(u) := \sum_{j=1}^{n'} \sup_{x \in K_m^\alpha, |\nu| \leq s} |\partial^\nu(\psi_\alpha^j \circ u|_{V_\alpha} \circ \psi_\alpha^{-1}(x))|,$$

where $s \in \mathbb{N}_0$ and $m \in \mathbb{N}$, $\alpha \in A$. In this topology convergence in $\Gamma^\infty(V, F)$ translates to convergence in all $\Gamma^\infty(V_\alpha, F) \cong \mathcal{C}^\infty(\psi_\alpha(V_\alpha))^{n'}$, i.e. convergence of coefficient functions.

Now we look at the space $\Gamma_{c,K}(X, F) := \{u \in \Gamma(X, F) \mid \mathrm{supp}(u) \subseteq K\}$ for $K \subset\subset X$, which is a closed subspace of some $\Gamma(V, F)$, for an appropriate choice of $V$, and therefore a Fréchet space as well (cp. $\mathcal{D}(K) \subseteq \mathcal{D}(\Omega)$). Completely analogous to the $\mathbb{R}^n$ case we define $\Gamma_c(X, F)$ as the inductive limit of the Fréchet spaces $\Gamma_{c,K}(X, F)$ and as such it is a locally convex vector space that induces the Fréchet space topology on each of the $\Gamma_{c,K}(X, F)$. What is more every bounded subset of $\Gamma_c(X, F)$ is contained (and bounded) in $\Gamma_{c,K}(X, F)$ for some $K \subset\subset X$. Finally we may now define distributions (or also distributional densities):

**Definition 2.11.** Let $E$ denote an arbitrary vector bundle over $X$. *Distributions with values in* $E$ are continuous linear forms on $\Gamma_c(X, E^*)$. We denote the space of distributions with values in $E$ by $\mathcal{D}'(X, E \otimes \mathrm{Vol}(X))$.

Some notes on this definition:

(i) The most general version of this definition using $q$- and $(1-q)$-densities would define the space of distributions as the topological dual of the space $\Gamma_c(X, E^* \otimes \mathrm{Vol}^{1-q}(X))$ and denote the thus emergent space of *$E$-valued distributions of densitiy character $q$* by $\mathcal{D}'(X, E \otimes \mathrm{Vol}^q(X))$.

(ii) A special case of our definition that will be important for us is if $E = X \times \mathbb{C}$, when we obtain the space of *distributions on the manifold $X$*. (If furthermore $X = \Omega$ an open subset of $\mathbb{R}^n$ we recover the usual distribution space.) We denote this space simply by $\mathcal{D}'(X)$.

(iii) It should not be a surprise that again analogous to the $\mathbb{R}^n$ case there is a semi-norm characterisation of distributions with values in $E$ using the semi-norms $p_{s,m,\alpha}$ defined above: A linear functional $T$ on the space $\Gamma_c(X, E^*)$ is a distribution in the above sense iff

$$\forall K \subset\subset X \ \exists s \in \mathbb{N}_0, m \in \mathbb{N} \ \exists \alpha_1, \ldots, \alpha_l \ \exists C_K > 0 : |T(\varphi)| \leq C_K \sup_i p_{s,m,\alpha_i}(\varphi),$$

for all $\varphi \in \Gamma_{c,K}(X, E^*)$.

We are still left with checking that this definition of distributions indeed solves the issue of providing a natural way to identify (locally integrable) functions with distributions. To this end we consider the canonical vector bundle homomorphism

$$\mathrm{tr}_{(E)} : E \otimes E^* \to X \times \mathbb{C}$$

that is induced by the dual action of $v^* \in E_p^*$ on $v \in E_p$. Then we can assign to a pair $(f, u) \in \Gamma^0(X, E \otimes \mathrm{Vol}(X)) \times \Gamma_c(X, E^*)$ a density which may be integrated via the vector bundle homomorphism

$$(. \mid .) := \mathrm{tr}_{(E)} \otimes \mathrm{id}_{\mathrm{Vol}(X)} : E \otimes E^* \otimes \mathrm{Vol}(X) \to (X \times \mathbb{C}) \otimes \mathrm{Vol}(X) = \mathrm{Vol}(X).$$

Therefore we can define the action of some locally integrable section $f \in \Gamma^0(X, E \otimes \mathrm{Vol}(X))$ on a test function $\varphi \in \Gamma_c(X, E^*)$ by

$$T_f(\varphi) \equiv \langle f, \varphi \rangle := \int_X (f \mid \varphi).$$

By the properties of the integral the map $T_f$ is linear and continuous, and hence defines a distribution. Moreover the assignment $f \mapsto T_f$ is linear and injective.

Ultimately we close this paragraph by investigating the relationship between distributions on $\mathbb{R}^n$ and those on manifolds. As already mentioned before, locally these two concepts agree in a very nice way. More precisely we have the following theorem:

**Theorem 2.12.** (Local Description of Distributions) *Distributions with values in* $E \otimes \mathrm{Vol}(X)$ *can be identified with families* $(T_\alpha)_\alpha$ *of distributions* $T_\alpha \in \mathcal{D}'(\psi_\alpha(V_\alpha), \mathbb{R}^{n'})$ *satisfying the transformation law*

$$T_\alpha = (\Psi_\alpha \circ \Psi_\beta^{-1})^\wedge(T_\beta),$$

*where* $\Psi^\wedge := (\Psi_*^{-1})'$.

*Proof.* I will only give a rough outline of the proof, for details cf.[9, 3.1.4].

Let $(V_\alpha, \Psi_\alpha)$ denote a chart of $(E^*, X, \pi)$. We start by looking at the map $(\Psi_\alpha)_* := \Psi_\alpha \circ u \circ \psi_\alpha^{-1} : \Gamma(V_\alpha, E^*) \to \Gamma(\psi_\alpha(V_\alpha), \mathbb{R}^{n'})$. Then we restrict it to sections with compact support to obtain

$$(\Psi_\alpha)_* : \Gamma_c(V_\alpha, E^*) \to \Gamma_c(\psi_\alpha(V_\alpha), \mathbb{R}^{n'}),$$

which can be shown to be an isomorphism of (LF)-spaces (using that $\Psi$ is a vector bundle isomorphism). Then we consider the adjoint map of $(\Psi_\alpha)_*^{-1}$, namely

$$\begin{aligned} \Psi_\alpha^\wedge \equiv ((\Psi_\alpha)_*^{-1})' &: \Gamma_c(V_\alpha, E^*)' \to \Gamma_c(\psi_\alpha(V_\alpha), \mathbb{R}^{n'})' \\ \langle \Psi_\alpha^\wedge T, u\rangle &:= \langle T, (\Psi_\alpha)_*^{-1} u\rangle = \langle T, (\Psi_\alpha)^{-1} \circ u \circ \psi\rangle. \end{aligned}$$

But upon closer inspection we see that the topological isomorphism defined by $\Psi_\alpha^\wedge$ actually gives the desired identification, since

$$\Psi_\alpha^\wedge : \mathcal{D}'(V_\alpha, E \otimes \mathrm{Vol}(X)) \to \mathcal{D}'(\psi_\alpha(V_\alpha), \mathbb{R}^{n'}).$$

So we can assign to any distribution $T \in \mathcal{D}'(V_\alpha, E \otimes \mathrm{Vol}(X))$ a family $(T_\alpha)_\alpha$ of the respective *local expressions* $T_\alpha := \Psi_\alpha^\wedge(T|_{V_\alpha}) \in \mathcal{D}'(\psi_\alpha(V_\alpha), \mathbb{R}^{n'})$ satisfying $\langle T, u\rangle = \langle T_\alpha, u_\alpha\rangle$, as can be easily verified.

Straightforward calculation gives the transformation behaviour. $\square$

**Remark 2.13.** (i) This theorem will be useful when simplifying problems on manifolds to open subsets of $\mathbb{R}^n$. Especially once we have managed to localise distributions to chart domains with the use of partition of unity, it allows us to identify such a local distribution with a distribution on an open subset on $\mathbb{R}^n$, since in this case we only have one chart and hence only one local expression.

(ii) In addition to this theorem describing $E$-valued distributions locally, there is a theorem shedding light on the global structure of the distribution space. More precisely they can be viewed as sections in the bundle $E \otimes \mathrm{Vol}(X)$ with distributional coefficients, i.e.

$$\mathcal{D}'(X, E \otimes \mathrm{Vol}(X)) \cong \mathcal{D}'(X) \otimes_{\mathcal{C}^\infty(X)} \Gamma(X, E \otimes \mathrm{Vol}(X)).$$

# 3 Kernel Distributions

Kernel distributions are generalisations of kernels in integral transforms, which are maps of the type

$$f \mapsto kf(y) := \int_X k(x,y)f(x)dx. \tag{3.1}$$

Typically, integral transforms map a class of functions defined on $X$ to functions defined on $Y$ with the help of a kernel function $k$ that is defined on $X \times Y$.

In the more general distribution setting one therefore starts off with a kernel distribution $K \in \mathcal{D}'(X \times Y)$ and looks at the map

$$\varphi \mapsto K \cdot \varphi \in \mathcal{D}'(Y)$$
$$\langle K \cdot \varphi, \psi \rangle = \langle K, \varphi \otimes \psi \rangle$$

for $\varphi \in \mathcal{D}(X), \psi \in \mathcal{D}(Y)$. The kernel theorem states that in this way, $K$ gives rise to a continuous map $\mathcal{D}(X) \to \mathcal{D}'(Y)$ and that, conversely, every mapping of this type is generated by a kernel distribution $K \in \mathcal{D}'(X \times Y)$.

Especially the second part of this result is quite astonishing, since it does not hold in the classical case. For instance in an $L^2$ setting, by far not every bounded operator from $L^2(X)$ to $L^2(Y)$ can be represented by a kernel $k(x,y) \in L^2(X \times Y)$. Indeed, for $X = Y$ not even the identity map can be given by an $L^2$-kernel, since

$$f(y) = \int_X k(x,y)f(x)dx$$

is solved by the Delta-distribution $k(x,y) = \delta(x-y)$. Since we will show that kernels are unique even in the broader distribution setting, $\delta(x-y)$ has to be the unique solution to the equation. And $\delta(x-y)$ obviously does not lie in $L^2(X \times X)$.

So in the first paragraph of this chapter we will give the proof of the kernel theorem and state some immediate consequences. We will also look at a wide variety of operators and compute their kernels. In the second paragraph on *regular kernel distributions* we have a closer look at a special type of kernel map, which maps not only into $\mathcal{D}'(Y)$, but more specifically into the set of smooth functions on $Y$, $\mathcal{E}(Y)$. These operators will allow us to extend the domains of kernel maps by using the *transposed map* of kernel distributions. In doing so, we can gain kernel operators that map $\mathcal{E}'(X) \to \mathcal{D}'(Y)$, or even larger distribution spaces if additional conditions are fulfilled. Ultimately we briefly look at *regularising operators*, that show an even nicer behaviour. They are the $\mathcal{C}^\infty$-analogue to integral operators as introduced above.

For more information cf. Dieudonné [5, 23.9-23.11].

## 3.1 The Kernel Theorem

One more warning before stating the kernel theorem: By common abuse of notation both the kernel $K \in \mathcal{D}'(X \times Y)$ and the associated operator from $\mathcal{D}(X)$ to $\mathcal{D}'(Y)$ are denoted by $K$. Here we will however refrain from doing so until the kernel theorem is proved, that is when we know it is safe to equate the two. This is why in part (ii) of the theorem, the operator is initially denoted by $A$.

**Theorem 3.1.** (L.- Schwartz) *Let $X$, $Y$ be two differentiable manifolds.*

*(i) Let $K$ be a distribution on $X \times Y$. Then for every $\varphi \in \mathcal{D}(X)$ the linear form defined as $\psi \to \langle K, \varphi \otimes \psi \rangle$ is a distribution on $\mathcal{D}(Y)$ denoted by $K \cdot \varphi$.*

*Furthermore the map $\varphi \mapsto K \cdot \varphi : \mathcal{D}(X) \to \mathcal{D}'(Y)$ is linear and continuous with respect to the usual topologies, i. e. $\forall L \subset\subset X$ the restriction of $K \cdot \varphi$ to $\mathcal{D}(L)$ is a continuous map from the (F)-space $\mathcal{D}(L)$ into the space $\mathcal{D}'(Y)$ equipped with the weak topology.*

*(ii) On the other hand, let $A$ be a linear operator $\mathcal{D}(X) \to \mathcal{D}'(Y)$ with the property that for each compact subset $L$ of $X$ the restriction of $A$ to $\mathcal{D}(L)$ is a continuous map from the (F)-space $\mathcal{D}(L)$ into the space $\mathcal{D}'(Y)$ equipped with the weak topology. Then there exists a unique distribution $K \in \mathcal{D}'(X \times Y)$ such that for $\varphi \in \mathcal{D}(X)$ and $\psi \in \mathcal{D}(Y)$*

$$\langle K, \varphi \otimes \psi \rangle = \langle K \cdot \varphi, \psi \rangle = \langle A(\varphi), \psi \rangle. \tag{3.2}$$

*Proof.* (i) In order to show that $K \cdot \varphi$ is a distribution on $\mathcal{D}(Y)$ we need to verify that for every compact subset $H$ of $Y$ the map $\psi \mapsto \langle K \cdot \varphi, \psi \rangle = \langle K, \varphi \otimes \psi \rangle$ is a continuous linear form on $\mathcal{D}(H)$. However using a finite partition of unity $(f_\alpha)$ on $Y$ subordinated to some open cover consisting of chart domains of $Y$ with $\sum f_\alpha(y) = 1$ for all $y \in H$ we see that we can effectively limit our inspection to the domain of a chart of $Y$.

Similarly we may write $\varphi$ as the finite sum of functions with small support and therefore assume that $\operatorname{supp}(\varphi)$ lies in the domain of some chart of $X$.

All in all we see that it suffices to show the desired continuity when the supports of both $\psi$ and $\varphi$ lie within respective charts of $X$ and $Y$, which allows us to use Theorem 2.12 on local descriptions of distributions on a manifold and assume that $X$ and $Y$ are open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$. Therefore we need to show that

$$\exists C \geq 0 \ \exists N \in \mathbb{N} : \langle K, \varphi \otimes \psi \rangle \leq C \sum_{|\alpha| \leq N} \|D^\alpha \psi\|_{L^\infty(H)} \quad \forall \psi \in \mathcal{D}(H).$$

Because $K$ is a distribution on the product space we know that

$$\forall A \subset\subset X \times Y \ \exists \tilde{C} \geq 0 \ \exists \tilde{N} \in \mathbb{N} \ \forall \Phi \in \mathcal{D}(A) : \langle K, \Phi \rangle \leq \tilde{C} \sum_{|\gamma| \leq \tilde{N}} \|D^\gamma \Phi\|_{L^\infty(A)},$$

which gives for $\operatorname{pr}_2(A) \supseteq H$ and $\Phi = \varphi \otimes \psi$ (hence $D^\gamma(\varphi \otimes \psi)(x,y) = D^\alpha \varphi(x) D^\beta \psi(y)$ with $(\alpha, \beta) = \gamma$), that $\exists \tilde{C} \geq 0 \ \exists \tilde{N} \in \mathbb{N} \ \forall \psi \in \mathcal{D}(H)$ :

$$\begin{aligned} \langle K \cdot \varphi, \psi \rangle = \langle K, \varphi \otimes \psi \rangle &\leq \tilde{C} \sum_{|\beta| \leq \tilde{N}} \|D^\beta \varphi\|_{L^\infty(X)} \cdot \sum_{|\alpha| \leq \tilde{N}} \|D^\alpha \psi\|_{L^\infty(H)} \\ &\leq C \cdot \sum_{|\alpha| \leq \tilde{N}} \|D^\alpha \psi\|_{L^\infty(H)}, \end{aligned}$$

for $C = \tilde{C} \sum_{|\beta| \leq \tilde{N}} \|D^\beta \varphi\|_{L^\infty(X)}$. Upon setting $N = \tilde{N}$ we obtain continuity of $K \cdot \varphi$ on $\mathcal{D}(H)$. As the map is clearly linear we have shown that $K \cdot \varphi \in \mathcal{D}'(Y)$.

Since also linearity of the map $\varphi \mapsto K \cdot \varphi$ is obvious, it only remains to show that the map $\varphi \mapsto K \cdot \varphi : \mathcal{D}(X) \to \mathcal{D}'(Y)$ is continuous. Due to the (LF)-structure of $\mathcal{D}(X)$ this means that for any compact subset $L$ of $X$ we have to estimate the $\mathcal{D}'(Y)$-semi-norms of $K \cdot \varphi$ by the $\mathcal{D}(L)$-norms of $\varphi \in \mathcal{D}(L)$, i.e. we have to show that for any arbitrary but fixed $\rho \in \mathcal{D}(Y)$

$$\exists C_\rho \geq 0 \ \exists N \in \mathbb{N} : |\langle K \cdot \varphi, \rho \rangle| \leq C_\rho \sum_{|\alpha| \leq N} \|D^\alpha \varphi\| \quad \forall \varphi \in \mathcal{D}(L).$$

But this can be obtained exactly as before by interchanging the roles of $X$ and $Y$.

(ii) Now we have to prove the existence and uniqueness of a distribution $K$ on the product space $X \times Y$ that fulfils

$$\langle K, \varphi \otimes \psi \rangle = \langle K \cdot \varphi, \psi \rangle = \langle A(\varphi), \psi \rangle.$$

We begin by discussing uniqueness, which can be seen rather easily: Formula (3.2) determines the continuous functional $K$ on the dense subspace

$$\mathcal{D}(X) \otimes \mathcal{D}(Y) \subseteq \mathcal{D}(X \times Y)$$

(cf. for instance in [8, Theorem 4.3.1]) and therefore, if such a $K$ exists, it has to be unique.

The existence is a lot harder to prove and will require some work. It will be shown in two steps and before going into all detail I want to sketch what is going to happen:

(1) In Step 1 we will assume that $X$ and $Y$ are open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively and that $A$ maps any test function $\varphi \in \mathcal{D}(X)$ to a distribution that can be given by a continuous function $f_\varphi \in \mathcal{C}(Y) \subset \mathcal{D}'(Y)$. Being a continuous function its action on a test function $\psi \in \mathcal{D}(Y)$ is given by

$$\langle A(\varphi), \psi \rangle = \langle f_\varphi, \psi \rangle = \int_Y f_\varphi(y)\psi(y)dy,$$

and we want the kernel $K \in \mathcal{D}'(X \times Y)$ to act on splitting tensors in just the same way. To this end we introduce the distribution $F_y(\varphi) := f_\varphi(y)$, in accordance with the above integral, and note that

$$F_y((\varphi \otimes \psi)(.,y)) = f_\varphi(y)\psi(y).$$

From this we see that $K \in \mathcal{D}'(X \times Y)$ can actually only be defined as

$$K(\Phi) := \int_Y F_y \Phi(.,y)dy, \text{ for all } \Phi \in \mathcal{D}(X \times Y).$$

Showing that this definition works will take some effort.

(2) In the second part of the proof we effectively reduce the problem of defining a suitable distribution $K \in \mathcal{D}'(X \times Y)$ to the case we dealt with in Step 1. That $X$ and $Y$ may be assumed to be open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ can be justified by using Theorem 2.12 and the fact that localisations, e.g. to the elements of an open cover consisting of chart domains, uniquely determine a distribution.

Obtaining distributions that can be given by a continuous function and that thus allow a definition of $K$ as above is harder to achieve. Yet essentially, we use the structure theorem of distributions that allows us to write any compactly supported distribution $\Lambda \in \mathcal{D}'(Y)$ as

$$\Lambda = (\partial_1 ... \partial_m)^{q+2}(E_{q+2} * \Lambda),$$

where, for an appropriate choice of $q$ and $E_{q+2}$, $(E_{q+2} * \Lambda)$ is a continuous function. The partial derivatives do not bother us much, as we can get rid of them by moving them to the test functions.

Having set the basic line of action, we begin to work out the details:

(1) First we assume that $X$ is an open subset of $\mathbb{R}^n$ and $Y$ an open subset of $\mathbb{R}^m$ and that for every $\varphi \in \mathcal{D}(X)$ the distribution $A(\varphi)$ on $\mathcal{D}(Y)$ can be identified with a continuous complex function denoted by $f_\varphi$ on $Y$. So now

$$A : \mathcal{D}(X) \to \mathcal{C}(Y)$$
$$A(\varphi) = f_\varphi$$

and the premise of the theorem says that this map is continuous from $\mathcal{D}(X)$ to $\mathcal{C}(Y)$, when $\mathcal{C}(Y)$ is equipped with the weak topology induced by $\mathcal{D}'(Y)$. But we can use Lemma 2.4 to obtain that $A$ is continuous even with respect to the finer Fréchet space topology on $\mathcal{C}(Y)$. Having established this we consider for every $y \in Y$ the linear map

$$\varphi \mapsto f_\varphi(y) \in \mathbb{C},$$

which we now know to be continuous on $\mathcal{D}(X)$ and which thus defines an element of $\mathcal{D}'(X)$. We denote this element by $F_y(\varphi)$, thus

$$F_y(\varphi) := f_\varphi(y).$$

Note that we can apply $F_y \in \mathcal{D}'(X)$ to a test function $\Phi \in \mathcal{D}(X \times Y)$ in the following way: We have

$$F_y(\Phi(.,y)) = f_{\Phi(.,y)}(y)$$

where $(f_{\Phi(.,y)})_{y \in Y}$ is a family of continuous functions on $Y$.

What we are interested in are the properties of the map

$$y \mapsto F_y(\Phi(.,y)).$$

Due to the compact support of $\Phi$ also the support of $y \mapsto F_y(\Phi(.,y))$ is compact (since for $y \notin \mathrm{pr}_2(\mathrm{supp}(\Phi))$ the function $\Phi(.,y)$ vanishes identically and $A$ is linear). We claim that $y \mapsto F_y(\Phi(.,y))$ is continuous: Let $(y_p)$ be a sequence converging to $y \in Y$, $H$ :=$\mathrm{pr}_1(\mathrm{supp}(\Phi))$ and $L$ a compact subset of $Y$ with $y_p, y \in L$ for all $p$. Then $\Phi$ is uniformly continuous on $H \times L$ and therefore we know that

$$\forall \epsilon > 0\ \exists \delta > 0 : |(x', y') - (x'', y'')| \leq \delta \Rightarrow |\Phi(x', y') - \Phi(x'', y'')| \leq \epsilon,$$

which in particular gives for large $p$ that

$$|(x, y_p) - (x, y)| \leq \delta \Rightarrow |\Phi(x, y_p) - \Phi(x, y)| \leq \epsilon \qquad \forall x \in H.$$

Repeating this argument for the derivatives of $\Phi$ yields that $\Phi(., y_p) \to \Phi(., y)$ in $\mathcal{D}(X)$ and so also $f_{\Phi(.,y_p)} \to f_{\Phi(.,y)}$ in $\mathcal{C}(Y)$ because of the continuity of A. Therefore we obtain

$$\begin{aligned} |F_{y_p}(\Phi(., y_p)) - F_y(\Phi(., y))| &= |f_{\Phi(.,y_p)}(y_p) - f_{\Phi(.,y)}(y)| \\ &\leq |f_{\Phi(.,y_p)}(y_p) - f_{\Phi(.,y_p)}(y)| + |f_{\Phi(.,y_p)}(y) - f_{\Phi(.,y)}(y)| \end{aligned}$$

which tends to 0 for $p \to \infty$.

Now we are in a position to postulate that the action of the distribution $K$ we are seeking can be given by the following formula:

$$\Phi \mapsto \int_Y F_y(\Phi(., y))dy \qquad \forall \Phi \in \mathcal{D}(X \times Y). \tag{3.3}$$

We have already established that the integral exists because the function under the integral is continuous with compact support. It remains to show that the integral in fact defines a distribution, i.e. that it describes a continuous linear functional on $\mathcal{D}(X \times Y)$, and that it fulfils the desired property (3.2). Linearity of $K$ is clear from the definition (3.3).

We show (sequential) continuity: Let $M \subset\subset X \times Y$ and $\Phi_p$ be a sequence in $\mathcal{D}(M)$ tending to 0. Then the support of each function $\Phi_p(., y)$ is contained in $L := \mathrm{pr}_1(M)$ and for $y \notin H := \mathrm{pr}_2(M)$ we have that $\Phi_p(., y) \equiv 0$. Hence

$$\int_Y F_y(\Phi_p(., y))dy = \int_H F_y(\Phi_p(., y))dy.$$

Note that when $y$ runs through the compact set $H$ the distributions $F_y$ form a bounded set in $\mathcal{D}'(X)$: If $B$ is a bounded subset of $\mathcal{D}(X)$, continuity of the map $A$ gives that $\{f_\varphi | \varphi \in B\}$ is a bounded set in the Fréchet space $\mathcal{C}(Y)$. Hence we can have an upper bound on the action of $(F_y)_{y\in H}$ on the test functions in the set $B$, i.e.

$$\exists\, C>0: \sup_{y\in H}\sup_{\varphi\in B}|\{F_y(\varphi)\}| = \sup_{y\in H}\sup_{\varphi\in B}|\{f_\varphi(y)\}| \leq C.$$

Therefore the family of distributions $(F_y)_{y\in H}$ is in particular pointwise bounded on $\mathcal{D}(X)$ and thus, according to the uniform boundedness principle [Theorem 2.1] for the Fréchet space $\mathcal{D}(M)$, equicontinuous. Therefore the integrand in the definition of $K$,

$$K:\Phi\mapsto\int_Y F_y(\Phi(.,y))dy,$$

which converges to 0 for any fixed $y$, does so even uniformly for $y\in H$ and so $K$ is indeed continuous.

Finally $K$ satisfies (3.2), since replacing $\Phi$ by $\varphi\otimes\psi$, $\varphi\in\mathcal{D}(X)$ and $\psi\in\mathcal{D}(Y)$, yields that

$$\begin{aligned}K(\varphi\otimes\psi) = \int_Y F_y((\varphi\otimes\psi)(.,y))dy = \int_Y F_y(\varphi)\psi(y)dy =\\ = \int_Y f_\varphi(y)\psi(y)dy = \langle A(\varphi),\psi\rangle\end{aligned}$$

(2) *The general case:* First we justify why we may assume without loss of generality that $X=\mathbb{R}^m$, $Y=\mathbb{R}^n$ and why it is enough to seek a distribution $K\in\mathcal{D}'(U\times V)$, that satisfies (3.2) for all $\varphi\in\mathcal{D}(U)$ and $\psi\in\mathcal{D}(V)$, where $U\subseteq X$ resp. $V\subseteq Y$ relatively compact and open:

Let $(U_\alpha)$ resp. $(V_\beta)$ be covers of $X$ resp. $Y$ consisting of relatively compact open sets. Due to the uniqueness of the distribution $K\in\mathcal{D}'(X\times Y)$ and the fact that localisations uniquely determine a distribution (cf. [8, Theorem 1.4.3.]), it suffices to prove the existence of suitable $K_{\alpha,\beta}\in\mathcal{D}'(U_\alpha\times V_\beta)$ that fulfil property (3.2) on their domains. The uniqueness also guarantees that any $K_{\alpha,\beta}, K_{\gamma,\delta}$ are compatible on overlapping domains, i. e.

$$K_{\alpha,\beta}=K_{\gamma,\delta} \text{ on } \mathcal{D}((U_\alpha\cap U_\gamma)\times(V_\beta\cap V_\delta)).$$

Again we may assume that the $(U_\alpha)$ resp. the $(V_\beta)$ lie within the domains of some charts of $X$ resp. $Y$ and therefore use Theorem 2.12 to reduce our problem to the simpler case where they are relatively compact open subsets of $X=\mathbb{R}^n$ and $Y=\mathbb{R}^m$.

A note on continuity: We know that the bilinear form $(\varphi,\psi)\mapsto\langle A(\varphi),\psi\rangle$ is separately continuous on the product of the (F)-spaces $\mathcal{D}(\overline{U})$ and $\mathcal{D}(\overline{V})$. Thus joint continuity follows from Lemma 2.3, a property we need in the subsequent construction.

Let $V_0$ be a relatively compact open neighbourhood of $\overline{V}$ in $Y$. Then there exists a neighbourhood $W$ of the origin in $\mathcal{D}(\overline{U})$ such that the set

$$N := \{A(\varphi)|_{\mathcal{D}(V_0)} : \varphi \in W\}$$

is bounded in $\mathcal{D}'(V_0)$:

Continuity of the map $(\varphi, \psi) \mapsto \langle A(\varphi), \psi\rangle$ on $\mathcal{D}(\overline{U}) \times \mathcal{D}(\overline{V_0})$ gives the existence of 0-neighbourhoods $W$ in $\mathcal{D}(\overline{U})$ and $W'$ in $\mathcal{D}(\overline{V_0})$ such that $|\langle A(\varphi), \psi\rangle| \leq 1$ for $\varphi \in W,\ \psi \in W'$. The 0-neighbourhood $W'$ is absorbing, this means that for every bounded subset $B$ of $\mathcal{D}(V_0)$ there exists an $\alpha > 0$ such that $B \subseteq \alpha \cdot W'$, hence $\alpha^{-1}B \subseteq W'$. Therefore (by pulling out the factor $\alpha^{-1}$ and bringing it to the other side)

$$|\langle A(\varphi), \psi\rangle| \leq \alpha \text{ for all } \psi \in B$$

and $N$ is indeed bounded in $\mathcal{D}'(V_0)$.

From this and Remark 2.2 it follows that there exists a $\tilde{q} \in \mathbb{N}$ such that the distributions in $N$ are at most of order $\tilde{q}$. We can get this to even hold for all $\varphi \in \mathcal{D}(U)$, because there always exists a real number $\beta > 0$ such that $\beta\varphi \in W$ (and multiplication with a scalar cannot change the order of a distribution). Also, the order remains unchanged if we choose a cut-off function $h$ around $\overline{V}$ whose support is contained in $V_0$ and multiply $A(\varphi)|_{\mathcal{D}(V_0)}$ by it. So all in all $h \cdot A(\varphi)|_{\mathcal{D}(V_0)} \in \mathcal{D}'(V_0)$ for $\varphi \in \mathcal{D}(U)$ is at most of order $q$, where we define $q$ to be the smallest even number greater or equal to $\tilde{q}$.

Now we use the structure theorem for distributions (cf. Friedlander-Joshi [8, Theorem 5.4.1.]) to rewrite $h \cdot A(\varphi)$ in the following way:

$$h \cdot A(\varphi) = (\partial_1 ... \partial_m)^{q+2}(E_{q+2} * (h \cdot A(\varphi))).$$

Here $E_{q+2} \in \mathcal{C}^q(V_0)$ is the standard fundamental solution of the partial differential operator $(\partial_1 ... \partial_m)^{q+2}$. For an explicit definition confer [8, Section 5.4.]. The benefit of this is that for all $\varphi \in \mathcal{D}(U)$ the convolution $E_{q+2} * (h \cdot A(\varphi))$ yields a continuous function on $V_0$. Furthermore, the map

$$\varphi \mapsto (E_{q+2} * (h \cdot A(\varphi)))$$

is continuous from $\mathcal{D}(U)$ to $\mathcal{D}'(V_0)$, as

$$\langle E_{q+2} * (h \cdot A(\varphi)), \psi\rangle = \langle A(\varphi), h \cdot (\check{E}_{q+2} * \psi)\rangle.$$

Therefore we can use Step 1) of our proof to find a distribution $R$ on $U \times V_0$ such that for

$\varphi \in \mathcal{D}(U)$ and $\psi \in \mathcal{D}(V_0)$ we have

$$\langle E_{q+2} * (h \cdot A(\varphi)), \psi \rangle = \langle R, \varphi \otimes \psi \rangle.$$

But now for all $\varphi \in \mathcal{D}(U)$, and for all $\psi \in \mathcal{D}(V)$ (that are extended to $V_0$ trivially)

$$\begin{aligned}
\langle R, \varphi \otimes ((\partial_1 ... \partial_m)^{q+2}\psi)\rangle &= \langle E_{q+2} * (h \cdot A(\varphi)), (\partial_1 ... \partial_m)^{q+2}\psi)\rangle \\
&= (-1)^{m\cdot(q+2)} \langle (\partial_1 ... \partial_m)^{q+2}(E_{q+2} * (h \cdot A(\varphi))), \psi \rangle \\
&= \langle h \cdot A(\varphi), \psi \rangle \\
&= \langle A(\varphi), h \cdot \psi \rangle = \langle A(\varphi), \psi \rangle
\end{aligned}$$

And so finally $\langle A(\varphi), \psi \rangle = \langle ((\partial_1 ... \partial_m)^{q+2})'' R, \varphi \otimes \psi \rangle$ for all $\varphi \in \mathcal{D}(U)$ and $\psi \in \mathcal{D}(V)$, where $((\partial_1 ... \partial_m)^{q+2})''$ denotes differentiation with respect to the last $m$ coordinates in $\mathbb{R}^{m+n}$, and the theorem is proved. □

**Remark 3.2.** (On the kernel theorem)

(i) There is also a so-called elementary version of the existence proof in (ii), that in particular does not recourse to the structure theorem for distributions, but uses convolution and regularisation instead.

Roughly the idea of the proof is to assume one knew the kernel distribution $K$ that generates the map $A : \mathcal{D}(X) \to \mathcal{D}'(Y)$, and could convolute it with a mollifier

$$\Psi_\epsilon = \psi_{1,\epsilon} \otimes \psi_{2,\epsilon} = \frac{1}{\epsilon^{n+m}} \cdot \psi_1(\frac{x}{\epsilon}) \otimes \psi_2(\frac{y}{\epsilon}),$$

(for $\psi_1 \in \mathcal{D}(X)$ and $\psi_2 \in \mathcal{D}(Y)$ suitably) such that $\Psi_\epsilon \to \delta_{(0,0)}$ for $\epsilon \to 0$.

This would give for $\epsilon$ small enough that

$$\begin{aligned}
(K * \Psi_\epsilon)(x', y') &= \langle K(x,y), \psi_\epsilon((x', y') - (x, y))\rangle \\
&= \frac{1}{\epsilon^{m+n}} \langle K(x,y), \psi_1(\frac{x'-x}{\epsilon}) \otimes \psi_2(\frac{y'-y}{\epsilon})\rangle \\
&\overset{(3.2)}{=} \frac{1}{\epsilon^{m+n}} \langle A(\psi_1(\frac{x'-x}{\epsilon})), \psi_2(\frac{y'-y}{\epsilon})\rangle =: K_\epsilon
\end{aligned}$$

If we knew $K$ existed, we would also know that $K_\epsilon \to K$ for $\epsilon \to 0$. On the other hand if one can show that $K_\epsilon$ converges to some $K_0 \in \mathcal{D}'(X \times Y)$ (which it does), and that this limit fulfils equation (3.2), that would yield the existence of $K$.

In the end things boil down to showing the equation

$$\langle K_\epsilon, \phi_1 \otimes \phi_2 \rangle = \langle A(\phi_1 * \check{\psi}_{1,\epsilon}), \phi_2 * \check{\psi}_{2,\epsilon} \rangle,$$

which basically says that the mollifiers in the definition of $K_\epsilon$ behave nicely and can be moved to the test functions in the regular way. Together with

$$\langle K_\epsilon, \phi_1 \otimes \phi_2 \rangle \xrightarrow{\epsilon \to 0} \langle K_0, \phi_1 \otimes \phi_2 \rangle,$$

and

$$\langle A(\phi_1 * \check{\psi}_{1,\epsilon}), \phi_2 * \check{\psi}_{2,\epsilon} \rangle \xrightarrow{\epsilon \to 0} \langle A(\phi_1), \phi_2 \rangle$$

this finishes the proof. For the details see Hörmander [11, Theorem 5.2.1.].

(ii) On a more abstract level the kernel theorem can be read as stating that

$$L(\mathcal{D}(X), \mathcal{D}'(Y)) \cong \mathcal{D}'(X \times Y).$$

This functional analytic approach to the kernel theorem can be found in Trèves [18, Ch. 51, Theorem 51.7] (who is working on open sets instead of manifolds). Trèves uses closures of projective tensor products of nuclear spaces to prove the theorem of L. Schwartz, and the major advantage of his approach is that this particular result is actually only one of many kernel statements his theory easily yields. Other results deduced within his approach would be for instance that

$$L(\mathcal{S}(\mathbb{R}^n), \mathcal{S}'(\mathbb{R}^m)) \cong \mathcal{S}'(\mathbb{R}^{n+m}), \text{ or}$$

$$L(\mathcal{C}^\infty(X), \mathcal{E}'(Y)) \cong \mathcal{E}'(X \times Y).$$

These would be analogous kernel theorems for tempered distributions or for distributions with compact support.

(iii) On a historical note the tensor product of spaces of test functions and distributions was defined by Laurent Schwartz in 1950 in [16] and the kernel theorem was announced only shortly afterwards in [17]. In both publications Schwartz used a decomposition of test functions in $X_1 \times X_2$ into sums of tensor products of test functions in $X_1$ and $X_2$, i.e. he emphasised the fact that $\mathcal{D}(X_1) \widehat{\otimes} \mathcal{D}(X_2) \cong \mathcal{D}(X_1 \times X_2)$, where $\widehat{\otimes}$ denotes the completion of the tensor product. Later versions of the proof of the kernel theorem then used a decomposition by Fourier series expansion or, as outlined in (i), regularisation and convolution.

In relation to the kernel theorem also the Peetre theorem is worth mentioning. It is an addendum that tells us that if $K$ is a continuous linear map

$$K : \mathcal{D}(X) \to \mathcal{C}^\infty(X)$$

with $\operatorname{supp}(Ku) \subseteq \operatorname{supp}(u)$, then $K$ is a differential operator with $C^\infty$-coefficients. Details on the Peetre theorem can be found either in the original publication (cf. [14]) or also in Dieudonné [4, section 17.13].

(iv) The isomorphism given by the kernel theorem tells us essentially that we may identify the operator from $\mathcal{D}(X)$ to $\mathcal{D}'(Y)$ and the respective kernel distribution in $\mathcal{D}'(X \times Y)$. Therefore it seems natural and convenient to denote them both by $K$, even though strictly speaking this is a slight abuse of notation. We will from now on follow these conventions, seeing that we have shown the one-to-one correspondence.

(v) To prove the kernel theorem we used Theorem 2.12 on local descriptions of distributions on manifolds in order to obtain w. l. o. g. that $X$ and $Y$ are open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$. This trick works in most of the cases, and therefore we will now often simply assume $X$ and $Y$ to be open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ for convenience.

The only instances where we really need to work on $\mathbb{R}^n$ is when we discuss Fourier transform or convolution, because these cannot be properly localised. For material explicitly working on manifolds see for instance [1].

**Example 3.3.** (i) If $X$ and $Y$ are open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, and the kernel $K \in \mathcal{D}'(X \times Y)$ is a locally integrable function, then we regain the notion of integral transform as in (3.1), i.e.

$$\begin{aligned} K : \mathcal{D}(X) &\to L^1_{\text{loc}}(Y) \subseteq \mathcal{D}'(Y) \\ f \mapsto K \cdot f(y) &= \int_X K(x,y) f(x) dx, \end{aligned}$$

and in particular the integral exists.

Whenever $K$ is a function on $X \times Y$ with the property that $K(x,.) : y \mapsto K(x,y)$ is locally integrable on $X$, i. e. when $K$ allows us to define $K \cdot f$ like we did above via the integral, $K$ is called an *integral operator given by the kernel function $K$*. Here it is sometimes useful to distinguish between the function and the kernel operator, and in this case we denote the function by $K_0$ and the respective distribution and kernel operator by $K$.

(ii) *(The identity mapping)*

Let $X = Y$ be an open subset of $\mathbb{R}^n$. Then kernel of the identity map $I : \mathcal{D}(X) \to \mathcal{D}(X) \subseteq \mathcal{D}'(X)$ is the distribution

$$\langle I, \Phi \rangle = \langle \delta(x-y), \Phi(x,y) \rangle = \int_X \Phi(x,x)dx.$$

This can be either seen from formula (3.2) that says

$$\langle I, \varphi \otimes \psi \rangle = \langle I \cdot \varphi, \psi \rangle = \langle \varphi, \psi \rangle = \int_X \varphi(x)\psi(x)dx,$$

or, since the identity does indeed map into a subspace of $\mathcal{C}(X)$, by following the steps of the proof.

For $A = I$ we get that $f_\varphi = \varphi$ and that $F_y(\varphi) = \varphi(y)$. Similarly for $\Phi \in \mathcal{D}(X \times X)$ we have that $F_y(\Phi(.,y)) = (\Phi(.,y))(y) = \Phi(y,y)$ and therefore by (3.3)

$$\langle I, \Phi \rangle = \int_X \Phi(x,x)dx = \langle \delta(x-y), \Phi(x,y) \rangle .$$

(iii) *(Partial differential operators)*

Let again $X = Y$ an open subset of $\mathbb{R}^n$ and let $P$ be a differential operator with $\mathcal{C}^\infty$ coefficients defined on $X$,

$$P(x,\partial) = \sum_{|\alpha| \leq m} a_\alpha(x)\partial^\alpha.$$

Then the image of $P$ when applied to test functions is a subspace of $\mathcal{C}(X)$ and hence we can use (3.3) to compute

$$\langle K, \Phi \rangle = \int P(x,\partial_x)\Phi(x,y)|_{x=y}dy, \quad \forall \Phi \in \mathcal{D}(X \times X).$$

Using the kernel $\delta(x-y) \in \mathcal{D}'(X \times X)$ of the identity map we can rewrite this as

$$\langle \delta(x-y), P(x,\partial_x)\Phi(x,y) \rangle = \langle P^t(x,\partial_x)\delta(x-y), \Phi(x,y) \rangle,$$

where $P^t$ is the adjoint of $P$. So $K(x,y) = P^t(x,\partial_x)\delta(x-y)$.

In particular, if $m = 0$ and hence $P = g \in \mathcal{C}^\infty$, then we have that the kernel of the multiplication map $\varphi \mapsto g\varphi$ is $g(y)\delta(x-y)$.

If $P = \partial_j$, $j = 1, \dots n$, then the kernel is given by $-(\partial/\partial_{x_j})\delta(x-y)$.

(iv) *(Convolution)*

We can also via kernel theory give an alternative representation of the convolution of two distributions, one of which has compact support. Let again $X = Y = \mathbb{R}^n$ and $u \in \mathcal{D}'(X)$. Then we compute the kernel $K_u$ of the map

$$\varphi \mapsto (u * \varphi)(y) = \langle u(x), \varphi(y-x)\rangle,$$

which is a continuous map $\mathcal{D}(\mathbb{R}^n) \to \mathcal{C}^\infty(\mathbb{R}^n)$, when $\mathcal{C}^\infty(\mathbb{R}^n)$ is equipped with the weak topology induced by $\mathcal{D}'(\mathbb{R}^n)$. According to (3.3) $K_u$ is given by

$$\begin{aligned} \langle K_u, \Phi\rangle \quad &= \textstyle\int \langle u(x), \Phi(y-x, y)\rangle dy \\ &= \textstyle\int \langle u(y-x), \Phi(x, y)\rangle dy, \end{aligned} \tag{3.4}$$

hence $K_u = u(y-x)$.

The examples above can be seen as special cases of convolution, since the identity operator can be represented as convolution with $\delta$, where (3.4) immediately gives $I = \delta(y-x)$. Also $\partial^\alpha \varphi = (\partial^\alpha \varphi) * \delta = \varphi * (\partial^\alpha \delta)$ and so the kernel is given by $(\partial^\alpha \delta)(y-x) = (-1)^{|\alpha|}\partial_x^\alpha$, as computed above.

**Remark 3.4.** *(The transposed map)* So far in this chapter we have been looking at the map $\varphi \mapsto K \cdot \varphi : \mathcal{D}(X) \to \mathcal{D}'(Y)$ we obtain through

$$\langle K, \varphi \otimes \psi\rangle = \langle K \cdot \varphi, \psi\rangle$$

for $K \in \mathcal{D}'(X \times Y)$, $\varphi \in \mathcal{D}(X)$ and $\psi \in \mathcal{D}(Y)$. However we can also let $K$ act on $\psi$ first, i.e. look at the map

$$\begin{gathered} K^t : \mathcal{D}(Y) \to \mathcal{D}'(X) \\ \psi \mapsto \langle K, \varphi \otimes \psi\rangle = \left\langle K^t \cdot \psi, \varphi \right\rangle, \end{gathered}$$

which we call the *transposed map of* $K$. The maps $K$ and $K^t$ are adjoints of each other, in the following sense:

$$\langle K \cdot \varphi, \psi\rangle = \langle K, \varphi \otimes \psi\rangle = \langle K^t \cdot \psi, \varphi\rangle. \tag{3.5}$$

Naturally the kernel theorem also holds for the transposed map and we will see in the next section that the interplay between these two maps allows us to extend certain kernel transforms to larger domains.

**Example 3.5.** (i) If we return to convolution, formula (3.4) allows us to compute the transposed kernel $K_u^t$ of the map $K_u$.

$$\begin{aligned}\langle K_u^t \cdot \psi, \varphi\rangle &= \int \langle u(x), \varphi(y-x)\psi(y)\rangle dy \\ &= \int \langle u(x), \varphi(y)\psi(x+y)\rangle dy \\ &= \int \langle u(x), \psi(x+y)\rangle \varphi(y) dy,\end{aligned}$$

hence $(K_u^t \cdot \psi)(y) = \langle u(x), \psi(x+y)\rangle = (\check{u} * \psi)(y)$, and $K_u^t : \mathcal{D}(\mathbb{R}^n) \to \mathcal{C}^\infty(\mathbb{R}^n)$.

(ii) Let $X$ and $Y$ be manifolds and $\mu$ and $\nu$ be one-densities on $X$ resp. $Y$. Suppose that the kernel distribution $K$ is given by a continuous function $K_0$ on $X \times Y$ together with the volume element $\mu \otimes \nu$, i.e. a kernel whose action is given by $\langle K, \varphi \otimes \psi\rangle = \int_{X\times Y} K_0(x,y)\varphi(x)\psi(y)\mu(x) \otimes \nu(y)$. Then we obtain the kernel map $K : \mathcal{D}(X) \to \mathcal{D}'(Y)$

$$(K \cdot \varphi)(y) = \int_X K_0(x,y)\varphi(x) d\mu(x) \otimes \nu$$

for $\varphi \in \mathcal{D}(X)$ and almost all $y \in Y$, and the transposed kernel $K^t$ is given by

$$(K^t \cdot \psi)(x) = \int_Y K_0(x,y)\psi(y) d\nu(y) \otimes \mu.$$

We shall call distributions on $X \times Y$ whose volume element $\sigma \in \mathrm{Vol}(X \times Y)$ can be split up into densities $\sigma = \mu \otimes \nu$ for $\mu \in \mathrm{Vol}(X)$ and $\nu \in \mathrm{Vol}(Y)$ *integral operators*, they are the analogue on manifolds to the the operators defined in Example 3.3 on open subsets of $\mathbb{R}^n$.

To close this paragraph we now turn to support properties of kernel distributions. We will give a criterion for finding out whether a distribution $K \cdot \varphi$ is compactly supported or not. To this end we first introduce some notation that will allow us to conveniently work with subsets of $X$, $Y$ and $X \times Y$.

**Notation 3.6.** Let $R$ be a subset of $X \times Y$ and $A$ (resp. $B$) be a subset of $X$ (resp. $Y$). The we denote the set $\mathrm{pr}_2(R \cap (A \times Y))$, i.e. the set of all $y \in Y$ where there exists an

$x \in A$ such that $(x, y) \in R$, by $R(A)$. Analogously we define $R^{-1}(B) := \mathrm{pr}_1(R \cap (X \times B))$, the set of all $x \in X$ where there exists a $y \in B$ with $(x, y) \in R$.

We immediately have that if $A$ is a compact subset of $X$, $B$ is a compact subset of $Y$ and $R \subseteq X \times Y$ closed, then both the sets $R(A)$ and $R^{-1}(B)$ are closed in $Y$ resp. $X$. We show the statement for $R(A)$: Let $(y_n)$ be a sequence in $R(A)$ and let $y_n \to y$. Then we know from the definition of $R(A)$ that there is a sequence $(x_n)$ in $A$ such that $(x_n, y_n) \in R$ for all $n \in \mathbb{N}$. Since $A$ is compact we can have that $(x_n)$ converges to some $x \in A$, by choosing a subsequence if necessary. Then we obtain that $(x_n, y_n) \in R$ and moreover $(x_n, y_n) \to (x, y)$, which consequently has to lie in the closed subset $R$ too. Thus we can conclude that $y$ is indeed in $R(A)$. The statement for $R^{-1}(B)$ is shown analogously.

**Proposition 3.7.** *Let $K$ be a distribution on $X \times Y$ and $R = \mathrm{supp}(K)$. Then we have that*

$$\mathrm{supp}(K \cdot \varphi) \subseteq R(\mathrm{supp}(\varphi))$$

*for all $\varphi \in \mathcal{D}(X)$, resp. $\mathrm{supp}(K^t \cdot \psi) \subseteq R^{-1}(\mathrm{supp}(\psi))$ for all $\psi \in \mathcal{D}(Y)$.*

*Proof.* Let $A = \mathrm{supp}(\varphi)$, then we have to show that we have $\langle K \cdot \varphi, \psi \rangle = \langle K, \varphi \otimes \psi \rangle = 0$ for every test function $\psi \in \mathcal{D}(Y)$ whose support $B$ satisfies $B \cap R(A) = \emptyset$. But this follows immediately, since the support of $\varphi \otimes \psi$ is a subset of $A \times B$, which is then disjoint to the closed set $R$. The second statement is shown analogously. ☐

**Remark 3.8.** (i) If we have a closed subset $F$ of $Y$ which contains the support of $K \cdot \varphi$ for all $\varphi \in \mathcal{D}(X)$, we can even gain information on the support of the kernel distribution $K$. More precisely we have $\mathrm{supp}(K) \subseteq \mathrm{pr}_2^{-1}(F)$: Suppose to the contrary that there existed a $y \notin F$ and an $x \in X$ with $(x, y) \in \mathrm{supp}(K)$, then we could choose a neighbourhood $V$ of $y$ with $V \cap F = \emptyset$ and two test functions $\psi \in \mathcal{D}(V)$ and $\varphi \in \mathcal{D}(X)$ such that $\langle K, \varphi \otimes \psi \rangle \neq 0$. Therefore $y \in \mathrm{supp}(K \cdot \varphi) \subseteq F$, which is a contradiction to the choice of $y$.

(ii) Proposition 3.7 implies that if $R(L)$ is compact for every $L \subset\subset X$, the support of $K \cdot \varphi$ is also compact for all $\varphi \in \mathcal{D}(X)$, and therefore $K : \mathcal{D}(X) \to \mathcal{E}'(Y)$.

If we now consider the map

$$\langle K^t \cdot \psi, \varphi \rangle = \langle K, \varphi \otimes \psi \rangle = \langle \underbrace{K \cdot \varphi}_{\in \mathcal{E}'(Y)}, \psi \rangle,$$

we see that we can actually allow $\psi$ to be in $\mathcal{C}^\infty$ with non-compact support, since the right-hand side of the equation is defined for all $\psi \in \mathcal{C}^\infty$. So using a regularity property of the map $K$, we can extend the transpose $K^t$ to $\mathcal{E}(X) \to \mathcal{D}'(Y)$.

In the next chapter on regular kernels we look at a class of kernel distributions that allows us to extend the maps $K$ and $K^t$ even to distribution spaces. But before doing so we prove that the condition used above is not only sufficient but also necessary.

**Proposition 3.9.** *Let $K$ be a distribution on $X \times Y$ and $R = \operatorname{supp}(K)$. The distribution $K \cdot \varphi$ has compact support for all $\varphi \in \mathcal{D}(X)$ if and only if $R(L)$ is compact in $Y$ for all $L \subset\subset X$.*

*Proof.* That the condition is sufficient we saw in Remark 3.9, to see that it is necessary we argue by contradiction: We suppose that $\exists L \subset\subset Y$ such that $R(L)$ is not compact.

First choose a sequence of relatively compact open sets $V_k \subset Y$ with $\bar{V}_k \subset V_{k+1}$ and $Y = \bigcup V_k$ and a set $L \subset\subset X$ such that $\mathring{R}(L)$ is not relatively compact. Additionally choose a relatively compact, open neighbourhood $W$ of $L$ in $X$. Then we inductively construct a sequence of functions $(f_k)$ in $\mathcal{D}(\overline{W})$ and a sequence $(g_k)$ in $\mathcal{D}(Y)$ with the following properties:

(a) $\operatorname{supp}(g_k) \cap H_k = \emptyset$ for $H_k := \bigcup_{j<k} \operatorname{supp}(K \cdot f_j) \cup \overline{V}_k$ (which are all compact by assumption),

(b) $\langle K \cdot f_k, g_k \rangle \neq 0$.

This is possible since we know by assumption that there is a $y_k \in R(L)$ with $y_k \notin H_k$ and consequently an $x_k$ with $(x_k, y_k) \in R$. Then it follows that there is a function $f_k \in \mathcal{D}(\overline{W})$ and a function $g_k \in \mathcal{D}(Y)$ with support disjoint to $H_k$ such that

$$\langle K \cdot f_k, g_k \rangle = \langle K, f_k \otimes g_k \rangle \neq 0.$$

Subsequently we may define a sequence of positive real numbers $c_k$ so that

(a) the series $f = \sum_k c_k f_k$ converges in $\mathcal{D}(\overline{W})$,

(b) $c_k |\langle K \cdot f_k, g_j \rangle| \leq 4^{-k} c_j |\langle K \cdot f_j, g_j \rangle|$ for all $k > j$.

Such a sequence of $c_k$ exists since we can cover the compact set $\overline{W}$ by finitely many chart domains of $X$ and thus assume w.l.o.g. that $X = \mathbb{R}^n$. Then one can obtain (a) by choosing the $c_k$ small enough to guarantee $c_k |D^\alpha f_k(x)| \leq 2^{-k}$ for $x \in \overline{W}$ and $|\alpha| \leq k$. Proceed inductively to satisfy condition (b) by multiplying $c_k$ by a sufficiently small factor to ensure that the inequality holds for all $j \leq k$ (note that the right-hand-side is non-zero).

By continuity, $K \cdot f = \sum_k c_k(K \cdot f_k)$ with respect to the weak topology on $\mathcal{D}'(Y)$ and thus we have

$$\begin{aligned} |\langle K \cdot f, g_j \rangle| &= |c_j \langle K \cdot f_j, g_j \rangle + \sum_{k>j} c_k \langle (K \cdot f_k), g_j \rangle| \\ &\geq |c_j \langle K \cdot f_j, g_j \rangle| + |\sum_{k>j} c_k \langle (K \cdot f_k), g_j \rangle| \\ &\geq \frac{1}{2} c_j |\langle K \cdot f_j, g_j \rangle| > 0 \end{aligned}$$

for all $j$, which already contradicts the assumption that $K \cdot f$ is compactly supported. To see this remember that the $g_j$ were constructed in a way that their supports lie outside $\bigcup_{k<j} \overline{V_k}$, where the $V_k$ are a growing sequence of compact sets that cover $Y$. The above calculation yields that $\operatorname{supp}(K \cdot f)$ intersects the support of every $g_j$, hence it cannot be compact. □

**Remark 3.10.** In addition, the following statement holds:

Let $K$ be a distribution on $X \times Y$ and $R = \operatorname{supp}(K)$. Then the operator $K$ can be extended to a linear and continuous operator $\mathcal{E}(X) \to \mathcal{D}'(Y)$ if and only if $R^{-1}(H)$ is compact in $X$ for all $H \subset\subset Y$.

To see that the condition is sufficient is fairly straight-forward: Let $K$ be a distribution such that $R^{-1}(H)$ is compact in $X$ for all $H \subset\subset Y$. Then we have that for every test function $\psi \in \mathcal{D}(Y)$ the intersection $R \cap (X \times \operatorname{supp}(\psi))$ is compact, and thus $\langle K, f \otimes \psi \rangle$ is defined for all $f \in \mathcal{E}(X)$. Since we can choose a cut-off function $h$ over the set $R^{-1}(\operatorname{supp}(\psi))$ we obtain the the extended map $K : f \mapsto \langle K, hf \otimes \psi \rangle = \langle K, f \otimes \psi \rangle$ is continuous on $\mathcal{E}(X)$.

To see that the condition is also necessary requires some more work, for details cf. [5, 23.9.8]

## 3.2 Regular Kernels

In this paragraph let $X$ and $Y$ be open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ for reasons of simplicity.

To begin with, let us suppose for instance that $K^t \cdot \psi$ can be identified with a function in $\mathcal{C}^{(r)}(X)$ for all $\psi \in \mathcal{D}(Y)$. Then $K^t \cdot \psi$ can act not only on test functions $\varphi \in \mathcal{D}(X)$, but even on any distribution $u \in \mathcal{C}'^{(r)}(X)$, where $\mathcal{C}'^{(r)}(X)$ denotes the dual space of $\mathcal{C}^{(r)}(X)$ which is the space of all compactly supported distributions of order $m \leq r$. This is since we have that $\langle K^t \cdot \psi, u \rangle$ is defined for all $\psi \in \mathcal{D}(Y)$ and $u \in \mathcal{C}'^{(r)}(X)$. The map

$K^t : \mathcal{D}(Y) \to \mathcal{C}^{(r)}(X)$ is continuous with respect to the weak topology on $\mathcal{C}^{(r)}(X)$, but then by Proposition 2.4 also with respect to the stronger Fréchet space topology.

Therefore we can extend the operator $K$ to a map $\mathcal{C}'^{(r)}(X) \to \mathcal{D}'(Y)$ by setting

$$\langle K \cdot u, \psi \rangle := \langle K^t \cdot \psi, u \rangle \quad \forall u \in \mathcal{C}'^{(r)}, \forall \psi \in \mathcal{D}(X). \tag{3.6}$$

This definition also shows that the linear map $K : \mathcal{C}'^{(r)}(X) \to \mathcal{D}'(Y)$ is continuous with respect to the weak topologies on $\mathcal{C}'^{(r)}(X)$ and $\mathcal{D}'(Y)$. Moreover we obtain the following result, which extends Proposition 3.7 to the larger space $\mathcal{C}'^{(r)}(X)$ :

**Proposition 3.11.** *Let $K$ be as described above and set $R = \operatorname{supp}(K)$. Then we have for every distribution $u \in \mathcal{C}'^{(r)}(X)$ that*

$$\operatorname{supp}(K \cdot u) \subseteq R(\operatorname{supp}(u)).$$

*Additionally any compact subset $H$ of $Y$ that satisfies $\operatorname{supp}(u) \cap R^{-1}(H) = \emptyset$ is disjoint to $\operatorname{supp}(K \cdot u)$.*

*Proof.* Let $\psi \in \mathcal{D}(Y)$ be a test function whose support $B$ is disjoint to $R(\operatorname{supp}(u))$. Then $\operatorname{supp}(u) \times B$ is disjoint to $R$ and hence we can choose an open neighbourhood $U$ of $\operatorname{supp}(u)$ in $X$ such that even $(U \times B) \cap R = \emptyset$.

Now we regularise the distribution $u$ by convoluting it with a sequence $(\varphi_k) \in \mathcal{D}(\mathbb{R}^n)$ that gives $(u * \varphi_k) \in \mathcal{D}(\mathbb{R}^n)$ and $(u * \varphi_k) \to u$ in the weak topology on $\mathcal{C}'^{(r)}(X)$. This gives that $\langle K \cdot (u * \varphi_k), \psi \rangle \to \langle K \cdot u, \psi \rangle$ and together with the fact that $\operatorname{supp}(u * \varphi_k) \subseteq U$ for $k$ large enough we obtain that $\langle K \cdot (u * \varphi_k), \psi \rangle = 0$ and hence also $\langle K \cdot u, \psi \rangle = 0$. This proves the first assertion.

The second assertion now follows like this: $\operatorname{supp}(u) \cap R^{-1}(H) = \emptyset \Leftrightarrow (\operatorname{supp}(u) \times H) \cap R = \emptyset \Leftrightarrow R(\operatorname{supp}(u)) \cap H = \emptyset$. Therefore, according to the first statement, $\operatorname{supp}(K \cdot u) \cap H = \emptyset$. □

**Example 3.12.** Let $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$. We consider an integral operator given by the kernel function $K_0 \in \mathcal{C}^{(r)}(X \times Y)$. Then we have that $(K^t \cdot \psi) \in \mathcal{C}^{(r)}(X)$ for all $\psi \in \mathcal{D}(Y)$ since we have that

$$\partial_x^\alpha (K^t \cdot \psi)(x) = \int_Y (\partial_x^\alpha K_0(x, y)) \psi(y) dy \quad \forall |\alpha| \leq r.$$

Therefore $K \cdot u$ is defined for all $u \in \mathcal{C}'^{(r)}$, and in addition $K \cdot u$ is in $\mathcal{C}^{(r)}(Y)$ given by the formula

$$(K \cdot u)(y) = \langle u, K_0(., y) \rangle : \tag{3.7}$$

First of all $u$ is compactly supported, and so the right side of the equation is defined. Furthermore the right-hand side is in $\mathcal{C}^{(r)}(Y)$ since

$$D^\alpha \langle u, K_0(.,y)\rangle = \langle u, D^\alpha_{(y)} K_0(.,y)\rangle, \tag{3.8}$$

as can be easily verified (or seen in [8, Theorem 4.11.]). The equality in (3.7) is gained by the following calculation: $\forall \psi \in \mathcal{D}(Y)$ :

$$\langle K \cdot u, \psi\rangle := \langle u, K^t \cdot \psi\rangle = \langle K_0, u \otimes \psi\rangle = \langle \psi(y), \langle u(x), K_0(x,y)\rangle\rangle = \langle\langle u, K_0(.,y)\rangle, \psi(y)\rangle.$$

We now define regular kernels, a class of kernel distributions that yields even "nicer" spaces:

**Definition 3.13.** $K \in \mathcal{D}'(X \times Y)$ is called a *regular kernel distribution*, if the following conditions hold:

(i) For every test function $\varphi \in \mathcal{D}(X)$ we have $K \cdot \varphi \in \mathcal{E}(Y)$

(ii) For every test function $\psi \in \mathcal{D}(Y)$ we have $K^t \cdot \psi \in \mathcal{E}(X)$

In this case both $K \cdot u$ for $u \in \mathcal{E}'(X)$ and $K^t \cdot v$ for $v \in \mathcal{E}'(Y)$ define a distribution on $Y$ resp. on $X$ via

$$\langle K \cdot u, \psi\rangle := \langle u, K^t \cdot \psi\rangle \quad \forall \psi \in \mathcal{D}(Y)$$

and

$$\langle K^t \cdot v, \varphi\rangle := \langle v, K \cdot \varphi\rangle \quad \forall \varphi \in \mathcal{D}(X).$$

Furthermore the maps $K : \mathcal{E}'(X) \to \mathcal{D}'(Y)$ and $K^t : \mathcal{E}'(Y) \to \mathcal{D}'(X)$ are continuous with respect to the associated weak topologies. If $K$ is a regular kernel distribution, we also call the corresponding operators $K$ and $K^t$ *regular*.
In particular $K \cdot \delta_x$ is a distribution for every $x \in X$, defined by

$$\langle K \cdot \delta_x, \psi\rangle = \langle \delta_x, K^t \cdot \psi\rangle = (K^t \cdot \psi)(x)$$

for all $\psi \in \mathcal{D}(Y)$.

**Remark 3.14.** Treves [18, 51.8] gives a characterisation of such regular kernel distributions. He calls a kernel distribution *regular in $x$* if (ii) from above holds, and shows for these that

$$L(\mathcal{D}(Y), \mathcal{C}^\infty(X)) \cong \mathcal{C}^\infty(X, \mathcal{D}'(Y)).$$

So regular kernels in $x$ can be identified with the space of $\mathcal{C}^\infty$-functions of $x$ with values in the distribution space $\mathcal{D}'(Y)$. Since the same is true for the dual case, we have that the space of regular kernel distributions is isomorphic to the space

$$\mathcal{C}^\infty(X, \mathcal{D}'(Y)) \cap \mathcal{C}^\infty(Y, \mathcal{D}'(X)).$$

**Example 3.15.** (i) Once again we turn to convolution. In the previous section in Examples 3.3 and 3.5 we showed, that the kernel distribution $\langle K_u, \Phi\rangle = \int \langle u(x), \Phi(y - x, y)\rangle dy$ for $u \in \mathcal{D}'(\mathbb{R}^n)$ gives rise to the two operators

$$K_u : \mathcal{D}(\mathbb{R}^n) \to \mathcal{C}^\infty(\mathbb{R}^n),$$
$$(K_u \cdot \varphi)(y) = (u * \varphi)(y)$$

and

$$K_u^t : \mathcal{D}(\mathbb{R}^n) \to \mathcal{C}^\infty(\mathbb{R}^n)$$
$$(K_u \cdot \psi)(x) = (\check{u} * \psi)(x).$$

Therefore the kernel is regular and we can extend the operator $K_u$ to $\mathcal{E}'(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^n)$, thus regaining the full convolution map $* : \mathcal{D}'(\mathbb{R}^n) \times \mathcal{E}'(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^n)$.

(ii) Choose $K = u \otimes v$ for $u \in \mathcal{D}'(X)$ and $v \in \mathcal{D}'(Y)$. Then $K \cdot \varphi = \langle u, \varphi\rangle v$ and $K^t \cdot \psi = \langle v, \psi\rangle u$, for $\varphi \in \mathcal{D}(X)$ and $\psi \in \mathcal{D}(Y)$ as usual. Clearly $K$ is not regular for arbitrary distributions $u, v$.

If $u$ and $v$ are locally integrable functions $f$ on $X$ and $g$ on $Y$, we obtain that $K$ is an integral operator with kernel function $K_0(x, y) = f(x) \cdot g(y)$. Is this the case, $K$ is regular if and only if $f$ and $g$ are $\mathcal{C}^\infty$. Integral operators with $\mathcal{C}^\infty$-kernels $K_0$ are called *regularising*. They map $\mathcal{E}'(X) \to \mathcal{E}(Y)$ as can be seen from Example 3.12 and are unique with this property, as we will see in the next section.

A small warning: While the concepts of regular *kernel distributions* and regular distributions as defined in Section 2.3. coincide for integral operators, this is generally not the case. By way of illustration we have seen that the identity mapping $\mathrm{id}_{\mathcal{D}(X)}$, which is clearly regular, is generated by the kernel distribution $\delta \in \mathcal{D}(X \times X)$.

**Proposition 3.16.** *Let $K$ be a regular kernel distribution on $X \times Y$ and $R = \operatorname{supp}(K)$. Then the following are equivalent:*

*(i) $R^{-1}(H)$ is compact in $X$ for all $H \subset\subset Y$.*

*(ii) The operator* $K^t$ *maps* $\mathcal{D}(Y)$ *into* $\mathcal{D}(X)$.

*(iii) The operator* $K^t$ *maps* $\mathcal{E}'(Y)$ *into* $\mathcal{E}'(X)$.

*(iv) The operator* $K$ *can be extended to a continuous and linear map* $\mathcal{E}(X) \to \mathcal{E}(Y)$ *with respect to the Fréchet space topologies.*

*If any of the above conditions hold, we obtain that the operator* $K$ *can be extended to a linear map* $\mathcal{D}'(X) \to \mathcal{D}'(Y)$ *whose restriction to any bounded subset of* $\mathcal{D}'(X)$ *is continuous wit respect to the weak topologies.*

*Proof.* $(i) \Leftrightarrow (ii)$ is clear from the definition of regular kernels and Prop. 3.9.

$(i) \Leftrightarrow (iii)$: Since $K$ is regular we can have $K^t : \mathcal{E}'(Y) \to \mathcal{D}'(X)$ by $\langle K^t \cdot u, \varphi\rangle = \langle K \cdot \varphi, u\rangle$. The support of $K^t \cdot u$ is compact by Proposition 3.11 if and only if $(i)$ holds.

$(i) \Leftrightarrow (iv)$ Follows from Remark 3.10

Now we show the last statement: By $(i), \operatorname{supp}(K^t \cdot \psi)$ is compact for all $\psi \in \mathcal{D}(Y)$. Regularity of the kernel gives that $K^t \cdot \psi \in \mathcal{D}(X)$ and hence the operator $K$ can act on any distribution $u \in \mathcal{D}'(X)$ via $\langle K \cdot u, \psi\rangle := \langle u, K^t \cdot \psi\rangle$. Continuity of the map $u \mapsto K \cdot u$ is obvious from the definition. □

**Example 3.17.** (i) *(Partial differential operators)*

Let $X = Y$ and $P$ be a linear partial differential operator on $X$. Then $P$ is clearly a regular operator and the support of the kernel distribution is a subset of the diagonal of $X \times X$ (according to Example 3.3). Therefore we can extend the operator $P$ to any of the spaces as given in Proposition 3.16, in particular to $P : \mathcal{D}'(X) \to \mathcal{D}'(X)$, $u \mapsto P \cdot u$.

In this particular example we can go even further and allow $P \cdot u$ to act on non-compactly supported test functions $\psi \in \mathcal{C}^\infty(X)$, provided that the intersection of $\operatorname{supp}(\psi)$ and $\operatorname{supp}(u)$ is compact. Even though

$$\langle P \cdot u, \psi\rangle = \langle u, P^t \cdot \psi\rangle \tag{3.9}$$

in general need not hold, we can by-pass all difficulties by choosing a cut-off function $h \in \mathcal{D}(X)$ with $h \equiv 1$ on a neighbourhood of the intersection of the supports which we introduce as a factor on the right-hand side: To be more precise since $\operatorname{supp}(P \cdot u) \subseteq \operatorname{supp}(P)$, we have that

$$\langle P \cdot u, \psi\rangle = \langle P \cdot u, h\psi\rangle = \langle u, P^t \cdot (h\psi)\rangle = \langle u, P^t \cdot \psi\rangle,$$

where we use the Leibniz rule in the last equality.

If the intersection is not compact, the two sides of (3.9) may be defined and still disagree with each other: For instance let $X = (0,1)$, $u$ be a regular distribution, $P = \frac{d}{dx}$ and thus $P^t = -\frac{d}{dx}$. Then

$$\begin{aligned} \langle P \cdot u, \psi \rangle &= \langle u', \psi \rangle = \int_0^1 u'(x)\psi(x)dx; \\ \langle u, P^t \cdot \psi \rangle &= -\langle u, \psi' \rangle = -\int_0^1 u(x)\psi'(x)dx, \end{aligned}$$

which agrees only if the boundary terms from integration by parts vanish. For $u = 1$ and $\psi(x) = x$ for instance, the top equation yields 0 whereas the bottom one gives $-1$.

(ii) Let $K$ be a regular distribution on $X \times Y$. Then we define the linear map

$$K^{t*} : \psi \mapsto \overline{K^t \cdot \overline{\psi}},$$

and we call $K^{t*} : \mathcal{D}(Y) \to \mathcal{C}^\infty(X)$ the adjoint operator of the operator $K$. We get

$$\langle K^{t*} \cdot \psi, \varphi \rangle = \overline{\langle K^t \cdot \overline{\psi}, \overline{\varphi} \rangle} = \overline{\langle K, \overline{\varphi} \otimes \overline{\psi} \rangle} = \langle \overline{K}, \varphi \otimes \psi \rangle,$$

for $\varphi \in \mathcal{D}(X)$, $\psi \in \mathcal{D}(Y)$ and $\overline{K}$ the complex conjugate distribution of $K$. By this calculation $K^{t*}$ is the operator defined by the kernel distribution $\overline{K}$.

Let us now consider the special case that $X = Y$: For $f \in \mathcal{C}^\infty(X)$ an $g \in \mathcal{D}(X)$ (or $f \in \mathcal{D}(X)$ and $g \in \mathcal{C}^\infty(X)$ we get that the function $f\overline{g}$ is in $\mathcal{D}(X)$ and is in particular integrable. We may hence define

$$(f, g) := \int_X f(x)\overline{g(x)}dx,$$

which coincides with the regular $L^2$ scalar product if the functions $f$ and $g$ are in $L^2(X)$. Since we obtain $\mathcal{C}^\infty$ functions for $K \cdot \varphi$ and $K^t \cdot \psi$ we can rewrite the action of these distributions with the help of $(.,.)$ as $\langle K \cdot \varphi, \overline{\psi} \rangle = (K \cdot \varphi, \psi)$ and $\langle \varphi, K^t \cdot \overline{\psi} \rangle = (\varphi, K^t \cdot \overline{\psi}) = (\varphi, K^{t*} \cdot \psi)$. Therefore we can all in all rewrite condition (3.5) as

$$(K \cdot \varphi, \psi) = (\varphi, K^{t*} \cdot \psi).$$

## 3.3 Regularising Operators

Let again $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$ unless stated otherwise.

In this section we are going to have a close look at regularising integral operators. Remember that for $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$, integral operators were defined as given by kernel functions $K_0 \in L^1_{loc}(X \times Y)$, hence $y \mapsto K_0(x,y)$ is locally integrable on $Y$ for almost any $x$. (For manifolds integral operators are distributions in $\mathcal{D}'(X \times Y)$ given by a function $K_0$ and a volume element $\sigma$ that can be split up into $\sigma = \mu \otimes \nu$ for $\mu \in \mathrm{Vol}(X)$ and $\nu \in \mathrm{Vol}(Y)$.) We call the integral operator *regularising* if additionally $K_0$ is in $\mathcal{C}^\infty(X \times Y)$.

We return to Example 3.12 where we had that for a regularising operator $K$, $K \cdot u$ is defined for all $u \in \mathcal{E}'$ and given by the formula

$$(K \cdot u)(y) = \langle u, K_0(.,y) \rangle.$$

Furthermore $K \cdot u$ is in $\mathcal{E}(Y)$ since

$$D^\alpha \langle u, K_0(.,y) \rangle = \langle u, D^\alpha_{(y)} K_0(.,y) \rangle.$$

Hence regularising operators are indeed regular, and they have an even stronger property: Not only do they map $\mathcal{D}(X)$ into $\mathcal{E}(Y)$, they even map $\mathcal{E}'(X)$ into $\mathcal{E}(Y)$. Regularising integral operators are unique with this property, as we show in the subsequent theorem:

**Theorem 3.18.** *We equip the space $\mathcal{E}'(X)$ of compactly supported distributions on $Y$ with the weak topology by viewing it as the space of linear and continuous functionals on $\mathcal{E}(X)$.*

*(i) Let $K_0$ be a $\mathcal{C}^\infty$ function on $X \times Y$. Then $K \cdot u$ is a function for every $u \in \mathcal{E}'(X)$ given by the formula*

$$(K \cdot u)(y) = \langle u, K_0(.,y) \rangle. \tag{3.10}$$

*Furthermore the operator $K$ maps every bounded subset of $\mathcal{E}'(X)$ continuously into the Fréchet space $\mathcal{E}(Y)$.*

*(ii) On the other hand let $K$ be a linear map $\mathcal{E}'(X) \to \mathcal{E}(Y)$ that is continuous on any bounded subset of $\mathcal{E}'(X)$, then $K$ is the extension of a regularising integral operator to the space $\mathcal{E}'(X)$.*

*Proof.* (i) The first claim is clear from what we noted above.

We prove the second statement: We have to show continuity on bounded subsets of $\mathcal{E}'(X)$, and we initially note that such a set is metrizable with respect to the weak topology:

Boundedness of the set $B$ means that

$$\exists C > 0 \; \forall b \in B \; \forall f \in \mathcal{C}^\infty(X) : \langle b, f \rangle \leq C.$$

This in particular implies equicontinuity of the set of linear functions in $B$, and allows us to use Theorem 2.7 to obtain that $B$ is metrizable.

Due to this it suffices to show sequential continuity of the operator $K$: therefore let $(u_k)$ be a zero-sequence of distributions in $\mathcal{E}'(X)$. We have to show that the sequence $K \cdot u_k$ converges to 0 in $\mathcal{E}(Y)$, i.e.

$$\forall H \subset\subset Y \; \forall \nu \in \mathbb{N}^m : D^\nu(K \cdot u_k)(y) = \langle u_k, D^\nu(K_0(.,y)) \rangle \overset{(k\to\infty)}{\to} 0,$$

uniformly w.r.t. $y \in H$. The function $y \mapsto D^\nu K_0(.,y)$, $H \to \mathcal{E}(X)$ is continuous, and so the set of functions $(D^\nu K_0(.,y))_{y \in H}$ is compact in $\mathcal{E}(X)$. Furthermore the set of linear forms $u_k$ on $\mathcal{E}(X)$ is equicontinuous according to the Banach-Steinhaus theorem. This allows us to use Prop. 2.5 to obtain that the sequence of functions $y \mapsto \langle u_k, D^\nu K_0(.,y) \rangle$ converges to 0 on $H$ not only pointwise, but even uniformly for $k \to \infty$.

(ii) To prove this statement we need to find a way to define the kernel function $K_0$ of the integral operator we are looking for, and we achieve this with the help of the delta-distribution: By assumption $K \cdot \delta_x$ is a smooth function on $Y$ for all $x \in X$. Therefore we may set $K_0(x,y) := (K \cdot \delta_x)(y)$ for all $y \in Y$, $x \in X$ and we now prove that this is indeed the function we are looking for:

As a first step we show that $K_0$ is smooth: For $x' \to x$, where $x'$ stays in a compact neighbourhood $L(x)$ of $x$, we have that $\delta_{x'}$ converges to $\delta_x$ with respect to the weak topology on $\mathcal{E}'(X)$ and hence also $K \cdot \delta_{x'} \to K \cdot \delta_x$ in $\mathcal{E}(Y)$. Moreover the set of all $(\delta_{x'})_{x' \in L(x)}$ is bounded and consequently $(K \cdot \delta_{x'})_{x' \in L(x)}$ is bounded in $\mathcal{E}(Y)$. So we obtain for all $x, x' \in L(x)$, $y, y' \in H \subset\subset X$ that

$$\begin{aligned}
|K_0(x',y') - K_0(x,y)| &\leq |K_0(x',y') - K_0(x',y)| + |K_0(x',y) - K_0(x,y)| \\
&\leq |(K \cdot \delta_{x'})(y') - (K \cdot \delta_{x'})(y)| + |(K \cdot \delta_{x'})(y) - (K \cdot \delta_x)(y)| \\
&\leq \sup_{y \in H} \big(|D_y(K \cdot \delta_{x'})(y)|\big) \cdot |y - y'| + |(K \cdot \delta_{x'} - K \cdot \delta_x)(y)| \\
&\leq \sup_{y \in H, x' \in L(x)} \big(|D_y(K \cdot \delta_{x'})(y)|\big) \cdot |y - y'| + \\
&\qquad\qquad\qquad |(K \cdot \delta_{x'} - K \cdot \delta_x)(y)|,
\end{aligned}$$

where the factor in the first term is bounded and the second term converges for the reasons given above. Therefore $K_0$ is continuous.

We turn to the derivatives: Let $D_{x_j}$ denote differentiation with respect to the $j^{th}$ coordinate of $x$. Then $D_{x_j}\delta_x$ is the weak limit of of $\frac{1}{h}(\delta_{x+hx_j} - \delta_x)$ in $\mathcal{E}'(X)$ for $h \to 0$. As a result

$$\frac{1}{h}(K \cdot \delta_{x+hx_j} - K \cdot \delta_x) \to K \cdot D_{x_j}\delta_x$$

in $\mathcal{E}(Y)$ and therefore the derivative $D_{x_j}K_0 = K \cdot D_{x_j}\delta_x$ exists on $X \times Y$ and $y \mapsto D_{x_j}K_0(x,y)$ is a smooth function on $Y$. Joint continuity of this derivative in $x$ and $y$ can be gained as above (since again $D_{x_j}\delta_{x'}$ converges weakly to $D_{x_j}\delta_x$ for $x' \to x$ etc.). By induction we obtain that all the derivatives $D_x^\alpha K_0$ with respect to $x$ exist, are continuous on $X \times Y$ and that the functions $y \mapsto D_x^\alpha K_0(x,y)$ are smooth on $Y$.

On the other hand, as $K \cdot \delta_{x'}$ converges to $K \cdot \delta_x$ *in* $\mathcal{E}(Y)$ for $x' \to x$, we have that the derivatives $D_y^\alpha$ with respect to $y$ of $K_0(x,y)$ are continuous on $X \times Y$. We obtain mixed derivatives through substituting $K_0(x,y)$ by $D_x^\alpha K_0$ in the previous argument, and we have thus shown that $K_0 \in \mathcal{E}(X \times Y)$.

Ultimately we show that the integral operator given by the function $K_0$, which we denote by $K'$ for the moment, is indeed $K$, i.e. that $\forall u \in \mathcal{E}'(X) : K' \cdot u = K \cdot u$. We observe that for $\delta_x$ equality is immediate from the definition:

$$(K' \cdot \delta_x)(y) = K_0(x,y) = (K \cdot \delta_x)(y).$$

However, the delta distributions generate a dense subspace of the space of distributions, i.e. we have the following statement:

Let $u \in \mathcal{E}'(X)$. Then there exists a sequence of linear combinations $\sum_i c_{ip}\delta_{y_{ip}}$ such that

$$\sum_i c_{ip}\delta_{y_{ip}} \to u \quad (\text{for } p \to \infty)$$

in the weak topology. (For a proof of this see for instance [5, Theorem 23.11.1].)

So since $K'$ is linear and continuous, we have the equality $K' \cdot u = K \cdot u$ for any $u \in \mathcal{E}'(X)$. $\square$

**Remark 3.19.** This Theorem thus proves that regularising integral operators are really the only kernel operators that can be continuously (with respect to the proper topology) extended to the distribution space $\mathcal{E}'(X)$ and still yield $\mathcal{C}^\infty$-functions. And yet, certain kernel functions allow an extension to even the space $\mathcal{D}'(X)$, i.e. to distributions that need not be compactly supported. From Proposition 3.11 it is no surprise that this is possible only if the support of the kernel function $K_0$ fulfils certain properties.

**Theorem 3.20.** *(i) Let $K_0$ be a $\mathcal{C}^\infty$-function on $X \times Y$ and let $R = \operatorname{supp}(K_0)$ satisfy that $R^{-1}(H)$ is compact in $X$ for all $H \subset\subset Y$. Then formula (3.10),*

$$(K \cdot u)(y) = \langle u, K_0(.,y)\rangle,$$

*defines a function $K \cdot u \in \mathcal{E}(Y)$ for all $u \in \mathcal{D}'(X)$ and the operator $K : \mathcal{D}'(X) \to \mathcal{E}(Y)$ is continuous on every bounded subset of $\mathcal{D}'(X)$ with respect to the weak topology on $\mathcal{D}'(X)$ and the Fréchet space topology on $\mathcal{E}(Y)$.*

*(ii) On the other hand, let $K$ be an linear operator that maps every bounded subset of $\mathcal{D}'(X)$ continuously into $\mathcal{E}(Y)$. Then the action of $K$ is given by formula (3.10), where $K_0$ is a $\mathcal{C}^\infty$-function whose support $R$ satisfies*

$$\forall H \subset\subset X : R^{-1}(H) \text{ is compact in } Y.$$

*Proof.* (i) To begin with we have that for any distribution $u \in \mathcal{D}'(X)$ and any test function $\psi \in \mathcal{D}(Y)$

$$R \cap \operatorname{supp}(u \otimes \psi) \subseteq R^{-1}(\operatorname{supp}(\psi)) \times \operatorname{supp}(\psi),$$

and so it is a compact set. Hence

$$\langle K, u \otimes \psi\rangle = \langle u, K^t \cdot \psi\rangle = \langle K \cdot u, \psi\rangle$$

is defined for all $u \in \mathcal{D}'(X)$. By (3.8) $K \cdot u$ is a smooth function and it only remains to show continuity of the operator.

Choosing a sequence $(u_k)$ in $\mathcal{D}'(X)$ that converges to 0 in the weak topology, we have to show that $D^\alpha(K \cdot u_k)$ converges to 0 uniformly on any compact subset $H$ of $Y$. By assumption we know that the supports of the functions $D^\alpha K_0(.,y)$ for $y \in H$ are contained in a compact set $L = R^{-1}(H)$ and hence $(D^\alpha K_0(.,y))_{y \in H} \subseteq \mathcal{D}(L)$. The set is bounded there since it is bounded in $\mathcal{E}(X)$ and $\mathcal{D}(L)$ is a closed subspace of $\mathcal{E}(X)$. Since the set of restrictions of $u_k$ to the Fréchet space $\mathcal{D}(L)$ is continuous and in particular pointwise bounded, we obtain via the Banach-Steinhaus theorem that it is equicontinuous. Analogously to the proof of continuity in part (i) of Theorem 3.18, it follows that $D^\alpha(K \cdot u_k)(y) = \langle u_k, D^\alpha(K_0(x,y))\rangle$ converges to 0 uniformly.

(ii) Before we can apply Theorem 3.18 we have to restrict the given operator $K$ to $\mathcal{E}'(X)$, which is possible since the embedding of $\mathcal{E}'(X)$ into $\mathcal{D}'(X)$ is continuous with respect to the topologies we are working with on these two spaces. So formula (3.10) holds for the restriction of $K$ and we also obtain the $\mathcal{C}^\infty$-kernel $K_0$. Looking at Example 3.12 where we deduced the formula, we see that it also holds in the wider setting when $u \in \mathcal{D}'(X)$.

The support property of $K_0$ follows from continuity of the operator $K$ and Remark 3.10. $\square$

# 4 Microlocal Analysis

In this chapter we introduce the wave front set of a distribution, which is a refinement of the notion of the singular support and contains information about the singularities of a distribution. We start with a definition of $WF(u)$ for distributions $u$ on open subsets of $\mathbb{R}^n$ and work out some basic properties. Subsequently we turn to defining the wave front set for distributions on a manifold, where the main task is to render the usual definition coordinate-invariant. Finally we return to regular kernel distributions $K \in \mathcal{D}'(X \times Y)$ and prove a theorem that will clarify the relationships of the wave front sets of $K$, any distribution $u \in \mathcal{E}'(X)$ and the distribution $K \cdot u \in \mathcal{D}'(Y)$.

## 4.1 Wave Front Sets on $\mathcal{D}'(\Omega)$

Let $\Omega \subseteq \mathbb{R}^n$. We begin with a repetition of the definition of the singular support.

**Definition 4.1.** Let $u \in \mathcal{D}'(\Omega)$. Then we define the singular support of $u$ as

$$\operatorname{singsupp}(u) := \Omega \setminus \{x \in \Omega : \exists\, U \text{ open neighbourhood of } x : u|_U \in \mathcal{C}^\infty(U)\},$$

where $u|_U \equiv u|_{\mathcal{D}(U)}$.

So roughly speaking the singular support tells us where, locally, the action of a distribution $u$ cannot be emulated by any $\mathcal{C}^\infty$-function. The wave front set will additionally give an idea of some sort of direction where things go wrong at such points $x \in \operatorname{singsupp}(u)$, the question is how can we further analyse singular points to gain additional information about then.

Here the Paley-Wiener theorem comes into play: It tells us that given a $\mathcal{C}^\infty$-function $f$ with compact support, the Fourier transform of $f$, denoted by $\hat{f}$ or $\mathcal{F}(f)$, is rapidly decreasing, i.e.

$$\mathcal{F}(f)(\xi) = \int_{\mathbb{R}^n} e^{-i\langle x,\xi\rangle} f(x)dx = \mathcal{O}(|\xi|^{-N}) \quad \forall N \in \mathbb{N},$$

resp.

$$\forall n \in \mathbb{N}\ \exists C_n > 0 : |\mathcal{F}(f)(\xi)| \leq C_n(1+|\xi|)^{-n}.$$

Conversely, if $f \in \mathcal{E}'$ with $\hat{f}$ rapidly decreasing, $f$ is $\mathcal{C}^\infty$. This allows us to give the following characterisation of the singular support: $x \notin \operatorname{singsupp}(u) \Leftrightarrow$

$$\exists\ U \text{ open neighbourhood of } x : \mathcal{F}(\varphi u)(\xi) = \mathcal{O}(|\xi|^{-N}) \quad \forall \varphi \in \mathcal{D}(U), \forall N \in \mathbb{N}.$$

Here we have introduced a cut-off function $\varphi \in \mathcal{D}(U)$, which ensures that the product $\varphi u \in \mathcal{E}'(\mathbb{R}^n)$ (when extended trivially) and thus enables us to apply the Fourier transform.

In the given characterisation, $\mathcal{F}(\varphi u)(\xi)$ is rapidly decreasing for arbitrary $\xi \in \mathbb{R}^n$. However it can happen that even if $x$ is in the singular support of $u$, $\mathcal{F}(\varphi u)(\xi)$ is rapidly decreasing for certain $\xi$, for instance when $\xi = t\omega$ for some vectors $\omega \in S_{n-1} := \{\omega \in \mathbb{R}^n : |\omega| = 1\}$ and $t > 0$.

**Example 4.2.** Let $\mathbb{R}^n = \{(x_1, x_2) : x_1 \in \mathbb{R}^p, x_2 \in \mathbb{R}^{n-p}\}$ for some integer $1 < p < n$ and set $u = \delta \otimes 1$, where $\delta$ is the delta distribution on $\mathbb{R}^p$ and 1 denotes the distribution given by the constant function 1 on $\mathbb{R}^{n-p}$. Then we obtain for all $\Phi \in \mathcal{D}(\mathbb{R}^p \times \mathbb{R}^{n-p})$ and every vector $\xi = (\xi_1, \xi_2) \in \mathbb{R}^p \times \mathbb{R}^{n-p}$ that

$$\mathcal{F}(\Phi u)(\xi_1, \xi_2) = \int_{\mathbb{R}^{n-p}} e^{-i\langle x_2, \xi_2\rangle} \Phi(0, x_2) dx_2.$$

This function is rapidly decreasing if and only if $\xi_2 \neq 0$, therefore for all $\xi = t\omega$ where $\omega = (\omega_1, \omega_2) \in S_{n-1}$ such that $\omega_2 \neq 0$. Note especially that in this case $\mathcal{F}(\Phi u)(\xi_1, \xi_2)$ is rapidly decreasing even if $\Phi$ is non-zero on the singular support of $u$.

We want the wave front set of $u$ to contain only the pairs $(x, \xi)$ where we cannot cut off the support in $x$ in any way to make $\mathcal{F}(\varphi u)(\xi)$ rapidly decreasing. Thus we define:

**Definition 4.3.** (i) If $u \in \mathcal{D}'(\Omega)$, then the *wave front set* $\operatorname{WF}(u)$ of $u$ is defined as the complement in $T^*\Omega \setminus \{0\} = \Omega \times (\mathbb{R}^n \setminus \{0\})$ of the collection of all $(x_0, \xi_0) \in T^*\Omega \setminus \{0\}$ such that for some neighbourhood $U$ of $x_0$, $V$ of $\xi_0$ we have that for any $\varphi \in \mathcal{D}(U)$ and each $N$:

$$\mathcal{F}(\varphi u)(\tau\xi) = \mathcal{O}(\tau^{-N}) \quad \text{for } \tau \to \infty, \text{ uniformly in } \xi \in V.$$

(ii) A subset $\Gamma$ of $\Omega \times (\mathbb{R}^n \setminus \{0\})$ is called *a cone*, or *conic*, if

$$(x, \xi) \in \Gamma \Rightarrow (x, \tau\xi) \in \Gamma \quad \forall \tau > 0.$$

**Remark 4.4.** (i) By definition, the wave front set of any distribution $u$ is a closed and conic subset of $T^*\Omega \setminus \{0\}$. Additionally we immediately see that for $u, v \in \mathcal{D}'(\Omega)$ we have

$$\mathrm{WF}(u+v) \subseteq \mathrm{WF}(u) \cup \mathrm{WF}(v),$$

and for $g \in \mathcal{C}^\infty(X)$ we get that

$$\mathrm{WF}(g \cdot u) \subseteq \mathrm{WF}(u).$$

(ii) In the context of conic sets it is often useful to use the cosphere bundle $S^*\Omega$ instead of $T^*\Omega \setminus \{0\}$. $S^*\Omega$ is given as a quotient of $T^*\Omega \setminus \{0\}$ with respect to the equivalence relation

$$(x_0, \xi_0) \sim (y_0, \nu_0) \Leftrightarrow x_0 = y_0 \text{ and } \exists \tau > 0 : \xi_0 = \tau \nu_0.$$

More explicitly

$$T^*\Omega \setminus \{0\} = \Omega \times \{\xi \in \mathbb{R}^n : \xi \neq 0\};$$
$$S^*\Omega = \Omega \times \{\xi \in \mathbb{R}^n : |\xi| = 1\} = \Omega \times S^{n-1},$$

where $|\xi| = 1$ is simply the representative of the equivalence class. Hence a conic neighbourhood $W$ of a point $(x_0, \xi_0) \in T^*\Omega \setminus \{0\}$ can also be given by the product of a neighbourhood $U$ of $x$ in $\Omega$ and a neighbourhood $V$ of $\frac{\xi}{|\xi|}$ in $S^{n-1}$. $W$ is closed (resp. open) in $\Omega \times (\mathbb{R}^n \setminus \{0\})$ if and only if $U \times V$ is closed (resp. open) in $\Omega \times S_{n-1}$.

**Proposition 4.5.** *Let $u \in \mathcal{D}'(\Omega)$ and $\pi$ denote the projection $(x, \xi) \mapsto x$ from $T^*\Omega \setminus \{0\} \to \Omega$. Then we have that*

$$\operatorname{singsupp}(u) = \pi(\mathrm{WF}(u)).$$

*Proof.* ($\subseteq$): If $x \notin \pi(\mathrm{WF}(u)) \Rightarrow \nexists\, \xi \in \mathbb{R}^n \setminus \{0\}$ such that $(x, \xi) \in \mathrm{WF}(u) \Rightarrow$ there exist neighbourhoods $U_\xi$ of $x$ and $V_\xi$ of $\xi$ such that for any $\varphi \in \mathcal{D}(U)$ we have $\mathcal{F}(\varphi u)(\tau\xi)$ is rapidly decreasing for any $\xi \in V$ as $\tau \to \infty$. For every $\xi$ we intersect the neighbourhoods $V_\xi$ with $S^{n-1}$ and use compactness of the sphere to obtain a finite cover

$$S^{n-1} = \bigcup_{i=1}^{n} V_{\xi_i}.$$

By choosing $U := \bigcup_{i=1}^n U_{\xi_i}$ we get that $\mathcal{F}(\varphi u)(\tau\xi)$ is rapidly decreasing as $\tau \to \infty$ for all $\varphi \in \mathcal{D}(U)$ and $\xi$ arbitrary, and hence $x \notin \operatorname{singsupp}(u)$.

($\supseteq$): On the other hand if $x \notin \operatorname{singsupp}(u) \Rightarrow \exists\, U$ open neighbourhood of $x$ such that

$\mathcal{F}(\varphi u)(\xi)$ is rapidly decreasing for any $|\xi| \to \infty \Rightarrow x \notin \pi(\mathrm{WF}(u))$. $\square$

## 4.2 Wave Front Sets for Distributions on Manifolds

In this section we aim at defining the wave front set for distributions on manifolds, which basically means we have to strip Definition 4.3 of any references to specific coordinates. Yet there are also other steps that have to be taken, since for instance the Fourier transform is not designed for manifolds but only works on $\mathbb{R}^n$. Therefore we will address the arising issues one by one.

Fourier transform and coordinate transforms: Before looking at how to emulate a localised Fourier transform on manifolds, shortly consider the more general problem of coordinate-transformations of the $(x, \xi)$ variables. Given a chart change in the cotangent bundle $T^*\mathbb{R}^n$ given by a diffeomorphism $\kappa : \mathbb{R}^n \to \mathbb{R}^n$, then we get for $\varphi \in \mathcal{D}(\mathbb{R}^n)$ that

$$\mathcal{F}(u \circ \kappa)(\xi) = \int e^{-i\langle x,\xi\rangle} u(\kappa(x))dx = \int e^{-i\langle \kappa^{-1}(y),\xi\rangle} u(y)|\det d\kappa^{-1}(y)|dy.$$

If $\kappa$ is linear, we can write on $\mathbb{R}^n$ that $\kappa(x) = A \cdot x$ and thus we have

$$\mathcal{F}(u \circ \kappa)(\xi) = \frac{1}{|\det A|} \int e^{-i\langle y,(A^{-1})^t\xi\rangle} u(y)dy = \frac{1}{|\det A|} \mathcal{F}(u)((A^{-1})^t\xi).$$

Therefore $(A^{-1})^t\xi$ is the respective transformation in the cotangent bundle, where $\xi \mapsto (A^{-1})^t\xi$ is still linear. However if $\kappa$ was not linear to begin with, the correspondence in the cotangent bundle becomes less clear and it would be difficult to evaluate the asymptotic behaviour, for which we need at least homogeneity in $\xi$ of degree 1. We need to keep this in mind when starting to work on manifolds, and find a way that takes care of this problem.

Fourier transform and manifolds: Since $\varphi u$ is compactly supported for any $u \in \mathcal{D}'(X)$ and $\varphi \in \mathcal{D}(X)$ we can use the following formula for the Fourier transform

$$\mathcal{F}(\varphi u)(\xi) = \langle u, e^{-i\langle .,\xi\rangle}\varphi\rangle.$$

It is apparent that on manifolds, the term $\langle ., \xi\rangle$ will cause trouble since we no longer have a scalar product we can resort to, and therefore we have to find a suitable replacement for it. In order to do this, we look at the essential properties of the scalar product, which consequently any substitute should also have:

We initially note that $\langle ., \xi\rangle$ maps $\Omega \to \mathbb{R}$, $x \mapsto \langle x, \xi\rangle$, so we are looking for a substitute of the form

$$\psi : X \to \mathbb{R}.$$

Like the scalar product, $\psi$ should be $\mathcal{C}^\infty$. Furthermore, we consider the gradient of $\langle ., \xi \rangle$ with respect to $x$ and obtain that

$$D(\langle ., \xi \rangle) = \xi.$$

This is vital here due to the following argument: In the Fourier transform, the term $e^{-i\langle x, \xi \rangle}$ is an oscillating test function, producing waves that pass over the distribution $u$ and these waves propagate in the direction of $\xi$. (Which, if not immediately clear, can also be seen from the following line of argument: The fronts of such testing waves are to be found where the scalar product $\langle x, \xi \rangle$ remains constant, that is at level hypersurfaces of the map $\langle ., \xi \rangle$. The wave fronts are perpendicular to the direction in which the wave propagates, and level hypersurfaces of $\langle ., \xi \rangle$ are in turn perpendicular to the gradient of $\langle ., \xi \rangle$. Therefore the wave moves along $\xi$.) We want to locally remain in control of the direction of passing waves, and so we deduce the following condition:

$$d\psi(x_0) = \xi_0,$$

where $d$ denotes the exterior derivative $d : \mathcal{C}^1(X) \to \Gamma(X, T^*(X))$, the appropriate substitute for $D$ when acting on scalar functions on manifolds.

In the previous section we saw that it is vital for asymptotic analysis that we obtain an expression that is homogeneous of degree 1 in $\xi$. We ensure this by using the same trick as already before, instead of letting $|\xi|$ go to infinity, we use a factor $\tau \to \infty$ which we pull out of $\psi$ from the beginning, thus to a certain extent forcing homogeneity. After all this, we propose that our definition of the wave front set of a distribution on a manifold should look like this:

Let $X$ be a manifold and $u$ be a distribution on the manifold. Then $(x_0, \xi_0) \in T^*(X) \setminus \{0\}$ is not in the *wave front set of* $u$, denoted by $\mathrm{WF}(u)$, if for any real-valued function $\psi \in \mathcal{C}^\infty(X, \mathbb{R})$ with $d\psi(x_0) = \xi_0$ there is an open neighbourhood $U_0$ of $x_0$ such that for any $\varphi \in \mathcal{D}(U)$ we have

$$\langle u, e^{-i\tau\psi(.)}\varphi \rangle = \mathcal{O}(\tau^{-N}) \quad \tau \to \infty, N \in \mathbb{N}.$$

We argue why this would yield a valid generalisation of the concept of wave front sets on open subsets of $\mathbb{R}^n$, i.e. we show that the two concepts coincide on $\mathbb{R}^n$. The core of the argument will be provided by the next proposition, which is a variant of the method of stationary phase. It will help us investigate the asymptotic behaviour of integrals similar to what we want to use in the definition.

**Proposition 4.6.** *Let $m \in \mathbb{R}$, $\delta < 1$. Let $f$ be a smooth and real-valued function on $\mathbb{R}^n$, and $g \in \mathcal{C}^\infty(\mathbb{R}^n \times \mathbb{R}^+)$ with $g(x,t) = 0$ for $x \notin K \subset\subset \mathbb{R}^n$ and*

$$\partial_x^\alpha g(x,t) = \mathcal{O}(t^{m+\delta|\alpha|}) \quad \text{for } t \to \infty, \tag{4.1}$$

*uniformly in $x \in K$. Then we are interested in the asymptotic behaviour of the integral*

$$I(t) := \int_{\mathbb{R}^n} e^{itf(x)} g(x,t) dx. \tag{4.2}$$

*We set*

$$\Sigma_f = \{x \in K : Df(x) = 0\},$$

*which is the set of stationary points of $f$ with respect to the integration variables. Assume that for every $N$ there exists a neighbourhood $\Omega$ of $\Sigma_f$ in $K$ such that*

$$g(x,t) = \mathcal{O}(t^{-N}) \quad \text{for } t \to \infty, \text{ uniformly in } x \in \Omega. \tag{4.3}$$

*We have that*

$$I(t) = \mathcal{O}(t^{-N}) \quad \text{for } t \to \infty,$$

*for all $N \in \mathbb{N}$.*

*Proof.* For detail see [7, 1.2.], as only a sketch of the proof is provided here:

Using a cut-off around $\Sigma_f$ in $\mathcal{D}(\Omega)$ and the fall-off (4.3) of $g$ on $\Omega$ one can obtain the desired asymptotic behaviour of the integral by straightforward calculation. The case where $g$ is not rapidly decreasing, i.e. where therefore $Df(x) \neq 0$ for all $x \in K \cap \operatorname{supp} g$, is more difficult.

To show the fall-off of the integral in this case one can choose a partial differential operator $P = \Sigma c_j(x) \frac{\partial}{\partial x_j}$, $c_j \in \mathcal{C}^\infty(\mathbb{R}^n)$ such that $Pf = 1$ in a neighbourhood of $K \cap \operatorname{supp} g$. Then

$$\begin{aligned} I(t) = \int e^{itf(x)} g(x,t) dx &= (it)^{-1} \int P(e^{itf(x)}) g(x,t) dx \\ &= (it)^{-1} \int e^{itf(x)} P^t(g(x,t)) dx \end{aligned}$$

Thus by repeatedly applying integration by parts one can obtain for all $k$ that

$$I(t) = (it)^{-k} \int e^{itf(x)} (P^t)^k (g(x,t)) dx$$

, and so it follows from (4.1) that $I(t) = \mathcal{O}(t^{m+\delta k-k})$ for $t \to \infty$ and for all $k$. Since $\delta < 1$ this yields the desired fall-off. $\square$

The next proposition now guarantees that the proposed definition of the wave front set coincides with Definition 4.3 for distributions on $\mathbb{R}^n$.

**Proposition 4.7.** *Let $u \in \mathcal{D}'(\Omega)$. Then $(x_0, \xi_0) \notin \mathrm{WF}(u)$ (as defined in 4.3) if and only if for any real-valued $\mathcal{C}^\infty$-function $\psi$ on $\mathbb{R}^n$ with $D\psi(x_0) = \xi_0$ there is an open neighbourhood $U_0$ of $x_0$ such that for any $\varphi \in \mathcal{D}(U_0)$ we have*

$$\langle u, e^{-i\tau\psi(.)}\varphi \rangle = \mathcal{O}(\tau^{-N}) \quad \textit{for } \tau \to \infty.$$

*Proof.* The "if" part is trivial: Simply choose for $\psi$ the scalar product $\langle ., \xi \rangle : \mathbb{R}^n \to \mathbb{R}$ and rewrite $\langle u, e^{-i\tau\psi(.)}\varphi \rangle$ as $\mathcal{F}(\varphi u)(\tau\xi)$.

To show the other direction we assume that $(x_0, \xi_0) \notin \mathrm{WF}(u)$, i.e. there exist neighbourhoods $U$ of $x_0$, and $V$ of $\xi_0$ such that we have for each $\varphi \in \mathcal{D}(U)$ and each $N$:

$$\mathcal{F}(\varphi u)(\tau\xi) = \mathcal{O}(\tau^{-N}) \quad \text{for } \tau \to \infty, \text{ uniformly in } \xi \in V.$$

Let $\rho \in \mathcal{D}(U_0)$ be equal to 1 on a neighbourhood of $\mathrm{supp}(\varphi)$. Then we get

$$\begin{aligned}
\langle u, e^{-i\tau\psi(.)}\varphi \rangle &= \langle \varphi u, e^{-i\tau\psi(.)}\rho \rangle \\
&= \langle \mathcal{F}(\varphi u), \mathcal{F}^{-1}(e^{-i\tau\psi}\rho) \rangle \\
&= (2\pi)^{-n} \langle \mathcal{F}(\varphi u)(\xi), \int e^{i\langle x,\xi \rangle} e^{-i\tau\psi(x)} \rho(x) dx \rangle \\
&= (2\pi)^{-n} \int \int e^{i\langle x,\xi \rangle} e^{-i\tau\psi(x)} \rho(x) \mathcal{F}(\varphi u)(\xi) dx d\xi \\
&\overset{\xi \to \tau\xi}{=} (2\pi)^{-n} \tau^n \int \int e^{i\tau(\langle x,\xi \rangle - \psi(x))} \rho(x) dx \cdot \mathcal{F}(\varphi u)(\tau\xi) d\xi
\end{aligned}$$

Now we apply the last proposition, or rather its proof, to the integral

$$I(\xi) = \int e^{i\tau(\langle x,\xi \rangle - \psi(x))} \rho(x) dx.$$

To this end we start by looking at the stationary points of the oscillatory function: Setting

$f(x,\xi) = \langle x,\xi\rangle - \psi(x)$ and $D_x = (\frac{\partial}{\partial x_j})_{1\leq j\leq n}$ the derivative with respect to the integration variables, we obtain that $D_x f(x,\xi) = \xi - D\psi(x)$. Since $D\psi(x_0) = \xi_0$, we can choose a small neighbourhood $U_0$ of $x_0$ such that $|D_x f(x,\xi)| = |\xi - D\psi(x)| \geq \epsilon$ for $\xi \notin V$. Here $V$ is the neighbourhood of $\xi_0$ that we get according to the definition of the complement of the wave front set.

So for $\xi \notin V$ we can apply the method of integration by parts as in the proof of Prop. 4.6 where we can explicitly choose $P = |\xi - D\psi|^{-2}\langle \xi - D\psi, \frac{\partial}{\partial x}\rangle$. This method yields an estimate of the form

$$|I(\xi)| \leq C_k \cdot \tau^{-k}(1+|\xi|)^{-k}, \quad \text{for } \xi \notin V, \tau \geq 1, k \in \mathbb{N}.$$

On the other hand $|\mathcal{F}(\varphi u)(\tau\xi)| \leq C(1+|\tau\xi|)^l$ for some $l$, and

$$|\mathcal{F}(\varphi u)(\tau\xi)| \leq C'_k \cdot \tau^{-k}, \quad \text{for } \xi \in V, \tau \geq 1,$$

for any $k$ if $U_0 \subseteq U$, and $U$ again as in the definition. These two estimates together give the desired result that $\langle u, e^{-i\tau\psi(.)}\varphi\rangle = \mathcal{O}(\tau^{-N})$ for $\tau \to \infty$. □

Consequently we can now formulate a definition of the wave front set for distributions on manifolds.

**Definition 4.8.** Let $X$ be a manifold and $u$ be a distribution on the manifold. Then $(x_0,\xi_0) \in T^*(X) \setminus \{0\}$ is not in the *wave front set of* $u$, denoted by $\mathrm{WF}(u)$, if for any real-valued function $\psi \in \mathcal{C}^\infty(X,\mathbb{R})$ with $d\psi(x_0) = \xi_0$ there is an open neighbourhood $U_0$ of $x_0$ such that for any $\varphi \in \mathcal{D}(U)$ we have

$$\langle u, e^{-i\tau\psi(.)}\varphi\rangle = \mathcal{O}(\tau^{-N}) \quad \tau \to \infty, N \in \mathbb{N}.$$

**Remark 4.9.** (i) How we defined the wave front set on manifolds is not the only way one can go about to reach a coordinate-invariant definition. One can for instance also look at the behaviour of wave front sets under a pull back given by a diffeomorphism $f : \Omega_1 \to \Omega_2$ and use this to define the wave front set via charts. However unlike the definition we gave, this is not an intrinsic approach as one has to resort to charts, which we did not have to do. Details on this approach can be found for instance in Chazarain [1, 8.3] or also in Hörmander [11, 8.2].

(ii) The original approach to wave front sets for distributions on manifolds as proposed by Hörmander in [10] is yet another one. There, Hörmander used pseudo-differential calculus on manifolds to define the wave front set as the intersection of the characteristics of all pseudo-differential operators of order 0 that give a $\mathcal{C}^\infty$-function when applied to $u \in \mathcal{D}'(X)$.

(iii) A note on the next chapter: Having reached a definition of the wave front set on manifolds, we can from now on w.l.o.g. work on open subsets of $\mathbb{R}^n$ again, using the usual trick of local descriptions of distributions on manifolds. The results of the following section remain true on manifolds, but to keep proofs and notation simple we only discuss the case of open sets in all details.

## 4.3 Microlocal Analysis of Kernel Maps

The aim of this section will be to prove estimates on the wave front set of distributions that arise out of regular kernel maps, i.e. we prove a relation between $\mathrm{WF}(K)$ (for $K \in \mathcal{D}'(X \times Y)$ regular), $\mathrm{WF}(u)$ (for $u \in \mathcal{D}'(X)$) and $\mathrm{WF}(K \cdot u)$. As already mentioned, the previous section allows us to restrict our inspection to the distribution spaces on open subsets $X$ of $\mathbb{R}^n$ and $Y$ of $\mathbb{R}^m$. Before we can state and prove this theorem, we need to make some preparations, in which we follow Dieudonné [5, 23.12].

First of all we derive a useful formula for the Fourier transform of a distribution of the form $K \cdot u$:

**Proposition 4.10.** *Let $K$ be a regular kernel distribution in $\mathcal{D}'(X \times Y)$ with compact support and let $\mathcal{F}_n$ denote the Fourier transform on $\mathbb{R}^n$. The we have for every $u \in \mathcal{E}'(X)$ and for every $\eta \in \mathbb{R}^m$ that the function $\xi \mapsto (\mathcal{F}_{m+n}K)(-\xi, \eta)(\mathcal{F}_n u)(\xi)$ is integrable on $\mathbb{R}^n$ and we have the formula*

$$\mathcal{F}_m(K \cdot u)(\eta) = (2\pi)^{-n} \int_{\mathbb{R}^n} (\mathcal{F}_{m+n}K)(-\xi, \eta)(\mathcal{F}_n u)(\xi) d\xi.$$

*Proof.* Let $e_\xi(x) = e^{-i\langle x, \xi\rangle}$ and $e_\eta(y) = e^{-i\langle y, \eta\rangle}$ for $(x, \xi) \in \mathbb{R}^n \times \mathbb{R}^n$ and $(y, \eta) \in \mathbb{R}^m \times \mathbb{R}^m$. By assumption $K \cdot u \in \mathcal{E}'(Y)$ for any $u \in \mathcal{E}'(X)$ (see Prop. 3.16) and $K^t \cdot e_\eta \in \mathcal{E}(X)$. Additionally we have that

$$\mathcal{F}_m(K \cdot u)(\eta) = \langle K \cdot u, e_\eta\rangle = \langle K^t \cdot e_\eta, u\rangle \quad \forall \eta \in \mathbb{R}^m.$$

Let $\psi \in \mathcal{D}(Y)$ be a cut-off around $\mathrm{supp}(K \cdot u)$, then we obtain

$$\langle K \cdot u, e_\eta\rangle = \langle K \cdot u, \psi e_\eta\rangle = \langle K^t \cdot (\psi e_\eta), u\rangle,$$

and since both sides of the last expression are compactly supported we have

$$\langle K^t \cdot (\psi e_\eta), u\rangle = \langle \mathcal{F}_n \left(K^t \cdot (\psi e_\eta)\right), \mathcal{F}_n^{-1}(u)\rangle. \tag{4.4}$$

Now we look at the expression $\mathcal{F}_n \left(K^t \cdot (\psi e_\eta)\right)$ and rewrite it as

$$\mathcal{F}_n \left(K^t \cdot (\psi e_\eta)\right)(\xi) = \langle K^t \cdot (\psi e_\eta), e_\xi\rangle = \langle K, e_\xi \otimes (\psi e_\eta)\rangle$$

and since $K$ is compactly supported we may assume $\psi \equiv 1$ on $\mathrm{pr}_2(\mathrm{supp}(K))$ such that

$$\langle K, e_\xi \otimes (\psi e_\eta)\rangle = \langle K, e_\xi \otimes e_\eta\rangle = (\mathcal{F}_{m+n}K)(\xi, \eta).$$

Putting things together we thus have that

$$\mathcal{F}_n \left(K^t \cdot (\psi e_\eta)\right)(\xi) = (\mathcal{F}_{m+n}K)(\xi, \eta),$$

and so since $K^t \cdot (\psi e_\eta) \in \mathcal{D}(X) \subseteq \mathcal{S}(\mathbb{R}^n)$ (again by Prop. 3.16) for any $\eta$, we get that $\xi \mapsto \mathcal{F}_{m+n}K(\xi, \eta) \in \mathcal{S}(\mathbb{R}^n)$.

Returning to equation (4.4) we have a brief look at the term $\mathcal{F}_n^{-1}(u)$: We know that $u \in \mathcal{E}'(X)$ and consequently $\mathcal{F}_n^{-1}(u)$ is a polynomially bounded function and thus we can write

$$\begin{aligned}
\langle \mathcal{F}_n \left(K^t \cdot (\psi e_\eta)\right), \mathcal{F}_n^{-1}(u)\rangle &= \langle \mathcal{F}_{m+n}K(., \eta), \mathcal{F}_n^{-1}(u)\rangle \\
&= \int \mathcal{F}_{m+n}K(\xi, \eta)(\mathcal{F}_n^{-1}u)(\xi)d\xi \\
&= (2\pi)^{-n} \int \mathcal{F}_{m+n}K(\xi, \eta)(\mathcal{F}_n u)(-\xi)d\xi \\
&= (2\pi)^{-n} \int \mathcal{F}_{m+n}K(-\xi, \eta)(\mathcal{F}_n u)(\xi)d\xi.
\end{aligned}$$

□

As a next step we turn to a lemma that will help us in constructions involving conic neighbourhoods. To this end we introduce the following notation: For any subset $V$ of $S_{n-1}$ be denote the set of all $t\xi \in \mathbb{R}^n \setminus \{0\}$ for $\xi \in V$ and $t > 0$ by $\mathbb{R}_+^* V$.

**Lemma 4.11.** *Let $W$ be a closed cone in $X \times (\mathbb{R}^n \setminus \{0\})$, $G$ be an open and conic neighbourhood of $W$ in $X \times (\mathbb{R}^n \setminus \{0\})$ and let $x \in \mathrm{pr}_1 W$. Then there exists an open neighbourhood $U$ of $x$ in $\mathrm{pr}_1 W$ and an open set $\Omega \subseteq S_{n-1}$ such that*

*(a)* $\{x\} \times \Omega \supseteq W \cap (\{x\} \times S_{n-1})$,

*(b)* $U \times \mathbb{R}_+^* \Omega \subseteq G$,

*(c)* $U \times \mathbb{R}_+^* \Omega$ *is a conic neighbourhood of* $W \cap (U \times \mathbb{R}^n \setminus \{0\})$ *in* $\operatorname{pr}_1 W \times (\mathbb{R}^n \setminus \{0\})$.

*Proof.* We choose the open set $\Omega \subseteq S_{n-1}$ such that we have

$$W \cap (\{x\} \times S_{n-1}) \subseteq \{x\} \times \Omega \subseteq G.$$

Now we prove by contradiction that there has to exist a neighbourhood $U$ of $x$ in $\operatorname{pr}_1 W$ that satisfies (c) in the statement (by our choice of $\Omega$ (a) and (b) both hold).

So we assume that for all open neighbourhoods $U$ of $x$ in $\operatorname{pr}_1 W$, $U \times \mathbb{R}_+^* \Omega$ is not a conic neighbourhood of $W \cap (U \times \mathbb{R}^n \setminus \{0\})$. This implies that there exists a sequence $x_k \to x$ in $\operatorname{pr}_1 W$ and a sequence $\xi_k$ in $S_{n-1}$ with

$$(x_k, \xi_k) \in W \quad \text{and} \quad \xi_k \notin \Omega.$$

However since $S_{n-1}$ is compact, we can suppose that the sequence $\xi_k$ converges to some $\xi \in S_{n-1}$ with $\xi \notin \Omega$. But then $(x, \xi)$ is in $W$ and by the definition of $\Omega$ this implies that $\xi \in \Omega$, which contradicts the construction of the sequence. $\square$

**Remark 4.12.** One more observation on the wave front set:

From the definition of the wave front set we directly get: Let $u \in \mathcal{D}'(X)$ and suppose that $x_0 \in \operatorname{singsupp}(u)$. Moreover, let $U_0$ be an open neighbourhood of $x_0$ in $X$ and let $\Omega \subseteq S_{n-1}$ open be such that

$$(U_0 \cap \operatorname{singsupp}(u)) \times (\mathbb{R}_+^* \Omega) \text{ is a conic neighbourhood of } \operatorname{WF}(u) \cap (U_0 \times (\mathbb{R}^n \setminus \{0\}))$$

in $\operatorname{singsupp}(u) \times (\mathbb{R}^n \setminus \{0\})$. Then we have that for every $\xi \in S_{n-1} \setminus \Omega$ there is a neighbourhood $U_\xi$ of $x_0$ in $U_0$ and a neighbourhood $V_\xi$ of $\xi$ in $S_{n-1}$ so that we have
$\forall \varphi \in \mathcal{D}(U_\xi)\ \forall k \in \mathbb{N}\ \exists C_k > 0\ \forall t > 0\ \forall \eta \in V_\xi :$

$$|\langle u, e^{-it\langle \cdot, \eta \rangle} \varphi \rangle| \leq C_k (1+t)^{-k}.$$

If we cover the compact set $S_{n-1} \setminus \Omega$ by finitely many neighbourhoods $V_{\xi_j}$ and if we choose $U \subseteq \bigcap_j U_{\xi_j}$, then we have that for every $\varphi \in \mathcal{D}(U)$ and every $k \in \mathbb{N}$ there exists a $C_k > 0$ such that $\forall t > 0$ and even $\forall \eta \in S_{n-1} \setminus \Omega$ :

$$|\langle u, e^{-it\langle \cdot, \eta \rangle} \varphi \rangle| \leq C_k (1+t)^{-k}.$$

We now introduce the setting of the theorem:
Let $X$ (resp. $Y$) be an open subset of $\mathbb{R}^n$ (resp. $\mathbb{R}^m$) and let $K$ be a regular kernel distribution in $\mathcal{D}'(X \times Y)$, i.e. the associated operator $K$ maps $\mathcal{E}'(X) \to \mathcal{D}'(Y)$. We additionally assume the following:

**(WF1)** $\mathrm{WF}(K) \cap \{((x,y),(\xi,0)) \in ((X \times Y) \times (\mathbb{R}^{n+m} \setminus \{0\}))\} = \emptyset$

**(WF2)** $\mathrm{WF}(K) \cap \{((x,y),(0,\eta)) \in ((X \times Y) \times (\mathbb{R}^{n+m} \setminus \{0\}))\} = \emptyset$

Ultimately we denote the set of all $((x,y),(\xi,\eta)) \in ((X \times Y) \times (\mathbb{R}^{n+m} \setminus \{0\}))$ for which

$$((x,y),(-\xi,\eta)) \in \mathrm{WF}(K)$$

by $\mathrm{WF}'(K)$.

**Theorem 4.13.** *Under the given assumptions we have for all $u \in \mathcal{E}'(X)$*

$$\mathrm{WF}(K \cdot u) \subseteq \mathrm{WF}'(K)(\mathrm{WF}(u)). \tag{4.5}$$

Remember that the notation on the right-hand side describes the set of all

$$(y,\eta) \in Y \times (\mathbb{R}^m \setminus \{0\})$$

for which there exists some $(x,\xi) \in \mathrm{WF}(u)$ such that $((x,y),(\xi,\eta)) \in \mathrm{WF}'(K)$ (resp. $((x,y),(-\xi,\eta)) \in \mathrm{WF}(K)$).

*Proof.* We will prove the relation (4.5) by showing that if a point $(y_0,\eta_0)$ is not an element of the right-hand side, i.e. $(y_0,\eta_0) \notin \mathrm{WF}'(K)(\mathrm{WF}(u))$, it follows that $(y_0,\eta_0) \notin \mathrm{WF}(K \cdot u)$. To be more precise, let $(y_0,\eta_0)$ with w.l.o.g. $|\eta| = 1$ be such that there exists no pair $(x,\xi)$ that fulfils

- $(x,\xi) \in \mathrm{WF}(u)$,
- $(x,y_0) \in \mathrm{singsupp}(K) =: S$,
- $((x,y_0),(-\xi,\eta_0) \in \mathrm{WF}(K)$.

Then it follows that $(y_0,\eta_0) \notin \mathrm{WF}(K \cdot u)$.

Since $u \in \mathcal{E}'(X)$ it is compactly supported and we may choose a cut-off function $h$ around its support such that $u = hu$. Then clearly it follows that $K \cdot u = K \cdot (hu)$, and we may interpret $K \cdot u$ as the image of $u$ under the operator that is given by the kernel distribution $(h \circ pr_1) \cdot K$. Therefore we can assume that the singular support $S$ of $K$ satisfies for any compact subset $H \subset\subset Y$ that $S^{-1}(H)$ is compact in $X$, and so we can in particular assume that $S^{-1}(U_0) \subset\subset X$ for any compact neighbourhood $U_0$ of $y_0$. We proceed by showing the following statement, which is slightly stronger than $(y_0, \eta_0) \notin \mathrm{WF}'(K)(\mathrm{WF}(u))$:

(∗) There exist a compact neighbourhood $U_0'$ of $(y_0, \eta_0)$ in $Y \times S_{m-1}$, a conic neighbourhood $G'$ of $\mathrm{WF}(u)$, a conic neighbourhood $G''$ of $\mathrm{WF}(K)$ and $c_2 > c_1 > 0$ such that we have for any $((x, y), (-\xi, \eta)) \in G''$ with $(y, \eta) \in U_0'$ that $(x, \xi) \notin G'$ and $c_1 \leq |\xi| \leq c_2$.

Let $U_0$ be a compact neighbourhood of $y_0$ and consider the conic sets

$$Z'' = (S^{-1}(U_0) \times U_0) \times (\mathbb{R}^{n+m} \setminus \{0\}) \cap \mathrm{WF}(K),$$

and

$$Z' = S^{-1}(U_0) \times (\mathbb{R}^n \setminus \{0\}) \cap \mathrm{WF}(u).$$

Moreover, let $(Z_k'')$ resp. $(Z_k')$ be a decreasing sequence of closed conic neighbourhoods of $Z''$ resp. $Z'$ such that

$$Z'' = \bigcap Z_k'' \qquad \mathrm{pr}_1(Z_k'') \subseteq S^{-1}(U_0) \times U_0$$

resp.

$$Z' = \bigcap Z_k' \qquad \mathrm{pr}_1(Z_k') \subseteq S^{-1}(U_0).$$

By this construction there exists a $K \in \mathbb{N}$ such that $Z_k'' \subseteq G''$ resp. $Z_k' \subseteq G'$ for any neighbourhoods $G''$ of $\mathrm{WF}(K)$ resp. $G'$ of $\mathrm{WF}(u)$ and for all $k \geq K$. We argue by contradiction and suppose that for all neighbourhoods $G''$ of $\mathrm{WF}(K)$ resp. $G'$ of $\mathrm{WF}(u)$ and $((x, y), (-\xi, \eta)) \in G''$ with $(y, \eta) \in U_0'$ either $(x, \xi) \notin G'$ or $c_1 \leq |\xi| \leq c_2$ is violated:

So we suppose there exists a sequence of points $((x_k, y_k), (-\xi_k, \eta_k)) \in Z_k''$ (which has to be in any neighbourhood $G''$ for $k$ large) with $(y_k, \eta_k) \to (y_0, \eta_0)$ (therefore $(y_k, \eta_k) \in U_0$ for $k$ large enough), which satisfies one of the following conditions:

1. $|\xi_k| \to 0$ for $k \to \infty$,
2. $|\xi_k| \to \infty$ for $k \to \infty$, or
3. $(x_k, \xi_k) \in Z_k'$ for all $k$ ($\Rightarrow (x_k, \xi_k) \in G'$ for large $k$).

Since $y_k \in U_0$ for large $k$ we have that $x_k \in S^{-1}(U_0)$ (as $(x_k, y_k) \in \mathrm{pr}_1(Z_k'') \subseteq S^{-1}(U_0) \times U_0$) and so by compactness of $S^{-1}(U_0)$ we can have $x_k \to x \in S^{-1}(U_0)$.

- If $|\xi_k| \to 0$, then $\lim((x_k, y_k), (-\xi_k, \eta_k)) = ((x, y_0), (0, \eta_0)) \in Z'' \subseteq \mathrm{WF}(K)$ in contradiction to our assumptions on the distribution $K$.

- On the other hand if $|\xi_k| \to \infty$, one could assume that the sequence $\xi_k/|\xi_k|$ converged to some $\xi \in S_{n-1}$, and using that $Z''_k$ resp. $Z''$ is conic we would obtain that

$$\lim((x_k, y_k), (-\frac{\xi_k}{|\xi_k|}, \frac{\eta_k}{|\xi|_k})) \overset{\eta_k \to \eta_0}{=} ((x, y_0), (\xi, 0)) \in Z'' \subseteq \mathrm{WF}(K),$$

  again in contradiction to the assumptions on $K$.

So we may assume that $c_1 \leq |\xi_k| \leq c_2$ for any $k$ and two constants $c_2 > c_1 > 0$ and consequently the sequence $((x_k, y_k), (-\xi_k, \eta_k))$ has to fulfil property 3, i.e. $(x_k, \xi_k) \in Z'_k$ for all $k$. We again get from compactness that the sequence $\xi_k$ takes some limit $\xi \in \mathbb{R}^n \setminus \{0\}$ and so we furthermore have that $\lim(x_k, \xi_k) = (x, \xi) \in Z' \subseteq \mathrm{WF}(u)$ and likewise $((x, y_0), (-\xi, \eta_0)) \in \mathrm{WF}(K)$. This contradicts the choice of $(y_0, \eta_0) \notin \mathrm{WF}'(K)(\mathrm{WF}(u))$ and so we have proved $(*)$.

Choose $G'$ and $G''$ as in $(*)$. We apply Lemma 4.11 twice: Once for $W = \mathrm{WF}(u)$ and $G = G'$, and once for $W = \mathrm{WF}(K)$ and $G = G''$. Then we have that for every $x \in S^{-1}(y_0)$ exist a relatively compact neighbourhood $V_x$ of $x$ in $X$, a relatively compact neighbourhood $U_x \subseteq U_0$ of $y_0$ in $Y$, an open subset $\Omega'_x$ of the sphere $S_{n-1}$ and an open subset $\Omega''_x$ of the sphere $S_{n+m-1}$ such that

- $(V_x \cap \mathrm{pr}_1(\mathrm{WF}(u))) \times (\mathbb{R}^*_+ \Omega'_x)$ is a conic neighbourhood of $\mathrm{WF}(u) \cap (V_x \times (\mathbb{R}^n \setminus \{0\}))$ that is contained in $G'$,

- $((V_x \times U_x) \cap \mathrm{pr}_1(\mathrm{WF}(K))) \times (\mathbb{R}^*_+ \Omega''_x)$ is a conic neighbourhood of $\mathrm{WF}(K) \cap ((V_x \times U_x) \times (\mathbb{R}^{m+n} \setminus \{0\}))$ that is contained in $G''$.

We cover $S^{-1}(y_0)$ by finitely many open sets $V_j = V_{x_j}$; furthermore let $U$ be a neighbourhood of $y_0$ that is contained in the intersection of the sets $U_{x_j}$ and let $\Omega$ be a neighbourhood of $\eta_0$ in $S_{m-1}$ that satisfies $U \times \Omega \subseteq U'_0$; finally we set

$$\Gamma'_j := \mathbb{R}^*_+ \Omega'_{x_j}$$

and

$$\Gamma''_j := \mathbb{R}^*_+ \Omega''_{x_j}.$$

After these constructions we return to our main argument, which is proving that $(y_0, \eta_0) \notin \mathrm{WF}(K \cdot u)$: Let $\varphi \in \mathcal{D}(U)$ be a test function on $Y$, we show that for any $k \in \mathbb{N}$ there exists a constant $C_k > 0$ such that

$$|\mathcal{F}_m(\varphi(K \cdot u))(t\eta)| \leq C_k t^{-k} \qquad \forall \eta \in \Omega, t \geq 1. \tag{4.6}$$

Let $(\psi_j)$ be a finite sequence of positive functions in $\mathcal{D}(X)$ with $\forall j : \operatorname{supp}(\psi_j) \subseteq V_j$ and $\sum_j \psi_j(x) = 1$ on a neighbourhood of $S^{-1}(y_0)$. By linearity it suffices to show (4.6) when $\varphi(K \cdot u)$ is replaced by the distributions $\varphi(K \cdot (\psi_j u)) := K_j \cdot u$, where the kernel distribution $K_j$ is compactly supported and regular. Therefore we can apply Proposition 4.10 to obtain that

$$\mathcal{F}_m(K_j \cdot u)(t\eta) = (2\pi)^{-n} \int_{\mathbb{R}^n} (\mathcal{F}_{m+n} K_j)(-\xi, t\eta)(\mathcal{F}_n u)(\xi) d\xi.$$

Hence we have to find an upper bound for the right-hand-side integral with respect to $\eta \in \Omega$ and $t \geq 1$. Since both $u$ and $K_j$ are compactly supported, we have that their Fourier transforms are in $\mathcal{O}_M(\mathbb{R}^n)$, the space of functions whose derivatives can be bounded by polynomials (see Section 2.3.). Hence there exists $C > 0$ and $p \in \mathbb{N}$ such that

$$|\mathcal{F}_{m+n} K_j(-\xi, t\eta)| \leq C(1 + |\xi| + t)^p,$$
$$|\mathcal{F}_n u(\xi)| \leq C(1 + |\xi|)^p$$

holds.

By choice of $\Gamma'_j$ and $\Gamma''_j$ as being contained in $\operatorname{pr}_2(G')$ and $\operatorname{pr}_2(G'')$, $(*)$ and the fact that $\eta \in \Omega \subseteq \operatorname{pr}_2(U'_0)$ we have that either $(-\xi, \eta) \in \Gamma''_j$ or $\eta \in \Gamma'_j$, but both cannot hold at the same time. So at least one of the points is not contained in the neighbourhood $\Gamma'_j$ resp. $\Gamma''_j$ of $\operatorname{pr}_2(\mathrm{WF}(u))$ resp. of $\operatorname{pr}_2(\mathrm{WF}(K))$ and we have the following two cases:

(a) We have

$$\left(-\frac{\xi}{t}, \eta\right) \in \Gamma''_j \quad \text{ and } \quad \frac{\xi}{t} \notin \Gamma'_j,$$

and so also $\xi \notin \Gamma'_j$ and $c_1 t \leq |\xi|$ (see $(*)$). Then by Remark 4.12 there exists for any $k \in \mathbb{N}$ a constant $C'_k > 0$ such that we have for every $\xi \notin \Gamma'_j$ that

$$|\mathcal{F}_n u(\xi)| \leq C'_k (1 + |\xi|)^{-k}.$$

From this we get that

$$\begin{aligned}|\mathcal{F}_{m+n}K_j(-\xi,t\eta)\mathcal{F}u(\xi)| &\le C_k'C(1+|\xi|+t)^p(1+|\xi|)^{-k}\\ &\le CC_k'\left(1+\left(1+\frac{1}{c_1}\right)|\xi|\right)^p\cdot\\ &\qquad\qquad(1+|\xi|)^{-p-n-1}(1+|\xi|)^{-k+p+n+1}\\ &\le \tilde{C}(1+|\xi|)^{-n-1}(1+c_1t)^{-k+p+n+1}\end{aligned}$$

as soon as $k>p+n+1$.

(b) The second case is where

$$\left(-\frac{\xi}{t},\eta\right)\notin\Gamma_j'' \quad\Rightarrow\quad (-\xi,t\eta)\notin\Gamma_j''.$$

Then again by Remark 4.12 we have for any $k\in\mathbb{N}$ we have a constant $C_k''>0$ independent of $\eta\in\Omega$ such that

$$|\mathcal{F}_{m+n}K_j(-\xi,t\eta)|\le C_k''(1+t+|\xi|)^{-k}.$$

Similar to (a) we thus gain that

$$\begin{aligned}|\mathcal{F}_{m+n}K_j(-\xi,t\eta)\mathcal{F}u(\xi)| &\le C_k''C(1+|\xi|+t)^{-k}(1+|\xi|)^p\\ &\le C_k''C(1+|\xi|+t)^{-n-p-1}(1+|\xi|+t)^{-k+n+p+1}(1+|\xi|)^p\\ &\le C_k''C(1+|\xi|)^{-n-p-1}(1+t)^{-k+n+p+1}(1+|\xi|)^p\\ &\le \tilde{C}(1+|\xi|)^{-n-1}(1+t)^{-k+p+n+1},\end{aligned}$$

for $k>p+n+1$.

Consequently we see that in both cases we get an estimate of the type

$$|\mathcal{F}_{m+n}K_j(-\xi,t\eta)\mathcal{F}u(\xi)|\le C_k(1+|\xi|)^{-n-1}(1+t)^{-k+p+n+1}$$

for $\eta\in\Omega$, $t\ge 1$ and any $\xi\in\mathbb{R}^n$. Since $k$ is arbitrary, we can integrate over $\xi$ due to the factor $(1+|\xi|)^{-n-1}$, and have the desired fall-off in $t$ from the second factor $(1+t)^{-k+p+n+1}$. □

**Remark 4.14.** On the conditions ($\mathrm{WF}_1$) and ($\mathrm{WF}_2$): For the theorem we had to insist that the kernel $K$ was not only regular but also fulfilled:

**(WF1)** $\mathrm{WF}(K) \cap \{((x,y),(\xi,0)) \in ((X \times Y) \times (\mathbb{R}^{n+m} \setminus \{0\}))\} = \emptyset$

**(WF2)** $\mathrm{WF}(K) \cap \{((x,y),(0,\eta)) \in ((X \times Y) \times (\mathbb{R}^{n+m} \setminus \{0\}))\} = \emptyset$

These conditions are stronger than regularity of the kernel distribution $K$, in fact they imply it: Remember Remark 3.14, where we gave a characterisation of regular kernels by:

$$K \in \mathcal{D}'(X \times Y) \text{ regular } \Leftrightarrow K \in \mathcal{C}^\infty(X, \mathcal{D}'(Y)) \cap \mathcal{C}^\infty(Y, \mathcal{D}'(X)).$$

Hörmander shows in [10, 2.5.13] that

$$(\mathrm{WF}_1) \Rightarrow K \in \mathcal{C}^\infty(X, \mathcal{D}'(Y)),$$

and so likewise

$$(\mathrm{WF}_2) \Rightarrow K \in \mathcal{C}^\infty(Y, \mathcal{D}'(X)).$$

(These implications can also be deduced from Dieudonné [6, 23.65.5].) Therefore $(\mathrm{WF}_1)$ and $(\mathrm{WF}_2)$ together imply that $K$ is regular, but the converse is not true.

For any regular distribution $K \in \mathcal{D}'(X \times Y)$ we only have the following, weaker statement: Let $S = \mathrm{singsupp}(K)$, then we have for any $u \in \mathcal{E}'(X)$ that

$$\mathrm{singsupp}(K \cdot u) \subseteq S(\mathrm{singsupp}(u)).$$

For a proof of this see [5, 23.12.11].

**Example 4.15.** (i) *(Convolution)* Let $u \in \mathcal{E}'(\mathbb{R}^n)$ and denote by $K_u$ the convolution operator $\mathcal{E}'(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^n)$ given by the regular kernel distribution $K_u(x,y) = u(y-x) \in \mathcal{D}'(\mathbb{R}^n \times \mathbb{R}^n)$. Then we have:

(a) Since we know $K$ to be regular we may at least apply Remark 4.14 to the kernel operator $K_u$, and we easily obtain for

$$S = \mathrm{singsupp}(K_u) = \{(x,y) \in \mathbb{R}^{2n} : (y-x) \in \mathrm{singsupp}(u)\}$$

that

$$\begin{aligned}
\mathrm{singsupp}(u * v) \subseteq S(\mathrm{singsupp}(v)) &= \{y \mid \exists x \in \mathrm{singsupp}(v) : (x,y) \in S\} \\
&= \{y \mid \exists x \in \mathrm{singsupp}(v) : (y-x) \in \mathrm{singsupp}(u)\} \\
&= \{y \mid \exists x \in \mathrm{singsupp}(v) : y \in \{x\} + \mathrm{singsupp}(u)\} \\
&= \mathrm{singsupp}(u) + \mathrm{singsupp}(v).
\end{aligned}$$

This is a well-known result which we obtain here as an easy corollary, but without the preceding theory it is usually not so easy to prove.

(b) Yet we have even more: The kernel operator $K_u$ fulfils ($\mathrm{WF}_1$) and ($\mathrm{WF}_2$) since

$$\mathrm{WF}(K_u) = \{((x,y),(-\xi,\xi)) \mid (y-x,\xi) \in \mathrm{WF}(u)\},$$

which we do not prove here but refer the reader to Hörmander [11, p. 270] or Friedlander-Joshi [8, Section 11.4] for an elegant proof based on pull-back results for wave front sets. So we may apply the Theorem 4.13 to obtain that

$$\begin{aligned}
\mathrm{WF}(u*v) &\subseteq \mathrm{WF}'(K_u)(\mathrm{WF}(v)) \\
&= \{(y,\eta) \mid \exists (x,\xi) \in \mathrm{WF}(v) : (x,y,-\xi,\eta) \in \mathrm{WF}(K_u)\} \\
&= \{(y,\eta) \mid \exists (x,\xi) \in \mathrm{WF}(v) : \eta = \xi \wedge (y-x,\xi) \in \mathrm{WF}(u)\} \\
&\overset{z=y-x}{=} \{(x+z,\xi) \mid \exists (x,\xi) \in \mathrm{WF}(v) : (z,\xi) \in \mathrm{WF}(u)\}.
\end{aligned}$$

(ii) Let $X = Y$ and $g \in \mathcal{D}(X)$. Then the regular operator $K_g : \mathcal{D}(X) \to \mathcal{D}(X), \varphi \mapsto g\varphi$ is given by the kernel distribution

$$K_g = g(y)\delta_{x-y} : \Phi \mapsto \int_X g(x)\Phi(x,x)dx,$$

(see Example 3.3). Hence the support as well as the singular support of $K_g$ are given by the diagonal $\{(x,x) \in X \times X\}$. To find the wave front set of $K_g$ we look at the Fourier transform of $\Phi K_g$ for $\Phi \in \mathcal{D}(X \times X)$ near the singular support, which is given by

$$\mathcal{F}(\Phi K_g)(\xi,\eta) = \int_X e^{-i\langle x,\xi+\eta\rangle} g(x)\Phi(x,x)dx,$$

since $\Phi K_g \in \mathcal{E}'(X \times X)$. From this we get that

$$\mathrm{WF}(K_g) = \{((x,x),(\xi,-\xi)) \in (X \times X) \times (\mathbb{R}^{2n} \setminus \{0\}) : x \in \mathrm{supp}(g)\},$$

which obviously satisfies (WF1) and (WF2). Therefore we may apply Theorem 4.13 to regain that

$$\mathrm{WF}(g \cdot u) \subseteq \mathrm{WF}(u) \quad \text{for } u \in \mathcal{D}'(X), g \in \mathcal{D}(X),$$

which we already remarked in Remark 4.4. We get this result for any $u \in \mathcal{D}'(X)$ since $K_g$ is compactly supported and so we can cut-off $u$ outside $\mathrm{pr}_1(\mathrm{supp}(K_g))$.

As in Example 3.3 we can interpret the multiplication operator as a special case of a partial differential operator, and similarly to multiplication operators we get that we can apply the theorem and obtain the formula

$$\mathrm{WF}(P \cdot u) \subseteq \mathrm{WF}(u),$$

for $P$ a partial differential operator and $u \in \mathcal{E}'(X)$. So multiplication operators as well as partial differential operators do not increase the wave front set.

This statement remains true for even one more level of abstraction: For pseudo-differential operators and their kernels we have that

$$\mathrm{WF}(K_\Psi) \subseteq \{(x, x, -\theta, \theta) : x, \theta \in \mathbb{R}^n\},$$

(cf. [12, 18.1.16.]) and so

$$\begin{aligned}\mathrm{WF}(K_\Psi \cdot u) &\subseteq \mathrm{WF}'(K_\Psi)(\mathrm{WF}(u)) \\ &\subseteq \{(y, \eta) \mid \exists (x, \xi) \in \mathrm{WF}(u) : (x, y, -\xi, \eta) \in \mathrm{WF}(K_\Psi)\} \\ &= \mathrm{WF}(u).\end{aligned}$$

We no longer have this inclusion if we move on to Fourier integral operators, whose wave front sets are considerably more complex but can be described geometrically. For more information on Fourier integral operators see for instance Duistermaat [7, chapter II resp. IV].

# Bibliography


[1] J. Chazarain & A. Piriou: *Introduction to the Theory of Linear Partial Differential Equations.* North-Holland Publishing Company, Amsterdam, 1982.

[2] J. Dieudonné: *Grundzüge der modernen Analysis, Band 1.* Friedr. Vieweg&Sohn, Braunschweig, 1985.

[3] J. Dieudonné: *Grundzüge der modernen Analysis, Band 2.* Friedr. Vieweg&Sohn, Braunschweig, 1975.

[4] J. Dieudonné: *Grundzüge der modernen Analysis, Band 3.* Friedr. Vieweg&Sohn, Braunschweig, 1976.

[5] J. Dieudonné: *Grundzüge der modernen Analysis, Band 7.* Friedr. Vieweg&Sohn, Braunschweig, 1982.

[6] J. Dieudonné: *Grundzüge der modernen Analysis, Band 8.* Friedr. Vieweg&Sohn, Braunschweig, 1983.

[7] J.J. Duistermaat: *Fourier Integral Operators.* Birkhäuser, Boston, 1996.

[8] F.G. Friedlander & M. Joshi: *Introduction to the Theory of Distributions.* Cambridge University Press, 2nd edition 1998.

[9] M. Grosser, M. Kunzinger, M. Oberguggenberger &R. Steinbauer: *Geometric Theory of Generalized Functions with Applications to General Relativity.* Kluwer Academic Publishers, Dordrecht, 2001.

[10] L. Hörmander: *Fourier Integral operators. I.* Acta Math. 127 (1971), 79-183.

[11] L. Hörmander: *The Analysis of Linear Partial Differential Operators 1: Distribution Theory and Fourier Analysis: Vol1.* Springer Verlag, Berlin, 2nd Edition 1990.

[12] L. Hörmander: *The Analysis of Linear Partial Differential Operators 3.* Springer Verlag, Berlin, 1985.

[13] M. Kunzinger: *Differential geometry*, Lecture notes, University of Vienna, 2008, available online at http://www.mat.univie.ac.at/∼mike/teaching/ss08/dg.pdf

[14] J. Peetre: *Une caractérisation abstraite des opérateurs différentiels*. Math. Scand. 7 (1959), 211-218.

[15] H.H. Schaefer: *Topological Vector Spaces*. Springer Verlag, New York, 1966.

[16] L. Schwartz: *Théorie des distributions I, II*. Hermann, Paris, 1950-51.

[17] L. Schwartz: *Théorie des noyaux*. Proc. Int. Congr. Math. Cambridge 1950, I, 220-230.

[18] F. Treves: *Topological Vector Spaces, Distributions and Kernels*. Dover Publications, Inc., Mineola, New York, 2006.

# Curriculum Vitae

| | |
|---|---|
| Name: | Vera Ganglberger |
| Date of birth: | 7 June 1987 |
| Place of birth: | Linz, Austria |
| | |
| 1993-1997 | Öffentliche Volksschule 48 in Linz |
| 1997-2005 | Europagymnasium Auhof - Realgymnasium |
| | LISA- Linz International School Auhof |
| 2004-2005 | International Baccalaureate |
| June 2005 | Austrian Matura passed with distinction |
| August 2005 | awarded Diploma of the International Baccalaureate Organization |
| | |
| 2005- | University of Vienna: |
| | Mathematics and English for becoming a teacher (Lehramt) |
| 2006-2010 | University of Vienna: |
| | Diploma studies in mathematics |
| 2008/09 winter term | University of Leeds: |
| | ERAMSUS outgoing: Department of mathematics |